ST/ESA/STAT/SER.F/96

DEPARTAMENTO DE ASUNTOS
ECONÓMICOS Y SOCIALES

División de Estadística

Estudios de métodos Serie F No. 96

Encuestas de hogares en los países en desarrollo y en transición

Naciones Unidas
New York, 2007

Departamento de Asuntos Económicos y Sociales

El Departamento de Asuntos Económicos y Sociales (DAES) de la Secretaría de las Naciones Unidas es un punto de enlace fundamental entre las políticas mundiales en las esferas ambiental, económica y social y las medidas que se adoptan en el plano nacional. El Departamento realiza su labor en tres ámbitos principales relacionados entre sí: i) compila, genera y analiza una amplia gama de datos e información de carácter ambiental, económico y social a la que recurren los Estados Miembros de las Naciones Unidas para examinar problemas comunes y evaluar las opciones de política; ii) facilita las negociaciones de los Estados Miembros en numerosos órganos intergubernamentales sobre la adopción conjunta de medidas para abordar los problemas mundiales actuales y emergentes; y iii) presta asesoramiento a los Gobiernos interesados sobre las formas de plasmar los marcos normativos elaborados en las conferencias y las cumbres de las Naciones Unidas en programas de ámbito nacional y, mediante su asistencia técnica, ayuda a fomentar la capacidad nacional.

Nota

Las signaturas de los documentos de las Naciones Unidas se componen de letras mayúsculas y cifras. La mención de una de tales signaturas indica que se hace referencia a un documento de las Naciones Unidas.

ST/ESA/STAT/SER.F/96

PUBLICACIÓN DE LAS NACIONES UNIDAS
Número de venta: S.05.XVII.6

ISBN 978-92-1-361225-5

Copyright © Naciones Unidas 2005
Reservados todos los derechos

Prefacio

Las encuestas de hogares son una fuente valiosa de datos socioeconómicos. Muchas veces permiten extraer importantes indicadores para orientar y supervisar las políticas de desarrollo. En los países en desarrollo se han convertido en la forma dominante de recopilación de datos, que complementa, y en algunos casos llega incluso a sustituir, a otros programas de recopilación de datos y sistemas de registro civil.

En esta publicación se presentan "los últimos avances" en relación con varios aspectos importantes de las encuestas de hogares en los países en desarrollo y en transición, en particular el diseño de la muestra, la realización de las encuestas, los errores no muestrales, los costos de las encuestas y el análisis de sus datos. El principal objetivo de este manual es ayudar a los estadísticos encargados de las encuestas nacionales a diseñar encuestas de hogares de manera eficiente y fiable y permitir a los usuarios una mayor utilización de los datos generados por las encuestas.

Los 25 capítulos de la publicación son obra de prestigiosos expertos de todo el mundo, especializados en metodología de investigación sobre las encuestas. La mayoría de ellos tienen experiencia práctica en ayudar a los organismos nacionales de estadística de los países en desarrollo y en transición. Esta publicación presenta las siguientes características singulares:

- Atención especial a las necesidades de los países en desarrollo y en transición;
- Insistencia en las normas y características prácticas que pueden aplicarse a los diferentes países y en las distintas encuestas;
- Consideración de los costos de las encuestas, con inclusión de ejemplos empíricos de elaboración de presupuestos y análisis de costos desglosados por componentes detallados;
- Amplia consideración de los errores no muestrales;
- Examen de las técnicas, tanto básicas como avanzadas, de análisis de datos de encuestas de hogares, incluida una detallada complementación empírica de los últimos programas informáticos disponibles para el análisis de datos de encuestas complejas;
- Presentación de ejemplos de diseño, ejecución y análisis de datos procedentes de algunas encuestas de hogares realizadas en países en desarrollo y en transición;
- Presentación de varios estudios de casos sobre encuestas reales en gran escala realizadas en países en desarrollo y en transición y que pueden utilizarse como ejemplos para el diseño de encuestas semejantes.

La presente publicación está basada en anteriores iniciativas realizadas por la División de Estadística del Departamento de Asuntos Económicos y Sociales de las Naciones Unidas, para mejorar la calidad de la metodología de las encuestas y aumentar la capacidad de los sistemas nacionales de estadística. La más ambiciosa de estas iniciativas en los dos últimos decenios ha sido el Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares (PDCNEEH). Su objetivo era ayudar a los países en desarrollo a obtener datos demográficos y socioeconómicos imprescindibles mediante un sistema integrado de encuestas de hogares, con el fin de contribuir a la planificación del desarrollo, la formulación de

políticas y la ejecución de programas. El Programa contribuyó en buena medida al progreso estadístico de muchos países en desarrollo, sobre todo en África, que se beneficiaron de un aumento significativo del número y variedad de las encuestas realizadas en el decenio de 1980. Además, respaldó las actividades metodológicas que dieron lugar a la publicación de varios estudios técnicos y manuales. En el *Manual de encuestas sobre hogares* (Edición revisada)* se presentó un panorama general de las cuestiones relacionadas con el diseño y la ejecución de ese tipo de encuestas, seguido de una serie de publicaciones en las que se consideraban cuestiones y procedimientos en áreas específicas de la metodología de las encuestas y en relación con numerosas cuestiones y entre las que cabe distinguir las siguientes:

* Estudios de Métodos, No. 31 (Publicación de las Naciones Unidas, No. de venta: S.83.XVII.13).

- *Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares: Marcos de muestreo y diseños muestrales para Programas Integrados de Encuestas de Hogares, Versión preliminar* (DP/UN/INT-84-014/5E), Nueva York, 1986.
- *National Household Survey Capability Programme: Sampling Errors in Household Surveys* (UNFPA/UN/INT-92-P80-15E), Nueva York, 1993.
- *National Household Survey Capability Programme: Survey Data Processing: A Review of Issues and Procedures* (DP/UN/INT-81-041/1), Nueva York, 1982.
- *National Household Survey Capability Programme: Non-sampling Errors in Household Surveys: Sources, Assessment y Control: Preliminary Version* (DP/UN/INT-81-041/2), Nueva York, 1982.
- *Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares: Desarrollo y diseño de los cuestionarios de encuestas* (INT-84-014), Nueva York, 1985.
- *National Household Survey Capability Programme: Household Income y Expenditure Surveys: A Technical Study* (DP/UN/INT-88-X01/6E), Nueva York, 1989.
- *National Household Survey Capability Programme: Guidelines for Household Surveys on Health* (INT/89/X06), Nueva York, 1995.
- *National Household Survey Capability Programme: Sampling Rare and Elusive Populations* (INT-92-P80-16E), Nueva York, 1993.

En la presente publicación se actualizan y amplían los aspectos técnicos de las cuestiones y procedimientos incluidos en forma detallada en las citadas publicaciones, pero haciendo referencia exclusivamente a las encuestas en países en desarrollo y en transición.

Paul Cheung

Director

División de Estadística de las Naciones Unidas
Departamento de Asuntos Económicos y Sociales

Panorama general

La presente publicación está organizada en la forma siguiente: consta de dos partes divididas en un total de 25 capítulos. La primera parte tiene 21 capítulos repartidos en cinco secciones, de la A a la E. A continuación se resume brevemente el contenido de cada una de las secciones de la primera parte.

Sección A. Diseño y realización de las encuestas. Contiene tres capítulos. En el capítulo II se presenta un panorama general de las distintas cuestiones relacionadas con el diseño de encuestas de hogares en el contexto de los países en desarrollo y en transición. En los capítulos III y IV se analizan cuestiones relativas al diseño de los cuestionarios y a la realización de las encuestas, respectivamente, en los países en desarrollo y en transición.

Sección B. Diseño de la muestra. Esta sección contiene una introducción y tres capítulos en los que se tratan los aspectos específicos del diseño de la muestra. El capítulo V tiene como tema el diseño de muestras y marcos maestros. La utilización de los efectos del diseño en el análisis de la muestra se examina en el capítulo VI, mientras que en el capítulo VII se lleva a cabo un análisis empírico de los efectos del diseño en las encuestas realizadas en varios países en desarrollo.

Sección C. Errores no muestrales. Esta sección consta de introducción y cuatro capítulos sobre los distintos aspectos de la medición, evaluación y control de los errores no muestrales en los países en desarrollo y en transición. En el capítulo VIII se tratan los errores de falta de observación (falta de respuesta y falta de cobertura). Los errores de medición se consideran en el capítulo IX. En el capítulo X se presentan las directrices y procedimientos de garantía de la calidad aplicados a las encuestas mundiales de salud, un programa de encuestas realizadas en los países en desarrollo y patrocinado por la Organización Mundial de la Salud (OMS). En el capítulo XI se describe un estudio monográfico sobre la medición, evaluación y compensación de los errores no muestrales de encuestas de hogares realizadas en el Brasil.

Sección D. Costos de las encuestas. Esta sección contiene una introducción y tres capítulos. En el capítulo XII se presenta un marco general para analizar los costos de las encuestas en el contexto de los países en desarrollo y en transición. En el capítulo XIII se describe, sobre la base de datos empíricos, un modelo de costos para una encuesta de ingresos y gastos realizada en un país en desarrollo. En el capítulo XIV se examinan cuestiones relativas a la elaboración de un presupuesto para las innumerables fases y funciones de una encuesta de hogares y se incluyen varios ejemplos y estudios monográficos que permiten establecer comparaciones que ilustran las importantes cuestiones presupuestarias examinadas en el capítulo.

Sección E. Análisis de los datos de las encuestas. Esta sección contiene una introducción y siete capítulos dedicados al análisis de datos de encuestas. El capítulo XV contiene orientaciones detalladas para la gestión de los datos de encuestas de hogares. En el capítulo XVI se examina un análisis tabular básico de los datos de encuestas, con inclusión de varios ejemplos concretos. En el capítulo XVII se examina el uso de las encuestas de hogares sobre varios temas como instrumento para la reducción de la pobreza en los países en desarrollo. En el capítulo XVIII se considera el uso de méto-

dos estadísticos de múltiples variables para la elaboración de índices a partir de datos de encuestas de hogares. En el capítulo XIX se considera el análisis estadístico de los datos de encuestas, con especial atención a las técnicas básicas de un análisis basado en modelos, a saber: la regresión lineal múltiple, la regresión logística y los métodos en varios niveles. En el capítulo XX se presentan planteamientos más avanzados para el análisis de los datos de encuestas que tienen en cuenta los efectos de la complejidad del diseño en el análisis. Finalmente, en el capítulo XXI se analizan los diversos métodos utilizados en la estimación de errores muestrales para los datos de encuestas y se describen también técnicas prácticas de análisis de datos, y se comparan varios programas informáticos utilizados para analizar datos de encuestas de gran complejidad. Se hace también hincapié en la fuerte relación entre el diseño de la muestra y el análisis de datos. En el cederrón que acompaña a esta publicación se presenta información más detallada sobre la comparación de programas informáticos, con inclusión de sus distintos resultados.

La segunda parte de la publicación, que contiene cuatro capítulos precedidos de una introducción, está dedicada a estudios monográficos en los que se presentan ejemplos concretos de encuestas realizadas en países en desarrollo y en transición. En ellos se analizan en forma detallada y sistemática algunas encuestas pagadas por el usuario y patrocinadas por organismos internacionales y encuestas presupuestadas por los países y realizadas en el marco de programas de encuestas periódicas de los sistemas estadísticos nacionales. En el capítulo XXII se describe el programa de encuestas demográficas y de salud (DHS), y en el capítulo XXIII se trata el programa de encuestas del Estudio de medición de los niveles de vida (EMNV). El examen de ambas series de encuestas incluye el cálculo de los efectos de diseño de las estimaciones de varias características fundamentales. En el capítulo XXIV se analizan el diseño y la realización de encuestas de presupuestos de hogares, utilizando como ilustración una encuesta llevada a cabo en la República Democrática Popular Lao. En el capítulo XXV se exponen las características generales de diseño y realización de encuestas organizadas en países en transición, y se incluyen varios estudios de casos.

Agradecimientos

La preparación de una publicación de esta magnitud debe ser por fuerza fruto de la cooperación. La División de Estadística se ha beneficiado inmensamente de la impagable asistencia ofrecida por muchos consultores individuales y organizaciones de todo el mundo, tanto del sistema de las Naciones Unidas como ajenos a él. Estos consultores son expertos con considerable experiencia en el diseño, realización y análisis de encuestas complejas, y muchos de ellos han tenido amplio contacto con países en desarrollo y en transición.

Todos los capítulos han sido objeto de un proceso muy riguroso de examen colegiado. Primero, cada capítulo fue examinado por dos expertos en la materia correspondiente. Los capítulos revisados se recopilaron luego para preparar el primer borrador de la publicación, que fue revisado críticamente en la reunión de un grupo de expertos organizada en Nueva York por la División de Estadística, en octubre de 2002. Al final de la reunión se estableció un consejo de redacción para examinar la publicación y formular las recomendaciones finales acerca de su estructura y contenido. Esta fase del proceso de revisión dio lugar a una reestructuración y racionalización de toda la publicación para hacerla más coherente, más completa y más armoniosa. Se escribieron nuevos capítulos y los antiguos se revisaron teniendo en cuenta las recomendaciones de la reunión del grupo de expertos y del consejo de redacción. Luego, cada capítulo revisado fue sometido a una tercera ronda de examen por dos expertos antes de tomar la decisión final de incluirlo o no en la publicación. Después, un equipo de editores se encargó de la revisión final y completa de la publicación, para garantizar que el material presentado fuera técnicamente válido, coherente entre sí y fiel a los objetivos primarios de la publicación.

La División de Estadística reconoce con agradecimiento las valiosas contribuciones del Sr. Graham Kalton a esta publicación. El Sr. Kalton presidió tanto la reunión del grupo de expertos como el consejo de redacción, examinó muchos capítulos y aportó asesoramiento técnico y orientación intelectual al personal de la División durante todo el proyecto. El Sr. John Eltinge realizó una considerable labor de orientación en las fases iniciales de elaboración de las ideas que dieron lugar a esta publicación y en calidad de revisor de varios capítulos y orientador y colaborador de algunas investigaciones básicas que permitieron formular el marco de la presente publicación continuó desempeñando un papel decisivo en todos los aspectos del proyecto. Los Sres. James Lepkowski, Oladejo Ajayi, Hans Pettersson, Karol Krotki y Anthony Turner contribuyeron con su inestimable ayuda editorial en varios capítulos y ofrecieron orientación general y apoyo en varias fases del proyecto.

Muchos otros expertos colaboraron con este proyecto en calidad de autores de los capítulos, como revisores de los capítulos preparados por otros expertos o como ambas cosas. Otros contribuyeron al proyecto participando en la reunión del grupo de expertos y ofreciendo el fruto de su examen de todos los aspectos del proyecto inicial de la publicación. En la lista de la página XXV se presentan los nombres de todos los expertos que participaron en el proyecto, junto con las instituciones a las que pertenecen.

Habría sido difícil, si no imposible, alcanzar los ambiciosos objetivos del proyecto sin la enorme contribución de varios funcionarios de la División de Estadística en todas las fases. El Sr. Ibrahim Yansaneh hizo la propuesta para la publicación, contrató a otros participantes y coordinó todos los aspectos técnicos del proyecto, incluido el proceso editorial. Escribió

también varios capítulos y desempeñó la labor de jefe de redacción de toda la publicación. El Director y el Director Adjunto de la División de Estadística alentaron y dieron su apoyo institucional a lo largo de todas las fases del proyecto. El Sr. Stefan Schweinfest organizó todo los aspectos administrativos del proyecto. La Sra. Sabine Warschburger diseñó y mantuvo el sitio web del proyecto y la Sra. Denise Quiroga aportó sus magníficos servicios de secretaría facilitando la circulación de los numerosos documentos entre autores y editores, organizando y armonizando los heterogéneos formatos y estilos de redacción de los documentos y ayudando a respetar el calendario del proyecto.

Índice

Prefacio iii
Panorama general v
Agradecimientos vii
Lista de expertos colaboradores xxv
Autores xxvii
Revisores xxix

PRIMERA PARTE. DISEÑO, REALIZACIÓN Y ANÁLISIS DE ENCUESTAS

Capítulo I. Introducción 1

A. Encuestas de hogares en países en desarrollo y en transición 3
B. Objetivos de la presente publicación 4
C. Importancia práctica de los objetivos 5

SECCIÓN A. DISEÑO Y REALIZACIÓN DE LAS ENCUESTAS

Capítulo II. Panorama general de las cuestiones relacionadas con el diseño de la muestra para las encuestas de hogares en países en desarrollo y en transición

A. Introducción 9
   1. Diseños muestrales para encuestas en países en desarrollo y en transición 9
   2. Panorama general 10
B. Muestreo estratificado en varias etapas 10
   1. Estratificación explícita 11
   2. Estratificación implícita 11
   3. Selección muestral de las unidades primarias de muestreo 12
   4. Muestreo de UPM con probabilidad proporcional al tamaño 13
   5. Selección muestral de hogares 16
   6. Número de hogares que deben seleccionarse por UPM 16
C. Marcos muestrales 18
   1. Características de los marcos muestrales para las encuestas en los países en desarrollo y en transición 18
   2. Problemas del marco muestral y posibles soluciones 19
   3. Mantenimiento y evaluación de los marcos muestrales 20
D. Estimación del campo 21
   1. Necesidad de estimaciones del campo 21
   2. Asignación de muestras 21
E. Tamaño de la muestra 22
   1. Factores que influyen en las decisiones sobre el tamaño de la muestra 22
   2. Precisión de las estimaciones de la encuesta 22

3. Calidad de los datos 24
4. Costo y oportunidad 25
F. Análisis de la encuesta 25
1. Elaboración y ajuste de las ponderaciones de muestreo 25
2. Análisis de los datos de encuestas de hogares 27
G. Conclusiones 27
Anexo. Diagrama de flujo del proceso de una encuesta 30

Capítulo III. Panorama general del diseño de cuestionarios para las encuestas de hogares en los países en desarrollo

A. Introducción 31
B. Aspectos generales 32
1. Objetivos de la encuesta 32
2. Obstáculos 34
3. Algunos consejos prácticos 35
C. Los detalles 35
1. El concepto de módulo 35
2. Formato y coherencia 37
3. Otras orientaciones sobre algunos detalles del diseño del cuestionario 41
D. El proceso 41
1. Constitución de un equipo 41
2. Elaboración del primer borrador de cuestionario 42
3. Comprobación sobre el terreno y terminación del cuestionario 43
E. Observaciones finales 44

Capítulo IV. Consideraciones generales sobre la realización de encuestas de hogares en los países en desarrollo

A. Introducción 47
B. Actividades antes de que la encuesta se realice sobre el terreno 48
a) Financiamiento del presupuesto 48
b) Plan de trabajo 50
c) Establecimiento de una muestra de hogares 51
d) Redacción de los manuales de capacitación 52
e) Capacitación del personal sobre el terreno y del encargado de la entrada de datos 53
f) Plan de actividades sobre el terreno y de entrada de datos 53
g) Realización de una prueba piloto 54
h) Lanzamiento de una campaña publicitaria 54
C. Actividades durante la realización de la encuesta sobre el terreno 54
1. Comunicaciones y transporte 55
2. Supervisión y garantía de calidad 55
3. Gestión de datos 56
D. Actividades necesarias después de la conclusión del trabajo sobre el terreno y la entrada y procesamiento de los datos 56
1. Informe final 57
2. Preparación del conjunto final de datos y documentación 57
3. Análisis de datos 57
E. Observaciones finales 58

SECCIÓN B. DISEÑO DE LA MUESTRA

Introducción 61

Capítulo V. Diseño de marcos muestrales maestros y muestras maestras para las encuestas de hogares en países en desarrollo

A. Introducción 65
B. Marcos muestrales maestros y muestras maestras. Panorama general 66
1. Marcos muestrales maestros 66
2. Muestras maestras 67
3. Resumen y conclusión 69
C. Diseño de un marco muestral maestro 71
1. Datos y materiales: evaluación de la calidad 71
2. Decisión sobre la cobertura del marco muestral maestro 72
3. Decisión sobre las unidades de marco básicas 72
4. Información sobre las unidades que deben incluirse en el marco 73
5. Documentación y mantenimiento de un marco muestral maestro 75
D. Diseño de muestras maestras 77
1. Elección de unidades primarias de muestreo para la muestra maestra 77
2. Combinación/división de zonas para reducir la variación en el tamaño de las UPM 78
3. Estratificación de las UPM y asignación de una muestra maestra a los estratos 80
4. Muestreo de las UPM 81
5. Durabilidad de las muestras maestras 82
6. Documentación 83
7. Utilización de una muestra maestra para encuestas de establecimientos 83
E. Observaciones finales 84

Capítulo VI. Estimación de los componentes de los efectos del diseño para su uso en el diseño muestral

A. Introducción 87
B. Componentes de los efectos del diseño 90
1. Estratificación 91
2. Conglomeración 95
3. Ponderación de los ajustes 98
C. Modelo para los efectos del diseño 101
D. Uso de los efectos del diseño muestral 105
E. Conclusiones 107

Capítulo VII. Análisis de los efectos del diseño en las encuestas de los países en desarrollo

A. Introducción 111
B. Las encuestas 112
C. Efectos del diseño 113
D. Cálculo de las tasas de homogeneidad 119
E. Consideraciones 121
Anexo. Descripción de los diseños muestrales para las 11 encuestas de hogares 122

SECCIÓN C. ERRORES NO MUESTRALES

Introducción 129

Capítulo VIII. Error de falta de observación en las encuestas de hogares en los países en desarrollo

A. Introducción 131
B. Marco para comprender los errores por falta de cobertura y por falta de respuesta 132
C. Error por falta de cobertura 135
   1. Causas de la falta de cobertura 135
   2. Error por falta de cobertura 138
D. Error por falta de respuesta 141
   1. Causas de la falta de respuesta en las encuestas de hogares 141
   2. Sesgo por falta de respuesta 142
   3. Medición del sesgo por falta de respuesta 144
   4. Reducción y compensación de la falta total de respuesta en las encuestas de hogares 145
   5. Falta de respuesta sobre una partida e imputación 147

Capítulo IX. Error de medición en las encuestas de hogares: causas y medición

A. Introducción 151
B. Causas del error de medición 153
   1. Efectos del cuestionario 154
   2. Efectos del método de recopilación de datos 156
   3. Efectos del entrevistador 158
   4. Efectos del informante 160
C. Planteamientos para cuantificar el error de medición 162
   1. Experimentos con aleatoriedad 162
   2. Métodos de investigación cognitiva 163
   3. Reentrevistas 164
   4. Estudios de comprobación de los registros 167
   5. Estudios sobre la varianza del entrevistador 169
   6. Codificación del comportamiento 169
D. Observaciones finales: el error de medición 170

Capítulo X. Garantía de calidad en las encuestas: Normas, directrices y procedimientos

A. Introducción 176
B. Normas y procedimientos de garantía de calidad 176
C. Aplicación práctica de las directrices sobre garantía de calidad: ejemplo de la Encuesta Mundial de Salud 178
   1. Selección de las instituciones encargadas de las encuestas 179
   2. Muestreo 180
   3. Traducción 183
D. Capacitación 185
E. Realización de la encuesta 188
F. Entrada de datos 191
G. Análisis de datos 194

H. Indicadores de calidad 195
1. Índice de desviación de la muestra 196
2. Tasa de respuesta 196
3. Tasa de datos omitidos 197
4. Coeficientes de fiabilidad para las entrevistas de prueba-comprobación 197
I. Informes de los países 197
J. Visitas sobre el terreno 198
K. Conclusiones 199

Capítulo XI. Presentación de informes y compensación de errores no muestrales para encuestas realizadas en el Brasil: práctica actual y desafíos futuros

A. Introducción 203
B. Práctica actual de registro y compensación de los errores no muestrales en las encuestas de hogares en el Brasil 206
1. Errores de cobertura 206
2. Falta de respuesta 210
3. Errores de medición y procesamiento 213
C. Desafíos y perspectivas 214
D. Recomendaciones sobre lecturas complementarias 216

SECCIÓN D. COSTOS DE LAS ENCUESTAS

Introducción 221

Capítulo XII. Análisis de las cuestiones relacionadas con los costos de las encuestas en los países en desarrollo y en transición

A. Introducción 223
1. Criterios para diseños muestrales eficientes 223
2. Componentes de las estructuras de costo en las encuestas de los países en desarrollo y en transición 224
3. Panorama general de este capítulo 225
B. Componentes del costo de una encuesta 226
C. Costos de las encuestas cuando se dispone de una infraestructura considerable 227
1. Factores relacionados con los preparativos 227
2. Factores relacionados con la recopilación y procesamiento de datos 227
D. Costos de las encuestas con poca o nula infraestructura previa 228
E. Factores relacionados con las modificaciones en los objetivos de la encuesta 229
F. Algunas advertencias sobre los informes relativos a los costos de las encuestas 229
G. Resumen y observaciones finales 230
Anexo. Marco de presupuestación para las encuestas por conglomerados de indicadores múltiples (MICS) del Fondo de las Naciones Unidas para la Infancia (UNICEF) 233

Capítulo XIII. Modelo de costos para una encuesta de ingresos y gastos

A. Introducción 235

B. Modelos de costos y estimaciones de costos 236
1. Modelos de costos 236
2. Estimaciones de costos 236
C. Modelos de costos para un diseño muestral eficiente 237
D. Estudio de un caso: Encuesta sobre gastos y consumo de la República Democrática Popular Lao de 2002 (LECS-3) 239
1. Diseño de la muestra para la LECS-3 239
2. Recopilación de datos en la LECS-3 239
3. Trabajo sobr el terreno 240
E. Modelo de costos para el trabajo sobre el terreno en la Encuesta sobre gastos y consumo de la República Democrática Popular Lao (LECS-3) 240
1. Estimaciones de costos 240
2. Modelos de costos 241
3. Tamaño óptimo de la muestra dentro de las aldeas 242
F. Observaciones finales 243

Capítulo XIV. Elaboración de un marco para la presupuestación de las encuestas de hogares en países en desarrollo

A. Introducción 245
B. Consideraciones preliminares 246
1. Fases de una encuesta 246
2. Calendario para la encuesta 247
3. Tipo de encuesta 248
4. Presupuestos frente a gastos 249
5. Estudios anteriores 249
C. Principales categorías contables dentro del marco presupuestario 250
1. Personal 251
2. Transporte 252
3. Equipo 252
4. Bienes fungibles 252
5. Otros costos 252
6. Ejemplos de presupuestación por categorías contables 253
D. Actividades fundamentales de la encuesta en el marco del presupuesto 255
1. Presupuestación para la preparación de la encuesta 255
2. Presupuestación de la realización de la encuesta 255
3. Presupuestación del procesamiento de datos de la encuesta 256
4. Presupuestación de los informes sobre la encuesta 256
5. Ejemplos de presupuestación de actividades de la encuesta 256
E. Sinopsis 257
F. Posibles limitaciones y dificultades presupuestarias 258
G. Sistema de registro y resúmenes 259
1. Diseño de la encuenta 259
2. Realización de la encuesta 260
3. Presentación de informes sobre la encuesta 260
4. Comparación entre gastos y presupuesto 260
H. Conclusiones 260
Anexo. Ejemplos de formularios para los registros diarios y semanales 261

SECCIÓN E. ANÁLISIS DE DATOS DE ENCUESTAS

Introducción 267

Capítulo XV. Guía para la gestión de datos de las encuestas de hogares

A. Introducción 271
B. Gestión de datos y diseño del cuestionario 272
C. Estrategias operacionales para la entrada y la edición de los datos 273
D. Criterios de control de calidad 276
E. Desarrollo del programa de entrada de datos 279
F. Organización y divulgación de los conjuntos de datos de la encuesta 281
G. Gestión de datos en el proceso de muestreo 284
H. Resumen de las recomendaciones 293

Capítulo XVI. Presentación de estadísticas descriptivas sencillas a partir de los datos de las encuestas

A. Introducción 297
B. Variables y estadísticas descriptivas 298
   1. Tipos de variables 298
   2. Estadísticas descriptivas sencillas 299
   3. Presentación de estadísticas descriptivas para una variable 301
   4. Presentación de estadísticas descriptivas para dos variables 304
   5. Presentación de estadísticas descriptivas para tres o más variables 306
C. Orientaciones generales para presentar estadísticas descriptivas 307
   1. Preparación de los datos 307
   2. Presentación de los resultados 308
   3. ¿Qué es lo que hace que un cuadro tenga o no interés? 309
   4. Uso de ponderaciones 311
D. Preparación de un informe general (compendio o informe estadístico) para una encuesta de hogares 312
   1. Contenido 312
   2. Proceso 313
E. Observaciones finales 313

Capítulo XVII. Uso de encuestas de hogares sobre varios temas para mejorar las políticas de reducción de la pobreza en los países en desarrollo

A. Introducción 315
B. Análisis descriptivo 316
   1. Definición de la pobreza 317
   2. Elaboración de un perfil de pobreza 317
   3. Utilización de los perfiles de pobreza para análisis básicos de las políticas 319
C. Análisis de regresión múltiple de los datos de encuestas de hogares 320
   1. Análisis de la demanda 321
   2. Uso de los servicios sociales 322
   3. Efectos de programas gubernamentales concretos 323
D. Resumen y observaciones finales 323

Capítulo XVIII. Métodos con múltiples variables para la elaboración de índices

A. Introducción 327
B. Algunas restricciones al uso de los métodos con múltiples variables 328
C. Descripción general de los métodos con múltiples variables 329
D. Gráficos y medidas sinópticas 330
E. Análisis por conglomerados 332
F. Análisis de componentes principales (ACP) 335
G. Métodos con múltiples variables en la creación de índices 337
1. Elaboración de modelos de gastos de consumo para la creación de una variable sustitutiva del ingreso 338
2. Análisis de componentes principales (ACP) utilizado para crear un índice "de riqueza" 340
H. Conclusiones 341

Capítulo XIX. Análisis estadístico de los datos de las encuestas

A. Introducción 345
B. Estadísticas descriptivas: ponderaciones y estimación de la varianza 347
C. Estadísticas analíticas 351
D. Observaciones generales sobre los modelos de regresión 352
E. Modelos de regresión lineal 355
F. Modelos de regresión logística 359
G. Uso de modelos en varios niveles 361
H. Modelos en apoyo de los procesos de las encuestas 365
I. Conclusiones 365

Capítulo XX. Planteamientos más avanzados para el análisis de los datos de encuestas

A. Introducción 369
1. Diseño muestral y análisis de datos 369
2. Ejemplos de los efectos (y de la ausencia de efectos) del diseño muestral en el análisis 370
3. Conceptos básicos 371
4. Efectos del diseño y su papel en el análisis de datos muestrales complejos 373
B. Planteamientos básicos del análisis de datos muestrales complejos 374
1. La especificación del modelo, como base del análisis 374
2. Posibles relaciones entre el modelo y el diseño muestral: diseños ilustrativos y no ilustrativos 375
3. Problemas en la utilización de los programas informáticos estándar para el análisis de muestras complejas 375
C. Análisis de regresión y modelos lineales 376
1. Efecto de las variables de diseño no incluidas en el modelo y estimadores ponderados de regresión 376
2. Comprobación del efecto del diseño en el análisis de regresión 378
3. Modelos en varios niveles en el marco de un diseño muestral ilustrativo 379
D. Análisis de datos categóricos 380
1. Modificaciones en las pruebas de ji-cuadrado para las pruebas de bondad del ajuste y de independencia 380
2. Generalizaciones para los modelos log-lineales 383

E. Resumen y conclusiones 385
Anexo. Definiciones formales y resultados técnicos 386

Capítulo XXI. Estimación de errores muestrales en los datos de encuestas

A. Diseños muestrales 392
B. Aspectos del análisis de datos de encuestas complejas 392
1. Análisis ponderados 392
2. Descripción general de la estimación de varianzas 393
3. Factores de corrección al muestreo de poblaciones finitas para muestreo sin reemplazamiento 393
4. Pseudoestratos y pseudo-UPM 394
5. Una aproximación común (*WR*) para la descripción de muchos planes de muestreo complejo 395
6. Técnicas de estimación de varianzas y variables de diseño de la encuesta 396
7. Análisis de datos de encuestas complejas 397
C. Métodos de estimación de varianzas 397
1. Linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas 397
2. El método de replicación para la estimación de varianzas 398
3. Replicación repetida y equilibrada (BRR) 399
4. Técnicas de replicación *jackknife* (JK) 400
5. Algunos errores comunes de los usuarios de *software* de estimación de varianzas 401
D. Comparación de la estimación de varianzas con distintos paquetes de *software* 401
E. El conjunto de datos de la encuesta de Burundi 404
1. Población de inferencia y parámetros de población 404
2. Plan de muestreo y recogida de datos 404
3. Procedimientos de ponderación y preparación para la estimación de varianzas 404
4. Tres ejemplos de análisis de datos de encuestas 405
F. Empleo de procedimientos ajenos a las muestras para analizar datos de encuestas 406
G. Procedimientos para análisis de encuestas en SAS 8.2 407
1. Descripción general de SURVEYMEANS y SURVEYREG 407
2. SURVEYMEANS 408
3. SURVEYREG 409
4. Ejemplos numéricos 409
5. Ventajas e inconvenientes y costo 410
H. SUDAAN 8.0 410
1. Descripción general de SUDAAN 410
2. DESCRIPT 412
3. CROSSTAB 413
4. Ejemplos numéricos 413
5. Ventajas e inconvenientes y costo 414
I. Procedimientos para análisis de encuestas en STATA 7.0 415
1. Descripción general de STATA 415

2. SVYMEAN, SVYPROP, SVYTOTAL, SVYLC 416
3. SVYTAB 417
4. Ejemplos numéricos 417
5. Ventajas e inconvenientes y costo 418
J. Procedimientos para encuestas en Epi-Info 6.04d y Epi-Info 2002 418
1. Descripción general de Epi-Info 418
2. Módulo CSAMPLE de Epi-Info Versión 6.04d (DOS) 419
3. Epi-Info 2002 (Windows) 419
4. Ejemplos numéricos 420
5. Ventajas e inconvenientes y costo 421
K. WesVar 4.2 421
1. Descripción general de WesVar 421
2. Uso de WesVar versión 4.2 422
3. Ejemplos numéricos 423
4. Ventajas e inconvenientes y costo 424
L. PC-CARP 424
M. CENVAR 425
N. IVEware (versión beta) 426
O. Conclusiones y recomendaciones 426

SEGUNDA PARTE. ESTUDIOS DE CASOS

Introducción 433

Capítulo XXII. Las encuestas demográficas y de salud

A. Introducción 437
B. Historia 438
C. Contenido 438
D. Marco muestral 439
E. Etapas de muestreo 440
F. Información sobre la falta de respuesta 442
G. Comparación de las tasas de falta de respuesta 443
H. Efectos del diseño muestral observados en las DHS 444
I. Realización de la encuesta 447
J. Preparación y traducción de los documentos de la encuesta 447
K. Prueba preliminar 448
L. Contratación del personal sobre el terreno 449
M. Capacitación de los entrevistadores 450
N. Trabajo sobre el terreno 450
O. Procesamiento de datos 452
P. Análisis y redacción del informe 453
Q. Divulgación 453
R. Utilización de los datos del DHS 454
S. Fortalecimiento de la capacidad 454
T. Enseñanzas aprendidas 455
Anexo. Tasas de respuesta de hogares y mujeres en 66 encuestas de 44 países, 1990-2000, en algunas regiones 458

Capítulo XXIII. Encuestas del Estudio de medición de los niveles de vida

A. Introducción 461
B. ¿Por qué una encuesta del EMNV? 462
C. Características fundamentales de las encuestas del EMNV 462
1. Contenido e instrumentos utilizados 463
2. Cuestiones relacionadas con la muestra 465
3. Organización del trabajo sobre el terreno 466
4. Calidad 467
5. Entrada de datos 468
6. Sostenibilidad 469
D. Costos de realización de una encuesta del EMNV 470
E. ¿Hasta qué punto ha influido en la calidad el diseño del EMNV? 471
1. Tasas de respuesta 471
2. Falta de respuesta sobre partidas concretas 472
3. Comprobaciones de coherencia interna 473
4. Efectos del diseño muestral 474
F. Usos de los datos de las encuestas del EMNV 475
G. Conclusiones 476
Anexo I. Encuestas del Estudio de medición de los niveles de vida efectuadas en varios países 478
Anexo II. Presupuestación de una encuesta del EMNV 480
Anexo III. Efecto del diseño muestral en la precisión y eficiencia de las encuestas del EMNV 481

Capítulo XXIV. Diseño de la encuesta y diseño de la muestra en las encuestas sobre el presupuesto de los hogares

A. Introducción 487
B. Diseño de la encuesta 488
1. Métodos de recopilación de datos en las encuestas sobre el presupuesto de los hogares 488
2. Problemas de medición 489
3. Períodos de referencia 490
4. Frecuencia de las visitas 490
5. Falta de respuesta 491
C. Diseño de la muestra 491
1. Estratificación y asignación de la muestra a los estratos 491
2. Tamaño de la muestra 492
3. Muestreo a lo largo del tiempo 492
D. Estudio de caso acerca de la Encuesta sobre gasto y consumo de la República Democrática Populr Lao, de 1997/1998 493
1. Condiciones generales para la realización de la encuesta 493
2. Temas incluidos en la encuesta y en los cuestionarios 494
3. Métodos de medición 494
4. Diseño de la muestra y trabajo sobre el terreno 495
E. Experiencias y enseñanzas aprendidas 495
1. Métodos de medición y falta de respuesta 495
2. Diseño muestral y errores de muestreo 496

3. Experiencias registradas en relación con el diario sobre el uso del tiempo 497
4. Uso de la LECS-2 para las estimaciones del PIB 497
F. Observaciones finales 498

Capítulo XXV. Encuestas de hogares en países en transición

A. Evaluación general de las encuestas de hogares en los países en transición 502
1. Introducción 502
2. Encuestas de hogares por muestreo en países de Europa central y oriental y la URSS antes del período de transición (1991-2000) 502
3. Encuestas de hogares en el período de transición 504
4. Encuestas sobre el presupuesto de los hogares 505
5. Encuestas de población activa 506
6. Características comunes de los diseños muestrales y realización de las HBS y las LFS 507
7. Observaciones finales 514
B. Encuestas de hogares por muestreo en países en transición: estudios de casos 514
1. Encuesta de población activa y Encuesta sobre el presupuesto de los hogares en Eslovenia 515
2. Encuesta de hogares por muestreo de Estonia 519
3. Diseño y realización de la Encuesta sobre el presupuesto de los hogares y la Encuesta de población activa en Hungría 522
4. Diseño y realización de las encuestas de hogares en Letonia 525
5. Encuestas de hogares por muestreo en Lituania 529
6. Encuestas de hogares en Polonia durante el período de transición 531

CUADROS

II.1. Efectos del diseño en combinaciones seleccionadas de tamaño de la muestra conglomerada y correlación intraclase 17
II.2. Tamaños óptimos de las submuestras para determinadas combinaciones de coeficiente de costo y correlación intraclase 18
II.3. Errores estándar e intervalos de confianza para estimaciones de la tasa de pobreza basadas en varios tamaños de muestra, suponiendo que el efecto del diseño es 2,0 23
II.4. Coeficiente de variación para estimaciones de la tasa de pobreza basadas en diversos tamaños de muestra, suponiendo que el efecto del diseño es 2,0 24
IV.1. Proyecto de presupuesto para una encuesta hipotética de 3.000 hogares 49
VI.1. Efectos del diseño debidos al muestreo no proporcional en el caso de los dos estratos 93
VI.2. Distribuciones de la población y tres asignaciones alternativas de la muestra entre las provincias (A-H) 106
VII.1. Características de las 11 encuestas de hogares incluidas en el estudio. 112

VII.2. Efectos estimados del diseño de ocho encuestas realizadas en África y en Asia sudoriental 114
VII. 3. Efectos estimados del diseño para estimaciones por países y por tipo de zona en algunas estimaciones de hogares (PNAD, 1999) 115
VII.4. Efectos estimados del diseño en algunas características individuales en el plano nacional y en varios subcampos (PNAD, 1999) 116
VII.5. Efectos estimados del diseño en algunas estimaciones de la PME correspondientes a septiembre de 1999 117
VII.6. Efectos estimados del diseño en algunas estimaciones de la PPV 117
VII.7. Comparaciones de los efectos del diseño en las diferentes encuestas .. 118
VII.8. Efectos globales del diseño divididos en efectos de la ponderación $(d_w^2(\overline{y}))$ y de la conglomeración $(d_{cl}^2(\overline{y}))$ 119
VII.9. Tasas de homogeneidad en las zonas urbanas y rurales 121
X.1. Lista sinóptica sobre la calidad del muestreo 183
X.2. Lista sinóptica para el examen de los procedimientos de traducción .. 185
X.3. Lista sinóptica para el examen de los procedimientos de capacitación 188
X.4. Lista sinóptica para el examen de la realización de la encuesta 190
X.5. Lista sinóptica para el proceso de entrada de datos 194
XI.1. Algunas características de las principales encuestas de hogares por muestreo del Brasil 207
XI.2. Estimaciones de la tasas de omisión en los censos de población del Brasil obtenidas a partir de las encuestas posteriores al empadronamiento de 1991 y 2000 209
XIII.1. Tiempo estimado para el trabajo sobre el terreno en una aldea 241
XIII.2. Costos estimados para la LECS-3 241
XIII.3. Tamaño óptimo de la muestra en las aldeas (m_{opt}) y eficiencia relativa del diseño efectivo (m=15) con diferentes valores de ρ 242
Cuadro de Gantt aplicado a un proyecto de calendario para una encuesta sobre el sector informal 247
XIV.1. Matriz de categorías contables frente a actividades de la encuesta 250
XIV.2. Matriz del tiempo de personal previsto (días) con respecto a las actividades de la encuesta 251
XIV.3. Costos desglosados por categorías contables, en proporción del presupuesto total. Encuesta sobre las metas para el final del decenio (1999-2000), algunos países africanos 254
XIV.4. Proporción del presupuesto asignado a categorías contables: Evaluación del impacto de los servicios a microempresas (AIMS), Zimbabwe (1999) 254
XIV.5. Costos de las actividades de la encuesta en proporción del presupuesto total. Encuesta sobre las metas de final del decenio (1999-2000), algunos países africanos 256
XIV.6. Costos de las actividades de la encuesta en proporción del presupuesto total: AIMS, Zimbabwe (1999) 258
XIV.7. Costos de las categorías contables por actividad de la encuesta en cuanto proporción prevista del presupuesto: AIMS, Zimbabwe (1999) 258
XIV.8. Costos de las categorías contables por actividad de la encuesta en cuanto proporción realizada del presupuesto: AIMS, Zimbabwe (1999) 258

XV.1. Datos de una encuesta de hogares almacenados en forma de sencillo archivo rectangular. 281
XVI.1. Distribución de la población por edad y sexo, Saipán, Comunidad de las Islas Marianas del Norte, abril de 2002: porcentajes de las filas 300
XVI.2. Distribución de la población por edad y sexo, Saipán, Comunidad de las Islas Marianas del Norte, abril de 2002: porcentajes de las columnas 300
XVI.3. Estadísticas sinópticas de los ingresos de hogares por grupo étnico, Samoa Americana, 1994 301
XVI.4. Fuentes de iluminación en los hogares vietnamitas, 1992-1993 302
XVI.5. Información sinóptica sobre el total de gastos de los hogares: Viet Nam, 1992-1993 304
XVI.6. Uso de servicios de salud por la población (de todas las edades) que visitó un centro de salud en las cuatro últimas semanas, desglosado por zonas rurales y urbanas de Viet Nam, en 1992-1993 304
XVI.7. Gasto total de los hogares por región en Viet Nam, 1992-1993 306
XVIII.1. Algunas técnicas con variables múltiples y su objetivo 330
XVIII.2. Datos sobre las explotaciones en que se aprecia la presencia o ausencia de diversas características 334
XVIII.3. Matriz de semejanzas entre ocho explotaciones agrícolas. 334
XVIII.4. Resultados de un análisis de componentes principales 336
XVIII.5. Variables utilizadas y sus ponderaciones correspondientes en la creación de un índice predictivo del gasto de consumo para la región del Kilimanjaro en la República Unida de Tanzanía 339
XVIII.6. Límites para dividir a la población en cinco quintiles de riqueza. 341
XIX.1. Estructura típica del diseño de las encuestas de hogares 346
XIX.2. Interpretación de las estimaciones del parámetro de regresión lineal cuando la variable dependiente son los ingresos del hogar procedentes de salarios, modelo 1 356
XIX.3. Ingresos estimables de los hogares resultantes de los salarios, modelo 1 356
XIX.4. Interpretación de las estimaciones del parámetro de regresión lineal cuando la variable dependiente son los ingresos del hogar procedentes de salarios, modelo 2 357
XIX.5. Interpretación de las estimaciones del parámetro de regresión logística cuando la variable dependiente es un indicador de los hogares por debajo del umbral de pobreza, modelo 4 360
XX.1. Sesgo y media cuadrática del estimador de cuadráticos mínimos ordinarios y varianzas de estimadores no sesgados para una población de 3.850 explotaciones agrícolas utilizando varios diseños de encuestas 378
XX.2. Cuadro de análisis de la varianza de comparación de regresiones ponderadas y no ponderadas 379
XX.3. Coeficientes de tres pruebas de ji-cuadrado iteradas y pruebas de muestreo aleatorio simple 381
XX.4. Tamaños asintóticos estimados de pruebas basadas en X^2 y en X_C^2 para algunas partidas seleccionadas de la Encuesta general de hogares de 1971 del Reino Unido de Gran Bretaña e Irlanda del Norte 382

XX.5. Tamaños asintóticos estimados de las pruebas basadas en X_I^2, $X_I^2/\hat{\delta}^2$ y en $X_I^2/\hat{\lambda}^2$ para la clasificación cruzada de variables seleccionadas de la Encuesta general de hogares del Reino Unido de Gran Bretaña e Irlanda del Norte 383

XX.6. Niveles de significación (NL) asintótica estimada de X^2 y estadísticos corregidos $X^2/\hat{\delta}^2$, $X^2/\hat{\lambda}^2$, $X^2/\hat{d}^2$ cuadro 2 x 5 x 4 y nivel de significación nominal $\alpha = 0{,}05$ 384

XXI.1. Comparación de procedimientos entre cinco paquetes de *software*: porcentaje estimado y número de mujeres que son seropositivas, con error estándar estimado; mujeres que han dado a luz recientemente, Burundi, 1988-1989..... 402

XXI.2. Atributos de ocho paquetes de *software* con capacidad de estimación de varianzas para datos de encuestas complejas..... 403

XXII.1. Valores medios $d(\overline{y})$ y $\widehat{\rho}$ en 48 encuestas demográficas y de salud (DHS), 1984-1993 446

XXIII.1. Contenido del cuestionario de hogares de Viet Nam, 1997-1998 464

XXIII.2. Ejemplos de módulos adicionales 464

XXIII.3. Controles de calidad en las encuestas del EMNV 467

XXIII.4. Tasas de respuesta en encuestas recientes del EMNV 472

XXIII.5. Frecuencia de la ausencia de datos sobre los ingresos en el EMNV y en las encuestas de población activa 472

XXIII.6. Hogares con agregados completos de consumo: ejemplos de encuestas recientes del EMNV..... 473

XXIII.7. Coherencia interna de los datos: relaciones correctas entre módulos . 474

XXIII.8. Ejemplos de efectos del diseño en las encuestas del EMNV..... 474

A.III.1. Variación de los efectos del diseño por variable, Ghana, 1987 482

AIII.2. Variación de los efectos del diseño a lo largo del tiempo, Ghana, 1987 y 1988..... 483

AIII.3. Variación de los efectos del diseño en los diferentes países..... 484

AIII.4. Descripción de las variables del análisis: nivel individual..... 484

AIII.5. Descripción de las variables del análisis: hogares 484

XXIV.1. Efectos del diseño en el consumo de los hogares y posesión de bienes duraderos..... 496

XXIV.2. Coeficiente entre el número efectivo y previsto de personas en la muestra del diario sobre el uso del tiempo 497

XXV.1. Nuevas encuestas sobre el presupuesto de los hogares y de población activa en algunos países en transición, 1992-2000: ... 505

XXV.2. Tamaño de la muestra, diseño de la muestra y métodos de estimación en la HBS y la LFS, 2000, en algunos países en transición 508

XXV.3. Tasas de falta de respuesta en la HBS en algunos países en transición, 1992-2000 511

XXV.4. Tasas de falta de respuesta en la LFS en algunos países en transición, 1992-2000 512

XXV.5. Estructura de costos de la HBS en Hungría en el año 2000 513

XXV. 6. Estructura de costos de la LFS en Hungría en el año 2000 513

FIGURAS

III.1. Ilustración de un formato de cuestionario 38
IV.1. Plan de trabajo para la preparación y realización de una encuesta de hogares 51
X.1. Procedimientos de garantía de calidad de la WHS 177
X.2. Proceso de entrada de datos y garantías de calidad. 192
X.3. Ejemplo de un índice de desviación de la muestra 196
XV.1. Segunda encuesta sobre el nivel de vida en Nepal. 283
XV.2. Utilización de una hoja de cálculo como marco muestral de la primera etapa 285
XV.3. Puesta en práctica de la estratificación implícita. 286
XV.4. Selección de una muestra PPT (primer paso) 287
XV.5. Selección de una muestra PPT (segundo paso) 288
XV.6. Selección de una muestra PPT (tercer paso). 289
XV.7. Selección de una muestra PPT (cuarto paso). 290
XV.8. Hoja de cálculo con las unidades primarias de muestreo seleccionadas. 291
XV.9. Cálculo de las probabilidades de selección en la primera etapa. 291
XV.10. Documentación de los resultados de la enumeración de hogares 292
XV.11. Documentación de la falta de respuesta 292
XV.12. Cálculo de probabilidades de la segunda etapa y ponderaciones de muestreo 293
XVI.1. Fuentes de iluminación en los hogares vietnamitas, 1992-1993 (gráfico de columnas). 302
XVI. 2. Fuentes de iluminación en los hogares vietnamitas, 1992-1993 (diagrama de sectores) 302
XVI.3. Distribución por edades de la población de Saipán, abril de 2002 (histograma). 303
XVI.4. Uso de servicios de salud por la población (de todas las edades) que visitó un centro de salud en las cuatro últimas semanas, desglosado por zonas urbanas y rurales de Viet Nam, en 1992-1993 305
XVIII.1. Ejemplo de un diagrama matricial entre cuatro variables. 331
XVIII.2. Dendograma formado por la matriz de semejanza entre explotaciones 335
XIX.1. Aplicación de ponderaciones y estimación estadística. 347
XX.1. Ninguna selección 371
XX.2. Selección en X: XL<X<XU. 371
XX.3. Selección en X: X<XL; X>XU 371
XX.4. Selección en Y: YL<Y<YU 371
XX.5. Selección en Y: Y<YL; Y>YU. 371
XX.6. Selección en Y: Y>U. 371
XXIII.1. Componentes de costos de una encuesta del EMNV 470

ESQUEMAS

XXIII.1. Relación entre objetivos del EMNV e instrumentos de las encuestas. 463
XXIII.2. Calendario de las actividades de cada equipo durante un mes 466

Lista de expertos colaboradores

Participantes en la reunión del Grupo de expertos sobre las características prácticas de las encuestas en los países en desarrollo y con economías en transición
(8-10 de octubre de 2002, Nueva York)

Savitri **Abeyasekera**
Universidad de Reading
Reading
Reino Unido de Gran Bretaña
e Irlanda del Norte

Oladejo O**. Ajayi**
Asesor estadístico
Ikoyi, Lagos
Nigeria

Jeremiah **Banda**
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Grace **Bediako**
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Donna **Brogan**
Universidad de Emory
Atlanta, Georgia
Estados Unidos de América

Mary **Chamie**
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

James R. **Chromy**
Instituto Research Triangle
Research Triangle Park
Carolina del Norte
Estados Unidos de América

Willem **de Vries**
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Paul **Glewwe**
Universidad de Minnesota
St. Paul, Minnesota
Estados Unidos de América

Ivo **Havinga**
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Rosaline **Hirschowitz**
Statistics South Africa
Pretoria, Sudáfrica

Gareth **Jones**
Fondo de las Naciones Unidas
para la Infancia
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Graham **Kalton**
Westat
Rockville, Maryland
Estados Unidos de América

Hiroshi **Kawamura**
DESA/División de Análisis de Políticas
de Desarrollo
Naciones Unidas
Nueva York, Nueva York

Erica **Keogh**
Universidad de Zimbabwe
Harare, Zimbabwe

Jan **Kordos**
Escuela de Economía de Varsovia
Varsovia, Polonia

James **Lepkowski**
Instituto para la Investigación Social
Ann Arbor, Michigan
Estados Unidos de América

Gad **Nathan**
Hebrew University
Jerusalén, Israel

Frederico **Neto**
DESA/División de Análisis de Políticas de Desarrollo
Naciones Unidas
Nueva York, Nueva York

Colm **O'Muircheartaigh**
Universidad de Chicago
Chicago, Illinois
Estados Unidos de América

Hans **Pettersson**
Statistics Suecia
Estocolmo, Suecia

Hussein **Sayed**
Universidad de El Cairo
Orman, Giza, Egipto

Michelle **Schoch**
Fondo de Población de las Naciones Unidas
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Stefan **Schweinfest**
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Pedro **Silva**
Fundaçao Instituto Brasileiro de Geografía e EstaTística (IBGE)
Río de Janeiro, Brasil

Anatoly **Smyshlyaev**
DESA/División de Análisis de Políticas de Desarrollo
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Diane **Steele**
Banco Mundial
Washington, D.C.
Estados Unidos de América

Sirageldin **Suliman**
DESA/UNSD
Nueva York, Nueva York
Estados Unidos de América

Shyam **Upadhyaya**
Integrated Statistical Services (INSTAT)
Kathmandú, Nepal

T. Bedirhan **Üstün**
Organización Mundial de la Salud
Ginebra, Suiza

Martin **Vaessen**
Programa de encuestas demográficas y de salud
ORC Macro (Empresa del grupo Opinion Research Corporation)
Calverton, Maryland
Estados Unidos de América

Ibrahim **Yansaneh**
Comisión de Administración Pública Internacional
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

Autores

Savitri **Abeyasekera**
Universidad de Reading
Reading, Reino Unido de Gran Gretaña e Irlanda del Norte

J. Michael **Brick**
Westat
Rockville, Maryland
Estados Unidos de América

Donna **Brogan**
Universidad de Emory
Atlanta, Georgia
Estados Unidos de América

Somnath **Chatterji**
Organización Mundial de la Salud
Ginebra, Suiza

James R. **Chromy**
Instituto Research Triangle
Research Triangle Park
Carolina del Norte
Estados Unidos de América

Paul **Glewwe**
Universidad de Minnesota
St. Paul, Minnesota
Estados Unidos de América

Hermann **Habermann**
Oficina del Censo de los Estados Unidos
Suitland, Maryland
Estados Unidos de América

Graham **Kalton**
Westat
Rockville, Maryland
Estados Unidos de América

Daniel **Kasprzyk**
Mathematica Policy Research
Washington, D.C.
Estados Unidos de América

Erica **Keogh**
Universidad de Zimbabwe
Harare, Zimbabwe

Jan **Kordos**
Escuela de Economía de Varsovia
Varsovia, Polonia

Thanh **Lê**
Westat
Rockville, Maryland
Estados Unidos de América

James **Lepkowski**
Universidad de Michigan
Ann Arbor, Michigan
Estados Unidos de América

Michael **Levin**
Oficina del Censo de los Estados Unidos
Washington, D.C.
Estados Unidos de América

Abdelhay **Mechbal**
Organización Mundial de la Salud
Ginebra, Suiza

Juan **Muñoz**
Consultor independiente
Santiago, Chile

Christopher J. L. **Murray**
Organización Mundial de la Salud
Ginebra, Suiza

Gad **Nathan**
Universidad Hebrea
Jerusalén, Israel

Hans **Pettersson**
Statistics Sweden
Estocolmo, Suecia

Kinnon **Scott**
Banco Mundial
Washington, D. C.
Estados Unidos de América

Pedro **Silva**
Fundaçao Instituto Brasileiro de Geografía e Estatística (IBGE)
Río de Janeiro, Brasil

Bounthavy **Sisouphantong**
Centro Nacional de Estadística
Vientiane
República Democrática Popular Lao

Diane **Steele**
Banco Mundial
Washington, D.C.
Estados Unidos de América

Tilahun **Temesgen**
Banco Mundial
Washington, D.C.
Estados Unidos de América

Mamadou **Thiam**
Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura
Montreal, Canadá

T. Bedirhan **Üstun**
Organización Mundial de la Salud
Ginebra, Suiza

Martin **Vaessen**
Programa de encuestas de salud y demográficas
ORC Macro
(Empresa del grupo Opinion Research Corporation)
Calverton, Maryland
Estados Unidos de América

Vijay **Verma**
Universidad de Siena
Siena, Italia

Ibrahim **Yansaneh**
Comisión de Administración Pública Internacional
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York

Revisores

Oladejo **Ajayi**
Asesor estadístico
Lagos, Nigeria

Paul **Biemer**
Instituto Research Triangle
Research Triangle Park
Carolina del Norte
Estados Unidos de América

Steven B. **Cohen**
Agency for Healthcare Research and Quality
Rockville, Maryland
Estados Unidos de América

John **Eltinge**
Oficina de Estadística del Departamento de Trabajo de los Estados Unidos
Washington, D.C.
Estados Unidos de América

Paul **Glewwe**
Universidad de Minnesota
St. Paul, Minnesota
Estados Unidos de América

Barry **Graubard**
Instituto Nacional del Cáncer
Bethesda, Maryland
Estados Unidos de América

Stephen **Haslett**
Universidad de Massey
Palmerston North
Nueva Zelandia

Steven **Heeringa**
Universidad de Michigan
Ann Arbor, Michigan
Estados Unidos de América

Thomas B. **Jabine**
Asesor estadístico
Washington, D.C.
Estados Unidos de América

Gareth **Jones**
Fondo de las Naciones Unidas para la Infancia
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

William D. **Kalsbeek**
Universidad de Carolina del Norte
Chapel Hill, Carolina del Norte
Estados Unidos de América

Graham **Kalton**
Westat
Rockville, Maryland
Estados Unidos de América

Ben **Kiregyera**
Oficina de Estadística de Uganda
Kampala, Uganda

Jan **Kordos**
Escuela de Económica de Varsovia
Varsovia, Polonia

Phil **Kott**
Departamento de Agricultura de los Estados Unidas
Servicio Nacional de Estadística Agrícola
Fairfax, Virginia
Estados Unidos de América

Karol **Krotki**
NuStats
Austin, Texas
Estados Unidos de América

James **Lepkowski**
Universidad de Michigan
Ann Arbor, Michigan
Estados Unidos de América

Dalisay **Maligalig**
Banco de Desarrollo Asiático
Manila, Filipinas

David **Marker**
Westat
Rockville, Maryland
Estados Unidos de América

Juan **Muñoz**
Consultor independiente
Santiago, Chile

Gad **Nathan**
Universidaad Hebrea
Jerusalén, Israel

Colm **O'Muircheartaigh**
Universidad de Chicago
Chicago, Illinois
Estados Unidos de América

Robert **Pember**
Organización Internacional del Trabajo
Oficina de Estadística
Ginebra, Suiza

Robert **Santos**
NuStats
Austin, Texas
Estados Unidos de América

Pedro **Silva**
Fundaçao Instituto Brasileiro
de Geografía e Estatística (IBGE)
Río de Janeiro, Brasil

Anthony G. **Turner**
Sampling Consultant
Jersey City, New Jersey
Estados Unidos de América

Ibrahim Y**ansaneh**
Comisión de Administración
Pública Internacional
DESA/UNSD
Naciones Unidas
Nueva York, Nueva York
Estados Unidos de América

PRIMERA PARTE

Diseño, realización y análisis de encuestas

Capítulo I
Introducción

Ibrahim S. Yansaneh*
Comisión de Administración Pública Internacional
Naciones Unidas, Nueva York

* Ex Jefe de la Unidad de Metodología y Análisis, DESA/UNSD.

RESUMEN

En este capítulo se presenta un breve panorama de las encuestas de hogares realizadas en países en desarrollo y en transición. Además, se esbozan los objetivos generales de la publicación y la importancia práctica de dichos objetivos.

Términos clave: Encuestas de hogares, características prácticas, diseños de encuestas complejas, costo de las encuestas, errores muestrales.

A. Encuestas de hogares en países en desarrollo y en transición

1. En los últimos decenios se ha observado un aumento de la demanda de datos demográficos y socioeconómicos vigentes y detallados sobre los hogares e individuos en los países en desarrollo y en transición. Estos datos se han convertido en elemento indispensable en el análisis de las políticas económicas y sociales, la planificación del desarrollo, la gestión de programas y la toma de decisiones en todos los niveles. Para atender esta demanda, las autoridades y otras partes interesadas han recurrido con frecuencia a las encuestas de hogares. En consecuencia, estas encuestas se han convertido en uno de los mecanismos más importantes para la recopilación de información sobre las poblaciones en los países en desarrollo y en transición. Ahora constituyen un componente central y estratégico en la organización de los sistemas nacionales de estadísticas y en la formulación de políticas. La mayor parte de los países cuentan ya con sistemas de recopilación de datos para las encuestas de hogares, aunque con diferentes niveles de experiencia e infraestructura. Las encuestas realizadas por los institutos nacionales de estadística tienen en general numerosos objetivos o son de carácter integrado y tratan de ofrecer datos fiables sobre diversas características demográficas y socioeconómicas de las distintas poblaciones. Las encuestas de hogares se están utilizando también para estudiar las pequeñas y medianas empresas y las pequeñas propiedades agrícolas en los países en desarrollo y en transición.

2. Además de las encuestas nacionales financiadas con cargo al presupuesto público, hay un gran número de encuestas de hogares realizadas en los países en desarrollo y en transición que son patrocinadas por organismos internacionales, con el fin de elaborar y supervisar estimaciones nacionales de características o indicadores de interés para dichos organismos, y también para establecer comparaciones internacionales de esos indicadores. La mayor parte de tales encuestas se realizan con carácter excepcional, pero hay cada vez mayor interés por establecer programas integrados de encuestas sobre temas diversos y con varias rondas, con

asistencia técnica de organizaciones internacionales, las Naciones Unidas y el Banco Mundial, en todas las fases del diseño, realización, análisis y divulgación de las encuestas. Como ejemplos destacados de encuestas de hogares realizadas por organismos internacionales en países en desarrollo cabe señalar las encuestas demográficas y de salud (DHS), realizadas por ORC-Macro para la Agencia de los Estados Unidos para el Desarrollo Internacional (USAID); las encuestas del Estudio de medición de los niveles de vida (EMNV), que cuentan con asistencia técnica del Banco Mundial, y las encuestas agrupadas de indicadores múltiples (MICS), llevadas a cabo por el Fondo de las Naciones Unidas para la Infancia (UNICEF). Estos programas de encuestas se aplican en diversos países en desarrollo de África, Asia, América Latina y el Caribe y el Oriente Medio. Los programas de DHS y EMNV se describen detalladamente en los estudios de casos incluidos en los capítulos V y VI, respectivamente. Véase también Banco Mundial (2000) donde se examinan con detalle otros programas de encuestas realizados por el Banco Mundial en países en desarrollo, incluidas las encuestas prioritarias y las encuestas sobre el cuestionario sobre indicadores básicos del bienestar. En cuanto a los detalles sobre las MICS, véase UNICEF (2000). El programa DHS es fruto de un programa anterior de encuestas, el Estudio mundial sobre la fecundidad (WFS), financiado por USAID y por el Fondo de Población de las Naciones Unidas (FNUAP), con apoyo de los gobiernos del Reino Unido de Gran Bretaña e Irlanda del Norte, los Países Bajos y el Japón. Véase Verma y otros (1980), donde se facilita más información sobre ese programa.

B. Objetivos de la presente publicación

3. La presente publicación ofrece un marco metodológico para la realización de encuestas en países en desarrollo y en transición. Dado el gran número de encuestas realizadas en estos países, siempre es necesario realizar estudios metodológicos en todas las fases del proceso de la encuesta y aplicar los mejores métodos disponibles para los productores y usuarios de datos de encuestas de hogares. Gran parte de esta labor metodológica se lleva a cabo con los auspicios de organismos internacionales y de la División de Estadística de las Naciones Unidas a través de sus publicaciones e informes técnicos. La presente publicación constituye el más reciente de estos esfuerzos.

4. La mayor parte de las encuestas llevadas a cabo en países en desarrollo y en transición se basan ahora en metodologías y procedimientos uniformes que se utilizan en todo el mundo. No obstante, muchas de estas encuestas se realizan en un entorno de fuertes restricciones presupuestarias en países con niveles muy heterogéneos de capacidad técnica e infraestructura para las encuestas. Es clara la constante necesidad no sólo de desarrollar y mejorar las metodologías subyacentes, sino también de transmitir dichas metodologías a los países en desarrollo y en transición. La mejor manera de lograrlo es mediante la cooperación técnica y el fortalecimiento de la capacidad de estadística. La presente publicación, que se ha preparado precisamente con ese fin, constituye un acervo básico de material técnico y de otras informaciones necesarias para el diseño y realización eficiente de encuestas de hogares, y para una utilización eficaz de los datos recopilados.

5. La publicación está destinada a todos los interesados en la producción y uso de datos de encuestas, en particular:

- Funcionarios de institutos nacionales de estadística;
- Consultores internacionales que ofrecen asistencia técnica a los países;
- Investigadores y otros especialistas dedicados al análisis de los datos de las encuestas de hogares;
- Profesores y alumnos interesados por los métodos de investigación de encuestas.

6. La publicación constituye una fuente global de datos y material y referencia sobre aspectos importantes del diseño, realización y análisis de encuestas por muestreo de hogares

en los países en desarrollo y en transición. Los lectores pueden utilizar la información metodológica y las orientaciones generales presentadas en la primera parte de la publicación, junto con los estudios de casos de la segunda parte, para diseñar las nuevas encuestas en esos países. Más en concreto, los objetivos de esta publicación son los siguientes:

a) Establecer una fuente central de datos y de material de referencia en la que se incluyan los aspectos técnicos del diseño, realización y análisis en los países en desarrollo y en transición;

b) Ayudar a los profesionales dedicados a la realización de encuestas a diseñar y realizar encuestas de hogares de manera más eficiente;

c) Presentar casos de estudios de varios tipos de encuestas que se han llevado a cabo o se están realizando en algunos países en desarrollo y en transición, haciendo hincapié en las características que pueden generalizarse y pueden ayudar a los profesionales a diseñar y realizar nuevas encuestas en esos o en otros países;

d) Examinar los componentes más detallados de tres características prácticas de las encuestas —los efectos del diseño, los costos y los errores no muestrales— y explorar la posibilidad de transferir estas características o sus componentes a diferentes encuestas y países;

e) Presentar directrices concretas para el análisis de datos obtenidos a partir de encuestas muestrales complejas y una comparación detallada de los tipos de programas informáticos para el análisis de los datos de las encuestas.

C. Importancia práctica de los objetivos

7. Las encuestas de hogares realizadas en países en desarrollo y en transición tienen muchas características en común. Además, hay con frecuencia semejanzas entre los países, sobre todo los de las mismas regiones, en lo que se refiere a las características clave de las poblaciones subyacentes. En la medida en que los diseños de las muestras para las encuestas de hogares y las características de la población subyacente son semejantes en diversos países, es de prever que algunas características prácticas o sus componentes sean también semejantes y quizá puedan aplicarse entre países.

8. La transferibilidad de las características prácticas de las encuestas ofrece varias ventajas. En primer lugar, la información sobre el diseño de una encuesta dada en un país determinado puede servir como orientación práctica para mejorar la eficiencia de la encuesta cuando se repita en el mismo país, o para mejorar la eficiencia de una encuesta semejante llevada a cabo en el mismo país o en otro. En segundo lugar, los países con escasa o nula infraestructura para las encuestas pueden beneficiarse inmensamente de los datos empíricos sobre las características del diseño muestral y la realización efectuada en otros países con mejor infraestructura y capacidad estadística general. En tercer lugar, pueden conseguirse importantes ahorros debido al hecho de que la costosa información relacionada con el diseño de las muestras puede "tomarse en préstamo" de una encuesta anterior. Además, la experiencia práctica resultante de una encuesta anterior puede utilizarse para multiplicar la eficiencia del diseño de la encuesta que es objeto de consideración.

9. Esta publicación tiene el objetivo general importante de promover el desarrollo de las encuestas de hogares de alta calidad en los países en desarrollo y en transición, además de considerar las cuestiones del costo y eficiencia del diseño y realización de encuestas, y está basada en anteriores iniciativas de las Naciones Unidas, como el Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares (PDCNEEH), finalizado hace más de un decenio. Los estudios de casos ofrecen una importante orientación sobre los aspectos del diseño y realización de las encuestas que han funcionado eficazmente en los países en desarrollo y en transición, sobre los inconvenientes que deben evitarse y sobre los pasos

que pueden darse para mejorar la eficiencia en lo que respecta a la fiabilidad de los datos de ellas y para reducir los costos generales de las mismas. El hecho de que todas las encuestas descritas en esta publicación se hayan llevado a cabo en países en desarrollo y en transición hace que este instrumento sea sumamente relevante y eficaz para el desarrollo estadístico en esos países.

10. El análisis y la divulgación de los datos de las encuestas se encuentran entre las áreas que necesitan un mayor fomento de la capacidad en los países en desarrollo y en transición. En muchos casos los análisis de datos de las encuestas no van más allá de las frecuencias y tabulaciones básicas. Los análisis adecuados de los datos y la divulgación oportuna de los resultados de dichos análisis garantizan que pueda disponerse fácilmente de la información necesaria para la formulación de políticas y la toma de decisiones sobre la asignación de recursos. Esta publicación contiene orientaciones prácticas sobre la manera de realizar análisis más complejos de microdatos y de explicar las complejidades del diseño en el análisis de los datos generados, la incorporación de los objetivos del análisis en la fase del diseño y la utilización de programas informáticos especiales para analizar datos de encuestas complejas.

11. En resumen, esta publicación constituye una fuente global de material de referencia sobre todos los aspectos de las encuestas de hogares realizadas en países en desarrollo y en transición. Es de esperar que el material técnico presentado en la primera parte, junto con los ejemplos concretos y estudios de casos de la segunda, sean útiles para los profesionales de las encuestas de todo el mundo en el diseño, realización y análisis de nuevas encuestas de hogares.

REFERENCIAS BIBLIOGRÁFICAS

Banco Mundial (2000). *Poverty in Africa: Survey databank*. Disponible en http://www4.worldbank.org/afr/poverty.

Fondo de las Naciones Unidas para la Infancia (UNICEF) (2000). *Manual de fin de decenio del UNICEF para la encuesta agrupada de indicadores múltiples.* Nueva York: UNICEF, febrero.

Verma, V., C. Scott y C. O'Muircheartaigh (1980). Sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society, Series A,* vol. 143, págs. 431-473. Con debate.

Sección A

Diseño y realización de encuestas

Capítulo II
Panorama general de las cuestiones relacionadas con el diseño de la muestra para las encuestas de hogares en países en desarrollo y en transición

Ibrahim S. Yansaneh*
Comisión de Administración Pública Internacional
Naciones Unidas, Nueva York

* Ex jefe de la Unidad de Metodología y Análisis, División de Estadística de las Naciones Unidas.

RESUMEN

En el presente capítulo se examinan las cuestiones clave vinculadas con el diseño de muestras nacionales, sobre todo para encuestas de hogares, en países en desarrollo y en transición. Se incluyen temas como los marcos muestrales, el tamaño de la muestra, el muestreo estratificado en varias etapas, la estimación del campo y el análisis de encuestas. Además, se presenta una introducción sobre todas las fases del proceso de las encuestas —que se tratan con mayor detalle en el resto de la publicación— al mismo tiempo que se hace hincapié en la relación de cada una de estas fases con el proceso de diseño de la muestra.

Términos clave: Diseño muestral complejo, marco muestral, población objetivo, estratificación, conglomeración, unidad primaria de muestreo.

A. Introducción

1. Diseños muestrales para encuestas en países en desarrollo y en transición

1. En este capítulo se presenta un panorama general de las cuestiones relacionadas con el diseño de muestras nacionales para las encuestas de hogares en países en desarrollo y en transición. Como en toda la publicación, se hace especial hincapié en las encuestas de hogares. Las encuestas sobre empresas y sobre la agricultura no se incluyen de manera explícita, aunque gran parte del material es también de interés en ambos casos.

2. Los diseños muestrales para encuestas de hogares en países en desarrollo y en transición tienen muchas características en común. La mayor parte de las encuestas están basadas en diseños muestrales probabilísticos estratificados por zonas en varias etapas. Estos diseños se utilizan sobre todo para la elaboración de marcos y para agrupar entrevistas en conglomerados con el fin de reducir los costos. La selección de muestras normalmente se lleva a cabo dentro de los estratos (véase la sección B). Las unidades seleccionadas en la primera etapa, conocidas en las publicaciones sobre muestras de encuestas como unidades primarias de muestreo (UPM), se elaboran con frecuencia partiendo de las zonas de empadronamiento identificadas y utilizadas en un censo nacional anterior de población

y vivienda, aunque también puede ser a partir de las zonas urbanas o de las aldeas en las áreas rurales. En algunos países, entre los posibles candidatos para las UPM se incluyen las zonas de supervisión de los censos o los distritos administrativos o sus subdivisiones. Las unidades incluidas en cada UPM seleccionada se conocen con el nombre de unidades de la segunda etapa; las que se seleccionan posteriormente serían unidades de la tercera etapa, y así sucesivamente. En lo que respecta a los hogares de los países en desarrollo y en transición, las unidades de la segunda etapa suelen ser unidades de vivienda u hogares, y las unidades seleccionadas en la segunda etapa suelen ser las personas. En general, las unidades seleccionadas en la última etapa de un diseño por etapas se conocen como unidades últimas de muestreo.

3. A pesar de las muchas semejanzas antes mencionadas, los diseños muestrales para encuestas en países en desarrollo y en transición no son idénticos en todos los países, y pueden presentar divergencias con respecto, por ejemplo, a las poblaciones objetivo, el contenido y las metas, el número de estratos de diseño, el tamaño de la muestra dentro de las UPM y el número de UPM seleccionadas en cada estrato. Además, las poblaciones subyacentes pueden variar con respecto a sus tasas de incidencia para determinadas características de la población y el grado de heterogeneidad y la distribución de sus poblaciones específicas dentro de cada estrato y entre ellos.

2. Panorama general

4. El capítulo está organizado de la manera siguiente: en la sección A se ha hecho una introducción general. La sección B tiene como tema los diseños muestrales estratificados en varias etapas. En primer lugar se describe el muestreo con probabilidad proporcional al tamaño. Luego se presenta el concepto de efecto del diseño en el contexto del muestreo por conglomerados. A continuación se examinan las opciones más adecuadas para el número de UPM y el número de unidades de la segunda etapa (viviendas, hogares, personas, etcétera) dentro de las UPM. Entre los factores que se tienen en cuenta figuran los requisitos de precisión previamente especificados para las estimaciones de las encuestas y las consideraciones prácticas resultantes de la organización del trabajo de campo. En la sección C se consideran los marcos muestrales y los problemas asociados y se proponen algunas posibles soluciones. En la sección D se aborda el tema de la estimación del campo y los diversos planes de asignación que pueden considerarse para atender las exigencias contradictorias resultantes del deseo de conseguir estimaciones de alcance nacional y subnacional. En la sección E se considera la determinación del tamaño de la muestra necesario para atender los niveles de precisión previamente especificados en lo que respecta al error estándar y al coeficiente de variación de las estimaciones. En la sección F se plantea el análisis de los datos de las encuestas y, en particular, se hace hincapié en el hecho de que el análisis adecuado de los datos de las encuestas debe tener en consideración las características del diseño muestral que generó esos datos. En la sección G se ofrece un resumen de algunas cuestiones importantes en el diseño de encuestas de hogares en países en desarrollo y en transición. En el anexo se presenta un diagrama de flujo en el que se reproducen los principales pasos necesarios en el proceso típico de una encuesta y las interrelaciones de los diferentes pasos del proceso.

B. Muestreo estratificado en varias etapas

5. La mayor parte de las encuestas realizadas en países en desarrollo y en transición están basadas en diseños por conglomerados estratificados en varias etapas. Las razones son dos. En primer lugar, la ausencia o mala calidad de las listas de hogares o direcciones obliga a seleccionar una muestra de unidades geográficas, y luego a elaborar listas de hogares o direcciones únicamente dentro de esas unidades seleccionadas. A partir de esas

listas podrán seleccionarse las muestras de hogares. En segundo lugar, el uso de diseños en varias etapas controla el costo de la recopilación de datos. En esta sección, examinamos los aspectos estadísticos y prácticos de las diversas etapas de un diseño típico por etapas.

1. Estratificación explícita

6. La estratificación suele aplicarse en cada una de las etapas del muestreo. No obstante, sus beneficios son especialmente notables en el muestreo de las UPM. Por ello es importante estratificar las UPM eficientemente antes de seleccionarlas.

7. La estratificación divide las unidades de la población en estratos o subgrupos mutuamente excluyentes y colectivamente exhaustivos. Luego, de cada estrato se seleccionan muestras independientes. Uno de los objetivos primarios de la estratificación es mejorar la precisión de las estimaciones de las encuestas. En ese caso la formación de los estratos debería ser tal que las unidades de un mismo estrato sean lo más homogéneas posible, y las unidades de los diferentes estratos, lo más heterogéneas posible con respecto a las características de interés para la encuesta. Otros beneficios de la estratificación son: la conveniencia administrativa y la flexibilidad, y la representación garantizada de los campos importantes y de las subpoblaciones especiales.

8. La experiencia obtenida del diseño muestral y del análisis de datos en muchos países ha revelado grandes diferencias en la distribución de las características de la población entre las distintas regiones administrativas y entre las áreas urbanas y rurales de cada país (véanse los capítulos XXII, XXIII y XXV de la presente publicación, donde se presentan ejemplos concretos). Ésta es una de las razones por las que, en lo que se refiere a las encuestas en esos países, los estratos explícitos están generalmente basados en regiones administrativas y áreas urbanas y rurales dentro de las regiones administrativas. Algunas regiones administrativas, como las capitales nacionales, quizá no tengan un componente rural, mientras que otras quizá no tengan un componente urbano. Es aconsejable examinar la distribución de frecuencias de los hogares y personas entre estos campos antes de concluir la elección de los estratos muestrales explícitos.

9. En algunos casos se desean estimaciones no sólo de alcance nacional sino también, por separado, acerca de cada región o subregión administrativa, como provincias, departamentos o distritos. La estratificación puede utilizarse para controlar la distribución de la muestra teniendo en cuenta esos campos de interés. Por ejemplo, en las encuestas demográficas y de salud (DHS) examinadas en el capítulo XXII los estratos iniciales están basados en las regiones administrativas cuyas estimaciones se desean. Dentro de la región se consigue una mayor estratificación mediante los componentes urbano/rural u otros tipos de subdivisión administrativa. Se imponen tasas de muestreo desproporcionadas entre los diferentes campos para garantizar la precisión adecuada en las estimaciones del campo. En general, la demanda de datos fiables sobre muchos campos requiere muestras generales de gran tamaño. La cuestión de la estimación del campo se examina en la sección D.

2. Estratificación implícita

10. Dentro de cada estrato explícito, al seleccionar las UPM se utiliza muchas veces una técnica conocida con el nombre de estratificación implícita. Antes de la selección de la muestra, las UPM de un estrato explícito se clasifican con respecto a una o más variables que se consideran fuertemente correlacionadas con la variable que interesa, y que se encuentran en todas las UPM del estrato. Luego se selecciona una muestra sistemática de UPM. La estratificación implícita garantiza que la muestra de UPM se distribuya entre las diferentes categorías de las variables de estratificación.

11. En el caso de muchas encuestas de hogares en países en desarrollo y en transición la estratificación implícita está basada en la ordenación geográfica de las unidades dentro de los estratos explícitos. Las variables de la estratificación implícita utilizadas a veces para la selección de las UPM incluyen la zona de residencia (de bajo ingreso, ingreso moderado, ingreso alto), la categoría de gasto (normalmente en quintiles), el grupo técnico y la zona de residencia en las áreas urbanas, y las zonas cultivadas, el número de aves de corral o de ganado vacuno que se posee, proporción de trabajadores no agrícolas, etcétera, en las zonas rurales. En las encuestas socioeconómicas las variables de la estratificación implícita incluyen la proporción de hogares clasificados como pobres, la proporción de adultos con estudios secundarios o superiores, y la distancia hasta el centro de una gran ciudad. Las variables utilizadas para la estratificación implícita suelen obtenerse normalmente de los datos del censo.

3. Selección muestral de las unidades primarias de muestreo

Características que deberían reunir las unidades primarias de muestreo

12. En el caso de las encuestas de hogares en los países en desarrollo y en transición, las UPM suelen ser unidades geográficas pequeñas dentro de los estratos. Si se dispone de información del censo, las UPM pueden ser las zonas de empadronamiento identificadas y utilizadas en el censo. A veces se utilizan también áreas semejantes o listas de la población local. En las zonas rurales, las aldeas pueden convertirse en UPM. En las zonas urbanas, las UPM algunas veces coinciden con los distritos o bloques.

13. Como las UPM afectan a la calidad de todas las fases posteriores del proceso de las encuestas, es importante garantizar que las unidades designadas como UPM sean de buena calidad y que se seleccionen para la encuesta de forma razonablemente eficiente. Para que las UPM puedan considerarse de buena calidad, deben, en general:

a) Tener límites claramente identificables que sean estables a lo largo del tiempo;
b) Abarcar por completo la población objetivo;
c) Ser de tamaño adecuado para el muestreo;
d) Contar con datos para la estratificación;
e) Que haya una cantidad elevada de ellas.

14. Antes de la selección de la muestra es preciso evaluar la calidad del marco muestral. Para un marco de zonas de empadronamiento, un primer paso es examinar los recuentos censales desglosados por campos de interés. En general, debería prestarse considerable atención al carácter de la UPM y a la distribución de los hogares e individuos entre las UPM en relación con toda la población y con los campos de interés. Un examen atento de estas distribuciones permitirá orientar las decisiones sobre la elección de UPM e identificará las unidades que necesitan ajuste para cumplir las especificaciones de una buena UPM. En general, una gran variabilidad en el número de hogares y personas entre las diferentes UPM y a lo largo del tiempo tendría efectos negativos en la organización del trabajo de campo. Si las UPM se seleccionan con el mismo grado de probabilidad, ello tendría también efectos negativos en la precisión de las estimaciones de las encuestas.

15. Muchas veces las opciones naturales para las UPM son inservibles porque carecen de uno o varios de los elementos mencionados. Dichas UPM deben modificarse o ajustarse antes de su utilización. Por ejemplo, si se considera que los límites de las zonas de empadronamiento no están bien definidos, podrían utilizarse como UPM unidades más grandes y más claramente definidas, como los distritos administrativos, las aldeas o comunas. Además, las UPM que se consideran demasiado grandes a veces se subdividen, o en otros casos se tratan como estratos, conocidos con frecuencia como selecciones de certeza o UPM autorrepresentativas (véase Kalton, 1983). Las UPM pequeñas normalmente se combinan

con las circundantes a fin de cumplir el requisito de contar con un número mínimo previamente especificado de hogares por UPM. El ajuste de las UPM de tamaño excesivamente grande o pequeño conviene realizarlo antes de la selección de las muestras.

16. Para garantizar una distribución equitativa de los hogares incluidos en la muestra dentro de las UPM, las UPM muy grandes a veces se dividen en varias subunidades de tamaño razonable, una de las cuales se selecciona al azar para ulteriores actividades de campo; por ejemplo, la lista de hogares. Es lo que se conoce con el nombre de *constitución de bloques* o *segmentación*. Conviene tener en cuenta que la selección y la segmentación de las UPM de tamaño excesivamente grande introducen una fase adicional de muestreo que debe tenerse en cuenta en el proceso de ponderación.

17. Las UPM muy pequeñas pueden combinarse también con UPM próximas en el marco de las UPM para cumplir el requisito del tamaño mínimo previamente especificado. No obstante, el trabajo que supone la combinación de pequeñas UPM se reduce considerablemente si la agrupación se lleva a cabo durante o después de la selección de las UPM. En cualquier caso, se trata de un proceso tedioso que requiere el seguimiento de normas estrictas y un gran esfuerzo de registro. En Kish (1965) se describe un procedimiento para combinar las UPM durante la selección de las muestras o una vez terminada. Una desventaja de este procedimiento es que no garantiza que las UPM seleccionadas para la agrupación sean contiguas, por lo que no es recomendable cuando haya un número considerable de UPM de tamaño suficiente.

Problemas relacionados con la falta de precisión en la medida del tamaño. Posibles soluciones

18. Uno de los problemas más frecuentes con los marcos de las zonas de empadronamiento que se utilizan como UPM —como suele ocurrir en los países en desarrollo y en transición— es que las medidas del tamaño pueden ser muy inexactas. Esas medidas suelen consistir en general en recuentos del número de personas o de hogares de las UPM, tomando como base el último censo de población. En algunos casos éstos están muy desfasados y pueden alejarse bastante de la realidad debido a factores como el crecimiento de las zonas urbanas y la contracción de otras áreas por efecto de la migración, las guerras y las catástrofes naturales. La medición inexacta del tamaño da lugar a una falta de control sobre la distribución de las unidades y el tamaño de las submuestras, lo que puede provocar graves problemas en las posteriores actividades sobre el terreno. Una solución es realizar una operación exhaustiva de enumeración para crear un marco de hogares en UPM seleccionadas antes de seleccionar los hogares. Otra solución es elegir las UPM con probabilidad proporcional al tamaño estimado. Ambos procedimientos se explican con mayor detalle en las secciones 4 y 5, *infra*. Otros problemas frecuentes asociados con la utilización de las zonas de empadronamiento como UPM son la falta de mapas de buena calidad y la cobertura incompleta de la población objeto, que es uno de los problemas relacionados con el marco muestral que se examinan en la sección C.

4. Muestreo de UPM con probabilidad proporcional al tamaño

19. Antes de la selección de las muestras, las UPM se estratifican expresa e implícitamente utilizando algunas de las variables enumeradas en los apartados B.1 y B.2 *supra*. En la mayor parte de las encuestas de hogares en los países en desarrollo y en transición las UPM se seleccionan con probabilidad proporcional a una medida del tamaño. Antes de la selección de las muestras se asigna a cada UPM una medida del tamaño, normalmente basada en el número de hogares o personas registradas durante el último censo o como consecuencia de una actualización reciente. Luego se selecciona una muestra independiente de las UPM dentro de cada estrato explícito con probabilidad proporcional a la medida del tamaño asignada.

20. El muestreo de probabilidad proporcional al tamaño (PPT) es una técnica que emplea datos auxiliares para aumentar de manera espectacular la precisión de las estimaciones de las encuestas, en particular si las medidas del tamaño son exactas y las variables de interés están correlacionadas con el tamaño de la unidad. Es la metodología preferida para el muestreo de las UPM en la mayor parte de las encuestas de hogares. El muestreo PPT arroja probabilidades desiguales de selección para las UPM. Fundamentalmente, la medida del tamaño de la UPM determina su probabilidad de selección. No obstante, cuando se combina con una fracción de submuestra adecuada para seleccionar los hogares dentro de las UPM seleccionadas puede dar lugar a una muestra autoponderada general de los hogares en la que todos los hogares tienen la misma probabilidad de selección, independientemente de las UPM en que estén ubicados. El principal atractivo del muestreo PPT es que puede dar lugar a muestras de tamaños aproximadamente iguales por UPM.

21. En las encuestas de hogares, un buen ejemplo de variable del tamaño PPT para la selección de las UPM es el número de hogares. Ciertamente el número de hogares de una UPM cambia a lo largo del tiempo y puede estar desfasado en el momento de selección de las muestras, pero hay varias formas de superar este problema, como se menciona en el párrafo 18. En las encuestas sobre explotaciones agrícolas, una medida del tamaño PPT que se utiliza con frecuencia es la superficie de la explotación. Esta elección se debe en parte a que los parámetros típicos de interés en esas encuestas, como el ingreso, la producción agrícola, las existencias de ganado y los gastos, están correlacionados con el tamaño de la explotación. En las encuestas sobre empresas las medidas PPT del tamaño suelen incluir el número de empleados, el número de establecimientos y el volumen anual de ventas. Como en el caso de la cantidad de de hogares, estas medidas PPT del tamaño también cambiarán probablemente a lo largo del tiempo, y este hecho debe tenerse en cuenta en el proceso de diseño.

22. Tomemos como ejemplo una muestra de hogares obtenida a partir de un diseño en dos etapas, con a UPM seleccionadas en la primera etapa y una muestra de hogares en la segunda etapa. Supongamos que la medida del tamaño (por ejemplo, el número de hogares en el momento del último censo) de la i^a UPM sea M_i. Si las UPM se seleccionan con PPT, la probabilidad P_i de seleccionar la i^a UPM viene dada por

$$P_i = a \times \frac{M_i}{\sum_i M_i}$$

23. Ahora supongamos que $P_{j|i}$ denota la probabilidad condicional de seleccionar el hogar j^o de la i^a UPM, dado que la i^a UPM se seleccionó en la primera etapa. En ese caso la ecuación de la selección para la probabilidad incondicional P_{ij} de seleccionar el hogar j^o de la i^a UPM en este diseño es

$$P_{ij} = P_i \times P_{j|\mathrm{i}}$$

24. Si se desea una muestra de hogares con igual probabilidad con una fracción de muestro general de $f=P_{ij}$, en ese caso los hogares deben seleccionarse con la tasa adecuada, inversamente proporcional a la probabilidad de selección de las UPM en que están ubicados, es decir,

$$P_{j|i} = \frac{f}{P_i}$$

25. Si las medidas del tamaño de las UPM son tamaños verdaderos, y no hay ningún cambio en la medida del tamaño entre la selección de la muestra y la recopilación de datos y

si se seleccionan hogares *b* en cada una de las UPM incluidas en la muestra, obtenemos una muestra autoponderada de los hogares con una probabilidad de selección dada por

$$P_{ij} = a \times \frac{M_i}{\sum_i M_i} \times \frac{b}{M_i} = \frac{a \times b}{\sum_i M_i} = f$$

donde *f* es una constante.

26. El problema de este procedimiento es que las medidas verdaderas del tamaño casi nunca se conocen en la práctica. No obstante, muchas veces es posible obtener estimaciones válidas, como los recuentos de la población y de hogares a partir de un censo reciente o alguna otra fuente fiable. Ello nos permite aplicar el procedimiento conocido como muestreo de probabilidad proporcional al tamaño estimado (PPTE). Hay dos opciones para el muestreo PPTE en un diseño en dos etapas en que los hogares se seleccionan en la segunda etapa: *a*) seleccionar los hogares con una tasa fija en cada una de las UPM incluidas en la muestra, o *b*) seleccionar un número fijo de hogares por UPM muestreada.

27. El muestreo PPTE de hogares con una tasa fija se lleva a cabo como sigue. Supongamos que los valores verdaderos de la medida del tamaño son denotados por N_i, y supongamos que los valores M_i son estimaciones válidas de N_i. Luego, aplicamos la tasa de muestreo b/M_i a la i^a UPM para obtener un tamaño de muestra de

$$b_i = \frac{b}{M_i} \times N_i$$

28. Téngase en cuenta que el submuestreo dentro de las UPM con una tasa fija (inversamente proporcional a las medidas del tamaño de las UPM) supone la determinación de una muestra por cada UPM muestreada de manera que, junto con la probabilidad de selección de UPM, obtenemos una muestra de hogares de igual probabilidad, independientemente del tamaño efectivo de las UPM. No obstante, este procedimiento no permite controlar los tamaños de la submuestra, ni, por tanto, el tamaño general de la muestra. Se incluirán en la muestra más hogares de la UPM junto a un número de hogares mayor del previsto, y en cambio, menos hogares de las UPM, con un número de hogares inferior al previsto. Ello tiene repercusiones en la organización del trabajo de campo. Además, si las medidas del tamaño están tan desfasadas que la variación en las muestras realizadas es extrema, quizá sea necesario cambiar la tasa de muestreo con el fin de obtener tamaños de muestra que sean un poco más homogéneos en las diferentes UPM, lo que supondría cierto grado de alejamiento de un diseño autoponderado.

29. El segundo procedimiento —la selección de un número fijo de hogares por UPM— evita la desventaja de los tamaños de muestras variables por UPM pero, por otro lado, no produce una muestra autoponderada. No obstante, si las medidas del tamaño se actualizan inmediatamente antes de la selección muestral de las UPM, pueden ofrecer aproximaciones bastante adecuadas que darán lugar a una muestra de hogares aproximadamente autoponderada.

30. En resumen, aun cuando el submuestreo dentro de las UPM con una tasa fija tenga como objetivo obtener muestras autoponderadas, hay circunstancias en las que este método no produce una muestra autoponderada de hogares. Por otro lado, aun cuando la selección de un número fijo de hogares dentro de las UPM muchas veces no arroje muestras autoponderadas, hay circunstancias en que este método da lugar a muestras aproximadamente autoponderadas de hogares. Cuando se produce un distanciamiento del diseño autoponderado, las ponderaciones deben utilizarse para compensar las probabilidades de selección diferencial resultantes en las distintas UPM.

5. Selección muestral de hogares

31. Una vez terminada la selección muestral de las UPM, se lleva a cabo un procedimiento cuyo objetivo es enumerar todos los hogares o todas las unidades de habitación o vivienda en cada UPM seleccionada. Algunas veces las enumeraciones incluyen las unidades de habitación, y luego todos los hogares de las unidades de habitación seleccionadas se incluyen si la muestra contiene una unidad de habitación. El objetivo de esta enumeración es crear un marco muestral actualizado a partir del cual puedan seleccionarse los hogares. Nunca se insistirá demasiado en la importancia de realizar este paso de manera eficaz. La calidad de la enumeración es uno de los factores más importantes que repercuten en la cobertura de la población objetivo.

32. Antes de la selección muestral en cada una de las UPM muestreadas, los hogares enumerados pueden clasificarse con respecto a las variables geográficas y de otro tipo que se consideran fuertemente correlacionadas con las variables de interés de la encuesta (véase el apartado B.2). Luego los hogares se muestrean a partir de la lista ordenada mediante un procedimiento de muestreo sistemático de igual probabilidad. Como se indica en el apartado B.4, los hogares pueden seleccionarse dentro de las UPM muestreadas con tasas de muestreo que generen probabilidades generales iguales de selección para todos los hogares o con tasas que generen un número fijo de hogares muestreados en cada UPM. Las ventajas y desventajas de estos planteamientos se examinan en el apartado B.4.

33. Con frecuencia las unidades últimas de muestreo son los hogares, y se recopila información sobre hogares seleccionados y sobre todos los miembros de esos hogares. En el caso de los módulos especiales que comprenden los ingresos y los gastos y cuyas unidades de análisis son los hogares, muchas veces se selecciona un declarante bien informado como representante del hogar. En las materias consideradas delicadas para las personas de los hogares (por ejemplo, el abuso doméstico) se selecciona una muestra aleatoria de personas (frecuentemente una sola persona) dentro de cada hogar muestreado.

6. Número de hogares que deben seleccionarse por UPM

34. Las unidades primarias de muestreo constan de conjuntos de hogares que están geográficamente conglomerados. En consecuencia, los hogares del mismo conglomerado normalmente suelen ser más parejos en lo que respecta a las características de la encuesta (por ejemplo, ingreso, educación, ocupación, etcétera) que los hogares en general. La conglomeración reduce considerablemente el costo de la recopilación de datos, pero las correlaciones entre las unidades del mismo conglomerado sobrevaloran la varianza (reducen la precisión) de las estimaciones de la encuesta, en comparación con un diseño en el que los hogares no estén conglomerados. Por ello el problema que se plantea al diseñar las encuestas es conseguir el justo equilibrio entre la reducción de costos y la correspondiente pérdida de precisión asociada con la conglomeración.

35. La inflación de la varianza de las estimaciones muestrales atribuibles a la conglomeración contribuye al llamado efecto del diseño. Éste representa el factor por el que debe multiplicarse la varianza de una estimación basada en una muestra aleatoria simple del mismo tamaño para tener en cuenta las complejidades del diseño de la muestra real debido a la estratificación, conglomeración y ponderación y se define como coeficiente de la varianza de una estimación basada en el diseño complejo con respecto a otra basada en una muestra aleatoria simple del mismo tamaño. Véanse los capítulos VI y VII de la presente publicación y las referencias citadas en ellos, en donde se explica con mayor detalle los efectos del diseño y su utilización en el diseño de la muestra. Una expresión del efecto del diseño (debido a la conglomeración) en una estimación [por ejemplo, una media estimada $(\bar{y})$] sería aproximadamente la siguiente:

$$D^2(\bar{y}) = 1 + (b-1)\rho$$

donde $D^2(\bar{y})$ denota el efecto del diseño para una media estimada $(\bar{y})$; ρ es la correlación intraclase, y b es el número de hogares que deben seleccionarse de cada conglomerado; es decir, el tamaño medio de la muestra conglomerada. La correlación intraclase es una medida del grado de homogeneidad (con respecto a la variable pertinente) de las unidades que integran un conglomerado. Como las unidades del mismo conglomerado suelen ser semejantes entre sí, la relación intraclase es casi siempre positiva. En las poblaciones humanas, una correlación intraclase positiva puede deberse al hecho de que los hogares del mismo conglomerado pertenecen a la misma clase de ingreso, pueden compartir las mismas actitudes hacia los problemas del momento y con frecuencia se ven expuestos a las mismas condiciones ambientales (clima, enfermedades infecciosas, catástrofes naturales, etcétera).

36. Si no se tienen en cuenta los efectos del diseño en las estimaciones de los errores estándar, se puede llegar a una interpretación inválida de los resultados de las encuestas. Conviene señalar que la magnitud de $D^2(\bar{y})$ está directamente relacionada con el valor de b, el tamaño de la muestra conglomerada y la correlación intraclase (ρ). Con un valor fijo de ρ, el efecto del diseño aumenta en forma lineal con b. Así, para lograr un bajo nivel de efectos del diseño, conviene utilizar una muestra conglomerada lo más pequeña posible. En el cuadro II.1 puede verse cómo repercuten en el efecto del diseño el tamaño medio de la muestra y la correlación intraclase. Por ejemplo, con un tamaño medio de la muestra conglomerada b de 20 unidades de habitación por UPM y ρ igual a 0,05, el efecto del diseño es 1,95. En otras palabras, este diseño de la muestra conglomerada arroja estimaciones con la misma varianza que las de una muestra (aleatoria simple) no conglomerada de aproximadamente la mitad del número total de hogares. Cuando los valores de ρ son mayores, la pérdida de precisión es todavía mayor, como puede verse en la parte derecha del cuadro II.1.

Cuadro II.1
Efectos del diseño en combinaciones seleccionadas de tamaño de la muestra conglomerada y correlación intraclase

Conglomerado	Correlación intraclase (ρ)								
Tamaño de la muestra (*b*)	0,005	0,01	0,02	0,03	0,04	0,05	0,10	0,20	0,30
1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
10	1,05	1,09	1,18	1,27	1,36	1,45	1,90	2,80	3,70
15	1,07	1,14	1,28	1,42	1,56	1,70	2,40	3,80	5,20
20	1,10	1,19	1,38	1,57	1,76	1,95	2,90	4,80	6,70
30	1,15	1,29	1,58	1,87	2,16	2,45	3,90	6,80	9,70
50	1,25	1,49	1,98	2,47	2,96	3,45	5,90	10,80	15,70

37. En general, el número óptimo de hogares que deben seleccionarse en cada UPM dependerá de la estructura de costos de la recopilación de datos y del grado de homogeneidad o conglomeración con respecto a las variables de la encuesta dentro de la UPM. Supongamos un diseño en dos fases con UPM seleccionadas en la primera fase y hogares seleccionados en la segunda. Supongamos también un modelo de costo lineal para el costo global relacionado con el muestreo de las UPM y los hogares, en la forma siguiente

$$C = aC_1 + abC_2$$

donde C_1 y C_2 son, respectivamente, el costo de una UPM adicional y el costo de un hogar adicional, y a y b denotan, respectivamente, el número de UPM seleccionadas y el número de hogares seleccionados por UPM (Cochran, 1977, pág. 280). En este modelo de costos, la opción óptima para b que reduce la varianza de la media de la muestra (véase Kish, 1965, apartado 8.3.b) viene dada aproximadamente por

$$b_{opt} = \sqrt{\frac{C_1(1-\rho)}{C_2 \quad \rho}}$$

38. En el cuadro II.2 se presenta el tamaño óptimo de submuestra (b) para diferentes coeficientes de costo C_1/C_2 y correlación intraclase. Téngase en cuenta que, en igualdad de condiciones, el tamaño óptimo de la muestra disminuye (es decir, la muestra se distribuye más ampliamente en los diferentes conglomerados) a medida que aumenta la correlación intraclase y que el costo de un hogar adicional aumenta en relación con el de una UPM.

39. El modelo de costo en la derivación del tamaño óptimo del conglomerado está muy simplificado, pero probablemente puede servir como orientación general. Como la mayor parte de las encuestas tienen varios objetivos e implican diferentes variables y, por consiguiente, diferentes valores de ρ, la elección de b muchas veces supone un grado de compromiso entre las diferentes propuestas óptimas.

Cuadro II.2
Tamaños óptimos de las submuestras para determinadas combinaciones de coeficiente de costo y correlación intraclase

Coeficiente de costo	Correlación intraclase				
(C_1/C_2)	0,01	0,02	0,03	0,05	0,08
4	20	14	11	9	5
9	30	21	17	13	10
16	40	28	23	17	14
25	50	35	28	22	17

40. En ausencia de una información precisa sobre los costos puede utilizarse el cuadro II.2 para determinar el número óptimo de hogares que deben seleccionarse en un conglomerado para diversas opciones de coeficiente de costos y correlación intraclase. Por ejemplo, si se sabe *a priori* que el costo de incluir una UPM es cuatro veces mayor que el de incluir un hogar, y que la correlación intraclase para una variable de interés es 0,05, conviene seleccionar aproximadamente nueve hogares en el conglomerado. Téngase en cuenta que el número óptimo de hogares que deben seleccionarse en un conglomerado no depende del presupuesto global disponible para la encuesta. El presupuesto total determina únicamente el número de UPM que se deben seleccionar.

41. En general, entre los factores que deben considerarse al determinar la asignación de muestras entre las diferentes UPM y entre los hogares dentro de las UPM se incluyen la precisión de las estimaciones de las encuestas (mediante el efecto del diseño), el costo de la recopilación de datos y la organización del trabajo sobre el terreno. Si los gastos de viaje son elevados, como en el caso de las áreas rurales, es preferible seleccionar un reducido número de UPM y muchos hogares en cada UPM. Por otro lado, si, como ocurre en las zonas urbanas, los gastos de viaje son más bajos, es más eficiente seleccionar muchas UPM y un menor número de hogares dentro de cada una de ellas. Por el contrario, en las zonas rurales, quizá sea más eficiente seleccionar más hogares por UPM. Estas decisiones deben tomarse de tal manera que se consiga una distribución eficiente de la carga de trabajo entre los entrevistadores y supervisores.

C. Marcos muestrales

1. Características de los marcos muestrales para las encuestas en los países en desarrollo y en transición

42. En la mayor parte de las encuestas de hogares, la población objetivo comprende la población civil no institucionalizada. Con el fin de mantener los datos deseados de esta población objetivo, muchas veces se realizan entrevistas por hogares. En general, sólo pue-

den incluirse en ellas las personas consideradas como residentes permanentes del hogar. Los residentes permanentes que se encuentran ausentes de forma temporal, por ejemplo las personas que han ido de vacaciones o están internadas temporalmente en un hospital y los estudiantes que viven fuera de casa durante el año escolar, generalmente se incluyen si se selecciona su hogar. Los alumnos que viven fuera de casa durante el año escolar no se incluyen en la encuesta si se les incluye en la muestra de su residencia escolar, ya que los datos de dichos alumnos se obtendrían del lugar de residencia permanente. Los grupos que se encuentran generalmente excluidos de las encuestas de hogares en los países en desarrollo y en transición son los miembros de las fuerzas armadas que viven en cuarteles o en casas privadas; las personas que se encuentran en centros de reclusión, hospitales, residencias para ancianos u otras instituciones; las personas sin hogar, y los nómadas. La mayor parte de estos grupos se excluyen, en general, debido a las dificultades prácticas que se encuentran normalmente para recopilar datos de ellos. No obstante, la decisión de excluir o no a un grupo debe tomarse teniendo en cuenta los objetivos de la encuesta.

2. Problemas del marco muestral y posibles soluciones

43. Como en otros tipos de encuestas, la calidad de los datos obtenidos de las encuestas de hogares depende en gran medida de la calidad del marco muestral a partir del cual se seleccionó la muestra para la encuesta. Por desgracia, los problemas relacionados con los marcos muestrales son una característica inevitable de las encuestas de hogares. En esta sección se analizan algunos de estos problemas y se proponen posibles soluciones.

44. Kish (1965, parágrafo 2.7) presenta una clasificación útil de cuatro problemas, junto con las posibles soluciones. Los cuatro problemas son: la falta de cobertura, los conglomerados de elementos, los espacios en blanco y las listas duplicadas, que examinaremos en el contexto de los diseños por etapas para encuestas realizadas en países en desarrollo y en transición.

45. Por "falta de cobertura" se entiende la situación en que el marco muestral no abarca toda la población objetivo, como consecuencia de lo cual algunas unidades de muestreo no tienen ninguna posibilidad de inclusión en la muestra. Este problema es un gran motivo de preocupación en las encuestas de hogares realizadas en países en desarrollo y en transición. Los efectos de la falta de cobertura se ponen de manifiesto por el hecho de que las estimaciones muestrales de los recuentos de población basados en la mayor parte de las encuestas en los países en desarrollo y en transición arrojan estimaciones de población inferiores a las de otras fuentes.

46. Hay tres niveles de falta de cobertura: la UPM, el hogar y la persona. En los países en desarrollo y en transición, la no cobertura de las UPM es un problema menos grave que la de los hogares y las personas pertenecientes a los hogares de la muestra. La falta de cobertura de las UPM se produce, por ejemplo, cuando algunas regiones de un país quedan deliberadamente excluidas de una encuesta porque son inaccesibles, se encuentran en guerra, han sufrido una catástrofe natural o por otras causas. Asimismo, algunas veces las zonas remotas con muy pocos hogares o personas se eliminan de los marcos muestrales para las encuestas de hogares por la pequeña proporción de población que representan y porque no influyen mucho en las cifras correspondientes. La falta de cobertura es un problema más grave cuando se trata de los hogares y personas. Los hogares o personas pueden quedar excluidos erróneamente de la encuesta debido a problemas más complejos de definición y de concepto sobre la estructura y composición de los hogares. Existe la posibilidad de interpretaciones contradictorias de estos temas por diferentes entrevistadores o por las personas responsables de elaborar las listas de hogares y miembros de los hogares. Por ello se necesitan instrucciones prácticas estrictas para indicar a los entrevistadores quién debe ser considerado miembro del hogar y qué debe considerarse como un hogar o una unidad de habitación. Para

resolver este problema, la calidad de la lista de hogares y personas que deben incluirse dentro de los hogares debería constituir un objetivo fundamental de la actividad metodológica y de la capacitación en los países en desarrollo y en transición.

47. El problema de los **espacios en blanco** se plantea cuando algunas listas del marco muestral no contienen ningún elemento de la población objetivo. En el caso de una lista de unidades de habitación, un espacio en blanco correspondería a una vivienda vacía. Este problema se plantea también en los casos en que la muestra incluye subgrupos particulares de población; por ejemplo, las mujeres que habían dado a luz el año anterior. Algunos hogares que se enumeraron e incluyeron en la muestra no tienen ninguna mujer que hubiera dado a luz el año precedente. Si es posible, las casillas en blanco pueden eliminarse del marco antes de la selección de la muestra. No obstante, en muchas circunstancias prácticas ello es poco eficaz en función de los costos. Una solución más práctica es identificar y eliminar los espacios en blanco después de la selección de la muestra. No obstante, la eliminación de los espacios en blanco significa que la muestra realizada será menor y de tamaño variable.

48. El problema de las **listas duplicadas** se plantea cuando algunas unidades de la población objetiva aparecen más de una vez en el marco muestral. Este problema puede producirse, por ejemplo, con las muestras de nómadas o de residentes que viven en un lugar sólo parte del año. Una manera de evitar la duplicación es designar como lista efectiva una lista única previamente especificada y considerar las demás listas como espacios en blanco. Una unidad sólo se incluye en la muestra si se incluye en la muestra la lista única. Por ejemplo, los nómadas que pastorean su rebaño yendo de un lugar a otro en busca de tierra de pastos y de agua para sus animales pueden incluirse en la muestra mientras se desplazan hacia los pozos de agua. Según los ciclos de abrevado de los animales (al parecer, los caballos tienen ciclos más largos que el ganado vacuno), es probable que algunos visiten con mayor frecuencia los pozos durante el período de recopilación de datos. Para evitar las duplicaciones, podría identificarse a los nómadas una sola vez —con ocasión de su primera visita al pozo después de una fecha determinada—, mientras que las visitas posteriores quedarían como espacios en blanco. De lo contrario habría que ajustar las ponderaciones de las unidades incluidas en la muestra para tener en cuenta las duplicaciones. Véase Yansaneh (2003), donde se presentan ejemplos concretos de cómo hacerlo.

49. El problema de los **conglomerados de elementos** se plantea cuando una lista única del marco de muestreo consiste realmente en unidades múltiples de la población objetivo. Por ejemplo, una lista de viviendas puede contener algunas viviendas con más de un hogar. En tales casos la inclusión de todos los hogares asociados con esa vivienda arrojará una muestra en la que los hogares tendrán la misma probabilidad de selección que la vivienda. Conviene observar que la práctica de seleccionar aleatoriamente una de las unidades del conglomerado lleva automáticamente a probabilidades desiguales de selección, que deberían compensarse mediante la ponderación.

3. Mantenimiento y evaluación de los marcos muestrales

50. La construcción y mantenimiento de marcos muestrales adecuados constituye una actividad costosa y lenta. Los países en desarrollo y en transición tienen la posibilidad de crear dichos marcos a partir de fuentes como los datos de los censos decenales. Conviene que todos los institutos nacionales de estadística den gran prioridad a la creación y mantenimiento de un marco muestral maestro de las zonas de empadronamiento definidas y utilizadas en el censo anterior. Este marco debería establecerse poco después de haberse terminado el censo, ya que el volumen de trabajo implicado aumenta en proporción con el tiempo transcurrido desde el censo. El marco debe contar con etiquetas apropiadas de otras zonas geográficas quizá más amplias que puedan utilizarse como unidades primarias de muestreo. Debería incluir también datos que puedan ser útiles para la estratificación, como

la composición étnica y racial, el gasto mediano o los quintiles de gastos, etcétera. Si se mantiene debidamente, el marco muestral maestro puede utilizarse para un sistema integrado de encuestas, incluidas las encuestas repetidas. En el capítulo V puede verse información más detallada sobre la construcción y mantenimiento de marcos muestrales maestros.

D. Estimación del campo

1. Necesidad de estimaciones del campo

51. En los últimos años, en la mayoría de los países ha aumentado la demanda de datos fiables tanto de alcance nacional como subnacional, debido sobre todo al hecho de que la mayor parte de los programas de desarrollo o de intervención son de nivel subnacional, como la región administrativa o el distrito. Para adoptar decisiones importantes acerca de la ejecución de los programas o la asignación de los recursos en el plano local se requieren datos precisos en ese ámbito.

52. En este contexto, entendemos por campo todo subconjunto de la población para el que se planifican estimaciones independientes en el diseño de las encuestas. Un campo podría ser un estrato, una combinación de estratos, una región administrativa o urbana o rural u otras subdivisiones dentro de esas regiones. Por ejemplo, las estimaciones de muchas encuestas nacionales se publican por separado según las diferentes regiones administrativas. Las regiones podrían entonces tratarse como campos, cada una de ellas con dos estratos (por ejemplo, subpoblaciones urbana y rural) o más. Los campos pueden ser subpoblaciones demográficas definidas por características como la edad, la raza y el sexo. No obstante, se plantea una complicación cuando los campos desbordan los límites de los estratos, como ocurre, por ejemplo, cuando un campo está integrado por hogares con acceso a los servicios de salud.

53. Es importante que el número de campos de interés para una encuesta concreta sea moderado. El tamaño de la muestra necesario para presentar estimaciones fiables sobre cada uno de los campos, cuando éstos son muy numerosos, sería necesariamente muy considerable. Los problemas asociados con las muestras de gran magnitud se examinan en la parte E.

2. Asignación de muestras

54. La presentación de estimaciones precisas sobre los ámbitos de interés requiere la asignación de muestras de tamaño adecuado a los diferentes campos. No obstante, se plantean conflictos cuando se desea el mismo grado de precisión para campos con tamaños de población muy diversos. Si se desean estimaciones con el mismo nivel de precisión para todos los campos (es decir, el mismo tamaño muestral por campo), la estrategia más eficiente es una asignación igual. No obstante, esta asignación puede provocar una grave pérdida de eficacia en las estimaciones nacionales. La asignación proporcional, que utiliza fracciones muestrales iguales en cada campo, suele ser la asignación más indicada para las estimaciones nacionales. Cuando los campos presentan grandes diferencias de tamaño y cuando se necesitan estimaciones tanto nacionales como por campos, es preciso llegar a una solución de compromiso entre la asignación igual y las fracciones muestrales iguales.

55. Kish (1988), propuso un compromiso entre la asignación proporcional y la igual, basado en una asignación proporcional a $n\sqrt{(W_h^2 + H^{-2})}$, en donde n es el tamaño de la muestra general, W_h es la proporción de la población en el estrato h y H es la cantidad de estratos. En los estratos muy pequeños el segundo término predomina sobre el primero, lo que impediría asignaciones a los pequeños estratos de tamaño demasiado reducido.

56. Un planteamiento alternativo es aumentar el tamaño de la muestra de los campos menores en la medida necesaria para atender los requisitos de precisión exigidos. Cuando

un campo es pequeño, la asignación proporcional dará lugar a un tamaño de la muestra para ese campo que puede ser demasiado pequeño para generar estimaciones lo bastante precisas. La solución es el sobremuestreo, o una muestra con una tasa superior, a partir de los campos menores.

57. En resumen, los encargados del diseño de las encuestas en los países en desarrollo y en transición muchas veces deben elegir entre realizar estimaciones precisas de alcance nacional o estimaciones precisas para los diferentes campos. Este problema se agrava cuando los campos de interés tienen tamaños muy diferentes. Una manera de eludir este dilema es definir los campos que son de tamaño aproximadamente igual, quizá combinando los campos existentes. Otra posibilidad es mantener separados los campos y permitir un nivel menor de precisión para los campos pequeños, o quizá no publicar ninguna estimación sobre los campos.

E. Tamaño de la muestra

1. Factores que influyen en las decisiones sobre el tamaño de la muestra

58. Tanto los productores como los usuarios de datos de encuestas prefieren muchas veces muestras de gran tamaño, un requisito que consideran necesario para que sea más "representativa", y para reducir los errores de muestreo y, por lo tanto, aumentar la fiabilidad de las estimaciones. Este argumento se expone sin tener casi en cuenta el posible aumento de los errores no muestrales que acompaña a las muestras de gran tamaño. A continuación examinamos los factores que deben tenerse en cuenta al determinar el tamaño adecuado de la muestra para una encuesta.

59. Los tres factores principales que deben orientar las decisiones acerca del tamaño adecuado de la muestra para una encuesta son los siguientes:

- Precisión (fiabilidad) de las estimaciones de la encuesta;
- Calidad de los datos recopilados por la encuesta;
- Costo, en tiempo y dinero, de la recopilación, elaboración y divulgación de los datos.

A continuación se examinan estos factores.

2. Precisión de las estimaciones de la encuesta

60. Los objetivos de la mayor parte de las encuestas en los países en desarrollo y en transición son, entre otros, la estimación del nivel de una característica (por ejemplo, la proporción de hogares clasificados como pobres) en un determinado momento y de la variación de ese nivel a lo largo del tiempo (por ejemplo, la variación de la tasa de pobreza entre dos puntos cronológicos diferentes). Examinamos la precisión de las estimaciones de las encuestas en el contexto de la estimación del nivel de una característica en un momento dado. Para el resto del análisis, utilizaremos como característica de interés el porcentaje de los hogares en situación de pobreza, que denominaremos "tasa de pobreza".

61. La precisión de una estimación se mide por su error estándar. La fórmula para el error estándar estimado de una tasa de pobreza estimada p en un campo dado, denominado $se(p)$, viene dada por

$$se(p) = \sqrt{d^2(p) \times (1 - \frac{n}{N}) \times \frac{p(100-p)}{n}}$$

donde n es el número global de hogares para el campo de interés; N es el número total de hogares en el campo, y $d^2(p)$ es el efecto del diseño estimado asociado con el diseño de la en-

cuesta compleja[1]. La proporción de la población que está en la muestra, n/N, recibe el nombre de fracción de muestreo, y el factor $[1-(n/N)]$ (proporción de la población no incluida en la muestra) se conoce como factor de corrección de población finita (cpf), que representa el ajuste efectuado en el error estándar de la estimación para tener en cuenta el hecho de que la muestra se selecciona sin sustitución a partir de una población finita.

62. Usaremos como ejemplo datos de Viet Nam. El número total de hogares, N, según el censo de población de 1999 es de 16.661.366. Véase en Glewwe y Yansaneh (2000) información más detallada sobre la distribución de los hogares tomando como base el censo de 1999. Téngase en cuenta que con un tamaño de población tan considerable el factor de corrección de población finita es insignificante en todos los casos. En el cuadro II.3 se presentan los errores estándar y los intervalos de confianza del 95% para las diversas estimaciones de la tasa de pobreza, suponiendo un efecto del diseño del 2,0. Un intervalo de confianza del 95% significa que hay un 95% de probabilidad de contener el valor verdadero. En el cuadro puede verse que con un tamaño de muestra dado los errores estándar aumentan a medida que crece la tasa de pobreza, alcanzando un máximo para $p = 50\%$. Los intervalos de confianza del 95% asociados aumentan también a medida que crece la tasa de pobreza, y alcanza su expresión mayor cuando la tasa de pobreza es del 50%. Así pues, en general, los campos con tasas de pobreza mucho menores o mayores del 50% tendrán estimaciones más precisas en relación con los campos con tasas de pobreza cercanas al 50%, con un tamaño de muestra y un efecto de diseño determinados[2]. Ello significa que los campos con tasas muy bajas o muy altas de pobreza requerirán un tamaño de muestra menor para conseguir el mismo error estándar que un campo con una tasa de pobreza próxima al 50%. Tomemos como ejemplo un tamaño de muestra de 500 hogares en un campo. Si este campo tiene una tasa de pobreza estimada de sólo el 5%, el intervalo de confianza es 5 ± 2,7%; con una tasa del 10%, el intervalo es 10 ± 3,7%; con una tasa del 25%, el intervalo es 25 ± 5,4%; finalmente, con una tasa de pobreza estimada del 50%, el intervalo es 50 ± 6,2%.

63. Naturalmente, si el tamaño de la muestra asciende a más de 500 hogares la amplitud del intervalo de confianza disminuye (en otras palabras, la estimación de la muestra resulta más precisa). No obstante, la reducción de esa amplitud es proporcional no al aumento del tamaño de la muestra sino a la raíz cuadrada de dicho aumento, en este caso $\sqrt{n/500}$, donde n es el nuevo tamaño de la muestra. Por ejemplo, en un campo con una tasa de pobreza del 25%, si el tamaño de la muestra se duplica, pasando de 500 a 1.000 hogares, se reduciría la amplitud del intervalo de confianza a razón de $\sqrt{2}$, es decir, pasaría de ± 5,4% a ± 3,8%. Estas reducciones deberían sopesarse atentamente teniendo también en cuenta la creciente complejidad de la gestión de las actividades de las encuestas, los costos de las mismas y los errores no muestrales.

[1] Aunque en la fórmula anterior n debería ser $n-1$ para $se(p)$, en la mayor parte de las aplicaciones prácticas n es lo bastante grande para que la diferencia entre n y $n-1$ sea insignificante

[2] En el caso de las tasas de pobreza superiores al 50% ($p > 50\%$), el error estándar es el mismo que para una tasa de pobreza de $100 - p$, y por lo tanto puede deducirse a partir del cuadro III.3. Por ejemplo, el error estándar de una tasa de pobreza estimada del 75% es el mismo que el de una tasa de pobreza estimada del 25%

Cuadro II.3
Errores estándar e intervalos de confianza para estimaciones de la tasa de pobreza basadas en varios tamaños de muestra, suponiendo que el efecto del diseño es 2,0

	Tasa de pobreza *(porcentaje)*									
	5		10		25		40		50	
Tamaño de la muestra	Error estándar	Intervalo de confianza	Error estándar	Intervalo de confianza	Error estándar	Intervalo de confianza	Error estándar	Intervalo de confianza	Error estándar	Intervalo de confianza
250	1,95	(1,2; 8,8)	2,68	(4,7; 15,3)	3,87	(17,4; 32,6)	4,38	(31,4; 48,6)	4,47	(41,2; 58,8)
500	1,38	(2,3; 7,7)	1,90	(6,3; 13,7)	2,74	(19,6; 30,4)	3,10	(33,9; 46,1)	3,16	(43,8; 56,2)
750	1,13	(2,8; 7,2)	1,55	(7,0; 13,0)	2,24	(20,6; 29,4)	2,53	(35,0; 45,0)	2,58	(44,9; 55,1)
1 000	0,97	(3,1; 6,9)	1,34	(7,4; 12,6)	1,94	(21,2; 28,8)	2,19	(35,7; 44,3)	2,24	(45,6; 54,4)
1 500	0,80	(3,4; 6,6)	1,10	(7,9; 12,1)	1,58	(21,9; 28,1)	1,79	(36,5; 43,5)	1,83	(46,4; 53,6)
2 000	0,44	(4,1; 5,9)	0,95	(8,1; 11,9)	1,37	(22,3; 27,7)	1,55	(37,0; 43,0)	1,58	(46,9; 53,1)

64. La precisión de las estimaciones de las encuestas se expresa muchas veces en términos del coeficiente de variación de la estimación del interés. Como en el caso anterior, consideramos únicamente la estimación del porcentaje de hogares clasificados como pobres en un país. El coeficiente estimado de variación de una estimación de la tasa de pobreza, designado por $cv(p)$, viene dado por

$$cv(p) = \frac{se(p)}{p} = \sqrt{d^2(p) \times (1 - \frac{n}{N}) \times \frac{(100 - p)}{np}}$$

65. En el cuadro II.4 se presentan los coeficientes estimados de variación para una tasa de pobreza estimada de muestras de diversos tamaños, suponiendo un efecto del diseño de 2,0, donde *cv* se expresa en forma de porcentaje. En el cuadro se observa que con un tamaño de muestra dado el coeficiente estimado de variación de la tasa de pobreza estimada disminuye de forma constante a medida que aumenta el porcentaje verdadero. Asimismo, con una tasa de pobreza dada el coeficiente de variación disminuye a medida que desciende el tamaño de la muestra. Con una muestra de 500, el coeficiente de variación es de aproximadamente 28% cuando p = 5%, 19% cuando p = 10%, 11% cuando p = 25%, 8% cuando p = 40%, 6% cuando p = 50%, 5% cuando p = 60%, 4% cuando p = 75%, 2% cuando p = 90%, y 1% cuando p = 95%. A medida que aumenta el tamaño de la muestra, el coeficiente estimado de variación disminuye en la forma correspondiente. Téngase en cuenta que a diferencia de los errores estándar que se presentan en el cuadro II.3, el coeficiente de variación del cuadro II.4 no es una función simétrica de la tasa de pobreza.

Cuadro II.4
Coeficiente de variación para estimaciones de la tasa de pobreza basadas en diversos tamaños de muestra, suponiendo que el efecto del diseño es 2,0

Tamaño de la muestra	Tasa de pobreza *(porcentaje)*								
	5	10	25	40	50	60	75	90	95
250	39	27	15	11	9	7	5	3	2
500	28	19	11	8	6	5	4	2	1
750	23	15	9	6	5	4	3	2	1
1 000	19	13	8	5	4	4	3	1	1
1 500	16	11	6	4	4	3	2	1	1
2 000	14	9	5	4	3	3	2	1	1

3. Calidad de los datos

66. Una consideración importante al determinar el tamaño de la muestra para una encuesta es la calidad de los datos que se van a recopilar. Es importante mantener datos de la mayor calidad posible para poder tener confianza en las estimaciones generadas a partir de ellos. Es imprescindible comprobar la calidad de los datos en todas las fases de realización de la encuesta. En consecuencia, conviene mantener el tamaño de la muestra dentro de límites razonables para que la comprobación y edición puedan realizarse de manera eficiente tanto en lo que respecta al tiempo como a los recursos.

67. Un factor relacionado con el tamaño de la muestra que influye en la calidad de los datos es el número de personas que trabajan en el estudio. Por ejemplo, los tamaños de muestra más pequeños requieren menos entrevistadores, lo que significa que éstos se pueden elegir de forma más selectiva. En particular, con un tamaño más pequeño es más probable que los entrevistadores estén mejor preparados y tengan más experiencia. Además, los entrevistadores estarán mejor entrenados, ya que al ser menos que los requeridos, la capitación

puede organizarse mejor y se puede destinar a ella un volumen proporcionalmente mayor de recursos. Además, se necesitará menos material de capacitación y los entrevistadores recibirán mayor atención individual durante la capacitación y sobre el terreno. Todo ello tendrá como resultado una cantidad menor de problemas en la recopilación de datos y en su posterior edición. Por consiguiente, los datos disponibles para el análisis tendrán más calidad, lo que permitirá a las autoridades confiar más en las decisiones adoptadas sobre la base de esos datos.

68. Además de las preocupaciones por la calidad de los datos recopilados, las muestras de mayor tamaño hacen que sea más difícil y costoso reducir los casos de ausencia de respuesta (véase el capítulo VIII). Es importante conseguir el nivel más bajo posible de falta de respuesta para reducir la posibilidad de grandes sesgos en las estimaciones de la encuesta (véase el apartado F.1). Estos sesgos podrían producirse si no conseguimos garantizar las respuestas de una parte apreciable de la población que puede ser considerablemente diferente de la incluida en la encuesta. Por ejemplo, las personas que viven en zonas urbanas y tienen ingresos relativamente altos muchas veces tienen menos probabilidad de participar en las encuestas de hogares. La no inclusión de un gran segmento de esta parte de la población puede dar lugar a una subestimación de características de la población como el promedio nacional de ingreso de los hogares, el nivel de instrucción y la alfabetización. Con una muestra más pequeña será mucho más fácil y más eficaz en función de los costos revisar los hogares que inicialmente decidan no participar y convencerlos de que lo hagan. Como el intento de persuadir a quienes inicialmente decidieron no participar puede resultar costoso y lento, es importante para la calidad de los datos de la encuesta asignar recursos adecuados y tiempo suficiente para convencer a quienes inicialmente se negaron a participar.

4. Costo y oportunidad

69. El tamaño de la muestra de una encuesta repercute claramente en su costo. En general, el costo global de una encuesta depende de los costos generales fijos y los costos variables asociados con la selección y procesamiento de cada unidad de la muestra en las diferentes fases de selección. Por ello, cuanto mayor sea la muestra, mayor será el costo de realización de la encuesta. En el capítulo XII se examinan con más detalle los componentes del costo de las encuestas de hogares. En los capítulos XIII y XIV se presentan ejemplos empíricos de cálculos de costo en encuestas concretas.

70. El tamaño de la muestra puede determinar también el punto cronológico en que los datos podrían ser objeto de análisis. Es importante que los datos y las estimaciones de la encuesta se presenten de forma oportuna, para que las decisiones normativas puedan basarse en datos razonablemente actualizados. Cuanto mayor sea la muestra, más tiempo se necesitará para depurar, editar y ponderar los datos necesarios para el análisis.

F. Análisis de la encuesta

1. Elaboración y ajuste de las ponderaciones de muestreo

71. Las ponderaciones muestrales son necesarias para compensar las probabilidades de selección desigual, la ausencia de respuesta y las diferencias conocidas entre la muestra y la población de referencia. Las ponderaciones deberían utilizarse en la estimación de las características de población de interés y también en la estimación de los errores estándar de las estimaciones generadas.

72. La ponderación básica de la unidad incluida en la muestra puede concebirse como el número de unidades de la población que están representadas por la unidad de la muestra a los efectos de la estimación. Por ejemplo, si la tasa de muestreo dentro de un estrato determinado

es 1 de 10, la ponderación básica de toda unidad incluida en la muestra del estrato es 10, es decir, la unidad de la muestra representa 10 unidades de la población, incluida la unidad misma.

73. La elaboración de ponderaciones de muestreo normalmente comienza con la elaboración de las ponderaciones básicas para las unidades de la muestra, con el fin de corregir la desigualdad en sus probabilidades de selección. En general, la ponderación básica de una unidad es el recíproco de su probabilidad de selección para incluirla en la muestra. En el caso de los diseños por etapas, la ponderación básica debe reflejar la probabilidad de selección en cada etapa. Las ponderaciones básicas de las unidades incluidas en la muestra se ajustan luego para compensar la ausencia de respuesta y de cobertura y para hacer que las estimaciones muestrales ponderadas correspondan a los totales conocidos de la población.

74. Cuando las ponderaciones ajustadas finales de todas las unidades de la muestra son iguales se dice que la muestra es autoponderada. En la práctica, las muestras no son autoponderadas por varias razones. En primer lugar, las unidades de muestreo se seleccionan con probabilidades desiguales de selección. De hecho, aun cuando las UPM se seleccionen muchas veces con probabilidad proporcional al tamaño y los hogares se seleccionen con una tasa adecuada dentro de la UPM para obtener un diseño autoponderado, todo ello puede anularse mediante la selección de una única persona para ser entrevistada en cada uno de los hogares de la muestra. En segundo lugar, la muestra seleccionada muchas veces presenta deficiencias, como la falta de respuesta y de cobertura, debido a problemas con el marco muestral (véase la sección C). En tercer lugar, la necesidad de estimaciones precisas sobre los campos y subpoblaciones especiales requiere con frecuencia el sobremuestreo de estos campos (véase la sección D).

75. Como ya se ha indicado, son pocos los casos en que se obtiene toda la información deseada de todas las unidades de la muestra. Por ejemplo, es posible que algunos hogares no faciliten ningún dato, o que faciliten sólo datos parciales, es decir, datos sobre algunas pero no todas las preguntas de la encuesta. El primer tipo de falta de respuesta se conoce como falta total de respuesta o de la unidad, mientras que la segunda se conoce como falta de respuesta parcial, es decir, sobre una partida. Si hay diferencias sistemáticas entre informantes y no informantes, las estimaciones acríticas basadas únicamente en quienes responden estarán sesgadas. Para reducir este potencial de sesgo muchas veces se introducen ajustes en el análisis con el fin de compensar la falta de respuesta. El método estándar para compensar la falta de respuesta sobre una partida es la imputación, que no se considera en el presente capítulo. En Yansaneh, Wallace y Marker (1998) y en la bibliografía de esa obra puede verse un examen general de los métodos de imputación y su aplicación a las encuestas grandes y complejas.

76. En lo que respecta a la falta total de respuesta, hay tres procedimientos básicos de compensación:

- Ajuste de las ponderaciones básicas;
- Selección de una muestra inicial mayor de lo necesario, para permitir una posible reducción del tamaño de la muestra debido a la falta de respuesta;
- Sustitución; es decir, un proceso que consiste en reemplazar un hogar que no responde por otro hogar que no estaba incluido en la muestra y que es semejante al que no respondió en lo que se refiere a las características de interés.

77. Conviene utilizar cierta forma de compensación para la falta total de respuesta en las encuestas de hogares, bien ajustando las ponderaciones básicas de los hogares que responden o bien mediante la sustitución. La ventaja de la sustitución es que ayuda a mantener bajo control el número de hogares participantes. No obstante, la sustitución supone que el entrevistador no se ve tan fuertemente presionado a obtener datos de los hogares originales de la muestra. Además, la sustitución de los hogares que no han res-

pondido requiere tiempo y se presta a error. Por ejemplo, puede realizarse una sustitución utilizando un hogar más accesible en vez del hogar específicamente designado para servir como sustituto de otro que no respondió. El procedimiento de ajustar las ponderaciones de la muestra en casos de falta de respuesta se utiliza más frecuentemente en las grandes encuestas de todo el mundo. Fundamentalmente, el ajuste transfiere las ponderaciones básicas de todas las unidades de la muestra que no han respondido a las unidades que sí lo han hecho. En el capítulo VIII se presenta un examen más detallado de la falta de respuesta y de cobertura en las encuestas de hogares, y de las formas prácticas de compensarlas (véase también la bibliografía citada en ese capítulo). En el capítulo XI y en los estudios de casos de la segunda parte (capítulos XXII, XXIII y XXV) se facilitan también detalles sobre encuestas concretas.

78. Pueden introducirse otros ajustes en las ponderaciones, si corresponde. Por ejemplo, si se dispone de totales de control fiables pueden utilizarse ajustes posteriores a la estratificación para conseguir que las distribuciones del muestreo ponderado de determinadas variables se corresponda con las distribuciones conocidas de la población. Véanse en Lehtonen y Pahkinen (1995) algunos ejemplos prácticos de cómo analizar los datos de las encuestas con posestratificación.

2. Análisis de los datos de las encuestas de hogares

79. Para que los datos de las encuestas de hogares puedan analizarse adecuadamente deben cumplirse varias condiciones. En primer lugar, la base de datos asociada debe contener información que refleje el proceso de selección de la muestra. En particular, debería incluir etiquetas adecuadas para los estratos del diseño de la muestra, las unidades primarias de muestreo, las unidades secundarias de muestreo, etcétera. En segundo lugar, deben señalarse las ponderaciones de la muestra para cada unidad del archivo de datos en que se refleje la probabilidad de selección de cada unidad de muestreo y se compensen la falta de respuesta y otras deficiencias de la muestra. En tercer lugar debe haber documentación técnica suficiente del diseño de la muestra para la encuesta que generó los datos. En cuarto lugar, los archivos de datos deben generar el formato y estructura adecuados y la información necesaria sobre la vinculación entre las unidades de muestreo y las distintas fases de la selección de la muestra. Finalmente, debe contarse con programas informáticos adecuados y con personal especializado para usarlo debidamente.

80. Se necesita un programa informático especial para calcular las estimaciones de los errores estándar de las estimaciones de la encuesta que reflejen las complejidades del diseño de la muestra realmente utilizado. Entre esas complejidades cabe señalar la estratificación, los conglomerados y el muestreo de probabilidad desigual (ponderación). Los programas estándar de estadística generalmente no pueden utilizarse para la estimación del error estándar con diseños de muestras de gran complejidad, ya que casi siempre suponen que los datos se han adquirido mediante muestreo aleatorio simple. En general, el uso de programas estándar de estadística infravalora los verdaderos errores estándar de las estimaciones de la encuesta. Existen ya varios programas de computadora para analizar los datos de encuesta obtenidos a partir de diseños de muestras complejos. Algunos de estos programas se examinan y comparan detenidamente en el capítulo XXI.

G. Conclusiones

81. Concluimos insistiendo en algunas cuestiones temáticas asociadas con el diseño de las encuestas de hogares en los países en desarrollo y en transición, a saber:

a) **Usos múltiples de la mayor parte de las encuestas de hogares.** En los países en desarrollo y en transición hay cada vez mayor interés por el establecimiento de

programas integrados de encuestas con fines múltiples, sobre diversos temas y en varias rondas, a diferencia de las encuestas especiales y que se realizan una sola vez. Desde el comienzo, los responsables del diseño de las encuestas deben reconocer los múltiples objetivos y las expectativas contradictorias a que se verán sometidos los datos generados por las encuestas. Estas exigencias contradictorias suponen ciertos requisitos, que muchas veces son difíciles de conseguir. Por ello, los diseñadores de las encuestas deben mantener conversaciones detalladas con los donantes, los responsables de las políticas, los productores de datos de los institutos nacionales de estadística y los usuarios de los datos en los diversos ministerios sectoriales del país. El objetivo de estos contactos preliminares es tratar de armonizar y racionalizar las exigencias contradictorias a que se ve sometido el diseño de la muestra antes de que finalice.

b) **Determinación del tamaño apropiado de la muestra.** Una de las principales cuestiones que deben abordarse desde el comienzo es la determinación del tamaño adecuado de la muestra. Cada vez hay mayor demanda de estimaciones precisas sobre las características de interés no sólo de alcance nacional y regional sino también provincial e incluso de nivel más bajo. Ello lleva inevitablemente a proponer muestras de gran tamaño. La mayor fiabilidad de las estimaciones que se consigue reduciendo el error de muestreo con muestras de gran tamaño vale mucho más que la garantía de calidad de los datos conseguida reduciendo los errores no muestrales. Conviene que el encargado de diseñar la encuesta realice un análisis de los beneficios en función de los costos de las distintas opciones sobre el tamaño de la muestra y el plan de asignación. Parte de ese análisis debería consistir en un examen de los errores no muestrales en las encuestas y sus efectos en la calidad global de los datos obtenidos. La demanda de muestras de gran tamaño debería considerarse teniendo también en cuenta los costos y beneficios asociados. Como se afirma en la sección D, es importante recordar que al asignar la muestra se preste atención prioritaria a los campos de interés;

c) **Documentación del diseño y realización de la encuesta.** En muchas encuestas la documentación del proceso de diseño y realización de la encuesta es escasa o nula. Para que un conjunto de datos sea útil a los analistas y otros usuarios, es absolutamente imprescindible que se documenten todos los aspectos del proceso de diseño que generó los datos, incluida la selección de la muestra, la recopilación de datos, la preparación de archivos de datos, la construcción de ponderaciones de la muestra, incluidos los posibles ajustes para compensar las imperfecciones de ésta, y, si es posible, especificaciones para la estimación de los errores estándar. Sin esa documentación no es posible realizar ningún análisis adecuado de los datos. La documentación de la encuesta es también imprescindible para la vinculación con otras fuentes de datos y para distintos tipos de pruebas y análisis complementarios;

d) **Evaluación del diseño de la encuesta.** Un aspecto muy importante del proceso de diseño de la encuesta es analizar y evaluar la eficacia del diseño una vez finalizada la encuesta. Deben asignarse recursos para esta importante actividad en el proceso de formulación del presupuesto general, ya en la fase de planificación. La evaluación del diseño actual de una encuesta puede ayudar a mejorar el diseño de otras que puedan hacerse en el futuro. Dicha evaluación puede revelar informaciones útiles; por ejemplo, si se consiguieron o no mejoras con una asignación desproporcionada o cuál fue el nivel de discrepancia, si lo hubo, entre las medidas actuales del tamaño y las obtenidas en el momento de selección de la muestra. Dicha información puede utilizarse luego para formular diseños más eficientes en encuestas futuras.

Reconocimientos

El autor desea agradecer las provechosas observaciones de varios revisores y editores, y en particular al Dr. Graham Kalton, por sus numerosas sugerencias, que dieron lugar a considerables mejoras en los borradores iniciales de este capítulo. Las opiniones expresadas son las del autor y no reflejan necesariamente las de las Naciones Unidas.

REFERENCIAS BIBLIOGRÁFICAS

Cochran, W. G. (1977). *Sampling Techniques,* tercera edición. Nueva York: Wiley.

Glewwe, P. e I. Yansaneh (2000). The Development of Future Household Surveys in Viet Nam. Report of Mission to the General Statistics Office, Viet Nam.

Kalton, G. (1983). *Introduction to Survey Sampling*. Quantitive Applications in the Social Sciences Series, Sage University Paper, No. 35. Beverly Hills, California: Sage Publications.

Kish, L. (1965). *Survey Sampling.* Nueva York: Wiley.

_____ (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society,* Series A, vol. 139, págs. 80-95.

_____ (1988). Multi-purpose sample design. *Survey Methodology*, vol. 14, págs. 19-32.

_____ (1995). Methods for desing effects. *Journal of Official Statistics,* vol. 11, págs. 55-77.

Lehtonen, R. y E. J. Pahkinen (1995). *Practical Methods for Design and Analysis of Complex Surveys*. Nueva York: Wiley.

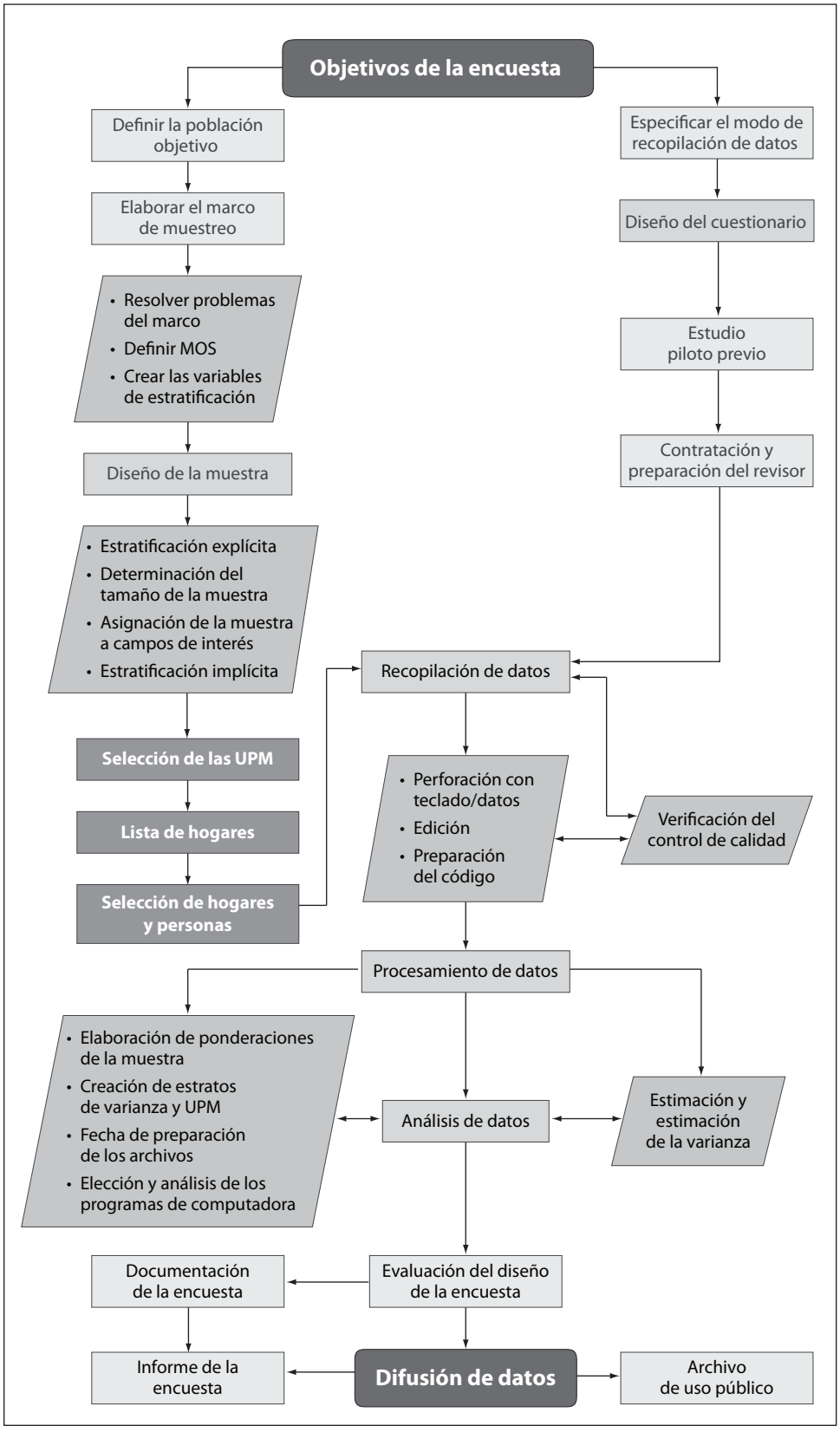
Lohr, Sharon (1999). *Sampling: Design and Analysis*. Pacific Grove, California: Duxbury Press.

Yansaneh, I. S. (2000). Sample Design for the 2000 Turkmenistan Mini-census Survey. Report of Mission to the National Institute for Statistics and Forecasting, Turkmenistán.

_____ (2003) Construction and use of sample weights (capítulo 7). En *Designing Samples for Household Surveys. Practical guidelines.* Nueva York: DESA/UNSD (de próxima aparición).

_____, L. Wallace y D. A. Marker (1998). Imputation methods for large complex datasets: an application to the NEHIS. En *Proceedings of the Survey Research Methods Section, American Statistical Association,* págs. 314-319. Alexandria, Virginia: American Statistical Association.

ANEXO Diagrama de flujo del proceso de una encuesta



Capítulo III
Panorama general del diseño de cuestionarios para las encuestas de hogares en los países en desarrollo

Paul Glewwe

Departmento de Economía Aplicada
Universidad de Minnesota
St. Paul, Minnesota, Estados Unidos de América

RESUMEN

Se examinan las cuestiones básicas referentes al diseño de cuestionarios para encuestas de hogares en los países en desarrollo. El capítulo se abre con el primer paso del diseño del cuestionario, que consiste en formular los objetivos de la encuesta y luego modificarlos teniendo en cuenta los obstáculos subyacentes. Después el autor advierte sobre muchos aspectos del diseño de los cuestionarios. Se ofrecen también recomendaciones acerca del ensayo sobre el terreno y sobre la conclusión del cuestionario.

Términos clave: Diseño del cuestionario, objetivos de la encuesta, obstáculos, prueba piloto.

A. Introducción

1. Las encuestas de hogares pueden ofrecer información abundante sobre muchos aspectos de la vida. No obstante, la utilización de sus datos depende mucho de la calidad de la encuesta, tanto por lo que se refiere al diseño del cuestionario como a la realización efectiva sobre el terreno. Si bien el diseño de los cuestionarios y la realización de las encuestas de hogares pueden parecer a primera vista tareas sencillas, en realidad requieren un arduo esfuerzo y mucho tiempo.

2. En el presente capítulo se presenta un panorama general básico sobre el proceso de diseño de un cuestionario de encuestas de hogares para su uso en los países en desarrollo. Esta presentación es simplemente introductoria, ya que el diseño de cuestionarios es un proceso muy complejo que no puede describirse con detalle en un capítulo de esta longitud. El objetivo es exponer las cuestiones más importantes y ofrecer algunas indicaciones útiles sobre cada una de ellas. Quienes deseen planificar una encuesta deberán consultar otros materiales para obtener asesoramiento más detallado. Un buen punto de partida es Grosh y Glewwe (2000), en donde se facilita información detallada sobre el diseño de las encuestas de hogares para los países en desarrollo. Aunque se escribió teniendo en cuenta un tipo específico de encuesta —Estudio de medición de los niveles de vida (EMNV), del Banco Mundial— gran parte de sus orientaciones son de interés para casi todos los tipos de encuestas de hogares. Otros análisis más generales, aunque menos recientes, del diseño de los cuestionarios pueden encontrarse en Casley y Lury (1987), Naciones Unidas (1985), Sudman y Bradburn (1982)

y Converse y Presser (1986). Puede verse un examen detallado del diseño de una encuesta sobre la mano de obra en Hussmanns, Merhan y Verma (1990).

3. A lo largo de todo este capítulo, se supone que el cuestionario de la encuesta será administrado por entrevistadores que visitarán a los encuestados en sus hogares, y que la unidad de muestreo es el hogar[1]. Como la mayor parte de las encuestas de hogares recopilan información sobre cada uno de los miembros integrantes del hogar, están basadas en muestras de individuos y en muestras de hogares.

[1] En algunas encuestas la unidad de muestreo es la vivienda, no el hogar, pero en tales casos algunos o todos los hogares de las viviendas incluidas en la encuesta se convierten en "unidades de información". Además, en algunas encuestas de hogares no pueden incluirse poblaciones específicas pertinentes, como los niños de la calle y los nómadas. Aun así, la mayor parte del material del presente capítulo se aplicará a las encuestas de esos tipos de población. Para más información sobre la manera de obtener muestras de dichas poblaciones, véase Naciones Unidas (1993).

4. El resto del capítulo está organizado de la forma siguiente: En la sección B se presentan los aspectos generales, es decir, los objetivos y obstáculos de la encuesta. En la sección C se ofrece asesoramiento para organizar la estructura del cuestionario de la encuesta, el formato y otros detalles del diseño del cuestionario. La Sección D contiene recomendaciones sobre el proceso general, desde la formación de un equipo hasta la comprobación sobre el terreno y la terminación del cuestionario. Una breve sección final permite ofrecer algunas observaciones a modo de conclusión.

B. Aspectos generales

5. Los cuestionarios de encuestas de hogares presentan enormes diferencias de contenido y extensión. La versión final de todo cuestionario es fruto de un proceso en el que se toman centenares, y hasta miles, de decisiones. Se necesita un marco general para garantizar un proceso ordenado y conseguir, en definitiva, que la encuesta alcance los objetivos establecidos. Para ello, los encargados del diseño de la encuesta deben ponerse de acuerdo en los objetivos de la misma y los obstáculos que encontrará. En esta sección se explica cómo establecer el marco general, comenzando con los principios básicos para ir pasando luego a orientaciones prácticas.

1. Objetivos de la encuesta

6. Los organismos gubernamentales y otras organizaciones realizan las encuestas de hogares con el fin de responder a sus preguntas acerca de la población[2]. Así pues, como los objetivos de la encuesta son obtener respuestas a dichas preguntas, el cuestionario debe contener los datos que puedan llevar a la respuestas. Dados los limitados recursos y la falta de tiempo de los encuestados, debería evitarse recopilar datos que no contribuyan a los objetivos de la encuesta. Por ello, el primer paso al diseñar una encuesta de hogares es ponerse de acuerdo sobre sus objetivos y formularlos por escrito.

[2] Estas preguntas generales, cuya respuesta desearía conocer la organización que realiza la encuesta, no son necesariamente las mismas que otras de carácter más específico del cuestionario de la encuesta que deben formularse a los miembros del hogar. En esta sección se considera el primer tipo de preguntas.

7. Para establecer los objetivos de la encuesta, los encargados de su diseño deben comenzar formulándose las preguntas cuyas respuestas desearía conocer la organización u organizaciones que patrocinan la encuesta. Pueden considerarse cuatro tipos de preguntas. Las más sencillas se refieren a las características fundamentales de la población en el momento actual. Podrían señalarse como ejemplo las siguientes preguntas:

- *¿Qué proporción de la población es pobre?*
- *¿Cuál es la tasa de desempleo?*
- *¿Cuál es la incidencia de la malnutrición entre los jóvenes?*
- *¿Qué cultivos son producidos por los hogares rurales en las diferentes regiones del país?*

8. Un segundo tipo de preguntas relaciona las características de los hogares con las políticas y programas gubernamentales, a fin de examinar la cobertura de los mismos. Un ejemplo de este tipo de preguntas sería el siguiente:

- *¿Qué proporción de los hogares participa en un programa determinado y cómo pueden compararse las características de estos hogares con las de los que no participan en el programa?*

9. Un tercer tipo de pregunta hace referencia a los *cambios* en las características de los hogares a lo largo del tiempo. Los organismos gubernamentales y las organizaciones muchas veces desean saber si las condiciones de vida de los hogares están mejorando o empeorando. Para responder a este tipo de pregunta se necesitan los datos de dos o más encuestas separados por un considerable período de tiempo, y los datos pertinentes deben recopilarse de la misma manera en ambas encuestas. Como se explica en Deaton y Grosh (2000), bastan pequeñas diferencias en la recopilación de los datos para que éstos no sean comparables y, por lo tanto puedan llevar a engaño.

10. El cuarto y último tipo de pregunta se refiere a los determinantes (causas) de las circunstancias y características de los hogares. Estas preguntas son difíciles de responder, pues lo que se desea saber no es sólo *qué* está ocurriendo, sino también *por qué* ocurre. No obstante, éstas son con frecuencia las preguntas más importantes, ya que tratan de comprender los efectos de las políticas o programas vigentes, y quizá de políticas o programas hipotéticos futuros, en las circunstancias y características de los hogares. Los economistas y sociólogos no siempre están de acuerdo en la forma de responder a estas preguntas, y algunas veces quizá ni siquiera se pongan de acuerdo en si es posible responder a una pregunta concreta. Si estos interrogantes revisten gran valor para los encargados de diseñar la encuesta, se necesita una planificación muy esmerada. No obstante, las cuestiones relacionadas con esta planificación van más allá del ámbito del presente capítulo (véanse los diferentes capítulos en Grosh y Glewwe (2000), donde se examina con detalle qué es lo que se requiere para responder a este tipo de pregunta).

11. Una vez que se ha llegado a un acuerdo sobre las preguntas que se deben responder, dichas preguntas pueden expresarse como objetivos de la encuesta. Por ejemplo, la presencia de una pregunta sobre la tasa actual de desempleo implica que uno de los objetivos de la encuesta es medir la incidencia del desempleo en la población económicamente activa. El siguiente paso es clasificar estos objetivos por orden de importancia. Si el número de objetivos es grande, es muy posible que la encuesta no pueda recopilar toda la información necesaria para conseguir todos ellos, por limitaciones presupuestarias, falta de capacidad y otros obstáculos. Cuando ello ocurre, deberían abandonarse los objetivos que tienen escasa prioridad (en relación con el esfuerzo necesario para recopilar la información requerida para poder alcanzarlos)[3]. En este proceso de decisión sobre los objetivos de la encuesta hay que comprobar si se pueden utilizar otros datos ya existentes para responder a la pregunta asociada con el objetivo. Todo objetivo que pueda alcanzarse utilizando datos disponibles de otra procedencia debería eliminarse de la lista de objetivos para la nueva encuesta. Este proceso de selección de un conjunto razonable de objetivos es más un arte que una ciencia, y los encargados del diseño de las encuestas deben tener también en cuenta factores como la experiencia anterior de recopilación de datos relacionados con el objetivo y la capacidad global del organismo que va a realizar la encuesta. No obstante, una vez que se han superado esas dificultades este planteamiento debería ayudar a los encargados de formulación de la encuesta a aprobar de común acuerdo una lista de objetivos que deberían conseguirse gracias a la encuesta.

[3] En vez de renunciar a un objetivo menos importante, existe la posibilidad de recopilar los datos necesarios para conseguirlo consultando sólo a una submuestra de hogares. Ello requerirá menos recursos, pero reducirá también la precisión de las estimaciones y podría complicar asimismo la realización de la encuesta sobre el terreno.

12. Un último punto que conviene señalar es que algunos responsables del diseño de encuestas prefieren expresar el conjunto de preguntas u objetivos por medio de cuadros que deben completarse utilizando los datos de la encuesta. Este planteamiento, denominado con frecuencia "plan de tabulación", funciona mejor con los tres primeros tipos de preguntas. En términos más generales, la forma en que se utilicen los datos recopilados en una encuesta de hogares para responder a las preguntas (alcanzar los objetivos) puede denominarse como

"plan de análisis de datos". Estos planes, que pueden ser muy detallados, deberían elaborarse una vez que se hayan aclarado los detalles de la encuesta de hogares (véase un examen más detallado en la sección C).

2. Obstáculos

13. El proceso de elección de los objetivos descritos más arriba tiene lugar en el contexto de una serie de problemas que limitan las opciones disponibles. Los encargados del diseño de las encuestas se enfrentan con tres grandes obstáculos. El primero y más evidente es la disponibilidad de recursos financieros para realizar la encuesta. Este obstáculo limita el número de hogares que se pueden incluir en la encuesta y el tiempo que los entrevistadores pueden dedicar a cada hogar (lo que, a su vez, limita el número de preguntas que se pueden formular en cada uno de ellos). En general, hay diferentes combinaciones de tamaño de la muestra (número de hogares encuestados) y volumen de información que se puede obtener de cada hogar, y cada presupuesto exige la búsqueda de una solución de compromiso en relación con estas dos características de la encuesta. En particular, con un determinado volumen de recursos financieros, la única forma de aumentar el tamaño de la muestra es disminuir el volumen de la información recopilado de cada hogar, y viceversa[4]. Indudablemente, ello tiene repercusiones en el número de objetivos de la encuesta y en la precisión de dichos objetivos (es decir, la precisión de las respuestas dadas a las preguntas subyacentes): una muestra pequeña puede permitir recopilar más datos por hogar, y de esa manera dar respuesta a más preguntas, pero la precisión de estas respuestas será menor debido al menor tamaño de la muestra. Una cuestión relacionada es que la calidad de los datos, en el sentido de exactitud de la información, se verá también afectada por los recursos disponibles. Por ejemplo, si se dispone de fondos para que cada entrevistador pueda dedicar más tiempo a completar un cuestionario de un determinado tamaño, el tiempo adicional podría utilizarse para volver al hogar y corregir los errores o contradicciones de los datos que se hayan observado después de terminada la entrevista.

[4] Cuando se cuenta con un presupuesto ajustado es difícil determinar la relación exacta entre información recopilada por hogar y número de hogares entrevistados. En particular, no es cierto que se pueda duplicar el tamaño de la muestra reduciendo a la mitad el cuestionario asignando un tiempo determinado a los entrevistadores, porque necesitan mucho tiempo para llegar a los hogares, presentarse e ir al siguiente hogar o zona de empadronamiento, y ese tiempo no se puede reducir aunque se abrevie el cuestionario.

14. El segundo obstáculo es la capacidad de la organización que realiza la encuesta. Las muestras de gran tamaño o los cuestionarios muy detallados pueden superar la capacidad de la organización cuando se quiere mantener un determinado nivel de calidad. Cuanto mayor sea la muestra, mayor será el número de personas entrevistadas y de personal que habrá que contratar y capacitar para la introducción de los datos (suponiendo que se pueda ampliar el tiempo necesario para completar la encuesta), lo que significa que la organización quizá tenga que reducir las calificaciones mínimas aceptables para los encargados de las encuestas y de la entrada de datos a fin de contratar los necesarios. De la misma manera, los cuestionarios más extensos requerirán mayor capacitación y personal más competente, siendo así que en los países en desarrollo muchas veces escasea el personal bien capacitado y con experiencia en realización de entrevistas y entrada de datos. Este obstáculo muchas veces no se reconoce debidamente, y en consecuencia muchas encuestas realizadas en los países en desarrollo han producido un gran volumen de datos pero de calidad dudosa y, por lo tanto, de utilidad incierta.

15. Un último obstáculo hace referencia a la voluntad y capacidad de los hogares entrevistados de facilitar la información deseada. En primer lugar, la disposición de los hogares a responder a las preguntas es limitada, por lo que la carga de la respuesta en los cuestionarios sumamente largos dará lugar probablemente a altas tasas de rechazo y/o a un gran número de datos incompletos o imprecisos. En segundo lugar, aun cuando los encuestados estén dispuestos a cooperar, quizá no sean capaces de responder a las preguntas complejas o que les obliguen a recordar acontecimientos que tuvieron lugar hace muchos meses o años. Ello repercute directamente en el diseño del cuestionario. Por ejemplo, quizá no sea posible obtener una estimación razonablemente precisa del ingreso de un hogar haciendo un número

de preguntas reducido, y quizá haya que presentar una larga serie de preguntas detalladas; así ocurre especialmente en los hogares agrícolas de las zonas rurales que cultivan numerosos productos, unos destinados al consumo directo y otros a la venta.

3. Algunos consejos prácticos

16. Los encargados del diseño de las encuestas deberán buscar distintas soluciones de compromiso entre los objetivos de la encuesta y los obstáculos encontrados hasta llegar a un conjunto de objetivos que sean viables, habida cuenta de esos obstáculos, y "óptimos", en el sentido de que constituyen los objetivos más importantes para la organización que realiza la encuesta. Una vez aclarado lo que es viable, quizá sea posible reducir los obstáculos obteniendo recursos financieros adicionales o impartiendo capacitación adicional a los futuros entrevistadores. La experiencia de otras encuestas recientemente terminadas en el mismo país debería servir de guía para determinar qué es viable y qué es poco realista. Como ya se ha mencionado, el logro del justo equilibrio es más un arte que una ciencia, y la experiencia local internacional constituye una buena guía para alcanzar ese equilibrio.

C. Los detalles

17. Una vez establecido el panorama general de los objetivos de la encuesta, los encargados del diseño deberán comenzar la labor detallada e inevitablemente tediosa de redactar el cuestionario, pregunta por pregunta. Una observación general que conviene hacer desde el principio es que se necesita un plan de análisis de datos. En él se explica con detalle qué datos se necesitan para alcanzar los objetivos (responder las preguntas) fijados para la encuesta. Los encargados del diseño deben consultar este plan constantemente cuando tratan de especificar los detalles del cuestionario. En algunos casos el plan de análisis de datos debe modificarse a medida que la labor detallada de diseño del cuestionario arroje nueva luz sobre la manera como se deben analizar los datos. Toda pregunta que no se utilice en el plan general de análisis de datos debe eliminarse del cuestionario.

18. Este capítulo es demasiado breve para describir con detalle cómo relacionar el diseño del cuestionario con los objetivos específicos y sus correspondientes planes de análisis de datos. En Grosh y Glewwe (2000) pueden verse varios capítulos sobre temas concretos en los que se ofrece asesoramiento mucho más detallado para los diferentes tipos de encuesta. En el resto de la presente sección se ofrece asesoramiento general pero muy útil sobre la manera de especificar los detalles del cuestionario de una encuesta de hogares.

1. El concepto de módulo

19. El cuestionario de una encuesta de hogares normalmente consta de varias partes, conocidas normalmente con el nombre de módulos. Un módulo consta de una o más páginas de preguntas en las que se recoge información sobre un tema concreto, como la vivienda, el empleo o la salud. Por ejemplo, la serie de encuestas demográficas y de salud que se examinan en el capítulo XXII tiene módulos sobre los métodos anticonceptivos, las preferencias en materia de fecundidad y la inmunización infantil. En términos más generales, en casi todos los cuestionarios de encuestas de hogares que contienen varias preguntas sobre un determinado tema, como la educación de cada uno de los miembros del hogar, conviene agrupar esas preguntas en una o varias páginas del cuestionario y considerar esa o esas páginas como el módulo correspondiente a ese tema; por ejemplo, las preguntas sobre educación representarían el "módulo de educación". De esta manera todo el cuestionario puede considerarse como un conjunto de módulos, quizá no menos de tres y no más de entre 15 y 20, dependiendo del número de temas incluidos en el cuestionario. Cada módulo contiene varias preguntas; algunas veces serán sólo 5 ó 6, pero en otras ocasiones puede haber

hasta 50 o incluso más de 100[5]. Los módulos muy grandes, como los que tienen más de 50 preguntas, deberían dividirse a su vez en submódulos relacionados con temas concretos. Por ejemplo, un gran módulo sobre el empleo podría dividirse en los siguientes submódulos: empleo primario, empleo secundario e historial de empleo. En cualquier caso, el número general de preguntas de un cuestionario debe reducirse al mínimo necesario para obtener la información deseada.

[5] Un módulo con más de cien preguntas puede dar lugar a que el tiempo de las entrevistas sea excesivo. Véase en la sección D un examen más detallado sobre la longitud del cuestionario.

20. El concepto de módulo tiene la ventaja de que permite dividir el diseño del cuestionario en dos pasos. El primero consiste en decidir qué módulos se necesitan, es decir, qué temas deberán incluirse en el cuestionario, y el orden que deberían seguir los módulos. El segundo consiste en elegir el diseño de cada módulo, pregunta por pregunta. Durante ambos pasos habrá que hacer referencia constante a los objetivos de la encuesta y al plan de análisis de datos.

21. La elección de los módulos y el grado de detalle de cada uno de ellos ofrecen un amplio abanico de posibilidades, según los objetivos de la encuesta y los obstáculos con que se enfrenta. No obstante, algunas orientaciones pueden ser válidas para la mayoría de las ocasiones. Por ejemplo, casi todas las encuestas de hogares recopilan información sobre el número de personas pertenecientes al hogar, junto con alguna información muy básica sobre ellas, como su edad, sexo y relación con el jefe del hogar. Estas preguntas pueden concentrarse en un módulo de una sola página, la "lista de hogares". Este módulo debería ser uno de los primeros —en la mayor parte de los casos el primero— del cuestionario. Muchos cuestionarios de encuestas de hogares formulan luego a los individuos que integran los hogares preguntas sobre temas como la educación, el empleo, la salud y la migración. Cada uno de estos temas, sobre los que se formulan cinco o más preguntas, debería colocarse en un módulo especial sobre dicho tema. Si sólo se formulan entre una y tres preguntas, quizá sea más conveniente incluirlas en la lista de hogares, o quizá en otro módulo dedicado a solicitar información a los individuos integrantes del hogar.

22. Casi todos los módulos de una encuesta de hogares pueden dividirse en dos tipos principales: los que formulan preguntas a los miembros individuales, como se ha mencionado antes, y los que hacen preguntas generales sobre el hogar. En cuanto al primer tipo, debe tenerse en cuenta que las preguntas que se formulan a los individuos componentes de los hogares no tienen que ser necesariamente las mismas para cada miembro; muchas encuestas de hogares tienen preguntas que se aplican únicamente a algunos miembros del hogar, como los niños de menos de cinco años de edad o las mujeres en edad de procrear. Como ejemplos del segundo tipo cabe señalar las preguntas acerca de las características de la vivienda del hogar y sobre los gastos del conjunto del hogar en alimentos y otros productos. Naturalmente, la extensión de cualquiera de estos módulos y el tipo de preguntas contenidas en ellos dependerá de los objetivos de la encuesta.

23. Finalmente, cabe formular algunas observaciones generales sobre el orden de los módulos en la encuesta de hogares. En primer lugar, ese orden debería corresponderse con el de realización de las entrevistas, de manera que el entrevistador pueda completar el cuestionario comenzando con la primera página y siguiendo, página por página, hasta el final del cuestionario. Quizá se necesiten excepciones en algunos casos, pero en general lo "natural" es que los módulos se ordenen de esa manera.

24. En segundo lugar, los primeros módulos del cuestionario deberían consistir en preguntas relativamente fáciles de contestar y relacionadas con temas que no sean delicados. La sugerencia de utilizar la lista de hogares como primer módulo coincide con esta recomendación, ya que la información básica sobre los miembros de los hogares normalmente no constituye un tema delicado. Si se comienza la entrevista de esa manera, los miembros del hogar podrán sentirse cómodos y el entrevistador podrá llegar a entablar una buena relación con ellos. Ello significa que la mayor parte de los módulos con información delicada

deberían presentarse al final del cuestionario. De esa manera el entrevistador tendrá todo el tiempo posible para ganarse la confianza de los miembros del hogar, lo que incrementará la probabilidad de que respondan a esas preguntas de forma exhaustiva y sincera. Además, si las preguntas delicadas hacen que los miembros del hogar den por terminada la entrevista, al menos se habrá obtenido ya el resto de la información.

25. Un tercer principio es el de agrupar los módulos que probablemente serán contestados por el mismo miembro del hogar. Por ejemplo, las preguntas sobre los gastos en alimentación y de otro tipo deberían ir juntas, ya que probablemente es una misma persona la que mejor pueda contestar a ese tipo de preguntas. De esa manera podrá responder a todos los módulos y terminar su participación, dejando que otros miembros del hogar respondan a los módulos restantes. La observación general es que conviene utilizar el tiempo de los miembros del hogar de manera eficiente, ya que no sólo lo agradecerán sino que cooperarán con mayor interés. Además, ello probablemente permitirá ahorrar tiempo al entrevistador, ya que tendrá que citar una sola vez a cada uno de los encuestados para que realicen su contribución a la entrevista.

2. Formato y coherencia

26. Una vez seleccionados los módulos y determinado su orden, hay que realizar la minuciosa y tediosa tarea de elegir las preguntas concretas y de redactarlas, palabra por palabra. Al llevar a cabo esta labor conviene comenzar examinando las anteriores encuestas de hogares sobre el mismo tema realizadas en el país, o quizá en un país vecino. En general, aunque las preguntas más indicadas y la mejor formulación dependerán del carácter y los objetivos de la nueva encuesta, pueden darse algunos consejos generales válidos para casi todas las encuestas de hogares.

27. La primera recomendación es que en casi todos los casos las preguntas deben aparecer escritas en el cuestionario de manera que el entrevistador pueda leerlas directamente durante la entrevista; de ese modo se garantiza que se formularán las mismas preguntas a todos los hogares. La otra posibilidad es preparar un formulario con un mínimo de texto, dejando que el entrevistador formule las preguntas utilizando sus propias palabras. Esta opción es menos aconsejable, ya que se presta a numerosos errores. Por ejemplo, supongamos que un módulo sobre el empleo tiene una "pregunta" que dice solamente "ocupación principal". El sentido no es claro. ¿Se refiere a la ocupación el día o la semana de la entrevista o durante los 12 últimos meses? Para las personas con dos ocupaciones, ¿la principal ocupación es la que le produce más ingreso o la que le lleva más horas del día? Esta confusión se puede evitar si la pregunta se redacta con todo detalle, como en el siguiente ejemplo: "Durante los últimos siete días, ¿qué tipo de trabajo realizó? Si ha realizado más de un tipo de trabajo, indique por favor a cuál dedicó más horas durante los últimos siete días". En la figura III.1 puede verse un ejemplo de página de cuestionario que recopila información sobre la vivienda (obsérvese que todas las preguntas están formuladas como frases completas). La ventaja de escribir el texto completo de las preguntas quedó demostrada en un estudio experimental de Scott y otros (1988): las preguntas que no se habían escrito con detalle produjeron entre 7 y 20 veces más errores que las que se formularon palabra por palabra.

28. La segunda recomendación está estrechamente relacionada con la primera: el cuestionario debería incluir definiciones precisas de todos los conceptos clave utilizados en él, sobre todo para que el entrevistador pueda consultar la definición cuando se presenten casos especiales. Asimismo, el cuestionario debería contener algunas instrucciones para el entrevistador; como ejemplo pueden verse las que acompañan a la pregunta 10 de la figura III.1. Deberían presentarse también instrucciones y explicaciones más detalladas en un manual para entrevistadores. El tema de los manuales se examina en el capítulo IV.

Figura III.1
Ilustración de un formato de cuestionario

1. ¿Es esta vivienda propiedad de un miembro de su hogar?

Sí 1
No ..... 2 (➧12)

2. ¿Cómo obtuvo su hogar esta vivienda?

Privatizada 1
Comprada a un particular 2
Nueva construcción 3
Cooperativa 4
Intercambiada ..... 5 (➧7)
Heredada ..... 6 (➧7)
Otro medio ..... 7 (➧7)

3. ¿Cuánto pagó por la vivienda?

4. ¿Realiza pagos periódicos por su vivienda?

Sí 1
No ..... 2 (➧7)

5. ¿Cuánto es el total de cada uno de esos pagos?

Cantidad (moneda)
Unidad de tiempo

6. ¿En qué año prevé que realizará su último pago?

Año

7. ¿Tiene usted un título legal o documento que acredite que es el propietario de la tierra?

Sí 1
No 2

8. ¿Tiene usted un título legal de la vivienda o un documento que acredite que es el propietario?

Sí 1
No 2

9. ¿Qué tipo de título tiene?

Título legal pleno, registrado 1
Título legal pleno, no registrado 2
Recibo de compra 3
Otro 4

10. ¿Qué persona tiene el título o documento de propiedad de esta vivienda?

Tome de la lista el código de identificación de esa persona

Primer código
Segundo código

11. ¿Podría vender usted la vivienda si quisiera?

Sí 1
No ..... 2 (➧14, página siguiente)

12. Si usted vendiera esta vivienda hoy ¿cuánto recibiría por ella?

Cantidad (moneda)

13. Estime, por favor, el dinero que podría recibir en concepto de alquiler si arrendara esta vivienda a otra persona

Cantidad (moneda)
Unidad de tiempo

➧ Pregunta 28, página siguiente

Unidades de tiempo: Día ..... 3 Mes ..... 6 Año ... 9
Semana ..... 4 Trimestre ... 7
Quincena ... 5 Semestre 8

29. Una tercera recomendación es preparar preguntas cortas y sencillas, en la medida de lo posible, utilizando términos habituales de la conversación diaria. Además, todas las preguntas deberían comprobarse atentamente para garantizar que no predisponen al encuestado ni le inducen a dar respuestas sesgadas. Si la pregunta es complicada, sería mejor dividirla en dos o tres. Por ejemplo, supongamos que se desea saber si una persona era empleada o autónoma (o ambas cosas) durante los siete últimos días. Para tratar de obtener toda esta información con una sola pregunta utilizando una jerga más bien técnica se podría llegar a una pregunta como la siguiente:

- *Durante los siete últimos días, ¿tenía usted trabajo asalariado o remunerado de otra manera, o trabajaba por su cuenta en una empresa familiar, realizaba ambos tipos de actividad simultáneamente o no tenía ninguno de esos dos tipos de trabajo?*

Esta pregunta podría dividirse en dos y formularse en términos menos técnicos:

1. *Durante los siete últimos días, ¿realizó usted algún trabajo remunerado para alguien que no es miembro de este hogar?*
2. *Durante los últimos siete días, ¿trabajó usted por su propia cuenta, por ejemplo, como agricultor o vendedor de bienes o servicios?*

Las preguntas 8, 9 y 10 de la figura III.1 ilustran también esta idea. Los encargados del diseño de las encuestas pueden tener la tentación de "abreviar" el cuestionario, combinando estas preguntas en una más larga:

- *¿Qué tipo de título legal o documento, en su caso, tiene usted que acredite la propiedad de esta vivienda, y quién es el miembro del hogar que posee realmente el título?*

No obstante, esta pregunta más larga podría confundir a muchos encuestados y, si ello ocurre, haría falta más tiempo para explicar la pregunta que para formular las tres por separado.

30. En cuarto lugar, el cuestionario debe concebirse de tal manera que las respuestas a casi todas las preguntas estén precodificadas. Estas preguntas reciben muchas veces el nombre de "preguntas cerradas". Por ejemplo, las respuestas a las preguntas cuya respuesta es *sí* o *no* pueden registrarse en el cuestionario como "1" (= Sí) o "2" (= No), respectivamente. De esa manera se facilitan las cosas al entrevistador, que sólo tiene que escribir una cifra, en vez de una palabra entera o una frase[6]. Además, y sobre todo, se elude el paso de la "codificación", que acontece en los casos de los cuestionarios con las respuestas de una o más palabras escritas manualmente por los entrevistadores (ilegibles en muchos casos) y que se entregan a un "codificador", que escribe luego los códigos numéricos correspondientes a esas respuestas. Este paso adicional puede producir nuevos errores, y en casi todos los casos se puede evitar. (En cualquier caso, la codificación de clasificaciones más complejas, como la ocupación y la rama de actividad, requiere conocimientos técnicos y tiempo, que el personal de campo normalmente no tiene, y se recomienda que se realice por personal especializado de las oficinas correspondientes tomando como base las descripciones escritas de los entrevistadores). En la figura III.1 todas las posibles respuestas están precodificadas y los códigos aparecen en la misma página que la pregunta (en general, inmediatamente después de ésta).

[6] Otra opción es que el entrevistador ponga una X o una marca de comprobación en el recuadro adjunto a la respuesta precodificada.

31. La quinta recomendación está relacionada con la tercera. El plan de codificación de las respuestas debería ser coherente para todas las preguntas. Por ejemplo, en casi todas las encuestas de hogares hay muchas preguntas cuya respuesta es *sí* o *no.* Los códigos numéricos de dichas preguntas del cuestionario deberían ser siempre los mismos, por ejemplo, "1"= *Sí* o "2"= *No.* Una vez que se establece esta norma (u otra parecida), debería utilizarse siempre que las respuestas a las preguntas del cuestionario sean *sí* o *no.* De esa manera el entrevistador comprenderá que debe utilizar el 1 o el 2 en todas las preguntas del cuestionario cuya respuesta pueda ser *sí* o *no.* Esto puede aplicarse también a otros tipos de respuestas. Muchos cuestionarios contienen preguntas cuyas respuestas se formulan en unidades de tiempo o de

distancia, como "¿Cuándo visitó usted por última vez a un médico?" o "¿Qué distancia hay entre su casa y la carretera más próxima?". Estas unidades podrían codificarse en la forma siguiente: 1 indicaría los minutos; 2, las horas; 3, los días; 4, las semanas, y así sucesivamente. En tal caso, la respuesta "10 días" se registraría con dos números: "10", que indica el total de días, y "3", que es el código de la unidad de tiempo, los días. De la misma manera, en lo que respecta a la distancia, el código 1 indicaría los metros y 2 los kilómetros. El esquema de codificación concreto puede diferir en las distintas encuestas; lo importante es que, en la medida de lo posible, todas las preguntas que requieren un código de este tipo utilicen el mismo esquema de codificación[7]. En la figura III.1 puede verse un ejemplo de puesta en práctica de esta recomendación. Efectivamente, los códigos relativos a las unidades de tiempo que aparecen al pie de la página se señalan una sola vez y se utilizan en dos preguntas de esa página: la 5 y la 13.

[7] Si bien importa poco que los códigos numéricos de conceptos sencillos, como las unidades de tiempo y de distancia, difieran de unas encuestas a otras en el mismo país, hay razones de peso para utilizar el mismo esquema de codificación en los conceptos más complejos, como los tipos de ocupación o de enfermedad, para garantizar la comparabilidad a lo largo del tiempo en las diferentes encuestas.

32. Este análisis de los esquemas de codificación plantea el interrogante de si el entrevistador debe decir a los encuestados las posibles respuestas a las preguntas o debe limitarse a leer la pregunta y no los códigos de respuesta. En general, es mejor este último método, porque los encuestados pueden decidirse por una respuesta sencillamente porque la oyeron antes, aun cuando otra de las respuestas que aparecen más tarde sea más precisa. Por otro lado, si son muchas las respuestas que se deben leer, los encuestados pueden cometer errores al elegir entre las muchas y diferentes respuestas posibles.

33. La sexta recomendación es incluir en el cuestionario "códigos de salto", que indican las preguntas que deben omitirse porque las respuestas a preguntas anteriores las hacen innecesarias. Por ejemplo, una encuesta puede incluir la pregunta "¿Ha buscado trabajo en los siete últimos días?". Si la respuesta es *sí,* el cuestionario puede preguntar cuáles son los métodos utilizados; si la respuesta es *no,* esa pregunta sería irrelevante. A la derecha de la primera pregunta deberían incluirse, instrucciones muy breves, como "SI LA RESPUESTA ES *NO*, PASAR A LA PREGUNTA *6*", para que el entrevistador no haga preguntas innecesarias, o adoptarse algunos códigos que reflejen esas instrucciones de forma más sucinta. Por ejemplo, la instrucción anterior podría convertirse en "EN CASO NEGATIVO, → P.6". En la figura III.1, este tipo de instrucciones son muy breves: se indican mediante números entre paréntesis después de los códigos de respuesta pertinentes. Por ejemplo, la indicación "(◗12)" después del código *NO* en la pregunta 1 indica que si la respuesta a esa pregunta es *No*, el entrevistador debe pasar a la pregunta 12.

34. Una última observación acerca del formato es que las preguntas deberían formularse de tal manera que permitan al encuestado responder con sus propias palabras. Por ejemplo, en una encuesta sobre la vivienda puede haber una pregunta sobre el alquiler pagado por ese concepto. Según sea el contrato de arrendamiento, algunos de los encuestados pagan una determinada cantidad cada semana, mientras que otros lo hacen una vez al mes o incluso una vez al año. Lo importante es dejar que el informante elija la unidad, por lo que la pregunta debería ser "¿Cuánto paga usted por el alquiler de su vivienda?" en vez de "¿Cuánto paga al mes por el alquiler de la vivienda?". El problema de esta última pregunta es que obliga al informante a calcular el alquiler que paga al mes. Es muy posible que sepa que tiene pagar, por ejemplo, 50 dólares por semana, pero puede cometer un error al multiplicar 50 dólares por 4,3 y, por lo tanto, dar una respuesta que no sea la correcta (217 dólares al mes). Lo mejor es elaborar el cuestionario de tal manera que el entrevistador pueda anotar los códigos numéricos de las diferentes unidades de tiempo, como se ilustra en la pregunta 5 de la figura III.1, de manera que, por ejemplo, 50 dólares por semana puedan registrarse como 50 en una casilla más 4 (código numérico de la semana) en un espacio adyacente. Cuando se analicen los datos, el investigador, que tiene muchas menos probabilidades de equivocarse que el informante, puede convertir fácilmente las cifras en unidades comunes, por ejemplo el alquiler pagado anualmente.

3. Otras orientaciones sobre algunos detalles del diseño del cuestionario

35. Finalmente, pueden ofrecerse algunas orientaciones más generales sobre el diseño del cuestionario. En primer lugar, cuando se trata de preguntas muy importantes, como el número de personas del hogar o las diferentes fuentes de ingreso, quizá convenga hacer una pregunta "de sondeo" que ayude al informante a recordar algo que podría haber olvidado. Por ejemplo, después de obtener una lista de todos los miembros del hogar, el entrevistador podría formular la siguiente pregunta:

- *Según la información que me ha dado, hay seis personas en este hogar. ¿Es eso correcto o hay alguien más que pertenezca a este hogar, por ejemplo, alguien que esté temporalmente fuera durante unos días o varias semanas?*

36. En segundo lugar, el cuestionario debería formularse de tal manera que cada hogar y cada persona del mismo tenga un número de código único que identifique a esa persona en todas las partes del cuestionario. Ello ayudará a los analistas de datos a encajar la información correspondiente a los mismos hogares y a las mismas personas. En casi todos los casos debe haber un solo cuestionario por hogar; en los casos excepcionales en que se utilicen dos o más cuestionarios, debe ponerse especial esmero en comprobar que se escriba el mismo código de hogar en todos los cuestionarios correspondientes a ese hogar.

D. El proceso

37. Hasta ahora se ha explicado cómo se diseñan los cuestionarios de encuestas de hogares, pero no se ha dicho prácticamente nada sobre quiénes deberán intervenir en ello y cómo pueden comprobar el cuestionario durante su preparación. En esta sección se dan algunas recomendaciones acerca del proceso utilizado para redactar, probar y ultimar el cuestionario.

1. Constitución de un equipo

38. Las encuestas de hogares implican casi siempre un gran número de decisiones y acciones, que normalmente resultan más complicadas de lo previsto inicialmente. Ello supone que una sola persona o incluso un grupo reducido no pueden tener el tiempo ni la especialización necesarios para diseñar con éxito un cuestionario. Por ello es preciso formar un equipo de expertos al comienzo mismo del proceso para no dejar de lado ninguno de los aspectos de la encuesta. Este equipo debería contar con representantes de todos los grupos más importantes.

39. Quizá lo más importante sea contar en el equipo con uno o varios miembros del grupo de encargados de la formulación de políticas, es decir, una o más personas que representen los intereses del grupo o grupos que tienen previsto utilizar la información recopilada en la encuesta para adoptar las políticas pertinentes. Aunque estas personas no sean técnicos en la materia, su presencia es necesaria para que puedan señalar (y recordar) a otros miembros del equipo los objetivos últimos de la encuesta. De esa manera se reforzará enormemente la comunicación entre quienes usan los datos y quienes los producen.

40. Un segundo grupo importante, integrado por investigadores y analistas de datos, utilizará la información para responder a las preguntas de interés que puedan formular las autoridades. Su función es elaborar el plan de análisis de datos con el fin de garantizar que éstos sean suficientes para poder dar respuesta a esas preguntas. En algunos casos la respuesta a las preguntas de las autoridades es una tarea sencilla, pero en otros puede resultar más bien complicada.

41. Finalmente, aunque no en orden de importancia, está el grupo de los recopiladores de datos, en que se incluyen los entrevistadores, supervisores y encargados de la entrada de datos (incluidos los especialistas en informática). Normalmente, estas personas son funcionarios de la organización que tiene la responsabilidad formal de recopilar los datos. Es imprescindible que los integrantes de este grupo cuenten con experiencia previa en recopilación de datos de encuestas de hogares. Son los que mejor saben qué tipos de preguntas pueden responder los hogares y cuáles no. Dentro de este grupo debería haber alguien con experiencia en la fase de entrada de datos. Algunas indicaciones sencillas de esos expertos pueden aumentar significativamente la precisión de los datos recogidos y reducir el tiempo necesario para preparar los datos para su análisis.

2. Elaboración del primer borrador de cuestionario

42. El primer borrador de casi todos los cuestionarios de encuestas de hogares se elabora en una serie de reuniones de los miembros del equipo. Como ocurre con todos los primeros borradores, el resultado tendrá inevitablemente numerosos errores. El planteamiento modular propuesto en este capítulo implica que el primer borrador estará integrado por una serie de módulos diferentes. Al integrarlos en el primer borrador es preciso comprobar varias cosas.

43. En primer lugar, el equipo de la encuesta debe verificar si el conjunto de los módulos recoge toda la información deseada. Puede ocurrir que se dé por supuesto que una pregunta fundamental para un módulo se ha incluido ya en otro, cuando en realidad no sea así. Es necesario organizar una reunión conjunta de todos los participantes en todos los módulos para comprobar que no se ha omitido ninguna información importante. Algo semejante podría decirse sobre las superposiciones. Cuando se integran todos los módulos es posible que algunas preguntas aparezcan por duplicado en diferentes módulos. Esta repetición debería eliminarse normalmente con el fin de ahorrar tiempo a los informantes y a los entrevistadores. El único caso en que no deben eliminarse las preguntas repetidas es cuando éstas permiten confirmar una información muy importante, por ejemplo, si una persona es realmente miembro del hogar. La edad de los miembros del hogar puede comprobarse incluyendo preguntas sobre la edad actual y la fecha de nacimiento, y el hecho de que alguien es realmente miembro del hogar puede verificarse preguntando si ha vivido en otros lugares durante los 12 últimos meses y, en caso afirmativo, cuántos meses ha vivido fuera (después de preguntar inicialmente cuántos meses lleva viviendo en el hogar donde se está realizando la entrevista).

44. En segundo lugar, es preciso comprobar el tamaño del cuestionario. En todos los países hay un límite a las preguntas que los informantes pueden estar dispuestos a responder en una encuesta de este tipo. Al mismo tiempo, hay cierta tendencia a formular un gran número de preguntas, con lo que el producto final es mucho más extenso de lo previsto originalmente. Para aclarar cuánto tiempo se necesita para entrevistar a un hogar típico (y cuánto tiempo podrían dedicar a la entrevista los informantes) puede realizarse una prueba sobre el terreno (véase más adelante), pero los entrevistadores y supervisores con experiencia pueden dar al equipo una idea aproximada sólo examinando el cuestionario. La eliminación de preguntas que recogerían información menos prioritaria es una parte dolorosa pero necesaria de la elaboración del primer borrador de cualquier cuestionario para una encuesta de hogares.

45. Finalmente, debería comprobarse el primer borrador del cuestionario para determinar la coherencia en los períodos de recuerdo. Por ejemplo, un objetivo de la encuesta puede ser recoger información sobre los ingresos de toda procedencia percibidos por los hogares en el último mes o año. Es preciso examinar el cuestionario para comprobar que todas las secciones que recogen datos sobre los ingresos tengan el mismo período

de recuerdo[8]. La principal excepción a esta norma es la de algunos casos particulares en que, como se ha explicado antes, debe darse a los informantes cierta flexibilidad para elegir el período de recuerdo que les resulte más fácil.

[8] Algunas encuestas incluyen puntos cronológicos de referencia; por ejemplo, cuando se pregunta acerca de las circunstancias existentes 5 ó 10 años antes. Estos puntos de referencia, que algunas veces incluyen un día, un mes o un año específico, deberían comprobarse también para ver si son coherentes en todo el cuestionario.

3. Comprobación sobre el terreno y terminación del cuestionario

46. Ningún cuestionario, por pequeño o sencillo que sea, debe terminarse sin una comprobación anterior en un pequeño número de hogares para descubrir los posibles problemas de diseño. En casi todos los casos un nuevo cuestionario tiene numerosos errores e inconvenientes que no se ven hasta que no se hace una prueba con algunos hogares representativos de la población objetivo. A continuación se ofrecen algunas normas generales. Puede verse un análisis más detallado en Grosh y Glewwe (2000) y en Converse y Presser (1986).

47. La comprobación del borrador del cuestionario sobre el terreno puede dividirse en dos etapas. La primera, conocida con frecuencia como prueba preliminar, consiste en probar algunas secciones (módulos) del cuestionario en un pequeño número de hogares (por ejemplo, de 10 a 15), para hacerse una idea aproximada de si las páginas del cuestionario funcionan adecuadamente. Esta operación puede realizarse más de una vez, comenzando con las primeras fases del proceso de diseño. La segunda fase es una prueba más completa sobre el terreno, denominada a veces prueba piloto. Se trata de una operación de más envergadura, con intervención de entre 100 y 200 hogares. Éstos deben pertenecer no a una pequeña área, sino a varias zonas que representen a la población pertinente. En las encuestas destinadas a zonas urbanas o rurales, la prueba piloto debe realizarse también en zonas urbanas y rurales. Debería llevarse a cabo también en diferentes partes del país o de la región donde se va a utilizar el cuestionario definitivo. Finalmente, la elección de los hogares debería hacerse de tal manera que se puedan comprobar todos los módulos al menos en 50 hogares, y en más si es posible. Ello significa, por ejemplo, que si el cuestionario tiene un módulo que recopila datos sobre pequeñas empresas familiares, al menos 50 de los hogares entrevistados para esta prueba piloto deberían tener una de esas empresas.

48. La mayoría de las pruebas piloto requieren un período de una a dos semanas para la realización de las entrevistas en los 100-200 hogares. Todos los miembros del equipo de la encuesta deberían participar en la prueba piloto y observar el mayor número de entrevistas posible. De hecho, las pruebas piloto ofrecen una excelente oportunidad de capacitación para quienes tienen poca experiencia en el diseño de cuestionarios de encuestas de hogares. Una información importante facilitada por la prueba piloto es la estimación del tiempo necesario para realizar un cuestionario[9]. No obstante, hay que tener en cuenta que la cifra obtenida sobreestimará (multiplicándolo al menos por dos) el tiempo necesario para entrevistar un hogar en la encuesta real, tanto porque los entrevistadores de la encuesta piloto han tenido poca experiencia con el proyecto de cuestionario como porque sufrirán demoras debido a los problemas del cuestionario que se corregirán en su versión final.

[9] Al realizar las pruebas preliminares y las pruebas piloto, el borrador de cuestionario debería incluir un espacio para anotar el tiempo de comienzo y de terminación de cada módulo, que se registrará para cada uno de los hogares entrevistados. De esa manera se sabrá cuánto tiempo de entrevistas se necesita para cada módulo.

49. Otro punto importante es que en los países donde se habla más de un idioma el cuestionario debe traducirse a todos los idiomas importantes, y las traducciones se incluirán en la prueba. Se trata de un requisito de importancia máxima. En particular, hay que evitar, en la medida de lo posible, que los entrevistadores tengan que improvisar la traducción durante la entrevista. Algunos estudios (por ejemplo, Scott y otros, 1988) han demostrado que en esos casos el número de errores es entre dos y cuatro veces mayor que cuando existe ya un cuestionario traducido al idioma del informante. Para comprobar la exactitud de la traducción, una persona o un grupo distinto del que realizó la traducción original debería realizar una "traducción inversa" del cuestionario a la lengua original. El resultado se comparará con el contenido del cuestionario inicial para determinar si la traducción reflejó claramente el contenido original; toda diferencia indicaría que algo "se perdió en la traducción". Una referencia útil para la traducción del cuestionario es Harkness, Van de Vijver y Mohler (2003).

50. Un aspecto final importante de la prueba piloto es que debe comprobar no sólo el borrador de cuestionario sino también todo el plan de trabajo sobre el terreno, incluidos los métodos de supervisión, la entrada de datos y los materiales escritos, como los manuales para entrevistadores (temas que se examinan con mayor detalle en el capitulo IV). Únicamente comprobando todo el proceso puede el equipo tener la garantía de que la encuesta está lista para su realización. Un último paso útil es realizar un análisis rápido de los datos recopilados en la prueba piloto para sacar a la luz los problemas que, de lo contrario, podrían pasarse por alto.

51. Inmediatamente después de la prueba piloto, el equipo de la encuesta debería celebrar varios días de reuniones para examinar los resultados y modificar el cuestionario de acuerdo con las enseñanzas aprendidas. Antes, debería realizar un análisis rápido de los datos de la prueba piloto que se mencionan en el párrafo anterior, que normalmente se presentará en forma de cuadros sencillos. En algunos casos puede haber tantos problemas que debería programarse una segunda prueba piloto, quizá no tan amplia como la primera, para verificar si los grandes cambios introducidos en el cuestionario funcionarán de hecho sobre el terreno. Todos los miembros del equipo deben estar presentes en esas reuniones de análisis, en las que deberían participar también todas o la mayoría de las personas que realizaron de hecho las entrevistas durante la prueba piloto.

52. En los últimos años se han realizado numerosas investigaciones sobre el diseño de los cuestionarios y se han elaborado nuevos y valiosos métodos para realizarlos de manera eficaz. Aunque todavía no se utilizan ampliamente en los países en desarrollo y en transición, es probable que su uso aumente notablemente en el futuro. No tenemos aquí espacio para describir estos métodos, pero alentamos al lector a que consulte las publicaciones pertinentes. Entre esos métodos figuran los grupos de discusión, las entrevistas cognitivas y la codificación del comportamiento. Esposito y Rothgeb (1997) y Biemer y Lyberg (2003) presentan una valiosa exposición general de esos métodos. Véase también Krueger y Casey (2000) sobre los grupos de discusión, Forsyth y Lessler (1991) sobre las entrevistas cognitivas y Fowler y Cannell (1996) sobre la codificación del comportamiento. En el capítulo IX de esta publicación se ofrece también información detallada sobre los grupos de discusión y la codificación del comportamiento en los apartados C.2 y C.6, respectivamente.

E. Observaciones finales

53. En el presente capítulo se han presentado algunas recomendaciones generales para el diseño de los cuestionarios de hogares en los países en desarrollo. Se ha insistido sobre todo en los cuestionarios aplicados a los hogares. Algunas encuestas de hogares recopilan también datos sobre la comunidad local en un "cuestionario sobre la comunidad" independiente. Estos cuestionarios no se tratan en este capítulo, por falta de espacio. Véanse en Frankenberg (2000) recomendaciones más detalladas sobre su diseño.

54. En el capítulo se abordan también otros temas, pero siempre en forma sucinta. Quienes estén interesados en planificar una encuesta deben consultar otros materiales a fin de obtener asesoramiento mucho más detallado. Las referencias bibliográficas que siguen son un buen punto de partida.

REFERENCIAS BIBLIOGRÁFICAS

Biemer, Paul P. y Lars E. Lyberg (2003). *Introduction to Survey Quality*. Nueva York: Wiley.

Casley, Dennis y Denis Lury (1987). *Data Collection in Developing Countries*. Oxford, Reino Unido: Clarendon Press.

Converse, Jean M. y Stanley Presser (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills, California: Sage Publications.

Deaton, Angus y Margaret Grosh (2000). Consumption. En *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study,* Margaret Grosh y Paul Glewwe, comps. Nueva York: Oxford University Press (para el Banco Mundial).

Esposito, James L. y Jennifer M. Rothgeb (1997). Evaluating survey data: making the transition from pretesting to quality assessment. En *Survey Measurement and Process Quality,* Lars E. Lyberg y otros, comps. Nueva York: Wiley.

Forsyth, Barbara H. y Judith T. Lessler (1991). Cognitive laboratory methods: a taxonomy. En *Measurement Errors in Surveys,* Paul P. Biemer y otros, comps. Nueva York: Wiley.

Fowler, F. J. y C. F. Cannell (1996). Using behavior coding to identify cognitive problems with survey questions. En *Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco, California: Jossey-Bass.

Frankenberg, Elizabeth (2000). Community and price data. En *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study,* Margaret Grosh y Paul Glewwe, comps. Nueva York: Oxford University Press (para el Banco Mundial).

Grosh, Margaret y Paul Glewwe, comps. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study*. Nueva York: Oxford University Press (para el Banco Mundial).

Harkness, Janet A., Fons J. R. van de Vijver y Peter Mohler (2003). *Cross-Cultural Survey Methods*. Nueva York: Wiley.

Hussmanns, R., F. Merhan y V. Verma (1990). *Surveys of Economically Active Population, Employment, Unemployment, and Underemployment*. An ILO Manual on Concepts and Methods. Ginebra: Organización Internacional del Trabajo.

Krueger, Richard A. y Mary Anne Casey (2000). *Focus Groups: A Practical Guide for Applied Research*. Thousand Oaks, California.: Sage Publications.

Naciones Unidas (1985). *Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares: Desarrollo y diseño de los cuestionarios de encuestas (*INT-84-014). Nueva York.

Naciones Unidas (1993). *National Household Survey Capability Programme*: *Sampling Rare and Elusive Populations* (INT-92-P80-16E). Nueva York.

Scott, Christopher y otros (1988). Verbatim questionnaires *versus* field translations or schedules: an experimental study. *International Statistical Review,* vol. 56, No. 3, págs. 259-278.

Sudman, Seymour y Norman M. Bradburn (1982). *Asking Questions. A Practical Guide to Questionnaire Design.* San Francisco, California: Jossey-Bass.

Capítulo IV

Consideraciones generales sobre la realización de encuestas de hogares en los países en desarrollo

Paul Glewwe
Departamento de Economía Aplicada
Universidad de Minnesota
St. Paul, Minnesota, Estados Unidos de América

RESUMEN

En el presente capítulo se examinan algunas cuestiones básicas relativas a la realización de encuestas de hogares en países en desarrollo, comenzando con las actividades que deben llevarse a cabo antes de poner en marcha la encuesta: elaboración de un presupuesto y plan de trabajo, preparación de la muestra, capacitación del personal y redacción de los manuales de capacitación, y preparación del plan de trabajo sobre el terreno. Se consideran también las actividades que tienen lugar mientras se lleva a cabo la encuesta: establecimiento y mantenimiento de sistemas de comunicación y transporte, elaboración de protocolos de supervisión y otras actividades que mejoren la calidad de los datos, y creación de un sistema de gestión de datos. El capítulo concluye con una breve sección sobre las actividades realizadas después del trabajo sobre el terreno, seguida de una breve conclusión.

Términos clave: Realización de la encuesta, presupuesto, plan de trabajo, muestra, capacitación, plan de trabajo sobre el terreno, comunicaciones, transporte, supervisión, gestión de datos.

A. Introducción

1. El valor de la información que aportan las encuestas de hogares depende mucho de la utilidad y la exactitud de los datos recopilados, que a su vez dependen de la forma en que se realiza la encuesta sobre el terreno. En este capítulo se presentan recomendaciones generales sobre la realización de las encuestas, que incluyen casi todos los aspectos del proceso de realización de una encuesta de hogares, con independencia del diseño del cuestionario.

2. Puede decirse que el cuestionario de una encuesta de hogares bien concebido (y los correspondientes planes de análisis de datos) representa el punto medio en el camino hacia una encuesta productiva. El punto final se alcanza mediante la realización eficaz de la encuesta. Ésta comienza no cuando los entrevistadores empiezan a entrevistar a los hogares asignados a ellos, sino varios meses —y quizá uno o dos años— antes. En la sección B de este capítulo se examinan las actividades que deben realizarse antes de poder comenzar con las entrevistas a los hogares. En la sección C se describen las actividades que tienen lugar mientras se lleva a cabo la encuesta; la sección D contiene un breve examen de las tareas que hay que llevar a cabo una vez finalizada la actividad sobre el terreno; en la sección final se ofrecen

algunas observaciones a modo de conclusión. Este capítulo puede ser una introducción útil sobre el tema, pero es demasiado breve para presentar con detalle todas las orientaciones pertinentes. Para que la encuesta logre sus objetivos, las personas responsables deben consultar materiales más detallados. Un buen punto de comienzo es Grosh y Muñoz (1996). Aunque se centra en el Estudio de medición de los niveles de vida (EMNV) del Banco Mundial, gran parte de sus indicaciones son también válidas para casi todos los tipos de encuestas de hogares. Otras dos referencias útiles son Casley y Lury (1987) y Naciones Unidas (1984).

3. A lo largo de todo este capítulo se supone que la encuesta es planificada y realizada por un equipo "básico" bien organizado, nombrado con ese fin. Se supone también que el cuestionario de la encuesta será aplicado por entrevistadores que visitarán a los encuestados en sus hogares y que la unidad de muestreo es el hogar[1]. Finalmente, el lector no debe olvidar que en este capítulo se tienen en cuenta los países en desarrollo, incluidas las economías en transición con ingresos bajos, como China y Viet Nam. De todas formas, la mayor parte de las recomendaciones se aplican también a las economías en transición más desarrolladas de Europa oriental y de la antigua Unión Soviética.

[1] En algunas encuestas la unidad de muestreo es la vivienda, no el hogar; en tales casos, algunos o todos los hogares de las viviendas incluidas en la muestra se convierten en "unidades de información" de la encuesta.

B. Actividades antes de que la encuesta se realice sobre el terreno

4. En toda encuesta de hogares, la primera tarea es crear un equipo básico que se encargue de gestionar todos los aspectos de la encuesta. En el capítulo III se explica con detalle a quiénes debe incluirse en el equipo. Una vez formado el equipo básico, antes de entrevistar a ningún hogar deben realizarse las ocho tareas siguientes:

a) Elaborar un presupuesto provisional y conseguir financiamiento;
b) Establecer un plan de trabajo para todas las actividades restantes;
c) Preparar una muestra de los hogares que se deben entrevistar;
d) Redactar manuales de capacitación;
e) Capacitación del personal sobre el terreno y de entrada de datos;
f) Preparación de un plan de actividades sobre el terreno y de entrada de datos;
g) Realización de una prueba piloto;
h) Lanzamiento de una campaña publicitaria.

Esta lista de tareas corresponde, aproximadamente, a su orden cronológico. A continuación se describe cada una de ellas.

a) Financiamiento del presupuesto

5. Los recursos financieros condicionan en muchos casos qué es lo que se puede hacer en una encuesta de hogares. Las limitaciones no siempre resultan evidentes. La primera tarea en casi todas las encuestas es establecer un proyecto de presupuesto basado en hipótesis acerca del número de hogares que deberán incluirse en la muestra y del tiempo de personal necesario para entrevistar a un hogar típico. Este presupuesto será aproximado, ya que algunos de los detalles del costo no pueden conocerse hasta tanto no se sepan los detalles del cuestionario, pero en la mayor parte de los casos el proyecto de presupuesto se parecerá bastante al resultado final (a no ser que se modifiquen significativamente los objetivos de la encuesta).

6. Una vez que se ha preparado el proyecto de presupuesto deben encontrarse los fondos necesarios. Si el financiamiento es incierto, probablemente habrá que aplazar la planificación detallada de la encuesta hasta que se consigan los fondos. De esa manera se evitarán pérdidas de tiempo de personal en el caso de que no pueda encontrarse financiamiento.

7. No es mucho lo que se puede decir de la preparación de un presupuesto cuando no se tiene más información sobre el carácter y tipo de la encuesta, pero cabe hacer algunas recomendaciones generales. En primer lugar, debería evaluarse la capacidad de la organización que efectuará la encuesta, y si carece de algunos de los conocimientos técnicos

necesarios, o si no cuenta con suficiente personal especializado en la preparación de muestras o en la utilización de las nuevas tecnologías de la información, habría que contratar consultores externos. Ello incrementará notablemente el costo de la encuesta, pero valdrá la pena. En segundo lugar, un buen modo de comenzar es considerar los presupuestos de encuestas semejantes previas en el país o en países semejantes. En tercer lugar, para evitar la carga de los costos imprevistos, debería añadirse expresamente un margen de protección de un 10%, como partida adicional del presupuesto. Esta partida recibe muchas veces el nombre de *costos imprevistos.* En los casos de gran incertidumbre acerca de los costos, quizá convenga establecer un fondo del 15%, o incluso del 20%, para esos costos.

8. Para situar el examen en un plano más concreto, en el cuadro IV.1 [versión modificada del cuadro 8.2 de Grosh y Muñoz (1996)] se presenta un proyecto de presupuesto para una encuesta hipotética. En este ejemplo se supone que en la encuesta se entrevistarán 3.000 hogares, y que la recopilación de datos durará un año. Además de un

Cuadro IV.1
Proyecto de presupuesto para una encuesta hipotética de 3.000 hogares *(dólares de los EE. UU.)*

Partida	Número	Tiempo	Costo por unidad	Costo total
Salarios básicos				
Jefe de proyecto	1	30 meses	800/mes	24 000
Encargado de datos	1	30 meses	600/mes	18 000
Director del trabajo sobre el terreno	1	30 meses	600/mes	18 000
Personal auxiliar/contable	3	24 meses	450/mes	32 400
Supervisores	4	14 meses	400/mes	22 400
Entrevistadores	12	13 meses	350/mes	54 600
Encargados de la entrada de datos	4	13 meses	300/mes	15 600
Conductores	2	13 meses	300/mes	7 800
Subtotal				**192 800**
Viáticos				
Jefe de proyecto	1	90 días	30/día	2 700
Encargado de datos	1	60 días	30/día	1 800
Director del trabajo sobre el terreno	1	90 días	30/día	2 700
Personal auxiliar	2	60 días	30/día	3 600
Personal encargado de las listas	10	60 días	15/día	9 000
Supervisores	4	290 días	15/día	17 400
Entrevistadores	12	270 días	15/día	48 600
Conductores	2	270 días	15/día	8 100
Subtotal				**93 900**
Material				
Adquisición de vehículos	2	–	20 000	40 000
Combustible y mantenimiento	2	13 meses	300/mes	7 800
Computadoras para la entrada de datos	4	–	1 000	4 000
Impresoras, estabilizadores, etcétera	5	–	1 000	5 000
Computadoras para el análisis de datos	3	–	1 500	4 500
Material de oficina/computadora	–	30 meses	350/mes	10 500
Fotocopiadora/fax	1 de cada	–	2 500	2 500
Subtotal				**74 300**
Costos de impresión				
Cuestionarios	3 500	–	2	7 000
Manuales de capacitación	40	–	5	200
Informes	500	–	5	2 500
Subtotal				**9 700**
Costos de los consultores				
Consultores extranjeros	5	Meses-persona	10 000/mes	50 000
Viático internacional	150	días	150/día	22 500
Viajes internacionales	8	viajes	2 000/viaje	16 000
Consultores locales	5	Meses-persona	3 000/mes	15 000
Subtotal				**103 500**
Imprevistos (10%)				47 400
TOTAL				**521 600**

Nota: El guión (–) indica que esa partida no se aplica.

equipo central (véase el capítulo III), hay cuatro equipos sobre el terreno, cada uno de ellos integrado por tres entrevistadores, un supervisor y un encargado de la entrada de datos. Dos conductores, con vehículos dedicados al proyecto, transportarán a los equipos a sus lugares de trabajo. Se supone que cada entrevistador trabajará 250 días a lo largo del año, entrevistando (por término medio) un hogar por día. En el cuadro se presentan los salarios hipotéticos de todo el personal, así como los viáticos ofrecidos a los miembros del equipo por cada día de trabajo sobre el terreno. Cada equipo sobre el terreno contará con una computadora para la entrada de datos; el equipo central tendrá tres computadoras para el análisis de datos. Se presentan también los costos hipotéticos de los consultores, tanto internacionales como locales. Naturalmente, este cuadro se presenta sólo a modo de ejemplo: el costo de una encuesta concreta dependerá del tamaño de la muestra, el número de personas contratadas, sus salarios y otras formas de remuneración, la proporción entre supervisores y entrevistadores, el número de hogares que puede abarcar un entrevistador en un solo día, la entrada de datos sobre el terreno o en un lugar centralizado, y muchos otros factores. Aquí se presenta para que pueda servir como lista de comprobación a fin de garantizar que en el proyecto de presupuesto se incluyan todos los costos básicos.

b) Plan de trabajo

9. Una vez conseguido el financiamiento necesario, la siguiente tarea es elaborar un plan de trabajo realista, que es fundamentalmente un calendario de actividades desde las primeras fases de la planificación de la encuesta hasta que finalice el trabajo sobre el terreno[2]. En el plan de trabajo se incluyen todas las actividades siguientes: gestión general (incluida la adquisición de equipo); elaboración del cuestionario; preparación de la muestra; asignación, contratación y capacitación del personal; entrada y gestión de datos; actividades sobre el terreno; y análisis de datos, procesamiento, documentación y preparación de informes. En cada una de estas áreas específicas debería prepararse una lista de las tareas que se deben realizar y las fechas de conclusión (en otras palabras, sus plazos). Debería hacerse hincapié en los momentos más importantes, como la prueba piloto y el primer día del trabajo sobre el terreno. Esta lista, que muchas veces puede presentarse en forma de gráfico, es el plan de trabajo de la encuesta.

[2] Se trata de un plan de trabajo general en el que se incluyen muchas tareas que deben realizarse antes de que comience el trabajo sobre el terreno (antes de que se entreviste a ningún hogar). Como se examina más adelante, se necesita también un "plan de actividades sobre el terreno y entrada de datos" más específico.

10. Es obvio que muchas de estas actividades están mutuamente relacionadas y, por lo tanto, deben estar coordinadas. Por ejemplo, muchas labores de gestión y análisis de datos no pueden comenzar mientras no se haya adquirido el equipo necesario y se haya asignado (o contratado) y capacitado el personal que la realizará. Hay que tener en cuenta que incluso los planes mejor concebidos deben modificarse cuando ocurren acontecimientos imprevistos. Considerados retrospectivamente, la mayor parte de los planes resultan demasiado optimistas, lo que significa que los retrasos son frecuentes. En la medida de lo posible, el calendario de las distintas actividades debe ser realista e incluir tiempos muertos que permitan a los participantes recuperarse cuando se produzcan los inevitables retrasos.

11. En la figura IV.1 [adaptación de la figura 8.1 de Grosh y Muñoz (1996)] se presenta un ejemplo de plan de trabajo. Comprende 30 meses. Los asteriscos (*) indican cuándo tienen lugar las diferentes actividades. En el diagrama se observa que los preparativos deben comenzar aproximadamente un año antes de que se proceda a realizar la encuesta sobre el terreno. El hecho de que la prueba piloto tenga lugar en el octavo mes significa que el proyecto de cuestionario, el personal capacitado y un esbozo del programa de entrada de datos deben estar ya listos para ese mes. El comienzo del trabajo sobre el terreno está previsto en el mes 12, y se supone que durará un año. El plan de trabajo supone también que se preparará un borrador de informe cuando se haya recopilado la mitad de los datos. Naturalmente, los planes de trabajo de una encuesta concreta serán diferentes de los que aquí se presentan. Esta versión sirve únicamente como lista de comprobación y trata de poner de manifiesto cómo debe coordinarse el calendario de las diferentes tareas.

Figura IV.1
Plan de trabajo para la preparación y realización de una encuesta de hogares

	Mes de la encuesta																													
Tarea	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
Gestión y logística																														
Nombrar el equipo central de la encuesta	*																													
Adquirir las computadoras		*				*	*	*	*																					
Adquirir los materiales para la encuesta						*	*	*																						
Publicidad							*	*	*	*	*																			
Adquirir/alquilar los vehículos					*	*	*	*	*																					
Preparación del cuestionario																														
Fijar los objetivos de la encuesta		*	*																											
Preparar el borrador de cuestionario			*	*	*																									
Reuniones sobre el borrador de cuestionario						*	*																							
Concluir la prueba piloto del borrador de cuestionario							*																							
Prueba piloto								*																						
Reuniones después de la prueba piloto									*																					
Imprimir la versión definitiva del cuestionario										*																				
Muestreo																														
Establecer el diseño y marco de la muestra		*	*																											
Elaborar una muestra (UPM)				*																										
Establecer el plan del trabajo sobre el terreno				*																										
Listas/mapas de las UPM					*	*	*	*																						
Personal y capacitación																														
Seleccionar y capacitar al personal de la prueba piloto						*	*																							
Preparar manuales de capacitación									*	*																				
Capacitación de los entrevistadores											*																			
Gestión de datos																														
Diseño del primer programa de entrada de datos						*	*	*																						
Versión definitiva del programa de entrada de datos									*	*																				
Redacción del manual de entrada de datos									*	*																				
Capacitación del personal de entrada de datos											*																			
Trabajo sobre el terreno												*	*	*	*	*	*	*	*	*	*	*	*							
Análisis y documentación																														
Borrador del plan de análisis																		*	*											
Analizar la primera mitad de los datos																				*	*									
Redactar el informe preliminar																						*								
Crear el primer conjunto completo de datos																								*	*					
Analizar los datos iniciales																										*	*			
Informe final y documentación																												*	*	*

c) Establecimiento de una muestra de hogares

12. En casi todas las encuestas de hogares hay una población que es objeto de especial interés; por ejemplo, la población de todo el país representada por los hogares de la encuesta. El proceso de selección de un conjunto de hogares que representen a la población en general se conoce con el nombre de muestreo, y el procedimiento para realizar el muestreo se denomina diseño de la muestra. Hay tantas cuestiones que deben considerarse al establecer una muestra, que no es ni siquiera posible enumerarlas todas en un panorama general tan breve como éste. En los capítulos II, V y VI de este volumen pueden verse recomendaciones detalladas al respecto. En Kalton (1983) se ofrece una presentación general del muestreo, mientras que en Kish (1965), Cochran (1977) y Lohr (1999) pueden encontrarse explicaciones mucho más exhaustivas.

13. Al referirnos al muestreo en este capítulo nos limitaremos a formular dos observaciones que deberá tener en cuenta el equipo encargado de la encuesta. En primer lugar, algunas veces es útil diseñar la muestra de manera que los hogares puedan entrevistarse a lo largo de un período de 12 meses. Ello permite obtener un promedio de la variación estacional de los fenómenos estudiados y utilizar los datos para el estudio de las pautas estacionales. En segundo lugar, y sobre todo, los encargados de planificar la encuesta deben evitar la tentación de incluir un número muy grande de hogares. Es lógico que deseen aumentar el tamaño de la muestra, sobre todo en los grupos de especial interés, ya que de esa manera se reduce el error de muestreo. No obstante, en muchos casos el aumento del tamaño va acompañado de un mayor número de errores "no muestrales" debido a la utilización de personal menos calificado y a una menor proporción entre supervisores y entrevistadores. Es muy posible, e incluso probable, que la reducción de los errores de muestreo debidos al mayor tamaño de la muestra se contrarreste por el mayor número de errores no muestrales.

d) Redacción de los manuales de capacitación

14. Quizá el componente más importante de la capacitación sea la preparación de manuales para todos los destinatarios: entrevistadores, supervisores y personal encargado de la entrada de datos. Se necesitan manuales distintos para cada uno de ellos, es decir, debe haber un manual para los entrevistadores, otro para los supervisores y un tercero para la entrada de datos. Los manuales son un elemento trascendental de la capacitación, y deben completarse antes de que ésta comience. Los manuales sirven sobre todo como material de referencia cuando la encuesta ha comenzado ya y tienen que contener toda la información necesaria para los diferentes tipos de personal sobre el terreno y encargado de la entrada de datos[3]. De hecho, los analistas de datos muchas veces utilizan manuales de capacitación para comprender mejor los datos que están analizando; ello supone que habrá que publicar ejemplares adicionales de todos los manuales, para uso de esos analistas. Por norma general, siempre que haya duda sobre si habría que incluir un material en el manual o excluirlo, sería mejor incluirlo.

15. Los manuales de capacitación deberían explicar el objetivo de la encuesta y las tareas básicas que va a realizar el personal al que se destina el manual. Deberían preverse procedimientos para los casos excepcionales, y en particular principios generales que habría que aplicar cuando surjan imprevistos. Los manuales deberían explicar también cómo rellenar los formularios que se consideran parte del trabajo (ello reviste especial importancia en el manual de los supervisores). Ya que incluso los manuales mejor preparados pueden tener errores u omisiones, deberían prepararse uno o varios conjuntos de instrucciones adicionales como complemento de los manuales que ya se hayan entregado al personal sobre el terreno y de entrada de datos.

e) Capacitación del personal sobre el terreno y del encargado de la entrada de datos

16. En algunos casos, la organización que realiza la encuesta tiene un gran número de entrevistadores, supervisores y personal de entrada de datos con abundante experiencia. Cuando la nueva encuesta es muy semejante a las realizadas anteriormente por la organización, se necesita dedicar poco tiempo a nuevas actividades de capacitación, quizá una o dos semanas para explicar los detalles del nuevo cuestionario y algunos cambios introducidos en los procedimientos que pueden acompañar a la nueva encuesta. Hay también casos en que la nueva encuesta es muy diferente de todo cuanto la organización ha realizado en el pasado. En la mayoría de esas situaciones las organizaciones deberán contratar al menos algunos nuevos trabajadores sobre el terreno y encargados de la entrada de datos, y los entrevistadores y supervisores contratados deben recibir capacitación general antes de darles a conocer los

[3] En el término "personal sobre el terreno" se incluyen los entrevistadores, supervisores, y otras personas que para realizar su labor viajan a las comunidades donde se entrevista a los hogares. Como se examina más adelante, es muy frecuente situar el personal de entrada de datos lo más próximo posible a esas comunidades. En las encuestas en las que el personal de entrada de datos viaja con el personal sobre el terreno podría incluirse también dentro de la denominación de personal sobre el terreno, pero no siempre ocurre así. El término "personal sobre el terreno y encargado de la entrada de datos" se utiliza en este capítulo para englobar ambas posibilidades.

aspectos específicos de la nueva encuesta. En general, estas situaciones requieren más de dos semanas de capacitación. Normalmente se necesitan de tres a cuatro semanas para garantizar que los entrevistadores y supervisores puedan realizar su labor de manera eficaz.

17. Si bien el carácter de la capacitación depende de la naturaleza de la encuesta, cabe hacer algunas observaciones generales. En primer lugar, en la capacitación debe incluirse un tiempo considerable a practicar entrevistas en hogares reales y utilizando el cuestionario. En segundo lugar, conviene insistir en la necesidad de comprender los objetivos de la encuesta y la forma en que los datos recopilados contribuirán a esos objetivos. La insistencia en este tipo de conocimientos, a diferencia de la capacitación del personal sobre el terreno y del encargado de la entrada de datos para que se limiten a seguir estrictamente las reglas, ayudará a los entrevistadores y supervisores a hacer frente a las cuestiones y problemas inesperados. En tercer lugar, es mejor capacitar a más individuos de los necesarios y aplicar algún tipo de prueba (con componentes escritos y de entrevistas prácticas). Los resultados de la prueba pueden utilizarse para seleccionar a los entrevistadores y supervisores que consigan un nivel más alto de resultados en la prueba. En cuarto lugar, la capacitación debería llevarse a cabo en una oficina central para que todo el personal sobre el terreno reciba la misma formación y para que la capacitación sea de la máxima calidad. Finalmente, es importante comprender que la calidad de la capacitación puede repercutir de manera decisiva en la calidad de la encuesta y, en definitiva, de los datos recogidos. Todo el equipo de la encuesta debe dar gran importancia a la capacitación y no limitarse a delegar esta tarea a uno o dos miembros.

f) Plan de actividades sobre el terreno y de entrada de datos

18. El trabajo efectivo de visitas a las zonas incluidas en la muestra y de entrevistas a los hogares pertinentes suele conocerse como trabajo sobre el terreno. Como éste debe estar estrechamente coordinado con la entrada de datos, en este capítulo se examinan ambos conjuntamente. El trabajo sobre el terreno debe comenzar lo antes posible después de la capacitación (menos de una semana, si las circunstancias lo permiten), con el fin de evitar que se olvide lo aprendido. Antes de comenzar es preciso trazar un plan muy detallado que permita establecer una correspondencia entre los hogares seleccionados (del plan de muestreo) y los entrevistadores, supervisores y personal de entrada de datos que van a realizar la labor. El personal de la encuesta normalmente se organiza por equipos dirigidos por un supervisor. Cada equipo se responsabiliza de una parte del total de la muestra y se encarga de garantizar que se entreviste a los hogares de la parte que se le ha asignado.

19. Al elaborar el plan del trabajo sobre el terreno conviene tener presentes varios principios. En primer lugar, hay que tener medios de transporte adecuados para el personal y para los suministros. La experiencia con encuestas de hogares en muchos países ha demostrado que los problemas logísticos más comunes son la falta de combustible y mantenimiento adecuado para los vehículos utilizados sobre el terreno. En segundo lugar, el plan debe ser realista, lo que quiere decir que debe estar basado en la experiencia anterior de otras encuestas de hogares en el país. Si se está experimentando un nuevo tipo de planteamiento, éste debe comprobarse durante la prueba piloto (véase el capítulo III). En tercer lugar, este plan debe ir acompañado de un proyecto de entrada de datos en el que se explique el proceso mediante el cual se incorporarán a las computadoras las informaciones procedentes de los cuestionarios terminados, para más adelante transformarlas en archivos maestros en la oficina central. En cuarto lugar, en las encuestas cuyas actividades sobre el terreno pueden durar varios meses debería tomarse un descanso[4] después de las primeras semanas para evaluar cómo está procediendo el trabajo sobre el terreno y la entrada de datos. Es muy probable que la experiencia obtenida en las primeras semanas permita formular sugerencias para modificar varios de los procedimientos de trabajo sobre el terreno y de entrada de datos; estos cambios podrían formularse por escrito y presentarse al personal sobre el terreno como adiciones a sus manuales, tal como se ha explicado antes. En quinto lugar, antes de concluir el plan de

[4] Esta pausa debería tener lugar durante un período de tiempo ordinario, para que la recopilación de datos no se interrumpa por un acontecimiento importante conocido con antelación.

trabajo sobre el terreno debería presentarse a supervisores y entrevistadores experimentados para que puedan hacer observaciones y sugerencias. Finalmente, los entrevistadores deben contar con tiempo suficiente en cada unidad primaria de muestreo (UPM) para realizar varias visitas a los hogares incluidos en la muestra y recopilar datos de las personas encuestadas mejor informadas; la alternativa de obtener respuestas "proporcionadas por terceros", es decir, otros miembros del hogar menos informados, reducirá probablemente la exactitud de los datos recogidos.

g) Realización de una prueba piloto

20. Todas las encuestas de hogares deben someter a prueba el diseño del cuestionario, los planes de trabajo sobre el terreno y de entrada de datos y todos los demás aspectos de la encuesta. Es lo que se conoce con el nombre de prueba piloto. Supone la realización de entrevistas en 100-200 hogares de todas las zonas del país que se incluirán en la encuesta. Como uno de los principales objetivos de la prueba piloto es evaluar el diseño del cuestionario, este tema se examina con detalle en el capítulo III. Después de terminada la prueba piloto se convoca una reunión de varios días en la que el equipo central y los participantes de la prueba piloto examinan los problemas identificados durante la misma. Los participantes en la reunión deben luego aprobar de común acuerdo un borrador final del cuestionario, los planes definitivos de trabajo y de entrada de datos y los demás aspectos de la encuesta.

h) Lanzamiento de una campaña publicitaria

21. Los responsables de las encuestas de hogares deberían dar a conocer en los medios de comunicación de masas el comienzo de una nueva encuesta, con el fin de sensibilizar a la opinión pública y de alentar a los hogares elegidos a que cooperen en las entrevistas. Otro beneficio de las campañas de publicidad es que elevan la moral del personal de las encuestas. En general no conviene gastar grandes sumas en publicidad, ya que la mayoría de los hogares que vean la información no serán entrevistados. No obstante, en algunos casos esta publicidad puede hacerse prácticamente sin costo, poniéndose en contacto con las emisoras de radio y televisión, los periódicos y otros medios de comunicación. Son especialmente útiles los artículos periodísticos, ya que los entrevistadores y supervisores pueden mantener ejemplares de los mismos para enseñar a los hogares que pongan en duda lo que los entrevistadores les dicen acerca de la encuesta.

22. Son también útiles las campañas de publicidad orientadas a objetivos más específicos, por ejemplo, octavillas distribuidas en las comunidades seleccionadas como UPM, o cartas dirigidas a los hogares que van a ser entrevistados. Las octavillas deben ser llamativas y de gran colorido, y tanto en ellas como en las cartas se debe insistir en la utilidad de los datos para mejorar las políticas gubernamentales. En las cartas debería hacerse también hincapié en que los datos son estrictamente confidenciales; en muchos países hay leyes que pueden citarse como garantía de la confidencialidad de la información. Finalmente, deben establecerse contactos con los dirigentes de las comunidades locales para que expliquen la importancia y los beneficios de la encuesta. Una vez convencidos de estos beneficios, las autoridades locales pueden convencer a los hogares reacios para que participen en la encuesta.

C. Actividades durante la realización de la encuesta sobre el terreno

23. Una vez terminados todos los preparativos, comienzan las entrevistas de los hogares. Cada país tiene una forma de realizar las encuestas. No obstante, pueden darse algunos consejos generales válidos para todos los países. Aquí suponemos que el trabajo sobre el terreno lo realizan equipos móviles.

1. Comunicaciones y transporte

24. Cada equipo sobre el terreno necesita acceso a una línea fiable de comunicación con la administración de la encuesta a fin de comunicar los progresos y los problemas y facilitar los datos a la oficina central lo antes posible. En los países en desarrollo muchas veces las posibilidades de comunicación son pobres, sobre todo en las zonas rurales. No obstante, en la mayor parte de los países el servicio telefónico ha mejorado hasta el punto de que cada equipo sobre el terreno puede tener acceso a un teléfono en el plazo de horas, o como máximo en uno o dos días. De hecho, los teléfonos celulares se están extendiendo en muchos países en desarrollo, aunque no siempre en las zonas rurales. Una opción sencilla es ofrecer teléfonos celulares a los equipos que trabajen en áreas adonde haya llegado esta tecnología. En cuanto a los equipos de las zonas remotas, los teléfonos por satélite pueden representar una inversión que merezca la pena.

25. El transporte fiable es también fundamental para el trabajo de los equipos sobre el terreno. El método utilizado variará de un país a otro, pero como mínimo cada equipo debería tener un medio de transporte seguro para poder desplazarse de una zona de trabajo a otra. Debe también planificarse el transporte en situaciones de emergencia, para aquellos casos en que un miembro del equipo se encuentre gravemente enfermo y necesite atención médica inmediata. Para el transporte tanto habitual como de emergencia debe preverse algún sistema de respaldo que pueda utilizarse si llegara a fallar el principal. Un medio de transporte fiable puede servir como medio alternativo de comunicación cuando falla todo lo demás.

2. Supervisión y garantía de calidad

26. La calidad del trabajo realizado por los entrevistadores es de importancia crucial en toda encuesta de hogares. Conseguir una encuesta con garantía de calidad no es una tarea fácil. Algunos entrevistadores quizá no puedan encargarse de este trabajo, y otros quizá no hagan todo lo posible si los incentivos son pocos o nulos. La clave para mantener la calidad es un sistema eficaz de supervisión del trabajo sobre el terreno.

27. Las siguientes recomendaciones ayudarán a los supervisores a observar y mantener con eficacia la calidad del trabajo realizado por los entrevistadores. En primer lugar, cada supervisor debe encargarse de un pequeño número de entrevistadores: no más de cinco, y si es posible sólo dos o tres. En segundo lugar, al menos la mitad del tiempo de cada supervisor debe dedicarse a comprobar la calidad del trabajo de los entrevistadores. En tercer lugar, debe prepararse una lista de comprobación relativamente corta para que los supervisores controlen los cuestionarios completados que presentan los entrevistadores; de esa manera se garantizará el seguimiento de algunas normas básicas para realizar las entrevistas en todos los hogares encuestados. *Cada* cuestionario debe comprobarse con respecto a las partidas de esa lista y debe llevarse un registro escrito de tales comprobaciones. En cuarto lugar, los supervisores deben realizar visitas de improviso a los entrevistadores, para observarlos mientras trabajan. De esa manera se garantizará que los entrevistadores se encuentren donde se supone que deben estar. Además, el supervisor debe observar al entrevistador mientras realiza la entrevista en un hogar, para verificar que sigue todos los procedimientos que se le han enseñado durante la capacitación. En quinto lugar, los supervisores deben seleccionar al azar algunos hogares para hacer nuevas visitas después de realizadas las entrevistas. Debería prepararse otra lista de comprobación más detallada para realizar una "minientrevista" sobre algunos puntos clave (por ejemplo, cuántas personas viven de hecho en el hogar) con el fin de garantizar que el entrevistador ha registrado correctamente la información básica en el cuestionario. En sexto lugar, en lo que se refiere a los equipos de personal que se desplazan de un lugar a otro, el plan de trabajo sobre el terreno debe organizarse de tal manera que el supervisor acompañe a los entrevistadores mientras van de un lugar a otro para realizar sus

entrevistas; después de todo, es muy poca la supervisión que se puede llevar a cabo cuando el supervisor se encuentra lejos de los entrevistadores.

28. Pueden formularse otras dos recomendaciones sobre la supervisión y la garantía de los datos. En primer lugar, debe considerarse seriamente la posibilidad de introducir los datos sobre el terreno utilizando computadoras portátiles, con programas que puedan comprobar la coherencia interna de los datos introducidos. Las contradicciones observadas pueden resolverse haciendo que el entrevistador regrese al hogar para obtener la información correcta[5]. En segundo lugar, los miembros del equipo básico deberían realizar visitas sin previo aviso a los diferentes equipos de la encuesta. Estas visitas son fundamentalmente un medio de observar a los supervisores, cuyo trabajo también debe ser objeto de comprobación.

[5] El uso de computadoras portátiles sobre el terreno no está carente de problemas. Algunos son los cortes de energía eléctrica, las averías por causa del polvo, el calor o la humedad excesiva, y naturalmente el alto costo de muchas de estas computadoras.

3. Gestión de datos

29. Una tarea fundamental en toda encuesta es introducir los datos y darles una forma que pueda someterse a un proceso de análisis de datos. En la mayor parte de los casos la entrada de datos se realiza ahora utilizando computadoras personales con programas específicos para ese fin. Los programas deben concebirse de tal manera que permitan comprobar la coherencia lógica de los datos. Si se encuentran contradicciones, se puede al menos comprobar el trabajo del personal de entrada de datos para determinar si la causa son algunos errores sencillos cometidos al introducir los datos. La adopción de un sistema todavía mejor, que permitiera al entrevistador regresar al hogar de la entrevista para corregir las contradicciones, sería posible si la entrada de datos se ha realizado sobre el terreno, pero resulta casi imposible si se ha llevado a cabo en la sede central de las organizaciones que realizan la encuesta.

30. El sistema de gestión de datos debe funcionar de tal manera que los datos lleguen a un lugar central lo antes posible. Ello es importante por dos razones. En primer lugar, el trabajo realizado durante la primera semana o el primer mes debe comprobarse inmediatamente para ver que no hay graves problemas en los datos que llegan a la oficina central. En segundo lugar, en casi todos los casos, cuanto antes llegue la información a manos de los analistas y las autoridades, más valiosa será.

31. Puede darse también algún asesoramiento más específico acerca de la gestión de datos. En primer lugar, debe mantenerse una contabilidad completa de todos los hogares incluidos en la muestra de tal manera que se pueda saber si se trata de unidades que han respondido, no han respondido o no reúnen los requisitos. Esta información es necesaria para la ponderación de los registros de datos de los informantes con vistas al análisis. En segundo lugar, el programa de entrada de datos debe comprobarse exhaustivamente antes de su utilización. Un momento excelente para ello es durante la prueba piloto del cuestionario. En tercer lugar, antes de presentar los datos a los investigadores y analistas debe examinarse cada una de las partes del conjunto de datos para comprobar que no se ha excluido por error ningún hogar ni que ninguno se ha incluido más de una vez. En cuarto lugar, debe prepararse un documento de información básica y presentarse a los analistas, para tener la certeza de que comprenden cómo deben utilizar los datos. Esto se explica con mayor detalle en la sección siguiente.

D. Actividades necesarias después de la conclusión del trabajo sobre el terreno y la entrada y procesamiento de los datos

32. Una vez terminadas todas las entrevistas, deben realizarse todavía algunas actividades para garantizar el éxito de una encuesta de hogares. Todas ellas tienen lugar normalmente en la sede central de la organización que recopiló los datos. La tarea más obvia es el análisis de datos, que se examina con detalle en otros lugares de esta publicación, pero hay que realizar también algunas otras importantes actividades al final del proceso.

1. Informe final

33. Todos los supervisores, y si es posible todos los entrevistadores y encargados de la entrada de datos, deberían participar en una reunión con el equipo básico para examinar los problemas encontrados, las ideas que permitirían eliminarlos en futuras encuestas y, en términos más generales, las posibles sugerencias para mejorar los resultados en el futuro. Esta reunión debería celebrarse inmediatamente después de terminada la encuesta y antes de que el personal encargado de la entrada de los datos olvide los detalles de su experiencia. Deben mantenerse registros detallados de las recomendaciones hechas para que puedan utilizarse cuando se organice la siguiente encuesta similar.

2. Preparación del conjunto final de datos y documentación

34. Los datos de casi todas las encuestas de hogares mantendrán su validez durante muchos años probablemente, y tanto el organismo que recopiló los datos como otros organismos de investigación (o investigadores individuales) pueden preparar numerosos informes y análisis en años sucesivos. Para evitar confusiones, debería prepararse una versión final oficial del conjunto de datos, que serviría como base para *todos* los análisis de *todas* las organizaciones e individuos que utilicen esos datos. Si es posible, esta versión final de los datos debería estar preparada antes de transcurridos dos o tres meses desde la recopilación de los datos. Por ello los datos recogidos sobre el terreno deben comprobarse y analizarse rigurosamente para descubrir posibles errores y anomalías que deban corregirse, o al menos señalarse. Naturalmente, algunos errores podrían salir a la luz después de transcurridos varios meses o incluso años, en cuyo caso podría prepararse un conjunto de datos revisado para todos los análisis posteriores.

35. Todo analista de datos se formula muchos interrogantes, que pueden ir desde preguntas triviales sobre cómo se han creado los archivos de datos hasta otras de mayor importancia acerca de la forma exacta de su recopilación. Con el fin de evitar verse inundados de solicitudes de aclaración que podrían ocupar una gran parte del tiempo del personal, los organismos que recogen los datos deberían preparar un documento en el que se explique cómo se recopilaron los datos, cómo se crearon los archivos y cómo se ha dado el formato. Esta documentación contendrá descripciones de los códigos que no se encuentren en los cuestionarios de la encuesta, así como explicaciones sobre los posibles casos en que la recopilación de datos no coincida con los planes iniciales. En teoría, el documento indicará la diferencia entre la muestra final y la prevista inicialmente; en otras palabras, cuántos hogares no pudieron encontrase o se negaron a participar y (en su caso) cómo se eligieron nuevos hogares en sustitución de aquéllos. Además de este documento, el paquete estándar de información para todo analista de datos debería incluir un ejemplar del cuestionario y todos los manuales de capacitación.

36. Una última cuestión referente a la documentación es, en muchos países, la traducción a otros idiomas. Hoy, muchos investigadores estudian países cuyas lenguas no comprenden, utilizando traducciones de los cuestionarios y otros documentos. En vez de dejar que muchos investigadores realicen sus propias traducciones, quizá inexactas, resulta más práctico traducir todo el material necesario para el análisis de datos a un idioma internacional común. El inglés es el primer candidato (otros son el francés y el español). Si bien ello puede representar una fuerte carga económica, quizá se pueda incluir el costo de esta traducción en el presupuesto inicial de la encuesta y pedir que los donantes aporten fondos específicamente con este fin.

3. Análisis de datos

37. Todos los datos se recopilan con vistas al análisis, por lo que no haría falta señalar que la actividad final después de la recopilación de los datos es su análisis. Como este tema

se examina en muchos otros capítulos, aquí nos limitaremos a señalar que el plan general de la encuesta debe formular una estimación realista del tiempo necesario para analizar los datos, e incorporar esta estimación en el calendario general de las actividades de la encuesta. Los analistas de datos casi siempre tardan más de lo previsto, pero es probable que las conclusiones basadas en los datos sean más precisas, y más útiles, cuanto más intensa sea la consulta entre el equipo de la encuesta y los individuos que analizaron los datos.

E. Observaciones finales

38. En este capítulo se han formulado algunas recomendaciones generales sobre la realización de encuestas de hogares en los países en desarrollo. Se han tratado muchos temas, pero brevemente; no podía ser de otra forma, dado que las encuestas de hogares son actividades complejas. Como la información facilitada en este capítulo es insuficiente para poder llevar a cabo una encuesta de hogares, quienes deseen realizarla deben consultar otros materiales para obtener orientación más detallada. Deberían leer las referencias bibliográficas citadas en la introducción del capítulo; además, siempre es conveniente examinar las experiencias de encuestas anteriores en el país en cuestión con las personas o grupos que las realizaron. La realización de encuestas puede ser una tarea tediosa, pero el esfuerzo, la atención al detalle y la aplicación de las orientaciones facilitadas en este capítulo pueden representar una diferencia espectacular en la calidad, y por lo tanto en la utilidad, de los datos recopilados.

REFERENCIAS BIBLIOGRÁFICAS

Casley, Dennis y Denis Lury (1987). *Data Collection in Developing Countries*. Oxford, Reino Unido: Clarendon Press.

Cochran, William (1977). *Sampling Techniques,* tercera edición. Nueva York: Wiley.

Grosh, Margaret y Juan Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No. 126. Washington, D.C.: Banco Mundial.

Kalton, Graham (1983). *Introduction to Survey Sampling*. Beverly Hills, California: Sage Publications.

Kish, Leslie (1965). *Survey Sampling*. Nueva York: Wiley.

Lohr, Sharon (1999). *Sampling: Design and Analysis*. Pacific Grove, California: Duxbury Press.

Naciones Unidas (1984). *Manual de encuestas sobre hogares (versión revisada*). Estudios de métodos, No. 31. No. de venta: S.83.XVII.13.

Sección B

Diseño de la muestra

Introducción

Vijay Verma
Universidad de Siena
Siena, Italia

1. En la sección A de la presente publicación se ha presentado una amplia introducción sobre las grandes cuestiones técnicas del diseño y realización de las encuestas de hogares. Además del diseño del cuestionario se han descrito a grandes rasgos las cuestiones relacionadas con la ejecución de la encuesta y el diseño de la muestra. La presente sección es un examen, en términos más específicos, de algunas cuestiones relativas al diseño de las muestras para las encuestas de hogares en el contexto de los países en desarrollo y en transición. Contiene tres capítulos, uno sobre el diseño de marcos muestrales maestros y muestras maestras para las encuestas de hogares y dos relativos a la estimación de los efectos del diseño y su utilización en el diseño de muestras.

2. El objetivo de una encuesta por muestreo es realizar estimaciones o deducciones de aplicabilidad general para una población que es objeto de estudio, basadas en observaciones realizadas sobre un número limitado (una muestra) de unidades de la población. Este proceso está sujeto a varios tipos de errores resultantes de diversas fuentes. En general, se establece una distinción entre los errores de muestreo y los errores no muestrales. No obstante, si se considera desde la perspectiva de todo el proceso de la encuesta hay una distinción más fundamental entre "errores de medición" y "errores de estimación". Los errores de medición —los que se producen cuando lo que se mide en las unidades incluidas en la encuesta no coincide con los valores reales (verdaderos) de dichas unidades— están relacionados con la exactitud de la medición en el plano de las unidades individuales enumeradas en la encuesta y se centran en el contenido sustantivo de ésta. Se distinguen de los errores en la estimación que se producen en el proceso de extrapolación de las unidades concretas enumeradas a toda la población estudiada, para lo cual se necesitan estimaciones o inferencias. Los errores de estimación —los que afectan a la posibilidad de generalización de las unidades observadas a la población objetivo— se centran en el proceso de diseño de la muestra y de ejecución. Entre estos errores se incluyen, aparte de la variabilidad del muestreo, diversos sesgos asociados con la selección de la muestra y con la realización de la encuesta, como los errores de cobertura y de falta de respuesta. Todos estos errores son un motivo básico de preocupación para los estadísticos. Muchas veces varias encuestas o rondas de encuestas tienen el mismo marco muestral, la misma muestra maestra, el mismo diseño de la muestra. En otros casos tienen incluso una muestra común de unidades. En tales situaciones los errores relativos al proceso de muestreo suelen ser comunes a dichas encuestas, y dependen menos de la materia considerada.

3. Esta diferencia entre medición y estimación orienta la selección de los temas considerados en esta sección. En los capítulos de la sección B se abordan dos importantes aspectos de la estimación: el marco muestral, que determina hasta qué punto llega la cobertura de la población de interés e influye en el costo y eficiencia de los diseños de muestreo que se puedan formular, y el efecto del diseño, que ofrece una medida cuantitativa de dicha eficiencia y puede ayudar a relacionar la estructura del diseño con los costos de la encuesta.

Naturalmente, hay otros aspectos del diseño, por lo que sería útil estudiar los capítulos de esta sección haciendo referencia al marco presentado en la sección anterior, en particular lo relativo a los principios básicos y a los métodos de diseño de la muestra que se presentan en el capítulo II.

4. En el capítulo V se examinan con gran detalle los conceptos de una muestra maestra y de un marco muestral maestro. La definición de la población a la que se deben generalizar los resultados de la muestra es un aspecto fundamental de la planificación y diseño de la encuesta. La población que debe ser objeto de la encuesta tiene que estar representada en una manera física a partir de la cual se puedan seleccionar las muestras necesarias. Esa representación es precisamente el marco muestral. En el caso más sencillo, el marco es meramente una lista explícita de todas las unidades de la población; en los diseños más complejos la representación del marco puede ser en parte implícita, pero no obstante da cuenta de todas las unidades. En la práctica, el marco necesario se define en relación con la estructura requerida de las muestras y el procedimiento para seleccionarlas. En los marcos por etapas, que en el caso de las encuestas de hogares están basados en las zonas correspondientes, la durabilidad del marco disminuye a medida que descendemos en la jerarquía de las unidades. En un extremo, el marco muestral primario representa una importante inversión para su utilización a largo plazo. En el otro extremo las listas de las unidades últimas (como direcciones, hogares y, especialmente, personas) deben ser objeto de actualizaciones frecuentes.

5. El marco para la primera fase de muestreo (denominado marco muestral primario) debe cubrir toda la población de unidades primarias de muestreo (UPM). Después de la primera etapa de selección, la lista de unidades en una de las etapas inferiores es necesaria solamente dentro de las unidades de la etapa superior seleccionadas en la etapa precedente. Por razones de economía y por comodidad, una o varias etapas de esta tarea pueden combinarse o compartirse entre distintas encuestas. La muestra resultante de las etapas compartidas se conoce como muestra maestra. El objetivo es ofrecer una muestra común de unidades hasta una determinada etapa, a partir de la cual pueden realizarse nuevos muestreos para las diferentes encuestas. Los objetivos al utilizar una muestra maestra son, entre otros, los siguientes:

a) Economizar, compartiendo los costos de elaboración y los de mantenimiento de marcos muestrales y los materiales entre diferentes encuestas;

b) Reducir el costo del diseño y selección de la muestra;

c) Simplificar el proceso técnico de selección de muestras individuales;

d) Facilitar vinculaciones sustantivas y operacionales entre diferentes encuestas, incluidas las rondas sucesivas de una encuesta permanente;

e) Facilitar y restringir y controlar, en caso necesario, la selección de muestras múltiples para varias encuestas a partir del mismo marco.

6. También es importante reconocer que en la práctica las muestras maestras tienen sus limitaciones:

a) La reducción de costos puede ser pequeña cuando el concepto de muestra maestra no puede ampliarse a las etapas inferiores de muestreo, cuando las unidades implicadas son menos estables y los marcos correspondientes y las listas deben actualizarse con frecuencia;

b) Puede conseguirse un ahorro razonable únicamente si la muestra maestra se utiliza para más de un encuesta, y preferiblemente en muchas de ellas;

c) El uso eficaz de una muestra maestra requiere planificación a largo plazo, lo que no es fácil en las circunstancias imperantes en los países en desarrollo;

d) La falta de flexibilidad en el diseño de encuestas individuales para acomodarse a una muestra maestra puede representar un problema;

e) Puede aumentar la complejidad técnica asociada con la elaboración de muestras individuales; en cualquier caso, se requiere el mantenimiento detallado y preciso de la documentación en una muestra maestra.

7. Es posible ampliar la idea de una muestra maestra para incluir no una muestra, sino toda la población de UPM. Éste es el concepto de marco muestral maestro que se examina en el capítulo V. La inversión en un marco de este tipo vale la pena cuando los marcos disponibles no abarcan totalmente la población de interés y/o no contienen información para seleccionar con eficiencia y facilidad las muestras. El uso de un marco muestral maestro reduce también los obstáculos relativos al tipo y al tamaño de las muestras que se pueden seleccionar a partir de una muestra maestra más restringida.

8. En los capítulos VI y VII se examina el importante concepto de efecto del diseño. El efecto del diseño (o su raíz cuadrada, que algunas veces se conoce con el nombre de factor de diseño) es una medida sinóptica global del efecto producido en la varianza de una estimación por los distintos niveles de complejidad del diseño. Para un estadístico dado, se calcula como el coeficiente entre su varianza con el diseño efectivo y la varianza que habría habido con una muestra aleatoria simple (MAS) del mismo tamaño. De esta manera, permite medir la eficiencia del diseño. Al considerar el coeficiente entre la varianza efectiva y la varianza de la MAS, el efecto del diseño elimina también el efecto de factores comunes a ambos, como el tamaño de la estimación en la escala de la medición, la varianza de la población y el tamaño de la muestra global. Eso significa que la medida es más "transferible" de una situación (encuesta, diseño) a otra. Estas dos características del efecto del diseño —como medida sinóptica y como medida transferible de la eficiencia del diseño— determinan la gran utilidad y uso generalizado de esta medida en las encuestas. El cálculo y el análisis de los efectos de diseño de muchos estadísticos, así como de las estimaciones sobre subpoblaciones diversas, son de valor incalculable para la evaluación de los diseños actuales y para el diseño de nuevas muestras.

9. Aunque la magnitud del efecto del diseño elimina algunas fuentes importantes de variación en la magnitud del error de muestreo antes mencionado, depende también de otras características del diseño como la cantidad y la manera de selección de los hogares o personas dentro de las áreas incluidas en la muestra. No hay un único efecto del diseño que describa la eficiencia de muestreo de "ese" diseño. Con un mismo diseño, diferentes tipos de variables y estadísticos pueden tener (y con frecuencia tienen) diferentes valores del efecto del diseño, lo mismo que ocurre con las diferentes estimaciones de la misma variable sobre subpoblaciones diferentes. Esta diversidad de los valores del efecto del diseño en las diferentes encuestas y dentro de ellas encuentra su ilustración en los numerosos resultados empíricos, que abarcan los diferentes tipos de variables de diez encuestas en seis países, como puede verse en el capítulo VII.

Capítulo V

Diseño de marcos muestrales maestros y muestras maestras para las encuestas de hogares en países en desarrollo

Hans Pettersson
Statistics Sweden
Estocolmo, Suecia

RESUMEN

En este capítulo se consideran cuestiones relativas al diseño de marcos muestrales maestros y muestras maestras. La sección A es una presentación general. La sección B es una breve exposición de las razones para elaborar y utilizar marcos muestrales maestros y muestras maestras; la sección C contiene un examen de las principales cuestiones que se plantean en el diseño de un marco muestral maestro; en la sección D se consideran las muestras maestras y las decisiones importantes que deben adoptarse durante la etapa de diseño (elección de las UPM, número de etapas de muestreo, estratificación, asignación de la muestra con respecto a los estratos, etcétera).

Términos clave: Marco muestral maestro, muestra maestra, diseño de la muestra, muestra por etapas.

A. Introducción

1. Las oficinas nacionales de estadística de los países en desarrollo suelen ser la principal fuente de estadísticas nacionales "oficiales". En este contexto, las oficinas nacionales de estadística deben considerar una gran variedad de necesidades de información en relación con las estadísticas demográficas, sociales y económicas. Por ello utilizan diferentes fuentes y métodos para recopilar los datos. Los datos y registros administrativos pueden consultarse, hasta cierto punto, pero las encuestas por muestreo constituirán siempre un importante método de recopilación. La mayor parte de las oficinas nacionales de estadística de los países en desarrollo realizan varias encuestas cada año. Algunas de éstas (el Estudio de medición de los niveles de vida, las encuestas demográficas y de salud o las encuestas agrupadas de indicadores múltiples, por ejemplo) son de diseño bastante uniforme, mientras que otras se adaptan a las demandas nacionales específicas. La necesidad de planificación y coordinación de las actividades de la encuesta ha estimulado los esfuerzos por integrarlas en programas de encuestas de hogares. La programación especial de las encuestas ha dejado paso en muchas oficinas nacionales de estadística a planes a largo plazo en los que las encuestas sobre temas diferentes se realizan de forma constante o a intervalos periódicos. El Programa de las Naciones Unidas para desarrollar la capacidad nacional de ejecutar encuestas por hogares (PDCNEEH) ha desempeñado un papel importante en este proceso.

2. Un programa de encuestas de hogares permite la integración del diseño y de operaciones de la encuesta de diversas maneras. Pueden utilizarse los mismos conceptos y definiciones para las variables que se presentan en varias encuestas. El intercambio de personal y de servicios entre las diferentes encuestas permite el empleo eficaz del personal y las instalaciones. La integración puede incluir también el uso de marcos muestrales y muestras comunes para todas las encuestas del programa. El establecimiento de un marco muestral maestro (MMM) y de una muestra maestra (MM) para las encuestas constituye muchas veces una parte importante de un programa integrado de encuestas de hogares.

3. El uso de un marco muestral maestro común de las unidades de área para la primera etapa de muestreo mejorará la eficiencia en función de los costos en un programa de encuestas de hogares. El costo de elaboración de un marco muestral suele ser alto; el establecimiento de un programa continuado de encuestas posibilita que la oficina nacional de estadística distribuya los costos de creación de un marco muestral entre las diferentes encuestas.

4. El reparto de costos puede llevarse todavía más allá si las encuestas seleccionan sus muestras como submuestras de una muestra maestra común seleccionada del MMM. El uso de una muestra maestra para todas o para la mayoría de las encuestas reduce el costo de selección de la muestra y de preparación de los marcos muestrales en la segunda y posteriores etapas de selección en cada encuesta. Esta reducción de costos con el MMM y las MM se observa también en las encuestas imprevistas especiales realizadas durante el período del programa de encuestas y también en el caso en que no exista ningún programa oficial de encuestas en la oficina nacional de estadística.

5. En este capítulo se consideran algunos temas relativos al diseño de marcos muestrales maestros y muestras maestras para encuestas de hogares. El manual *Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares: Marcos de muestreo y diseños muestrales para programas integrados de encuestas de hogares* (Naciones Unidas, 1986) contiene una descripción adecuada de los pasos que componen el proceso de diseño, preparación y mantenimiento de un marco muestral maestro y de una muestra maestra. El manual incluye un anexo con varios estudios de casos. El lector interesado puede consultar provechosamente esa publicación, donde se presenta un examen detallado del tema.

B. Marcos muestrales maestros y muestras maestras. Panorama general

1. Marcos muestrales maestros

6. Como se describe en el capítulo II, las muestras de hogares en los países en desarrollo normalmente se seleccionan en varias etapas de muestreo. Las unidades de muestreo utilizadas en la primera etapa se conocen con el nombre de unidades primarias de muestreo (UPM). Se trata de unidades de área. Pueden ser subdivisiones administrativas, como los distritos o barrios, o pueden ser zonas delimitadas para un objetivo específico, como las zonas de empadronamiento censal (ZE). La segunda etapa consiste en una muestra de unidades secundarias de muestreo (USM) seleccionadas dentro de las UPM previamente elegidas. Las unidades de muestreo de la última etapa se conocen como unidades últimas de muestreo (UUM). Se necesita un marco muestral —lista de unidades entre las que se selecciona la muestra— para cada una de las etapas de selección en una muestra en varias etapas. El marco de muestro para las unidades de la primera etapa debe abarcar de manera exhaustiva y sin superposiciones toda la población de la encuesta, pero los marcos de muestreo de la segunda etapa se necesitarían únicamente dentro de las UPM seleccionadas en la etapa precedente.

7. Si las UPM son unidades administrativas, quizá exista una lista de estas unidades. En general podría confeccionarse una fácilmente a partir de los registros administrativos y

utilizarla como marco muestral. Esta lista especial de UPM podría prepararse cada vez que se necesite una muestra. No obstante, cuando vaya a haber una serie de encuestas a lo largo de un cierto período, sería mejor preparar y mantener un marco muestral maestro que sea accesible en cada nueva encuesta. De esa manera los gastos serían bastante menores que si hubiera que preparar marcos muestrales para cada ocasión. Asimismo, el hecho de que el marco se utilice en varias encuestas justificará los costos de su elaboración y mantenimiento y el gasto de recursos para mejorar la calidad del marco.

8. Un marco muestral maestro es básicamente una lista de unidades de área en que se abarca todo el país. Con respecto a cada unidad puede haber información sobre su clasificación como urbana/rural, la identificación de unidades de nivel superior (por ejemplo, el distrito y la provincia a que pertenece la unidad), los recuentos de población y quizá otras características. En cada una de las unidades debe haber también información sobre sus límites. El MMM para las encuestas de hogares en la República Democrática Popular Lao, por ejemplo, contiene una lista de aproximadamente 11.000 aldeas. En cada una de ellas hay información sobre el número de hogares, número de hembras y varones, carácter urbano o rural (las subdivisiones administrativas de las zonas urbanas se denominan también aldeas) e información sobre el distrito y provincia a la que pertenece la aldea. Hay también información sobre si la aldea es accesible por carretera.

9. El tipo más común de MMM es aquel en que las ZE constituyen las unidades de marco básicas. Normalmente, sobre cada unidad hay información que la relaciona con las unidades de nivel superior (subdivisiones administrativas). De ese MMM es posible seleccionar muestras de ZE directamente. También se pueden seleccionar muestras de subdivisiones administrativas y elegir muestras de las ZE dentro de las subdivisiones seleccionadas.

10. Un MMM actualizado y flexible presenta otras ventajas además de las relacionadas con el costo y la calidad que se han examinado antes. Facilita la selección rápida y fácil de muestras para las encuestas de diferentes tipos y podría cumplir los diferentes requisitos para la muestra establecidos en las encuestas. Otra ventaja es que un MMM es un elemento valioso para el siguiente censo de población. El censo mismo requiere un marco semejante al que se utilizará para las encuestas de hogares. Elaborar el marco para el censo será bastante mas fácil si ya se ha utilizado durante el período entre censos un marco muestral y se ha mantenido debidamente. La situación ideal es aquella en que el nuevo MMM se planifica y elabora durante el período del censo y luego se actualiza plenamente durante el siguiente censo.

2. Muestras maestras

11. A partir de un marco muestral maestro es posible seleccionar las muestras para encuestas distintas de manera totalmente independiente. No obstante, en muchos casos es más ventajoso seleccionar una muestra de gran tamaño —una muestra maestra— y luego seleccionar submuestras de la misma para encuestas diferentes pero relacionadas. Muchas oficinas nacionales de estadística han decidido establecer una muestra maestra para atender las necesidades de sus encuestas de hogares.

12. Una muestra maestra es una muestra a partir de la cual se pueden seleccionar submuestras para atender las necesidades de más de una encuesta o ronda de encuestas (Naciones Unidas, 1986). Puede adoptar diversas formas. Una muestra maestra con diseño sencillo y más bien común es la integrada por UPM en las que éstas son ZE. La muestra se utiliza para la selección de una muestra en dos etapas en la que las unidades de muestreo de la segunda etapa (USM) son unidades de habitación u hogares.

13. El submuestreo puede llevarse a cabo de muchas maneras. Un submuestreo en el nivel primario (de UPM) arrojaría una submuestra única de las UPM de la muestra

maestra para cada encuesta, es decir, cada encuesta tendría una muestra diferente de ZE. El submuestreo en el nivel secundario daría una submuestra de las unidades de habitación de cada UPM de la muestra maestra, es decir, cada encuesta tendría la misma muestra de ZE pero diferentes muestras de unidades de vivienda dentro de las ZE. El submuestreo podría llevarse a cabo de manera independiente, o emplearse algún tipo de proceso de selección controlado para garantizar que la superposición entre muestras sea del nivel deseado. Otra manera de seleccionar las muestras a partir de una muestra maestra sería seleccionar reiteraciones independientes de la muestra. Una o varias de las reproducciones podrían seleccionarse como submuestra para cada encuesta. De esta manera se aseguraría que la muestra maestra se construyó desde el comienzo a partir del conjunto de reiteraciones totalmente independientes.

14. Una oficina nacional de estadística puede reducir notablemente los costos utilizando una muestra maestra, ya que los gastos se distribuirían entre todas las encuestas que utilizan la MM. En consecuencia, disminuirían los costos por encuesta. Como la selección de las unidades de la muestra maestra es básicamente una operación de oficina (sobre todo si existe un MMM de calidad), la reducción del costo en esta etapa puede ser modesta. Será mucho mayor cuando se distribuyan entre las encuestas los costos de preparación de mapas y submuestreo de marcos de unidades de vivienda dentro de las unidades de la muestra maestra. Las actividades sobre el terreno necesarias para establecer los marcos de muestreo suelen ser considerables, y el costo por encuesta de dichas actividades disminuirá casi proporcionalmente a la cantidad de encuestas que utilicen el mismo marco de submuestreo.

15. En algunos países las dificultades y los costos asociados con los viajes podrían hacer aconsejable, desde el punto de vista económico, la contratación de entrevistadores en el ámbito de las unidades primarias de muestreo de la MM o en sus proximidades y que se instalen allí durante todo el período de la encuesta. En ese caso se utilizan UPM relativamente grandes. De esa manera la utilización de una muestra maestra fija de esas UPM, en vez de seleccionar una nueva muestra para cada encuesta y hacer que los entrevistadores se desplacen o contratar nuevos entrevistadores, puede suponer grandes beneficios.

16. El uso de las mismas unidades de la muestra maestra reducirá el tiempo de espera hasta el comienzo de las encuestas en la zona. En muchos países en desarrollo el entrevistador tiene que conseguir permiso de las autoridades regionales y locales para realizar las entrevistas en el área. En países como la República Democrática Popular Lao y Viet Nam, por ejemplo, hay que obtener permiso de las autoridades administrativas locales, y a veces hasta del dirigente de la aldea. El tiempo necesario para "montar el negocio" se reducirá notablemente si se utilizan las mismas zonas para varias encuestas.

17. El uso de las UPM de la misma muestra maestra para varias encuestas reducirá el tiempo necesario para que el entrevistador encuentre los hogares. Cuando se dispone de mapas y submarcos muestrales de buena calidad, el entrevistador puede recorrer rápidamente la zona; en algunos casos quizá haya trabajado allí en una encuesta anterior. Puede introducirse una numeración permanente de unidades de habitación para facilitar la orientación en la zona. Así se ha hecho en algunas muestras maestras: Torene y Torene (1987) describen el caso de la muestra maestra de Bangladesh.

18. La MM permite superponer muestras en dos o más encuestas. Ello hace posible integrar los datos en los niveles más detallados mediante la vinculación de los datos de los hogares de las diferentes encuestas. No obstante, existe el riesgo de que se produzcan efectos negativos en la calidad de los resultados de la encuesta cuando se utilizan unidades muestrales en varias ocasiones. Los hogares que participan en varias rondas de una encuesta o en varias encuestas pueden resistirse a participar, o quizá estén menos motivados para responder con precisión en las encuestas sucesivas.

19. Una MM presenta ventajas (costos, integración y coordinación) para las encuestas periódicas en un programa de encuestas. Una MM permitirá también a la oficina nacional de estadística estar mejor preparada para realizar el muestreo en las encuestas especiales: las submuestras pueden seleccionarse rápidamente a partir de la MM cuando se necesiten para encuestas ocasionales.

20. Las ventajas de las muestras maestras son obvias, pero hay también algunas desventajas o limitaciones. El diseño de la muestra maestra representa siempre un compromiso entre diferentes requisitos de diseño resultantes de las encuestas incluidas en el programa. La muestra maestra será válida para las encuestas que tengan requisitos de diseño razonablemente compatibles con respecto a las estimaciones del campo y la distribución de la población objetivo dentro de esas zonas. El diseño elegido para la muestra maestra normalmente encajará de forma aceptable con la mayor parte de las encuestas incluidas en el programa, aunque no a la perfección. El diseño de la muestra maestra impone algunas limitaciones y requisitos (referentes al tamaño de la muestra, los conglomerados, la estratificación, etcétera) a las encuestas individuales, y en algunos casos quizá no sea fácil tenerlos en cuenta. El resultado será cierta pérdida de eficiencia en las diferentes encuestas.

21. Hay también encuestas con necesidades especiales de diseño que la muestra maestra no podrá satisfacer en absoluto, a saber:

- Encuestas destinadas a ciertas zonas regionales o locales cuando se necesita una muestra amplia para una zona pequeña (por ejemplo, encuestas utilizadas para evaluar los efectos de un proyecto de desarrollo en una zona local).
- Encuestas destinadas a subgrupos de población (por ejemplo, étnicos) irregularmente distribuidos.

22. Un ejemplo del primer tipo es la encuesta de cultivo de opio que se hace periódicamente en algunas zonas de cuatro provincias septentrionales de la República Democrática Popular Lao para evaluar el progreso del proyecto del gobierno nacional para reducir el cultivo de ese producto. En este caso, como la muestra maestra no podía atender los requisitos relativos al diseño de la muestra, se seleccionó para la encuesta una muestra independiente (la alternativa habría sido utilizar las UPM de la muestra maestra en las cuatro provincias y seleccionar UPM adicionales del marco muestral maestro).

23. En algunos casos la reducción de costos de una muestra maestra quizá no se haga plenamente realidad. Para extraer una submuestra de una muestra maestra que responda a las necesidades específicas de una encuesta individual y luego calcular las probabilidades de selección correctamente se requiere una gran preparación técnica. Ello puede ser una operación más complicada que seleccionar una muestra independiente. El hecho de que los estadísticos de muestreo sean escasos en muchas oficinas nacionales de estadística de los países en desarrollo puede entorpecer la aplicación de una muestra maestra, o incluso impedir, su establecimiento. Hay ejemplos de muestras maestras que no se utilizan suficientemente debido a la falta de conocimientos sobre muestreo del personal de la oficina nacional de estadística.

3. Resumen y conclusión

24. Las ventajas, desventajas y limitaciones examinadas hasta aquí pueden resumirse como sigue:

Marco muestral maestro

- Eficiente en función de los costos; permite a la oficina nacional de estadística repartir entre varias encuestas los costos de creación de un marco muestral.

- La calidad será normalmente mejor que la de los marcos muestrales especiales, ya que es más fácil convencer de invertir en mejorar la calidad de un marco que se va a utilizar durante un período de tiempo más largo.
- Simplifica el proceso técnico de elaborar muestras individuales; facilita la selección rápida y sin complicaciones de las muestras para las encuestas de diferentes tipos.
- Si se mantiene debidamente, el marco conservará su valor para los próximos censos de población.

Muestra maestra

- *Ahorro de costos*
  - Los costos de selección de las unidades de la muestra maestra se distribuirán entre todas las encuestas que la utilicen.
  - Los costos de preparación de mapas y marcos de submuestreo de las unidades de vivienda u hogares se repartirán entre las encuestas que utilicen la MM; no obstante, los marcos de submuestreo deberán actualizarse periódicamente para incorporar las unidades de habitación de nueva construcción y eliminar las inexistentes.
    - Clara ventaja resultante de utilizar una MM cuando los entrevistadores deban estar situados en la UPM o en sus proximidades debido a las dificultades y altos costos asociados a los viajes sobre el terreno.
- *Operaciones más eficientes*
  - El uso de las UPM de la misma muestra maestra para varias encuestas reducirá el tiempo necesario para iniciar las encuestas en la zona y el tiempo que necesita el entrevistador para encontrar a los encuestados.
  - La MM contribuye a la selección rápida y fácil de la muestra; las submuestras de la MM pueden seleccionarse rápidamente cuando son necesarias para las encuestas especiales.
- *Integración*
  - El hecho de que la MM permita tener muestras superpuestas en dos o más encuestas hace posible la integración de los datos de las encuestas.
- *Limitaciones, desventajas*
  - La MM no es válida para todas las encuestas. Mientras dure el programa de encuestas, la oficina nacional de estadística se encontrará en situaciones en que se presenten necesidades imprevistas que no puedan atenderse con una muestra maestra (esto representa una limitación más que una desventaja).
  - Cuando vuelven a utilizarse las unidades muestrales, sobre todo en los hogares, existe el riesgo de sesgos resultantes de los efectos condicionantes y de una falta de respuesta mayor, a causa de la carga acumulada.
  - La aplicación continuada de una MM requiere técnicas de muestreo que quizá no estén al alcance de la oficina nacional de estadística.

Conclusión

25. Es claro que los marcos muestrales maestros y las muestras maestras tienen muchos rasgos atractivos. Es conveniente que toda oficina nacional de estadística elabore y mantenga debidamente un marco muestral maestro que pueda atender las necesidades de sus encuestas de hogares, independientemente de si se organizan en un programa de encuestas o en forma esporádica. Para muchas oficinas nacionales de estadística representará una ventaja dar un paso más allá y elaborar y utilizar una muestra maestra para todas o para la mayoría de las encuestas de hogares.

C. Diseño de un marco muestral maestro

26. El programa nacional de encuestas de hogares define qué es lo que se debe esperar del diseño de marcos muestrales maestros y muestras maestras en lo que respecta al número previsto de muestras, cobertura de la población, estratificación y tamaño de las muestras, por ejemplo. La forma de atender estas demandas en la labor de diseño depende de las condiciones existentes para la construcción de los marcos en los respectivos países. El factor más importante es la disponibilidad de datos y otros materiales que puedan utilizarse para la construcción del marco. A continuación, en el apartado 1 se examinan brevemente los tipos de datos y materiales que son necesarios y los problemas de calidad que pueden presentarse en esos datos.

27. Cuando se han evaluado los datos y materiales disponibles, la oficina nacional de estadística tiene que decidir las características fundamentales del MMM en relación con los siguientes aspectos:

- Cobertura del MMM (véase el apartado C.2);
- Unidades de área que deberían servir como unidades de marco en el MMM (véase el apartado C.3);
- Información sobre las unidades del marco que deberían incluirse en el MMM (véase el apartado C.4).

28. La documentación completa y bien tramitada del marco y la existencia de procedimientos claros para su actualización, son imprescindibles para el uso eficiente del MMM (véase el apartado C.5).

1. Datos y materiales: evaluación de la calidad

29. La fuente de datos y los materiales más importantes serán normalmente los censos de población más recientes. Ello es obvio cuando la oficina nacional de estadística tiene previsto utilizar las zonas de empadronamiento censal como unidades del marco; de todas formas, aun cuando se utilicen otras unidades (administrativas), normalmente se necesitan datos de población o de hogares procedentes del censo. Los materiales básicos del censo son listas de las ZE con recuentos de población y de hogares y mapas esquemáticos de las ZE. Hay también mapas de zonas más amplias (distritos, regiones) en los que se marcan las ZE. Normalmente, las ZE se identifican mediante un código que indica su condición urbana/rural y la división y subdivisión administrativa a la que pertenecen. Algunas veces el código indica también si la ZE contiene población institucional (que vive en cuarteles militares, residencias de estudiantes, etcétera).

30. La calidad de los datos y materiales del censo varía considerablemente de un país a otro. Así ocurre especialmente con los mapas. Algunos países, como Sudáfrica, han digitalizado los mapas de las ZE almacenados en las bases de datos mientras que otros, como la República Democrática Popular Lao, no tienen buenos mapas en absoluto. En algunos países los mapas de las ZE son muchas veces muy elementales y difíciles de utilizar sobre el terreno. Como las ZE pueden consistir de hecho en listas de localidades más que en unidades espaciales propiamente dichas, las poblaciones dispersas fuera de las localidades enumeradas quizá no se incluyan en dichos marcos. Un problema especial relacionado con la calidad, que resulta algo molesto para los encargados de establecer los marcos, es la dificultad de recuperar los materiales del censo, en especial los mapas. Éstos pueden ser de buena calidad, pero no sirven de mucho si existen problemas para su recuperación. El hecho de que sea práctica común que los mapas de las ZE se "sepulten" en un archivo después del censo, algunas veces en forma no muy ordenada, dificulta su búsqueda. Son también frecuentes los casos en que algunos mapas de las ZE no aparecen en los archivos.

31. En general, la calidad del material del censo se deteriora con el paso del tiempo. Así ocurre sin duda con los recuentos de población para las ZE cuyo crecimiento demográfico y migración repercutirá en ellas desigualmente. Asimismo, los cambios en las unidades administrativas, como la modificación de los límites o la división/fusión de unidades, hará que la información del censo quede casi obsoleta. En cualquier caso, la información estará inevitablemente desfasada si el último censo tuvo lugar siete u ocho años antes.

32. Un primer paso en el diseño del MMM debe ser identificar y evaluar los diferentes materiales disponibles para la construcción del marco, incluyendo tanto los materiales del censo como otros datos/materiales. Aun cuando el censo de población deba ser la principal fuente de materiales, hay otras que quizá sean necesarias para actualizar o complementar esos datos. Conviene formularse las siguientes preguntas: ¿Qué datos/materiales pueden obtenerse y cuál es su nivel de precisión? ¿Hasta qué punto continúan siendo válidos los datos y con qué frecuencia se actualizan? Los mapas deben evaluarse para determinar hasta qué punto son detallados y especifican los límites de las subdivisiones administrativas. Debería hacerse lo posible por estimar la proporción de mapas esquemáticos de las ZE que cumplen las normas de calidad exigidas.

33. En esta etapa del trabajo es también importante obtener o preparar una descripción precisa y completa de la estructura administrativa del país y una lista actualizada de sus divisiones y subdivisiones administrativas.

2. Decisión sobre la cobertura del marco muestral maestro

34. Una decisión que debe tomarse en las fases iniciales es la relativa a la cobertura del MMM. ¿Deberían excluirse del marco algunas partes muy remotas y escasamente pobladas? La decisión de la mayor parte de los países de preparar marcos muestrales maestros de cobertura nacional completa generalmente tiene sentido, ya que cuando de las encuestas periódicas del programa se excluyen algunas partes remotas y poco pobladas, pueden surgir situaciones en que se necesite una encuesta especial para obtener información sobre esos lugares. Un caso particular es el de los grupos nómadas y tribus de las montañas, que presentan dificultades para la obtención de muestras y para establecer contactos sobre el terreno. En algunos países estos grupos están excluidos de la población objetivo de los programas de encuestas de hogares.

35. Debe decidirse también sobre la cobertura de la población institucional. En algunos países las grandes instituciones se definen como zonas de empadronamiento especiales (internados, grandes hospitales, cuarteles militares y residencias de mineros). En ese caso sería posible excluir estas zonas del marco, aunque en general es mejor mantenerlas, ya que así se podría decidir sobre su cobertura en encuestas posteriores.

3. Decisión sobre las unidades de marco básicas

36. Las unidades de marco son unidades de muestreo incluidas en el marco muestral maestro. Las unidades de marco básicas son las de nivel más bajo en el marco muestral maestro. En general, conviene que correspondan a zonas pequeñas, para poder agruparlas en unidades de muestreo mayores si los problemas de costo de una encuesta obligaran a ello.

37. Las zonas de empadronamiento censal (ZE) son muchas veces la mejor opción para las unidades de marco básicas. Presentan varias ventajas. Su demarcación se lleva a cabo para conseguir zonas de tamaño parecido en términos de población, lo que representa una ventaja en algunas situaciones de muestreo. Las ZE se cartografían y en la mayoría de los casos el mapa se complementa con una descripción de los límites. Normalmente se dispone de mapas básicos en los que se indica la ubicación de la ZE dentro de las divisiones administrativas. En los censos se producen listas computadorizadas de ZE, que pueden utilizarse como

punto de partida para un MMM. Son muchas las razones que aconsejan utilizar las ZE como unidades de marco, pero en algunos casos hay problemas de calidad, como los examinados en el apartado C.1, que pueden aconsejar la búsqueda de otras soluciones.

38. Algunos países tienen subdivisiones administrativas que son lo bastante pequeñas como para servir como unidades de marco básicas, y puede haber situaciones en que esas unidades presenten ventajas con respecto a las ZE en cuanto unidades de marco básicas, como en el caso del MMM mantenido por el Centro Nacional de Estadística de la República Democrática Popular Lao. Las ZE se habían venido considerando unidades de marco básicas, pero se observó que la documentación de las ZE era difícil de recuperar y, en general, de calidad más bien deficiente, lo que obstaculizaba la utilización de los límites de las ZE sobre el terreno. En ese contexto, se decidió utilizar las aldeas como unidades de marco básicas. Las aldeas de la República Democrática Popular Lao son unidades administrativas bien definidas. No obstante, no son unidades geográficas de área en sentido riguroso. Los límites entre las aldeas son poco claros y no hay mapas propiamente dichos, pero no cabe duda de qué hogares pertenecen a una u otra aldea.

39. Aunque existen, no son frecuentes los casos en que se utilizan unidades diferentes de las ZE como unidades de marco básicas. Un ejemplo es el de Tailandia, donde las ZE de las zonas municipales se subdividen en bloques, y el empadronamiento de la población y de los hogares para el censo se lleva a cabo en cada uno de esos bloques, que se utilizaron como unidades de marco básicas en la parte municipal del MMM.

40. Las unidades de marco básicas, sean ZE o unidades de otro tipo, difieren de tamaño en lo que respecta al número de hogares y de habitantes de la zona. Aun cuando la intención sea crear ZE que no acusen una variación demasiado grande en lo que respecta al tamaño de la población, se producen desviaciones de esta norma por varias razones (por ejemplo, pueden construir ZE menores en términos de población en zonas escasamente pobladas adonde es difícil desplazarse). El resultado suele ser una variación sustancial del tamaño de la ZE, con algunos casos extremos tanto por arriba como por abajo. En Viet Nam, por ejemplo, el número medio de hogares por zona de empadronamiento es 100. El número de hogares en las 166.000 ZE varía desde un mínimo de dos hasta un máximo de 304 (Glewwe y Yansaneh, 2001). Aproximadamente el 1% de las ZE tiene 50 hogares o menos. En la República Democrática Popular Lao, la proporción de ZE pequeñas es todavía mayor: el 6% de las ZE tiene menos de 25 hogares. Esta variación demográfica en el tamaño de las zonas que se utilizan como unidades de marco básicas no suele ser en general un problema, pero las unidades muy pequeñas no pueden utilizarse como unidades de muestreo. Pueden aceptarse en el MMM, pero en el caso de las muestras basadas en el MMM estas ZE deben asociarse con las ZE adyacentes para constituir unidades de muestreo válidas.

4. Información sobre las unidades que deben incluirse en el marco

41. Una lista sencilla de las unidades de marco básicas constituye un marco muestral rudimentario, pero la posibilidad de obtener muestras eficientes de dicho marco es limitada. La utilidad del marco mejorará notablemente si contiene datos suplementarios sobre las unidades de marco que podrían utilizarse para elaborar diseños muestrales eficientes. Los datos suplementarios pueden ser de tres tipos:

a) Información que permite agrupar las unidades de marco básicas en unidades mayores. Una manera de incrementar el potencial de muestreo eficiente a partir del marco es permitir el muestreo de diferentes tipos de unidades a partir del marco. Por ello es conveniente que el marco contenga información que haga posible formar unidades mayores y de esa manera disponer de mayor flexibilidad al elegir las unidades de muestreo a partir del marco;

b) Información sobre el tamaño de las unidades. La eficiencia de las muestras del marco aumentará también si en cada una de las unidades se incluye una medición del tamaño. Ello reviste especial importancia cuando hay una gran variación en el tamaño de las unidades;

c) Otra información complementaria. La información que pueda utilizarse para la estratificación de las unidades o como variables auxiliares en la etapa de estimación aumentará la eficiencia de las muestras del MMM.

Información que permite agrupar las unidades de marco básicas en unidades mayores

42. En algunas encuestas la mejor alternativa para las UPM son las zonas pequeñas, como las zonas de empadronamiento. En otras, la consideración de los costos y de los errores de muestreo aconsejará la adopción de UPM que sean considerablemente mayores que las ZE. Estas UPM de mayor tamaño podrían formarse a partir de grupos de ZE próximas. Otra posibilidad es utilizar como UPM unidades administrativas semejantes a los barrios y distritos. En todos estos casos, es necesario que el marco muestral maestro ofrezca posibilidades para la construcción de estas UPM de mayor magnitud. Por ello es importante que los registros de la unidad de marco en los MMM contengan información sobre las unidades de nivel superior a que pertenece la unidad de marco.

43. Un modelo de diseño de un marco muestral maestro utilizado por muchos países es el que usa las zonas de empadronamiento censal como unidades de marco básicas y ordena las unidades geográficamente en unidades mayores (administrativas) de acuerdo con una estructura jerárquica. Las muestras pueden tomarse del MM de diferentes maneras: *a*) muestreando las ZE; *b*) agrupando las ZE para formar UPM de tamaño adecuado y muestreando las UPM, y *c*) muestreando las subdivisiones administrativas en la primera etapa y con muestreos posteriores en etapas adicionales hasta el nivel de la ZE. La estructura jerárquica del marco muestral maestro de Viet Nam contiene los siguientes niveles:

- Provincias
  - Distritos
    - Comunas (rurales), barrios (urbanos)
      - Aldeas (rurales), bloques (urbanos)
        - Zonas de empadronamiento censal.

44. La flexibilidad en la elección de las unidades de muestreo aumenta si se asignan identificadores a todas las unidades de marco (tanto unidades de marco básicas como de orden superior), tomando como base la proximidad geográfica. Ello permite utilizar las unidades de marco como elementos básicos para formar UPM del tamaño deseado a partir de unidades adyacentes. Este proceso resultaría necesario en los casos de Viet Nam y de la República Democrática Popular Lao descritos en la sección anterior. Otra ventaja de un identificador basado en la proximidad geográfica es que partiendo del marco muestral maestro pueden elegirse muestras geográficamente dispersas utilizando el muestreo sistemático en unidades de muestreo ordenadas geográficamente.

Medidas del tamaño de las unidades de marco

45. La inclusión de medidas del tamaño de las unidades de marco reviste especial importancia cuando hay una gran variación en el tamaño de las unidades de marco. Normalmente, las medidas de tamaño son recuentos de población, de hogares o de unidades de habitación dentro de la misma unidad de marco. Es importante observar que las medidas de tamaño no deben ser exactas. De hecho, casi siempre tienen cierto grado de imprecisión, ya que están basadas en datos de un punto cronológico anterior, el hecho de que la población

está en constante cambio hace que la información esté cada vez más desfasada. Los errores en las medidas de tamaño no dan lugar a sesgos en las estimaciones de la encuesta pero reducen la eficiencia de las medidas de tamaño, sobre todo cuando se utilizan en la etapa de estimación. Por ello debería hacerse lo posible para que esas medidas sean lo más exactas posible.

46. Las medidas del tamaño se utilizan sobre todo en la selección muestral de unidades de marco con probabilidad proporcional al tamaño. Otros usos de las medidas de tamaño son:

- Determinar la asignación de las UPM de la muestra a los estratos;
- Formar estratos de unidades clasificadas por tamaños;
- Como variables auxiliares para estimaciones de regresión o del coeficiente;
- Formar unidades de muestreo de tamaño adecuado.

Otra información complementaria para las unidades de marco

47. Debería considerarse la posibilidad de incluir en el marco información complementaria sobre las unidades de marco que pudieran obtenerse con costos razonables. La información sobre la densidad demográfica, los grupos étnicos predominantes, la actividad económica principal y el ingreso medio en las unidades de marco son variables que muchas veces resultan útiles para la estratificación.

48. En el marco muestral maestro de Namibia se incluyó una clasificación aproximada del nivel de ingresos —ingreso alto, ingreso mediano e ingreso bajo— para las unidades de marco básicas urbanas (ZE) en Windhoek, la capital, lo que permitió formar dos estratos de nivel de ingresos en el subcampo urbano de Windhoek. Otro ejemplo es el marco muestral maestro de la República Democrática Popular Lao, donde las unidades de marco rurales tienen información sobre si la unidad está próxima o no a una carretera. Las muestras para las encuestas de hogares que utilizan el marco muestral maestro se estratifican teniendo en cuenta si tienen o no acceso a una carretera.

5. Documentación y mantenimiento de un marco muestral maestro

Documentación

49. Un requisito imprescindible para poder utilizar el marco es disponer de documentación puesta al día, precisa y fácilmente accesible del marco muestral maestro. Si la documentación es deficiente, los beneficios del marco no podrán hacerse plenamente realidad. El núcleo de la documentación es una base de datos que contenga todas las unidades marco. El contenido de los registros de las unidades de marco debería ser el siguiente:

- Un identificador primario, que debería ser numérico. Debería tener un código que identifique de manera exclusiva todas las divisiones y subdivisiones administrativas en que está ubicada la unidad del marco. Será una ventaja si las unidades de marco están numeradas en orden geográfico. Normalmente, los códigos ZE tienen esas características. Los identificadores totalmente numéricos son mejores que los nombres o códigos alfanuméricos. En muchos casos, los sistemas existentes de geocodificación de las fuentes administrativas y del censo son válidos como identificadores primarios.
- Un identificador secundario, que será el nombre de la aldea (u otra subdivisión administrativa) en que esté ubicada la unidad de marco. Los identificadores secundarios se utilizan para ubicar la unidad de marco en los mapas y sobre el terreno.
- Algunas características de la unidad, como la medida del tamaño (población, hogares), el carácter urbano/rural, la densidad de población, etcétera. Deberían in-

cluirse todos los datos referentes a la unidad que puedan obtenerse con un costo razonable y sean de calidad aceptable. Las características podrían utilizarse para la estratificación, asignando probabilidades de selección y como variables auxiliares en la estimación.

- Datos operacionales, información sobre los cambios ocurridos en las unidades e indicación del uso de la muestra.

50. El marco debe ser de fácil acceso y utilización para las distintas manipulaciones, como clasificación, filtrado y producción de estadísticas sinópticas que puedan facilitar el diseño de la muestra y la estimación. Para ello lo mejor es almacenar el marco en una base de datos informatizada. Debería evitarse el uso de formatos a los que sólo puedan acceder los especialistas. Una sencilla hoja de datos de Excel será muchas veces la solución adecuada. Excel es fácil de utilizar, son muchos quienes saben usarlo y tiene funciones de clasificación, filtrado y agregación que son necesarias cuando se preparan muestras a partir del marco. La mayor parte de los demás programas informáticos pueden importar las hojas de datos.

Mantenimiento del marco muestral maestro

51. Las rutinas para el mantenimiento del marco están estrechamente vinculadas con la documentación del marco muestral maestro (MMM). Durante el tiempo de utilización del MMM se producirán cambios que influirán tanto en el número como en la definición de las unidades de marco. La cantidad de trabajo necesaria para mantener un marco muestral maestro depende sobre todo de la estabilidad de las unidades de marco. Hay dos tipos de cambios que pueden producirse en las unidades de marco: cambios en sus límites y cambios en sus características.

52. Los cambios en los límites de las unidades de marco afectan sobre todo a las subdivisiones administrativas. Éstas pueden sufrir cambios en sus límites, sobre todo en los niveles inferiores, debido a decisiones políticas o administrativas. Muchas veces esos cambios se introducen en respuesta a transformaciones importantes de la población de las zonas afectadas. Se crean nuevas unidades mediante la división o integración de las unidades ya existentes o mediante un proceso más complicado de redistribución de las unidades. Asimismo, los límites de las unidades existentes pueden modificarse sin necesidad de crear unidades nuevas. Si hay cambios frecuentes en las subdivisiones administrativas, deberán asignarse considerables recursos para que la información del marco esté actualizada y continúe siendo precisa.

53. Los cambios que afectan a los límites de las unidades de marco deben registrarse en el MMM. Debe establecerse un sistema de recopilación de información acerca de los cambios administrativos para poder efectuar su seguimiento.

54. Entre los cambios en las características de la unidad de marco se incluyen desde los más sencillos, como el cambio de nombre, hasta otros de mayor importancia, como los de la medida del tamaño (población o número de hogares/unidades de habitación) o la clasificación como zona urbana/rural. Estos cambios no deben reflejarse necesariamente en el MMM. No obstante, como ya se ha dicho, la información desfasada sobre las medidas del tamaño provoca una pérdida de eficiencia en las muestras seleccionadas a partir del marco. Las medidas de actualización del tamaño para todo el marco serían muy costosas y en general poco eficientes en función del costo, pero en lo que respecta a las zonas periurbanas de rápido crecimiento es buena idea actualizar las medidas de tamaño periódicamente.

55. Los cambios en las medidas de tamaño para las unidades de marco resultan problemáticos cuando hay cambios grandes y repentinos en la población, lo que puede ocurrir, por ejemplo, en las zonas de ocupantes sin título cuando las autoridades locales deciden acabar con la ocupación. Estos cambios notables deben reflejarse en el marco muestral. Un ejemplo de cambio menos llamativo pero ciertamente problemático (para

el marco muestral) es la migración promovida en la República Democrática Popular Lao, cuyo gobierno está alentando a los miembros de que viven en las aldeas de zonas montañosas a trasladarse a otras con mayor acceso a servicios básicos. Como consecuencia de este proceso, el número de aldeas ha disminuido aproximadamente un 10% en dos años. Sin duda estos cambios deben incluirse en el marco de la muestra.

56. Existe el riesgo de que el mantenimiento del MMM se descuide cuando una oficina nacional de estadística cuenta con pocos recursos y tiene dificultades para atender la demanda de resultados estadísticos. Por ello, es importante que la oficina nacional de estadística elabore planes y procedimientos para actualizar el marco en las primeras etapas y que se asignen recursos suficientes con ese fin.

D. Diseño de muestras maestras

57. Una muestra maestra es aquella a partir de la cual pueden seleccionarse submuestras para atender las necesidades de más de una encuesta o ronda de encuestas (Naciones Unidas, 1986). El objetivo debería ser establecer muestras para las encuestas de hogares que tienen necesidades de diseño razonablemente compatibles con respecto a los ámbitos del análisis y las distribuciones de sus poblaciones objetivo dentro de esas zonas. Una muestra maestra se define en función del número de etapas de muestreo y el tipo de unidades que sirven como unidades últimas de muestreo (UUM). Una muestra maestra seleccionada en dos etapas en zonas de empadronamiento como unidades de la segunda etapa se denominaría *muestra maestra de zonas de empadronamiento en dos etapas.* Si las ZE se seleccionaran directamente en la primera etapa, tendríamos una *muestra maestra de ZE en una sola etapa*. Ambas opciones constituyen diseños muestrales maestros comunes en países en desarrollo.

58. En los apartados D.1 a D.4 se examinan algunos pasos importantes en la elaboración de una muestra maestra, y en los apartados D.5 y D.6 se analizan cuestiones referentes a la documentación y mantenimiento de la muestra maestra. Finalmente, el apartado D.7 trata sobre la utilización de la muestra maestra para encuestas que no tienen como objetivo primario los hogares.

1. Elección de unidades primarias de muestreo para la muestra maestra

59. El MMM constituye el marco para la selección de la muestra maestra. La unidad de marco básica en el MMM podría utilizarse en algunos casos como unidad primaria de muestreo para la muestra maestra. En otras ocasiones podemos optar por constituir UPM mayores que las unidades de marco básicas del MMM. En tales casos se suelen utilizar como UPM unidades administrativas bien definidas (condados, barrios, etcétera), pero algunas veces las UPM se han construido utilizando las unidades de marco. En este caso las unidades adyacentes se agrupan en UPM del tamaño adecuado. Un ejemplo es la muestra maestra de Lesotho, donde se formaron las UPM combinando ZE adyacentes del censo en grupos integrados por 300-400 hogares. Las 3.055 ZE del censo se agruparon en 1.038 grupos de ZE que deberían servir como UPM (Pettersson, 2001a).

60. Hay varios factores relacionados con la eficiencia estadística, los costos y los procedimientos operacionales que conviene tener en cuenta al decidir cuál debe ser la unidad primaria de muestreo. Suponiendo que las unidades de marco básicas del MMM sean las ZE, ¿en qué circunstancias preferiríamos utilizar como UPM unidades mayores que las ZE?

- Si sabemos que las demarcaciones de una parte significativa de las ZE son de mala calidad, podemos decidir utilizar como UPM unidades mayores, ya que éstas tienen en general una demarcación de límites más estable y clara.

- Cuando el desplazamiento entre las diferentes zonas es difícil y/o costoso. Las dificultades y los costos relacionados con las visitas sobre el terreno podrían hacer aconsejable desde el punto de vista económico la contratación de entrevistadores dentro de las UPM de la muestra o en sus proximidades y destinarlos allí durante todo el período de la encuesta. Para ello harían falta UPM más bien grandes.
- Cuando el uso de las UPM para muestras es tan frecuente que una UPM pequeña, como una ZE, se agota rápidamente. Este problema podría solucionarse utilizando unidades mayores como UPM o manteniendo como UPM las ZE y rotando la muestra de ZE. La primera opción es preferible cuando el costo de introducción y de puesta en marcha de la encuesta en la zona sea elevado.
- Cuando por razones de control de costos y eficiencia de muestreo sea habitual introducir una o más etapas de muestreo que afecten a unidades mayores que las unidades de marco básicas. Si, por ejemplo, las unidades de marco básicas son ZE, podemos decidir utilizar como UPM unidades mayores, por ejemplo barrios, y luego seleccionar ZE u otras unidades de área dentro de la UPM en la etapa siguiente.
- Cuando, como ocurre en algunas encuestas, las variables individuales y del hogar están vinculadas con variables comunitarias. Un ejemplo es el de una encuesta de salud en que las variables de salud individual están vinculadas con otras variables referentes a las instalaciones sanitarias de la aldea o comuna. Otro ejemplo es una encuesta sobre el nivel de vida en la que las variables del hogar están vinculadas con las variables comunitarias en las escuelas, carreteras, abastecimiento de agua, saneamiento, precios locales, etcétera. Si la muestra maestra debe servir para varias encuestas de este tipo, presenta ventajas utilizar la comunidad (aldea, comuna, barrio, etcétera) como UPM. Si se utiliza la comunidad como UPM, podemos garantizar que la submuestra de USM estará bien distribuida en el conjunto de la comunidad.

61. Las unidades de área de gran tamaño no son válidas como UPM, ya que su número es demasiado reducido. No tendría sentido tomar una muestra de una población de 50-100 unidades. Si es posible, el número de UPM de la población debería ser de más de 1.000, a fin de que una muestra del 10% permita contar con más de 100 UPM para la muestra. Una fracción muy superior al 10% no permitirá aprovechar la reducción de costos asociada al muestreo. Un número de muestras de UPM muy inferior a 100 aumentaría la varianza. Conviene también señalar que podría ser eficiente utilizar diferentes tipos de UPM en diferentes partes de la población; por ejemplo, ZE en las zonas urbanas y en las unidades mayores de las zonas rurales.

2. Combinación/división de zonas para reducir la variación en el tamaño de las UPM

62. Cuando se ha tomado una decisión referente al tipo de unidad que debe servir como UPM (y, en el caso de que haya varias etapas, qué unidad debería servir como USM), podemos observar que hay "valores atípicos" que son mucho mayores o menores de lo deseable.

Unidades de muestreo muy pequeñas

63. Las UPM muy pequeñas de la muestra maestra presentan problemas. A la hora de decidir qué tamaño debe considerarse aceptable hay que tener en cuenta la carga de trabajo prevista para la muestra maestra. El Departamento de Estadística de Sudáfrica (Statistics South Africa), que está utilizando ZE censales como UPM para su muestra maestra, decidió establecer la cantidad de 100 hogares como tamaño mínimo de las UPM. Las ZE que tienen menos de 100 hogares se han asociado con ZE próximas durante la preparación

del MMM. Para su muestra maestra, la Oficina Nacional Central de Estadística de Namibia aplicó la norma de que las UPM deberían contener al menos 80 hogares. En el censo se formaron 2.162 ZE. Después de unir las ZE pequeñas a las adyacentes se mantuvieron 1.696 UPM. De las 1.696 UPM, 405 se formaron mediante la agregación de varias ZE; cada una de las 1.291 restantes tenía una sola ZE.

64. La integración de las pequeñas ZE antes de la selección puede ser agotadora, si el número de pequeñas ZE es considerable. Un ejemplo ilustrativo es el de Viet Nam. Para sus encuestas, la Oficina General de Estadística de Viet Nam quería una muestra de zonas con al menos 70-75 hogares. Aproximadamente el 5% de las ZE (= 8.000 ZE) tienen menos de 70 hogares (Pettersson, 2001b). La combinación de aproximadamente 8.000 ZE con las ZE adyacentes fue un trabajo tedioso e interminable.

65. Una manera de reducir el esfuerzo que supone integrar pequeñas unidades de área en UPM de tamaño adecuado es realizar esta operación únicamente cuando se dé el caso de que la zona pequeña (UPM) se seleccione para formar parte de una muestra. Kish (1965) propuso un procedimiento para vincular pequeñas UPM con UPM próximas durante el proceso de selección, o una vez terminado éste.

66. Otra manera de reducir el trabajo que supone la integración de pequeñas unidades es agregar a la primera fase prevista una etapa de muestreo adicional. En algunos casos, en vez de utilizar las unidades de área deseadas como UPM podríamos utilizar como UPM zonas más amplias. En las UPM seleccionadas realizamos la operación de combinar pequeñas unidades (nuestras UPM inicialmente previstas) en unidades de tamaño adecuado. Esta labor se realiza únicamente dentro de las unidades de la primera etapa seleccionadas. En consecuencia, el trabajo sería mucho menor que si se utilizaran las zonas más pequeñas como unidades de la primera etapa. Esta alternativa supone una etapa adicional como complemento de la primera etapa prevista, lo que puede repercutir en la eficiencia del diseño. No obstante, si seleccionamos sólo una USM por UPM seleccionada en la segunda etapa, la muestra equivaldrá en efecto a la muestra prevista de una sola etapa de unidades de área. Esta fue la solución adoptada en el caso de Viet Nam. Se decidió utilizar como UPM unidades administrativas mayores —las comunas— en vez de las ZE. Dentro de las comunas seleccionadas, las ZE de tamaño insuficiente se vincularon con ZE adyacentes para formar unidades de tamaño aceptable. De esta manera se redujo el trabajo que suponía la vinculación de las ZE pequeñas con las ZE adyacentes. En vez de integrar 8.000 ZE, sólo hubo que agrupar aproximadamente 1.400 ZE de las 1.800 comunas seleccionadas. En la segunda fase se seleccionaron tres ZE (o grupos de ZE, en el caso de las ZE pequeñas) dentro de las comunas seleccionadas.

Unidades de área muy grandes

67. En el otro extremo puede haber casos de unidades de área que sean demasiado grandes —en lo que respecta a la población o a la extensión geográfica— para que puedan servir como UPM. En ambos casos los costos de enumeración serán mucho mayores que en las unidades de área normales (ZE o algunas otras unidades de área). Se presentarán problemas en ambos casos si algunas de las UPM muy grandes se seleccionan para la muestra maestra. Para reducir el trabajo de preparación de los marcos de listas de hogares en estas unidades grandes podemos colocarlas en estratos independientes y seleccionar estas UPM con tasas de muestreo inferiores; podríamos mantener las tasas de muestreo generales incrementando las tasas de muestreo dentro de las UPM.

68. Otra manera de resolver el problema de una gran UPM es dividirla en varios segmentos y seleccionar uno de ellos aleatoriamente. El problema es un poco más sencillo que en el caso de las UPM pequeñas, sobre todo porque no tenemos que hacer nada antes de

la selección de la muestra maestra. Sólo habrá que intervenir cuando se dé el caso de que se seleccione una UPM de gran tamaño para la muestra maestra.

69. Otro problema distinto es el de las UPM que han aumentado o disminuido notablemente desde la fecha del censo. Siempre habrá cambios en la población a medida que pasen los años, lo que hace que las mediciones del tamaño de las UPM sean cada vez menos precisas. El efecto general es el crecimiento de las varianzas. No obstante, no se introduce ningún sesgo. El problema puede ser grave cuando ocurren cambios espectaculares en algunas UPM debido, por ejemplo, al desalojo de las zonas suburbanas o a la nueva construcción en gran escala en algunos lugares. Los procedimientos para tener en cuenta estos cambios deben concebirse como parte del mantenimiento de la muestra maestra. En el manual *Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares* (Naciones Unidas, 1986) se consideran dos estrategias: la sustitución y la revisión de la muestra.

3. Estratificación de las UPM y asignación de una muestra maestra a los estratos

Estratificación

70. Las UPM de la muestra maestra se estratifican muchas veces en las principales divisiones administrativas del país (provincias, regiones, etcétera) y, dentro de esas divisiones, en partes urbanas y rurales. Otros factores comunes de estratificación son el nivel de urbanización (metrópolis, grandes ciudades, ciudades, aldeas) y las características socioeconómicas y ecológicas. En la muestra maestra de Lesotho, las UPM se estratifican en 10 regiones administrativas y en 4 zonas agroeconómicas (tierras bajas, laderas, montañas y valle del río Senqu), lo que da como resultado 23 estratos que reflejan las diferentes formas de vida en las zonas rurales.

71. Es posible definir estratos de la "periferia urbana" en las zonas rurales próximas a las grandes ciudades. En ellos se incluirán los hogares rurales que, hasta cierto punto, dependen del sector moderno. En las grandes ciudades podría llevarse a cabo una estratificación secundaria teniendo en cuenta el tipo de vivienda, el nivel de ingreso o algunas otras características socioeconómicas.

72. Una técnica común utilizada para alcanzar una estratificación más profunda dentro de los principales estratos es ordenar las UPM dentro de los estratos de acuerdo con un criterio de estratificación y seleccionar la muestra sistemáticamente (estratificación implícita). Una de las ventajas de la estratificación implícita es que no tienen que definirse los límites de los estratos.

Asignación de la muestra maestra

73. La asignación de las UPM de la muestra maestra a los estratos podría adoptar diferentes formas:

- Asignación proporcional a la población de los estratos;
- Asignación igual a los estratos;
- Asignación proporcional a la raíz cuadrada de la población de los estratos.

74. Muchas muestras maestras se asignan a los estratos de modo proporcional a la población (número de personas u hogares) de los estratos. La asignación proporcional es una estrategia acertada en muchas situaciones. No obstante, la asignación proporcional atribuye una pequeña proporción de la muestra a los estratos pequeños. Ello puede ser un problema cuando los principales estratos son regiones administrativas (por ejemplo, provincias) del país para el que se requieren estimaciones independientes y cuando los tamaños de estas

regiones presentan diferencias enormes (como suele ocurrir con frecuencia). La demanda de asignación igual de la muestra entre las diferentes provincias podría ser muy fuerte entre los altos cargos gubernamentales de las provincias (al menos en las provincias pequeñas). Cuando las provincias son de tamaño muy diverso, la asignación igual dará lugar a una variación sustancial de las fracciones de muestreo. En la muestra maestra de la República Democrática Popular Lao construida en 1997 se decidió utilizar la asignación igual entre los 19 estratos provinciales con el fin de conseguir la misma precisión para las estimaciones de las provincias. El resultado fue que las provincias más pequeñas obtuvieron fracciones de muestreo diez veces mayores que las correspondientes a las provincias más pobladas.

75. Una asignación proporcional estricta en los ámbitos urbano/rural dará lugar a pequeñas muestras urbanas en los países con poca población urbana. La muestra maestra preparada por el Instituto Nacional de Estadística de Camboya se asigna proporcionalmente a las provincias y a las zonas urbanas/rurales. La muestra de 600 UPM consta de 512 UPM rurales y 88 urbanas. En algunas encuestas la muestra urbana se ha considerado demasiado pequeña y se ha exigido un muestreo adicional de UPM urbanas. Quizá sea conveniente cierto sobremuestreo del ámbito urbano en la muestra maestra.

76. Un compromiso entre la asignación proporcional y la asignación igual es la *asignación según la raíz cuadrada,* en la que la muestra se asigna proporcionalmente a la raíz cuadrada del tamaño del estrato. Esta forma de asignación se ha utilizado para las muestras maestras en Viet Nam y en Sudáfrica. Kish (1988) ha propuesto un compromiso alternativo basado en una asignación proporcional a $n\sqrt{(W_h^2 + H^{-2})}$, donde n es el tamaño global de la muestra, W_h es el tamaño relativo del estrato h y H es el número de estratos. En los estratos muy pequeños el segundo término domina al primero, con lo que se evita que las asignaciones a los estratos pequeños sean demasiado pequeñas.

77. Otro compromiso sería contar con una muestra maestra grande, adecuada para las estimaciones de nivel provincial, y una submuestra de una muestra más amplia, que se designaría sobre todo para las estimaciones nacionales. Un ejemplo es la muestra maestra de 1996 de Filipinas, integrada por 3.416 UPM en una muestra ampliada para estimaciones de nivel provincial con una submuestra de 2.247 UPM designada como muestra maestra básica en los casos en que se necesitaban únicamente estimaciones de nivel regional.

4. Muestreo de las UPM

78. El método más frecuente consiste en seleccionar las UPM de la muestra maestra con probabilidad proporcional al tamaño (PPT). En este caso la probabilidad de seleccionar una UPM es proporcional a la población de ésta, lo que significa que una UPM grande tiene mayor probabilidad de estar incluida en la muestra.

79. El método tiene algunas ventajas prácticas cuando las UPM varían considerablemente de tamaño. En primer lugar, este método podría dar lugar a muestras autoponderadas. En segundo lugar, genera muestras de tamaño aproximadamente igual dentro de las UPM, lo que a su vez implica cargas de trabajo aproximadamente iguales para los entrevistadores, lo que constituye una situación deseable desde la perspectiva del trabajo sobre el terreno. En el capítulo II puede verse información más detallada sobre el muestreo PPT y sus ventajas y limitaciones.

80. Una muestra PPT puede seleccionarse de diferentes modos. Un método frecuente es la selección sistemática dentro de los estratos. Si las UPM se enumeran en algo parecido a un orden geográfico dentro de los estratos, el resultado sería una buena distribución geográfica de la muestra dentro de los estratos principales (véase información más detallada en el capítulo II). Las muestras maestras de Lesotho, la República Democrática Popular Lao y Viet Nam están seleccionadas con PPT sistemática y un punto de comienzo aleatorio dentro de cada estrato.

Submuestras interpenetrantes

81. Un medio alternativo es la selección de un conjunto de submuestras interpenetrantes. Submuestra interpenetrante es la formada por un conjunto de submuestras cada una de las cuales constituye por sí misma una muestra probabilística de la población objetivo.

82. La posibilidad de utilizar submuestras interpenetrantes en el submuestreo de la muestra maestra presenta ciertas ventajas. Las submuestras ofrecen flexibilidad en cuanto al tamaño de la muestra. La muestra de una encuesta determinada puede estar formada por una o varias de las submuestras. Las submuestras pueden utilizarse también para la sustitución de las muestras en la encuestas de varias rondas.

83. El uso de submuestras interpenetrantes en el diseño de la muestra maestra no es tan frecuente como el de la selección sistemática simple. Un ejemplo de muestra maestra que utiliza muestras interpenetrantes es el de la Oficina de Estadística de Nigeria (Ajayi, 2000).

5. Durabilidad de las muestras maestras

84. La calidad de la muestra maestra se deteriora a lo largo del tiempo, pero el hecho de que las medidas de tamaño utilizadas para asignar las probabilidades de selección queden desfasadas a medida que cambia la población no sería un problema si este cambio de la población fuera más o menos uniforme en todas las unidades del marco muestral maestro. No es eso lo que ocurre normalmente. El crecimiento demográfico y la migración se producen con ritmos cambiantes en las diferentes zonas: muchas veces hay un crecimiento lento, o incluso un descenso, en algunas zonas rurales, y un fuerte crecimiento en algunas áreas suburbanas de las ciudades. Cuando tiene lugar este crecimiento desigual, las medidas del tamaño utilizadas en la selección de la muestra maestra dejarán de reflejar la distribución relativa de la población incluida en la encuesta. Ello da lugar a un mayor número de errores de muestreo en las estimaciones procedentes de la muestra maestra. Asimismo, los cambios en los límites administrativos y en las clasificaciones (por ejemplo, la distinción entre zonas urbanas y rurales) pueden hacer que la estratificación quede desfasada.

85. El marco muestral maestro normalmente se revisa por completo después de cada censo de población; en general, cada diez años. Durante el período intercensal el marco debería actualizarse periódicamente. La disponibilidad de un marco muestral maestro bien mantenido y actualizado permite seleccionar periódicamente muestras maestras del todo nuevas a partir del marco muestral maestro. Se plantea entonces la siguiente pregunta: ¿Cuánto tiempo debe mantenerse una muestra maestra sin cambios significativos? La durabilidad de una muestra maestra depende, hasta cierto punto, de condiciones locales como la migración interna y el ritmo de cambio en las unidades administrativas. Por ello no es posible ofrecer una recomendación general válida para todas las situaciones. Muchas veces la eficiencia de una muestra maestra se habrá deteriorado sustancialmente después de tres o cuatro años. La decisión de utilizar la muestra maestra sin ajustes durante un período de tiempo más largo debe ser examinada más atentamente.

86. Básicamente son dos las estrategias disponibles para evitar el deterioro de la eficiencia de la muestra maestra. Una consiste en seleccionar una muestra maestra totalmente nueva a intervalos periódicos. En Lesotho, por ejemplo, la muestra maestra se sustituye cada tres años. La segunda estrategia es mantener la muestra maestra durante un período de tiempo más largo pero introducir ajustes periódicos para compensar los efectos de los cambios en el marco y en la unidades de la muestra. Estos ajustes pueden incluir la creación de estratos independientes de fuerte crecimiento y la especificación de normas para manejar los cambios en las divisiones administrativas que influyen en las unidades

de muestreo o estratos. Aunque esta estrategia de revisión se ha utilizado en la muestra maestra de Australia, parece que se emplea sólo raras veces en los países en desarrollo. Una razón es probablemente la complejidad que presenta desde el punto de vista del muestreo, ya que requiere gran esmero y conocimientos técnicos en el diseño y en la ejecución.

6. Documentación

87. Gran parte de la labor de documentación se ha realizado ya si la muestra maestra se ha seleccionado a partir de un marco muestral maestro bien documentado. No obstante, algunas veces la documentación es uno de los puntos débiles de las muestras maestras en los países en desarrollo. La información puede ser dispersa y algunas veces escasa, lo que dificulta el seguimiento de la selección de la muestra y el cálculo de las probabilidades de muestreo. Los procedimientos y las probabilidades de selección para todas las unidades de la muestra maestra en cada una de las etapas deben documentarse perfectamente. Debería haber registros en los que se indique qué unidades de la muestra maestra se han utilizado en las muestras para las diferentes encuestas. Para las unidades de muestreo debe utilizarse un sistema estándar de números de identificación.

88. La documentación de la muestra maestra debería incluir también medidas de desempeño de la misma en lo que respecta a los errores de muestreo y los efectos del diseño en las estimaciones importantes. Estas medidas son útiles para la planificación del tamaño y la asignación de las muestras en nuevas encuestas basadas en la muestra maestra. Los procedimientos para el cálculo de las varianzas correctas y los efectos del diseño pueden encontrarse ahora en muchos programas informáticos de análisis estadístico (véanse los detalles en el capítulo XXI).

89. La documentación debería incluir también materiales auxiliares para la muestra maestra. Si los marcos muestrales secundarios (MMS) se han preparado para las UUM de la muestra maestra, estos marcos deberían formar parte de la documentación. Los MMS consistirán en unidades de área, como bloques o segmentos, o en unidades de una lista dentro de las UUM de una muestra maestra.

7. Utilización de una muestra maestra para encuestas de establecimientos

90. El objetivo principal de una muestra maestra es poder contar con muestras para las encuestas de hogares en el programa de encuestas permanentes (y de las posibles encuestas especiales que encajen con el diseño de la muestra maestra). Así pues, la muestra se concebiría fundamentalmente de manera que pueda aplicarse a un conjunto básico de encuestas de hogares. En general, la muestra no será eficiente para el muestreo de otros tipos de unidades. No obstante, en algunas ocasiones quizá sea posible utilizar la muestra maestra para encuestas relacionadas con el estudio de características de las unidades económicas, como las empresas domésticas, las empresas por cuenta propia y las propiedades agrícolas en pequeña escala.

91. En la mayor parte de los países en desarrollo, una gran proporción de los establecimientos económicos en los sectores de los servicios, el comercio y la agricultura están estrechamente asociados con los hogares privados. Dichos establecimientos suelen ser muy numerosos y de pequeño tamaño y están ampliamente dispersos en toda la población. En muchos casos hay una correspondencia unívoca entre dichos establecimientos y los hogares, siendo los hogares más que los establecimientos los que pueden servir como unidades últimas de muestreo. Una muestra maestra de hogares puede utilizarse para las encuestas de ese tipo de establecimientos. Para ello muchas veces habrá que renunciar a los diseños autoponderados. Verma (2001) examina cómo mejorar la eficiencia del diseño de la muestra en las encuestas de las unidades económicas.

92. No obstante, normalmente hay algunos grandes establecimientos que no están asociados con los hogares. Suelen ser más bien pocos, pero representan una proporción importante de muchas estimaciones sobre las cifras totales (producción, número de empleados, etcétera). Asimismo, en muchos casos están irregularmente distribuidos con respecto a la población general. Como la muestra maestra de las zonas no incluirá estas grandes unidades de manera eficiente, se necesita para ellos un marco de muestreo independiente. En muchos casos este marco podría construirse a partir de los registros de los organismos gubernamentales (por ejemplo, los encargados de los impuestos o la concesión de licencias). A partir de esta lista, deberían seleccionarse para la encuesta todas las unidades muy grandes y una muestra del resto de las unidades, junto con otra muestra de establecimientos tomada de las UPM de la muestra maestra.

93. Un caso especial de encuesta de establecimientos es aquel en el que una encuesta de hogares está vinculada con una "encuesta de la comunidad". Por ejemplo, en una encuesta de salud la encuesta de los individuos/hogares puede complementarse con una encuesta de servicios de atención de salud que incluya amplias zonas alrededor de cada una de las zonas de la muestra original (por ejemplo, zonas de empadronamiento). Los datos de la encuesta complementaria pueden tener un doble objetivo: pueden vincularse con los datos del hogar y utilizarse para análisis de la calidad y accesibilidad de los servicios locales, o bien pueden emplearse para preparar estimaciones nacionales del número y tipo de servicios de salud. En el caso del primer objetivo, los hogares/individuos continúan siendo la unidad de análisis: no se plantea ninguna nueva cuestión de muestreo. El segundo objetivo puede presentar más complicaciones. Si la zona más amplia en torno a la de la muestra original se considera como una unidad de mayor tamaño (distrito, comuna, zona de supervisión del censo, etcétera) integrada por varias zonas junto con la incluida en la muestra, la situación es fácil. La muestra resultante sería el equivalente de una muestra de zonas de mayor tamaño con probabilidad de selección de la zona mayor igual a la suma de probabilidades de selección de las zonas menores contenidas dentro de la zona mayor. En cambio, si la zona de mayor tamaño se basa en la norma "a menos de *x* kilómetros de la zona de la muestra original", la determinación de las probabilidades de selección es más compleja.

E. Observaciones finales

94. El diseño y ejecución de las encuestas de hogares es una tarea importante para todas las oficinas nacionales de estadística. Muchas oficinas nacionales de estadística de los países en desarrollo realizan varias encuestas cada año. La necesidad de planificación y coordinación de las actividades de la encuesta ha estimulado los esfuerzos por integrar éstas en los programas de encuestas de hogares. La idea de un programa integrado de encuestas de hogares se está llevando a cabo ahora en muchas oficinas nacionales de estadística.

95. Una parte importante del trabajo en un programa de encuestas es el diseño de muestras para las diferentes encuestas. En este capítulo se han planteado las cuestiones fundamentales referentes al diseño y elaboración de marcos muestrales maestros y muestras maestras. Se han descrito las ventajas de un marco muestral maestro bien diseñado y se ha mantenido que toda oficina nacional de estadística que ejecute un programa de encuestas de hogares debería poseer y mantener debidamente un marco muestral maestro que puede atender las necesidades de las encuestas de hogares en el programa de encuestas y las necesidades de encuestas especiales que pudieran surgir durante el período del programa de encuestas. Además, muchas oficinas nacionales de estadística pueden dar un paso más y diseñar y utilizar una muestra maestra para la totalidad o la mayoría de las encuestas del programa de encuestas y quizá para encuestas especiales imprevistas.

96. En este capítulo se ha presentado también un panorama general de los principales pasos que deben darse al elaborar los marcos muestrales maestros y las muestras maestras

y se han presentado ejemplos de algunos países en desarrollo. La peculiaridad de la publicación no favorece tratar detalladamente todas las cuestiones importantes relacionadas con la elaboración de los marcos muestrales maestros y las muestras maestras. El lector interesado en profundizar de un modo más exhaustivo debería consultar el correspondiente manual de las Naciones Unidas (véase Naciones Unidas, 1986).

REFERENCIAS BIBLIOGRÁFICAS

Ajayi, O. O. (2000). *Survey methodology for the sample census of agriculture in Nigeria with some comparisons of experiences in other countries.* Presentación realizada en el Seminario Internacional sobre los resultados del Censo Agrícola de China, que tuvo lugar en Beijing del 19 al 22 de septiembre de 2000.

Glewwe, P. e I. Yansaneh (2001). *Recommendations for Multi-Purpose Household Surveys from 2002 to 2010.* Report of Mission to the General Statistics Office, Viet Nam.

Kish, L. (1965). *Survey Sampling.* Nueva York: John Wiley and Sons.

_____1988). Multi-purpose sample design. *Survey Methodology*, vol. 14, págs. 19-32.

Naciones Unidas (1986). *Programa para desarrollar la capacidad nacional de ejecutar encuestas por hogares: Marcos de muestreo y diseños muestrales para programas integrados de encuestas de hogares.* Versión preliminar. DP/UN/INT-84-014/5E, Nueva York.

Pettersson, H. (1994). *Master Sample Design: Report from a Mission to the National Central Statistics Office, Namibia.* Mayo 1994. International Consulting Office, Statistics Sweden.

_____ (2001a) *Sample Design for Household and Business Surveys: Report from a Mission to the Bureau of Statistics, Lesotho, 21 May-2 June 2001*. International Consulting Office, Statistics Sweden.

_____ (2001b). *Recommendations Regarding the Design of a Master Sample for the Household Surveys of GSO: Report of Mission to the General Statistics Office, Viet Nam*. International Consulting Office, Statistics Sweden.

Rosen, B. (1997). *Creation of the 1997 Lao Master Sample: Report from a Mission to the National Statistics Centre, Lao PDR*. International Consulting Office, Statistics Sweden.

Torene, R. y L. G. Torene (1987). The practical side of using master samples: the Bangladesh experience. *Bulletin of the International Statistical Institute: Proceedings of the 46th Session, Tokyo, 1987*, vol. LII-2, págs. 493-511.

Verma, V. (2001). Sample design for national surveys: surveying small-scale economic units. *Statistics in Transition*, vol. 5, No. 3 (diciembre de 2001), págs. 367-382.

Capítulo VI

Estimación de los componentes de los efectos del diseño para su uso en el diseño muestral

Graham Kalton, J. Michael y Brick Thanh Lê
Westat
Rockville, Maryland
Estados Unidos de América

RESUMEN

El efecto del diseño —el coeficiente de la varianza de una estadística con un diseño muestral complejo y la varianza de esa estadística con una muestra aleatoria simple o una muestra irrestricta del mismo tamaño— es un valioso instrumento para el diseño de muestras. No obstante, el efecto del diseño observado en una encuesta no debe adoptarse automáticamente para el diseño de otra. Un efecto del diseño representa el efecto combinado de varios componentes, como la estratificación, los conglomerados, las desiguales probabilidades de selección y los ajustes de ponderación para los casos de falta de respuesta y de no cobertura. En vez de limitarse a importar el efecto del diseño global de una encuesta anterior, debería prestarse gran atención a los diversos componentes implicados. En este capítulo se examinan los efectos del diseño debidos a componentes individuales y se describen modelos que pueden utilizarse para agrupar esos componentes en un efecto del diseño global. A partir de los componentes, el encargado del diseño de la muestra puede elaborar estimaciones de efectos de diseño globales para diseños alternativos y luego utilizar estas estimaciones para orientar la elección de un diseño eficiente para la encuesta que se esté planificando.

Términos clave: Estratificación, conglomerado, ponderación, coeficiente de correlación intraclase.

A. Introducción

1. Como puede verse en otros capítulos de esta publicación, las encuestas nacionales de hogares en países en desarrollo y en transición emplean diseños muestrales complejos, como el muestreo en varias etapas, la estratificación y unas probabilidades de selección frecuentemente desiguales. La consecuencia del uso de un diseño muestral complejo es que los errores de muestreo de las estimaciones de la encuesta no pueden calcularse utilizando las fórmulas de los textos habituales de estadística, basadas en el supuesto de que las variables observadas son variables aleatorias distribuidas de manera independiente e idéntica (*iid*). Dicho supuesto no es válido en las observaciones seleccionadas mediante diseños muestrales complejos, lo que significa que se necesita un planteamiento diferente para estimar los errores de muestreo contenidos en las estimaciones de las encuestas.

2. Las varianzas de las estimaciones de las encuestas procedentes de diseños muestrales complejos pueden estimarse mediante algún método de remuestreo como el *jackknife* o la reiteración repetida compensada, o mediante el método de linealización por series de Taylor [véanse, por ejemplo, Wolter (1985); Rust (1985); Verma (1993); Lehtonen y Pahkinen (1994); Rust y Rao (1996)]. Existen varios programas informáticos especializados para realizar los cálculos [véase el examen de muchos de ellos realizado por Lepkowski y Bowles (1996), disponible también en http://www.fas.harvard.edu/~stats/survey-soft/iass.html; y el resumen de programas informáticos de análisis de encuestas, preparado por la Survey Research Methods Section of the American Statistical Association, disponible en http://www.fas.harvard.edu/~stats/survey-soft/survey-soft html]. Cuando las varianzas se calculan por un procedimiento que tiene en cuenta el diseño muestral complejo, las estimaciones resultantes de la varianza son diferentes de las que se obtendrían de la aplicación de las fórmulas estándar para las variables *iid.* En muchos casos las varianzas asociadas con un diseño complejo son mayores —muchas veces, bastante mayores— que las obtenidas a partir de fórmulas estándar.

3. Las fórmulas de varianza encontradas en los textos de estadística habituales pueden aplicarse a una sola forma de diseño muestral, a saber, el muestreo irrestricto (conocido también como muestreo aleatorio simple con sustitución). Con este diseño las unidades de la población incluidas en la encuesta se seleccionan independientemente y con probabilidad igual. Se realiza el muestreo de las unidades con sustitución, lo que significa que una unidad puede aparecer más de una vez en la muestra. Supongamos que una muestra irrestricta de tamaño n arroja valores $y_1, y_2, ..., y_n$ para la variable y. La varianza de la media de la muestra $\overline{y} = \sum y_i / n$ es

$$V_u(\overline{y}) = \sigma^2 / n \tag{1}$$

donde $\sigma^2 = \sum^N (Y_i - \overline{Y})^2 / N$ es la varianza del elemento de los valores N y de la población $(Y_1, Y_2, ...Y_N)$ y $\overline{Y} = \sum Y_i / N$. Esta varianza puede estimarse a partir de la muestra mediante la formula

$$v_u(\overline{y}) = s^2 / n \tag{2}$$

donde $s^2 = \sum^n (y_i - \overline{y})^2 / (n-1)$. Estas mismas fórmulas se encuentran en los textos clásicos de estadística.

4. Por norma general, las muestras de las encuestas se seleccionan sin sustitución, ya que las estimaciones de la encuesta son más precisas (es decir, tienen varianzas más bajas) cuando las unidades pueden incluirse en la muestra una sola vez. Con el muestreo aleatorio simple sin sustitución, conocido en general con el nombre de muestreo aleatorio simple (MAS), las unidades se seleccionan con probabilidad igual, y es igualmente probable que la muestra esté integrada por todos los conjuntos posibles de unidades distintas n de la población de unidades N. Con un MAS de tamaño n, la varianza y la estimación de la varianza para la media de la muestra $\overline{y} = \sum y_i / n$ vienen dadas por

$$V_0(\overline{y}) = (1-f) S^2 / n \tag{3}$$

y

$$v_0(\overline{y}) = (1-f) s^2 / n \tag{4}$$

donde $f = n/N$ es la fracción de muestreo, $S^2 = \sum^N (Y_i - \overline{Y})^2 / (N-1)$ y $s^2 = \sum^n (y_i - \overline{y})^2 / (n-1)$. Cuando N es grande, como suele ocurrir en la investigación de encuestas, σ^2 y S^2 son aproximadamente iguales. Así pues, la principal diferencia entre la varianza para la media de muestro irrestricto en la ecuación (1) y la del MAS en (3) es el factor $(1-f)$, conocido como corrección de población finita (cpf). En la mayor parte de las situaciones reales, la fracción de muestreo n/N es pequeña y puede tratarse como 0. Cuando

así ocurre, el término cpf en (3) y (4) es aproximadamente 1, y se puede hacer caso omiso de la distinción entre muestreo con y sin sustitución.

5. Las fórmulas de varianza presentadas más arriba no son aplicables a diseños muestrales complejos, pero sirven como útiles puntos de referencia y comparación para las varianzas de las estimaciones a partir de diseños complejos. Kish (1965) acuñó el término "efecto del diseño" para denotar el coeficiente entre la varianza de cualquier estimación, por ejemplo, z, resultante de un diseño complejo y la varianza de z que se daría con un MAS o una muestra irrestricta del mismo tamaño[1]. Téngase en cuenta que el efecto del diseño hace referencia a la estimación de una encuesta específica z, y es diferente en las distintas estimaciones de una encuesta dada. Téngase también en cuenta que z puede ser una estimación digna de interés; por ejemplo, una media, una proporción, un total o un coeficiente de regresión.

6. El efecto del diseño depende tanto de la forma del diseño complejo de la muestra empleado como de la estimación de la encuesta considerada. A fin de tener en cuenta ambas características, empleamos la notación $D^2(z)$ para el efecto del diseño de la estimación z, donde

$$D^2(z) = \frac{\text{Varianza de } z \text{ con diseño complejo}}{\text{Varianza de } z \text{ con una muestra irrestricta del mismo tamaño}} = \frac{V_c(z)}{V_u(z)} \qquad (5)$$

El término cuadrático de esta notación se emplea para hacer posible el uso de $D(z)$ como raíz cuadrada del efecto del diseño. Una notación simple para $D(z)$ es útil, ya que representa el multiplicador que debería aplicarse al error estándar de z en el diseño de una muestra irrestricta para obtener su error estándar en el marco del diseño complejo como, por ejemplo, en el cálculo de un intervalo de confianza.

7. Un concepto útil directamente relacionado con el efecto del diseño es el de "tamaño real de la muestra", o n_{eff}. El tamaño real de la muestra es el tamaño de una muestra irrestricta que produciría el mismo nivel de precisión para la estimación de la encuesta que el conseguido por el diseño complejo. Así pues, el tamaño real de la muestra viene dado por

$$n_{eff} = n / D^2(z) \qquad (6)$$

8. La anterior definición de $D^2(z)$ es para estudios teóricos, donde se conocen las verdaderas varianzas $V_c(z)$ y $V_0(z)$. En las aplicaciones prácticas estas varianzas se estiman a partir de la muestra, y $D^2(z)$ se estima luego mediante $d^2(z)$. Así,

$$d^2(z) = \frac{v_c(z)}{v_u(z)} \qquad (7)$$

donde $v_c(z)$ se estima utilizando un procedimiento adecuado para el diseño complejo, y $v_u(z)$ se estima mediante una fórmula para el muestreo irrestricto con parámetros desconocidos estimados a partir de la muestra. Así, por ejemplo, en el caso de la media de la muestra

$$v_u(z) = s^2 / n \qquad (8)$$

y, en el caso de las grandes muestras, s^2 puede estimarse mediante

$$\frac{\sum w_i (y_i - \bar{y})^2}{\sum w_i}$$

donde y_i y w_i son el valor y y la ponderación de la unidad de la muestra i, y $\bar{y} = \sum w_i y_i / \sum w_i$ es la estimación ponderada de la media de la población. En el caso de una proporción p, para una n grande

$$v_u(p) = \frac{p(1-p)}{n-1}$$

[1] Más exactamente, Kish (1982) definió *Deff* como este coeficiente con un denominador de la varianza del MAS, y *Deft*^{2} como coeficiente con un denominador de la varianza de la muestra irrestricta. La diferencia entre *Deff* y *Deft*^{2} está determinada por el hecho de que se incluya o no el término cpf (1–*f*). Como ese término tiene un efecto insignificante en la mayor parte de las encuestas nacionales de hogares, la distinción entre *Deff* y *Deft*^{2} no suele tener importancia práctica, y por lo tanto no se considerará en el resto de este capítulo. Supondremos en todo momento que el término cpf puede pasarse por alto. Véase también Kish (1995). Skinner definió un concepto diferente pero relacionado: el efecto de especificación errónea (*meff*), que según él es más adecuado para su uso en el análisis de datos de encuestas (véase, por ejemplo, Skinner, Holt y Smith (1989), cap. 2). Como en este capítulo se considera el diseño de la muestra más que el análisis, no se tendrá en cuenta dicho concepto.

o

$$v_u(p) = \frac{p(1-p)}{n}$$

donde p es la estimación ponderada de la proporción de la población.

9. Al definir los efectos del diseño y los efectos estimados del diseño conviene plantearse también otra cuestión. Muchas encuestas emplean diseños muestrales con probabilidades de selección desiguales, y en ese caso los grupos pueden estar representados de manera desproporcionada en la muestra. Por ejemplo, en una encuesta nacional de hogares, el 50% de una muestra de 2.000 hogares puede estar seleccionado a partir de zonas urbanas y el otro 50% de zonas rurales, mientras que sólo el 30% de los hogares de la población están en zonas urbanas. Consideremos el efecto del diseño para una media estimada de, por ejemplo, los hogares urbanos. El denominador de (8) es s^2/n. La pregunta es cómo se debe calcular n. Un planteamiento es utilizar el tamaño efectivo de la muestra urbana: 1.000 en este caso. La alternativa es emplear el tamaño previsto de la muestra en las zonas urbanas para un MAS de n = 2.000, que en este caso es 0,3×2000 = 600. El primero de estos planteamientos, que está condicionado a un tamaño efectivo de 1.000, es el más frecuentemente empleado, y es el que utilizaremos en este capítulo. No obstante, la opción de calcular los efectos del diseño basados en el segundo planteamiento puede ser útil en algunos programas de estimación de la varianza. Como los dos planteamientos pueden producir valores muy diferentes, es importante tener conciencia de la distinción entre ellos y seleccionar la opción adecuada.

10. El concepto de efecto del diseño ha demostrado ser un instrumento valioso en el diseño de muestras complejas. Los diseños complejos implican una combinación de varios componentes del diseño, como la estratificación, el muestreo en varias etapas y la selección con probabilidades desiguales. El análisis de los efectos del diseño para cada uno de estos componentes de manera individual permite extraer conclusiones útiles sobre sus efectos en la precisión de las estimaciones de las encuestas y, por lo tanto, puede contribuir a la elaboración de diseños muestrales eficientes. En la sección B examinamos los efectos del diseño para los componentes individuales. Al diseñar una muestra compleja es útil elaborar modelos que prevean los efectos globales del diseño resultantes de una combinación de componentes. Examinamos brevemente estos modelos en la sección C. Presentamos un ejemplo hipotético ilustrativo del uso de los efectos del diseño para el diseño de la muestra en la sección D y concluimos con algunas observaciones generales en la sección E.

B. Componentes de los efectos del diseño

11. En este apartado se consideran los efectos del diseño resultantes de los siguientes componentes del diseño muestral complejo: estratificación proporcionada y desproporcionada; conglomerado; probabilidades desiguales de selección, y ajustes de ponderación de la muestra en función de la falta de respuesta, así como ajustes de ponderación de la población en función de la falta de cobertura y para mejorar la precisión. Estos distintos componentes se examinan por separado en la presente sección; sus efectos conjuntos se analizan en la sección C. El principal estadístico considerado es una estimación de la media de la población $\bar{Y}$ (por ejemplo, el ingreso medio). Como una proporción de la población P (por ejemplo, la proporción de personas que viven en situación de pobreza) es de hecho un caso especial de una media aritmética, el tratamiento incluye también una proporción. Las proporciones son probablemente los estadísticos más ampliamente utilizados en los informes de encuestas, y por lo tanto se analizarán por separado cuando convenga. Los resultados de muchas encuestas están relacionados con subgrupos de la población total, como las mujeres de 15 a 44 años de edad, o las personas que viven en zonas rurales. Por ello se considerarán las repercusiones de la ponderación y el conglomerado en los efectos del diseño de las estimaciones del subgrupo.

1. Estratificación

12. Comenzamos considerando el efecto del diseño en la media de la muestra en una muestra estratificada en una sola etapa con muestreo aleatorio simple dentro de los estratos. La media de la muestra estratificada viene dada por

$$\bar{y}_{st} = \sum_h \frac{N_h}{N} \sum_i \frac{y_{hi}}{n_h} = \sum_h W_h \bar{y}_h$$

donde n_h es el tamaño de la muestra seleccionada a partir de las unidades N_h del estrato h, $N=\sum N_h$ es el tamaño de la población, $W_h = N_h/N$ es la proporción de la población en el estrato h; y_{hi} es el valor de la unidad de la muestra i en el estrato h, y $\bar{y}_h = \sum_i y_{hi} / n_h$ es la media de la muestra del estrato h. En la práctica, $\bar{y}_{st}$ se calcula como una estimación ponderada, donde a cada unidad de la muestra se asigna una ponderación básica que es la inversa de su probabilidad de selección (ignorando, por el momento, los ajustes de ponderación de la muestra y la población). Cada unidad del estrato h tiene una probabilidad de selección de n_h/N_h y por ende una ponderación básica de $w_{hi}=w_h=N_h/n_h$. Así pues, $\bar{y}_{st}$ puede expresarse como

$$\bar{y}_{st} = \frac{\sum_h \sum_i w_{hi} y_{hi}}{\sum_h \sum_i w_{hi}} = \frac{\sum_h \sum_i w_h y_{hi}}{\sum_h n_h w_h} \tag{9}$$

Suponiendo que se pueda hacer caso omiso de la corrección de la población finita, la varianza de la media estratificada viene dada por

$$V(\bar{y}_{st}) = \sum_h \frac{W_h^2 S_h^2}{n_h} \tag{10}$$

donde $S_h^2 = \sum_i (Y_{hi} - \bar{Y}_h)^2 / (N_h - 1)$ es la varianza de la unidad de población dentro del estrato h.

13. La magnitud de $V(\bar{y}_{st})$ depende de la forma en que se distribuye la muestra entre los estratos. En los numerosos casos en que se utiliza una asignación proporcional, de manera que el tamaño de la muestra en un estrato sea proporcional al tamaño de la población de dicho estrato, las ponderaciones para todas las unidades de la muestra son las mismas. La media estratificada se reduce a la media no ponderada simple $\bar{y}_{prop} = \sum\sum y_{hi} / n$, donde $n=\sum n_h$ es el tamaño global de la muestra, y su varianza se reduce a

$$V(\bar{y}_{prop}) = \frac{\sum W_h S_h^2}{n} = \frac{S_w^2}{n} \tag{11}$$

donde S_w^2 denota la varianza media de la unidad dentro del estrato. El efecto del diseño para $\bar{y}_{prop}$ en una muestra estratificada proporcional se obtiene luego utilizando la varianza de la media para una muestra aleatoria simple de la ecuación (3), ignorando el término cpf, y con la siguiente definición del efecto del diseño de la ecuación :

$$D^2(\bar{y}_{prop}) = \frac{S_w^2}{S^2} \tag{12}$$

Dado que la varianza de la unidad dentro del estrato ya no es la varianza de la unidad global (siempre que los valores de N_h sean de tamaño considerable), el efecto del diseño para la media de una muestra proporcionada no es mayor de 1. Así pues, la estratificación proporcional no puede dar lugar a una pérdida de precisión, y en general permite aumentarla. Se consigue una mayor precisión cuando las medias de los estratos $\bar{Y}_h$ difieren: cuanto mayor sea la variación entre las medidas, mayor será la precisión.

14. En muchos casos se necesita una muestra estratificada no proporcional para que la encuesta pueda ofrecer estimaciones sobre campos concretos. Por ejemplo, un objetivo de la encuesta puede ser presentar estimaciones fiables sobre cada región de un país, cada una de las cuales puede tener una población diferente. Para conseguir ese objetivo quizá sea necesario asignar a las regiones menores tamaños de muestra que sean sustancialmente mayores de los que les corresponderían con un muestreo estratificado proporcional. Los costos de la recopilación de datos que difieren considerablemente en función de los estratos pueden ser otra razón para renunciar a la asignación proporcional. Un diseño óptimo en este caso sería el que asigne a los estratos con costos más bajos de recopilación de datos un tamaño de muestra superior al proporcional.

15. La mayor precisión conseguida con la estratificación proporcional no se aplica necesariamente con respecto a una asignación no proporcional de la muestra. Para evitar complicaciones innecesarias suponemos que las varianzas de la población dentro del estrato son constantes, en otras palabras, que $S_h^2 = S_c^2$ en todos los estratos. Este supuesto muchas veces es razonable en las encuestas nacionales de hogares cuando se utiliza una estratificación no proporcional por las razones antes mencionadas. En este supuesto, la ecuación se simplifica como sigue:

$$V(\bar{y}_{st}) = S_c^2 \sum_h \frac{W_h^2}{n_h} = \frac{S_c^2}{N} \sum_h W_h \, w_h \tag{13}$$

El efecto del diseño en este caso es

$$D^2(\bar{y}_{st}) = \frac{S_c^2}{S^2} = \frac{n}{N} \sum_h W_h \, w_h \tag{14}$$

16. Además de suponer varianzas constantes dentro del estrato como las utilizadas para llegar a la ecuación (14), muchas veces es razonable asumir que las medias del estrato son aproximadamente iguales, es decir, que $\bar{Y}_h = \bar{Y}$ en todos los estratos. Con este nuevo supuesto, $S_c^2 = S^2$ y el efecto del diseño se reduce a

$$D^2(\bar{y}_{st}) = \frac{n}{N} \sum_h W_h w_h = n \sum_h \frac{W_h^2}{n_h} \tag{15}$$

Kish (1992)[2] presenta el efecto del diseño debido a la asignación no proporcional como

$$D^2(\bar{y}_{st}) = (\sum_h W_h w_h)(\sum_h W_h / w_h) \tag{16}$$

Esta fórmula es muy útil para el diseño muestral. No obstante, no debe aplicarse acríticamente, sin tener en cuenta el fundamento de sus supuestos básicos (véase *infra*).

[2] En esta referencia se resumen muchos de los resultados en forma muy útil. Muchas de las relaciones eran bien conocidas, y se publicaron varios decenios antes. Véase, por ejemplo, Kish (1965) y Kish (1976).

17. Como ejemplo sencillo de la aplicación de la ecuación (16), pensemos en un país con dos regiones: una tendría el 80% de la población total y la otra el 20% (por lo tanto, $W_1 = 4W_2$). Supongamos que se realiza una encuesta en que se asigna el mismo tamaño a las muestras de las dos regiones ($n_1 = n_2 = 1.000$). Puede utilizarse cualquiera de las expresiones anteriores para calcular el efecto del diseño resultante de la asignación no proporcional para la media nacional estimada (suponiendo que las medidas y las varianzas unitarias sean las mismas en las dos regiones). Por ejemplo, utilizando la ecuación (16) y teniendo en cuenta que $w_1 = 4w_2$, el efecto del diseño es

$$D_w^2(\bar{y}_{st}) = (4W_2 \cdot 4w_2 + W_2 \cdot w_2)\left(\frac{4W_2}{4w_2} + \frac{W_2}{w_2}\right) = 1{,}36$$

ya que $W_2 = 0{,}2$. La asignación no proporcional utilizada para conseguir una precisión aproximadamente igual para las estimaciones de todas las regiones da lugar a una media estimada para todo el país con un tamaño real de la muestra de $n_{eff} = 2.000/1{,}36 = 1.471$.

18. En el cuadro VI.1 puede verse el efecto del diseño debido a la asignación no proporcional para algunas tasas de sobremuestreo comúnmente utilizadas cuando hay sólo dos estratos. Las cifras que se encuentran al comienzo de cada columna son los coeficientes de las ponderaciones en los dos estratos, que son equivalentes a los inversos de los coeficientes de las tasas de muestreo en los dos estratos. Las subpartidas son las proporciones de la población del primer estrato. Como el efecto del diseño es simétrico en torno a 0,5, los valores para W_1>0,5 pueden obtenerse utilizando la línea correspondiente a $(1-W_1)$. Para ilustrar el uso del cuadro, consideremos el ejemplo anterior. El valor de la fila donde $W_1 = 0{,}20$ y la columna donde el coeficiente de sobremuestreo es 4 da $D^2(\bar{y}_{st}) = 1{,}36$. En el cuadro puede observarse que los efectos del diseño aumentan a medida que crece el coeficiente de las tasas de muestreo y la proporción de la población en los estratos se acerca al 50%. Cuando las tasas de muestreo de los estratos son muy diferentes, el efecto del diseño para la media global puede ser muy grande y, por lo tanto, el tamaño real de la muestra es pequeño. La asignación no proporcional da lugar a una muestra muy ineficiente para estimar el estadístico de población global en este caso.

Cuadro VI.1
Efectos del diseño debidos al muestreo no proporcional en el caso de los dos estratos

	Coeficiente entre w_1 y w_2							
W_1	1	2	3	4	5	8	10	20
0,05	1,00	1,02	1,06	1,11	1,15	1,29	1,38	1,86
0,10	1,00	1,05	1,12	1,20	1,29	1,55	1,73	2,62
0,15	1,00	1,06	1,17	1,29	1,41	1,78	2,03	3,30
0,20	1,00	1,08	1,21	1,36	1,51	1,98	2,30	3,89
0,25	1,00	1,09	1,25	1,42	1,60	2,15	2,52	4,38
0,35	1,00	1,11	1,30	1,51	1,73	2,39	2,84	5,11
0,50	1,00	1,13	1,33	1,56	1,80	2,53	3,03	5,51

19. Muchas encuestas nacionales tratan de presentar estimaciones sobre todo el país, pero también sobre sus distintas regiones. Normalmente, las regiones tienen tamaños muy diversos. En tal caso, se plantea un conflicto al determinar una asignación muestral adecuada en las diferentes regiones, como confirman los resultados anteriores. En los supuestos de varianzas unitarias y medias iguales dentro de las regiones, la asignación óptima para las estimaciones nacionales es una asignación proporcional, mientras que para las estimaciones regionales lo es una muestra de igual tamaño en cada región. El uso de la asignación óptima con un solo fin dará lugar a una muestra inadecuada para el otro objetivo. En cambio, una asignación de compromiso puede funcionar de forma razonablemente satisfactoria para ambos fines (véase el apartado D).

20. La ecuación (16) se utiliza ampliamente en el diseño muestral para evaluar el efecto del uso de una asignación no proporcional en las estimaciones nacionales. No obstante, los usuarios deben prestar atención a los supuestos de varianzas y medias iguales dentro del estrato en que está basada. Consideremos primero la situación en que las medias son diferentes pero las varianzas no. En este caso, el efecto del diseño resultante de la estratificación no proporcional viene dado por la ecuación (14), con el factor adicional $S_c^2 = S^2$. Este factor es inferior a 1, y por lo tanto el efecto del diseño no es tan grande como el asociado con la ecuación (16). El efecto del diseño representa, no obstante, el efecto global de la estratificación y la asignación no proporcional. Para medir sólo el efecto de la asignación no proporcional, la comparación que debe hacerse es entre la muestra estratificada no proporcional y una muestra estratificada proporcional del mismo tamaño. El coeficiente entre la varianza de $\bar{y}_{st}$ para el diseño no proporcional y el de $\bar{y}_{prop}$ es, en función de las ecuaciones (11) y (13) con $S_w^2 = S_c^2$,

$$R = V(\bar{y}_{st})/V(\bar{y}_{prop}) = (\Sigma_h W_h w_h)(\Sigma_h W_h / w_h)$$

Así pues, en este caso la fórmula de la ecuación (16) puede interpretarse simplemente como el efecto de la asignación no proporcional.

21. El supuesto de las varianzas unitarias iguales dentro del estrato reviste mayor importancia. Los resultados anteriores indican que una asignación no proporcional da lugar a una pérdida de precisión en las estimaciones globales cuando las varianzas unitarias dentro del estrato son iguales, pero ello no ocurre necesariamente cuando las varianzas unitarias dentro del estrato son desiguales. De hecho, cuando las varianzas dentro del estrato son desiguales, las fracciones de muestreo óptimas que deben utilizarse son proporcionales a las desviaciones estándar en los estratos [véase, por ejemplo, Cochran (1977)]. Este tipo de asignación no proporcional se utiliza ampliamente en las encuestas de empresas. Puede dar lugar a aumentos considerables de la precisión con respecto a una asignación proporcional cuando las desviaciones estándar dentro del estrato difieren notablemente.

22. En las encuestas de hogares el supuesto de varianzas iguales, o aproximadamente iguales, dentro del estrato es muchas veces razonable. Un tipo de estimación cuyas varianzas dentro del estrato pueden ser desiguales es una proporción. Una proporción es la media de una variable que considera sólo los valores 1 y 0, que corresponden a tener o no tener la característica dada. La varianza unitaria de dicha variable es $\sigma^2 = P(1-P)$, donde P es la proporción de la población que tiene la característica. Así pues, la varianza unitaria en el estrato h con una proporción P_h que tiene la característica es $S_h^2 = P_h(1-P_h)$. Si P_h varía en los estratos, lo mismo ocurrirá con S_h^2. No obstante, la variación en S_h^2 es sólo ligera para las proporciones comprendidas entre 0,2 y 0,8, desde un máximo de 0,25 para $P_h = 0{,}5$ a un mínimo de 0,16 para $P_h = 0{,}2$ o 0,8.

23. Para ilustrar el efecto de la variabilidad en las proporciones del estrato y, por lo tanto, las varianzas del estrato, volvamos a nuestro ejemplo con los dos estratos con $W_1 = 0{,}8$, $W_2 = 0{,}2$ y $n_1 = n_2$, y consideremos dos conjuntos diferentes de valores para P_1 y P_2. Para el caso 1 supongamos que $P_1 = 0{,}5$ y $P_2 = 0{,}8$. Entonces el efecto global del diseño, calculado utilizando las ecuaciones (10) y (11), es $D^2(\bar{y}_{st}) = 1{,}35$, y el coeficiente de las varianzas para los diseños no proporcionales y proporcionales es $R = 1{,}43$. En el caso 2 supongamos que $P_1 = 0{,}8$ y $P_2 = 0{,}5$. Entonces, $D^2(\bar{y}_{st}) = 1{,}16$ y $R = 1{,}26$. Los valores obtenidos para $D^2(\bar{y}_{st})$ y R en estos dos casos pueden compararse con el efecto del diseño de 1,36 que se ha obtenido en el supuesto de varianzas iguales dentro del estrato. En ambos casos los efectos globales del diseño son inferiores a 1,36 debido al aumento de la precisión como consecuencia de la estratificación. En el caso 1 el valor de R es mayor que 1,36 porque el estrato 1, a cuya muestra se aplica la tasa más baja, tiene la mayor varianza dentro del estrato. En el caso 2 ocurre lo contrario: el estrato 2, que ha sido objeto de sobremuestreo, tiene la mayor varianza dentro del estrato. Por consiguiente, este sobremuestreo va en la dirección necesaria para conseguir una mayor precisión. De hecho, en este caso la asignación óptima sería muestrear el estrato 2 con una tasa 1,25 veces mayor que la del estrato 1. Aun cuando las proporciones del estrato difieran considerablemente en estos ejemplos y, en consecuencia, las varianzas dentro del estrato presenten también diferencias apreciables, los valores de R obtenidos —1,26 y 1,43— son razonablemente próximos a 1,36. Estos cálculos ilustran el hecho de que la medida aproximada del efecto del diseño resultante de la ponderación producida a partir de la ecuación es adecuada para la mayor parte de las actividades de planificación aun cuando las variaciones dentro del estrato difieran en cierto grado.

24. Finalmente, consideremos un ejemplo más extremo con $P_1 = 0{,}05$ y $P_2 = 0{,}5$, y con $W_1 = 0{,}8$, $W_2 = 0{,}2$ y $n_1 = n_2$. En este caso $D^2(\bar{y}_{st}) = 0{,}67$ y $R = 0{,}92$. Este ejemplo demuestra que la estratificación no proporcional puede conseguir una mayor precisión. No obstante, dados los supuestos en que está basado, la ecuación no puede producir un valor menor que 1. Así pues, la ecuación no debería aplicarse indiscriminadamente sin tener en cuenta los supuestos subyacentes.

2. Conglomeración

25. Consideremos ahora otro componente importante del efecto global del diseño en las encuestas de población más generales: a saber, el efecto del diseño debido a la conglomeración en las muestras en varias etapas. Las muestras se conglomeran para reducir los costos de recopilación de datos, ya que es antieconómico enumerar y muestrear hogares distantes esparcidos por todo un país o región. Normalmente se emplean dos o más etapas de muestreo: la primera etapa, o unidades primarias de muestreo (UPM,) son zonas geográficas claramente definidas que en general se muestrean con probabilidades proporcionales a los números estimados de hogares o personas que contienen. Dentro de las UPM seleccionadas pueden establecerse una o más etapas adicionales de muestreo de la zona y luego, en las subáreas finalmente seleccionadas, se hacen listas con las unidades de habitación y se muestrean los hogares de las listas. Para una encuesta de hogares se recogen datos para los hogares muestreados. En una encuesta de personas se compila una lista de personas de los hogares seleccionados y se selecciona a todas ellas o a una muestra de las que reúnen los requisitos para la encuesta. En estas páginas suponemos que se trata de una encuesta de hogares con sólo dos etapas de muestreo (UPM y hogares). No obstante, la inclusión de etapas múltiples es un proceso directo.

26. En la vida real, las UPM son siempre de tamaño variable (es decir, cambia el número de las unidades que contienen) y por esta razón en las muestras se aplica el procedimiento de probabilidad proporcional al tamaño estimado (PPTE). El tamaño de las muestras seleccionado a partir de las UPM seleccionadas varía también en general entre las UPM. No obstante, para facilitar la presentación, comenzamos suponiendo que la población consta de A UPM (por ejemplo, distritos de empadronamiento censal), cada una de las cuales contiene B hogares. Se selecciona una muestra aleatoria simple de a UPM y en cada una de las UPM seleccionadas se selecciona a su vez una muestra aleatoria simple de $b \leq B$ hogares (el caso especial en que $b=B$ representa una muestra conglomerada en una sola etapa). Suponemos que el factor de corrección de población finita en una sola etapa es insignificante. El diseño de la muestra para seleccionar los hogares utiliza el método de probabilidad igual de selección (mpis), de manera que la media de la población pueda estimarse mediante la media de la muestra no ponderada simple $\bar{y}_{cl} = \sum_{\alpha}^{a} \sum_{\beta}^{b} y_{\alpha\beta} / n$, donde $n = ab$ y el subíndice cl denota el conglomerado. La varianza de $\bar{y}_{cl}$ puede escribirse así:

$$V(\bar{y}_{cl}) = \frac{S^2}{n}[1+(b-1)\rho] \tag{17}$$

donde S^2 es la varianza unitaria de la población y ρ es el coeficiente de correlación intraclase que mide la homogeneidad de la variable y en las UPM. En la práctica, las unidades de una UPM suelen ser algo semejantes entre sí en relación con casi todas las variables, aunque el grado de semejanza suele ser bajo. Por ello, ρ es casi siempre de signo positivo y de pequeña magnitud.

27. El efecto del diseño en esta situación sencilla es el siguiente:

$$D^2(\bar{y}_{cl}) = 1+(b-1)\rho \tag{18}$$

Este resultado básico revela que el efecto del diseño de conglomeración de la muestra dentro de las UMP depende de dos factores: del tamaño de la submuestra dentro de las UPM seleccionadas (b) y de la correlación intraclase ρ. Como ρ es en general de signo positivo, el efecto del diseño resultante de la conglomeración es, por norma general, mayor que 1.

28. Una característica importante de la ecuación —y de otras semejantes que se presentan más adelante— es que depende de ρ, que es una medida de homogeneidad dentro de

[3] El examen de esta sección se aplica a la medida de homogeneidad dentro de los conglomerados, independientemente de que sean iguales o desiguales.

las UPM para una variable concreta[3]. El valor de ρ es próximo a cero en muchas variables (por ejemplo, edad y sexo), y pequeño pero no insignificante en otros (por ejemplo, $\rho = 0{,}03$ a 0.05), pero puede ser elevado en ciertos casos (por ejemplo, acceso a un dispensario en la aldea —la UPM—, en que todas las personas de la localidad o tienen acceso o no lo tienen). Teóricamente es posible que ρ sea de signo negativo, pero no es probable que eso ocurra en la práctica (aunque las estimaciones muestrales de ρ son muchas veces negativas). Con frecuencia, ρ está inversamente relacionado con el tamaño de la UPM debido a que los conglomerados mayores suelen ser más diversos, sobre todo cuando las UPM son zonas geográficas. Estos tipos de relaciones se explotan en el diseño óptimo de las encuestas, en donde se utilizan UPM mayores y más diversas cuando existe esa posibilidad. Las estimaciones de ρ para las variables clave son necesarias para la planificación de los diseños de muestreo. Estas estimaciones suelen estar basadas en las de encuestas anteriores para variables y UPM iguales o semejantes, y se admite la transferibilidad de los valores de ρ entre variables y UPM semejantes.

29. En condiciones reales las UPM no son de igual tamaño y no se obtienen mediante el método de muestreo aleatorio simple. En la mayor parte de los diseños de muestras nacionales de hogares las muestras estratificadas de las UPM se seleccionan utilizando el muestreo PPTE. En consecuencia, la ecuación no se aplica directamente. No obstante, sirve como modelo útil para el efecto del diseño resultante de la conglomeración para diversos diseños muestrales mpis con una modificación adecuada con respecto a la interpretación de ρ.

30. Consideremos en primer lugar una muestra PPT no estratificada de UPM en la que se conocen las medidas exactas del tamaño. En este caso la combinación de una muestra PPT de a UPM y una muestra mpis de b hogares de cada UPM incluida en la muestra da lugar a un diseño mpis global. Con este diseño la ecuación continúa siendo válida, pero ahora ρ se interpreta como un indicador sintético de la homogeneidad dentro de los conglomerados últimos creados por el diseño de la submuestra (Kalton, 1979). El valor de ρ, por ejemplo, para un diseño de la submuestra que seleccione b hogares mediante muestreo sistemático es diferente de un diseño de la submuestra que divida cada UPM de la muestra en subáreas que contengan b hogares en cada caso y seleccione una sola subárea (es probable que el valor de ρ sea mayor en este último caso). Así pues, esta ampliación está relacionada con el muestreo PPT pero también con varias formas alternativas de diseño de la submuestra.

31. Veamos ahora la estratificación de las UPM. Kalton (1979) señala que el efecto del diseño debido a la conglomeración en un diseño mpis global en el que se seleccione una muestra estratificada de a UPM y en el que se incluyan en la muestra b unidades elementales con igual probabilidad dentro de cada una de las UPM seleccionadas puede expresarse aproximadamente como sigue:

$$D^2(\bar{y}_d) = 1 + (b-1)\bar{\rho} \qquad (19)$$

donde $\bar{\rho}$ es la medida media de homogeneidad dentro del estrato, siempre que la homogeneidad dentro de cada estrato sea aproximadamente de la misma magnitud. Los beneficios de la estratificación eficaz de las UPM pueden ser considerables cuando b es de tamaño significativo, ya que la medida global de la homogeneidad se sustituye en la ecuación (19) por una medida menor de la homogeneidad dentro del estrato. En otros términos, la reducción del efecto del diseño de $(b-1)(\rho - \bar{\rho})$ resultante del muestreo estratificado de las UPM puede ser grande cuando b es considerable.

32. Hasta ahora hemos supuesto una muestra mpis en que el tamaño de cada una de las UPM seleccionadas es el mismo, b. Estas condiciones se cumplen cuando se muestrean UPM del mismo tamaño con probabilidad igual y cuando se muestrean UPM de tamaño distinto mediante muestreo PPT exacto. No obstante, en la práctica no se da ninguna de esas dos situaciones. En realidad, se muestrean UPM de tamaño desigual mediante PPTE, con medidas estimadas del tamaño que son hasta cierto punto inexactas. En

este caso la aplicación de las tasas de submuestreo en las UPM de la muestra para obtener un diseño global mpis da lugar a cierta variación en el tamaño de la submuestra. Siempre que la variación de los tamaños de la submuestra no sea considerable, la ecuación (19) puede utilizarse todavía como medida aproximativa, sustituyéndose b por el tamaño medio de la submuestra, es decir,

$$D^2(\bar{y}_d) = 1 + (\bar{b} - 1)\bar{\rho} \tag{20}$$

donde $\bar{b} = \sum b_\alpha / a$, y b_α es el número de unidades elementales de la UPM α. La ecuación (20) ha demostrado ser de gran utilidad práctica en situaciones en las que el número de unidades muestreadas de cada una de las UPM es relativamente constante.

33. Cuando la variación en los tamaños de la submuestra por UPM es considerable, la aproximación asociada con la ecuación (20) resulta inadecuada. Holt (1980) amplía la anterior aproximación para tratar las submuestras de tamaño desigual sustituyendo $\bar{b}$ en la ecuación (20) por un tamaño ponderado de la submuestra media. El efecto del diseño debido a la conglomeración con tamaños de conglomerado desiguales puede expresarse así:

$$D^2(\bar{y}_d) = 1 + (b' - 1)\bar{\rho} \tag{21}$$

donde $b' = \sum b_\alpha^2 / \sum b_\alpha$ (la cantidad b' puede considerarse como la media ponderada $b' = \sum k_\alpha b_\alpha / \sum k_\alpha$, donde $k_\alpha = b_\alpha$). Igual que antes, la aproximación supone un diseño de la muestra mpis.

34. Como ejemplo, supongamos que hay cinco UPM muestreadas con submuestras de 10, 10, 20, 20 y 40 hogares, y supongamos que $\bar{\rho} = 0{,}05$. El tamaño medio de las submuestras es $b = 20$, donde $b' = 26$. En este ejemplo el efecto del diseño debido a la conglomeración es, por lo tanto, 1,95 utilizando la aproximación (20), mientras que con la aproximación (21) es 2,25.

35. Verma, Scott y O'Muircheartaigh (1980) y Verma y Lê (1996) indican otra manera de expresar este ajuste, que es adecuada cuando los tamaños de las submuestras son muy diferentes para los distintos campos (por ejemplo, las zonas urbanas y rurales). Con dos campos, supongamos que los hogares b_1 se muestrean en cada una de las UPM muestreadas a_1 en un campo, siendo $n_1 = a_1 b_1$, y que los hogares b_2 se muestrean en las UPM muestreadas restantes a_2 en el otro campo, siendo $n_2 = a_2 b_2$. En este caso, con esta notación, $b' = (n_1 b_1 + n_2 b_2)/(n_1 + n_2)$.

36. Hasta ahora hemos examinado los efectos del diseño resultantes de la conglomeración para las estimaciones de las medias (y proporciones) en el caso de la población total. Gran parte de este tratamiento es igualmente aplicable a las estimaciones de subgrupos, siempre que se preste gran atención a los puestos subyacentes. Conviene introducir entre las UPM una triple clasificación de tipos de subgrupos, de acuerdo con sus distribuciones. En un extremo hay subgrupos que están distribuidos de manera uniforme entre las UPM, que se conocen con el nombre de "clases cruzadas". Por ejemplo, los subgrupos de edad/sexo generalmente son clases cruzadas. En el otro extremo hay subgrupos, cada uno de los cuales se concentra en un subconjunto de UPM que reciben el nombre de "clases segregadas". Es probable que los subgrupos urbanos y rurales sean de este tipo. En un punto intermedio se encuentran subgrupos que están algo concentrados por UPM. Se trata de las "clases mixtas".

37. Las clases cruzadas siguen la distribución de la muestra total entre las UPM. Si la muestra total está distribuida de manera bastante equitativa entre las UPM, puede utilizarse la ecuación (20) para calcular un efecto aproximado del diseño resultante de la conglomeración y esa ecuación puede utilizarse también para una clase cruzada. No obstante,

en este caso se produce un cambio importante: $\overline{b}$ representa ahora el tamaño medio de la submuestra de clases cruzadas por UPM. Como consecuencia de este cambio los efectos del diseño para las estimaciones de las clases cruzadas son menores que en las estimaciones de la muestra global.

38. Las clases segregadas constituyen todas las unidades de un subgrupo de las UPM en el total de la muestra. Como el tamaño de la muestra de las subclases para una clase segregada es igual que para el total de la muestra en ese subconjunto de UPM, en general no hay ninguna razón para prever que el efecto del diseño para una estimación de una clase segregada sea inferior al de la estimación para una muestra total. El efecto del diseño para una estimación de una clase segregada será distinto del correspondiente a una estimación de la muestra total sólo si el tamaño medio de la submuestra por UPM en la clase segregada es distinto del de la muestra total o si no coincide el grado de homogeneidad (por ejemplo, una diferencia del sintético ρ debida a diferentes diseños de la submuestra en una clase segregada y en otros lugares). Si la muestra total se distribuye uniformemente entre las UPM, la ecuación (20) puede aplicarse de nuevo, siendo b y ρ los valores para el conjunto de UPM en la clase segregada.

39. La distribución desigual de una clase mixta entre las UPM implica que la ecuación (20) no es aplicable en este caso. Para estimar el efecto del diseño resultante de la conglomeración en una estimación a partir de una clase mixta puede utilizarse la ecuación (21), siendo b_α el número de miembros de la clase mixta muestreados en la UPM α.

3. Ponderación de los ajustes

40. Como se ha observado previamente en el apartado B.1 (Estratificación), las probabilidades desiguales de selección entre estratos con estratificación no proporcional obligan a recurrir a la ponderación en el análisis de los datos de las encuestas. Las ecuaciones (15) y (16) presentan el efecto del diseño resultante de la estratificación no proporcional y las correspondientes ponderaciones desiguales en los supuestos de que las medias de los estratos y las varianzas unitarias sean siempre iguales. Consideremos ahora formas alternativas de estas fórmulas, de más fácil aplicación para determinar los efectos de las ponderaciones en la etapa del análisis. Antes de nada, observamos los factores que generan la necesidad de ponderaciones variables en el análisis de las encuestas [véase también Kish (1992)]. En primer lugar, como ya hemos señalado, en el análisis se necesitan ponderaciones variables para compensar las probabilidades desiguales de selección asociadas con la estratificación no proporcional. En términos más generales, son necesarias para compensar las probabilidades desiguales de selección, cualquiera que sea su causa. Las ponderaciones que compensan las probabilidades desiguales de selección son los inversos de las probabilidades de selección, y muchas veces se conocen con el nombre de ponderaciones básicas. Éstas se ajustan con frecuencia para compensar la falta de respuesta y hacer que los totales de la muestra ponderada se correspondan con los totales conocidos de la población. En consecuencia, las ponderaciones del análisis final suelen ser hasta cierto punto variables.

41. Incluso sin el sobremuestreo de algunos campos, los diseños de la muestra normalmente se alejan de mpis debido a problemas del marco. Por ejemplo, si los hogares se seleccionan con probabilidad igual a partir de un marco de hogares y luego un miembro del hogar se selecciona aleatoriamente en cada uno de los hogares seleccionados, los miembros del hogar se incluyen en la muestra con probabilidades desiguales y por lo tanto se necesitan ponderaciones en el análisis, como compensación. Estas ponderaciones dan lugar a un componente del efecto del diseño, como se examina a continuación. Conviene señalar de pasada que este efecto de ponderación puede evitarse incluyendo en la muestra a todos los miembros del hogar seleccionados. No obstante, este procedimiento introduce otra etapa de con-

glomeración, con un efecto agregado de conglomeración debido a la semejanza de muchas características de los miembros del hogar [véase Clark y Steel (2002), donde se describen los efectos del diseño asociados con estos métodos alternativos de selección de personas en los hogares muestreados].

42. Otro caso frecuente de diseño distinto del mpis resultante de un problema del marco es aquel en el que se utiliza un diseño muestral en dos etapas y las UPM se muestrean con probabilidades proporcionales al tamaño estimado (PPTE). Si las medidas del tamaño son razonablemente precisas, el tamaño de la muestra por UPM seleccionada en un diseño global mpis es aproximadamente el mismo para todas las UPM. No obstante, si el tamaño estimado de una UPM seleccionada representa claramente una infravaloración, el diseño mpis requiere un número de unidades de esa UPM muy superior a la media. Como muchas veces no es posible recopilar datos de encuesta para un número tan elevado, puede establecerse una muestra más pequeña, lo que daría lugar a probabilidades desiguales de selección y a la necesidad de ponderaciones compensatorias.

43. Prácticamente en todas las encuestas se dan casos de falta de respuesta. Un planteamiento común utilizado para reducir el posible sesgo por falta de respuesta supone ajustar diferencialmente las ponderaciones básicas de los informantes. El procedimiento consiste en identificar subgrupos de la muestra que tengan tasas de respuesta diferentes e "inflar" las ponderaciones de los encuestados en cada subgrupo mediante el inverso de la tasa de respuesta en dicho subgrupo (Brick y Kalton, 1996). Estos ajustes hacen que las ponderaciones varíen con respecto a las ponderaciones básicas, y muchas veces lo que ocurre es que aumenta el efecto del diseño de una estimación.

44. Cuando se dispone de información sobre la población procedente de otra fuente, las ponderaciones ajustadas en función de la falta de respuesta pueden ajustarse todavía más para conseguir que las estimaciones ponderadas de la muestra correspondan a la información sobre la población. Por ejemplo, si una fuente externa permite disponer de estimaciones válidas sobre el tamaño de la población regional, puede hacerse que las estimaciones de estas poblaciones regionales en la muestra coincidan con las estimaciones externas. Este tipo de ajuste de la ponderación de la población se realiza muchas veces mediante un tipo de ajuste de posestratificación y puede ayudar a compensar la falta de cobertura y a aumentar la precisión de algunas estimaciones de la encuesta. No obstante, representa una mayor variabilidad en las ponderaciones, lo que puede influir negativamente en la precisión de las estimaciones de la encuesta que no están relacionadas con las variables de población empleadas en el ajuste.

45. En este contexto, consideremos ahora una generalización del efecto del diseño para la estratificación no proporcional con el fin de evaluar los efectos generales de las ponderaciones variables. Kish (1992) presenta otro modo de expresar el efecto del diseño para una media estratificada que es muy útil para calcular el efecto de la estratificación no proporcional en la etapa de análisis. La ecuación siguiente es una representación diferente de las ecuaciones (15) y (16), y por lo tanto está basada en los mismos supuestos de igualdad de las medias de los estratos y de las varianzas unitarias, en particular en estas últimas. Como se calcula a partir de la muestra, el efecto del diseño se designa como $d^2(\bar{y}_{st})$ y

$$d^2(\bar{y}_{st}) = \frac{n\sum_h \sum_i w_{hi}^2}{(\sum_h \sum_i w_{hi})^2} = 1 + cv^2(w_{hi}) \tag{22}$$

donde $cv(w_{hi})$ es el coeficiente de variación de las ponderaciones,

$$cv^2(w_{hi}) = \sum\sum(w_{hi} - \bar{w})^2 / n\bar{w}^2$$

y $\bar{w} = \sum\sum w_{hi} / n$ es la media de las ponderaciones.

46. Una forma más general de esta ecuación viene dada por

$$d^2(\bar{y}_{st}) = \frac{n\sum_j w_j^2}{(\sum_j w_j)^2} = 1 + cv^2(w_j) \tag{23}$$

donde cada una de las unidades n de la muestra tiene su propia ponderación w_j (j=1, 2, ..., n). El efecto del diseño debido a la ponderación desigual que se observa en la ecuación (23) depende del supuesto de que las ponderaciones no están relacionadas con la variable de la encuesta. La ecuación puede representar una medida razonable del efecto de la ponderación diferencial para probabilidades desiguales de selección si los supuestos subyacentes continúan siendo válidos, al menos en forma aproximada [véase Spencer (2000), donde se describe el efecto aproximado del diseño para el caso en que las probabilidades de selección estén correlacionadas con la variable de la encuesta].

47. Los ajustes en función de la falta de respuesta generalmente se introducen dentro de las clases definidas por las variables auxiliares conocidas para los informantes y los no informantes. Para que pueda contribuir a reducir el sesgo por falta de respuesta, las variables medidas en la encuesta deben diferir en las distintas clases de ponderación. No obstante, la divergencia no es en general grande, sobre todo en la varianza unitaria. En consecuencia, la ecuación (23) es muy utilizada para examinar el efecto de los ajustes de ponderación en función de la falta de respuesta en la precisión de las estimaciones. Este examen puede llevarse a cabo calculando la ecuación (23) únicamente con las ponderaciones básicas o con las ponderaciones ajustadas en función de la no respuesta. Si este último cálculo da lugar a un valor mucho mayor que el primero ello significa que esos ajustes están provocando una pérdida sustancial de precisión en las estimaciones, en cuyo caso quizá convenga modificar los ajustes de ponderación contrayendo las clases de ponderación o reduciendo las ponderaciones demasiado grandes para reducir la pérdida de precisión.

48. Si bien la ecuación (23) es razonable con respecto a la mayor parte de los ajustes de ponderación de la muestra en función de la falta de respuesta, muchas veces no ofrece una aproximación válida sobre el efecto de los ajustes de ponderación de la población. En particular, cuando las ponderaciones están posestratificadas o calibradas con arreglo a totales conocidos de control procedentes de una fuente externa, el efecto del diseño para la media de y no se puede estimar adecuadamente con la ecuación (23) cuando y está fuertemente correlacionado con uno o varios de los totales de control. Por ejemplo, supongamos que las ponderaciones están posestatificadas para controlar los totales del número de personas de un país, desglosado por sexo. Consideremos el caso extremo en que los datos de la encuesta se utilizan para estimar la proporción de mujeres de la población. En caso de correlación perfecta entre la variable y la variable de control, la proporción estimada no está sujeta a error de muestreo y, por lo tanto, tiene varianza cero. En la práctica, la correlación no será perfecta, pero puede ser considerable en algunas de las variables de la encuesta. Cuando la correlación es notable, la posestratificación o calibración con arreglo a los totales conocidos de población puede mejorar notablemente la precisión de las estimaciones de la encuesta, pero esta mejora no se pondrá de manifiesto mediante el uso de la ecuación (23). Por el contrario, la ecuación (23) indicará una pérdida de precisión.

49. De lo dicho se deduce que la ecuación (23) no debe utilizarse para estimar los efectos del diseño resultantes de ajustes en la ponderación de la población para estimaciones basadas en variables que estén estrechamente relacionadas con las variables de control. No obstante, en la mayor parte de las encuestas generales de población de los países en desarrollo las variables de control fiables son pocas o inexistentes, y las relaciones entre las que pudiera haber disponibles y las variables de las encuestas no suelen ser muy sólidas. En consecuencia, el problema de sobreestimar sustancialmente los efectos del diseño resultantes de la ponderación utilizando la ecuación (23) no debería presentarse con frecuencia. En cualquier

caso, la exposición anterior puede servir como advertencia de que la ecuación (23) no debe aplicarse indiscriminadamente.

50. Concluiremos esta exposición sobre los efectos del diseño asociados a la ponderación con algunos comentarios sobre los efectos de la ponderación en las estimaciones de los subgrupos. Todos los resultados presentados en este apartado y en el apartado B.1 pueden aplicarse directamente para obtener los efectos de diseño asociados a las estimaciones de los subgrupos sencillamente restringiendo los cálculos a los miembros del subgrupo. No obstante, no conviene inferir los efectos de la ponderación para las estimaciones de los subgrupos a partir de los resultados de toda la muestra. Para que esta inferencia sea válida, la distribución de las ponderaciones en el subgrupo debe ser semejante a la de toda la muestra. Así ocurre algunas veces, pero no siempre. En particular, cuando se utiliza la estratificación no proporcional para conseguir tamaños de muestra adecuados para ciertos campos (subgrupos), los efectos del diseño para las estimaciones de toda la muestra serán superiores a 1 (en los supuestos de medidas y varianzas iguales). No obstante, los efectos del diseño resultantes de la ponderación para las estimaciones del campo pueden ser iguales a 1, ya que dentro de los campos se utilizan probabilidades de selección iguales.

C. Modelo para los efectos del diseño

51. En el apartado anterior se han presentado algunos resultados sobre los efectos del diseño asociados con la ponderación y la conglomeración, por separado, con especial insistencia en los efectos del diseño en las medias y las proporciones. En este apartado se amplían esos resultados considerando los efectos del diseño debidos a una combinación de ponderación y conglomeración y los efectos del diseño para algunos otros tipos de estimaciones.

52. Se han utilizado varios modelos para representar los efectos del diseño en el caso de estas ampliaciones. Los modelos se han empleado tanto en el diseño como en el análisis de diseños muestrales complejos (Kalton, 1977; Wolter, 1985). Históricamente, los modelos han desempeñado un papel importante en el análisis. No obstante, su uso con ese fin está probablemente en descenso. Su utilización primaria —e importante— en el futuro, en la planificación de nuevos diseños, será precisamente el tema de los párrafos que siguen.

53. En los últimos años se han producido grandes avances en la capacidad de cálculo y en los programas informáticos para estimar los errores de muestreo resultantes de diseños muestrales complejos. Antes de que se realizaran esos avances, la tarea de calcular errores de muestreo válidos para las estimaciones de muestras complejas había sido laboriosa y lenta. Por ello era frecuente calcular los errores de muestreo directamente sólo en relación con un número relativamente pequeño de estimaciones y utilizar el efecto del diseño u otros modelos para inferir los errores de muestreo de otras estimaciones. La situación ha mejorado ahora espectacularmente, por lo que ya no es un gran obstáculo el cálculo directo de errores de muestreo para estimaciones muy numerosas. Además, es de prever que en el futuro se produzcan mejoras tanto en la capacidad de cálculo como en los programas y que, por lo tanto, desaparezca prácticamente el uso de los modelos de efectos del diseño con este fin.

54. Otra razón para utilizar modelos de errores de muestreo en la etapa de análisis es ofrecer un medio para resumir brevemente los errores de muestreo en los informes de las encuestas, con lo que se eliminaría la necesidad de presentar un error de muestreo para cada una de las estimaciones. En algunos casos puede argumentarse también que las estimaciones del error de muestreo a partir de un modelo pueden ser preferibles a las estimaciones directas del error de muestreo, ya que son más precisas. Hay algunos casos en que este último argumento tiene validez (por ejemplo, al estimar el error de muestreo para una estimación de una región en la que el número de UPM de la muestra es muy pequeño). No obstante, en

general, el uso de modelos para señalar los errores de muestreo por una de esas dos razones es cuestionable. La validez de las estimaciones del modelo depende de la validez de los modelos y, cuando se han efectuado comparaciones de los errores de muestreo y basados en modelos, muchas veces las comparaciones han suscitado serias dudas acerca de la validez de los modelos [véase, por ejemplo, Bye y Gallicchio (1989)]. Asimismo, mientras que los modelos del error de muestreo pueden representar un medio conciso de resumir los errores de muestreo en los informes de la encuesta, imponen a los usuarios la carga de realizar cálculos de errores de muestreo a partir de los modelos. Nuestra conclusión general es que el efecto del diseño y otros modelos del error de muestreo desempeñarán en el futuro un papel limitado en el análisis de encuestas.

55. Por el contrario, los modelos del efecto del diseño continuarán desempeñando un papel importante en el diseño de la muestra. Éste presupone una comprensión adecuada de las consecuencias de una asignación no proporcional de la muestra y de los efectos de la conglomeración en la precisión de los diferentes tipos de estimaciones de encuestas. Evidentemente, al determinarse el tamaño de la muestra necesario para conseguir la precisión adecuada en las estimaciones de la encuesta se debe tener en cuenta el efecto resultante de un diseño dado. Asimismo, la estructura de un diseño de la muestra eficiente puede desarrollarse examinando los resultados de modelos para diferentes diseños. Téngase en cuenta que las estimaciones de parámetros desconocidos, como ρ, son condición necesaria para aplicar los modelos en la etapa del diseño. Este requisito revela la necesidad de preparar estimaciones de esos parámetros a partir de encuestas anteriores, como se ilustra en la sección siguiente.

56. Comenzamos describiendo los modelos para inferir los efectos de la conglomeración en las muestras mpis en una serie de estadísticas que van más allá de las medias y las proporciones consideradas en el apartado B.3 (Ponderación de los ajustes). Para presentar estos modelos regresamos a las medias de los subgrupos ya examinadas, haciendo la debida distinción entre las clases cruzadas, las clases segregadas y las clases mixtas. En una clase cruzada, designada como *d*, que esté uniformemente distribuida entre las UPM, el efecto del diseño para la media de una clase cruzada viene dado aproximadamente por la ecuación (20), que adopta la siguiente fórmula:

$$D^2(\bar{y}_{d:d}) = 1 + (\bar{b}_d - 1)\bar{\rho}_d \tag{24}$$

donde $\bar{b}_d$ denota el tamaño medio de la muestra de la clase cruzada por UPM y $\bar{\rho}_d$ es el índice sintético de la homogeneidad de *y* en las UPM para la clase cruzada. Un modelo muy utilizado supone que la medida de homogeneidad para la clase cruzada es la misma que para el total de la población. En otras palabras, que $\bar{\rho}_d = \bar{\rho}$. En ese caso el efecto del diseño para la media de la clase cruzada puede estimarse mediante

$$d^2(\bar{y}_{d:d}) = 1 + (\bar{b}_d - 1)\hat{\rho} \tag{25}$$

donde $\hat{\rho}$ es una estimación de $\bar{\rho}$ a partir de la muestra completa dada por

$$\hat{\rho} = \frac{d^2(\bar{y}_d) - 1}{\bar{b} - 1} \tag{26}$$

57. Una ampliación frecuente de este planteamiento consiste en calcular los $\hat{\rho}$ para un conjunto de estimaciones comparables que impliquen variables relacionadas y, siempre que los $\hat{\rho}$ sean bastante semejantes, utilizar alguna forma de promedio de las mismas para estimar $\bar{\rho}$ y, por lo tanto, también $\bar{\rho}_d$ para las estimaciones del subgrupo para todas las variables. Este planteamiento se ha aplicado con frecuencia para ofrecer modelos del efecto

del diseño con el fin de resumir los errores de muestreo en los informes de las encuestas. Es también la base de una forma de función de la varianza generalizada utilizada con este fin (Wolter, 1985, pág. 204).

58. Un caso especial de este planteamiento se produce con las estimaciones de encuestas que son proporciones del subgrupo que pertenecen a diferentes categorías de una variable categórica, como las proporciones de los diferentes subgrupos que han alcanzado diferentes niveles de educación o que se encuentran en diferentes categorías profesionales. Muchas veces se supone que los valores de $\bar{\rho}$ para las diferentes clasificaciones son semejantes, por lo que el valor de $\bar{\rho}$ debe estimarse únicamente para una clasificación, y que, una vez estimado, $\hat{\rho}$ pueda aplicarse luego a todas las demás clasificaciones. El supuesto de un $\bar{\rho}$ común es matemáticamente correcto cuando hay sólo dos categorías (por ejemplo, hogar con electricidad y hogar sin electricidad), pero no tiene que ser necesariamente así cuando hay más de dos categorías. Un ejemplo sería el de las estimaciones de la proporción de trabajadores dedicados a la agricultura y a la minería. El valor de $\bar{\rho}$ para los trabajadores agrícolas es casi con toda certeza muy inferior al de los mineros, ya que la minería se concentra probablemente en un reducido número de zonas. Por ello no se puede aplicar indiscriminadamente el supuesto de un valor común $\bar{\rho}$ para todas las clasificaciones.

59. Cuando las varianzas para las medias de las clases cruzadas resultantes de la ecuación (25) se han comparado con las obtenidas directamente, se ha observado que suelen ser subestimaciones. Esta comprobación puede deberse al hecho de que los subgrupos, aun cuando se clasifiquen como clases cruzadas, no están distribuidos de forma totalmente homogénea entre las UPM. Un recurso que se ha utilizado para resolver este problema es modificar la ecuación (25), con lo que resulta

$$d^2(\bar{y}_{d:d}) = 1 + k_d(\bar{b}_d - 1)\hat{\rho} \tag{27}$$

donde k_d>1. Kish (1995), que basó su estudio en numerosos análisis empíricos, propone valores de k_d=1,2 ó 1,3; Verma y Lê (1996) dejan que k_d varíe con el tamaño de la clase cruzada (siendo k_d siempre mayor de 1). Una posible solución alternativa sería sustituir $\bar{b}_d$ en (25) con $b'_d = \sum b_{d\alpha}^2 / \sum d_{d\alpha}$ de acuerdo con la ecuación (21).

60. Consideremos ahora brevemente los efectos del diseño para las estadísticas analíticas. La forma más sencilla y más ampliamente utilizada de estadística analítica es la diferencia entre dos medias o proporciones de subgrupo. Generalmente se ha observado que el efecto del diseño para la diferencia entre dos medias es mayor de 1 pero inferior al que se obtiene tratando como independientes las dos medias del subgrupo (Kish y Frankel, 1974; Kish, 1995). Cuando se expresa en términos de varianzas, se obtiene

$$V(\bar{y}_{u:d}) + V(\bar{y}_{u:d'}) < V(\bar{y}_{d:d} - \bar{y}_{d:d'}) < V(\bar{y}_{d:d} + \bar{y}_{d:d'}) \tag{28}$$

donde *d* y *d'* representan los dos subgrupos. La varianza de la diferencia en las medias suele ser inferior al margen superior cuando ambos subgrupos están representados en la misma UPM. Esta característica origina una covarianza entre las dos medias que es casi siempre positiva, y esa covarianza positiva reduce luego la varianza de la diferencia. Este efecto no se produce cuando los subgrupos son clases segregadas que se encuentran en diferentes conjuntos de UPM; en este caso se aplica el tramo superior. En el supuesto de que las varianzas de las unidades en los dos grupos sean iguales (en otras palabras, $S_d^2 = S_{d'}^2$), esta desigualdad se reduce a

$$1 < D^2(\bar{y}_d - \bar{y}_{d'}) < \frac{n_{d'} D^2(\bar{y}_d) + n_d D^2(\bar{y}_{d'})}{n_d + n_{d'}}$$

61. Un caso especial de diferencia entre dos proporciones se plantea cuando una y otra están basadas en la misma variable de varias categorías, como ocurre, por ejemplo, cuando se pide a los encuestados que elijan entre diversas alternativas y el analista está inte-

resado en saber si una de las alternativas goza de mayor aceptación que la otra. Kish y otros (1995) examinaron los efectos del diseño para esas diferencias y comprobaron empíricamente que $d^2(p_d - \rho_{d'}) = \left[d^2(\rho_d) + d^2(\rho_{d'}) \right]/4$ en este caso especial.

62. La comprobación citada de que los efectos del diseño resultantes de la conglomeración suelen ser menores en lo que respecta a las diferencias de las medias que a las medias globales se generaliza a otras estadísticas analíticas. En Kish y Frankel (1974) se presentan algunas de las primeras pruebas empíricas y algunas sugerencias sobre modelos para los efectos del diseño correspondiente a los coeficientes de regresión múltiple. Los efectos del diseño para los coeficientes de regresión son semejantes a los de las diferencias entre medias. Su correspondencia con las expectativas puede observarse comprobando que la pendiente de una progresión lineal simple de *y* sobre *x* puede estimarse de manera bastante eficiente mediante $b = (\bar{y}_u - \bar{y}_l)/(\bar{x}_u - \bar{x}_l)$, donde las medias de *y* y *x* se calculan para los tercios superior (u) e inferior (l) de la muestra basada en la variable *x.* En Skinner, Holt y Smith (1989) y Lehtonen y Pahkinen (1994) se describen los efectos del diseño en el análisis de regresión y en otras formas de análisis, y en Korn y Graubard (1999) los efectos de los diseños muestrales complejos en la precisión del análisis de los datos de la encuesta.

63. Concluimos este apartado con algunas observaciones acerca del difícil problema de descomponer un efecto del diseño global en componentes, debido a la ponderación y a la conglomeración. El cálculo del efecto del diseño $d^2(\bar{y}) = v_c(\bar{y})/v_u(\bar{y})$ engloba los efectos combinados de la ponderación y de la conglomeración. No obstante, al utilizar los datos de una encuesta para planificar otra en el futuro, es preciso separar los dos componentes del efecto del diseño. Por ejemplo, la encuesta futura podría utilizar el mpis mientras que en la encuesta actual quizá se haya producido el sobremuestreo de ciertos campos. Asimismo, aun cuando no hubiera cambios en las UPM ni en la estratificación, para la encuesta futura quizá se prefiera cambiar el tamaño de la submuestra por UPM. Kish (1995) examina este problema, que no tiene una solución única ni sencilla. Aquí presentamos un planteamiento que puede utilizarse únicamente cuando las ponderaciones son total o aproximadamente aleatorias. En tal caso el efecto del diseño global puede descomponerse aproximadamente en un producto de los efectos del diseño resultantes de la ponderación y la conglomeración, en cuyo caso

$$d^2(\bar{y}) = d_w^2(\bar{y}) \cdot d_{cl}^2(\bar{y}) \tag{29}$$

donde $d_w^2(\bar{y})$ es el efecto del diseño resultante de la ponderación de acuerdo con la ecuación (23), y $d_{cl}^2(\bar{y})$ es el efecto del diseño resultante de la conglomeración que se observa en las ecuaciones (20) o (21). Hay poca justificación teórica de la ecuación (29); no obstante, Gabler, Haeder y Lahiri (1999), utilizando un enfoque basado en modelos, deducen el efecto del diseño dado por la ecuación (29) como límite superior. Utilizando la ecuación (29) con la ecuación (20), $\bar{\rho}$ puede estimarse mediante

$$\hat{\bar{\rho}} = \frac{\left[d^2(\bar{y})/d_w^2(\bar{y}) \right] - 1}{\bar{b} - 1} \tag{30}$$

Como se verá más adelante, a la hora de la planificación, la estimación del parámetro $\bar{\rho}$ es más importante que la estimación del efecto del diseño resultante de la conglomeración, ya que es más transferible entre los diferentes diseños. El efecto del diseño resultante de la conglomeración en una encuesta puede aplicarse directamente al planificar otra sólo si el tamaño de la submuestra por UPM se mantiene intacto.

D. Uso de los efectos del diseño muestral

64. Los modelos para los efectos del diseño examinados anteriormente en este capítulo pueden servir como instrumentos útiles para planificar un nuevo diseño muestral. No obstante, deben respaldarse con datos empíricos, sobre todo en lo relativo al índice sintético de la homogeneidad $\bar{\rho}$. Estos datos pueden obtenerse analizando los efectos del diseño para encuestas anteriores semejantes. De ahí el valor de la acumulación de datos sobre los efectos del diseño.

65. Son muchos los datos sobre los efectos del diseño en las encuestas demográficas de fecundidad y salud que se pueden encontrar en los detallados análisis de errores muestrales realizados para los programas de encuestas WFS y DHS. El programa WFS había realizado 42 encuestas en 41 países entre 1974 y 1982. En 1984 le siguió el programa DHS, con más de 120 encuestas realizadas hasta la fecha en 66 países; las encuestas se repiten en la mayoría de los países cada tres-cinco años. Véanse en Verma y Lê (1996) los análisis sobre los errores de muestreo de DHS, y en Kish, Groves y Krotki (1976) y Verma, Scott y O'Muircheartaigh (1980) otros análisis semejantes de los errores de muestreo de WFS. Una conclusión importante de los análisis de errores de muestreo para estos programas es que las estimaciones de $\bar{\rho}$ para una estimación dada son bastante transferibles de un país a otro, siempre que los diseños de la muestra sean comparables. Así pues, al diseñar una nueva encuesta en un país pueden utilizarse los datos sobre errores de muestreo de una encuesta semejante en un país próximo, siempre que se adopten las medidas necesarias para comprobar la comparabilidad del diseño muestral.

66. El ejemplo que sigue ilustra el uso de los efectos del diseño al elaborar el diseño de la muestra para una encuesta nacional hipotética. En él suponemos que se trata de una muestra PPT en dos etapas, con los distritos de empadronamiento censal como UPM y los hogares como unidades de la segunda etapa. Suponemos que el estadístico clave es la proporción de hogares en situación de pobreza, que a los efectos de la planificación sería del 25% y semejante en todas las provincias del país. Las especificaciones iniciales son que la estimación de esta proporción debe tener un coeficiente de variación de no más del 5% para la nación y no más del 10% para cada una de las ocho provincias. Además, la muestra debería permitir obtener estimaciones precisas para una serie de estadísticas de subgrupos nacionales distribuidos de manera uniforme en las ocho provincias. Si se utilizó el muestreo aleatorio simple, el coeficiente de variación sería

$$CV = \sqrt{\frac{1-P}{nP}}$$

donde P es la proporción de hogares en situación de pobreza (25% en este caso). Esta fórmula puede utilizarse también con un diseño muestral complejo, pero sustituyendo n por el tamaño real de la muestra, $n_{eff} = n/D^2(p)$.

67. La primera cuestión que debe abordarse es cómo distribuir la muestra entre las provincias. En el cuadro VI.2 se presenta la distribución de la población entre las provincias (W_h), junto con una asignación proporcional de la muestra entre ellas, una asignación igual del tamaño de la muestra para cada provincia y una asignación de compromiso de la muestra que se sitúa entre la asignación proporcional y la igual. En este punto se utiliza un tamaño total arbitrario de la muestra de 5.000 hogares. Si es necesario, puede revisarse más tarde.

Cuadro VI.2
Distribuciones de la población y tres asignaciones alternativas de la muestra entre las provincias (A–H)

	A	B	C	D	E	F	G	H	Total
W_h	**0,33**	**0,24**	**0,20**	**0,10**	**0,05**	**0,04**	**0,02**	**0,02**	**1,00**
Asignación proporcional	1 650	1 200	1 000	500	250	200	100	100	5 000
Asignación igual del tamaño de la muestra	625	625	625	625	625	625	625	625	5 000
Asignación de compromiso de la muestra	1 147	879	767	520	438	427	411	411	5 000

68. En igualdad de condiciones, la asignación proporcional es la más adecuada para producir estimaciones nacionales y estimaciones de subgrupos cuando éstos se encuentran uniformemente distribuidos entre las provincias. Por otro lado, la asignación igual del tamaño de la muestra es la más idónea para producir estimaciones provinciales. Como puede verse en el cuadro VI.2, estas dos asignaciones difieren notablemente, como consecuencia de la gran diferencia de tamaño de las provincias en la línea W_h. La asignación proporcional arroja en las provincias pequeñas (E, F, G y H) muestras que son demasiado pequeñas para permitir el cálculo de estimaciones fiables para ellas. Por el contrario, la asignación igual del tamaño de la muestra reduce la precisión de las estimaciones nacionales. Esa pérdida de precisión puede calcularse a partir de la ecuación (15), que en este caso se simplifica a $H\sum W_h^2 = 1{,}77$, donde H es el número de provincias. Así pues, considerando los efectos de la asignación no proporcional únicamente (es decir, excluyendo los efectos de la conglomeración), el tamaño de la muestra de 5.000 para las estimaciones nacionales se reduce a un tamaño real de la muestra de 5.000/1,77= 2.825.

69. El que la considerable pérdida de precisión para las estimaciones nacionales (en particular, para los subgrupos) resultante del uso de la asignación igual sea aceptable o no depende de la importancia relativa de las estimaciones nacionales y provinciales. Muchas veces las estimaciones nacionales son lo bastante importantes como para que esta pérdida resulte inaceptable. En tal caso, una asignación de compromiso que se sitúe entre la asignación proporcional y la asignación igual puede servir para atender las necesidades de las estimaciones tanto nacionales como provinciales. La asignación de compromiso de la última fila del cuadro VI.2 se calcula de acuerdo con una asignación propuesta por Kish (1976, 1988) para los casos en que las estimaciones nacionales y provinciales son de igual importancia. Dicha asignación, dada por $n_h \propto \sqrt{W_h^2 + H^{-2}}$, aumenta los tamaños de la muestra para las provincias pequeñas de manera considerable con respecto a la asignación proporcional, pero no tanto como la asignación igual. El efecto del diseño para la ponderación desigual en esta asignación es 1,22, frente a 1,77 para la asignación igual del tamaño de la muestra. Supondremos que en la encuesta se adopta la asignación de compromiso.

70. La siguiente cuestión que debe abordarse es cómo determinar el número de UPM y el número deseado de hogares que deben seleccionarse por UPM. Como se ha analizado en el capítulo II, mediante el uso de un modelo de costo simple, el número óptimo de hogares que debe seleccionarse por UPM de la muestra viene dado por

$$b_{opt} = \sqrt{C^* \frac{(1-\rho)}{\rho}}$$

donde C^* es el coeficiente entre el costo de agregar una UPM a la muestra y el costo de agregar un hogar. El modelo de costo está excesivamente simplificado, y no convendría utilizar indiscriminadamente la fórmula para b_{opt}; no obstante, puede servir como orientación útil.

71. Supongamos que la estructura organizativa del trabajo sobre el terreno aconseja el uso del modelo de costo simple y que un análisis de la estructura de costo indica que C^* es aproximadamente 16. Además, supongamos que una encuesta anterior, en la que se utilizaron las mismas UPM, ha producido una estimación de $\rho = 0{,}05$ para una característica que está fuertemente correlacionada con la pobreza. Aplicando estos números a la fórmula anterior se obtiene $b_{opt} = 17{,}4$, que en aras de la simplicidad redondeamos a 17. Muchas veces, en la práctica, el coeficiente de costo C^* no es constante en todo el país; por ejemplo, puede ser mucho más bajo en las zonas urbanas que en las rurales. Si es así, pueden utilizarse valores diferentes en las distintas partes del país. En adelante prescindi-

remos de esta complicación. Pueden encontrarse ejemplos de esas diferencias en varios de los capítulos de esta publicación que describen los diseños muestrales nacionales.

72. Con ρ=0,05 y b=17, el efecto del diseño resultante de la conglomeración es

$$D^2(p) = 1+(b-1)\rho = 1{,}80$$

Este efecto del diseño debe tenerse en cuenta al determinar la precisión de las estimaciones provinciales. Por ejemplo, el tamaño real de la muestra de 411 hogares en la provincia H es 411/1,80 = 228. Por ello el coeficiente de variación para la proporción de hogares en situación de pobreza en la provincia H es 0,11. Si este nivel de precisión se considerara insuficiente, habría que aumentar el tamaño de la muestra de la provincia H (y también en el caso de G).

73. El efecto del diseño para las estimaciones nacionales debe combinar los efectos del diseño para la conglomeración y la asignación no proporcional entre las provincias. Por lo tanto, para la proporción nacional global de hogares en situación de pobreza, el efecto del diseño estimado puede obtenerse a partir de la ecuación (29) como 1,22×1,80 = 2,20. Por ello el tamaño efectivo de la muestra correspondiente al tamaño de una muestra real de 5.000 hogares es 2.277, y el coeficiente de variación para la estimación nacional de la proporción de hogares en situación de pobreza es 0,036. Muchas veces ocurre que el tamaño de la muestra global es más que suficiente para cumplir los requisitos de precisión para las estimaciones de la población total. Más preocupante es el nivel de precisión para los subgrupos de población. En este caso el efecto del diseño resultante de la conglomeración para las clases cruzadas uniformemente distribuidas entre las UPM es menor que para la muestra total, como se describe en el apartado C. Por ejemplo, conseguiremos una clase cruzada que comprenda un tercio de la población. En este caso, si se aplica la fórmula (27) con k_d= 1,2 y b_d= 17/3 se obtiene un efecto del diseño por conglomerados de 1,23. Si se combina este efecto con el de la asignación no proporcional entre las provincias se obtiene un efecto del diseño global para la estimación de la clase cruzada de 1,22×1,23=1,50, y un tamaño real de la muestra de 5.000/(3×1,50) = 1.111. El coeficiente estimado de variación para la estimación de la clase cruzada es, por lo tanto, 0,05.

74. Los cálculos realizados siguiendo las filas indicadas más arriba pueden utilizarse para evaluar la precisión probable de las estimaciones de las encuestas más importantes, y los tamaños de la muestra pueden modificarse para atender los requisitos deseados. En las estimaciones finales de los tamaños de la muestra debe preverse un margen para la falta de respuesta. Por ejemplo, con una tasa de respuesta bastante uniforme del 90% en el conjunto del país, los tamaños de la muestra calculados más arriba deben aumentar un 11%. Asimismo, el efecto del diseño puede aumentar algo como consecuencia de la variación adicional en las ponderaciones como consecuencia de ajustes debidos a la falta de respuesta. Al calcular las fracciones de muestreo que deberán utilizarse para generar los tamaños deseados de la muestra debe preverse también un margen para la falta de cobertura. Con una tasa de cobertura del 90%, las fracciones del muestreo deben aumentar un 11%.

E. Conclusiones

75. La comprensión de los efectos del diseño y de sus componentes es de gran utilidad para elaborar los diseños muestrales de nuevas encuestas. Por ejemplo:

- Las magnitudes del efecto del diseño global para las estimaciones de las encuestas principales pueden utilizarse para determinar el tamaño deseado de la muestra. El tamaño de la muestra necesario para conseguir el nivel especificado de precisión en cada una de las estimaciones principales puede calcularse en relación con

una muestra irrestricta, y este tamaño de la muestra puede luego multiplicarse por el efecto del diseño de la estimación para obtener el tamaño deseado de la muestra para dicha estimación con el diseño muestral complejo. El tamaño final de la muestra puede luego elegirse examinando el tamaño deseado de la muestra para cada una de las estimaciones (quizá, aceptándose el mayor de los tamaños de la muestra).

- Cuando debe utilizarse un diseño muestral estratificado no proporcional para obtener estimaciones de campo del nivel de precisión necesario, la pérdida resultante de precisión para las estimaciones de la muestra total y de los subgrupos que abarcan diferentes campos puede evaluarse calculando el efecto del diseño debido a las ponderaciones variables. Si se comprueba que la pérdida es demasiado grande, podría ser conveniente un cambio en los requisitos sobre el campo que dé lugar a ponderaciones menos viables.
- Si el efecto del diseño resultante de la conglomeración es muy grande para algunas estimaciones clave de la encuesta, debería considerarse la posibilidad de aumentar el número de UPM muestreadas (a) con un tamaño de submuestra menor (b).

76. Si bien las fórmulas presentadas en este capítulo son útiles en el diseño muestral, no deberían aplicarse indiscriminadamente. Como se ha observado en varios lugares, las fórmulas se deducen a partir de una serie de supuestos y simplificaciones. Los usuarios deben tener en cuenta estas características y considerar si las fórmulas les permitirán obtener aproximaciones razonables para su situación.

77. La estimación de los efectos del diseño resultantes de la conglomeración requiere estimaciones de los valores de ρ para las variables fundamentales de la encuesta. Estas estimaciones son inevitablemente imperfectas, pero quizá sean suficientes para tener estimaciones razonables. Un error que suponga el uso de un valor de ρ mayor del previsto da lugar a la especificación de un tamaño de muestra mayor; por ello, ésta es una estrategia más bien conservadora.

78. Finalmente, conviene señalar que el objetivo de utilizar estos modelos del efecto del diseño es conseguir un diseño muestral eficiente. Si los modelos no llegan a funcionar con exactitud, el resultado será cierta pérdida de eficiencia. No obstante, el uso de modelos inadecuados para obtener el diseño de la muestra no afecta a la validez de las estimaciones de la encuesta. Con un muestreo probabilístico, las estimaciones de la encuesta continúan siendo estimaciones válidas de los parámetros de la población.

REFERENCIAS BIBLIOGRÁFICAS

Brick, J. M. y G. Kalton (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, vol. 5, págs. 215-238.

Bye, B. y S. Gallicchio (1989). A note on sampling variance estimates for Social Security program participants from the Survey of Income and Program Participation. *United States Social Security Bulletin*, vol. 51, No. 10, págs. 4-21.

Clark, R. G. y D. G. Steel (2002). The effect of using household as a sampling unit. *International Statistical Review*, vol. 70, págs. 289-314.

Cochran, W. G. (1977). *Sampling Techniques,* 3a. ed. Nueva York: Wiley.

Gabler, S., S. Haeder y P. Lahiri (1999). A model based justification of Kish's formula for design effects for weighting and clustering. *Survey Methodology,* vol. 25, págs. 105-106.

Holt, D. H. (1980). Discussion of the paper by Verma, V., C. Scott and C. O'Muircheartaigh: sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society,* Series A, vol. 143, págs. 468-469.

Kalton, G. (1977). Practical methods for estimating survey sampling errors. *Bulletin of the International Statistical Institute*, vol. 47, No. 3, págs. 495-514.

_____ (1979). Ultimate cluster sampling. *Journal of the Royal Statistical Society,* Series A, vol. 142, págs. 210-222.

Kish, L. (1965). *Survey Sampling*. Nueva York: Wiley.

_____ (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society*, Series A, vol. 139, págs. 80-95.

_____ (1982). Design effect. En *Encyclopedia of Statistical Sciences*, vol. 2, S. Kotz y N. L. Johnson, comps., Nueva York: Wiley, págs. 347-348.

_____ (1988). Multi-purpose sample designs. *Survey Methodology*, vol. 14, págs. 19-32.

_____ (1992). Weighting for unequal P_i . *Journal of Official Statistics*, vol. 8, págs. 183-200.

_____ (1995). Methods for design effects. *Journal of Official Statistics*, vol. 11, págs. 55-77.

_____ y M. R. Frankel (1974). Inference from complex samples. *Journal of the Royal Statistical Society*, Series B, vol. 36, págs. 1-37.

_____ y otros (1995). Design effects for correlated (p_i–p_j). *Survey Methodology*, vol. 21, págs. 117-124.

_____ y otros (1976). Sampling Errors in Fertility Surveys. *World Fertility Survey Occasional Paper*, No. 17. La Haya: International Statistical Institute.

Korn, E. L. y B. I. Graubard (1999). *Analysis of Health Surveys*. Nueva York: Wiley.

Lehtonen, R. y E. J. Pahkinen (1994). *Practical Methods for Design and Analysis of Complex Surveys.* Edición revisada. Chichester, Reino Unido: Wiley.

Lepkowski, J. M. y J. Bowles (1996). Sampling error software for personal computers. *Survey Statistician*, vol. 35, págs. 10-17.

Naciones Unidas (1993). *National Household Survey Capability Programme: Sampling Errors in Household Surveys.* UNFPA/UN/INT-92-P80-15E. Nueva York: División de Estadística de las Naciones Unidas. Publicación preparada por Vijay Verma.

Rust, K. F. (1985). Variance estimation for complex estimators in sample surveys. *Journal of Official Statistics*, vol.1, págs. 381-397.

_____ y J. N. K. Rao (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, vol. 5, págs. 283-310.

Skinner, C. J., D. Holt y T. M. F. Smith, comps. (1989). *Analysis of Complex Survey*s. Chichester, Reino Unido: Wiley.

Spencer, B. D. (2000). An approximate design effect for unequal weighting when measurements may correlate with selection probabilities. *Survey Methodology*, vol. 26, págs. 137-138.

Verma, V. y T. Lê (1996). An analysis of sampling errors for the Demographic and Health Surveys. *International Statistical Review*, vol. 64, págs. 265-294.

Verma, V., C. Scott y C. O'Muircheartaigh (1980). Sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society,* Series A, vol. 143, págs. 431-473.

Wolter, K. M. (1985). *Introduction to Variance Estimation*. Nueva York: Springer-Verlag.

Capítulo VII
Análisis de los efectos del diseño en las encuestas de los países en desarrollo

Hans Pettersson
Statistics Sweden
Estocolmo, Suecia

Pedro Luis do Nascimento Silva
Escola Nacional de Ciências Estatísticas/
Instituto Brasileiro de Geografia e Estatística (ENCE/IBGE)
Río de Janeiro, Brasil

RESUMEN

En este capítulo se presentan los efectos del diseño de 11 encuestas de hogares de siete países y, en relación con tres encuestas que son más bien similares en cuanto al diseño, se comparan los efectos del diseño y las tasas de homogeneidad (*roh*) para las estimaciones de consumo y posesión de bienes duraderos de los hogares. Se concluye con unas consideraciones sobre la posibilidad de estimaciones de *roh* entre diferentes encuestas.

Términos clave: Efectos del diseño, eficiencia, tasas de homogeneidad, diseño de la encuesta, diseño de la muestra, conglomeración.

A. Introducción

1. Todavía no se ha extendido demasiado la práctica de calcular los efectos del diseño como resultado estándar de las encuestas de hogares en los países en desarrollo. Hay una excepción en lo que respecta a algunas encuestas estandarizadas, como las del Estudio de medición de los niveles de vida (EMNV) y las encuestas demográficas y de salud (DHS). En esos casos, se han calculado los efectos del diseño y se han comparado entre diferentes países (véanse los capítulos XXII y XXIII). Se ha llevado a cabo un análisis inicial comparativo de gran amplitud sobre 35 encuestas realizadas en el marco de la Encuesta mundial sobre la fecundidad (WFS) (Verma, Scott y O'Muircheartaigh, 1980).

2. En este capítulo se presentan los efectos del diseño de 11 encuestas correspondientes a 7 países. La selección de las encuestas ha sido subjetiva y se ha basado sobre todo en su disponibilidad. Las encuestas son de Brasil (3), Camboya (1), la República Democrática Popular Lao (1), Lesotho (1), Namibia (2), Sudáfrica (2) y Viet Nam (1). Son de diferentes tipos y abarcan diversos temas. Algunas tienen varios objetivos, otras son sobre la mano de obra, hay una sobre el nivel de vida y otra de contenido demográfico. Los efectos del diseño se han calculado en relación con algunas características, sobre todo para la planificación de la encuesta. El objetivo principal de este capítulo es ofrecer al lector una idea general de los niveles de efectos del diseño experimentados en varias encuestas.

3. En relación con tres encuestas, de diseño semejante, se ha hecho un análisis más profundo y se han comparado los efectos del diseño y las tasas de homogeneidad en relación con algunas variables referentes al consumo de los hogares y el acceso a los bienes duraderos, con miras a conocer el comportamiento de variables (aproximadamente) iguales en diferentes poblaciones y analizar las semejanzas y posibles pautas en las constataciones realizadas.

B. Las encuestas

4. Las encuestas cuyos efectos del diseño se registran en este capítulo son:

- Encuesta sobre gastos y consumo de la República Democrática Popular Lao, 1997/98 (LECS);
- Encuesta socioeconómica de Camboya, 1999 (CSES);
- Encuesta sobre ingresos y gastos de los hogares de Namibia, 1993/94 (NHIES);
- Encuesta demográfica intercensal de Namibia, 1995/96 (NIDS);
- Encuesta de hogares con diversos objetivos en Viet Nam, 1999 (VMPHS);
- Encuesta de mano de obra de Lesotho, 1997 (LFS);
- Encuesta de hogares de octubre, 1999, República de Sudáfrica (OHS);
- Encuesta de mano de obra de febrero, 2000, República de Sudáfrica;
- PNAD (Pesquisa Nacional por Amostra de Domicílios) 1999, Brasil;
- PME (Pesquisa Mensal de Emprego) de septiembre de 1999, Brasil;
- PPV (Pesquisa de Padrões de Vida) 1996/97, Brasil.

5. En el cuadro VII.1 se resumen las principales características del diseño de las 11 encuestas. Se utilizaron diseños estándar en dos etapas con probabilidad proporcional al tamaño (PPT) en todas las encuestas, excepto en la de Viet Nam, que tuvo lugar en tres etapas.

Cuadro VII.1
Características de las 11 encuestas de hogares incluidas en el estudio

Encuesta	Número de etapas	Muestra de la primera etapa: Número de UPM seleccionadas en la muestra	Tamaño de la UPM: Número medio de hogares por UPM	Tamaño del conglomerado: Número de hogares seleccionados por UPM (o USM, si hay dos etapas)	Tamaño de la muestra: Número de hogares en la encuesta	Asignación de la muestra entre estratos
Encuesta sobre gastos y consumo de la República Democrática Popular Lao, 1997/98 (LECS)	1	R: 348 U: 102	R: 51 U: 87	R: 20 U: 20	R: 6 960 U: 2 040	No proporcional
Encuesta socioeconómica de Camboya, 1999 (CSES)	1	R: 360 U: 240	R: 154 U: 243	R: 10 U: 10	R: 3 600 U: 2 400	Aproximadamente proporcional
Encuesta sobre ingresos y gastos de los hogares de Namibia, 1993/94 (NHIES)	1	R: 123 U: 96	R: 152 U: 148	R: 20 U: 20	R: 2 685 U: 1 712	Aproximadamente proporcional
Encuesta demográfica intercensal de Namibia, 1995/96 (NIDS)	1	R: 120 U: 82	R: 152 U: 148	R: 50 U: 50	R: 5 600 U: 3 900	Aproximadamente proporcional
Encuesta de hogares con diversos objetivos en Viet Nam, 1999 (VMPHS)	2	839 UPM, (2 USM seleccionadas en cada UPM)	R: 1 417 U: 2 579 USM: R: 99 U: 105	R: 15 U: 15	25 170	No proporcional
Encuesta de mano de obra de Lesotho, 1997 (LFS)	1	R: 80 U: 40	R: 370 U: 341	R: 33 (promedio) U: 25 (promedio)	R: 2 600 U: 1 000	Aproximadamente proporcional
Encuesta de mano de obra de 2000, República de Sudáfrica	1	R: 426 U: 1 148	R: min 100 * U: min 100 *	R: 10 U: 5	R: 4 059 U: 5 646	No proporcional
Encuesta de hogares de octubre de 1999, República de Sudáfrica	1	R: 1 273 U: 1 711	R: 110-120 U: 80-100	R: 10 U: 10	R: 10 923 U: 15 211	No proporcional
PNAD, 1999, Brasil	1 ó 2	7 019	250	13	93 959	No proporcional
PME, septiembre de 1999, Brasil, BRASIL	1	1 557	250	20	30 535	No proporcional
PPV, 1996-1997, Brasil	1	554	250	R: 16 U: 8	4 944	Muy desproporcionada

Nota: R = rural, U = urbana.

* Mínimo de 100.

La PNAD empleó también muestreo en tres etapas para pequeños municipios no metropolitanos, pero éstos contenían sólo aproximadamente un tercio de la población incluida en la encuesta. La mayor parte de las encuestas utilizaron como UPM zonas de empadronamiento censal (con alguna modificación de pequeñas ZE en algunos casos). En estos casos eran frecuentes UPM con un promedio de 90-150 hogares. Tres encuestas se alejaban de esa pauta. Las dos encuestas de Lesotho tenían UMP mucho mayores: eran grupos de ZE con un tamaño medio de 340-370 hogares. En el otro extremo, las UPM rurales de la República Democrática Popular Lao tenían un promedio de sólo 50 hogares.

6. Los tamaños de la muestra dentro de las UPM (tamaño de los conglomerados) eran de unos 20 hogares en varias de las encuestas. La encuesta demográfica intercensal de Namibia se distingue por una gran muestra de 50 hogares de cada UPM. En el extremo inferior estaban la encuesta PPV del Brasil, en la que se seleccionaron 8 hogares por UPM urbana, y las dos encuestas de Sudáfrica y la encuesta de Camboya, con 10 hogares seleccionados de cada UPM. La mayoría de las encuestas tenían el mismo tamaño de conglomeración en las zonas urbanas y en las rurales.

7. La mayor parte de las encuestas estaban estratificadas expresamente en zonas urbanas/rurales dentro de las divisiones administrativas (provincias, regiones). La LFS de Lesotho tenía una estratificación adicional en zonas agroecológicas, y la LECS de la República Democrática Popular Lao una nueva estratificación, en la que se especificaba si la aldea tenía acceso a la carretera o no. La PNAD y la PME del Brasil estaban estratificadas sólo de forma implícita en zonas urbanas y rurales, con selección sistemática PPT de las UPM después de la clasificación de acuerdo con la ubicación.

8. La selección sistemática fue el sistema utilizado para la selección de los hogares dentro de las unidades de área últimas en todas las encuestas, excepto en la PPV, en que los hogares se seleccionaron mediante muestreo aleatorio simple.

9. Una característica importante de muchos de los diseños muestrales es que utilizaron asignaciones muestrales no proporcionales entre las provincias con el fin de obtener estimaciones provinciales de precisión adecuada. Las ponderaciones necesarias en el análisis para compensar las asignaciones no proporcionales fueron muy variables en algunos casos. Por ejemplo, el coeficiente entre la ponderación de muestreo mayor y menor en la PPV de Brasil fue de aproximadamente 40. En el anexo al final de este artículo se da información más detallada sobre los diseños muestrales para las encuestas.

C. Efectos del diseño

10. En los cuadros VII.2 a VII.6 pueden verse los efectos del diseño $(d^2(\bar{y}))$ en una selección de estimaciones de cada encuesta (en el capítulo VI se describe cómo se calcula el efecto del diseño). Los efectos del diseño se han calculado utilizando el programa *Software for the Statistical Analysis of Correlated Data* (SUDAAN) o StATA. En algunos casos los efectos del diseño fueron facilitados por las oficinas nacionales de estadística[1].

11. La variación en los efectos del diseño es considerable, como cabría prever dadas las diferencias en el diseño de la muestra y las variables entre las encuestas y la variación debida a las condiciones específicas de la población de cada país. Algunos efectos son muy elevados. Los efectos del diseño en el intervalo 6 a 10 para las variables de los hogares no son excepcionales en los resultados presentados en los cuadros VII.2 a VII.6, y hay algunos efectos en el intervalo de 10 a 15. Téngase en cuenta que estos resultados reflejan los efectos de los diseños muestrales conglomerados estratificados y complejos y las asignaciones no proporcionales entre las provincias (en su caso). Los efectos del diseño presentados en los cuadros VII.2 a VII.6 permiten ilustrar los niveles de los efectos del diseño que se han experimentado en algunas encuestas socioeconómicas y demográficas de hogares en países en desarrollo.

[1] El profesor David Stoker, de Statistics South Africa, compiló los efectos del diseño para la encuesta de mano de obra y la encuesta de hogares de octubre, de la República de Sudáfrica. Los efectos del diseño para la encuesta de hogares con diversos objetivos de Viet Nam fueron aportados por Nguyen Phong, Director del Departamento de Estadística Social y Ambiental, Oficina General de Estadística de Viet Nam. Los efectos del diseño para la encuesta de ingresos y gastos de hogares de Namibia fueron calculados por Alwis Weerasinghe, del Centro Nacional de Estadística de Namibia. Los correspondientes a las encuestas del Brasil fueron calculados por el Dr. Pedro Silva, del IBGE. En las demás encuestas los efectos del diseño fueron calculados por el Dr. Hans Pettersson, tomando como base los datos facilitados por los institutos nacionales de estadística.

12. En el cuadro VII.2 se presentan las estimaciones de los efectos del diseño en ocho encuestas realizadas en África y Asia sudoriental, en lo que respecta al nivel nacional y al ámbito urbano y rural. La mayor parte de los efectos del diseño afectaban a las variables socioeconómicas de los hogares. Los efectos del diseño de tres de las encuestas se refieren sobre todo a las variables de la mano de obra en el plano individual. El efecto del diseño medio global en el plano nacional es 4,2. Hay una variación bastante amplia en los efectos, que van desde 1,3 hasta 8,1, pero la mayoría de ellos se encuentran en el intervalo de 2,0 a 6,0. El promedio de los efectos del diseño para los subcampos urbano y rural es de 4,1 y 4,0, respectivamente. Las diferencias en las variables y el diseño de la muestra dificultan la exploración de los resultados para determinar posibles diferencias generales entre los tipos de variables (por ejemplo, socioeconómicas/mano de obra) o campos (urbano/rural) en el cuadro. En el cuadro VII.7 se comparan algunos de los efectos del diseño.

Cuadro VII.2
Efectos estimados del diseño de ocho encuestas realizadas en África y en Asia sudoriental

		Zonas urbanas	Zonas rurales	Total nacional
Encuesta sobre gastos y consumo de la República Democrática Popular Lao, 1997/98	Consumo mensual total por hogar	3,8	7,8	5,4
	Consumo mensual de alimentos por hogar	4,4	6,8	5,8
	Proporción de hogares con acceso a vehículo automóvil	1,3	3,3	2,1
	Proporción de hogares con acceso a la televisión	3,1	6,8	5,4
	Proporción de hogares con acceso a la radio	2,7	4,8	4,5
	Proporción de hogares con acceso al vídeo	3,9	6,1	5,5
Encuesta socioeconómica de Camboya, 1999	Consumo mensual total por hogar	2,0	2,0	1,4
	Consumo mensual de alimentos por hogar	3,1	3,2	3,2
	Proporción de hogares con acceso a la televisión	2,4	2,2	2,6
Encuesta sobre ingresos y gastos de los hogares de Namibia, 1993/94	Consumo total anual de los hogares	2,9	1,9	2,5
	Ingreso total anual de los hogares	2,9	2,8	2,8
	Proporción de hogares con acceso a la televisión	6,0	4,6	4,1
	Proporción de hogares con acceso a la radio	2,7	2,1	2,4
	Proporción de hogares con acceso al teléfono	6,2	4,6	4,5
Encuesta demográfica intercensal de Namibia	Proporción de hogares con acceso a la televisión	14,7	4,1	6,6
	Proporción de hogares que usan electricidad para la iluminación	4,4	3,9	4,2
	Proporción de hogares que sufrieron el fallecimiento de un miembro del hogar durante los 12 últimos meses	2,1	4,3	2,3
Encuesta de hogares con diversos objetivos en Viet Nam, 1999	Tasa de pobreza	..	..	7,1
Encuesta de mano de obra de Lesotho, 1997	Tasa de empleo	5,6	3,1	6,6
	Proporción de la población de 10 años o más que no ha asistido a la escuela	4,6	5,9	5,5
	Proporción de agricultores de subsistencia	6,3	4,4	8,1
	Proporción de trabajadores por cuenta propia	3,0	1,4	2,4
Encuesta de hogares de octubre, 1999, República de Sudáfrica	Tasa de empleo	4,0	3,6	3,8
Encuesta de mano de obra, 2000, República de Sudáfrica	Tasa de empleo	2,5	3,4	2,8

Nota: Dos puntos suspensivos (..) indica que faltan datos.

13. En el cuadro VII.3 se presentan las estimaciones de los efectos del diseño sobre un número de estimaciones por hogares de la PNAD del Brasil.

Cuadro VII. 3
Efectos estimados del diseño para estimaciones por países y por tipo de zona en algunas estimaciones de hogares (PNAD, 1999)

Variable	Nacional	Zonas metropolitanas	Grandes municipios	Otras zonas
Proporción con suministro general de agua corriente	9,80	6,60	6,74	10,73
Proporción con agua de manantial	9,24	4,04	4,19	9,43
Proporción con alcantarillado adecuado	9,04	6,36	5,87	11,59
Proporción con suministro general de agua corriente	8,48	5,16	4,79	9,40
Proporción con al menos un baño	8,34	1,51	7,20	7,76
Proporción con tierra en propiedad	8,10	11,53	4,49	7,09
Proporción con electricidad	7,92	1,03	4,43	7,27
Proporción con material adecuado para los muros	7,43	6,17	5,01	6,84
Proporción con agua corriente en al menos una habitación	7,09	4,74	5,45	7,04
Proporción con material adecuado para el tejado	5,68	2,91	2,41	5,65
Número medio de habitaciones por hogar	5,32	6,26	4,50	5,09
Proporción con teléfono	4,80	5,59	4,44	5,91
Proporción con frigorífico	4,59	1,53	2,77	5,02
Proporción con lavadora	4,34	3,98	3,49	6,25
Proporción con televisión en color	4,31	1,77	2,76	4,88
Proporción con congelador	3,83	3,55	2,68	4,67
Proporción con filtro de agua	3,39	2,50	2,07	4,37
Proporción con radio	3,01	1,46	1,62	3,29
Proporción con televisión en blanco y negro	2,79	1,50	1,30	2,93
Alquiler medio	2,52	3,09	2,01	3,39
Proporción de hogares en propiedad	2,46	3,18	1,74	2,30
Proporción de hogares en alquiler	2,32	2,71	1,78	2,51
Número medio de habitaciones utilizadas como dormitorios	2,14	2,37	1,72	2,09

14. Los efectos del diseño varían entre 2 y 10 en las estimaciones de alcance nacional, con un valor medio de 5,5. Son más elevados en variables como la proporción de hogares con abastecimiento general de agua corriente, proporción con agua de manantial y proporción con alcantarillado adecuado. Este resultado era previsible, dado el alto grado de conglomeración que suelen presentar estas variables. Los efectos del diseño son más bajos en algunas variables "económicas", como el alquiler medio, la proporción de hogares en propiedad o en alquiler y el número medio de habitaciones utilizadas como dormitorios. Como también era previsible, los efectos del diseño son en general más bajos en las zonas metropolitanas y en los grandes municipios, donde el diseño es el muestreo conglomerado en dos etapas, que en otros lugares, donde el diseño es más conglomerado (muestreo por conglomerados en tres etapas).

15. En el cuadro VII.4 se presentan los efectos del diseño en un conjunto de variables medidas individualmente.

Cuadro VII.4
Efectos estimados del diseño en algunas características individuales en el plano nacional y en varios subcampos (PNAD, 1999)

Variable	Nacional	Zonas metropolitanas	Grandes municipios	Otras zonas
Proporción de raza = blanca	15,97	11,97	8,14	19,97
Proporción de raza=negra o de color	15,75	12,23	8,44	19,41
Proporción de trabajadores remunerados	8,44	4,45	5,81	7,49
Proporción del trabajadores por cuenta propia	7,65	3,73	5,51	6,66
Proporción con seguridad social	6,59	2,93	3,28	8,45
Proporción de analfabetos	6,33	3,67	4,37	7,10
Ingreso medio de la ocupación principal	5,54	7,16	4,45	6,38
Proporción con prestaciones para vivienda	5,23	3,80	3,00	5,54
Proporción con prestaciones de transporte	4,93	2,94	2,78	9,10
Proporción con prestaciones de salud	4,90	3,76	2,29	8,79
Proporción que trabaja (10 + años)	4,79	1,97	1,67	7,08
Proporción con prestaciones alimentarias	3,35	2,60	2,08	4,60
Proporción de niños que trabajan (5-9 años)	3,27	1,25	2,04	3,00
Proporción de empleadores	2,87	2,80	1,54	2,63
Proporción de personas escolarizadas	1,88	1,75	1,57	1,94
Proporción de prestaciones escolares	1,87	1,85	1,74	2,22

16. Los efectos del diseño para las estimaciones de alcance nacional varían desde aproximadamente 2 hasta 16, con un promedio de 6,2. Los efectos del diseño son bastante altos en las variables relativas a la raza, elevados en las relacionadas con el empleo o los ingresos y bajos en variables como la proporción de personas escolarizadas o que reciben prestaciones de educación. También en este caso los efectos del diseño son mayores en otras áreas con diseño en tres etapas. Los efectos del diseño de las variables de los hogares son en general más bajos que los relativos a los individuos, como era de prever dado que el número de personas es mayor que el de los hogares encuestados por UPM. Las variaciones sustanciales en los efectos del diseño para las diferentes variables son previsibles porque presentan diferentes grados de conglomeración. Estos efectos del diseño más bien elevados se explican también por el uso de una asignación muestral no proporcional entre estratos, lo que da lugar a ponderaciones diversas.

17. Los efectos del diseño en la PME del Brasil se recogen en el cuadro VII.5 con respecto a una selección de estimaciones publicadas cada mes. Se obtuvieron los valores correspondientes a septiembre de 1999, ya que presentan el mismo período de referencia que los de la PNAD de 1999.

18. Aunque no se señalan aquí, se calcularon los efectos del diseño para las mismas estimaciones correspondientes a otros meses de la serie y se comprobó que variaban poco de un mes a otro. La muestra de zonas de empadronamiento es fija a lo largo de todo el decenio y el tamaño de las muestras varía también poco a corto plazo. Los efectos del diseño son mayores en el caso del ingreso medio de la ocupación principal y sólo moderados en la proporción de personas analfabetas y de empleadores. No es de extrañar que estén en consonancia con los valores observados para estimaciones similares calculadas a partir de la PNAD para las zonas metropolitanas, ya que fundamentalmente se adoptó el mismo diseño muestral para la PME y la PNAD, con excepción de una muestra mayor por UPM en la PME. Los efectos del diseño son inferiores a 2,5 en las otras variables. El hecho de que los efectos del diseño para variables comparables estimadas a partir de la PME sean en general más bajos que los de la PNAD se debe a que la asignación muestral es más próxima a la proporcional en la PME que en la PNAD.

Cuadro VII.5
Efectos estimados del diseño en algunas estimaciones de la PME correspondientes a septiembre de 1999

Variable	Recife	Salvador	Belo Horizonte	Río de Janeiro	São Paulo	Pôrto Alegre	Todos
Ingreso medio de la ocupación principal	3,43	4,47	2,49	4,44	4,89	4,79	6,23
Proporción de empleadores	2,00	2,16	3,06	2,53	2,33	2,27	3,34
Proporción de analfabetos	4,23	4,43	1,86	2,69	2,11	2,13	3,24
Tasa de desempleo	1,64	2,62	1,98	2,06	1,65	1,67	2,43
Proporción con empleo registrado	1,61	1,87	1,66	1,50	1,40	1,75	2,02
Proporción de personas económicamente activas	1,59	1,99	1,78	1,61	1,31	1,40	1,96
Proporción de trabajadores remunerados	1,51	1,67	1,43	1,37	1,34	1,55	1,88
Proporción de trabajadores por cuenta propia	1,53	2,26	1,60	1,47	1,19	1,14	1,78
Proporción de personas escolarizadas	1,41	1,57	1,64	1,24	1,26	1,49	1,72

19. Los efectos del diseño en la PPV del Brasil pueden verse en el cuadro VII.6, con respecto a una pequeña selección de estimaciones obtenidas a partir de esa encuesta.

Cuadro VII.6
Efectos estimados del diseño en algunas estimaciones de la PPV

Parámetro de la población estimado	Estimación del efecto del diseño
Número de personas de más de 14 años que son analfabetas	4,17
Proporción de personas de más de 14 años que son analfabetas	3,86
Número de personas que calificaron su estado de salud como "malo"	3,37
Proporción de hogares alquilados	2,97
Número medio de personas por hogar	2,64
Número de personas entre 7 y 14 años que son analfabetas	2,64
Proporción de personas entre 7 y 14 años que son analfabetas	2,46
Número de mujeres de 12-49 años que tuvieron algún hijo que nació muerto	2,03
Número de mujeres de 12-49 años que tuvieron hijos	2,02
Número de mujeres de 12-49 años que tuvieron hijos que nacieron vivos	2,02
Coeficiente de dependencia (número de personas de 0-14 años más número de personas de 65 o más años, dividido por el número de personas de 15-64 años)	1,99
Número medio de niños nacidos por mujer de 12-49 años	1,26

20. En las estimaciones aquí consideradas los efectos del diseño varían entre 1,3 y 4,2. Los valores relativamente pequeños de estos efectos del diseño se deben al menor grado de conglomeración en la PPV, en la que se seleccionaron sólo 8 hogares por UPM y reflejan también el hecho de que se consideraron sobre todo variables incluidas en los bloques demográficos y educativos del cuestionario, más dos variables relativas a los hogares.

21. Ahora seleccionaremos, a partir de los cuadros VII.2 a VII.6, un conjunto de estimaciones que aparecen en más de una encuesta. Los efectos del diseño se presentan en el cuadro VII.7. Dichos efectos se han agrupado en tres categorías: *a*) consumo del hogar e ingreso del hogar; *b*) bienes duraderos de los hogares, y *c*) empleo y ocupación. Dentro de cada categoría hemos agrupado las estimaciones que tienen aproximadamente las mismas definiciones.

Cuadro VII.7
Comparaciones de los efectos del diseño en las diferentes encuestas

Tema/característica	Zonas urbanas	Zona rural	Nacional	Observaciones
Consumo, ingreso de los hogares (variables de los hogares)				
• Consumo mensual total (República Democrática Popular Lao: LECS)	3,8	7,7	5,4	
• Consumo mensual total (Camboya: CSES)	2,0	2,0	1,4	El tamaño del conglomerado en la CSES es la mitad del de LECS y NHIES
• Consumo total de los hogares (Namibia: NHIES)	2,9	1,9	2,5	
• Consumo mensual de alimentos (República Democrática Popular Lao: LECS)	4,4	6,8	5,8	
• Consumo mensual de alimentos (Camboya: CSES)	2,5	3,3	3,3	
Bienes duraderos de los hogares (variables de los hogares)				
• Proporción de hogares con acceso a la televisión (República Democrática Popular Lao: LECS)	3,1	6,8	5,4	
• Proporción de hogares con acceso a la televisión (Camboya: CSES)	2,4	2,2	2,6	
• Proporción de hogares con acceso a la televisión (Namibia: NHIES)	6,0	4,6	4,1	
• Proporción de hogares con acceso a la televisión (Namibia: NIDS)	14,7	4,1	6,6	El hecho de que el tamaño del conglomerado en la NIDS sea más del doble que en las otras encuestas explica los grandes efectos del diseño en las zonas urbanas (pero no el bajo efecto del diseño en las zonas rurales)
• Proporción de hogares con televisión en color (Brasil: PNAD)	..	..	4,3	
• Proporción de hogares con acceso a la radio (República Democrática Popular Lao: LECS)	2,7	4,8	4,5	
• Proporción de hogares con acceso a la radio (Camboya: CSES)	2,1	2,8	3,4	
• Proporción de hogares con acceso a la radio (Namibia: NHIES)	2,7	2,1	2,4	
• Proporción de hogares con acceso al teléfono (Namibia: NHIES)	6,2	4,6	4,5	
• Proporción de hogares con acceso a teléfono (Brasil: PNAD)	–	–	4,8	
Empleo, ocupación (variables de las personas)				
• Tasa de empleo (Sudáfrica: OHS)	4.0	3.6	3.8	La diferencia de los efectos del diseño en las zonas urbanas entre la LFS y la OHS de Sudáfrica es consecuencia del menor tamaño de los conglomerados en la zona urbana en la LFS (5 hogares, frente a 10 en la OHS)
• Tasa de empleo (Sudáfrica: LFS)	2.5	3.4	2.8	
• Tasa de empleo (Lesotho: LFS)	5.6	3.1	6.6	
• Tasa de empleo (Brasil: PNAD)	–	–	4.8	

Nota: Dos puntos suspensivos (..) indica que faltan datos. El guión (–) indica que los datos no se aplican.

22. Los efectos del diseño en las estimaciones de alcance nacional varían entre 1,4 y 6,6, con un valor mediano de 4,3. Algunos de los efectos del diseño son muy elevados. Destaca un valor de 14,7 en la proporción de hogares urbanos con acceso a la televisión en la NIDS de Namibia. La adopción del gran conglomerado de 50 hogares contribuye a este valor elevado; si el número de hogares hubiera sido 20, como en la NHIES, el efecto del diseño habría sido 6,7, en consonancia con el efecto del diseño de 6,0 en la NHIES. Se trata de un efecto todavía considerable, y no hay ninguna contribución apreciable de las ponderaciones de la variable en este caso. Los efectos del diseño en la mayoría de las estimaciones rurales de la LECS son también altos. En la NHIES, algunos de los efectos del diseño en las zonas urbanas son elevados en el caso de los bienes duraderos.

23. En todas las encuestas, con excepción de las dos de Sudáfrica y la de Camboya, hay claras diferencias entre las zonas urbanas y rurales. En las encuestas de la República Democrática Popular Lao y del Brasil (véanse los cuadros VII.2 a VII.6) los efectos del diseño en las zonas urbanas son mayores que en las rurales (la mayor parte de las encuestas tenían el mismo tamaño de conglomerado en las zonas urbanas y rurales, por lo que las diferencias no son resultado del diferente tamaño de los conglomerados).

24. Entre los efectos del diseño se incluyen los de la estratificación, la ponderación desigual, el tamaño de los conglomerados y la homogeneidad de los conglomerados (véase un examen más detenido de los efectos en el capítulo VI). Las encuestas del cuadro VII.7 pueden ser semejantes a grandes rasgos en sus diseños muestrales, pero hay claras diferencias en cuanto a la estratificación, los tamaños de los conglomerados, la asignación de la mues-

tra, etcétera. Ello dificulta la comparación de los efectos del diseño entre las encuestas, aun cuando se trate de la misma estimación. Para lograr una mejor comparabilidad, es conveniente eliminar los efectos del tamaño del conglomerado y de la ponderación.

D. Cálculo de las tasas de homogeneidad

25. El análisis puede continuar con un conjunto menor de encuestas y variables, utilizando un reducido número de estimaciones de consumo y posesión de bienes duraderos de los hogares, a partir de LECS, CSES y NHIES, tres encuestas con diseños semejantes. Todas ellas utilizaron diseños en dos etapas con ZE como unidades primarias de muestreo. Las UPM se estratificaron de forma más o menos semejante por provincias y entre zonas urbanas/rurales dentro de las provincias. Los hogares se seleccionaron mediante muestreo sistemático dentro de las ZE. No obstante, había diferencias en la asignación de la muestra con respecto a los estratos. La encuesta de la República Democrática Popular Lao tenía una asignación igual con respecto a las provincias, mientras que las otras dos encuestas tenían asignaciones casi proporcionales en relación con las provincias. El objetivo del análisis es examinar el efecto de los diseños muestrales complejos en la precisión de la estimación (aproximadamente) idéntica en diferentes poblaciones y explorar las semejanzas y posibles pautas en las tasas de homogeneidad.

26. Un primer paso es eliminar de los efectos del diseño los efectos resultantes de ponderaciones desiguales. En el cuadro VII.8 los efectos del diseño se han dividido en componentes debidos a la ponderación y a la conglomeración. Estos componentes se calculan utilizando las ecuaciones 23 y 20 del capítulo VI. La igualdad del tamaño de la muestra dentro de las provincias en la encuesta LECS representa una variación sustancial en las ponderaciones de muestreo. Por consiguiente, los efectos del diseño debidos a la ponderación son más bien elevados en las estimaciones de la LECS. La NHIES presenta cierto grado de sobremuestreo en las regiones pobladas y en las zonas urbanas, lo que da lugar a efectos del diseño debidos a la ponderación por encima de 1,0 pero considerablemente más bajos que los efectos en el caso de la LECS. La CSES presenta también sobremuestreo en las zonas urbanas.

Cuadro VII.8
Efectos globales del diseño divididos en efectos de la ponderación ($d_w^2(\bar{y})$) y de la conglomeración ($d_{cl}^2(\bar{y})$)

Tema/Característica	Zona urbana			Zona rural		
	Global $d^2(\bar{y})$	Ponderación $d_w^2(\bar{y})$	Conglomeración $d_{cl}^2(\bar{y})$	Global $d^2(\bar{y})$	Ponderación $d_w^2(\bar{y})$	Conglomeración $d_{cl}^2(\bar{y})$
Consumo e ingreso de los hogares						
• Consumo mensual total (LECS)	3,8	1,60	**2,4**	7,7	1,55	**5,0**
• Consumo mensual total (CSES)	2,0	1,11	**1,8**	2,0	1,16	**1,7**
• Consumo total doméstico de los hogares (NHIES)	2,9	1,20	**2,4**	1,9	1,23	**1,5**
• Consumo mensual de alimentos (LECS)	4,4	1,60	**2,8**	6,8	1,55	**4,4**
• Consumo mensual de alimentos (CSES)	2,5	1,11	**2,3**	3,3	1,16	**2,8**
• Ingreso total de los hogares (NHIES)	2,9	1,20	**2,4**	2,8	1,23	**2,3**
Bienes duraderos de los hogares						
• Proporción de hogares con acceso a la televisión (LECS)	3,1	1,60	**2,0**	6,8	1,55	**4,4**
• Proporción de hogares con acceso a la televisión (CSES)	1,9	1,11	**1,7**	1,8	1,16	**1,6**
• Proporción de hogares con acceso a la televisión (NHIES)	6,0	1,20	**5,0**	4,6	1,23	**3,7**
• Proporción de hogares con acceso a la radio (LECS)	2,7	1,60	**1,7**	4,8	1,55	**3,1**
• Proporción de hogares con acceso a la radio (CSES)	2,1	1,11	**1,9**	2,3	1,16	**2,0**
• Proporción de hogares con acceso a la radio (NHIES)	2,7	1,20	**2,3**	2,1	1,23	**1,7**
• Proporción de hogares con acceso al vídeo (LECS)	3,9	1,60	**2,4**	6,1	1,55	**3,9**
• Proporción de hogares con acceso al teléfono (NHIES)	6,2	1,20	**5,2**	4,6	1,23	**3,7**

27. Las tres encuestas utilizaron un diseño en el que se seleccionó un número constante de hogares a partir de cada UPM (utilizando el muestreo sistemático). Este tamaño constante de los conglomerados contribuye también a la variación en las ponderaciones, ya que las imperfecciones en las medidas del tamaño de las UPM generarán una variación en las ponderaciones globales de la muestra.

28. Los efectos de diseño de la conglomeración, $d_{cl}^2(\bar{y})$ dependen del tamaño de la muestra conglomerada. Las encuestas de la República Democrática Popular Lao y de Namibia tenían conglomerados de 20 hogares, mientras que la encuesta de Camboya tenía 10 hogares por conglomerado. Para eliminar los efectos del diferente tipo de conglomerado al comparar los resultados de las encuestas hemos calculado las tasas de homogeneidad (ρ, *roh*) para las estimaciones del cuadro VII.8 (véase la ecuación 30 del capítulo VI). Los resultados pueden verse en el cuadro VII.9. Las *roh* miden la homogeneidad interna de las UPM (zonas de empadronamiento) en las variables de la encuesta. Lo que se debe examinar aquí es si hay semejanzas en los niveles y pautas de las *roh* entre los países.

29. Como la homogeneidad puede diferir entre los conglomerados urbanos y rurales, los valores de *roh* se han calculado por separado en esas dos partes de la población. Los resultados pueden verse en el cuadro VII.9, donde llaman la atención los resultados siguientes:

- Las pautas de las diferencias entre las zonas urbanas y rurales expresadas en *roh* son diferentes en los tres países. Las *roh* para los conglomerados urbanos de la República Democrática Popular Lao son sistemáticamente mucho más bajos que las *roh* de los conglomerados rurales. El coeficiente medio urbano/rural es 0,4. En la encuesta de Namibia las diferencias se producen en sentido contrario: las *roh* urbanas son por término medio 1,9 mayores que las *roh* rurales. En la encuesta de Camboya, no hay ninguna pauta clara entre los conglomerados urbanos y rurales en las *roh*.
- Las *roh* de los conglomerados rurales son elevadas en la LECS (en el intervalo de 0,110 a 0,209, con un valor mediano de 0,178). Las *roh* para los conglomerados urbanos son muy inferiores (en el intervalo de 0,036 a 0,092, con un valor mediano de 0,072).
- La *roh* correspondiente a consumo mensual de alimentos es alta en las zonas rurales de Camboya (0,204). Esta *roh* es bastante más elevada que la correspondiente al consumo mensual total y también superior a las *roh* correspondientes a las estimaciones de los bienes duraderos de los hogares.

30. Las grandes diferencias entre las *roh* urbanas y rurales en la República Democrática Popular Lao se deben sobre todo a las altas *roh* de las zonas rurales. Estos resultados se corresponden con los de una LECS anterior realizada en el país. Los altos valores de *roh* en las zonas rurales pueden explicarse teniendo en cuenta el hecho de que las aldeas rurales son pequeñas y más bien homogéneas en términos económicos. Asimismo, las zonas urbanas tienen muy poca segregación por nivel de ingresos, lo que hace que sean de carácter mixto en términos socioeconómicos. La estacionalidad presente en el consumo mensual total y en el consumo mensual de alimentos puede contribuir también a explicar estas variables. Cada UPM se visita durante un mes, y las muestras de las UPM se distribuyen durante 12 meses. En consecuencia, hay un "conglomerado estacional" por encima del conglomerado geográfico. Hay razones para pensar que esta estacionalidad es algo más firme en las zonas rurales.

31. En Namibia, muchas de las UPM rurales en las zonas de agricultura comercial son más bien heterogéneas, con una combinación de hogares agrícolas de ingreso alto y hogares de jornaleros de ingreso bajo. En las zonas urbanas, por el contrario, hay una fuerte segregación por nivel de ingresos que sólo se han tenido en cuenta parcialmente en la estratificación. Estas circunstancias pueden explicar las mayores *roh* para el consumo del hogar e ingreso del hogar en las zonas urbanas.

Cuadro VII.9
Tasas de homogeneidad en las zonas urbanas y rurales

Tema/Característica	Zonas urbanas	Zonas rurales	Coeficiente urbano/rural
Consumo e ingreso de los hogares			
• Consumo mensual total (LECS)	0,072	0,209	0,3
• Consumo mensual total (CSES)	0,089	0,080	1,1
• Consumo total doméstico de los hogares (NHIES)	0,071	0,025	2,9
• Consumo mensual de alimentos (LECS)	0,092	0,178	0,5
• Consumo mensual de alimentos (CSES)	0,139	0,204	0,7
• Ingreso total de los hogares (NHIES)	0,071	0,058	1,2
Bienes duraderos de los hogares			
• Acceso a la televisión (LECS)	0,049	0,178	0,3
• Acceso a la televisión (CSES)	0,079	0,061	1,3
• Acceso a la televisión (NHIES)	0,200	0,125	1,6
• Acceso a la radio (LECS)	0,036	0,110	0,3
• Acceso a la radio (CSES)	0,100	0,109	0,9
• Acceso a la radio (NHIES)	0,063	0,032	1,9
• Proporción de hogares con acceso al vídeo (LECS)	0,076	0,154	0,5
• Acceso al teléfono (NHIES)	0,208	0,125	1,7

32. A las explicaciones mencionadas hay que agregar otras dos. La primera es que los efectos del diseño (y, por consiguiente, las *roh*) para las variables del consumo son más bien sensibles a los valores extremos. La eliminación de algunos de estos valores extremos modificará en algunos casos el efecto del diseño en forma considerable. La segunda es que los valores *roh* reflejan algo más que las simples medidas de la homogeneidad de los conglomerados. Reproducen también los efectos de la varianza asociada al entrevistador, cuando diferentes entrevistadores o equipos de entrevistadores realizan las entrevistas en las distintas UPM.

E. Consideraciones

33. No es posible apreciar semejanzas entre los países en los niveles o pautas de las *roh* en el cuadro VII.9. Los resultados no ofrecen gran consuelo a un estadístico del muestreo que desee utilizar las *roh* de una encuesta semejante en otro país al diseñar la muestra para una encuesta. Parece que las condiciones de población específicas del país pueden contribuir fuertemente a determinar el grado de homogeneidad de los conglomerados en variables socioeconómicas como las aquí estudiadas. Se trata sin duda de un estudio muy limitado, y la única conclusión general que se puede extraer es la necesidad de actuar con cautela al "importar" una *roh* de una encuesta en otro país. Los resultados aconsejan también tener en cuenta la necesidad de calcular y documentar los efectos del diseño y las *roh* de la encuesta en curso para que puedan utilizarse en el diseño de la siguiente.

34. Las conclusiones del estudio, aunque inciertas, no son las habituales en los análisis de este tipo. Los estudios de las encuestas demográficas y de salud han revelado que las estimaciones de la *roh* para una estimación dada son bastante transferibles de un país a otro, siempre que los diseños muestrales sean comparables (véase el capítulo XXII). De la misma manera, el estudio realizado con una serie de encuestas muestrales de fecundidad llegó también a la conclusión de que había semejanzas en las pautas de las *roh* en los diferentes países. Puede ocurrir que las *roh* de las variables demográficas "se comporten mejor" y sean más transferibles que las *roh* para las variables socioeconómicas.

ANEXO

Descripción de los diseños muestrales para las 11 encuestas de hogares

A continuación se describen brevemente los diseños muestrales de las 11 encuestas:

Encuesta sobre gastos y consumo de la República Democrática Popular Lao, 1997/98 (LECS)

Sirvieron como UPM las zonas de empadronamiento censal (ZE). Las UPM se estratificaron en 18 provincias y zonas urbanas/zonas rurales. Las ZE rurales se estratificaron asimismo en "con acceso a la carretera" y "sin acceso a la carretera". Se seleccionaron muestras iguales de 25 UPM con PPT sistemática en cada provincia (450 UPM en total) (Rosen, 1997). Se seleccionaron 20 hogares en cada UPM, con lo que se obtiene una muestra de 9.000 hogares. La asignación igual de la muestra en las diferentes provincias dio lugar a una gran variación en las ponderaciones de muestreo en los hogares.

Encuesta socioeconómica de Camboya, 1999 (CSES)

Las aldeas sirven como UPM. Se excluyeron algunas comunas y aldeas que no pudieron visitarse por razones de seguridad; la zona excluida representaba el 3,4% del número total de hogares del país.

Las aldeas se agruparon en cinco estratos basados en zonas ecológicas. Phnom Penh se trató como un estrato aparte, y los sectores rurales y urbanos se consideraron como estratos separados. Así pues, se crearon 10 estratos a partir de las 4 zonas geográficas (Phnom Penh, Llanuras, Tonle Sap, Costa y Meseta/Montaña). De cada estrato se formaron cuatro submuestras independientes de aldeas. La muestra se asignó en forma aproximadamente proporcional a los estratos.

Se seleccionaron seiscientas aldeas con muestreo circular sistemático con PPT. En cada aldea se seleccionaron 10 hogares (Instituto Nacional de Estadística, Reino de Camboya, 1999).

Encuesta de ingresos y gastos de los hogares de Namibia, 1993/94 (NHIES)

Las UPM fueron básicamente las zonas de empadronamiento censal. Algunas ZE pequeñas se combinaron con ZE adyacentes antes de la selección. El tamaño medio de las UPM era de aproximadamente 150 hogares. Se llevó a cabo una primera estratificación de acuerdo con las divisiones en zonas urbanas/rurales y en 14 regiones y una estratificación secundaria en el ámbito urbano, en la que se definieron "estratos urbanos" y "estratos urbanos pequeños" (semiurbanos). La muestra se asignó de manera aproximadamente proporcional a los estratos. No obstante, se introdujo un ligero sobremuestreo de las zonas urbanas. Se seleccionó una muestra de 96 UPM urbanas y 123 rurales utilizando un procedimiento sistemático de PPT (Pettersson, 1994).

Encuesta demográfica intercensal de Namibia, 1995/96 (NIDS)

El diseño fue el mismo que el de la NHIES. Se seleccionó una muestra de 82 UPM urbanas y 120 rurales. En esta encuesta, en cada UPM se seleccionó una muestra más bien grande de 50 hogares, lo que dio una muestra total de 9.500 hogares (Pettersson, 1997).

Encuesta de hogares con diversos objetivos en Viet Nam, 1999 (VMPHS)

En las zonas rurales se utilizaron como UPM las comunas. En las zonas urbanas se utilizaron los barrios. La estratificación se llevó a cabo distinguiendo entre zonas urbanas/rurales y provincias (61 provincias). Se seleccionaron 839 comunas con PPT. La muestra era básicamente de igual tamaño para cada provincia, pero a las provincias grandes se asignaron muestras algo mayores. Las unidades secundarias de muestreo (USM) fueron las aldeas dentro de las comunas, y los bloques dentro de los barrios. Se seleccionaron dos USM dentro de cada comuna seleccionada. En cada USM se seleccionaron 15 hogares. En total, se seleccionaron aproximadamente 25.000 hogares (Phong, 2001).

Encuesta de mano de obra de Lesotho, 1997 (LFS)

Se utilizó una muestra en dos etapas. Las unidades primarias de muestreo fueron grupos de zonas de empadronamiento. El tamaño medio de las UPM era de 370 hogares. Las UPM se estratificaron de acuerdo con la división entre zonas urbanas/rurales, las regiones (10) y las zonas agroeconómicas (4), para obtener en total 33 estratos. La muestra se asignó en forma proporcional a los estratos, con dos excepciones: en dos pequeños estratos se aplicó un sobremuestreo muy considerable. Se utilizó un procedimiento sistemático de PPT para seleccionar 120 UPM. Dentro de la UPM se seleccionaron entre 15 y 40 hogares utilizando el muestreo aleatorio sistemático para generar una muestra total de 3.600 hogares. En la encuesta se incluyeron todos los miembros de los hogares que reunían las condiciones apropiadas (Pettersson, 2001).

Encuesta de hogares de octubre, 1999, República de Sudáfrica (OHS)

Las zonas de empadronamiento censal (ZE) sirvieron como UPM. Durante el proceso de selección, las ZE con menos de 80 hogares se combinaron con ZE próximas de la lista, utilizando un método propuesto por Kish (1965). El tamaño medio de las UPM era de 80-100 hogares en las UPM urbanas y de 110-120 hogares en las UPM rurales. Las UPM se estratificaron por provincias (nueve, en total). La muestra se asignó a los estratos con el procedimiento de la raíz cuadrada. Dentro de cada provincia se llevó a cabo una nueva estratificación en función de los consejos de distrito (y consejos metropolitanos). Se seleccionó una muestra de 2.984 UPM mediante muestreo sistemático con PPT: 1.711 en zonas urbanas y 1.273 en zonas rurales. En cada UPM se estableció una muestra sistemática de 10 "puntos de visita" (aproximadamente el mismo número que para el de los hogares) (Stoker, 2001).

Encuesta de mano de obra de febrero, 2000, República de Sudáfrica

Esta encuesta fue la primera que utilizó una nueva muestra maestra elaborada al final de 1999 sobre la base del censo de 1996. La muestra estaba integrada por 2.000 UPM (más adelante, en ese mismo año la muestra se amplió a 3.000 UPM). Las zonas de empadronamiento censal sirvieron como UPM, y las ZE con menos de 100 hogares se vincularon con ZE próximas. Las UPM se estratificaron en función de las nueve provincias. La muestra se asignó a los estratos mediante el procedimiento de la raíz cuadrada. En cada UPM se formaron conglomerados integrados por 10 puntos de visita, y cada conglomerado abarcaba una UPM completa. Se seleccionó un conjunto de conglomerados que deberían utilizarse en las futuras encuestas de mano de obra.

Como consecuencia de los problemas presupuestarios se decidió reducir la encuesta de mano de obra a 10.000 puntos de visita. Ello se llevo a cabo de la manera siguiente: de todas las UPM urbanas se seleccionaron sólo cinco puntos de visita del conglomerado identificado. En la muestra rural se elaboró una submuestra sistemática con PPT que contenía el 50% de las UPM rurales a partir del conjunto de UPM rurales, y en las UPM formadas el conglomerado identificado de 10 puntos de visita constituía parte de la muestra (Stoker, 2001).

PNAD (*Pesquisa Nacional por Amostra de Domicílios*) 1999, Brasil

La PNAD abarca anualmente una muestra de aproximadamente 115.000 hogares, que representan la totalidad del Brasil, con excepción de las zonas rurales de la región septentrional (Amazonia). La estratificación se realizó con criterios geográficos y dio lugar a 36 estratos. Éstos representaban 18 de los estados como un solo estrato en cada caso, y los otros 9 estados restantes se subdividieron en dos estratos cada uno. Luego se formó un estrato con las UPM ubicadas en la zona metropolitana de la capital del estado y otro estrato con las UPM restantes del estado. En los estratos formados por zonas metropolitanas el diseño era un muestreo mediante conglomerados en dos etapas, en el que las UPM eran zonas de empadronamiento censal —seleccionadas mediante muestreo sistemático con PPT— con tamaños iguales al número de hogares privados obtenidos en el último censo de población. Antes de la selección de las UPM se clasificaron por código geográfico, lo que dio lugar a una estratificación implícita por municipio y por condición urbana-rural.

En los estratos que no eran zonas metropolitanas, las UPM eran municipios. Se estratificaron por tamaño y código geográfico, lo que permitió formar estratos de población aproximadamente igual (utilizando datos del último censo de población disponible). Luego se seleccionaron dos municipios (UPM en estos estratos) en cada estrato utilizando muestreo sistemático con PPT, con la población total como medida del tamaño. Antes de la selección sistemática se declaró que algunos municipios eran UPM "indudables", debido a su gran población, y por lo tanto se incluyeron sin dudar en la muestra de municipios. Dentro de cada municipio seleccionado, las ZE se seleccionaron utilizando muestreo sistemático con PPT, con tamaños iguales al número de hogares privados obtenidos en el último censo de población. En la última etapa de selección se seleccionaron los hogares dentro de las ZE mediante muestreo sistemático a partir de listas actualizadas cada año. En la encuesta se incluyeron todos los miembros de los hogares seleccionados. Debería haberse seleccionado una muestra objetivo de 13 hogares a partir de cada ZE. No obstante, con el fin de reducir la variación de las ponderaciones resultante del desfase de las medidas del tamaño, en cada ZE se utilizaron fracciones muestrales constantes en vez de *tamaños* muestrales constantes, lo que dio lugar a conglomerados diversos.

La asignación de la muestra se hizo en forma no proporcional a los estratos, y el coeficiente entre la ponderación mayor y la menor fue aproximadamente igual a 8.

PME (*Pesquisa Mensal de Emprego*) de septiembre 1999, Brasil

La PME es una encuesta de mano de obra que comprende una muestra mensual de unos 40.000 hogares en las seis mayores áreas metropolitanas del Brasil, a partir de la cual se deducen los principales indicadores corrientes de la mano de obra. El diseño muestral es el mismo que en la PNAD en los estratos de las zonas metropolitanas, con excepción del conglomerado objetivo, que es de 20 en la PME, mientras que en la PNAD era de 13.

PPV (*Pesquisa de Padrões de Vida*) 1996/97, Brasil

La PPV tenía como objetivo medir los niveles de vida utilizando el planteamiento elaborado en la familia de encuestas del Estudio de medición de los niveles de vida (EMNV) realizadas en varios países bajo el patrocinio del Banco Mundial (Grosh y Muñoz, 1996). La encuesta del Brasil, llevada a cabo en 1996-1997, investigó una gran cantidad de características demográficas, sociales y económicas usando una muestra de 4.944 hogares seleccionados de 554 ZE en las regiones nororiental y sudoriental del Brasil. El diseño era una muestra conglomerada y estratificada en dos etapas. La estratificación supuso dos pasos: primero se formaron diez estratos geográficos para identificar las seis áreas metropolitanas de Fortaleza, Recife, Salvador, Belo Horizonte, Río de Janeiro y São Paulo, y otros cuatro estratos que cubrían el resto de las regiones nororiental y sudoriental, subdivididas en zonas de empadronamiento urbanas y rurales. En cada uno de estos 10 estratos geográficos, las ZE se subdividieron en 3 estratos de acuerdo con el ingreso medio del jefe de hogar registrado en el censo de población de 1991. Así se formó un total de 30 estratos.

El tamaño total de la muestra se fijó en 554 ZE, 278 para la región nororiental y 276 para la sudoriental. La asignación de la ZE dentro de los estratos era proporcional al número de ZE en cada estrato. La selección de la ZE se llevó a cabo utilizando una PPT con procedimiento de sustitución y se adoptó el número de hogares privados por ZE como medida del tamaño. En cada una de las ZE urbanas seleccionadas, se seleccionó una toma fija de 8 hogares, mediante muestreo aleatorio simple, sin sustitución. La toma de la encuesta por ZE rural se fijó en 16 hogares, por su mayor eficiencia en función de los costos.

A pesar de su pequeño tamaño en comparación con la PNAD y la PME, la encuesta PPV ofrece información útil sobre los efectos del diseño, ya que utilizó la estratificación del ingreso directo de las ZE y tomas menores por ZE que las otras encuestas. Otra característica distintiva se debe al hecho de que la estimación utilizó sólo las ponderaciones estándar de la probabilidad inversa de selección y que no hubo ningún intento de calibración con respecto a las proyecciones de la población. La variación de las ponderaciones de la muestra en la PPV fue considerable; hubo una diferencia de 40 veces entre la ponderación mayor y la menor.

REFERENCIAS BIBLIOGRÁFICAS

Grosh, M. y Muñoz, J. (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No. 126. Washington, D.C.: Banco Mundial.

Kish, L. (1965). *Survey Sampling*. Nueva York: Wiley.

National Institute of Statistics, Kingdom of Cambodia (1999). *Cambodia Socio-Economic Survey 1999: Technical Report on Survey Design and Implementation.* Phnom Penh.

Pettersson, H. (1994). *Master Sample Design: Report from a Mission to the National Central Statistics Office, Namibia, May 1994*. International Consulting Office, Statistics Sweden.

_____ (1997). *Evaluation of the Performance of the Master Sample 1992-96: Report from a Mission to the National Central Statistics Office, Namibia, May 1997*. International Consulting Office, Statistics Sweden.

_____ (2001). *Sample Design for Household and Business Surveys: Report from a Mission to the Bureau of Statistics, Lesotho May 21-June 2, 2001*. International Consulting Office, Statistics Sweden.

Phong, N. (2001). Personal correspondence concerning sample design for the Viet Nam Multipurpose Household Survey 1999.

Rosen, B. (1997). *Creation of the 1997 Lao Master Sample. Report from a Mission to the National Statistics Centre, Lao PDR*. International Consulting Office, Statistics Sweden.

Stoker, D. (2001). Personal correspondence concerning sample design for the October Household Survey and Labour Force Survey in the Republic of South Africa.

Verma, V., C. Scott y C. O'Muircheartaigh (1980). Sample designs and Sampling Errors for the World Fertility Survey. *Journal of the Royal Statistical Society,* Series A, vol. 143, parte 4, págs. 431-473.

Sección C

Errores no muestrales

Introducción

James Lepkowski
Universidad de Michigan
Ann Arbor, Michigan
Estados Unidos de América

1. En la secciones y capítulos anteriores se han examinado fundamentalmente errores de muestreo que se presentan cuando se toma una muestra de probabilidad representativa de una población. En esta sección se consideran otros errores que ocurren en las encuestas de hogares. Algunos de ellos, como el error de muestreo, varían en las diferentes muestras posibles o en las posibles repeticiones del proceso de medición. Otros errores son fijos y sistemáticos y no varían de una muestra a otra.

2. En el marco del diseño de la muestra, los errores variables normalmente se conocen como varianza de muestreo. Hay errores de muestreo fijos, algunos de los cuales se han mencionado ya, que suelen conocerse con el nombre de sesgos. La exclusión deliberada de un subgrupo de la población, por ejemplo, significa la falta de cobertura de ese subgrupo, error que estará presente y será de la misma magnitud, independientemente de la muestra que se seleccione.

3. Los errores no muestrales suponen errores de observación cuando no se logra obtener datos de una unidad de muestreo o variable, o errores de medición que se producen cuando se recogen los valores para la variable de la encuesta. Los errores que no son de observación suelen ser de carácter fijo, y dan lugar a consideraciones sobre el sesgo en las estimaciones de la encuesta. Los errores de medición algunas veces son fijos, pero pueden ser también variables.

4. Entre las causas de los errores que no son de observación hay dos de especial importancia: la no cobertura y la falta de respuesta. En el muestreo de probabilidad debe haber una población bien definida de elementos, cada uno de los cuales tiene una probabilidad de selección distinta de cero. La falta de cobertura se produce cuando un elemento de la población no tiene posibilidad real de ser seleccionado; en otras palabras, no puede llegar a formar parte de la muestra seleccionada. La falta de respuesta hace referencia a la situación en la que no se recopilan datos de una parte de la respuesta incluida en la muestra. Ello puede ocurrir porque un hogar o una persona se niegan a cooperar en absoluto o porque existen dificultades lingüísticas o problemas de salud, o debido a que no hay nadie en casa durante el período de la encuesta.

5. Los errores de medición pueden ser de diferentes procedencias: los encuestados, los entrevistadores, los supervisores e incluso los sistemas de procesamiento de datos. Los errores de medición de los encuestados pueden producirse cuando uno de ellos olvida la información y da una respuesta incorrecta, o deforma la información al responder a una pregunta delicada. Este tipo de error es probable que constituya un sesgo cuando el encuestado olvida sistemáticamente o distorsiona una respuesta siempre de la misma manera, indepen-

dientemente de cuándo se le formule. Estos errores pueden ser también variables. Algunos encuestados olvidan a veces una pregunta en un momento y la recuerdan en otro.

6. Los encargados del diseño de las encuestas deben considerar cuatro aspectos en relación con estos errores. El primero supone una definición atenta del error y un examen de las fuentes de tal error en el proceso de la encuesta, que incluye la parte del proceso de la encuesta que parece ser la causa de que se genere este tipo de error. El segundo supone la medición del tamaño del error, problema particularmente difícil. El tercero hace referencia a los procedimientos que deben elaborarse para reducir la magnitud del error, aunque su aplicación requiere muchas veces recursos adicionales. Finalmente, hay que decir que en todas las encuestas se producen errores no muestrales. Los encargados del diseño tratan de compensar esos errores en los resultados de la encuesta.

7. En los capítulos VIII y IX de esta sección se examinan, desde el punto de vista teórico, el error que no es no de observación y el error de medición, respectivamente, con ilustraciones de muchos tipos diferentes de dichos errores. En los capítulos X y XI se analizan de forma más detallada esos errores. En el capítulo X; se considera el efecto global en la calidad de los resultados de la encuesta, y en el capítulo XI se presenta un estudio monográfico de estos errores en el Brasil.

Capítulo VIII
Error de falta de observación en las encuestas de hogares en los países en desarrollo

James Lepkowski
Universidad de Michigan
Ann Arbor, Michigan
Estados Unidos de América

RESUMEN

La falta de observación en una encuesta se produce cuando no se realizan o no pueden realizarse mediciones de parte de la población objetivo o de la muestra. La falta de observación puede ser completa, en cuyo caso no se realiza ninguna medición de una unidad (por ejemplo, un hogar o una persona), o parcial, cuando se realizan algunas mediciones deseadas de una unidad, pero no todas. En este capítulo se examinan dos fuentes de falta de observación: la falta de cobertura y la falta de respuesta. Se dice que hay falta de cobertura cuando las unidades de población no tienen la posibilidad de ser seleccionadas para la encuesta. La falta de respuesta ocurre cuando un hogar o persona seleccionada para la encuesta no participa en ella o lo hace pero no da información completa. En este capítulo se examinan las causas, consecuencias y medidas para corregir los errores de falta de observación. La falta de cobertura y la falta de respuesta pueden dar lugar a estimaciones sesgadas cuando la parte de la población o de la muestra que queda excluida es diferente de la parte observada. Como los sesgos pueden ser graves, se examinan algunas medidas correctivas y ajustes para los casos de falta de cobertura y falta de respuesta.

Términos clave: Falta de respuesta, falta de cobertura, sesgo, población objetivo, marco muestral, tasas de respuesta.

A. Introducción

1. La falta de observación en la investigación de las encuestas se produce cuando se omite la medición de una parte de la población objetivo de la encuesta. La omisión puede ser completa, en cuyo caso no se realiza ninguna medición, o parcial, cuando se llevan a cabo algunas mediciones deseadas pero no todas.

2. Una fuente obvia de falta de observación es el proceso de muestreo. Sólo en un censo —una encuesta cuyo objetivo sea realizar mediciones de todos los componentes de la población— no hay ninguna falta de observación resultante de la elaboración de la muestra. La falta de observación asociada al muestreo da lugar a diferentes errores, que se examinan en los capítulos VI y VII de la presente publicación. Por ello no nos ocuparemos aquí de este factor de falta de observación.

3. En este capítulo se examinan otras dos causas de falta de observación: la falta de cobertura y la falta de respuesta. Como se explicará luego con mayor detalle, la falta de co-

bertura se produce cuando hay unidades de la población objetivo que no tienen la oportunidad de ser incluidas en la muestra de la encuesta; la falta de respuesta tiene lugar cuando una unidad incluida en la muestra no llega a participar en la encuesta, absoluta o parcialmente. En este capítulo se considerarán las causas de la falta de observación, sus posibles consecuencias, las medidas que pueden adoptarse para atajar el problema y los métodos para reducir el sesgo de las estimaciones de las encuestas que se pueda generar. Entre las consecuencias de la falta de cobertura y de respuesta se incluye la posibilidad de sesgo en los resultados de la encuesta. Si la parte de la población que queda excluida varía de la observada, habrá diferencias ente los resultados de la encuesta y lo que ocurre realmente en la población. Tales diferencias son sesgos por falta de observación, y pueden ser graves.

4. Naturalmente, es posible que aun cuando no se realicen mediciones de una parte de la población no se den sesgos por falta de observación. Si bien el registro de los casos de falta de observación es sencillo, descubrir el sesgo por falta de observación presenta dificultades. Estas dificultades son las que hacen que la consideración del sesgo por falta de observación no sea frecuente entre los temas de investigación. Es posible encontrar ejemplos en los que la falta de observación no represente ninguna diferencia en toda una encuesta, o en lo que respecta a la mayor parte de las preguntas de ella. Es posible también encontrar ejemplos en los que la falta de observación haya dado lugar a sesgos considerables en las estimaciones de la encuesta resultantes de una pregunta concreta, o a sesgos sustanciales en las estimaciones resultantes de un conjunto de preguntas, en cuyo caso habría que recelar de todos los resultados de la encuesta.

5. Se han realizado numerosas investigaciones sobre la falta de observación. La amplitud del tema obliga a que en este capítulo presentemos únicamente una introducción sobre el carácter de los errores por falta de cobertura y por falta de respuesta en las encuestas de hogares. El lector puede consultar las referencias bibliográficas, que examinan el tema con mayor detalle. En el apartado siguiente se presenta un marco para distinguir entre la falta de cobertura y la falta de respuesta, y en los que le siguen abordaremos por separado cada una de estas causas de error.

B. Marco para comprender los errores por falta de cobertura y por falta de respuesta

6. Para comprender la diferencia entre la falta de cobertura y la falta de respuesta hay que entender la naturaleza de las poblaciones y los marcos muestrales. La población objetivo es la colección de elementos sobre los que el diseñador de la encuesta desea conseguir estimaciones. Por ejemplo, a veces se pide al responsable del diseño que elabore una muestra para estudiar la participación en la mano de obra de las personas de 15 años o más que viven en un determinado país. Indudablemente, esta población tiene límites geográficos que están bien definidos (las fronteras del país), y límites en lo que respecta a las características de las unidades, determinados por la edad.

7. Hay otros aspectos implícitos en la definición de la población objetivo, como el significado de "persona que vive en el país". Muchas encuestas utilizan una definición de residencia, según la cual una persona debe haber vivido en el país la mayoría del año anterior, o si acaba de llegar al país debe tener la intención de quedarse en él permanentemente. Algunas partes de la población pueden quedar excluidas para un determinado tema de la encuesta. Por ejemplo, las personas en prisiones, en la cárcel o en cuarteles pueden quedar excluidas en algunas encuestas sobre las condiciones económicas. Así, puede ocurrir que las instituciones se excluyan porque tienen personas que no forman parte de la base conceptual para la medición que se va a realizar. Hay también una dimensión temporal implícita en la definición de población objetivo. Probablemente, lo que interesa en la encuesta es la participación actual en la población laboral activa y no la evolución histórica del individuo. En

tal caso la encuesta trata de realizar estimaciones sobre las características de la población tal como existe en un determinado momento.

8. La población objetivo es también la población de inferencia. En definitiva, se considera que los resultados de la encuesta se refieren a una población concreta. Muchas veces las encuestas se conciben para medir las características de las personas de un determinado país. Independientemente de que éstas se incluyan o no en el proceso de muestreo, el informe final de la encuesta puede hacer afirmaciones infundadas sobre toda la población. Por ejemplo, aun cuando la encuesta excluya a las personas que viven en instituciones, en el informe final quizá se afirme que los resultados se aplican a todas las personas que viven en el país. El lector poco informado puede llegar a suponer que los resultados también representan a las personas que viven en instituciones, aun cuando no se incluyeran en el proceso de muestreo. Por ello, al describir la encuesta en las publicaciones pertinentes, es importante incluir indicaciones detalladas y completas acerca de la población objetivo y la población de la encuesta.

9. La población objetivo difiere muchas veces de otra población importante: el conjunto de elementos a partir de los cuales se obtiene realmente la muestra, que es lo que se conoce como marco muestral. El marco muestral es la colección de materiales utilizados para extraer la muestra, y quizá no se corresponda exactamente con la población objetivo. Por ejemplo, en algunos países, se utilizan como marco muestral registros de direcciones preparados y mantenidos por un organismo de seguridad pública, como la policía. Pero algunos hogares de la población no se encuentran en esos sistemas administrativos. En tales circunstancias el marco muestral difiere de la población objetivo.

10. En otros casos el marco difiere de la población objetivo por cuestiones estructurales o por una decisión deliberada. Una parte de la población puede quedar excluida del marco por motivos administrativos o de costos. Por ejemplo, puede haber una región, varios distritos o una provincia de un país donde se estén produciendo disturbios civiles. Los organismos de seguridad pública pueden limitar las entradas y salidas de la región. El encargado de diseñar la encuesta quizá deje deliberadamente la región fuera del marco, aun cuando existan materiales para elaborar la muestra en la región.

11. También el costo puede influir en la decisión de excluir una parte de la población. En muchos países, las personas que viven en zonas remotas y escasamente pobladas se ven excluidas del marco muestral por el alto costo que representaría su inclusión en la muestra. Además, en los países con muchas lenguas autóctonas se requieren fuertes desembolsos para poder contar con las traducciones oportunas y con entrevistadores que puedan hablar todas las lenguas, por lo que los encargados de diseñar la encuesta y los patrocinadores de la misma quizá excluyan expresamente a los miembros de la población que no hablan una de las lenguas principales del país. En este caso quizá no sea posible excluir a una persona hasta que se haya identificado el hogar y se haya comprobado la capacidad lingüística de sus miembros. La exclusión se lleva a cabo mediante un proceso selectivo en el hogar.

12. Por otro lado, al diseñar la encuesta quizá se prefiera clasificar este tipo de problema como falta de respuesta, es decir, como falta de cobertura debida a la exclusión lingüística o como falta de respuesta debida a la incapacidad de comunicación. La decisión sobre la forma de clasificar las "exclusiones lingüísticas" depende en parte de la magnitud del problema. Por ejemplo, en un país la encuesta puede limitarse a las poblaciones que hablan una de las varias lenguas oficialmente reconocidas. Esta decisión puede excluir a un número considerable de personas que no hablan esas lenguas. Por el contrario, en otro país donde prácticamente todos hablan una de las lenguas oficiales es posible que se contacte pero no se entreviste a pequeños grupos de población que hablen idiomas no oficiales a los que no se han traducido los cuestionarios. En el primer caso quizá convenga, con la debida documentación, clasificar a los grupos lingüísticos excluidos como un caso de falta de cobertura. En el segundo, la ausencia de entrevistas podría clasificarse como falta de respuesta.

13. La falta de cobertura se produce cuando hay elementos en la población objetivo que no corresponden a las listas del marco muestral. En las encuestas de hogares se presentan problemas típicos de falta de cobertura cuando las unidades de habitación no se incluyen en un lista preparada durante las operaciones sobre el terreno, cuando se utilizan listas administrativas de hogares que están desfasadas o son inexactas o cuando se omite de las listas de residentes a individuos de un hogar.

14. La falta de cobertura se produce cuando se niega a un elemento de la población la oportunidad de ser seleccionado para la muestra de la encuesta, mientras que la falta de respuesta se debe a un intento fallido de recopilar datos de una unidad de la muestra que reúna los debidos requisitos y pertenezca a la población objetivo. La falta de cobertura se debe a errores o problemas en el marco utilizado para la selección de la muestra; la falta de respuesta se produce una vez que se han elaborado los marcos y se han seleccionado los elementos de la muestra a partir del marco. Por ejemplo, supongamos que en un hogar muestreado un residente varón del hogar está ausente en el momento de la entrevista porque durante la semana tiene empleo temporal fuera de la aldea donde se encuentra el hogar. Si ese residente no está enumerado en la lista del hogar durante la entrevista inicial porque el informante del hogar se olvidó de él, se ha producido un caso de falta de cobertura. Por el contrario, si un residente está incluido en la lista pero se encuentra fuera durante el período de las entrevistas y la encuesta acepta únicamente datos comunicados por el propio residente y, en consecuencia, no se recopilan datos de él, ese residente se incluye entre los no declarantes.

15. La falta de cobertura afecta normalmente a unidades enteras, como los hogares o las personas. La falta de respuesta puede afectar a unidades enteras o a datos concretos. Por ejemplo, habría una falta de cobertura si no se incluye un hogar en la lista de una aldea porque se encuentra encima de una tienda. Toda la unidad está ausente del marco. La falta de respuesta podría producirse cuando el hogar, aun incluido en la lista, se niega a participar en la encuesta, o cuando algunos miembros del hogar cooperan y facilitan datos mientras que otros no se encuentran en casa o se niegan a responder. Estas dos formas de falta total de respuesta de una unidad, sea el hogar o la persona, contrasta con el caso en que un miembro del hogar responde a todas las preguntas de la encuesta con excepción de a un subconjunto. Por ejemplo, el declarante de un hogar puede negarse a facilitar datos sobre sus ingresos en la economía informal, quizá por temor a una intervención administrativa oficial por sus ingresos no declarados. Esta forma última de ausencia de respuesta se conoce como falta de respuesta sobre una partida. Téngase en cuenta que el tipo de falta de respuesta en este caso depende también de si la unidad de análisis es la persona o el hogar: la falta de respuesta de una persona es una falta de respuesta sobre una partida para el análisis por hogares, pero una falta total de respuesta para el análisis por persona.

16. Es importante también considerar las soluciones de compromiso entre falta de cobertura y la falta de respuesta. Si bien podrían identificarse muchas causas de falta de cobertura o de respuesta en una encuesta dada mediante un estudio detenido —y quizá exista el deseo de reducir el grado de uno u otro de esos problemas—, ello supondrá el gasto de los escasos y limitados recursos de la encuesta. Luego podría haber competencia por estos recursos en lo que respecta a la reducción de estas dos causas de error.

17. Por ejemplo, supongamos que en un país con 40 idiomas o dialectos importantes el instrumento de la encuesta se traduce a cinco idiomas que se hablan en los hogares del 80% de la población. El sexto grupo lingüístico más hablado representa el 3% de la población. Al mismo tiempo, supongamos que en la encuesta se prevén dos visitas a un hogar durante un período de dos días con el fin de encontrar a alguno de sus miembros, y que se sabe que el 10% de los hogares visitados dos veces no responderán a la encuesta porque no va a haber nadie en casa durante los dos días de las entrevistas. El encargado de diseñar la encuesta tiene que tomar una decisión acerca de los recursos. Podrían destinarse más fondos

a traducir el instrumento al sexto idioma para cubrir un 3% adicional de la población que habla esa lengua. La otra alteernativa sería destinar más fondos para que los entrevistadores dedicaran un tercero o cuarto día en cada aldea a realizar visitas a los hogares para tratar de encontrar en casa a una proporción mayor de miembros del hogar.

18. La decisión sobre cómo utilizar los posibles recursos adicionales, bien para la traducción o para aumentar el número de visitas a los hogares, dependerá de los sesgos previstos y de los costos y recursos implicados. Los sesgos dependen tanto del nivel de la falta de cobertura como de la falta de respuesta y de las diferencias entre las poblaciones incluidas y no incluidas o entre las personas de la muestra que responden y las que no responden.

19. Estos tipos de soluciones de compromiso entre costo y error son frecuentes cuando se diseña una encuesta. No tiene sentido considerar detalladamente en este capítulo el tipo de datos necesario para las soluciones de compromiso o la forma en que deben realizarse. En el diseño de la mayor parte de las encuestas las decisiones se toman con información limitada y de una manera informal.

C. Error por falta de cobertura

1. Causas de la falta de cobertura

20. Las causas de la falta de cobertura en las encuestas de hogares dependen de los materiales del marco utilizados para seleccionar la muestra. Como muchas encuestas de hogares en los países en desarrollo y en algunos países en transición utilizan métodos de muestreo de zonas, aquí nos limitaremos a considerar los problemas de marco y de falta de cobertura en las encuestas de hogares basadas en las muestras de zonas.

21. El muestreo de zonas suele ir unido a la selección en múltiples etapas. La etapa primaria, y algunas veces la secundaria, de selección, implican zonas geográficas que pueden considerarse conglomerados de hogares. En una fase de selección posterior, debe obtenerse o crearse una lista de hogares para un conjunto de zonas geográficas relativamente pequeñas. En la última etapa de selección se crea en cada zona incluida en la muestra una lista de personas o residentes en los hogares. Así pues, hay tres tipos de unidades que deben considerarse al examinar la falta de cobertura en dichas encuestas: las unidades geográficas, los hogares y las personas. Como se examina más adelante, estas unidades pueden ser también causas independientes de falta de respuesta en las encuestas de hogares.

22. La falta de cobertura de las unidades geográficas como consecuencia de las deficiencias existentes en el marco muestral es algo que se produce sólo rara vez, ya que la mayor parte de los marcos están basados en materiales del censo que abarcan toda la extensión geográfica de una población. La falta de cobertura de una zona geográfica es algo que puede ocurrir, pero de manera más sutil, como se ha dicho antes. Una encuesta puede concebirse de manera que permita hacer inferencias sobre toda la población de un país o de una región de un país, y las referencias a la población en el informe final pueden incluir de hecho a la población que vive en toda la superficie, aunque quizá la muestra no se haya seleccionado teniendo en cuenta todo el país.

23. Por ejemplo, durante el diseño de la encuesta se pueden identificar algunas zonas geográficas con una proporción limitada de la población cuya cobertura resultaría sumamente costosa. Entonces se toma la decisión deliberada de excluir del marco a esas zonas geográficas. No obstante, al comunicar los resultados de la encuesta no se explica, o se menciona sólo en forma sumaria, que se ha prescindido de esas áreas. Los lectores del informe pueden tener la impresión de que los resultados de la encuesta se aplican a todo el país o región, cuando de hecho una parte de la población no está incluida. En la práctica, la magnitud del error de falta de cobertura resultante de estas situaciones suele ser irrelevante, y en general no se tiene en cuenta.

24. Es importante tener presente que continúa habiendo una diferencia entre una población objetivo deseada (es decir, la población que vive en toda la extensión geográfica del país) y una "población de la encuesta" restringida que vive en la zona geográfica incluida. No obstante, existe el peligro de que debido a una documentación incompleta el usuario de los datos pueda tener la impresión de que la muestra de la encuesta cubre toda la población, cuando de hecho no es así.

25. Una causa más importante de falta de cobertura está relacionada con los hogares. La mayoría de las encuestas consideran que los hogares son un conjunto de personas que normalmente residen en una unidad de habitación. Así pues, hay dos componentes importantes: la definición de residente habitual y la de unidad de habitación.

26. Las definiciones de unidad de habitación son complejas en la medida en que tienen en cuenta si una estructura física está destinada a vivienda y si las personas que residen en ella viven y comen separadamente de otras dentro de la misma estructura (como en las estructuras con varias unidades; por ejemplo, los edificios de apartamentos). El vivir por separado significa que los residentes tienen acceso directo al lugar de residencia desde el exterior de la estructura o desde un vestíbulo o pasillo compartido. La posibilidad de "comer por separado" implica normalmente la existencia de un lugar para servir y preparar la comida, o la libertad completa de los residentes de elegir lo que comen.

27. Es difícil aplicar este tipo de definición amplia a las muchas y diversas situaciones que se registran en los distintos países, o en las diferentes regiones de un país. La mayor parte de las unidades de habitación se identifican fácilmente, por ejemplo, las unidades de vivienda unifamiliares o separadas; varias unidades de vivienda distintas que comparten una pared pero que tienen entradas diferentes, y los apartamentos en los edificios con estructuras múltiples. No obstante, hay muchas unidades de habitación que son difíciles de clasificar o de localizar. Por ejemplo, en las zonas urbanas de tugurios, las unidades de habitación independientes pueden ser difíciles de identificar cuando las personas viven en estructuras construidas con materiales reciclados o de desecho. Las unidades de vivienda se ubican a veces en lugares que no pueden identificarse mediante una inspección casual de las entradas desde una calle, callejón o camino.

28. En las zonas rurales, a veces es fácil identificar una estructura destinada a habitación, pero los complejos mecanismos sociales dentro de la estructura pueden dificultar la identificación de las unidades de habitación independientes. Por ejemplo, en un grupo tribal se utilizan para habitación casas corridas con una sola entrada y con compartimientos donde puede dormir cada unidad familiar, con una zona común para la preparación de los alimentos para las comidas familiares colectivas o individuales; es decir, los compartimientos individuales no son propiamente unidades de habitación, ya que no tienen una entrada independiente ni su propio lugar para cocinar y comer. En este caso el concepto de hogar como grupo de personas que residen normalmente en una unidad de habitación específica es más difícil de aplicar. No está claro si lo que debe tratarse como unidad de habitación es toda la estructura o cada uno de sus compartimientos. En la práctica, la totalidad de la casa se trata como una unidad de habitación o vivienda, y si forma parte de la muestra se incluyen en la encuesta todos los hogares identificados durante la enumeración de hogares sobre el terreno.

29. Hay también locales de habitación que no se consideran unidades de vivienda. Los locales institucionales ocupados por individuos que están bajo la atención o custodia de otros, como los orfanatos, prisiones o cárceles o los hospitales, no se consideran unidades de habitación. Los dormitorios de estudiantes, monasterios y conventos y refugios para personas sin hogar son tipos especiales de habitación que no ofrecen necesariamente la atención o la custodia asociadas con una institución. Los locales de habitación para situaciones transitorias o estacionales son también un problema. Por ejemplo, puede haber unidades de

habitación independientes en una zona agrícola para los trabajadores temporeros contratados para una o varias campañas al año. Supuestamente, estos residentes estacionales viven habitualmente en otro lugar, y no deben considerarse como parte de un hogar en la unidad estacional.

30. Para el muestreo de zonas por etapas en los países en desarrollo hay que enumerar en un determinado momento del proceso las viviendas de pequeñas zonas geográficas, como un bloque en una ciudad o una zona de empadronamiento en un área rural. Muchas veces se producen situaciones de falta de cobertura cuando el personal de la encuesta a tiempo parcial visita sobre el terreno las unidades de habitación y se encuentra con las complejas situaciones descritas más arriba. La identificación de la mayor parte de las unidades de habitación es sencilla, pero la ausencia de algunas de ellas puede ser todavía un fenómeno común en la medida en que el personal a tiempo parcial tiene limitada experiencia para aplicar a una situación compleja una definición que consta de varios componentes.

31. El problema de la falta de cobertura en la enumeración de las unidades de habitación se complica todavía más por la dimensión temporal. Una unidad de habitación puede estar desocupada, o en construcción, en el momento de la enumeración. Si la encuesta se va a realizar más adelante, este tipo de unidades quizá deba incluirse en la enumeración. En las encuestas en las que las mismas unidades de habitación se utilizan en más de una ocasión para una única encuesta, o en varias encuestas diferentes, es normal tratar de incluir las unidades de construcción que están desocupadas o en construcción.

32. En las encuestas realizadas en países en transición es posible utilizar una lista ya preparada por una autoridad administrativa. No obstante, es preciso evaluar atentamente su calidad para las encuestas de hogares. Los mismos problemas acabados de mencionar podrían ocurrir con respecto a las listas administrativas.

33. Así pues, el proceso de enumeración de las unidades de habitación puede generar falta de cobertura de determinados tipos de hogares. Este problema puede ser difícil de identificar sin una inversión considerable de recursos adicionales para la encuesta.

34. Finalmente, dentro de una unidad de habitación de la muestra la enumeración de las personas que son residentes habituales forma también parte del proceso de enumeración de los hogares. Deben darse a los entrevistadores normas claras sobre a quién deben incluir en la unidad de habitación como residente habitual. En el caso de las unidades de habitación, la mayor parte de las situaciones son claras. La mayoría de las personas encontradas residen en la unidad de vivienda en el momento del encuentro, y ése es su único lugar de residencia. Habrá otras que se encuentren ausentes pero que tienen allí su único lugar de residencia.

35. En cambio, hay personas para las cuales la unidad de habitación es sólo una de las varias en que viven. El personal a tiempo parcial sobre el terreno debe tomar una decisión sobre si esa unidad es el lugar habitual donde reside la persona. Es también difícil para los informantes de los hogares señalar con precisión cuál es la situación de algunos residentes, y si la información es ofrecida por un sustituto quizá no sea del todo exacta.

36. Los informantes pueden tener también razones personales para excluir deliberadamente a algunas personas que son residentes habituales. Por ejemplo, una persona puede estar viviendo en una unidad de habitación y su presencia podría determinar que la vivienda deje de recibir prestaciones públicas que ya recibe. Asimismo, un informante puede excluir deliberadamente a un residente que no quiere ser identificado por los organismos públicos o privados debido a problemas financieros (por ejemplo, deuda) o legales (por ejemplo, actividades delictivas).

37. Los informantes a veces no incluyen a alguien en el hogar por razones culturales o por supuestos reales o infundados. Un informante puede no señalar la presencia de un lactante de menos de un año porque en su cultura se considera que no tienen edad suficiente

todavía para considerarlos como personas. Pueden excluirlos también porque creen que la organización de la encuesta no está interesada en recopilar datos sobre niños de tan poca edad, o simplemente pueden olvidarse de incluir a alguien, sea un lactante o de más edad.

38. La falta de cobertura en las encuestas de hogares puede deberse, pues, a diferentes circunstancias relacionadas con la definición y las operaciones. El problema es determinar hasta qué punto la falta de cobertura da lugar a un error en los resultados de la encuesta.

2. Error por falta de cobertura

39. Supongamos que la encuesta trata de estimar la media de alguna característica Y de una población de N personas, N_{nc} de las cuales no se incluyen en el marco muestral de la encuesta. Supongamos que la media de la población de tamaño N es $\overline{Y}$, que $\overline{Y}_c$ es la media de los incluidos en el marco muestral y que $\overline{Y}_{nc}$ es la media de los no incluidos en el marco. El error asociado con la falta de cobertura se conoce como sesgo de falta de cobertura de la media de la muestra, $\overline{y}_c$, que está basada únicamente en los incluidos en la muestra y que de hecho estima $\overline{Y}_c$ más que $\overline{Y}$.

40. El sesgo de la media de la muestra, $\overline{y}_c$, depende de dos componentes: la proporción de la población no incluida, N_{nc}/N, y la diferencia en las medias de la característica Y entre las personas incluidas y no incluidas. Por ello,

$$B(\overline{y}_c) = (N_{nc} / N)(\overline{Y}_c - \overline{Y}_{nc})$$

41. Esta formulación del sesgo de falta de cobertura permite comprender cómo abordan el problema de la falta de cobertura los encargados del diseño de las encuestas. Para mantener bajo el nivel de error asociado con la falta de cobertura, o para reducir sus efectos, deben conseguir pequeñas diferencias entre las personas incluidas y no incluidas, o tener una pequeña proporción de personas no incluidas en la encuesta.

42. Una dificultad importante con esta formulación es que en la mayoría de las encuestas no se conocen ni la diferencia $(\overline{Y}_c - \overline{Y}_{nc})$ ni la proporción (N_{nc}/N) no incluida. Además, la tasa de falta de cobertura (N_{nc}/N) puede variar también en las diferentes subclases. La diferencia puede variar entre las distintas variables y entre las subclases de personas (como una región, o un subgrupo, definido mediante una característica demográfica como la edad). Así pues, el error de falta de cobertura no se debe a la encuesta, sino a la característica individual y al estadístico estimado.

43. En muchas organizaciones gubernamentales de encuestas, se exigen con frecuencia estimaciones de un total. El sesgo de falta de cobertura asociado con un total no sólo depende de las diferencias entre las unidades incluidas y no incluidas en la característica pertinente, sino también del número (y no la tasa) de las no incluidas. Es decir, con un total estimado de informantes $\hat{Y}_r = N\overline{y}_r$, el sesgo es $B(\hat{Y}_r) = N_{nc}(\overline{Y}_r - \overline{Y}_m)$.

Reducción, medición y notificación del error de falta de cobertura

44. Hay cuatro posibles medios de abordar el error de falta de cobertura en las encuestas de hogares:

- Reducir el nivel de falta de cobertura mejorando los procedimientos sobre el terreno.
- Crear procedimientos para medir la magnitud de falta de cobertura y señalarlo en la encuesta.
- Tratar de compensar el error de falta de cobertura con ajustes estadísticos.
- Explicar con el mayor detalle posible en el informe de la encuesta cuáles son las características de la falta de cobertura.

45. Para reducir el error de falta de cobertura en las encuestas de hogares, normalmente se recurre al uso de marcos múltiples o a métodos que permiten mejorar los procesos de enumeración de la encuesta. La utilización de marcos múltiples para las unidades de habitación es mayor que para las personas. Los marcos requieren disponer de listas separadas de aquellas unidades de habitación que planteen problemas específicos para su enumeración sobre el terreno.

46. Por ejemplo, supongamos que las unidades de habitación estacionales para los trabajadores agrícolas presentan dificultades de enumeración sobre el terreno en un determinado país. Supongamos también que un organismo encargado de la producción agrícola, la educación o el bienestar social tiene una lista con el número y el tipo de unidades de habitación estacionales en las explotaciones agrícolas o en las empresas que emplean y alojan a trabajadores temporeros. La lista de unidades de habitación de esta fuente alternativa puede utilizarse como marco independiente. Los entrevistadores sobre el terreno que preparan listas de unidades de habitación recibirían una lista con las explotaciones agrícolas o empresas de las que el organismo cuenta ya con listas disponibles, y se les indicaría que no enumeraran las unidades de habitación estacionales en ese lugar. Luego las muestras de unidades de habitación para la encuesta se seleccionarían a partir de la lista de unidades de habitación preparada por el entrevistador y de la lista mantenida por el organismo gubernamental. Indudablemente, seguirá habiendo falta de cobertura en ambas listas, y quizá puedan producirse también casos de "sobrecobertura"; pero el uso de ambos marcos puede reducir el nivel de falta de cobertura y el error asociado a ésta.

47. Es también importante considerar los métodos para mejorar los procesos de enumeración. Cuando se dispone de fuentes administrativas con listas de unidades de habitación, pueden comprobarse mediante una actualización sobre el terreno antes de seleccionar la muestra. Una posibilidad es enviar a los entrevistadores a las zonas geográficas con una lista de unidades de habitación sacada de fuentes administrativas, con instrucciones sobre la manera de comprobar y agregar o borrar unidades de habitación de la lista a medida que van examinando el lugar.

48. Se puede también capacitar a los entrevistadores para que utilicen sobre el terreno un procedimiento de "intervalo semiabierto" para localizar las unidades de habitación no incluidas en las listas administrativas o poner al descubierto las listas realizadas sobre el terreno en las que faltan unidades. Dicho procedimiento supone la selección de una unidad de habitación de una lista de direcciones, una visita del entrevistador a la unidad muestreada y un orden de lista implícito o explícito. En la unidad se indica al entrevistador que se informe sobre las posibles unidades de habitación adicionales que podría haber entre la unidad de habitación seleccionada y la que aparece a continuación en la lista.

49. La siguiente unidad de la lista se define mediante una especie de ruta previamente determinada a través de una zona geográfica. En el caso de en un bloque urbano, por ejemplo, se indicará a los entrevistadores que preparan una enumeración, que comiencen en una esquina determinada y que continúen en el sentido de las agujas del reloj. La lista de unidades de habitación debe recopilarse en ese mismo orden.

50. Si un entrevistador encuentra una unidad de habitación que no aparece en la lista y está entre la unidad de habitación seleccionada y la siguiente de la lista, se debe agregar a la muestra la unidad de habitación que se ha pasado por alto y tratar de conseguir una entrevista. Si son varias las unidades que se encuentran en esa situación, el entrevistador debería solicitar a la oficina central nuevas instrucciones, para evitar trastornos en las operaciones sobre el terreno.

51. Dentro de los hogares, una forma de mejorar los procedimientos de enumeración puede ser establecer secuencias de preguntas, administradas por el entrevistador al informante, con el fin de localizar a las personas omitidas. Por ejemplo, se puede indicar al

entrevistador que pregunte si hay algún lactante que no se ha incluido en la lista de residentes habituales. También puede mejorarse la enumeración de hogares si los entrevistadores reciben orientaciones sobre la elección de los informantes adecuados o instrucciones para que repitan al informante los nombres de la lista de personas y que no se pase a nadie por alto.

52. La medición del sesgo de falta de cobertura es también importante y representa un problema de difícil solución. ¿Cómo puede una organización identificar las unidades que no se incluyen en ninguna de sus listas? Como la medición de la falta de cobertura puede ser una tarea costosa, sólo se lleva a cabo en forma ocasional.

53. Una forma habitual de evaluar el error de falta de cobertura es comparar los resultados de la encuesta, en particular las variables que se prestan a comparación, con los datos procedentes de fuentes externas o independientes. Para evaluar la magnitud de la falta de cobertura, una encuesta puede comparar la distribución de las personas de su muestra por edad y género con la obtenida de un censo reciente o de los registros administrativos. Las diferencias en las distribuciones indicarán los problemas de falta de cobertura. Para evaluar el error de falta de cobertura asociado con una variable puede establecerse una comparación de los valores del estadístico de interés con una fuente independiente. Por ejemplo, el salario total y los ingresos salariales registrados en una encuesta, en relación con la muestra total y con los subgrupos principales, pueden compararse con los informes administrativos sobre los salarios y los ingresos salariales. En un estudio clásico, Kish y Hess (1950) compararon la distribución de las unidades de habitación en una encuesta con datos de un censo reciente sobre la distribución de las unidades de habitación por bloques. Esa comparación permitió comprender el carácter del problema de la falta de cobertura en la recopilación de datos de las encuestas.

54. Un sistema más costoso para evaluar el error de falta de cobertura es recurrir a un sistema dual de medición o a procedimientos de confrontación de casos afines. Los censos emplean métodos de sistema dual para evaluar la cobertura [véase, por ejemplo, Marks (1978)]. En un censo se compara una encuesta independiente con los resultados de un censo para identificar los problemas de falta de cobertura. La evaluación del alcance de la falta de cobertura depende de una confrontación caso por caso de la muestra de la encuesta con los elementos del censo para determinar qué elementos de la muestra no aparecían en el censo. Estos procedimientos están estrechamente relacionados con los métodos de "muestreo de captura-recuperación" utilizados en los estudios ambientales de las poblaciones animales.

55. Como las encuestas de hogares se ven afectadas universalmente por el error de falta de cobertura, muchas emplean como procedimientos estadísticos los ajustes de control de poblaciones o la posestratificación para ajustar los resultados de la encuesta y compensar el error de falta de cobertura. Estos ajustes son muy semejantes al método esbozado antes para evaluar la magnitud del error de falta de cobertura. La distribución de la muestra por edad y género, por ejemplo, puede compararse con la de una fuente externa, tal que un censo reciente o las proyecciones de población. Cuando la distribución de la muestra es baja (o alta) en un grupo de edad-género, puede aplicarse una ponderación a todos los datos de las personas de la muestra de dicho grupo de edad-género para aumentar (reducir) su contribución a los resultados de la encuesta. Se necesitarán estimadores ponderados para realizar adecuadamente las ponderaciones del análisis.

56. Como consideración final sobre la falta de cobertura cabe señalar la importancia que reviste una información adecuada para toda organización de estadística. Los informes de análisis deben ofrecer definiciones claras de la población objetivo, y en particular de las posibles exclusiones. El marco debería describirse con bastante detalle para que el lector vea cómo pueden producirse errores de falta de cobertura, e incluso realizar una evaluación provisional sobre la magnitud del posible error. Sería útil incluir como referencias o apéndices evaluaciones de calidad del marco, como las comprobaciones de la calidad de las listas ad-

ministrativas o de unidades de habitación, o una comparación entre las listas originales de personas dentro de las unidades de habitación y las listas obtenidas de entrevistas realizadas, para evaluar el control de calidad.

57. Un problema más difícil es señalar las posibles tasas de cobertura o sesgos de falta de cobertura en la población y las subclases de población. Estos tipos de evaluación sólo pueden realizarse con respecto a las encuestas en curso cuando en algún momento se ha intentado evaluar la magnitud del problema de falta de cobertura. Es muy difícil, por no decir imposible, realizar esas evaluaciones en las encuestas transversales que se llevan a cabo una sola vez.

58. Finalmente, si se efectúan ajustes de control de población o de posestratificación, la documentación de la encuesta debe contener una descripción de los procedimientos de ajuste y de su magnitud para subgrupos importantes de la población.

D. Error por falta de respuesta

59. El error por falta de respuesta presenta algunos paralelos con el error por falta de cobertura en lo que se refiere a definiciones, medición, reducción, compensación e información. Por eso, esta sección se organiza en forma muy semejante a la anterior. No obstante, es importante tener claro que se trata de problemas muy distintos, con diferentes causas y en algunos casos con soluciones distintas. Mientras que en el caso de la falta de cobertura los encargados del diseño de la encuesta casi siempre lo único que saben es la ubicación y características generales de la proporción no incluida de la población, en la falta de respuesta conocen al menos la información sobre el marco relativa a quienes no responden. Parece también que la falta de respuesta es un fenómeno más extendido en las encuestas de hogares, y por lo tanto su contribución al sesgo de las estimaciones de la encuesta puede ser mayor.

60. Como ya se ha señalado, en las encuestas de hogares se identifican con frecuencia dos tipos de falta de respuesta, una total y otra relativa a una de las partidas. Estos dos tipos tienen repercusiones muy diferentes en los resultados de la encuesta. Los métodos utilizados para medir, reducir y notificar, así como para compensar las faltas de respuesta, son también en cierta manera distintos. Podría dedicarse una sección especial a cada uno de los dos tipos, pero aquí se ha preferido tratarlos conjuntamente.

1. Causas de la falta de respuesta en las encuestas de hogares

61. En las encuestas de hogares la falta total de respuesta puede ocurrir en diferentes tipos de unidades. Como ocurre con la falta de cobertura, la falta de respuesta puede afectar a unidades de muestreo primarias o secundarias. Por ejemplo, una unidad primaria de muestreo podría ser un distrito o un subdistrito de un país. Las condiciones atmosféricas o las catástrofes naturales podrían impedir la realización de la encuesta en un distrito o subdistrito que se ha seleccionado en una fase de muestreo primaria o secundaria. La unidad se incluye en la encuesta, pero durante la realización de ésta no es posible recopilar datos de ninguno de los hogares de la unidad.

62. La falta de respuesta de los hogares es más frecuente. Es posible que una unidad de habitación elegida para la encuesta esté ocupada y que se intente realizar una entrevista, pero que cuando se presente el entrevistador algunos acontecimientos adversos impidan la recopilación de datos. Un miembro del hogar puede negarse a participar a título personal o en cuanto representante de toda la unidad.

63. Aun cuando una unidad de habitación esté ocupada, es posible que sus residentes se encuentren fuera de casa durante todo el período de la encuesta. En algunos países en desarrollo se plantean problemas serios con unidades de vivienda donde consta que vive alguien pero que se encuentran cerradas durante todo el período de recopilación de datos.

64. En muchos países, aunque en las unidades de habitación haya individuos en el momento de la recopilación de datos, los problemas lingüísticos pueden representar un obstáculo. Quizá el cuestionario no se ha traducido a la lengua que se habla en ese hogar, o el entrevistador no habla la lengua local. Para evitar casos de falta de respuesta se pueden contratar intérpretes locales que acompañen a los entrevistadores a las casas. Hay otras encuestas que rechazan esta práctica por temor a que la traducción no sea correcta o que no sea la misma en todos los hogares. No obstante, los hogares que no pueden aportar respuestas por problemas lingüísticos pueden clasificarse entre las unidades que no han respondido. Algunas organizaciones lo que hacen es excluir de la encuesta a los hogares que no hablan ninguna de las lenguas a las que se ha traducido el cuestionario. En ese caso el problema de los hogares sería un caso de falta de cobertura más que de falta de respuesta. La decisión que tome la organización —considerar que se trata de una falta de cobertura o de una falta de respuesta— debe indicarse claramente en la documentación de la encuesta.

65. Puede darse también el caso de falta total de respuesta en el plano personal. En las encuestas que permiten que otra persona responda a las preguntas pueden recopilarse datos de otros miembros del hogar acerca de personas de éste que no se encuentran presentes en el momento de la entrevista. En cambio, en las encuestas que exigen que sea el interesado quien responda a algunas o a todas las preguntas, quienes no se encuentran en casa durante la encuesta, se niegan a participar o tienen algún obstáculo (por ejemplo, lingüístico) que impide la entrevista figuran como no declarantes. A veces la falta de respuesta se debe a problemas de salud, sean permanentes (la sordera o la ceguera) o temporales (un episodio de una enfermedad aguda grave).

66. En cuanto a los hogares con problemas lingüísticos, algunas organizaciones consideran que cuando una persona tiene dificultades lingüísticas o problemas de salud permanentes habría que hablar de falta de cobertura, mientras que si las afecciones son temporales sería un caso de falta de respuesta (Seligson y Jutkowitz, 1994). No hay normas comúnmente aceptadas para decidir cómo hacer esa clasificación. En una encuesta de ingresos o gastos las personas con afecciones de salud temporales son lo bastante escasas como para que la organización las considere como no incluidas. En cambio, en una encuesta sobre las condiciones de salud las respuestas de estos individuos pueden ser lo bastante diferentes como para desaconsejar su exclusión. En ese caso podrían clasificarse entre los casos de falta de respuesta. Dada la ausencia de una práctica ampliamente aceptada, es importante que las organizaciones expliquen claramente en los informes de la encuesta cómo se han resuelto esos casos en una encuesta dada.

2. Sesgo por falta de respuesta

67. Son muchas más las investigaciones sobre el problema de la falta de respuesta en las encuestas de hogares que sobre la falta de cobertura [véanse, por ejemplo, los estudios de Groves y Couper (1998), y Lessler y Kalsbeek (1992)]. Ello se debe a varios factores.

68. La falta de cobertura es en cierto sentido menos visible que la falta de respuesta. Los hogares o personas no incluidos sencillamente no pueden ser objeto de estudio, mientras que las unidades que no han respondido pueden observarse y contarse; es más, se les podría animar a participar.

69. En los países desarrollados se presupone que la falta de cobertura es menos importante que la falta de respuesta, ya que las tasas de aquéllas son menores que las de ésta. En los países en desarrollo, donde las tasas de falta de respuesta son más bajas y las de falta de cobertura son muy superiores a las de los países desarrollados, puede ocurrir lo contrario. Conviene recordar que el sesgo por falta de cobertura en la media de una muestra puede atribuirse a dos factores: la magnitud de la tasa de falta de cobertura y la diferencia entre las medias de los grupos de población incluidos y no incluidos. De la misma manera, en el caso

de la falta de respuesta, la magnitud del sesgo por falta de respuesta en la media de una muestra puede atribuirse a la proporción de la población que no responde, y a la diferencia en las medias de población entre los grupos que responden y los que no lo hacen.

70. Como en el caso de la falta de cobertura, supongamos que la encuesta trata de estimar alguna característica Y, y que la media de población $\overline{Y}$ consta de una media de las personas que responden, por ejemplo $\overline{Y}_r$, y una media de las que no responden, $\overline{Y}_{nr}$. Supongamos que N_{nr} denota el número de personas que no responderían si se incluyeran en la muestra. El sesgo de la media de la muestra para los declarantes $\overline{y}_r$ sería entonces $B(\overline{y}_r) = (N_{nr} / N)(\overline{Y}_r - \overline{Y}_{nr})$. En cuanto a la falta de cobertura, al diseñar la muestra debe procurarse que la tasa de falta de respuesta sea pequeña, o prever pequeñas diferencias entre los hogares y personas que responden y los que no lo hacen. Este marco general puede utilizarse para comprender también la falta de respuesta sobre partidas concretas. El problema del sesgo por falta de respuesta sobre una partida es más complicado, porque muchas veces las partidas se consideran en forma conjunta, y la falta de respuesta sobre una partida es la unión de varias faltas de respuesta sobre varias partidas.

71. Mientras que en la falta de cobertura no se conoce ni la diferencia ni la tasa, la tasa de no respuesta se puede estimar con encuestas bien planteadas. Éstas mantienen registros detallados de la disposición de cada unidad de la muestra que se desea estudiar, sean hogares, personas o partidas de datos individuales, y permiten estimar la tasa de no respuesta directamente a partir de los datos. Pueden tener también datos que observar si las tasas de respuesta difieren entre las subclases importantes, sobre todo las subclases geográficas relativas a los hogares.

72. Para evaluar las diferencias entre los declarantes y los no declarantes se requiere una actividad más detallada de recopilación y medición. Muchas veces, durante la recopilación de datos de la encuesta, resulta imposible tratar de medir las características de los no declarantes. No obstante, a lo largo de la encuesta pueden realizarse estudios especiales para conseguir respuestas de las unidades no declarantes.

73. La falta de respuesta en las últimas oleadas de las encuestas por panel ofrece más información para estudiar los efectos del posible sesgo por falta de respuesta y acomodarse a él que la falta de respuesta en las encuestas transversales o que se realizan una sola vez. Encuestas por panel son las que realizan el seguimiento de las mismas unidades y recopilan datos de las unidades del panel repetidamente a lo largo del tiempo. Parte de las unidades pueden omitirse durante el seguimiento, lo que daría lugar a una falta de respuesta por desgaste a lo largo de la encuesta. No obstante, las investigaciones sobre la falta de respuesta en estos casos utilizan los datos recopilados en las oleadas anteriores para conocer mejor las diferencias entre los declarantes y los no declarantes, lo que serviría como base para los ajustes que se describen a continuación. En Lepkowski (1988) pueden verse algunas técnicas para compensar la falta de respuesta del panel.

74. La disponibilidad de algo más de información sobre los no declarantes que sobre las personas no incluidas y el uso posible de modelos de comportamiento para estudiar y compensar la falta de respuesta han generado también más investigaciones sobre ésta que sobre la falta de cobertura. Cuando se mantienen atentamente los registros de todas las unidades de la muestra, y no sólo de los que no responden, los datos de la muestra permiten realizar comparaciones directas entre los que responden y los que no. Además, la falta de respuesta se debe en general al comportamiento del hogar o de la persona: es un fenómeno de autoselección. Al diseñar la encuesta se puede consultar gran cantidad de publicaciones de sociología, psicología y psicología social para estudiar de qué forma los individuos y los grupos toman decisiones acerca de la participación en diversas actividades. Es posible estudiar los modelos de comportamiento (siempre que se disponga de datos sobre los no declarantes), con el fin de comprender los determinantes de la falta de respuesta en una encuesta.

3. Medición del sesgo por falta de respuesta

75. Para medir el sesgo por falta de respuesta es preciso medir las tasas de falta de respuesta y las diferencias entre los declarantes y los no declarantes en relación con las variables de la encuesta. El cálculo de la tasa de falta de respuesta de los hogares o personas a partir de los hogares de la muestra requiere a su vez una definición de los posibles resultados para todos los casos incluidos en la muestra y una especificación de cómo deben utilizarse esos resultados para calcular la tasa. Por ejemplo, las entrevistas terminadas y parciales (las que tienen datos suficientes para aportar información sobre los conceptos clave del estudio) muchas veces aparecen agrupadas.

76. Entre los casos admisibles de falta de entrevistas están las personas u hogares que se encuentran en la población y se han identificado al realizar la encuesta pero de los cuales no se recopilaron datos. Por ejemplo, si una encuesta se limita a las personas de 15 años o más, serían casos de falta de entrevista admisible las personas de 15 años o más de quienes no se recopilaron datos. Normalmente hay al menos tres causas de falta de entrevistas: rechazos (Ref), es decir, personas u hogares con los que se ha entablado contacto pero que se niegan a participar en el estudio; falta de contactos (NC), es decir, personas u hogares que reúnen los requisitos pero con los que no se ha podido entablar contacto durante la recopilación de datos; y otros (Oth), es decir, cuando la entrevista no se realiza por alguna otra razón, como una dificultad idiomática o un problema de salud. Finalmente, hay también casos que no reúnen los requisitos (Inelig) para la encuesta (por ejemplo, las personas de menos de 15 años) o de los que se desconoce si pueden incluirse en la encuesta o no (Unk).

77. La tasa de respuesta en este conjunto simplificado de resultados puede calcularse de varias maneras. Un método comúnmente aceptado de cálculo de la tasa de respuesta (en que "Int" significa el número de entrevistas completas y parciales de una encuesta) es

$$\overline{R} = \frac{\text{Int}}{\text{Int} + \text{Ref} + \text{NC} + \text{Oth} + \varepsilon \times \text{Unk}}$$

Aquí se estima que es admisible cierta proporción, ε, de los casos de admisibilidad desconocida. Con frecuencia esta admisibilidad estimada se calcula a partir de los datos existentes utilizando la tasa de admisibilidad conocida (los casos con resultados Int, Ref, NC y Oth) entre todos los casos cuya admisibilidad se ha determinado. Por consiguiente

$$\hat{\varepsilon} = \frac{\text{Int} + \text{Ref} + \text{NC} + \text{Oth}}{\text{Int} + \text{Ref} + \text{NC} + \text{Oth} + \text{Inelig}}$$

78. Las encuestas de hogares que entrevistan repetidamente a los mismos hogares, o a un panel de personas seleccionadas de entre una muestra de hogares, tienen consideraciones adicionales sobre la falta de respuesta que influyen en el cálculo de las tasas de respuesta. Estas encuestas longitudinales por panel tienen una falta total de respuesta en la oleada inicial de entrevistas lo mismo que en una encuesta transversal, y además es posible que en oleadas posteriores no consigan obtener datos de algunos miembros del panel. Los cálculos sobre la tasa de respuesta deben tener en cuenta las pérdidas debidas a la falta de respuesta tanto en la oleada inicial como en las posteriores de recopilación de datos. Entre los objetivos de la presente publicación no entra examinar el cálculo de las tasas de respuesta en las encuestas por panel. Puede verse más información al respecto en el sitio web de la American Association for Public Opinion Research (www.aapor.org, sección Survey Methods).

79. Las mediciones de las diferencias entre las medias de informantes y no informantes, u otros estadísticos, son más difíciles de obtener. Se pueden comparar los resultados de la encuesta con los de fuentes externas en relación con algunas variables con el fin de determinar si hay una gran diferencia entre la encuesta y la fuente externa en cuanto al valor de una estimación; no obstante, este planteamiento puede ser difícil de aplicar porque puede

haber diferencias en las definiciones y en la metodología entre la encuesta y la fuente externa que compliquen la interpretación de las posibles diferencias observadas. En otras palabras, la diferencia entre las estimaciones de la encuesta y las de la fuente externa pueden atribuirse a causas distintas de la falta de respuesta.

80. La medición de las diferencias entre informantes y no informantes resulta costosa. En principio, con recursos suficientes, algunas veces se supone que pueden obtenerse respuestas de los casos de falta de respuesta. No obstante, son pocas las veces que se dispone de recursos para intentar obtener datos de cada caso de falta de respuesta. Como alternativa, puede establecerse una segunda fase o una doble muestra para los no informantes, y todos los recursos restantes de la encuesta se destinarían a recopilar datos de esa submuestra.

81. Estadísticamente no hay muchas publicaciones sobre el muestreo en dos fases para la falta de respuesta acerca de una serie de características del diseño (véase, por ejemplo, Cochran, 1977, párrafo 13.6). Cuando de la muestra de falta de respuesta en dos fases se obtiene una respuesta completa es posible determinar una fracción óptima de muestreo en la segunda fase —suponiendo escasez de recursos— que reduzca la varianza de muestreo de una estimación en dos fases de la media.

4. Reducción y compensación de la falta total de respuesta en las encuestas de hogares

82. La reducción de la falta total de respuesta se consigue muchas veces a través de métodos especiales que parecen constituir medios sensatos de reducir las tasas de respuesta. Más recientemente, se han formulado teorías generales basadas en principios sociológicos y psicológicos [véase Groves y Couper (1998)] a partir de las cuales pueden deducirse métodos de reducción de la falta de respuesta gracias a una mejor comprensión de cómo funciona la falta de respuesta en las encuestas de hogares. No entra dentro de los objetivos de este capítulo describir estos marcos teóricos más generales. Más bien, lo que interesa aquí es describir varias técnicas cuya eficacia en la reducción de la no respuesta se ha demostrado en algunos estudios experimentales.

83. Las visitas repetidas son un procedimiento estándar en la mayor parte de las encuestas por muestreo. Los entrevistadores no se limitan a intentar una vez el contacto con un hogar o con una persona, sino que vuelven para tratar de conseguir una entrevista completa. El número de visitas repetidas, su calendario y las técnicas del entrevistador para convencer a quienes se resisten a participar son temas que han sido objeto de investigación sobre el terreno. No obstante, no hay una única norma recomendada al respecto. Las diferencias entre los países en cuanto a las tasas de respuesta, la aceptación pública de las encuestas y la movilidad de la población impiden establecer una teoría unificada sobre las visitas repetidas. La aceptación pública de las encuestas sobre diferentes temas dificulta el establecimiento de normas al respecto, incluso en un solo país, cuando se trata de encuestas de diferentes tipos. No obstante, siempre es recomendable utilizar a los mejores entrevistadores para resolver la difícil situación de las negativas a colaborar.

84. No hay pruebas empíricas de que una técnica, ni siquiera las visitas repetidas, consigan tasas de respuesta más altas en las encuestas de hogares. Muchas veces se emplea una combinación de técnicas diversas. Las encuestas de hogares gestionadas por entrevistadores que avisan previamente por teléfono o por carta, la personalización de la correspondencia, la información sobre quién patrocina las encuestas y la explicación gráfica a los posibles informantes sobre la forma como se utilizan los datos son procedimientos de eficacia comprobada que aumentan las tasas de respuesta. El uso de incentivos en las encuestas de los países en desarrollo y en transición es discutible, y en muchos países se desaconseja su utilización, aunque se están generalizando en las encuestas de los países desarrollados [véase Kulka (1995), donde se examinan las investigaciones sobre esta técnica].

85. Las tasas de respuesta pueden mejorarse también prestando atención a la técnica de la entrevista. La capacitación de los entrevistadores para que sepan acomodarse a las diferentes reacciones de los informantes puede mejorar notablemente las tasas de respuesta. Los incentivos pagados a los entrevistadores de acuerdo con la producción supervisada y la calidad de su trabajo cuando supera los objetivos de la encuesta han conseguido también efectos beneficiosos en las tasas de respuesta.

86. Es inevitable que en toda encuesta de hogares haya unidades que no respondan. Muchas veces en el diseño de la encuesta se introducen ajustes en el tamaño cuando se prevé la falta de respuesta de algunas unidades, y se calculan ponderaciones compensatorias como medio de ajuste en la estimación y el análisis.

87. El ajuste del tamaño de la muestra en función de la posible falta de respuesta requiere, antes de la recopilación de datos, una estimación de la tasa prevista de la falta total de respuesta. Esta estimación muchas veces es específica de una encuesta, teniendo en cuenta los datos de la experiencia anterior con la población pertinente, el tema de la encuesta y otros factores. En una encuesta transversal y que se realice una sola vez, la estimación supone muchas veces que se va a repetir la experiencia de otras encuestas. En las encuestas transversales repetidas en las que se entrevista a la misma población a intervalos regulares o irregulares, los datos para estimar las tasas previstas de respuesta pueden obtenerse fácilmente. En las encuestas por panel, en las que las unidades muestrales son objeto de seguimiento a lo largo del tiempo, habrá que prever no sólo la tasa de respuesta total en la primera oleada inicial, sino también la falta de respuesta de desgaste en oleadas posteriores, en las que algunos sujetos que cooperaron anteriormente no pueden ser entrevistados (debido a su negativa, o a la incapacidad de localizarlos o a otros factores).

88. El ajuste del tamaño de la muestra aumenta el tamaño requerido por motivos de costos o de precisión, con el fin de contar con unidades suficientes para obtener el resultado deseado. Por ejemplo, supongamos que se necesita un tamaño final de muestra de 1.000 entrevistas terminadas y que hay una falta de respuesta prevista del 20%. Con el fin de obtener las 1.000 entrevistas completas, se propone una muestra de $1.000/(1–0{,}2) = 1.250$. En la medida en que la tasa de respuesta sea la prevista, el tamaño final de la muestra permitirá conseguir aproximadamente el número total deseado de entrevistas terminadas. A los entrevistadores se les asigna un número de unidades a las que deben entrevistar y se les imparten instrucciones para que obtengan respuestas del mayor número de ellas que sea posible. No se permite ninguna sustitución.

89. Otro planteamiento para resolver el problema de la falta de respuesta es la sustitución. En este caso se deja en manos del entrevistador la decisión de si ponerse o no en contacto con una unidad; en otras palabras, es el criterio subjetivo del entrevistador, y no una selección objetiva de probabilidad, lo que determina cuáles son las unidades de la muestra con las que se establecen contactos. Los métodos de sustitución pueden llevar en este contexto al tamaño exacto de la muestra. No obstante, hay bastantes pruebas [véase, por ejemplo, Stephan y McCarthy (1958), donde se trata un procedimiento conexo no probabilístico: el muestreo por cuotas] de que los métodos de sustitución dan lugar a muestras que no se corresponden con las distribuciones de población conocidas.

90. Los ajustes estadísticos pueden aplicarse a los datos finales de la encuesta con el fin de compensar en parte el posible sesgo por falta de respuesta. El tipo más frecuente de compensación supone la formulación de ponderaciones de ajuste para la falta de respuesta.

91. Para estas ponderaciones es preciso contar con la misma información sobre todos los informantes y todos los no informantes. Como se sabe poco acerca de estos últimos, el tipo de variables disponibles para este tipo de ajustes es limitado en la mayoría de las encuestas de hogares. En la mayor parte de los casos la información primaria conocida acerca de los no declarantes es su ubicación geográfica; es decir, dónde se encontraba el hogar.

92. Por ejemplo, supongamos que en una encuesta de hogares se utiliza un método de muestreo de zona en el que las zonas de empadronamiento censal se seleccionan en la primera etapa. Durante la recopilación de datos no todos los hogares elegidos para la encuesta en una determinada zona de empadronamiento aportan los datos previstos. Un esquema simple de ajuste de la ponderación por la falta de respuesta asignaría mayor ponderación a todos los hogares de la zona de empadronamiento que han respondido, en compensación de los que no lo han hecho. Si respondió el 90% de los hogares de una zona de empadronamiento, las ponderaciones de los hogares que respondieron se multiplicarían por 1/0,9 = 1,11. Si en otra zona respondió el 80%, el resultado se multiplicaría por 1/0,8 = 1,25. Las ponderaciones de todos los hogares que respondieron en la zona de empadronamiento se multiplican por el mismo número. Los hogares que no han respondido se eliminan de la muestra final, lo que significa que cada uno de ellos recibe una ponderación de cero.

93. En algunos casos el ajuste de la ponderación puede basarse en la comparación de los datos administrativos con los datos que aparecen en las respuestas de la encuesta. Por ejemplo, es posible que los datos administrativos se hayan utilizado para seleccionar la muestra. Podrían asignarse ponderaciones a los declarantes de la muestra con el fin de que las distribuciones de los declarantes ponderados en relación con algunas variables clave se correspondan con las distribuciones registradas en los datos administrativos.

94. Los ajustes por falta de respuesta pueden basarse también en un modelo. Cuando se indica simplemente si los hogares de la muestra han respondido o no y se dispone de datos sobre los hogares declarantes y no declarantes, el cálculo de regresión permite determinar la calidad del declarante o del no declarante a partir de las variables disponibles. Luego podrán utilizarse los coeficientes de regresión logística para prever la probabilidad de respuesta de cada hogar. Puede utilizarse la inversa de las probabilidades previstas, en forma semejante al caso anterior, para calcular una ponderación, que es lo que se conoce a veces como ponderación de la propensión a responder. Como las ponderaciones calculadas directamente a partir de las probabilidades previstas suelen ser muy variables, en muchos casos las probabilidades previstas se agrupan en clases, y se asigna una ponderación única a cada clase utilizando el inverso del punto medio, la mediana o la probabilidad prevista media, o la tasa de respuesta ponderada de una clase.

5. Falta de respuesta sobre una partida e imputación

95. Un tema que ha suscitado recientemente más investigaciones es el de la falta de respuesta sobre una partida [véase, por ejemplo, el reciente examen de Groves y otros (2002)]. En este caso hay muchos datos disponibles sobre cada caso de falta de respuesta. Esos datos ofrecen la oportunidad de comprender mejor el fenómeno y una posibilidad de medición, reducción y compensación tomando como base modelos estadísticos más complejos.

96. Por ejemplo, supongamos que el 90% de los declarantes en una encuesta de hogares sobre la salud y la disponibilidad de servicios de salud responden a todas las preguntas, mientras que el 10% responde a todas las preguntas menos a una, relativa a los salarios y a los ingresos salariales del mes anterior. La información disponible del 90% que presentó datos completos puede utilizarse para elaborar modelos estadísticos y comprender la relación entre salud y asistencia de salud y salarios e ingresos salariales. A su vez, esos modelos pueden emplearse para proponer métodos destinados a reducir el nivel de falta de respuesta sobre los salarios y los ingresos salariales para compensar o prever los valores ausentes sobre esa partida.

97. La sustitución de los valores ausentes se conoce con el nombre de imputación, que se utiliza en las encuestas desde hace decenios. [Véase en Kalton y Kasprzyk (1986) y Brick y Kalton (1996) el examen de los procedimientos de imputación utilizados en las encuestas de hogares y de otro tipo.] La imputación es un procedimiento para compensar

los valores de las partidas ausentes. La idea básica es sustituir estos valores por otro que se ha previsto utilizando la información disponible sobre el sujeto (hogar o persona), por ejemplo, o la recibida de otros sujetos de la encuesta.

98. La imputación puede llevarse a cabo mediante un modelo de regresión. En el caso de una variable Y de una encuesta, puede proponerse un modelo para Y que "prevea" Y utilizando un conjunto de p otras variables X_1, ..., X_p de la encuesta. Este modelo podría expresarse así:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi} + \varepsilon_i$$

El modelo se ajusta al conjunto de sujetos con respecto a los cuales constan la variable Y y las variables "predictoras" X_1, ..., X_p. Luego se prevé el valor de Y para los casos ausentes, utilizando los parámetros estimados obtenidos al ajustar el modelo anterior. El valor previsto de la variable Y para la unidad i^a viene dado por:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \cdots + \hat{\beta}_p X_{pi}$$

99. Este modelo de regresión para la imputación se pone en práctica de varias formas. La predicción de la regresión puede incluir un "residual" previsto que deberá agregarse al valor previsto. La técnica de imputación secuencial conocida con el nombre de *hot deck* aplica una forma de imputación de regresión que incorpora eficazmente, como caso que debe imputarse, una cantidad residual "tomada en préstamo" de otro caso del archivo de datos con valores semejantes en X_1, ..., X_p.

100. Los avances recientes en el terreno de la imputación han considerado también el problema resultante del hecho de que la imputación introduce una variabilidad adicional en las estimaciones que utilizan los valores imputados. Esta variabilidad puede explicarse mediante procedimientos de estimación de la varianza, como la estimación de la varianza *jackknife*, o mediante modelos para el proceso de imputación, o con un procedimiento de imputación múltiple en el que la imputación se repite numerosas veces y la variabilidad entre los valores imputados se incluye en la estimación de la varianza.

101. Hay algunas técnicas que pueden utilizarse para reducir el nivel de falta de respuesta sobre una partida. Se puede capacitar a los entrevistadores para que investiguen toda respuesta incompleta o no codificable dada a alguna pregunta del cuestionario. Es un hecho que los encargados del diseño agregan preguntas de seguimiento a algunas partidas que permiten un examen más a fondo cuando se recibe una respuesta como "no sé" o "no deseo responder a esa pregunta". Por ejemplo, las preguntas sobre los ingresos tienen una tasa de falta de respuesta parcial mayor que otras partidas. Las encuestas referentes a los ingresos incorporan a veces una secuencia de preguntas sobre determinadas partidas para "desglosar" una serie de tramos por medio de los cuales se pueden averiguar los ingresos. Si el encuestado no quiere responder o desconoce el total de los ingresos, las preguntas posteriores podrían ser las siguientes: ¿Se trata de un ingreso superior a XXX unidades? ¿... entre YYY unidades y XXX unidades?, etcétera. Estas preguntas permiten establecer una serie de tramos entre los cuales se puedan registrar los ingresos.

102. Las organizaciones que realizan encuestas de hogares deben examinar periódicamente la frecuencia de la falta de respuestas sobre partidas concretas en el conjunto de la encuesta para determinar la importancia del problema. Las tasas de falta de respuesta de este tipo no suelen publicarse, excepto en relación con algunas partidas concretas. El usuario muchas veces tiene que determinar hasta qué punto esa falta de respuesta puede representar un problema para su análisis. En la documentación de las encuestas deberían incluirse las tasas de falta de respuestas sobre partidas concretas, sobre todo cuando éstas son de mayor importancia o cuando concentran altas tasas de falta de respuesta.

Reconocimientos

El autor desea agradecer a Kenneth Coleman, candidato al doctorado de ciencias en el programa de metodología de encuestas de la Universidad de Michigan, su valiosa colaboración en el examen de los métodos de encuesta en América Latina y América del Sur.

REFERENCIAS BIBLIOGRÁFICAS

Brick, J. M. y G. Kalton (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, vol. 5, págs. 215-238.

Cochran, W. G. (1977). *Sampling Techniques.* 3a. ed. Nueva York: John Wiley and Sons.

Groves, R. M. (1989). *Survey Errors and Survey Costs*. Nueva York: John Wiley and Sons.

_____ y M. P. Couper (1998). *Non-response in Household Interview Surveys*. Nueva York: John Wiley and Sons.

Groves, R. M. y otros (2002). *Survey Non-response*. Nueva York: John Wiley and Sons.

Kalton, G. y D. Kasprzyk (1986). The treatment of missing survey data. *Survey Methodology*, vol. 2, págs. 1-16.

Kish, L. e I. Hess (1950). On non-coverage of sample dwellings. *Journal of the American Statistical Association*, vol. 53, págs. 509-524.

Kulka, R. (1995). *The use of incentives to survey "hard-to-reach" respondents: a brief review of empirical research and current research practices.* Seminar on New Directions in Statistical Methodology. Statistical Policy Working Paper, No. 23. Washington, D.C.: U.S. Office of Management and Budget, págs. 256-299.

Lessler, J. y W. Kalsbeek (1992). *Non-sampling Error in Surveys*. Nueva York: John Wiley and Sons.

Lepkowski, James M. (1988). The treatment of wave non-response in panel surveys. En *Panel Survey Design and Analysis*, D. Kasprzyk, G. Duncan and M. P. Singh, comps. Nueva York: Wiley and Sons.

Marks, E. S. (1978). The role of dual system estimation in census evaluation. En *Developments in Dual System Estimation of Population Size and Growth*, K. J. Krotki, ed. Edmonton, Alberta, University of Alberta Press.

Seligson, M. A. y J. Jutkowitz (1994). *Guatemalan Values and the Prospects for Democratic Development*. Arlington, Virginia: Development Associates, Inc.

Capítulo IX
Error de medición en las encuestas de hogares: causas y medición

Daniel Kasprzyk
Mathematica Policy Research
Washington, D.C.
Estados Unidos de América

RESUMEN

En este capítulo se describen las causas fundamentales del error de medición que se encuentra en las encuestas por muestreo y los métodos utilizados habitualmente para cuantificar el error de medición. Se examinan cuatro causas —el cuestionario, el modo de recopilación de datos, el entrevistador y el informante—, y se describe la forma en que se produce el error de medición en las encuestas por muestreo debido a esas causas. Se describen los métodos utilizados para cuantificar el error de medición, como los experimentos aleatorizados, los estudios de investigación cognitiva, los estudios repetidos de medición y los estudios de verificación de registros para ilustrar la aplicación del método.

Términos clave: Error de medición, causas del error de medición, métodos para cuantificar el error de medición.

A. Introducción

1. Los datos de las encuestas de hogares se recopilan con diversos métodos. En el proceso de recopilación de esos datos está implícito el supuesto de que las características y conceptos que se miden pueden definirse con precisión, pueden obtenerse mediante un conjunto de procedimientos bien definidos y tienen valores verdaderos con independencia de la encuesta. El error de medición es, pues, la diferencia entre el valor de una característica ofrecida por el informante y el valor verdadero (pero desconocido) de esa característica. Por lo tanto, el error de medición está relacionado con la observación de la variable a través del proceso de recopilación de datos de la encuesta y, en consecuencia, algunas veces se conoce con el nombre de "error de observación" (Groves, 1989).

2. El presente capítulo está basado en el capítulo sobre el error de medición de un documento de trabajo preparado por un subcomité sobre medición y notificación de la calidad de los datos de encuestas del United States Federal Committee on Statistical Methodology (2001). Por ello, muchas de las referencias y ejemplos son sobre investigaciones realizadas en los Estados Unidos de América y en otros países desarrollados. No obstante, el examen se aplica a todas las encuestas, independientemente de dónde se lleven a cabo. Por eso este capítulo puede ser útil también para quienes realizan encuestas en países en desarrollo y en transición.

3. Son numerosas las publicaciones sobre el error de medición en las encuestas por muestreo [véanse Biemer y otros (1991) y Lyberg y otros (1997)], donde se examinan algu-

nas cuestiones importantes sobre este tema. El error de medición puede dar lugar tanto a sesgos como a errores variables (varianza) en una estimación a lo largo de ensayos repetidos de la encuesta. El sesgo de medición o el sesgo de respuesta se producen como dirección o pauta sistemática en la diferencia entre las respuestas de los informantes a una pregunta y los valores verdaderos. Por ejemplo, los informantes pueden tender a no señalar los ingresos percibidos de un segundo o un tercer empleo, lo que daría lugar a ingresos registrados inferiores a los ingresos efectivos de algunos informantes. La varianza se produce si los valores se comunican de manera diferente cuando se repiten las preguntas sobre las unidades (hogares, personas, entrevistadores y cuestionarios) que son las causas de error. La varianza de respuesta simple refleja la variación aleatoria de la respuesta de un informante a una pregunta de la encuesta cuando la pregunta se repite (es decir, es posible que los informantes den respuestas diferentes a la misma pregunta si ésta se hace varias veces). Los efectos variables que los entrevistadores pueden tener en las respuestas de los informantes pueden ser una causa de error variable, que es lo que se conoce como varianza del entrevistador. Ésta es una forma de varianza de respuesta correlacionada que se produce porque los errores de respuesta están correlacionados para las unidades de la muestra entrevistadas por el mismo entrevistador.

4. En las publicaciones se observan varios planteamientos generales para estudiar el error de medición. Uno de ellos compara las respuestas de la encuesta con otros datos, potencialmente más precisos, de otra procedencia. Los datos podrían ser de una unidad de muestreo individual, como en un "estudio de comprobación de registros". Un ejemplo sencillo: si se preguntara a los informantes cuál era su edad, las respuestas podrían comprobarse examinando los certificados de nacimiento. No obstante, debemos reconocer que incluso en este caso sencillo no se debe dar por descontado que los certificados de nacimiento están por completo libres de error. De todas formas, un método de estudiar el error de medición en una encuesta por muestreo es comparar las respuestas de la encuesta con los datos de otras fuentes independientes y válidas. Un medio alternativo de evaluar el error de medición utilizando datos de otra procedencia es realizar el análisis en forma agregada, es decir, comparar las estimaciones basadas en la encuesta con las estimaciones de población de otra fuente. Un segundo planteamiento consiste en obtener mediciones repetidas de algunas de las unidades. Normalmente, ello significa un programa de nuevas entrevistas y supone la comparación de las respuestas de la entrevista original con las obtenidas en una segunda entrevista realizada poco después de la primera. El tercer planteamiento para estudiar los errores de medición consiste en seleccionar submuestras aleatorias de toda la muestra de la encuesta y aplicar tratamientos diferentes, como preguntas o cuestionarios alternativos o diferentes modos de recopilación de datos. Finalmente, el error de medición se puede evaluar en contextos cualitativos. Entre esos métodos se incluyen los grupos de discusión y un contexto de laboratorio controlado, como un laboratorio de investigación cognitiva.

5. En este capítulo se describen las causas principales de error de medición que se encuentran en las encuestas por muestreo y en sus mediciones. Los procedimientos para cuantificar el error de medición son costosos y muchas veces de difícil aplicación. Por esta razón, y porque constituye una práctica recomendable, los responsables de las encuestas hacen más hincapié en tratar de controlar las fuentes de medición de error mediante una buena planificación y prácticas adecuadas de aplicación de la encuesta. Entre esas prácticas se incluyen la comprobación de los materiales, cuestionarios y procedimientos de la encuesta, la formulación y comprobación de conceptos bien definidos y operacionalmente viables, la insistencia especial en resolver los problemas de recopilación de datos para los subgrupos difíciles de alcanzar, el rigor en la contratación de personal sobre el terreno calificado y la elaboración y ejecución de programas intensivos de capacitación e instrucciones bien especificadas y claramente redactadas para el personal sobre el terreno. El control de los errores no muestrales, y en particular el error de medición, requiere un examen detenido. Véase, por ejemplo, el informe de las Naciones Unidas (1982) en que se incluye una lista de comproba-

ción para controlar los errores no muestrales en las encuestas de hogares. En este capítulo no se trata este tema; se insiste más bien en la descripción de las causas fundamentales de error de medición en las encuestas por muestreo y en las formas habituales de cuantificar el error de medición.

6. Siguiendo a Biemer y otros (1991), se examinaran cuatro causas de error: el cuestionario, el modo de recopilación de datos, el entrevistador y el informante. Una parte significativa del capítulo se dedicará a describir de qué forma se produce el error de medición en las encuestas por muestreo debido a estas fuentes de error. Luego se examinan algunos planteamientos para cuantificar el error de muestreo. Entre ellos se incluyen los experimentos aleatorizados, los estudios de investigación cognitiva, los estudios de medición repetida y los estudios de comprobación de registros. Para cuantificar el error de medición se requiere siempre la adopción de algunas medidas adicionales antes, durante y después de la realización de la entrevista. Entre los inconvenientes frecuentemente citados para iniciar estudios que cuantifiquen las causas específicas de error de medición se encuentran el tiempo y los gastos necesarios para realizar el estudio. No obstante, los estudios del error de medición son sumamente valiosos tanto para cuantificar el nivel de error en la encuesta vigente como para indicar dónde podrían introducirse mejoras para las encuestas futuras. Estos estudios son particularmente útiles en los programas de encuestas repetidas.

B. Causas del error de medición

7. En Biemer y otros (1991) se identifican cuatro causas fundamentales del error de medición:

- *Cuestionario:* Efecto del diseño del cuestionario, su diseño gráfico, temas incluidos y formulación de las preguntas.
- *Método de recopilación de datos:* Efecto de la forma de administración del cuestionario al informante (por ejemplo, por correo, en persona o en forma de diario). Es posible que los informantes respondan de distinta manera según que estén en presencia de un entrevistador, se encuentren solos o utilicen un diario.
- *Entrevistador:* Efecto que el entrevistador tiene en la respuesta a una pregunta. El entrevistador puede inducir a error en las respuestas de la encuesta si no lee las partidas en la forma prevista, si no hace las debidas averiguaciones cuando se encuentra con una respuesta inadecuada o si añade otra información que puede confundir o desorientar al informante.
- *Informante:* Efecto debido al hecho de que los informantes con experiencias, conocimientos y aptitudes diferentes pueden interpretar el significado de las partidas de un cuestionario de distintas maneras.

8. Estas cuatro causas son fundamentales en la realización de una encuesta por muestreo. El cuestionario es el método utilizado para solicitar formalmente información al informante encuestado. La forma de recopilación de datos representa la manera en que el cuestionario se aplica o presenta (autoadministrado o en persona). El entrevistador, en el caso de que el cuestionario se aplique en persona, es quien se encarga de presentar el cuestionario. El informante es quien recibe la solicitud de información. Cada uno de ellos puede introducir errores en el proceso de medición. La mayoría de las encuestas consideran estas causas por separado, en el supuesto de que lleguen a tenerlas en cuenta. No obstante, las causas pueden interrelacionarse mutuamente; por ejemplo, las características de los entrevistadores y las de los informantes pueden interrelacionarse e introducir errores que no se observan en cada una de las fuentes por separado. A continuación se examinan las formas en que puede presentarse el error de medición en el contexto de estas cuatro causas.

1. Efectos del cuestionario

9. El cuestionario es el instrumento de recopilación de datos para obtener información del informante. Durante los 20 últimos años, los principios básicos del diseño de cuestionarios, considerados antes más un arte que una ciencia, han sido objeto de detenidos estudios (Sirken y otros, 1999; Schwarz, 1997; Sudman, Bradburn, y Schwarz, 1996; Bradburn y Sudman, 1991). El cuestionario o las características del cuestionario, es decir, la forma en que se formulan las preguntas o el formato del mismo, pueden influir en la forma en que responden los individuos a la encuesta. En esta sección describimos las formas en que el cuestionario puede introducir errores en el proceso de recopilación de datos.

Problemas de especificación

10. Al planificar una encuesta muchas veces surgen problemas porque los objetivos y conceptos de la investigación y la información recopilada en el cuestionario son ambiguos, no están bien definidos o presentan contradicciones. Las preguntas del cuestionario, tal como están formuladas, quizá no permitan recabar la información deseada para alcanzar los objetivos de la investigación. Los problemas de especificación de los datos pueden surgir debido a que los cuestionarios y las instrucciones de la encuesta están mal formulados, a que las definiciones son ambiguas o a que el concepto deseado es difícil de medir. Por ejemplo, una encuesta podría preguntar sobre "los servicios de maternidad recibidos durante el embarazo" pero no especificar a qué embarazo o a qué período de tiempo se refiere la pregunta. Puede haber ambigüedad en preguntas tan básicas como "¿Cuántos empleos tiene?" si no se especifica debidamente la naturaleza del empleo: temporal o permanente, de plena dedicación o a tiempo parcial. Los conceptos analíticos combinados, como el ingreso total de una persona, quizá no se comuniquen en su integridad si no se identifican y definen claramente al informante cuáles son los componentes individuales del ingreso.

Formulación de las preguntas

11. Las preguntas del cuestionario deben estar formuladas con precisión y claridad para que el informante las interprete en la forma deseada por quienes han diseñado la encuesta. Como el cuestionario es una forma de comunicación entre el recopilador de datos y el informante, puede haber posibles causas de error. En primer lugar, quien diseñó el cuestionario quizá no haya formulado claramente el concepto que trata de medir. En segundo lugar, aun cuando el concepto esté claramente formulado, quizá no se presente debidamente en la pregunta o conjunto de preguntas; incluso si el concepto es claro y está fielmente representado, la interpretación del informante quizá no sea la deseada por quien diseñó el cuestionario. Las diferencias lingüísticas y culturales o en cuanto a experiencias y contexto entre quien diseñó el cuestionario y el informante pueden contribuir a una interpretación errónea de las preguntas. Estas diferencias pueden ser particularmente importantes en los países en desarrollo y en transición que tienen grupos étnicos diferentes. En Vaessen y otros (1987) se examinan los problemas lingüísticos en los países multilingües.

12. Hay al menos dos niveles en la comprensión de una pregunta formulada en una encuesta por muestreo. El primer nivel es el de la simple comprensión del significado literal de la pregunta. ¿Está el informante familiarizado con las palabras incluidas en la pregunta? ¿Dispone de información que corresponda a su comprensión de esas palabras y permita ofrecer una respuesta significativa? No obstante, para responder a una pregunta, el informante debe inferir también el objetivo del cuestionario; es decir, para responder a la pregunta, el informante debe determinar el sentido pragmático de la pregunta (Schwarz, Groves y Schuman, 1995). Este segundo elemento es el que hace que la formulación de las preguntas sea una tarea más difícil y más compleja que la de simplemente preparar partidas que requieran un bajo nivel de comprensión. Para conseguir un instrumento bien diseñado es necesaria la

aportación de los informantes, es decir, su interpretación y forma de entender las preguntas. Los métodos de investigación cognitiva ofrecen medios útiles de conseguir esa información (véase el apartado C.2).

Longitud de las preguntas

13. El sentido común y la práctica de una buena redacción recomiendan el uso de preguntas breves y sencillas para conseguir una interpretación más clara. No obstante, las investigaciones revelan que las preguntas más largas pueden recibir de los informantes respuestas más detalladas y precisas que las preguntas breves, al menos en lo que respecta al comportamiento relacionado con los síntomas y las visitas médicas (Marquis y Cannell, 1971) y el consumo de alcohol y drogas (Bradburn, Sudman y asociados, 1979). Las preguntas más largas pueden facilitar más información o claves para ayudar al informante a recordar, y pueden ofrecerle más tiempo para pensar sobre la información que se le solicita.

Longitud del cuestionario

14. Los investigadores y analistas desean siempre formular la mayor cantidad posible de preguntas, mientras que los especialistas en metodología de las encuestas reconocen que los cuestionarios demasiado largos pueden provocar errores. Un informante puede perder concentración o cansarse, según las características personales del interesado (edad, o estado de salud, por ejemplo), la importancia del tema, la relación con el entrevistador, el diseño del cuestionario y la forma de la entrevista.

Orden de las preguntas

15. Los investigadores han observado que el orden de las preguntas influye en la respuesta (Schuman y Presser, 1981), sobre todo en las encuestas sobre aptitudes y opiniones. Se han observado fenómenos de asimilación —cuando las preguntas sucesivas se orientan en la misma dirección que las precedentes— y de contraste, cuando las respuestas sucesivas se orientan en dirección contraria a las precedentes. Los informantes pueden utilizar también la información derivada de las partidas anteriores acerca del significado de los términos para responder mejor a preguntas posteriores.

Categorías de respuesta

16. Las categorías de respuesta a la pregunta pueden influir en las respuestas cuando dan a entender al informante qué es importante para quien ha formulado el cuestionario. El informante deduce que las categorías incluidas con una partida son consideradas como las más importantes para el autor del cuestionario. Ello puede dar lugar a confusión acerca del objetivo de la pregunta, si las categorías de respuesta no parecen adecuadas al informante. El orden de las categorías puede influir también en las respuestas. Es posible que los informantes se confíen durante una entrevista y elijan sistemáticamente el mismo punto de una escala de respuestas, respondan a las primeras opciones más que a las últimas o elijan las últimas respuestas ofrecidas.

17. El efecto producido por el orden de las categorías de respuesta puede estar también influido por la manera en que se lleva a cabo la entrevista. Si las partidas son autoadministradas, es más probable que se recuerden y acepten las categorías de respuesta que aparecen en primer lugar en la lista (efecto de primacía), ya que el informante tiene más tiempo para procesarlas. Si las partidas son administradas por el entrevistador, es más probable que se recuerden las categorías que aparecen en los últimos lugares (efecto de cercanía o proximidad).

Formatos abiertos y cerrados

18. Un formato de pregunta en el que se presente a los informantes un conjunto especificado de opciones de respuesta (formato cerrado) puede arrojar respuestas diferentes que si no se les dan esas opciones (formato abierto) (Bishop y otros, 1988). Es menos probable que se registre una pregunta dada en un formato abierto que cuando se incluye como opción en un formato cerrado (Bradburn, 1983). El formato cerrado puede recordar a los informantes que deben incluir algo que de lo contrario no recordarían. Las opciones de respuesta pueden indicar a los informantes el nivel o tipo de respuestas que se consideran adecuadas [véase Schwarz, Groves y Schuman (1995) y Schwarz y Hippler (1991)].

Formato del cuestionario

19. El "aspecto" concreto de un cuestionario autoadministrado, es decir, su formato y tipografía, pueden favorecer o dificultar la precisión de las respuestas. El hecho de que los informantes puedan desorientarse con un diseño gráfico inadecuado podría dar lugar a que no se comprenda bien cómo se pasa de una pregunta a otra, o a una interpretación inadecuada de las preguntas y de las instrucciones. En Jenkins y Dillman (1997) se recogen algunos principios para diseñar cuestionarios autoadministrados para la población de los Estados Unidos. Debería procederse con cautela al transferir estos principios a otro país, ya que habría que considerar antes los factores culturales y lingüísticos específicos de éste.

2. Efectos del método de recopilación de datos

20. Para identificar la forma más apropiada de recopilación de datos hay que tomar decisiones sobre varias cuestiones relativas a los métodos de la encuesta. Los recursos financieros influyen muchas veces de manera significativa en esta decisión. No obstante, el contenido del cuestionario, la población objetivo, las tasas de respuesta deseadas, la longitud del período de recopilación de datos y el error de medición previsto son consideraciones todas ellas importantes a la hora de decidir el modo de recopilación de datos más indicado. Si bien los avances de la tecnología han permitido una mayor utilización del teléfono con ese fin, hay otras modalidades que ofrecen opciones muy diversas para la realización de una encuesta. En Lyberg y Kasprzyk (1991) puede verse una descripción general de los diferentes métodos de recopilación de datos junto con las causas de error de medición de los mismos. A continuación se presenta un resumen de su exposición.

Entrevistas personales

21. La entrevista personal es el método principal de recopilación de datos en los países en desarrollo y en transición. En la mayor parte de los casos, un entrevistador aplica un cuestionario estructurado a los informantes e introduce las respuestas de éstos en el cuestionario escrito. El uso de este método de entrevista personal con papel y lápiz tiene una larga historia. Los avances recientes en la producción de computadoras portátiles ligeras ha hecho posible la entrevista personal con ayuda de computadora (EPAC). Los entrevistadores visitan la casa de los informantes y realizan entrevistas utilizando computadoras portátiles en vez de cuestionarios de papel. Véanse en Couper y otros (1998) algunas de las cuestiones relacionadas con la EPAC. La ventaja más obvia de esta metodología es la relativa al control de calidad y la reducción del error de respuesta. Los entrevistadores introducen las respuestas en un archivo electrónico. El programa informático utilizado garantiza que se sigan correctamente las pautas de salto de unas preguntas a otras y que sea posible introducir las respuestas y corregir los errores en el momento de la entrevista. En consecuencia, se ahorra tiempo y recursos en la etapa de depuración de datos de la encuesta.

22. Con las entrevistas personales es posible realizar entrevistas complejas, se pueden utilizar ayudas visuales para ayudar al informante a responder las preguntas y los entrevistadores capacitados y hábiles pueden establecer una relación que les permita llegar a respuestas más completas y precisas. No obstante, los entrevistadores pueden influir en las respuestas de los informantes, con lo que se produciría un sesgo en las estimaciones o un efecto de varianza del entrevistador, como se examina en el apartado C.3. Los entrevistadores pueden influir en la respuesta mediante una combinación de rasgos de personalidad y de comportamiento. Un motivo especial de preocupación es el relacionado con las características o actos que no están bien vistos socialmente. Es posible que los informantes se resistan a comunicar esas características o actos al entrevistador. DeMaio (1984) observa que el factor de la aceptabilidad social parece encerrar dos elementos: la idea de que hay cosas buenas y cosas malas, y el hecho de que los informantes quieren aparecer como "buenos" y responden a las preguntas para dar esa impresión.

23. Otra posible causa de error de medición relacionada con las entrevistas personales en las encuestas de hogares es la posible presencia de otros miembros del hogar durante la entrevista, lo que podría condicionar las respuestas del informante, sobre todo cuando se trata de cuestiones delicadas. Por ejemplo, es posible que al informante le cueste responder verazmente a las preguntas relacionadas con el uso de drogas ilegales cuando está presente otro miembro del hogar. Incluso preguntas aparentemente inocuas pueden considerarse como delicadas en presencia de otro miembro del hogar (por ejemplo, las preguntas relacionadas con la vida conyugal o la fecundidad).

Encuestas autoaplicadas

24. Las causas de error de medición en los cuestionarios de las encuestas autoaplicadas son diferentes que las de las entrevistas directas. Obviamente, en las entrevistas autoaplicadas no hay ningún efecto del entrevistador y existe menor riesgo de los efectos de la "aceptabilidad social". Ofrecen también un medio de formular preguntas sobre temas delicados o embarazosos sin poner en situación de compromiso al informante. Otra ventaja es que, si es necesario, se pueden administrar simultáneamente a más de un informante en un hogar (Dillman, 1983). Por otro lado, tienen el riesgo de sesgo sistemático si la población objetivo está formada por personas con poca o ninguna instrucción o con dificultades para leer y escribir. Este sesgo puede observarse en las respuestas a las preguntas "abiertas", que quizá sean menos exhaustivas y detalladas que las obtenidas en las encuestas realizadas por entrevistadores. Este método de recopilación de datos quizá no sea el más aconsejable en países con bajas tasas de alfabetización; por otro lado, aun cuando la población objetivo tenga un nivel de instrucción relativamente alto, los informantes pueden interpretar erróneamente las preguntas e instrucciones. En general, las tasas de respuesta sobre temas concretos son menores que en las encuestas autoaplicadas, pero cuando se consiguen respuestas los datos suelen ser de calidad superior. Las encuestas autoaplicadas, quizá más que otras formas de recopilación de datos, se benefician especialmente cuando el cuestionario está bien diseñado y el formato y redacción de las preguntas es claro. Un tema específico de este tipo de encuestas es la encuesta por correo autoaplicada, en la que se pide a los informantes que cumplimenten ellos mismos el cuestionario, cuya entrega y recuperación se realiza por correo (Dillman, 1978; 1991; 2000).

Encuestas a través de diarios

25. Las encuestas en forma de diario son formas autoadministradas que se utilizan en el caso de temas que requieren una información detallada del comportamiento a lo largo de un período de tiempo (por ejemplo, los gastos, el uso del tiempo y el tiempo dedicado a la televisión). Para reducir o evitar los errores de recuerdo se aconseja al informante que

utilice un diario y registre las respuestas sobre un hecho o tema poco después de que se haya producido el evento. El éxito de este sistema depende de que el informante intervenga activa y comprometidamente en el registro de la información y cumplimente luego un formulario que requiere un trabajo bastante tedioso. Este mecanismo supone también que la población objetivo debe ser capaz de leer e interpretar las preguntas del diario, condición que no se cumple en los países con bajas tasas de alfabetización. El procedimiento de recopilación de datos normalmente requiere que los entrevistadores se pongan en contacto con el informante para entregarle el diario, conseguir su cooperación y explicarle detalladamente los procedimientos de registro de datos. El entrevistador regresa después de un tiempo determinado previamente para recoger el diario y, si éste no se ha terminado, para ayudar al informante a finalizarlo.

26. En Lyberg y Kasprzyk (1991) se identifican varias causas de error de medición en esta modalidad. Por ejemplo, es posible que los informantes poco responsables no registren algunos acontecimientos, aun cuando sean recientes. El diario mismo, por su diseño y formato y por la complejidad de las preguntas, puede presentar al informante dificultades prácticas significativas. Además, es posible que los informantes cambien su comportamiento como consecuencia de la utilización del diario; por ejemplo, el hecho de tener que enumerar las compras en un diario de gastos puede hacer que el informante cambie sus costumbres de compra. Puede encontrarse un análisis sobre los errores de medición en las encuestas de gastos y, en particular, en lo que se refiere al componente de los diarios, en Neter (1970) y Kantorowitz (1992). En Silberstein y Scott (1991) se comparan los datos procedentes de entrevistas directas y entrevistas a través de diarios.

Observación directa

27. La observación directa, en cuanto método de recopilación de datos, significa que el entrevistador debe recopilar datos utilizando sus propios sentidos físicos (vista, oído), su perspicacia (tacto y capacidad de comprobación) o los dispositivos de medición. Este método se utiliza en muchas áreas; por ejemplo, en las encuestas agrícolas, para estimar los rendimientos de las cosechas ("estimación ocular"), y en las encuestas de hogares, para determinar la calidad de la vivienda de los informantes. Los observadores introducen errores de medición de la misma manera que sucede con los entrevistadores; por ejemplo, es posible que los observadores no comprendan bien los conceptos ni la información que deben registrar o que cambien sus pautas de registro de la información con el tiempo por exceso de confianza o fatiga.

3. Efectos del entrevistador

28. El entrevistador desempeña un papel fundamental en muchas encuestas por muestreo. Como parte fundamental del proceso de recopilación de datos, su comportamiento puede influir en la calidad de los datos de la encuesta. No obstante, el entrevistador es sólo uno de los componentes del proceso de recopilación cuyo comportamiento puede tratar de controlar el responsable de la investigación o de administración de la encuesta. Por consiguiente, se han establecido estrategias —mediante selección y contratación, capacitación y seguimiento del desempeño laboral— para reducir el error asociado con el papel del entrevistador (Fowler, 1991). Debido a diferencias individuales, cada entrevistador resuelve las situaciones de manera distinta; algunos, por ejemplo, quizá no formulen las preguntas exactamente tal como están redactadas, no sigan las normas de salto de preguntas o no traten de profundizar en las respuestas de la forma adecuada. A veces no siguen las indicaciones al pie de la letra, bien por decisión deliberada o porque las indicaciones no son claras. Sin ser conscientes de ello, los entrevistadores pueden modificar la inflexión o el tono de su voz o mostrar otros cambios en sus modales.

29. El entrevistador puede introducir errores, sea por exceso o por defecto. Cuando ambos tipos de error se compensan entre todos los entrevistadores, el sesgo resultante global será pequeño. No obstante, los errores de los distintos entrevistadores pueden ser considerables y en la misma dirección, con lo que se produciría un sesgo notable. La variación en los sesgos de los distintos entrevistadores da lugar a lo que se conoce como varianza del entrevistador, que puede tener graves efectos en la precisión de las estimaciones de la encuesta.

Varianza correlacionada del entrevistador

30. En los primeros años sesenta, Kish (1962) formuló un planteamiento basado en el coeficiente de correlación entre entrevistadores, denotado por ρ, para evaluar el efecto de la varianza del entrevistador en las estimaciones de la encuesta. El valor de ρ, que se define como coeficiente entre el componente de la varianza del entrevistador y la varianza total de la variable de una encuesta, se estima mediante análisis simple de la varianza.

31. En las encuestas directas bien aplicadas, ρ suele ser de aproximadamente 0,02 en la mayor parte de las variables. Aunque ρ es de pequeña magnitud, el efecto en la precisión de la estimación puede ser grande. La varianza de la media de la muestra se multiplica por $1 + \rho (n-1)$, donde n es el tamaño del volumen de trabajo del entrevistador medio. Una ρ de 0,02 con un volumen de trabajo de 10 entrevistas aumenta la varianza un 18%, y una carga de trabajo de 25 da lugar a una varianza un 48% más alta. Así pues, incluso valores pequeños de ρ pueden reducir significativamente la previsión de las estadísticas de la encuesta. Por motivos prácticos y económicos, los entrevistadores suelen tener una gran carga de trabajo, de ahí que un entrevistador que aporte un sesgo sistemático influirá en los resultados obtenidos de un número considerable de encuestados, y el efecto sobre la varianza puede ser notable.

Características del entrevistador

32. Las investigaciones no ayudan demasiado a identificar las características indicativas del buen entrevistador. En el Reino Unido de Gran Bretaña e Irlanda del Norte, Collins (1980) no encontró ninguna base para recomendar que la contratación de entrevistadores se concentrara más en las mujeres que en los hombres, o en personas de la clase media, o en personas de mediana edad, más que en los jóvenes. Weiss (1968), estudiando una muestra de madres de Nueva York acogidas al sistema de beneficencia, validó la precisión de varias partidas y comprobó que la semejanza entre el entrevistador y el informante con respecto a la edad, la educación y la condición socioeconómica no se manifestaba en una mejora de los informes. En Sudman y otros (1977) se estudiaron las expectativas de los entrevistadores acerca de la dificultad de obtener información delicada y se observaron los escasos efectos en cuanto a la relación entre dificultades previstas y reales de las entrevistas. Groves (1989) examinó varios estudios y llegó a la conclusión de que, en general, pueden producirse efectos demográficos cuando las mediciones están relacionadas con las características demográficas, pero no en otros casos; por ejemplo, puede haber un efecto basado en la raza del entrevistador si las preguntas están relacionadas con ese tema.

Métodos para controlar los errores del entrevistador

33. Hasta cierto punto, el administrador de la encuesta puede controlar los errores del entrevistador por medio de capacitación, supervisión o seguimiento y organizando su carga de trabajo. Un programa de capacitación de duración suficiente para incluir los conocimientos y técnicas sobre cómo entrevistar y para facilitar información sobre la encuesta específica de que se trate ayuda a introducir cierto grado de estandarización en el proceso de las entrevistas (Fowler, 1991). Muchos creen que esa estandarización reduce los efectos achacados a la función del entrevistador.

34. La supervisión y el seguimiento de los resultados, cuyo objetivo es controlar el desempeño mediante la observación y estadísticas de los resultados e identificar las preguntas problemáticas, constituyen otro componente del sistema de control de calidad de los entrevistadores. Se llevan a cabo programas de repetición de las entrevistas y de observaciones sobre el terreno para evaluar el desempeño de determinados entrevistadores. Las observaciones sobre el terreno se realizan utilizando amplias listas de codificación o guías detalladas para los observadores, y el supervisor comprueba si los procedimientos se siguen adecuadamente. Por ejemplo, podrían observarse el aspecto y la conducta del entrevistador, la forma como se presenta a sí mismo e introduce la encuesta, la manera como se formulan las preguntas y se registran las respuestas, el uso de tarjetas indicativas y sondeos neutrales, y la utilización adecuada del manual de los entrevistadores. En otros casos pueden hacerse grabaciones (audiovisuales o de audio) y luego codificar y analizar el comportamiento del entrevistador (Lyberg y Kasprzyk, 1991).

35. Otra manera de reducir el efecto de la varianza del entrevistador es reducir el volumen de trabajo medio; no obstante, ello supone que se puede contar con entrevistadores adicionales de la misma calidad. En Groves y Magilavy (1986) se examina cuál es el volumen óptimo de trabajo de los entrevistadores en función de los costos de contratación y de capacitación del entrevistador y la magnitud de la correlación entre entrevistadores. Como esta correlación varía entre los estadísticos de la misma encuesta, es muy difícil establecer qué es lo que constituye un volumen de trabajo óptimo.

36. Los efectos del entrevistador pueden reducirse evitando problemas de diseño en el cuestionario, dando instrucciones y definiciones claras e inequívocas, capacitando a los entrevistadores para que sigan las instrucciones y reduciendo la dependencia de la variable "habilidad de los entrevistadores para conseguir respuestas".

4. Efectos del informante

37. Los informantes pueden contribuir al error de medición si no ofrecen respuestas precisas. Groves (1989) señala los modelos tradicionales del proceso de entrevista (Kahn y Cannell, 1957) y las perspectivas de la ciencia cognitiva sobre la respuesta en las encuestas. Hastie y Carlston (1980) identifican cinco etapas secuenciales en la formación y presentación de respuestas por los informantes:

- *Codificación de la información,* que supone el proceso de acumular recuerdos o mantener conocimientos.
- *Comprensión de la pregunta de la encuesta*, que agrupa el conocimiento de las palabras y frases del cuestionario y la impresión del informante acerca del objetivo de la encuesta, el contexto y la forma de la pregunta y el comportamiento del entrevistador al preguntar.
- *Recuperación de la información de la memoria*, que supone el intento por parte del informante de buscar la información pertinente en su memoria.
- *Juicio sobre la respuesta adecuada*, que supone la elección por el informante de las respuestas alternativas a una pregunta tomando como base la información recuperada.
- *Comunicación de la respuesta*, en que se tienen en cuenta tanto las influencias en la presentación de informaciones adecuadas después de que el informante haya recuperado la información pertinente como la capacidad del informante para responder.

38. Son muchos los aspectos del proceso de la encuesta que afectan a la calidad de las respuestas del informante que emanan de este proceso en cinco etapas. A continuación se presentan algunos ejemplos de los factores que influyen en los efectos del informante.

Normas sobre el informante

39. Las normas sobre el informante que definen los criterios utilizados para determinar la persona o personas que pueden responder al cuestionario desempeñan un papel importante en el proceso de respuesta. Si una encuesta recopila información sobre los hogares, el conocimiento de las respuestas a las preguntas puede variar según los diferentes informantes del hogar. Las encuestas que recopilan información sobre los individuos dentro de los hogares incluidos en la muestra pueden recurrir a la autoinformación o a la información mediante sustitutos. Las diferencias entre una y otra varían de acuerdo con la materia (por ejemplo, la autoinformación es mejor en las encuestas sobre actitudes). En Naciones Unidas (1982) se describe el resultado de una prueba piloto de los efectos de la respuesta por sustitutos en las partidas demográficas de la encuesta demográfica turca. En Blair, Menon, y Bickart (1991) se hace un examen de las investigaciones sobre las diferencias entre autoinformación e información mediante sustitutos.

Preguntas

40. La formulación y complejidad de la pregunta y el diseño del cuestionario pueden influir en que el informante entienda o no la pregunta y en cómo la entiende (véase información más detallada en el apartado B.1). La disponibilidad del informante a ofrecer las respuestas correctas está influida por el tipo de preguntas formuladas, por la dificultad de determinar las respuestas y por la opinión del informante acerca de la aceptabilidad social de las respuestas.

Entrevistadores

41. Los indicios visuales (por ejemplo, edad, género, forma de vestir, expresión facial) y auditivos (por ejemplo, tono de voz, ritmo, inflexión) del entrevistador pueden influir en la comprensión de la pregunta por el informante.

Período de recuerdo

42. En general, el tiempo reduce la capacidad de recordar los hechos y los datos. La memoria se pierde y el resultado es que los informantes tienen más dificultad para recordar una actividad cuando ha transcurrido un largo período de tiempo entre el acontecimiento y la encuesta. Por ejemplo, en algunos países incluidos en el Estudio Mundial sobre la Fecundidad es probable que las fechas de los partos más recientes se recuerden con más precisión que las de los que tuvieron lugar muchos años antes (Singh, 1987). Los encargados del diseño de la encuesta pueden tratar de establecer períodos de recuerdo que reduzcan el error cuadrático medio total en lo que respecta al error de muestreo y a los posibles sesgos; por ejemplo, Huang (1993) comprobó que la mayor precisión conseguida aumentando el tamaño de la muestra y pasando de un período de referencia de cuatro meses a otro de seis meses no compensaría el alto sesgo debido a los problemas de recuerdo. Eisenhower, Mathiowetz y Morganstein (1991) examinan el uso de las ayudas nemotécnicas (por ejemplo, calendarios, mapas, diarios) para reducir el sesgo de recuerdo. Mathiowetz (2000) señala los resultados de un metaanálisis en el que se comprueba la hipótesis de que la calidad de los informes retrospectivos está en función de lo dilatado del período de recuerdo.

Efecto telescopio

43. Se produce un efecto telescopio cuando los informantes señalan un acontecimiento como si hubiera ocurrido dentro del período de referencia cuando en realidad tuvo lugar fuera de él. Las técnicas de delimitación (por ejemplo, realización de una entrevista

inicial con el único fin de establecer una fecha de referencia, o el uso de una fecha o acontecimiento significativo como referencia del período de referencia) pueden permitir reducir este tipo de efecto (Neter y Waksberg, 1964).

Encuestas por panel/longitudinales

44. Otros factores relacionados con el informante contribuyen al error en las encuestas por panel o longitudinales. En primer lugar, puede obtenerse una impresión errónea de cambio cuando un informante da distintas respuesta a preguntas idénticas o semejantes en dos puntos diferentes y las respuestas se deben a la variación aleatoria de la respuesta a las mismas preguntas más que a un cambio real. En Kalton, McMillen y Kasprzyk (1986) pueden verse ejemplos de error de medición en sucesivas oleadas de una encuesta longitudinal. Se citan la edad, la raza, el sexo y la ocupación y la rama de actividad como variables en las que se observó un error de medición en la Encuesta de los Estados Unidos sobre los ingresos y la participación en programas. En el perfil de calidad de dicha encuesta se examinan ésta y otras cuestiones sobre el error de medición identificadas en la encuesta (Oficina del Censo de los Estados Unidos, 1998). Las técnicas de entrevistas dependientes, en las que se utilizan para la entrevista en curso las respuestas de una entrevista anterior, pueden reducir la incidencia de cambios injustificados. Hill (1994) comprobó que las entrevistas dependientes habían conseguido una mejora neta en las medidas de cambio en la ocupación y en el sector de empleo pero que posiblemente se omitan algunos casos de cambio verdadero, lo que significa que debe utilizarse en forma selectiva. En Mathiowetz y McGonagle (2000) se examinan las prácticas actuales dentro de un entorno de entrevistas con ayuda de computadora y pruebas empíricas sobre los efectos de las entrevistas dependientes en la calidad de los datos.

45. El sesgo de condicionamiento de panel o de "tiempo de inclusión en la muestra" es otra fuente posible de error en las encuestas por panel. Por condicionamiento se entiende el cambio de respuesta que se produce cuando un informante ya ha tenido una o más entrevistas. Woltman y Bushery (1977) investigaron el sesgo del tiempo de inclusión en la muestra en el Estudio nacional sobre las víctimas de delitos hecho en los Estados Unidos, en el que se comparan los informes sobre las víctimas de delitos presentados por individuos con diferente cantidad de entrevistas previas, que habían sido entrevistados el mismo mes, y comprobaron que en general aparecían tasas decrecientes de incidencia de delitos según aumentaba la cantidad de entrevistas anteriores. Kalton, Kasprzyk y McMillen (1989) examinan también esta causa de error.

C. Planteamientos para cuantificar el error de medición

46. Existen varios planteamientos generales para cuantificar el error de medición. Con el fin de estudiar los sesgos de medición pueden administrarse diferentes tratamientos, como preguntas o cuestionarios alternativos o una forma diferente de recopilación de datos, a submuestras de la muestra total de la encuesta seleccionadas aleatoriamente. El error de medición puede estudiarse en contextos cualitativos, como los grupos de discusión o los laboratorios de investigación cognitiva. Otro planteamiento puede ser la repetición de mediciones de la unidad muestral, como las realizadas en un programa de reentrevistas de una encuesta. Finalmente, hay "estudios de comprobación de registros", que comparan las respuestas de la encuesta con datos más precisos de otra procedencia para estimar el error de medición. Estos planteamientos se examinan a continuación.

1. Experimentos con aleatoriedad

47. Un método frecuentemente utilizado para estimar los errores de medición es el experimento aleatorio (*randomized experiments*), que ha recibido distintos nombres: *interpenetrated samples, split-sample experiments, split-panel experiments, random half-sample*

experiments y *split-ballot experiments*, que consiste en aplicar a submuestras aleatorias del mismo diseño tratamientos diferentes relacionados con el error específico que se desea medir. Para estudiar los errores se incluyen y comparan muchas entidades diferentes que se consideran como explicación del error (por ejemplo, muchos entrevistadores diferentes en las estimaciones de la varianza del entrevistador). Para estudiar los sesgos se comparan dos o tres tratamientos (por ejemplo, dos formas diferentes de recopilación de datos) y se considera como preferente uno de los métodos. Las comprobaciones sobre el terreno realizadas antes de la encuesta incluyen con frecuencia experimentos aleatorios para evaluar métodos, procedimientos y cuestionarios alternativos.

48. Por ejemplo, puede utilizarse un experimento aleatorio para comprobar el efecto de la longitud del cuestionario. Se asignan unidades muestrales de manera aleatoria a uno de los dos grupos utilizados, uno de los cuales recibiría una versión "breve" de las preguntas y el otro la versión "larga". Suponiendo que se dispone de una fuente de datos independiente, las respuestas para cada grupo pueden compararse luego con las estimaciones de la fuente de datos, que se supone que es exacta y fiable. De la misma manera, pueden evaluarse los efectos del orden de las preguntas invirtiendo el orden de la pregunta fijado en un cuestionario alternativo administrado a muestras aleatorias. Este método se utilizó para una encuesta realizada en la República Dominicana dentro del programa mundial de encuestas demográficas y de salud; el cuestionario básico se utilizó para dos tercios de la muestra y el cuestionario experimental para el tercio restante. El objetivo era determinar las diferencias de respuesta resultantes de la administración de dos conjuntos de preguntas (Westoff, Goldman y Moreno, 1990).

2. Métodos de investigación cognitiva

49. Durante los últimos 20 años el uso de métodos de investigación cognitiva para la reducción del error de medición ha crecido rápidamente. Estos métodos se utilizaron inicialmente para comprender mejor los procesos mentales de los informantes, pero se usan cada vez más como complemento de las comprobaciones sobre el terreno de tipo tradicional (Schwarz y Sudman, 1996; Sudman, Bradburn y Schwarz, 1996). Los informantes dicen al encargado de diseñar el cuestionario cómo interpretan sus distintas partidas. Éste es un procedimiento costoso y que requiere gran cantidad de mano de obra por informante, y por ello la comprobación cognitiva se realiza con muestras pequeñas. Una deficiencia de las entrevistas cognitivas es que se realizan con muestras pequeñas no aleatorias. El encargado de diseñar el cuestionario debe tener en cuenta que las conclusiones revelan los problemas que podrían presentarse pero que no son necesariamente representativas de los posibles informantes de la encuesta.

50. Muchos de los métodos más utilizados dependen de protocolos verbales (Willis, Royston y Bercini, 1991). Se pide a los informantes que rellenen el proyecto de cuestionario y que describan cómo interpretan cada partida. Un entrevistador realizará investigaciones acerca de determinadas palabras, definiciones, pautas de salto de preguntas u otros elementos del cuestionario sobre los que desee recibir observaciones específicas del informante, al que se pide que pregunte todo lo que no le resulte claro. A veces se le pide que lo haga según rellena el cuestionario ("reflexión en voz alta simultánea"), y en otras ocasiones se organiza una sesión posterior para oír sus observaciones ("reflexión en voz alta retrospectiva"). A veces se incorporan sondeos para investigar la claridad de diferentes partidas o elementos del cuestionario en posteriores entrevistas. La ventaja de esta técnica es que no está sujeta a sesgos impuestos por el entrevistador. La desventaja es que no funciona bien en el caso de informantes que no se sienten cómodos cuando tienen que verbalizar sus opiniones o que no están acostumbrados a ello (Willis, 1994).

51. Una técnica afín consiste en que el entrevistador interrogue al informante acerca de alguna característica de la pregunta inmediatamente después de que el informante haya

terminado las respuestas de una partida (Nolin y Chandler, 1996). Este planteamiento depende menos de la capacidad y preparación del informante para verbalizar sus pensamientos, pero limita la investigación a aquellas partidas sobre las que el diseñador de la encuesta considera que puede formular preguntas. Además, puede introducir un sesgo del entrevistador, ya que los sondeos dependen de éste. En la medida en que el concepto de sondeo es diferente de la realización de una entrevista, algunos lo consideran un planteamiento artificicioso (Willis, 1994).

52. Otros planteamientos permiten al informante completar el instrumento de la encuesta con preguntas formuladas en grupos de discusión. Estos planteamientos tienen la ventaja de la interacción de los miembros del grupo, lo que puede dar lugar a la exploración de áreas que quizá no puedan tratarse en las entrevistas individuales.

53. La convocatoria de paneles de expertos (un grupo pequeño de expertos reunidos para examinar un cuestionario), puede ser un medio eficaz para sacar a la luz los problemas del cuestionario (Czaja y Blair, 1996). Los profesionales del diseño de encuestas y los expertos en las diferentes materias reciben el cuestionario varios días antes de reunirse con quienes lo diseñaron. En una reunión del grupo, sus componentes examinan y formulan observaciones sobre él pregunta por pregunta.

54. Los métodos de investigación cognitiva se utilizan ahora ampliamente en el diseño de cuestionarios y para reducir el error de medición en las encuestas en los países desarrollados. Sudman, Bradburn y Schwarz (1996) resumen las principales conclusiones relacionadas con la metodología de las encuestas. Tucker (1997) analiza las cuestiones metodológicas en la aplicación de la psicología cognitiva a la investigación sobre encuestas.

3. Reentrevistas

55. Las reentrevistas —medición repetida de la misma unidad en una encuesta mediante entrevistas— permiten volver a formular las mismas preguntas de la entrevista inicial (o algunas de ellas). Suelen realizarse con una submuestra pequeña de las unidades muestrales de una encuesta (normalmente, en torno al 5%) y tienen uno o varios de los siguientes objetivos:

- Averiguar qué entrevistadores falsifican los datos;
- Conocer qué entrevistadores no entienden bien los procedimientos y deberían recibir capacitación para corregir ese fallo;
- Estimar la varianza simple de la respuesta;
- Estimar el sesgo de la respuesta.

56. Los dos primeros objetivos permiten obtener información sobre los errores de medición resultantes de efectos asociados al entrevistador. Los dos últimos aportan informaciones sobre los errores de medición debidos al efecto conjunto de las cuatro causas: entrevistador, cuestionario, informante y forma de recopilación de datos.

57. A continuación se examinan los requisitos específicos de diseño para cada uno de los cuatro tipos de reentrevista [véase Forsman y Schreiner (1991)]. Además, se presentan algunos métodos para analizar los datos de las reentrevistas, junto con las limitaciones de los resultados.

Reentrevista para averiguar cuándo hay falsificaciones de los entrevistadores

58. Los entrevistadores pueden falsificar los resultados de las encuestas de varias maneras; por ejemplo, cuando el entrevistador inventa las respuestas a algunas o a todas las preguntas o si deliberadamente no sigue los procedimientos establecidos. Para averiguar esas falsificaciones se elige una muestra de reentrevistas que realiza generalmente el personal de

supervisión. Por tasa de falsificación se entiende la proporción de entrevistas falsificadas por los entrevistadores descubiertas mediante la reentrevista oportuna. Schreiner, Pennie y Newbrough (1988) descubrieron que eso sucedió en un 0,4% en la Encuesta de población de los Estados Unidos, en el 0,4% en el Estudio nacional sobre las víctimas de delitos realizado en los Estados Unidos, y en el 6,5% en la Encuesta de vivienda y viviendas vacantes en la ciudad de Nueva York, todas ellas realizadas por la Oficina del Censo de los Estados Unidos.

Reentrevista para evaluar a los entrevistadores

59. Los programas de reentrevistas que identifican a los entrevistadores que no consiguen niveles aceptables de calidad se conocen con el nombre de reentrevistas para evaluar a los entrevistadores. El objetivo es averiguar qué entrevistadores no entienden correctamente los procedimientos de la encuesta, para reentrenarlos. La mayor parte de las características de diseño de este tipo de reentrevista son iguales a las señaladas en el párrafo anterior. Pueden utilizarse cuadros de tolerancia, basados en la teoría de control de la calidad estadística, con el fin de determinar si el número de diferencias en la reentrevista después de la conciliación es superior a un límite específico aceptable. Los programas de reentrevistas de la Oficina del Censo de los Estados Unidos utilizan niveles de tolerancia de calidad aceptables que oscilan entre el 6% y el 10% (Forsman y Schreiner, 1991).

Reentrevista para determinar la varianza simple de la respuesta

60. La reentrevista para determinar la varianza simple de la respuesta es una repetición independiente de los procedimientos de la entrevista original. En la medida de lo posible se repiten exactamente todas las directrices, procedimientos y procesos de la primera entrevista. La muestra de la reentrevista es una submuestra representativa del diseño de la muestra original. Se utilizan los entrevistadores, la forma de recopilación de datos, las normas para los informantes y los cuestionarios de la entrevista inicial. En la práctica estos supuestos no se dan siempre. Por ejemplo, si el cuestionario original es demasiado largo se utiliza un subconjunto del cuestionario de la entrevista inicial. *No* se concilian las diferencias entre la entrevista original y la reentrevista.

61. Un estadístico estimado a partir de este tipo de reentrevista es la tasa de diferencia bruta (*GDR*), que es la diferencia cuadrática media entre las respuestas de la entrevista original y de la reentrevista. La *GDR* dividida por 2 es una estimación no sesgada de la varianza simple de la respuesta (*SRV*). En el caso de las características que tienen dos posibles resultados, la *GDR* es igual al porcentaje de casos que tenían respuestas diferentes en la entrevista original y en la reentrevista. Brick, Rizzo y Wernimont (1997) ofrecen normas generales para interpretar la varianza de la respuesta medida en función de la *GDR*.

62. Otro estadístico es el índice de incongruencia (*IOI*), que mide la proporción de la varianza de la población total atribuida a la varianza simple de la respuesta. Por ello,

$$IOI = \frac{GDR}{s_1^2 + s_2^2}$$

donde s_1^2 es la varianza de la muestra para la entrevista original y s_2^2 es la varianza de la muestra para la reentrevista.

63. El valor del *IOI* muchas veces se interpreta como sigue:

- Un *IOI* de menos de 20 es una varianza relativa de la respuesta *baja;*
- Un *IOI* entre 20 y 50 es una varianza relativa de la respuesta *moderada;*
- Un *IOI* por encima de 50 es una varianza relativa de la respuesta *alta.*

64. Las medidas de la varianza de la respuesta, la *GDR* y el *IOI*, ofrecen a los usuarios de los datos información sobre la fiabilidad y coherencia de respuesta de las preguntas de

una encuesta. En Naciones Unidas (1982) pueden encontrarse ejemplos del uso de la *GDR* y el *IOI* en determinadas variables procedentes de una encuesta de fecundidad en el Perú, en relación con un error no muestral en las encuestas de hogares. Dentro de la segunda fase del Programa de encuestas demográficas y de salud se llevó a cabo en Pakistán, con una submuestra de mujeres entrevistadas en la encuesta principal (Curtis y Arnold, 1994), un programa de reentrevistas para evaluar la coherencia nacional de las respuestas. En Westoff, Goldman y Moreno (1990) se describe un estudio de reentrevistas realizado dentro del Programa de encuestas demográficas y de salud en la República Dominicana, que merece especial mención por la necesidad que hubo de adoptar varios compromisos, como la restricción de las reentrevistas a un reducido número de zonas geográficas y a un conjunto de la población objetivo. En Oficina del Censo de los Estados Unidos (1985), donde se examinan los procedimientos de evaluación del censo, se describen encuestas de reentrevistas en la India, realizadas con el objetivo de determinar la varianza de la respuesta.

65. En Feindt, Schreiner y Bushery (1997) se describen los esfuerzos de una encuesta periódica por mejorar constantemente los cuestionarios utilizando un programa de reentrevistas. Cuando las preguntas tienen altas tasas de discrepancia descubiertas en la reentrevista puede emprenderse una investigación de mejora del cuestionario utilizando métodos de investigación cognitiva. Estos métodos pueden identificar la causa de los problemas y proponer posibles soluciones. Durante la siguiente ronda de entrevistas puede efectuarse una reentrevista sobre las cuestiones revisadas para determinar si se ha logrado una mayor fiabilidad. Luego este proceso se repite con el resto de las preguntas con problemas.

Reentrevista para determinar el sesgo de la respuesta

66. Una reentrevista cuyo objetivo es medir el sesgo de la respuesta trata de obtener las respuestas verdaderas o correctas en una submuestra representativa del diseño de la muestra original. Con el fin de obtener las respuestas verdaderas se recurre a los entrevistadores y supervisores más experimentados. Además, en las reentrevistas se recurre al informante que dispone de mayores conocimientos; otra opción es sean los propios miembros del hogar quienes respondan. En la reentrevista se vuelven a hacer las mismas preguntas de la entrevista original, y las diferencias entre ambas se señalan al informante para establecer la "verdad". Otro planteamiento consiste en utilizar una serie de preguntas de sondeo en sustitución de las preguntas originales con el fin de obtener respuestas precisas, y luego conciliar las diferencias con el informante. Puede verse un examen de las encuestas de reentrevistas realizadas para obtener estimaciones del sesgo de respuesta en el informe en el que se describen los procedimientos de evaluación del censo presentado por la Oficina del Censo de los Estados Unidos (1985).

67. La reconciliación con el fin de establecer la verdad tiene sus limitaciones. Los informantes pueden transmitir a sabiendas información falsa y hacerlo por sistema tanto en la entrevista original como en la reentrevista, lo que significará que la conciliación no permitirá obtener las estimaciones "verdaderas". En un estudio sobre la calidad de los datos de reentrevistas de la encuesta permanente de población de los Estados Unidos, Biemer y Forsman (1992) determinaron que nada menos que el 50% de los errores de la entrevista original no habían aflorado en la conciliación de la reentrevista.

68. El sesgo de la respuesta se estima calculando la tasa de diferencia neta (*NDR*), considerándose que la diferencia media entre la respuesta de la entrevista original y la respuesta de la reentrevista conciliada representa la respuesta "verdadera". En este caso,

$$NDR = \frac{1}{n}\sum_{i=l}^{n}(y_{Oi} - y_{Ti})$$

donde n es el tamaño de la muestra de la reentrevista; y_0 es la respuesta de la entrevista original e y_T es la respuesta de la reentrevista después de la conciliación, que se supone que es la respuesta verdadera.

69. La *NDR* aporta información sobre la precisión de una pregunta e indica también las preguntas que ofrecen resultados sesgados. La existencia de este sesgo debe considerarse cuando se analizan los datos y se interpretan los resultados. Brick y otros (1996) utilizaron una reentrevista intensiva para obtener una mejor comprensión de la perspectiva y razones de las respuestas del informante, lo que dio lugar a estimaciones del sesgo de respuesta. Aunque trabajaron con una muestra pequeña, los autores llegaron a la conclusión de que el método podía descubrir y cuantificar los sesgos. Se formularon estimaciones en las que se habían corregido los sesgos y que ilustran los posibles efectos en las estimaciones cuando se dispone de medidas del sesgo.

4. Estudios de comprobación de los registros

70. Un estudio de comprobación de los registros compara las respuestas de la encuesta en casos individuales con los valores obtenidos de una fuente externa, que en teoría contendría los valores verdaderos para las variables de la encuesta. Estos estudios se emplean para estimar los sesgos de respuesta resultantes del efecto combinado de las cuatro causas del error de medición (entrevistador, cuestionario, informante y modalidad de recopilación de datos).

71. Groves (1989) describe los tres tipos de diseño utilizados para los estudios de comprobación de registros:

- Comprobación inversa de los registros;
- Comprobación *a posteriori* de los registros;
- Comprobación total de los registros.

72. En la *comprobación inversa* de los registros, se selecciona la muestra de la encuesta a partir de una fuente con datos precisos sobre características importantes del estudio. Luego la estimación del sesgo de la respuesta se basa en una comparación entre las respuestas de la encuesta y la fuente de datos exactos.

73. Muchas veces la fuente del registro es una enumeración de unidades (hogares o personas) con una determinada característica; por ejemplo, las que reciben una forma particular de ayuda estatal. En tal caso el estudio de verificación inversa de los registros no mide los errores de información por exceso (es decir, unidades que dicen tener una característica que de hecho no tienen). Estos estudios pueden medir únicamente la proporción de registros de la fuente que señalan correctamente esa característica o cometen el error de no señalarla. Por ejemplo, la United States Law Enforcement Assistance Administration (1972) realizó un estudio de este tipo para evaluar los errores sobre las víctimas de delitos registradas. Se tomaron muestras de los registros del departamento de policía y se establecieron contactos con las víctimas. Durante las entrevistas, las víctimas mencionaron el 74% de los delitos que constaban en los registros del departamento de policía.

74. En una *comprobación a posteriori* de los registros se consultan los sistemas de registros externos que contienen información precisa sobre los informantes, una vez obtenidas las respuestas de la encuesta. Las estimaciones del sesgo de respuesta se basan en una comparación de las repuestas de la encuesta con los valores de los sistemas de registro. Estos estudios ofrecen la oportunidad de cuantificar los casos de error por exceso en la información. Una dificultad de este tipo de estudios es que obligan a ponerse en contacto con los organismos encargados de mantener los registros y obtener permiso de los informantes para conseguir esta información. Si la respuesta de la encuesta indica que la unidad no tiene una característica determinada, puede ser difícil buscar esa unidad en el sistema de registro. Por ello estos estudios ofrecen sólo una capacidad limitada de medir los casos de error por defecto. Chaney (1994) describe un estudio de este tipo en el que se compararon las informaciones presentadas por los mismos profesores acerca de su preparación académica con

las actas de los centros universitarios. Los datos indicaban que las informaciones facilitadas sobre el tipo y el año de la obtención del título y las áreas generales de especialización eran precisas en su mayoría; en cambio, la información sobre los cursos y las horas acreditadas era menos exacta.

75. Una *comprobación total de los registros* combina las características de las dos modalidades anteriores. Se selecciona una muestra a partir de un marco que comprenda toda la población y se localizan los registros de todas las procedencias que tengan relación con los casos de la muestra. En consecuencia, los errores de información por exceso y por defecto pueden cuantificarse comparando las respuestas de la encuesta con todos los registros (es decir, del marco muestral y de las fuentes externas) de los informantes de la encuesta. Aunque ese tipo de comprobación de registros evita los inconvenientes de las otras dos modalidades, requiere una base de datos que comprenda todas las unidades de la población y todos los eventos correspondientes de esas unidades. Marquis y Moore (1990) ofrecen una descripción detallada del diseño y análisis de un estudio realizado para estimar los errores de medición en la Encuesta de los Estados Unidos sobre los ingresos y la participación en programas. En ese estudio se confrontan los datos sobre las cantidades percibidas como prestaciones de ocho programas federales y estatales en cuatro estados con los registros administrativos sobre los mismos programas. El perfil de calidad de la encuesta (Oficina del Censo de los Estados Unidos, 1998) permite obtener una información resumida del diseño y del análisis.

76. Los tres tipos de estudios de comprobación de registros tienen algunas limitaciones comunes debidas a los tres siguientes supuestos que, en la práctica, son poco realistas y no están nunca justificados. Primero, que los sistemas de registro están siempre libres de errores de cobertura, falta de respuesta o ausencia de datos; segundo, que los registros individuales de estos sistemas son completos, precisos y libres de errores de medición; tercero, que los errores de cotejo (errores que se producen al tratar de cotejar los registros de la encuesta sobre los informantes) son inexistentes o mínimos.

77. El sesgo de respuesta de una característica puede estimarse determinando la diferencia media entre la respuesta de la encuesta y el valor de la comprobación de registros de dicha característica, de acuerdo con la siguiente fórmula:

$$\text{Sesgo de respuesta} = \frac{1}{n}\sum_{i=l}^{n}(Y_i - X_i)$$

donde: n es el tamaño de la muestra del estudio de comprobación de registros; Y_i = respuesta de la encuesta para la i^a persona de la muestra, y X_i = valor de comprobación de registros de i^a persona de la muestra.

78. Las medidas del sesgo de respuesta tomadas de un estudio de comprobación de registros ofrecen información acerca de la exactitud de una pregunta y descubren las preguntas que producen estimaciones sesgadas. Estas medidas pueden utilizarse también para evaluar alternativas de varias características del diseño de la encuesta, como el diseño del cuestionario, los períodos de recuerdo, las formas de recopilación de datos y las técnicas de delimitación. Cash y Moss (1972) presentan los resultados de un estudio de comprobación inversa de registros en tres condados de Carolina del Norte acerca de la comunicación de accidentes de automóvil. Se realizaron entrevistas en hogares de la muestra en los que había algún miembro que se había visto implicado en un accidente de automóvil en los 12 meses anteriores a la entrevista. En el estudio se comprobó que los accidentes no notificados representaban sólo el 3,4% de los accidentes ocurridos en los 3 meses próximos a la entrevista, mientras que en los que habían tenido lugar entre 9 y 12 meses antes la proporción de omisiones era superior al 27%.

5. Estudios sobre la varianza del entrevistador

79. Para estudiar la varianza del entrevistador, las asignaciones de entrevistadores deben hacerse en forma aleatoria, de manera que las diferencias de los resultados obtenidos por los diferentes entrevistadores puedan atribuirse a los efectos provocados por ellos mismos.

80. Esta varianza se estima asignando cada entrevistador a informantes diferentes pero similares, es decir, que tengan los mismos atributos con respecto a las variables de la encuesta. En la práctica, esta equivalencia se garantiza mediante la asignación aleatoria. La muestra se divide en subconjuntos aleatorios, cada uno de los cuales representaría a la misma población, y cada entrevistador se ocupa de un subconjunto diferente de la muestra. En este contexto, cada entrevistador realiza una pequeña encuesta con todos los atributos esenciales de las encuestas de gran magnitud, con la única diferencia de su tamaño. O'Muircheartaigh (1982) describe la metodología utilizada en el Estudio Mundial sobre la Fecundidad para medir la varianza de la respuesta debida a los entrevistadores y ofrece estimaciones de la varianza de la respuesta sobre las encuestas realizadas en el Perú (1984b) y Lesotho (1984a).

81. En las encuestas de entrevistas directas las asignaciones interpenetradas de entrevistadores se definen con criterios geográficos para evitar los gastos de viaje. Las zonas asignadas tienen un tamaño ajustado al volumen de trabajo del entrevistador. Se marcan pares de zonas de asignación que se confían a una pareja de entrevistadores a cada uno de los cuales se asigna al azar la mitad de las unidades de vivienda. Así cada entrevistador entrevista en dos zonas de asignación y cada una de estas zonas se confía a dos entrevistadores distintos. El diseño consiste en un experimento (comparación de los resultados de dos entrevistas en cada una de las dos zonas) que se repite por un número igual al de parejas de entrevistadores. Bailey, Moore, y Bailar (1978) presentan como ejemplo de interpenetración para las entrevistas personales las que se llevaron a cabo en ocho ciudades de los Estados Unidos para el Estudio nacional sobre las víctimas de delitos.

6. Codificación del comportamiento

82. El desempeño del entrevistador, sea durante el entrenamiento o en el terreno, puede evaluarse con un sistema de codificación del comportamiento. Personal debidamente capacitado observa una muestra de entrevistas, codifica sus diferentes aspectos o graba una muestra de entrevistas; la codificación se hace a partir de esas grabaciones. Se asignan códigos para registrar las principales actividades verbales y comportamientos del entrevistador, como la formulación de preguntas, el uso de preguntas de sondeo y el resumen de las respuestas. Los códigos se pueden clasificar según cómo lee la pregunta el entrevistador, si las preguntas se hacen en forma correcta y completa, si se introducen pequeños cambios y omisiones y si el entrevistador modifica notablemente el texto de la pregunta o no la termina. El sistema de codificación indica si los sondeos orientaron al informante hacia una respuesta determinada, definieron mejor la pregunta o no influyeron en la respuesta, si las respuestas se resumieron fielmente o no, y si otros comportamientos fueron adecuados o inadecuados. Los resultados codificados reflejan hasta qué punto el entrevistador utilizó los métodos que se le habían enseñado durante su capacitación; en otras palabras, se entiende por comportamiento "adecuado" o "inadecuado" el que ha sido desaconsejado durante la capacitación del entrevistador. Para establecer y mantener un alto nivel de fiabilidad en cada una de las entrevistas codificadas, un segundo profesional debería codificar independientemente una submuestra de entrevistas.

83. Un sistema de codificación del comportamiento puede indicar a los nuevos entrevistadores cuáles de sus técnicas son aceptables y cuáles no lo son, y serviría de base para que los entrevistadores y supervisores examinen la labor realizada sobre el terreno y analicen los problemas identificados en la codificación. Además, facilita una evaluación del desem-

peño del entrevistador, que puede compararse con el de otros y con el suyo propio durante otras entrevistas codificadas (Cannell, Lawson, y Hauser, 1975).

84. Oksenberg, Cannell y Blixt (1996) describen un estudio en el que se grabó el comportamiento del entrevistador, que luego se codificó y analizó para identificar los problemas del entrevistador y del informante en el contexto de la Encuesta nacional de gastos médicos de 1987 realizada por el Organismo de los Estados Unidos para la Investigación y la Calidad de la Atención de Salud, cuyo objetivo era averiguar si el comportamiento del entrevistador se había alejado de los principios y técnicas que se le dieron durante su capacitación. Los autores concluyeron que en muchos casos los entrevistadores no leyeron las preguntas como estaban redactadas y que en ocasiones las hicieron de tal manera que podían condicionar las respuestas. Los entrevistadores no habían realizado los sondeos necesarios, y cuando los hicieron con frecuencia eran inadecuados y condicionaban en parte la respuesta.

D. Observaciones finales: el error de medición

85. El error de medición se da en eel proceso de recopilación de datos y se debe a cuatro causas principales: el cuestionario, el método o modalidad de recopilación de datos, el entrevistador y el informante. Para cuantificar la existencia y magnitud de un tipo específico de error de medición se necesita una planificación previa y una consideración atenta del problema. Si no se llevan a cabo estudios en pequeña escala (es decir, con muestras limitadas) hay que realizar estudios especiales que requieren la aleatoridad de las submuestras, reentrevistas y comprobaciones de registros. Estos estudios suelen ser costosos y requieren la ayuda de un estadístico para el análisis de datos. No obstante, si hay bastante preocupación de que el tema no se resolverá adecuadamente durante las preparaciones de la encuesta o si la fuente del error es particularmente patente en la encuesta que se está realizando, los responsables de la encuesta deberían adoptar medidas para diseñar estudios especiales con el fin de cuantificar la fuente principal de error o que cree más problemas.

86. Nunca se insistirá demasiado en la importancia de realizar estudios para comprender y cuantificar el error de medición en una encuesta. Ello resulta especialmente importante si los conceptos que se van a medir son nuevos y complejos. Los análisis que los usuarios efectúan dependen de si tienen o no datos de buena calidad y de si comprenden o no su carácter y limitaciones. Los estudios sobre el error de medición requieren un compromiso explícito del programa de encuestas, por el mucho tiempo y recursos que hay que dedicarles. No obstante, el compromiso no acaba con la realización de los estudios. Éstos deben analizarse y sus resultados deberán comunicarse a los analistas para que evalúen a su vez el efecto del error de medición en su resultado. En los informes metodológicos y técnicos [véanse, por ejemplo, los informes metodológicos y analíticos elaborados por el Programa de encuestas demográficas y de salud (Stanton, Abderrahim y Hill, 1997; Institute for Resource Development/Macro Systems Inc., 1990; Curtis, 1995)] suelen encontrarse estudios especiales sobre análisis de pruebas y experimentos y evaluaciones de la calidad de los datos. Finalmente, los resultados de los estudios sobre el error de medición son importantes para mejorar la próxima aplicación de la encuesta. La introducción de mejoras significativas en la medición depende en buena medida de la disponibilidad de conocimientos y de los resultados de las encuestas anteriores. Para mejorar en el futuro la calidad de los datos de las encuestas se necesita el compromiso de los profesionales de la investigación sobre encuestas.

REFERENCIAS BIBLIOGRÁFICAS

Bailey, L., T. F. Moore and B.A. Bailar (1978). An interviewer variance study for the eight impact cities of the National Crime Survey Cities Sample. *Journal of the American Statistical Association,* vol. 73, págs. 16-23.

Biemer, P. P., y G. Forsman (1992). On the quality of reinterview data with application to the current population survey. *Journal of the American Statistical Association*, vol. 87, págs. 915-923.

Biemer, P. P. y otros, comps. (1991). *Measurement Errors in Surveys*. Nueva York: John Wiley and Sons.

Bishop, G. F. y otros (1988). A comparison of response effects in self-administered and telephone surveys. En *Telephone Survey Methodology*, R. M. Groves y otros, comps. Nueva York: John Wiley and Sons, págs. 321-340.

Blair, J., G. Menon y B. Bickart (1991). Measurement effects in self vs. proxy responses to survey questions: an information-processing perspective. En *Measurement Errors in Surveys,* P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 145-166.

Bradburn, N. M. (1983). Response effects. En *Handbook of Survey Research*, P. H. Rossi, J. D. Wright y A. B. Anderson, comps. Nueva York: Academic Press, págs. 289-328.

_____ y S. Sudman (1991). The current status of questionnaire design. En *Measurement Errors in Surveys*, P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 29-40.

_____ y asociados (1979). *Improving Interviewing Methods and Questionnaire Design: Response Effects to Threatening Questions in Survey Research*. San Francisco, California: Jossey-Bass.

Brick, J. M., L. Rizzo y J. Wernimont (1997). *Reinterview Results for the School Safety and Discipline and School Readiness Components*. Washington, D.C.: United States Department of Education, National Center for Education Statistics. NCES 97-339.

Brick, J. M. y otros (1996). *Estimation of Response Bias in the NHES: 95 Adult Education Survey*. Working Paper, No. 96-13. Washington, D.C., United States Department of Education, National Center for Education Statistics.

Cannell, C. F., S. A. Lawson y D. L. Hauser (1975). *A Technique for Evaluating Interviewer Performance*. Ann Arbor, Michigan: Universidad de Michigan, Survey Research Center.

Cash, W. S. y A. J. Moss (1972). Optimum recall period for reporting persons injured in motor vehicle accidents. *Vital and Health Statistics*, vol. 2, No. 50. Washington, D.C.: Public Health Service.

Chaney, B. (1994). *The Accuracy of Teachers' Self-reports on Their Post Secondary Education: Teacher Transcript Study, Schools and Staffing Survey*. Working Paper, No. 94-04. Washington, D.C.: United States Department of Education, National Center for Education Statistics.

Collins, M. (1980). Interviewer variability: a review of the problem. *Journal of the Market Research Society*, vol. 22, No. 2, págs. 77-95.

Couper, M. P. y otros, comps. (1998). *Computer Assisted Survey Information Collection*. Nueva York: John Wiley and Sons.

Curtis, S. L. (1995). *Assessment of the Data Quality of Data Used for Direct Estimation of Infant and Child Mortality in DHS-II Surveys*. Occasional Papers, No. 3. Calverton, Maryland: Macro International, Inc.

_____ y F. Arnold (1994). *An Evaluation of the Pakistan DHS Survey Based on the Reinterview Survey*. Occasional Papers, No. 1. Calverton, Maryland: Macro International, Inc.

Czaja R. y J. Blair (1996). *Designing Surveys: A Guide to Decisions and Procedures*. Thousand Oaks, California: Pine Forge Press (una empresa de Sage Publications).

DeMaio, T. J. (1984). Social desirability and survey measurement: a review. En *Surveying Subjective Phenomena,* C. F. Turner y E. Martin, comps. Nueva York: Russell Sage, págs. 257-282.

Dillman, D. A. (1978). Mail and Telephone Surveys: *The Total Design Method*. Nueva York: John Wiley and Sons.

_____ (1983). Mail and other self-administered questionnaires. En *Handbook of Survey Research,* P. Rossi, R. A. Wright y B. A. Anderson, comps. Nueva York: Academic Press, págs. 359-377.

_____ (1991). The design and administration of mail surveys. *Annual Review of Sociology*, vol. 17, págs. 225-249.

_____ (2000). Mail and Internet Surveys: *The Tailored Design Method*. Nueva York: John Wiley and Sons.

Eisenhower, D., N. A. Mathiowetz y D. Morganstein (1991). Recall error: sources and bias reduction techniques. En *Measurement Errors in Surveys,* P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 127-144.

Feindt, P., I. Schreiner y J. Bushery (1997). Reinterview: a tool for survey quality management. En *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginia: American Statistical Association, págs. 105-110.

Forsman, G. e I. Schreiner (1991). The design and analysis of reinterview: an overview. En *Measurement Errors in Surveys*, P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 279-302.

Fowler, F. J. (1991). Reducing interviewer-related error through interviewer training, supervision and other means. En *Measurement Errors in Surveys*, P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 259-275.

Groves, R. M. (1989). *Survey Errors and Survey Costs*. Nueva York: John Wiley and Sons.

_____ y L. J. Magilavy (1986). Measuring and explaining interviewer effects. *Public Opinion Quarterly*, vol. 50, págs. 251-256.

Hastie, R. y D. Carlston (1980). Theoretical issues in person memory. En *Person Memory: The Cognitive Basis of Social Perception*, R. Hastie y otros, comps. Hillsdale, New Jersey: Lawrence Erlbaum, págs. 1-53.

Hill, D. H. (1994). The relative empirical validity of dependent and independent data collection in a panel survey. *Journal of Official Statistics*, vol. 10, No. 4, págs. 359-380.

Huang, H. (1993). *Report on SIPP Recall Length Study*. De uso interno de la Oficina del Censo de los Estados Unidos, Washington, D.C.

Institute for Resource Development/Macro Systems, Inc. (1990). *An Assessment of DHS-1 Data Quality*. Demographic and Health Surveys Methodological Reports, No. 1. Columbia, Maryland: Institute for Resource Development/Macro Systems, Inc.

Jenkins, C. y D. Dillman (1997). Towards a theory of self-administered questionnaire design. En *Survey Measurement and Process Quality*, L. Lyberg y otros, comps. Nueva York: John Wiley and Sons, págs. 165-196.

Kahn, R. L. y C.F. Cannell (1957). *The Dynamics of Interviewing*. Nueva York: John Wiley and Sons.

Kalton, G., D. Kasprzyk and D.B. McMillen (1989). Non-sampling errors in panel surveys. En *Panel Surveys*, D. Kasprzyk y otros, comps. Nueva York: John Wiley and Sons, págs. 249-270.

Kalton, G., D. McMillen y D. Kazprzyk (1986). Non-sampling error issues in SIPP. En *Proceedings of the Bureau of the Census Second Annual Research Conference*. Washington, D.C., págs. 147-164.

Kantorowitz, M. (1992). *Methodological Issues in Family Expenditure Surveys*, Vitoria-Gasteiz, Comunidad Autónoma de Euskadi: Euskal Estatistika-Erakundea, Instituto Vasco de Estadística.

Kish, L. (1962). Studies of interviewer variance for attitudinal variables. *Journal of the American Statistical Association*, vol. 57, págs. 92-115.

Lyberg, L. y D. Kasprzyk (1991). Data Collection Methods and Measurement Errors: An Overview. En *Measurement Errors in Surveys*, P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 237-258.

_____, P. Biemer, M. Collins, E. D. de Leeuw, C. Dippo, N. Schwartz y D. Trewin (1997). En *Survey Measurement and Process Quality*. Nueva York: John Wiley and Sons.

Marquis, K. H. y C. F. Cannell (1971). Effects of some experimental techniques on reporting in the health interview. En *Vital and Health Statistics*, Washington, D.C.: Public Health Service, Series 2 (Data Evaluation and Methods Research), No. 41.

_____ y J. C. Moore (1990). Measurement errors in SIPP program reports. En *Proceedings of the Bureau of the Census 1990 Annual Research Conference*. Washington, D.C., págs. 721-745.

Mathiowetz, N. (2000). The effect of length of recall on the quality of survey data. En *Proceedings of the 4th International Conference on Methodological Issues in Official Statistics*. Estocolmo: Statistics Sweden. Disponible en www.scb.seGrupp/Omscb_Dokument/Mathiowetz.pdf (consultado el 3 de junio de 2004).

_____ y K. McGonagle (2000). An assessment of the current state of dependent interviewing in household surveys. *Journal of Official Statistics*, vol. 16, págs. 401-418.

Naciones Unidas (1982). *National Household Survey Capability Programme: Non-sampling errors in Household Surveys: Sources, Assessment, and Control: Preliminary Version.* DP/UN/INT-81-041/2. Nueva York: Naciones Unidas, Departamento de Cooperación Técnica para el Desarrollo y Oficina de Estadística.

Neter, J. (1970). Measurement errors in reports of consumer expenditures. *Journal of Marketing Research*, vol. VII, págs. 11-25.

_____ y J. Waksberg (1964). A study of response errors in expenditure data from household interviews. *Journal of the American Statistical Association*, vol. 59, págs. 8-55.

Nolin, M. J. y K. Chandler (1996). *Use of Cognitive Laboratories and Recorded Interviews in the National Household Education Survey*. Washington, D.C.: United States Department of Education, National Center for Education Statistics. NCES 96-332.

Oficina del Censo de los Estados Unidos (1985). *Evaluating Censuses of Population and Housing*. Statistical Training Document. Washington, D.C. ISP-TR-5.

Oksenberg, L., C. Cannell y S. Blixt (1996). Analysis of interviewer and respondent behavior in the household survey. *National Medical Expenditure Survey Methods*, 7. Rockville, Maryland: Agency for Health Care and Policy Research, Public Health Service.

O'Muircheartaigh, C. (1982). *Methodology of the Response Errors Project*. WFS Scientific Reports, No. 28. Voorburg, Países Bajos: International Statistical Institute.

_____ (1984a). *The Magnitude and Pattern of Response Variance in the Lesotho Fertility Survey*. WFS Scientific Reports, No. 70. Voorburg, Países Bajos: International Statistical Institute.

_____ (1984b). *The Magnitude and Pattern of Response Variance in the Perú Fertility Survey*. WFS Scientific Reports, No. 45. Voorburg, Países Bajos: International Statistical Institute.

Schreiner, I., K. Pennie y J. Newbrough (1988). Interviewer falsification in Census Bureau Surveys. En *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginia: American Statistical Association, págs. 491-496.

Schuman, H. y S. Presser (1981). *Questions and Answers in Attitude Surveys*. Nueva York: Academic Press.

Schwarz, N. (1997). Questionnaire design: the rocky road from concepts to answers. En *Survey Measurement and Process Quality*, L. Lyberg y otros, comps. Nueva York: John Wiley and Sons, págs. 29-46.

_____, R. M. Groves y H. Schuman (1995). *Survey Methods*. Survey Methodology Program Working Paper Series. Ann Arbor, Michigan, Institute for Survey Research, University of Michigan.

_____ y H. Hippler (1991). Response alternatives: the impact of their choice and presentation order. En *Measurement Errors in Surveys*, P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 41-56.

_____ y S. Sudman (1996). *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco, California: Jossey-Bass.

Silberstein, A. y S. Scott (1991). Expenditure diary surveys and their associated errors. En *Measurement Errors in Surveys*, P. Biemer y otros, comps. Nueva York: John Wiley and Sons, págs. 303-326.

Singh, S. (1987). Evaluation of data quality. En *The World Fertility Survey: An Assessment*, J. Cleland and C. Scott, comps. Nueva York: Oxford University Press, págs. 618-643.

Sirken, M. y otros (1999). *Cognition and Survey Research*. Nueva York: John Wiley and Sons.

Stanton, C., N. Abderrahim y K. Hill (1997). *DHS Maternal Mortality Indicators: An Assessment of Data Quality and Implications for Data Use*. Demographic and Health Surveys Analytical Report, No. 4. Calverton, Maryland: Macro International, Inc.

Sudman, S., N. Bradburn and N. Schwarz (1996). *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, California: Jossey-Bass.

_____ y otros (1977). Modest expectations: the effect of interviewers' prior expectations on response. *Sociological Methods and Research*, vol. 6, No. 2, págs. 171-182.

Tucker, C. (1997). Methodological issues surrounding the application of cognitive psychology in survey research. *Bulletin of Sociological Methodology*, vol. 55, págs. 67-92.

_____ (1998). *Survey of Income and Program Participation (SIPP) Quality Profile*, 3a. ed. Washington, D.C.: United States Department of Commerce.

United States Federal Committee on Statistical Methodology (2001). *Measuring and Reporting Sources of Error in Surveys*, Statistical Policy Working Paper, No. 31. Washington, D.C.: United States Office of Management and Budget. Disponible en www.fcsm.gov (consultado el 14 de mayo de 2004).

United States Law Enforcement Assistance Administration (1972). *San José Methods Test of Known Crime Victims*. Statistics Technical Report No.1. Washington, D.C.

Vaessen, M. y otros (1987). Translation of questionnaires into local languages. En *The World Fertility Survey: An Assessment*, J. Cleland y C. Scott, comps. Nueva York: Oxford University Press, págs. 173-191.

Weiss, C. (1968), Validity of welfare mothers' interview response. *Public Opinion Quarterly,* vol. 32, págs. 622-633.

Westoff, C., N. Goldman y L. Moreno (1990). *Dominican Republic Experimental Study: An Evaluation of Fertility and Child Health Information*. Princeton, New Jersey: Office of Population Research; and Columbia, Maryland: Institute for Resource Development/ Macro Systems, Inc.

Willis, G.B. (1994). *Cognitive Interviewing and Questionnaire Design; A Training Manual*. Cognitive Methods Staff Working Paper, No. 7. Hyattsville, Maryland: National Center for Health Statistics.

_____, P. Royston y D. Bercini (1991). The use of verbal report methods in the development and testing of survey questions. En *Applied Cognitive Psychology*, vol. 5, págs. 251-267.

Woltman, H. F. y J. B.Bushery (1977). *Update of the National Crime Survey Panel Bias Study*. Internal United States Bureau of the Census report, Washington, D.C.

Capítulo X
Garantía de calidad en las encuestas: normas, directrices y procedimientos

T. Bedirhan Üstun, Somnath Chatterji,
Abdelhay Mechbal y Christopher J. L. Murray

En representación de los colaboradores de la Encuesta Mundial de Salud (WHS)*
Organización Mundial de la Salud
Pruebas Científicas e Información para las Políticas
Ginebra, Suiza

* Puede verse la lista completa de colaboradores de la WHS en el sitio web de la Encuesta: www.who.int/whs/

RESUMEN

La calidad de una encuesta es de importancia primordial para obtener resultados exactos, fiables y válidos. Los equipos encargados de las encuestas deben aplicar regularmente procedimientos sistemáticos de garantía de calidad para evitar prácticas inaceptables y reducir los errores en la recopilación de datos. El establecimiento de estrategias eficaces y eficientes para mejorar la calidad de una encuesta ayudará tanto a mejorar la recopilación oportuna de datos de alta calidad como a la validez de los resultados. La "garantía de calidad" puede considerarse también como un instrumento de organización para aplicar normas operacionales previamente definidas acerca de la estructura, el proceso y el resultado de la encuesta. Los equipos deberían atenerse a normas explícitas de calidad y seguir los procedimientos establecidos para hacer realidad dichas normas. Los procedimientos deben ser transparentes, ser supervisados de manera sistemática y comunicarse debidamente en el marco de la documentación general sobre la aplicación y los resultados de la encuesta. Es también importante que la encuesta se mida y resuma mediante indicadores cuantificables, en la medida de lo posible.

En el presente capítulo se esboza un planteamiento sistemático para aplicar medidas de garantía de calidad, que vayan más allá de los simples mecanismos de control. Una importante encuesta internacional —la Encuesta Mundial de Salud (WHS), realizada por numerosas instituciones en 71 países diferentes— permite ilustrar la aplicación de un programa total de garantía de calidad. La encuesta se concibió con el fin de obtener datos comparables para evaluar las diferentes dimensiones de los sistemas de salud en los países participantes utilizando muestras nacionalmente representativas. De acuerdo con la importancia de los resultados de la WHS se establecieron rigurosos procedimientos de garantía de calidad en los que intervinieron expertos internacionales convocados para formar un grupo de examen colegiado externo y para ayudar a los países a alcanzar normas de calidad acordadas de común acuerdo y viables con respecto a las siguientes áreas: metodología de selección de muestras, logro de tasas de respuesta aceptables, tratamiento de los datos omitidos, cálculo de las medidas de fiabilidad y comprobaciones para determinar la comparabilidad de los datos entre diferentes subgrupos de población y países.

Términos clave: Garantía de calidad, indicadores de calidad, Encuesta Mundial de Salud, datos omitidos, tasa de respuesta, muestreo, fiabilidad, comparabilidad entre poblaciones, comparaciones internacionales.

A. Introducción

1. Una de las características básicas en relación con el diseño y realización de una encuesta es su "calidad" (Lyberg y otros, 1997). En todas las iniciativas de recopilación de datos, los resultados dependen de los materiales utilizados: no es posible escribir derecho con líneas torcidas. Además de la calidad de los instrumentos y técnicas de análisis de la encuesta, la calidad de los resultados de la encuesta depende sobre todo de la forma en que se aplica, en particular de la utilización de métodos acertados de muestreo y de la administración adecuada del cuestionario.

2. Para alcanzar la máxima calidad, todos los equipos de las encuestas deben atenerse a un conjunto uniforme de directrices. En ellas se identifican los siguientes aspectos:

a) Normas de calidad que deben cumplirse en cada paso de una encuesta;

b) Procedimientos de garantía de calidad en los que se identifiquen las medidas que deben adoptarse expresamente para supervisar la realización de la encuesta en contextos reales;

c) Evaluación del proceso de garantía de calidad que mida los efectos de las normas de garantía de calidad en los resultados de la encuesta y los procedimientos para mejorar la pertinencia y eficiencia del proceso global de garantía de calidad (Biemer y otros, 1991).

3. El objetivo general de las directrices es contribuir a mejorar la calidad más que evaluar *a posteriori* la ejecución de la encuesta. Como toda encuesta es una gran inversión en la que intervienen numerosas partes y produce importantes resultados que influyen en las políticas de una nación, es fundamental que la calidad ocupe un lugar importante en las operaciones. La garantía de calidad se entiende como un proceso continuo a lo largo de toda la encuesta, desde la preparación y muestreo hasta la recopilación y análisis de datos y la redacción del informe. Las directrices tratan también de conseguir una mejor comprensión del diseño de la encuesta entre los usuarios. Lo que se pretende con los procedimientos estándar es ayudar a conseguir los siguientes objetivos:

- Una recopilación de datos que sea relevante y significativa para las necesidades del país;
- Datos que puedan compararse dentro de un país y entre países diferentes para determinar las semejanzas y diferencias entre poblaciones;
- Aplicación práctica de la encuesta basada en los protocolos aceptados;
- Reducción al mínimo posible de los errores en la recopilación de datos;
- Mejora a lo largo del tiempo de la capacidad de recopilación de datos.

B. Normas y procedimientos de garantía de calidad

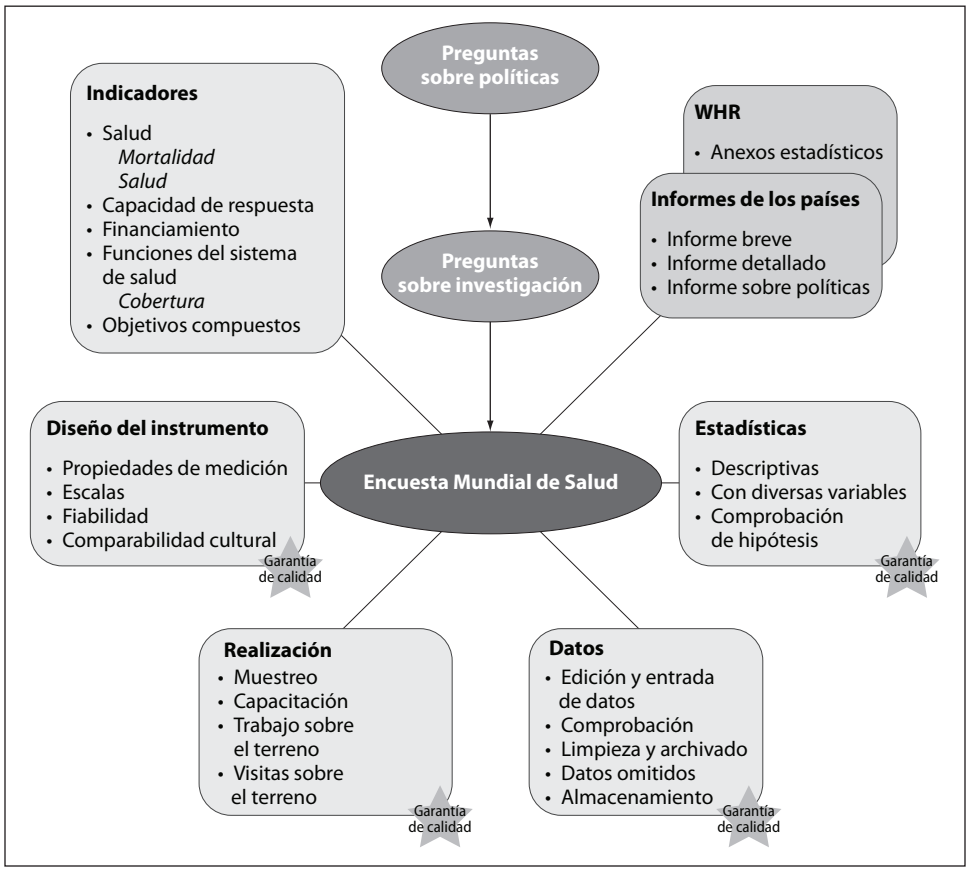
4. Por garantía de calidad (Statistics Canada, 1998) se entiende cualquier método o procedimiento para recopilar, procesar o analizar datos de encuestas con el fin de mantener o aumentar su fiabilidad o validez. Podría entenderse de varias maneras, semejantes pero distintas. En el presente capítulo utilizamos el paradigma de gestión de la calidad total, que examina el proceso de la encuesta en cada una de sus fases y trata de esbozar un planteamiento tanto para reducir los errores muestrales y no muestrales como para mejorar la pertinencia y viabilidad de la encuesta y la capacidad del país de realizar encuestas. Para lograr este objetivo, sin perder de vista la orientación práctica, en este capítulo se utilizarán las normas y procedimientos de garantía de calidad de la Encuesta Mundial de Salud (WHS) (Organización Mundial de la Salud, 2002) en la que se hace referencia a todos los pasos necesarios, a saber:

- Selección de las instituciones encargadas de la encuesta;
- Muestreo;
- Traducción;
- Capacitación;
- Realización de la encuesta;
- Entrada y captación de datos;
- Análisis de datos;
- Indicadores de la calidad;
- Informes de los países;
- Visitas sobre el terreno.

5. En la figura X.1 se presenta el ciclo completo de la WHS y se especifican los pasos arriba mencionados en cada una de las fases de realización de la encuesta. Las directrices sobre la garantía de calidad, redactadas por un gran número de participantes en las WHS y por expertos internacionales, tratan de establecer prácticas óptimas y viables con el fin de conseguir y supervisar una encuesta de buena calidad. Cada paso supone cierto examen de la calidad. Por ejemplo, es importante que los instrumentos de la encuesta tengan posibilidades adecuadas de medición, que el muestreo sea representativo de la población objetivo y que los datos sean limpios y completos.

6. Este conjunto de procedimientos constituye simplemente un ejemplo para ilustrar el "concepto de garantía de calidad" aplicado al diseño y realización de la encuesta, en

Figura X.1
Procedimientos de garantía de calidad de la WHS



cuanto proceso y como instrumento para mejorar los resultados de la encuesta en lo que se refiere a su pertinencia, exactitud, coherencia y comparabilidad. Todo equipo encargado del diseño y realización de una encuesta podría utilizar un planteamiento semejante teniendo en cuenta los objetivos específicos de su propia encuesta y la viabilidad de las normas de garantía de calidad propuestas en este capítulo. Sobre todo la calidad debería ser objeto de atención expresa y debería orientarse y supervisarse dentro de un contexto operacional. Los resultados del proceso de garantía de la calidad deberían comunicarse tanto en términos cuantitativos, utilizando indicadores adecuados cuando sea posible la medición (por ejemplo, coeficientes de muestreo, tasas de respuesta, datos omitidos, fiabilidad de la aplicación en caso de repetición), como en términos cuantitativos que resuman la estructura, el proceso y resultados de la encuesta.

C. Aplicación práctica de las directrices sobre garantía de calidad: ejemplo de la Encuesta Mundial de Salud

7. La estrategia de garantía de calidad total descrita más arriba se ha aplicado en el marco de la WHS con el fin de mejorar la calidad de las encuestas, en particular en varios países en desarrollo de Asia y del África subsahariana. En este apartado se utilizarán como orientación concreta las normas, procedimientos y sistemas de información sobre la garantía de calidad. Otros equipos de la encuesta pueden utilizar el presente ejemplo según sus objetivos. Por lo que nosotros podemos saber, ésta es la primera vez en que se aplican de forma sistemática los procedimientos de garantía de calidad en encuestas internacionales, y los organismos encargados de la ejecución y sus colaboradores los han valorado positivamente como medio para organizar su trabajo e informar al respecto. Según los datos iniciales, era posible detectar errores desde el comienzo y prevenirlos, así como aumentar el número de actividades terminadas y la precisión y eficiencia de los resultados.

8. La Organización Mundial de la Salud (OMS) ha puesto en marcha la Encuesta Mundial de Salud (WHS) como plataforma de recopilación de datos de la vida real para obtener información sobre la salud de las poblaciones y los sistemas sanitarios en forma continuada (Üstün y otros, 2003a, 2003b; Valentine, de Silva y Murray, 2000; Organización Mundial de la Salud, 2000). La WHS responde a la necesidad de los países de contar con un sistema detallado y sostenible de información sobre la salud y recopila datos a través de encuestas para medir los parámetros esenciales de salud de la población, además de agrupar procedimientos e instrumentos estándar para encuestas generales de población con el fin de presentar datos comparables entre los diferentes Estados miembros de la OMS. Estos métodos e instrumentos son de carácter modular y se han perfeccionado mediante el examen científico de las publicaciones, de amplias consultas con expertos internacionales y de pruebas piloto en gran escala realizadas en más de 63 países y en 40 idiomas (Üstün y otros, 2003a, 2003c; 2001). La WHS debe evolucionar progresivamente a medida que se vaya aplicando, gracias a la aportación constante de colaboradores muy distintos, como autoridades, instituciones encargadas de encuestas, científicos y otras partes interesadas. Los países y la OMS son propietarios conjuntos de los datos, y se ha manifestado el compromiso por la recopilación de datos a largo plazo, el fortalecimiento de la capacidad y la utilización de los resultados de las encuestas para orientar la formulación y aplicación de la política de salud.

9. En este capítulo se examinan sistemáticamente los diferentes pasos del proceso de las encuestas, con excepción del diseño y comprobación del cuestionario, que se examina en otro lugar (véase Üstün y otros, 2003b), y se presentan las normas de garantía de calidad de la WHS en cada una de las esferas. Se trata de normas aconsejables por medio de las cuales se podrá aumentar la eficiencia e impedir la adopción de prácticas inaceptables. La calidad debe ser objeto de la mayor atención, debido a la creciente importancia de los datos de la WHS para los Estados miembros de la OMS y sus repercusión en las políticas de salud. Por

ello, la WHS ha formulado directrices generales sobre la práctica de sus encuestas con el fin de aumentar la fiabilidad y validez de las encuestas de la salud, buscando que en lo posible los errores prevenibles queden reducidos al mínimo. Las directrices sobre garantía de calidad adoptadas se convertirán en instrumentos básicos de organización para la WHS y servirán también para organizar la encuesta y preparar y planificar su ejecución. Por ello, en este capítulo se presenta una orientación general sobre los aspectos fundamentales que deben ser objeto de particular atención para garantizar la recopilación de datos de buena calidad.

10. Estas directrices servirán también como pauta de evaluación para los responsables de las encuestas y los asesores sobre garantía de calidad (red de expertos internacionales con amplia experiencia en encuestas que se encargan del examen colegiado de todo el proceso). Dichos expertos realizarán visitas a los países para respaldar sus esfuerzos por aplicar la WHS y efectuarán una evaluación estructurada y detallada del proceso, que ayudará a los países a evaluar la calidad de manera sistemática y a identificar qué aspectos de las encuestas pueden mejorarse.

1. Selección de las instituciones encargadas de las encuestas

11. Para realizar una encuesta nacional se necesita un gran acervo de conocimientos, técnicas, recursos y personas experimentadas. Por ello, la actividad de las encuestas se ha organizado según diferentes estilos y tradiciones en los distintos países y sectores. Para garantizar que la WHS sea realizada en un país por un grupo competente, es importante establecer un método que permita determinar cuáles son las instituciones competentes y especificar las normas como si fueran condiciones contractuales. La práctica habitual de la WHS es consultar a los ministerios de salud, las oficinas regionales y los funcionarios de enlace y representantes de la OMS en los países, para conocer cuáles son esas instituciones competentes. Dado el tamaño y complejidad de la encuesta, la viabilidad debería demostrarse mediante un proceso de licitación contractual de acuerdo con los reglamentos de la OMS. El proceso comienza con requerimiento de información para dar con las instituciones competentes que puedan participar en las licitaciones sobre la WHS según las especificaciones técnicas de muestreo, entrevistas y recopilación de datos [las especificaciones técnicas para la WHS pueden consultarse en su sitio en la web: (www.who.int/whs)]. Las ofertas presentadas se evalúan en función de una serie de criterios, antes de proceder a la selección final.

12. Entre los criterios para evaluar las normas de desempeño de las instituciones que podrían encargarse de las encuestas figuran los siguientes:

- Su historial (es decir, su experiencia en al menos cinco grandes encuestas nacionales en el pasado reciente con muestras de 3.000 o más unidades);
- Su capacidad para realizar todo el proceso de la encuesta (muestreo, capacitación, recopilación de datos y análisis);
- Su experiencia en diferentes formas de recopilación de datos, incluidas las entrevistas directas (y otras formas posibles, como el teléfono, el correo, la computadora, etcétera);
- Documentación sobre encuestas anteriores (incluidos los aspectos relacionados con la medición de la representación de la muestra, cobertura de la población del país, calidad de las entrevistas, costo y tipo de capacitación, garantía de calidad y otros procedimientos);
- Antecedentes en la programación habitual del calendario de las encuestas y capacidad de completarlas dentro del marco cronológico establecido;
- Potencial para establecer y utilizar una infraestructura adecuada con respecto a los sistemas de información de la salud, en estrecha colaboración con el Ministerio de Salud, los organismos nacionales de estadística y otras instituciones.

13. El procedimiento de licitación contractual es útil para localizar la mejor oferta posible en lo que respecta a la actividad y los costos, y permite una evaluación comparativa de todos los posibles proveedores en un país. De esta manera, la OMS y el Ministerio de Salud pueden identificar la institución que mejor podría encargarse de la encuesta con el fin de desarrollar la capacidad para encuestas posteriores e incorporar los datos de la WHS en el sistema de información sobre la salud. El proceso contractual permite también aplicar sanciones cuando no se alcanzan los resultados previstos ni se garantiza debidamente la calidad. Deberían alentarse las ofertas hechas a través de consorcios, para tener la certeza de que los socios pertinentes (por ejemplo, el ministerio de salud junto con la oficina nacional de estadística) colaborarán mutuamente para garantizar el acceso a un marco muestral adecuado.

14. Utilizando la lista de criterios descritos más arriba, debería llevarse a cabo un examen atento de las diferentes propuestas presentadas. Este análisis comparativo debería documentarse adecuadamente.

15. En resumen, es importante no sólo identificar una institución que reúna las especificaciones técnicas de la encuesta deseada en el país correspondiente sino también ofrecer a dicha institución el apoyo técnico necesario, con el fin de alcanzar los resultados deseados. En el caso de encuestas nacionales en gran escala, muchas veces es preciso crear dentro del país una asociación entre grupos, instituciones y personas que tengan la especialización necesaria en el diseño, capacitación, ejecución, procesamiento de datos, análisis y redacción de informes.

2. Muestreo

16. La calidad de la muestra condiciona la de toda la encuesta. Si el diseño o la realización de la muestra, o ambos, presentan deficiencias, no se puede hacer mucho para compensar la escasa representatividad de la muestra o corregir la falta de información. Los resultados de la encuesta presentarán sesgos de signo desconocido y muchas veces de magnitud no cuantificable.

17. Para el caso de que haya una gran variedad de aplicaciones sobre el terreno, la OMS y un grupo de expertos técnicos internacionales han establecido un conjunto de directrices con el fin de conseguir una muestra válida para la WHS [las *WHS Sampling Guidelines for Participating Countries* pueden consultarse en el sitio web de la OMS (www.who.int/whs)]. Las normas de muestreo científico están basadas en métodos de selección probabilística y son ampliamente conocidas y aceptadas (Üstün y otros, 2001; Kish, 1995a). No obstante, estas normas no suelen aplicarse, por problemas de instrumentación, falta de supervisión de la ejecución de los procedimientos de muestreo sobre el terreno y/o costo elevado de su aplicación en determinados contextos y condiciones.

18. En las directrices de la OMS se hace hincapié en los principios científicos del muestreo para encuestas como normas explícitas de calidad, se presentan ejemplos de planes acertados de muestreo y se señalan muestras de garantía de calidad a las que deben atenerse los países. La OMS y los asesores técnicos ofrecen apoyo técnico a los países en caso necesario. A continuación se señalan algunos aspectos importantes del muestreo en la WHS:

a) La muestra de la WHS debería centrarse en la población *de facto* (es decir, todas las personas que viven en el país, incluidos los trabajadores extranjeros, inmigrantes y refugiados), y no en la población *de jure* (ciudadanos del país, únicamente). Es importante crear una buena representación como "miniatura" de la población global del país. Con este fin, es fundamental representar a todas las personas que viven en el país y contar con la cobertura geográfica completa del mismo;

b) El tamaño de la muestra debe ser suficiente para poder ofrecer estimaciones válidas (sólidas) de las cantidades pertinentes de alcance nacional o subnacional, según los objetivos de la encuesta; al mismo tiempo, debe buscarse un equilibrio entre

la necesidad de muestras de menor tamaño para conseguir mejores estimaciones y evitar el aumento correspondiente de los costos. El deseo de realizar muestras de gran tamaño no debe hacerse a expensas de la calidad. Por varios motivos, quizá deba contarse con una representación suficiente de las minorías (por ejemplo, las minorías étnicas u otros subgrupos) en cuyo caso sería recomendable el sobremuestreo (es decir, que exista la posibilidad de ofrecer una mayor probabilidad de selección). Si una subpoblación debe estar sobrerrepresentada en la muestra por razones científicas, deberían indicarse detalladamente las especificaciones correspondientes. En caso de sobremuestreo, en la fase del análisis de datos debería aplicarse una ponderación diferencial para corregir la distorsión provocada por ese sobremuestreo;

c) En la WHS se considera aceptable un marco muestral (es decir, una lista de zonas geográficas, hogares o individuos a partir de los cuales se selecciona la muestra, que podría tomarse de una lista de población informatizada, un censo reciente, las listas electorales, etcétera) con una cobertura del 90% de todos los subgrupos importantes. Los países deberán utilizar el marco muestral más reciente disponible. Si éste tiene ya dos o más años, en la penúltima fase de selección muchas veces es preciso elaborar una lista de hogares para actualizar el marco. Pueden utilizarse métodos de recuento rápido para actualizar las medidas del tamaño de las unidades primarias de muestreo antes de la selección. Entre esos métodos se incluyen el recuento en puntos seleccionados donde no se disponga de un marco actualizado debido a una cartografía anticuada o a otras razones. Además de los métodos de recuento rápido en zonas de muestreo seleccionadas, el marco puede seleccionarse consultando otras fuentes, como las direcciones postales de las oficinas locales de correo, las listas de las compañías de facturación de los servicios de agua o electricidad, etcétera Es fundamental que la población se pondere científicamente de acuerdo con el censo más reciente;

d) La muestra de la WHS se propone el objetivo de considerar a todos los miembros adultos de la población general de 18 años o más[1]. En la mayoría de los casos se toman los datos del censo más reciente como marco muestral. Los hogares se seleccionan utilizando un procedimiento de muestreo estratificado por conglomerados en varias etapas. Luego se selecciona un individuo por hogar mediante un procedimiento de selección aleatoria [por ejemplo, el método del cuadro de Kish (Kish, 1995a), o métodos alternativos, como el del último en nacer, y el método de Trohdahl/Carter/Bryant (Bryant, 1975)]. En esta etapa podrían utilizarse también cuadros numéricos aleatorios, siempre que los números de selección estén atentamente documentados. Cualquiera que sea la técnica utilizada, debería hacerse lo posible por reducir el sesgo de selección durante la aplicación real sobre el terreno. Los países deberían tratar de diseñar el plan de muestreo más sencillo posible que cumpla los objetivos de medición de la encuesta. En el caso de un diseño demasiado complejo, la aplicación puede presentar dificultades y será difícil controlar los errores. La viabilidad y la existencia de una trayectoria de datos para supervisar el diseño del muestro son condiciones clave para la calidad;

e) La WHS utiliza la definición de hogar de las Naciones Unidas[2]; no obstante, puede haber muchas variaciones en esta definición, debido a las circunstancias locales. La posible repercusión que las variaciones en la definición del hogar pueden tener en el muestreo debería explicarse en los informes de los países. Si éstos utilizaran un marco muestral de hogares, deberían utilizar en la encuesta la misma definición de hogar que se utilizó en el marco original;

f) La WHS utiliza una estrategia de muestreo científico, que comprende una probabilidad conocida de selección distinta de cero para todo individuo incluido en

[1] Actualmente, la WHS incluye únicamente a los adultos. En futuros proyectos se intentará elaborar una encuesta que incluya también a los niños.

[2] Para las Naciones Unidas, un hogar es un grupo de personas que viven bajo el mismo techo y comparten los servicios de cocina y alimentación (en otras palabras, comen en la misma fuente). Para la WHS, una persona forma parte del hogar aunque esté internada por problemas de salud. Estas personas deben incluirse en los registros de los hogares.

la encuesta. El uso de métodos probabilísticos estrictos en cada fase del muestreo es de importancia crucial y permite extrapolar los datos de la muestra a toda la población. De lo contrario, los resultados de la encuesta no serán representativos ni válidos;

g) Incluir las poblaciones institucionalizadas (internadas en una institución) en una encuesta de población general presenta dificultades, pues habría que elaborar marcos independientes. En relación con las entrevistas en instituciones (hospitales, asilos, cuarteles y cárceles) se plantean también numerosas cuestiones éticas. Dadas las grandes diferencias de la institucionalización en los distintos países, no es posible encontrar una solución única. Lo que hace la WHS es incluir a las personas institucionalizadas por un problema de salud siempre que sea posible entrevistarlas mientras dure la encuesta. Luego se utilizan las tasas de población institucional del censo para comprobar la concordancia de las obtenidas en la encuesta. Ello representa un motivo de especial preocupación para la WHS, ya que las personas que viven en instituciones como asilos, hospitales para personas internadas durante largos períodos, etcétera, se encontrarán probablemente en peores condiciones de salud que quienes no están en instituciones, razón por la cual deben incluirse en la muestra, para reducir así la posibilidad de infravalorar los problemas de salud;

h) En las directrices sobre muestreo de la WHS se explica claramente qué se entiende por falta total de respuesta de la unidad y el cálculo de las tasas de falta de respuesta en relación con las muestras deseadas y conseguidas. La estrategia de muestreo de la WHS no permite la sustitución de la falta de respuesta por otro hogar o individuo;

i) Los resultados de la encuesta sobre el muestreo deben indicar los errores estándar de las variables importantes, para que los usuarios puedan ver el error de medición en términos estadísticos;

j) El uso de sistemas de información geográfica (SIG) puede ser útil para mejorar la calidad de los resultados, ya que permitiría verificar la ejecución del plan de muestreo sobre el terreno; en otras palabras, comprobar que las entrevistas se han llevado a cabo en un determinado lugar y que no se trata de entrevistas realizadas en la calle o ficticias (De Lepper, Scholten y Stern, 1995). Los SIG pueden aportar también un valor adicional a los datos asociando la información sobre la distancia a los centros de salud, a los lugares de abastecimiento de agua y a otros recursos ambientales, por ejemplo, con parámetros de salud cuantificados (como el estado de salud, las enfermedades o los factores de riesgo) de la encuesta. Las calidades de dispersión de un parámetro pueden reproducirse en un mapa, lo que revelería gráficamente las desigualdades en materia de salud. Con esa mira, la WHS ha utilizado el sistema mundial de determinación de la posición (GPS) y ha digitalizado los mapas para la geocodificación de los datos siguiendo ciertas directrices (véase www3.who.int/whosis/gis). Se han adoptado algunas medidas jurídicas para mantener la confidencialidad de la información personal, ya que la geocodificación de la información puede violar las normas de protección de los datos.

Evaluación del muestreo

19. La estrategia de muestreo debería evaluarse antes del comienzo de la encuesta, para determinar si la estratificación es la adecuada y si la representación de la población y el tamaño y distribución de los conglomerados seleccionados son los previstos. En el informe deberían documentarse atentamente los procedimientos exactos utilizados sobre el terreno, señalándose además los posibles desvíos con respecto al diseño, para que los usuarios puedan estar mejor informados acerca de la calidad de los resultados de la encuesta.

20. Durante la recopilación de datos, los encargados sobre el terreno y/o en la sede deben supervisar atentamente la selección de hogares e individuos para comprobar, por ejemplo, si los cuadros de Kish se utilizan adecuadamente y si se han completado los registros de hogares.

21. Una vez recopilados los datos se utilizan los elementos de medición del análisis de datos (que se examinan más adelante) para evaluar la calidad de los datos por medio de:

- Un estadístico sinóptico que denominaremos "índice de desviación muestral" (SDI);
- Fiabilidad del proceso de prueba y comprobación para determinar la "estabilidad" del instrumento con respecto al uso de diferentes entrevistadores;
- Información sobre el grado de falta de respuesta y de omisión de datos.

22. Los procedimientos se describen con mayor detalle en la sección sobe el análisis de datos. El cuadro X.1 contiene una lista sinóptica detallada sobre la calidad del muestreo.

Cuadro X.1
Lista sinóptica sobre la calidad del muestreo

- Panorama general de la composición de la población (urbana/rural, minorías, idiomas, grupos sobremuestreados).
- Marco muestral y número de etapas de muestreo:
  - ¿Comprenden los marcos muestrales todas las poblaciones objetivo?
  - ¿Es reciente el marco muestral?
- Estratificación dentro del marco muestral.
- Unidades de muestreo en cada etapa: probabilidad de selección conocida.
- Tamaño de las unidades muestrales en cada etapa: cerciorarse de que todas las unidades muestrales tengan un tamaño que supere el mínimo determinado previamente.
- Comprobación del tamaño "sobre el terreno" de las unidades, verificación de la existencia de uno o más hogares por "dirección" seleccionada y determinación de cómo hacer la selección dentro de ellos.
- Tamaño de la muestra elegida.
- Ponderación de la probabilidad en el caso del hogar.
- Ponderación de la probabilidad en el caso del informante.
- Capacitación en el uso y aplicación adecuada del cuadro de Kish (o instrumento alternativo).
- Comprobación del procedimiento para la selección del informante en el hogar.
- Informe sinóptico sobre el muestreo en la realización práctica, desviaciones, ponderaciones, errores estándar.

3. Traducción

23. Para poder realizar comparaciones significativas de datos entre diversas culturas se necesita un instrumento pertinente que asigne un valor al mismo concepto para diferentes países. El instrumento de la WHS se ha elaborado teniendo en cuenta el examen científico de los actuales instrumentos de encuesta, consultas en gran escala con expertos y comprobaciones sistemáticas sobre el terreno en un estudio de encuestas en numerosos países (Üstün y otros, 2003a). Hemos registrado las características de los instrumentos de la encuesta, la pertinencia y la aplicabilidad cultural en otros lugares (Üstün y otros, 2003b). En cualquier otra encuesta, los encargados del diseño deben tratar de contar con los mejores instrumentos y medidas y comprobar que su instrumento es válido para el fin que se persigue, ofrece posibilidades satisfactorias de medición y ha superado las pruebas piloto de comprobación de su viabilidad y estabilidad.

24. Una vez que se cuenta con un instrumento adecuado, la traducción es uno de los elementos fundamentales para garantizar versiones equivalentes de las preguntas en diferentes idiomas. Dado el carácter multicultural de las sociedades en que vivimos, es fundamental que dispongamos de buenas traducciones que midan los mismos conceptos de la encuesta.

25. Muchas veces en un país el instrumento se traduce a numerosos idiomas, dependiendo del tamaño de los diferentes grupos lingüísticos del país. Se sugiere que todo grupo lingüístico que constituya más del 5% de la población sea entrevistado en su propio idioma. En el caso de los informantes entrevistados en un idioma para el que no se ha realizado una traducción oficial, debe hacerse hincapié en la comprensión de los conceptos clave. Los entrevistadores utilizan una de las traducciones existentes en el país para formular las preguntas en el idioma pertinente sin necesidad de traducción, utilizando las directrices generales. Un nuevo desafío para las grandes encuestas de carácter multinacional es que en muchos países de África y de Asia hay lenguas no escritas, y no se puede disponer de textos. En tales casos se recomienda la preparación de una traducción estándar de acuerdo con las directrices, y su transliteración con un texto de otra lengua conocida en el país, que se utilizará para preparar la versión escrita.

26. Las directrices para la traducción de los instrumentos de la WHS están basadas en la amplia experiencia de la OMS en la preparación y realización de estudios internacionales con múltiples socios y expertos lingüísticos. Las Directrices de Traducción de la WHS, que pueden consultarse en el sitio web de la WHS (www.who.int/whs), hacen hincapié en la importancia de mantener la equivalencia de conceptos y garantizar un procedimiento que identifique posibles inconvenientes y evite la distorsión del significado. En esas directrices se insiste en los siguientes aspectos:

- La traducción debería tratar de presentar un cuestionario localmente inteligible;
- El objetivo original de las preguntas debería traducirse al idioma local con la terminología más equivalente que sea posible;
- Las especificaciones pregunta por pregunta deberían tratar de comunicar el significado original de las preguntas y las opciones de respuesta precodificadas;
- El cuestionario debería ser traducido inicialmente por expertos en salud y en encuestas familiarizados con los conceptos fundamentales de la materia. Algunos términos clave seleccionados o que resultaron problemáticos durante la primera traducción directa deberían ser traducidos de nuevo por expertos lingüísticos, que luego formularían observaciones sobre las posibles interpretaciones de los términos y propondrían alternativas. Un grupo editorial, bajo la supervisión del oficial responsable de la encuesta en ese país, debería examinar la traducción y la traducción inversa e informar a la OMS sobre la calidad de la traducción;
- Para mejorar la calidad debería haber grupos de discusión y utilizar métodos lingüísticos cualitativos, como serían la elaboración de un inventario de expresiones locales y comparar expresiones entre idiomas. La OMS ha realizado ya estudios sistemáticos de traducción y entrevistas cognitivas en ciertos idiomas e incorporado los resultados de esos estudios al texto actual del cuestionario de la WHS. Se recomienda todavía la aplicación a sujetos locales de "entrevistas cognitivas" (estudios posteriores para explorar lo que los sujetos consideraron que era el significado de las preguntas) utilizando el cuestionario traducido. Es imprescindible traducir al idioma local todos los documentos de la WHS (el cuestionario de la WHS, las especificaciones pregunta por pregunta, el manual de la encuesta y los manuales de capacitación). El programa de entrada de datos puede dejarse en inglés. No obstante, si el país ha traducido el cuestionario de la WHS utilizando medios electrónicos de acuerdo con las especificaciones de la OMS, el programa de entrada de datos puede generarse automáticamente en el idioma correspondiente;

- Cada país de la WHS debería presentar al final de la fase piloto un informe sobre la calidad de la labor de traducción. Para los casos que resultaron ser especialmente difíciles de traducir, deberían solicitarse formularios de evaluación lingüística en los que se describan las dificultades de traducción;
- Los asesores de garantía de calidad del país deberían prestar especial atención a los pasos de la ejecución en el proceso de traducción y comprobar la lista de términos clave con el responsable principal de la encuesta en el país;
- En los países con muchos dialectos y/o lenguas no escritos deberían examinarse con la OMS protocolos específicos de traducción.

Evaluación de la traducción

27. Antes del comienzo de las entrevistas piloto en la WHS debería presentarse a la OMS una traducción completa del cuestionario. Ésta será comprobada por expertos competentes en los respectivos idiomas, que harán las recomendaciones oportunas al país, en caso necesario.

28. Debería examinarse la lista de términos clave que son objeto de traducción inversa, junto con un informe sobre el proceso de traducción y las cuestiones que plantea. Los formularios de evaluación lingüística (Üstün y otros, 2001) deberían ser examinados sistemáticamente por el encargado de la encuesta en el país y posteriormente por la OMS para determinar las cuestiones con problemas especiales y poder llegar a una solución común entre las diferentes lenguas, si fuera posible.

29. Se mantendrán conversaciones con los entrevistadores para llegar a un punto común de acuerdo sobre los procedimientos que emplear sobre el terreno cuando no se entienda bien un término, una frase o una pregunta. En estas conversaciones se examinará hasta qué punto los entrevistadores deben "explicar" e "interpretar" las preguntas a los informantes.

Cuadro X.2
Lista sinóptica para el examen de los procedimientos de traducción

- Idiomas hablados en el país; cobertura de los principales grupos lingüísticos.
- ¿Quién participó en el proceso de traducción?
- ¿Se tradujeron todos los materiales necesarios?
  - Cuestionario
  - Apéndice
  - Guía para la administración (sólo cuando los entrevistadores no tengan conocimientos de inglés)
  - Manual de la encuesta (sólo cuando los entrevistadores no tengan conocimientos de inglés)
  - Códigos de resultados
- ¿Qué problemas se plantearon en la traducción?
- ¿Qué protocolo se aplicó (por ejemplo, envío de la traducción completa a la OMS o sólo de los términos clave)?
- ¿Se completaron los formularios de evaluación lingüística?

D. Capacitación

30. La capacitación del equipo de la encuesta es un requisito clave para la calidad. Se trata de un proceso constante que debería llevarse a cabo antes y a lo largo del proceso de recopilación de datos, y finalizar con una sesión dedicada a un detallado examen *a posteriori*, una vez finalizado el trabajo sobre el terreno.

31. La capacitación debería darse al personal de todos los niveles del equipo de la encuesta, desde los entrevistadores al equipo docente y los supervisores, así como el del equipo central que gestiona el proceso en todo el país. De esa manera se garantizará que todas las personas implicadas sabrán claramente cómo pueden contribuir a garantizar la calidad de los datos.

32. La capacitación tiene los siguientes objetivos:

- Conseguir una aplicación uniforme de los materiales de la encuesta;
- Explicar las razones del estudio y su protocolo;
- Motivar a los entrevistadores;
- Presentar sugerencias prácticas;
- Mejorar la calidad global de los datos.

33. Para conseguir parte del objetivo de la capacitación, la OMS ha organizado talleres regionales para la WHS destinados a los principales investigadores de todos los países participantes y ha elaborado distintos materiales de capacitación, incluido un vídeo y un CD en que se recogen todas las cuestiones relacionadas con la capacitación.

Selección de entrevistadores

34. Es importante contar con entrevistadores experimentados y con personas familiarizadas con el tema de la encuesta.

35. Los entrevistadores deben haber terminado al menos un período completo de escolarización dentro de su país y dominar el idioma nacional principal. Cada país debe decidir qué otros niveles de formación se requieren y cuáles son las evaluaciones formales que se llevarán a cabo antes de la selección.

36. Los países deben decidir también si los entrevistadores deben ser profesionales de la salud o no. Las características de los entrevistadores (edad, sexo, educación, formación profesional, situación de empleo, experiencia anterior en encuestas, etcétera) deberán registrarse en una base de datos independiente. Esta información puede vincularse luego con los números de identificación de los entrevistadores para cada cuestionario terminado y se puede realizar un análisis del desempeño de cada entrevistador.

Longitud, métodos y contenido de la capacitación

37. La capacitación debe ser lo bastante larga como para que los entrevistadores puedan familiarizarse con las técnicas para una entrevista eficaz y con el contenido del cuestionario que se va a utilizar. En el caso de los trabajadores experimentados, la capacitación será más breve que en el de los que tienen menos experiencia.

38. La duración recomendada de la capacitación para la WHS es de entre tres y cinco días; tres serían los días adecuados para los entrevistadores experimentados que requieren capacitación sólo acerca del cuestionario. Se recomienda el período más largo para todos los demás entrevistadores.

39. Todas las actividades de capacitación debenser realizadas en la medida de lo posible por el mismo equipo, para conseguir una capacitación uniforme de todos los entrevistadores dentro de una sesión o de los diferentes grupos en diferentes momentos y lugares. Para reducir los costos y organizar actividades de alcance regional, la capacitación puede descentralizarse y organizarse como proceso en cascada. No obstante, esta reducción de los costos se ve luego contrarrestada por las desventajas de una capacitación diluida o poco uniforme.

40. Se recomienda encarecidamente la introducción de una sesión de refuerzo durante el período de recopilación de datos. Convendría organizarla hacia la mitad del período

de recopilación de datos para la WHS. De esa manera se podrían examinar los distintos aspectos de esa actividad, con particular atención a las actividades que están resultando complejas y difíciles o a las directrices que los entrevistadores no están aplicando suficientemente. Esta sesión podría servir también para el intercambio de opiniones sobre lo que se ha conseguido y sobre los aspectos positivos, con inclusión de observaciones de los supervisores y el equipo central a los entrevistadores, y viceversa.

41. Los métodos de capacitación deberían incluir en la medida de lo posible escenificaciones de entrevistas (con un mínimo de una por entrevistador). Este método permite asimilar las técnicas de entrevista de manera más eficaz. Para conseguir buenos resultados, antes de la capacitación deberían prepararse diferentes guiones que permitan ilustrar las distintas estructuras de articulación de la entrevista, el carácter de las explicaciones permitidas y los previsibles problemas durante una entrevista con informantes difíciles.

42. Además de la escenificación de posibles situaciones, antes de comenzar la recopilación efectiva de datos debería ofrecerse al menos una oportunidad de realizar una entrevista con un informante de la vida real ajeno al grupo de entrevistadores. Estas actividades prácticas deberían grabarse en la medida de lo posible, para luego reproducirlas y comentarlas durante las sesiones de capacitación. Se alienta a los países de la WHS a que preparen un vídeo de capacitación estándar semejante al de la OMS, si es posible. Deberían hacerse comentarios sobre cada actividad práctica o escenificación de entrevistas.

43. Deberán darse los materiales de capacitación a todos los entrevistadores, que utilizarán como material de referencia. El material suministrado deberá examinarse exhaustivamente durante la capacitación y, en su caso, traducirse a los idiomas utilizados en el país.

44. El contenido de las actividades de capacitación deberá incluir:

- Cuestiones administrativas;
- Planificación del trabajo sobre el terreno;
- Examen de todos los materiales presentados;
- Procedimientos de contacto, formularios de consentimiento y confidencialidad.

En la realización de una entrevista deberían incluirse los siguientes elementos:

- Procedimientos de las entrevistas sobre el terreno;
- Supervisión sobre el terreno y procedimientos de presentación de informes;
- Estructura del equipo de la encuesta y función de todos los miembros integrantes.

Evaluación de la capacitación

45. La evaluación de la capacitación debería producirse en distintos niveles. Es preciso evaluar a los entrevistadores para determinar si son capaces de realizar sus funciones con eficacia y qué apoyo necesitan, en su caso. Los entrevistadores pueden a su vez evaluar la capacitación ofrecida y a quienes la imparten. Debería haber una evaluación continua durante el período inicial de recopilación de datos y al finalizar el trabajo sobre el terreno.

46. Los supervisores deben ser evaluados igualmente por el equipo central. Conviene mencionar aquí que el carácter de la capacitación debe adaptarse a las tareas que los supervisores prevén de mayor complicación, como la respuesta ante quienes se niegan a responder, la comprobación cruzada y la verificación de algunas entrevistas seleccionadas, y la edición de entrevistas. Deberán elaborarse protocolos detallados para estos procedimientos, que se explicarán claramente durante el proceso de capacitación.

47. Puede realizarse una evaluación formal de los entrevistadores al final del período de capacitación, y podría entregarse un certificado a los que superaran las pruebas con éxito. Esta decisión debe ser adoptada y aplicada independientemente por cada uno de los países.

Cuadro X.3
Lista sinóptica para el examen de los procedimientos de capacitación

- Número de sesiones de capacitación.
- Número de días de capacitación.
- ¿Quién realizó la capacitación y qué competencia tenía en capacitación y en encuestas de salud?
- ¿Qué documentación se utilizó?
- Componentes prácticos: observación de escenificación de situaciones en contexto real.
- Problemas experimentados en la capacitación.
- Evaluación de la capacitación.

E. Realización de la encuesta

48. La planificación y gestión de la realización de la encuesta es una tarea compleja desde el punto de vista logístico, pero también en otros sentidos. Requiere gran preparación, programación y capacidad de sortear las dificultades sobre el terreno con el fin de alcanzar la muestra deseada. Estratégicamente, la realización de la encuesta es un elemento clave, del que depende la calidad de los datos. Por ello, es de gran importancia prestar suma atención a la calidad de esta fase y supervisarla en tiempo real para poder resolver los problemas conforme se vayan presentando.

49. La forma concreta como se realiza de hecho la encuesta es, dentro de todo el proceso, el paso que más contribuye a determinar la calidad. Una organización fuerte y centralizada de la encuesta en cada país ayudará a garantizar la calidad. Cada paso (impresión de los cuestionarios, elaboración de listas de muestras, inscripción de los objetos, envío de equipos de entrevistadores, realización de la supervisión diaria sobre el terreno, edición de los cuestionarios, etcétera) deberá planificarse y examinarse atentamente para comprobar su calidad. Más en concreto:

a) Cada equipo de la encuesta deberá preparar un plan de ejecución central y un calendario de tareas en los que se especifiquen claramente los detalles de la logística de la encuesta. El plan debería especificar cuántos entrevistadores se necesitan para cubrir una proporción determinada de la muestra en una región dada con cierto número de visitas (incluidas las repetidas) y una tasa de éxito determinada. En consecuencia, debería tener en cuenta la tasa de falta de respuesta prevista y las entrevistas incompletas, y la presencia del equipo de la encuesta en un determinado lugar;

b) Cada equipo de la encuesta debe tener un supervisor que examine y coordine la labor de los entrevistadores, además de ofrecer capacitación y apoyo sobre el terreno. El coeficiente ideal de supervisores-entrevistadores para la WHS varía entre 1:5 y 1:10, según el país y los lugares;

c) Los supervisores deberían establecer la labor prevista al comienzo de cada día de trabajo, en colaboración con los entrevistadores, y examinar los resultados al final del día. En ese examen, los entrevistadores informarán a sus supervisores acerca de las entrevistas y de los resultados. Los supervisores deben examinar las entrevistas terminadas para comprobar que la selección de los informantes en el hogar se ha realizado correctamente y que el cuestionario está terminado y debidamente codificado;

d) Deberá mantenerse un diario para supervisar los progresos de la encuesta en cada centro en los países de la WHS. Los elementos que deben registrarse son los siguientes:

- Número de informantes contactados, entrevistas realizadas y entrevistas incompletas;
- Tasas de respuesta, rechazo y falta de contacto;
- Número de visitas repetidas y resultados de las visitas.

Debe mantenerse información sobre cada entrevistador para que el supervisor pueda observar su labor de forma constante. Esta base de datos sobre el entrevistador puede utilizarse luego para recoger observaciones individuales y ayudará en las futuras contrataciones;

e) Al comienzo del período de la WHS cada país debe realizar una encuesta piloto, de una o dos semanas de duración. Sería una especie de ensayo general de la encuesta principal. Luego, el 50% de la muestra piloto volvería a ser entrevistada por otro entrevistador, para comprobar la estabilidad de la aplicación de la entrevista. El período piloto debería evaluarse con espíritu crítico y examinarse con la OMS. Los datos de esta experiencia deberían analizarse rápidamente para determinar los posibles problemas concretos de ejecución. Como el instrumento que se va a utilizar en la entrevista habrá sido ya sometido a amplias pruebas antes de este ensayo final, lo que se pretende ahora es identificar las cuestiones lingüísticas y de viabilidad de menor importancia y permitir una mejor planificación de la fase principal. También podrían salir algunos errores concretos en las pautas de salto de unas preguntas a otras, etcétera. Las observaciones recibidas a través de la prueba piloto permitirán corregir estos errores e introducir pequeños ajustes. Previa consulta con la OMS, debería comenzar el estudio principal;

f) Debe reconocerse la utilidad de imprimir los cuestionarios y buscar un formato práctico (por ejemplo, codificación en color de los conjuntos de rotaciones, plastificación de las tarjeras de los informantes). Todos los países deberían enviar a la OMS una ejemplar de los documentos impresos;

g) De acuerdo con las especificaciones del contrato de la WHS, el 10% de los informantes debería ser controlado de nuevo en forma aleatoria por los supervisores u otros equipos. Esta comprobación puede realizarse por teléfono o en persona, y su objetivo es comprobar que la entrevista inicial se ha realizado debidamente. La entrevista de confirmación debería abarcar la información demográfica básica y toda la información no recogida en la entrevista inicial;

h) De acuerdo con las especificaciones del contrato de la WHS, antes de transcurridos siete días después de la primera entrevista otro entrevistador deberá aplicar toda la entrevista a una muestra aleatoriamente seleccionada del 10% de la muestra total de informantes, para poder evaluar la fiabilidad del cuestionario (los informantes de esta segunda prueba no deben ser los mismos que los de la segunda comprobación mencionados en el inciso *g*);

i) Las tasas de respuesta deben supervisarse de forma continua y cada centro debe emplear una combinación de estrategias diversas para incrementar la participación en la encuesta y reducir la falta de respuesta. Por ejemplo, la publicidad en la televisión, en la radio, en los periódicos o en los medios de comunicación locales, el envío de cartas o tarjetas a los participantes, la solicitud de asistencia de los trabajadores sanitarios locales, la concesión de incentivos a la participación, la negociación con las autoridades tradicionales locales o con otras instancias reconocidas, etcétera, son todas ellas técnicas de relaciones públicas que se pueden utilizar para lograr una mayor respuesta. El uso de cada uno de los métodos debe decidirlo el país en cuestión;

Cuadro X.4
Lista sinóptica para el examen de la realización de la encuesta

Encuesta piloto

- ¿Dónde se llevó a cabo la encuesta piloto?
- ¿Qué capacitación se ofreció para ella?
- ¿Hubo problemas en la entrada de datos?
- Análisis de datos: véanse los resultados; ¿qué problemas se experimentaron?
- ¿Algún cambio en la metodología a raíz de la encuesta piloto?
- ¿Algún cambio en la traducción a raíz de la encuesta piloto?

Encuesta principal

- Número de entrevistadores, supervisores y coordinadores centrales:
  - ¿Cómo se lleva a cabo la supervisión? Comentarios.
- Logística:
  - Viajes: ¿Fue fácil o difícil llegar hasta el hogar? ¿Qué tipo de transporte se utilizó?
  - Organización del equipo.
- Procedimientos de contacto:
  - ¿Fue fácil o difícil ponerse en contacto con el informante?
  - ¿Cuántas visitas de contacto se realizaron?
  - ¿Cuál fue la tasa de rechazo y cuál fue la principal razón para no hacer la entrevista?
- Pago de los entrevistadores.
- Firma y registro del formulario de consentimiento (como parte del cuestionario o como hoja aparte).
- Procedimientos de comprobación sobre el terreno por parte de los supervisores.
- Procedimientos de comprobación centralizados.
- Devolución de los cuestionarios a la oficina central y seguridad.
- Comprobación final sobre el cuestionario y procedimiento para corregir los errores.
- Procedimientos de comprobación y supervisión.

Informes de situación sobre la producción semanal

- Evaluar el proceso de las entrevistas.
- Examinar las tasas de respuesta, rechazo y falta de contacto: comprobar la tasa de respuesta.
- Supervisar los resultados y garantizar que se lleva a cabo la recopilación de datos.

Verificación de los registros

- ¿Se registra con detalle el número de contactos (contacto/intento de contacto)?
  - ¿Se ha verificado al menos el 10% de las entrevistas de cada entrevistador para garantizar que algunas respuestas se mantienen constantes (edad, educación, composición del hogar) y que se ha llevado a cabo la entrevista?
    - Comprobar el número de entrevistas ya realizadas y la planificación del calendario de entrevistas.
    - Verificar que los códigos de resultado final de las entrevistas terminadas y de los rechazos se han asignado correctamente.
    - Comprobar que se han firmado los formularios de consentimiento con conocimiento de causa.

Todos los datos de identificación deducidos de los cuestionarios y del programa de entrada de datos.

Proyecto de informe con recomendaciones sobre las medidas que deberían adoptarse.

j) Cada encuesta debe aspirar a la tasa de respuesta más alta posible. En las especificaciones del contrato de la WHS se requiere una tasa global de respuesta de al menos el 75%. Este umbral no significa que una vez alcanzado el 75% deba interrumpirse la encuesta. Significa simplemente el mínimo aceptable, según acuerdo común de los colaboradores de la WHS, teniendo en cuenta las entrevistas anteriores en muchos países diferentes. En numerosos casos las tasas de respuesta de la WHS han sido más altas. Puede haber diferencias de un país a otro y deben realizarse comparaciones con otras encuestas del mismo país. En los cálculos, en todos los países debe utilizarse la misma definición de entrevista terminada. Durante los procedimientos de depuración de los datos se utiliza un algoritmo para determinar hasta qué punto puede darse por terminada una encuesta, tomando como base un conjunto de variables fundamentales;

k) Segundas visitas: De acuerdo con las especificaciones del contrato de la WHS, los equipos de las encuestas deben intentar hasta 10 segundas visitas (con inclusión de las llamadas telefónicas o las notas o tarjetas depositadas para indicar que el entrevistador ha visitado el lugar). El número medio de estas visitas depende de la tasa de respuesta. Cada centro debe examinar los beneficios que representa cada nueva visita adicional y consultar con la OMS acerca del número suficiente para ese país concreto;

l) La realización de la encuesta depende en gran medida de los recursos disponibles. Cada encuesta debe evaluarse en el contexto del país. Es fundamental comparar con otras encuestas semejantes del mismo país. En la evaluación deben tenerse en cuenta las costumbres y las tradiciones locales. A la hora de determinar la repercusión en la calidad hay que buscar una solución de compromiso entre un número de entrevistadores que realice más entrevistas durante un período de tiempo más largo o contar con un número mayor de entrevistadores que hagan menos entrevistas durante un período más breve.

F. Entrada de datos

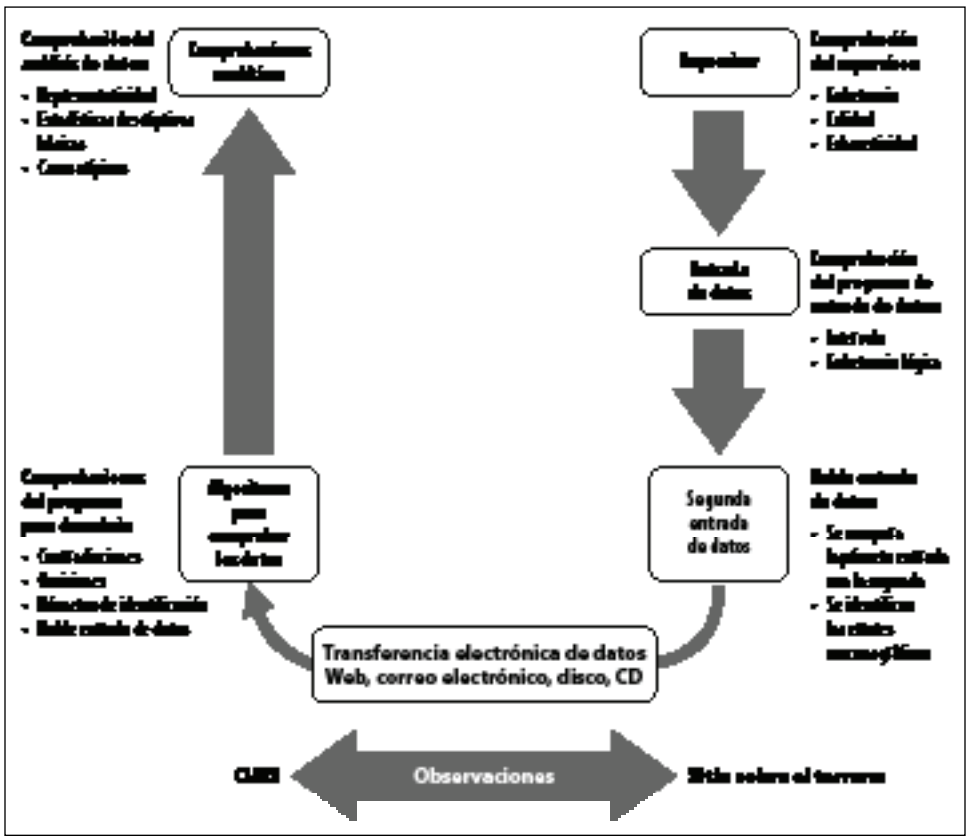
50. El resultado duradero de una encuesta son los datos. Es importante captar los datos con precisión y en forma oportuna. El proceso de la WHS está planificado de manera que haya una entrada local e inmediata de datos y una coordinación central. Es fundamental que los datos se transfieran a sistemas informatizados lo antes posible después de la recopilación. De esta manera pueden realizarse fácilmente comprobaciones rutinarias estándar mediante el uso de las computadoras locales. Los errores encontrados en ese momento pueden resolverse mientras la encuesta continúa sobre el terreno.

51. En la figura X.2 *infra* se describe el flujo de datos en la WHS y las medidas de garantía de calidad relacionadas con el mismo. Las tareas realizadas en el país se presentan a la derecha y las de la OMS a la izquierda.

52. Una vez aplicada la entrevista, se dan los siguientes pasos:

- El supervisor comprueba el formulario del cuestionario antes de que comience la entrada de datos.
- La entrada de datos (o captación/registro de datos) se realiza utilizando el programa de entrada de datos de la OMS. Éste comprueba los intervalos (por ejemplo, los intervalos de la variable de respuesta autorizada) y trata de determinar la coherencia lógica de los códigos relacionados (por ejemplo, una enfermedad no puede durar más que la propia edad; los hombres no pueden tener problemas ginecológicos; etcétera).

Figura X.2
Proceso de entrada de datos y garantías de calidad



- La segunda entrada de datos se realiza con el fin de identificar los errores mecanográficos y las cuestiones que se han saltado accidentalmente.
- Los datos se envían a la OMS en lotes utilizando el correo electrónico, un cederrón o un disquete.
- Una vez que los datos llegan a la OMS, los programas buscan las incoherencias, los valores omitidos, los problemas con los números de identificación o los casos de prueba/comprobación. Estos programas generan un informe que se remite a los países. Asimismo, las posibles correcciones recibidas de los países en cuestión se aplican a los datos.
- Los analistas de datos tratan de comprobar la representatividad, las estadísticas descriptivas clásicas y los casos atípicos. La representatividad se determina comparando la distribución por edad-sexo de la muestra realizada con la distribución de la población prevista. Se utilizan estadísticas descriptivas básicas para determinar las distribuciones de la respuesta e identificar las posibles distribuciones sesgadas, los resultados anómalos y los casos atípicos.
- La OMS envía sus observaciones a los países. Éstos presentarán, en caso necesario, correcciones y/o explicaciones de conformidad con las observaciones recibidas.

53. Cuestiones importantes sobre la calidad en la entrada de datos:

- La entrada de datos debería hacerse utilizando un programa concebido para ese fin, que permita efectuar controles de calidad. El uso de otros programas que no incluyan esas características puede representar una desventaja.

- Los formularios de entrevistas cumplimentados deberían ser comprobados por el supervisor antes de que comience la entrada de datos.
- El programa de entrada de datos debe ser accesible únicamente a los miembros del equipo encargado y a nadie más. Se trata de un requisito fundamental para mantener el carácter confidencial de la información.
- Se necesita una doble entrada de datos para evitar fallos mecanográficos o de edición. El programa de entrada de datos identifica la doble entrada una vez terminada la segunda fase.
- Los países deben proceder con cautela al introducir el número de identificación. Con ese fin, se envía a los países una lista de números de identificación válidos. El programa tiene un dígito de suma de control para comprobar que el código de identificación se introduce correctamente. La utilización de números de identificación correctos es especialmente importante en los casos de repetición de las entrevistas, ya que ese número se utiliza para emparejar los casos iniciales con las repeticiones.
- Los datos deben presentarse a la OMS periódicamente; por ejemplo, cada día o cada semana.
- Una vez que la OMS comienza a recibir datos de los países, se hacen las debidas comprobaciones y se envían observaciones a los países mientras continúa la recopilación de datos.
- Se aplican algunas normas para mantener la integridad y la precisión de los datos, lo que significa, por ejemplo, que se llevan a cabo comprobaciones para determinar si se utiliza el mismo informante dos veces y cuál es el alcance de los datos omitidos.

54. Los datos de identificación se separarán de los cuestionarios, y el programa de entrada de datos mantendrá la información confidencial en un archivo independiente, en caso necesario. Es responsabilidad del país proteger la confidencialidad. La seguridad de los datos durante su transferencia a través de Internet se garantiza utilizando un mecanismo criptográfico.

Evaluación de la entrada de datos

55. Deberían supervisarse y examinarse atentamente los siguientes aspectos (véase el cuadro X.5 en la página siguiente):

- Número de personal para entrada de datos y su capacitación;
- Número de formularios utilizados por día y persona, incluidas las tasas de error;
- Procedimientos de comprobación y supervisión de la entrada de datos;
- Tiempo transcurrido entre la terminación de la entrevista sobre el terreno y la entrada de datos;
- Número y periodicidad de las entrevistas terminadas que se envían a la OMS y problemas encontrados en el envío de los datos.

56. Aunque varios problemas de la entrada de datos se pueden reducir mediante entrevistas con ayuda de computadora, en la que los datos se introducen a medida que avanza la entrevista, los programas informáticos requerirán que lleven incorporados medios de comprobación para garantizar la aplicación correcta de la entrevista con todas las normas de elusión y bifurcación y la introducción de datos coherentes dentro de los intervalos especificados.

Cuadro X.5
Lista sinóptica para el proceso de entrada de datos

- ¿Quiénes son los encargados de la entrada de datos?
- ¿Cuál es la tasa de terminación y de errores cometidos por el personal encargado de la entrada de datos? ¿Necesita este personal algún tipo de readiestramiento?
- Observar el proceso de entrada de datos. ¿Cuál es el sistema utilizado para supervisar el número de cuestionarios asignados a cada entrevistador?
- Examinar el análisis de datos y el cálculo de la matriz de calidad de los datos, y la necesidad de ulterior apoyo.

Cuestionarios

Elegir varios cuestionarios terminados de cada entrevistador y comprobar que:

- Se han borrado los nombres de los cuestionarios.
- La página informativa inicial se ha separado del cuestionario.
- Los hogares de las listas se seleccionaron aleatoriamente de una manera correcta y el registro se completó apropiadamente.
- La escritura es legible y clara.
- Las opciones se han registrado adecuadamente (por ejemplo, las opciones se rodean con un círculo, pero no se marcan con una señal ni se subrayan o se tachan).
- Las preguntas abiertas se contestan, en caso necesario.
- Las preguntas abiertas se registran textualmente.
- Los saltos entre preguntas se efectúan correctamente.
- Las preguntas que deben ser respondidas por las mujeres son respondidas sólo por ellas.

- Doble entrada de datos.
- Uso del programa de entrada de datos:
  - Verificar la confidencialidad y seguridad de los datos.
  - ¿Se realiza la doble entrada de datos?
  - Comprobar la codificación en la base de datos y compararla con la copia impresa.
  - Comprobar si hay errores de intervalo, de coherencia, de itinerario, etcétera.
  - Comprobar la magnitud de los datos omitidos.

G. Análisis de datos

57. Antes del análisis de los datos sustantivos de la WHS, hay varias comprobaciones sistemáticas sobre la calidad de los datos. La compilación de estas comprobaciones se conoce con el nombre de "métrica de la encuesta WHS" y ofrece una serie de indicadores sumarios sobre la calidad de los datos.

58. Los componentes de la métrica de la encuesta son los siguientes:

- Exhaustividad, en la que se incluyen la tasa de respuesta (teniendo en cuenta los hogares cuya admisibilidad quizá sea desconocida, en cuyo caso debe realizarse una estimación sobre la proporción de hogares admisibles o, si estos hogares se excluyen del cálculo de las tasas de respuesta, debe ofrecerse una justificación clara de por qué se considera que esos hogares no tenían informantes admisibles) y los cuestionarios incompletos o la falta de respuesta a partidas completas. La frecuencia de la omisión de datos se calcula por partidas comparando diferentes informantes, y por informantes comparando todas las partidas. Ello ayuda a identificar los problemas de realización de la encuesta, en particular las partidas con dificultades del cuestionario.

- Índice de desviación de la muestra (SDI), que indica hasta qué punto la representatividad de la muestra se desvía de la población objetivo. Si esta medida revela una desviación significativa, el análisis debería estratificarse. El SDI puede evaluarse formalmente utilizando el estadístico chi^2. Si se dan casos de sobremuestra intencionada en algunos subgrupos importantes, debería tenerse en cuenta para ajustar el SDI aplicando el factor de sobremuestreo intencionado.
- Fiabilidad, que indica la posibilidad de reproducir los resultados utilizando el mismo instrumento de medición al mismo informante en diferentes momentos y con diferentes entrevistadores. Este análisis utiliza los datos del protocolo de prueba/comprobación realizado en el 50% de las entrevistas piloto y en el 10% de toda la muestra.
- Comparación con otros encargados de validación de encuestas; en otras palabras, comparación con los resultados de otras encuestas, como el censo, las encuestas y los datos de servicio, además de los del sector privado y público.

59. Estos elementos se explican con mayor detalle en el siguiente apartado. El procesamiento de datos se realiza por países, cuando se dispone de la capacidad necesaria, o en la sede de la OMS.

60. El análisis ulterior de los datos de los países se considera fundamental para garantizar el uso eficaz de los resultados. En la sede y en las oficinas regionales de la OMS se identificarán los países que requieren apoyo para efectuar el análisis completo de los datos y se establecerán mecanismos para dárselo.

Evaluación del análisis de datos

61. La evaluación de este aspecto de la encuesta requiere un examen sobre la disponibilidad de personal especializado en el país capaz de realizar el análisis y sobre el nivel de apoyo que pueden ofrecerse mutuamente los países.

H. Indicadores de calidad

62. Conviene resumir la información sobre la garantía de calidad mediante indicadores. Éstos podrán utilizarse luego para evaluar otros factores contextuales que repercuten en la calidad de la encuesta, con lo que se cerraría el ciclo relativo a la calidad. Por lo que podemos saber, no ha habido un conjunto sistemático de indicadores propuestos para supervisar la calidad de una encuesta en forma sumaria y presentar la información correspondiente. La WHS utiliza algunos indicadores cuantificables que se explican más adelante, y una evaluación cualitativa estructurada mediante un proceso de examen colegiado, en forma de informe, sobre la garantía de calidad.

63. En general, toda encuesta de hogares está sujeta a dos tipos de errores: muestrales y no muestrales. El error muestral se produce porque las encuestas no se aplican a toda la población sino sólo a una muestra. Se ve afectado por el tamaño de la muestra, por la variabilidad registrada en la población en lo que respecta a las cantidades y por otros aspectos del diseño de la muestra, como los efectos de la aglomeración y la estratificación. Por el contrario, los errores no muestrales se ven afectados por factores como la naturaleza de los conceptos asociados a una determinada materia, la precisión y el grado de exhaustividad del marco muestral, la correspondencia entre los procedimientos efectivos de selección sobre el terreno y el diseño muestral deseado y los errores en la realización de la encuesta. Este último factor implica problemas como el mal diseño del cuestionario, los errores del entrevistador al formular las preguntas y los errores o información inadecuada en las respuestas a las mismas, los errores de entrada de datos y otros relacionados con el procesamiento, la falta de

respuesta y las técnicas incorrectas de estimación. A continuación se presentan algunos de los errores no muestrales que se pueden medir y cuantificar.

64. Con respecto al seguimiento del resultado final de los datos de la encuesta, actualmente se utilizan los siguientes indicadores estándar para supervisar la calidad de los datos de la WHS.

1. Índice de desviación de la muestra

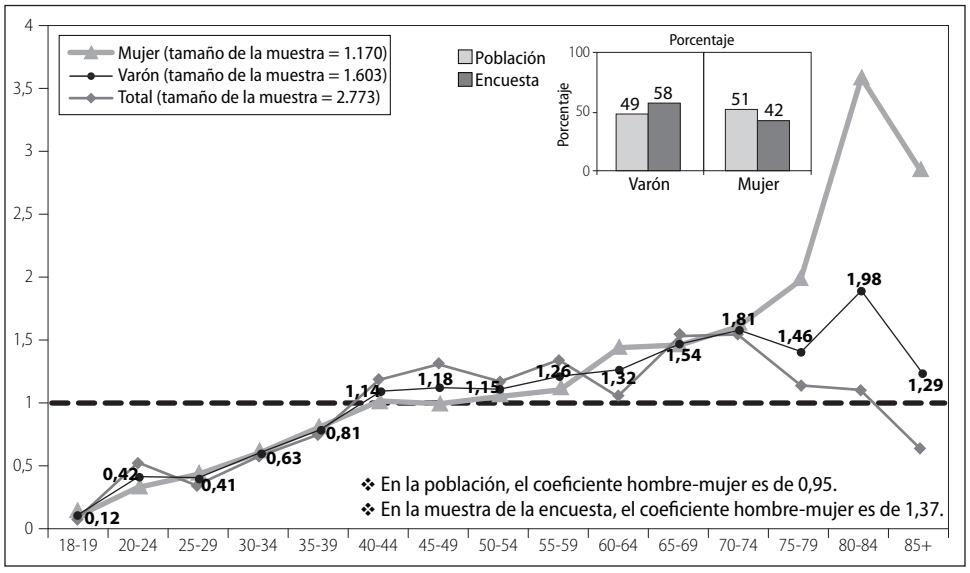
65. El índice de desviación de la muestra (SDI)[3], indica la proporción de estratos de edad y sexo dentro de la muestra en comparación con los datos de población de una fuente independiente, suponiéndose que esta última es la norma. La WHS ha utilizado, como fuente independiente, la base de datos de población de las Naciones Unidas, pero podría utilizarse cualquier otra fuente de datos de población más reciente y fiable. El SDI es un indicador de los datos de la muestra en lo que respecta a su representatividad (es decir, hasta qué punto la muestra representa a la población en general). Un coeficiente de 1 revela que la muestra de la encuesta se corresponde con las características de la población en general con respecto a una categoría por edad o sexo, mientras que las desviaciones de 1 indican errores de muestreo, por exceso o por defecto, con respecto a esa categoría por edad o sexo.

66. El valor previsto de 1 (representatividad ideal) se observa muy pocas veces en las encuestas, debido a errores de muestreo. En la figura X.3 se presenta el SDI de una de las encuestas. En ella se aprecia la representación insuficiente en las edades más jóvenes y la sobrerrepresentación en las más avanzadas, sobre todo en el caso de los hombres de más edad.

2. Tasa de respuesta

67. La tasa de respuesta revela la tasa de terminación de entrevistas en la muestra seleccionada, es decir, el número de entrevistas terminadas entre las personas u hogares que reúnen las condiciones para su inclusión (un "hogar" seleccionado que resulta ser una vivienda abandonada, por ejemplo, no reúne las condiciones). En este indicador se observa el comportamiento de la encuesta con respecto al logro del ideal de una respuesta del 100%. En general se considera como mínimo aceptable una tasa de respuesta del 60%, aunque en la WHS se pide al menos un 75%.

Figura X.3
Ejemplo de un índice de desviación de la muestra



[3] $SDI = \sum_{a=1}^{a} |1 - índice_a|$,

donde a = categorías de edad e *índice* es el coeficiente entre la muestra dentro de esa categoría de edad y la población de la misma categoria de edad tomada de la base de datos de población de las Naciones Unidas u otra fuente actualizada, como el censo nacional. El índice (*index*) revela hasta qué punto la muestra representa a la población en lo que se refiere a la distribución por edad o sexo. El índice puede comprobarse también mediante las pruebas de pi-asterisco (π^*) o el ji^2, para determinar la homogeneidad.

3. Tasa de datos omitidos

68. Por tasa de datos omitidos se entiende la proporción de partidas ausentes en la entrevista de un informante. La WHS mide la proporción de personas que no llegan a completar un mínimo aceptable de partidas (por ejemplo, el 10% en las entrevistas directas en los hogares) para determinar la calidad de las entrevistas. Se identifican también las partidas problemáticas con un alto nivel de respuestas omitidas (más del 5%) por los informantes.

4. Coeficientes de fiabilidad para las entrevistas de prueba-comprobación

69. Los coeficientes de fiabilidad para las entrevistas de prueba-comprobación revelan la estabilidad de la administración de las entrevistas con respecto a la variabilidad de la respuesta en dos ocasiones distintas. Éstas se calculan como tasas de concordancia corregidas en forma aleatoria (es decir, el estadístico kappa para los coeficientes de relación categórica e intraclase en las variables continuas). Este indicador revela hasta qué punto una determinada partida/pregunta de la entrevista obtiene los mismos resultados cuando se repite la misma entrevista. En general, una puntuación de más de 0,4 es aceptable; de más de 0,6 es buena, y de más de 0,8 es considerada excelente (Cohen, 1960; Fleiss, 1981).

70. El principal indicador de la calidad de una encuesta en lo que respecta al error presente en los datos resultante del componente de muestreo es el error estándar estimado de cada uno de los estadísticos clave de la encuesta. Revela el intervalo estimado de error de muestreo (por ejemplo, más o menos el 3%) en torno a una estimación dada. Cuando es posible, se calcula una medida relacionada, a saber, los coeficientes del efecto del diseño para las muestras por conglomerados en varias etapas de la WHS. Este coeficiente es la relación entre la varianza de la muestra efectiva y la de una muestra aleatoria simple hipotética del mismo tamaño. Como una muestra aleatoria simple auténtica no resulta práctica en las encuestas en gran escala, debido a los costos (incluidos los de transporte), normalmente se calcula la varianza de muestreo (cuadrado del error estándar) para compararla con una muestra aleatoria (Kish, 1995b). En general, se considera que un efecto del diseño entre 1 y 6 es aceptable para los indicadores relacionados con la WHS.

I. Informes de los países

71. Una característica importante de la garantía de calidad está relacionada con el producto final en forma de notificación de los datos, debido al efecto de la encuesta en términos de valor añadido a nuestra base de conocimientos y como fuente de nuevas orientaciones para la formulación de políticas. La presentación adecuada de informes está sin duda estrechamente relacionada con la pertinencia de la WHS para las necesidades del país. Los resultados de la WHS se presentarán en informes de diversos tipos, a saber:

a) Informes nacionales para cada uno de los países de la WHS:

i) Resumen ejecutivo para las autoridades y el publico;

ii) Informe detallado para los investigadores y otros usuarios científicos;

b) Informes regionales e internacionales sobre cuestiones específicas.

72. La plantilla inicial para el informe sobre un país [párrafo 71 *a*) *supra*] comprende los siguientes elementos:

- Introducción (por ejemplo, información para orientar las políticas y sobre los sistemas de salud).
- Exposición sobre la realización de la encuesta (por ejemplo, descripción de la encuesta, métodos de muestreo, capacitación, recopilación y elaboración de datos,

procedimientos de garantía de calidad, descripción de los elementos métricos de la encuesta).

- Panorama general de los resultados de la encuesta y repercusiones para la formulación de políticas (por ejemplo, informaciones que se puedan aprovechar para el sistema de salud, características de la población y de los hogares, cobertura de las intervenciones sanitarias, salud de la población, capacidad de respuesta de los sistemas de salud, gastos sanitarios).
- Conclusiones: recomendaciones específicas para la política de salud y seguimiento de los objetivos de desarrollo del milenio en el país.

73. La plantilla se elaborará con mayor detalle en colaboración interactiva con los países, oficinas regionales y otras partes interesadas.

74. Debe elaborarse claramente una estrategia de divulgación para el informe nacional a través de los medios de comunicación, talleres y otros actos. Hay que conseguir que las diferentes partes interesadas utilicen la información generada en los debates normativos a raíz de la encuesta.

75. Los propios países deberán ser los principales responsables de generar sus propios informes nacionales. La OMS ayudará con apoyo técnico y proveyendo los datos fundamentales así como los instrumentos para preparar y examinar estos informes junto con los equipos encargados de su elaboración.

76. La WHS puede ayudar a obtener información sobre diferentes aspectos de la salud de la población y los sistemas sanitarios. Estos elementos incluyen muchos componentes del marco de evaluación de los resultados del sistema de salud. Además, las encuestas aportan información detallada sobre otros aspectos, como factores de riesgo específicos, funciones de los sistemas de salud, epidemiología de enfermedades específicas y servicios sanitarios. Por ello es importante, extraer la mayor información posible de los datos de la WHS.

77. Algunos países quizá deseen utilizar los datos de la WHS para un análisis subnacional. En la mayoría de los casos, para ello se necesitan muestras de mayor tamaño. En otros, los datos de la WHS pueden utilizarse junto con otras fuentes de datos, como el censo y otras encuestas.

78. A la larga, es de prever que la estructura modular de la WHS permita la integración de varias encuestas sobre la salud y los sistemas sanitarios en una sola encuesta.

Evaluación de los informes nacionales

79. El análisis de los datos y la redacción de los informes nacionales representa la culminación de la realización de la encuesta. La calidad de los informes y la forma en que se examinan los resultados determinará de qué manera se aplican las futuras rondas de encuestas y la repercusión que los resultados tendrán en la formulación y seguimiento de las políticas dentro del país.

J. Visitas sobre el terreno

80. Los países de la WHS saben por adelantado qué se espera de ellos en cuanto a la realización de la encuesta y los procedimientos de garantía de calidad. Es importante documentar la labor realizada sobre el terreno. Para lograr este objetivo, la OMS contratará asesores independientes de garantía de calidad que realicen visitas sobre el terreno en cada uno de los países. Estas visitas constituirán de hecho un examen colegiado externo del proceso de realización de la encuesta y registrarán de manera independiente el cumplimiento de las normas de garantía de calidad. Ofrecerán asimismo una oportunidad para averiguar los problemas y solucionarlos lo antes posible. El equipo del país y el asesor de garantía de

calidad prepararán juntos una evaluación estructurada de la calidad global de la encuesta, junto con las directrices de la OMS.

81. La garantía de la calidad es un proceso; no se puede reducir al hecho aislado de una visita. La relación entre los asesores de garantía de calidad y los equipos del país debe entenderse como un proceso a largo plazo en tres fases: antes, durante y después de la visita.

82. Antes de la visita, los países y los asesores de garantía de calidad deben preparar un archivo para la visita en el que se incluirán el formato básico de las directrices de garantía de calidad de la OMS tal como se esbozan en este documento y todos los aspectos de la lista de comprobación para la visita sobre el terreno. En ese archivo se recogerá toda la información básica disponible con respecto al lugar, la institución de la encuesta, el diseño muestral, la experiencia local, los instrumentos y los programas de capacitación utilizados localmente, así como una plantilla para el informe nacional de la WHS. La información no disponible se obtendrá durante la visita.

83. Los responsables del país en la sede de la OMS y los asesores de garantía de calidad estarán en comunicación directa con el investigador principal o el responsable de la encuesta en el país para hacer del proceso de garantía de calidad parte integrante de la realización de la encuesta. De esa manera se creará una cultura de garantía de calidad en las encuestas. El objetivo del proceso de garantía de calidad no es supervisar o controlar, sino lograr la calidad en la WHS ofreciendo asistencia y apoyo.

84. Para que las visitas sobre el terreno produzcan el máximo efecto deberían programarse hacia el final de las actividades de capacitación y al comienzo de la recopilación de datos. En la visita deberían considerarse todos los aspectos del proceso de la encuesta, es decir, diagnosticar los problemas, proponer soluciones, tener en cuenta el contexto local y prestar apoyo y establecer una relación constante.

85. La función de los asesores de garantía de calidad en sus visitas a los países es averiguar los problemas y señalar los aspectos positivos en la realización de la encuesta. Su principal tarea es examinar el proceso de realización de la WHS utilizado por el país e identificar cualquier desviación de las normas de garantía de calidad previstas. Es fundamental su dictamen acerca de si esta desviación es significativa y cómo podría corregirse. Los asesores deberían prestar también apoyo directamente mediante conversaciones con la sede de la OMS o adoptando las medidas oportunas para que otra entidad ofrezca el apoyo pertinente.

86. Los asesores de garantía de calidad realizarán su evaluación de acuerdo con una lista de comprobación estructurada en la que se incluirán los diversos pasos por orden de importancia. Esta evaluación debería incluir el análisis de los aspectos cuantitativos de la encuesta (siempre que se hayan introducido ya algunos datos cuando tenga lugar la visita), en el que se incluyan indicadores sobre la calidad de los datos.

87. La evaluación de la garantía de calidad se examinará conjuntamente con el equipo del país y la OMS. Los países deberán saber por adelantado qué se espera de ellos en lo que respecta a los procedimientos de garantía de calidad.

88. Al informe sobre la visita le sigue el informe nacional de la WHS, que es el producto final de la visita sobre el terreno y del apoyo al país. La visita debería iniciar el proceso de preparación del informe nacional y explorar estrategias específicas para su elaboración, con indicación expresa de cómo utilizar las conclusiones para la formulación de políticas.

K. Conclusiones

89. La garantía de la calidad es una cuestión básica en la realización de una encuesta. Es necesario y posible especificar mecanismos de garantía de calidad en cada paso de una encuesta. Si estos mecanismos se definen en términos operativos, se podrán cuantificar y se podrá supervisar la calidad global de la encuesta.

90. El establecimiento de la garantía de calidad requiere un cambio en la mentalidad de los realizadores de la encuesta, ya que el examen y evaluación de cada uno de los pasos son requisitos obligatorios.

91. Es imprescindible realizar una evaluación constante de los indicadores de calidad durante toda la encuesta. El proceso no debería considerarse meramente como una actividad *post hoc*; debería utilizarse también para introducir correcciones sobre la marcha cuando sea necesario, localizando los problemas e interviniendo adecuadamente. Esta mejora constante de la calidad o gestión de la calidad total en el proceso de producción debe integrarse en todas las encuestas.

92. La disponibilidad de instrumentos informáticos permite ahora establecer un sistema de gestión y seguimiento de la encuesta que haga posible la supervisión constante del proceso, lo que ayuda a considerar los datos con mayor confianza.

93. Es importante documentar las cuestiones críticas (por ejemplo, las relativas a la realización de la encuesta, la capacitación, etcétera) de forma sistemática en lo que se refiere a los informes cualitativos y a los indicadores cuantitativos (a saber, el índice de desviación de la muestra, las tasas de respuesta, la proporción de datos omitidos y la fiabilidad de la prueba-comprobación) con el fin de proveer a los usuarios de los datos información esencial sobre la calidad de una encuesta.

94. El resultado deseado del proceso de garantía de calidad es conseguir una encuesta con datos de la mejor calidad, de modo que los resultados puedan luego clasificarse como válidos, fiables y comparables.

95. La aplicación continuada de estos procedimientos de garantía de calidad permitirá establecer normas para iniciativas internacionales de recopilación de datos. Los métodos para la supervisión de esas normas evolucionarán de forma constante.

Reconocimientos

Deseamos expresar nuestro reconocimiento por la participación de los siguientes expertos de encuestas de varios países e instituciones en la elaboración de las directrices de garantía de calidad de la WHS:

Dr. Farid Abolhassani, República Islámica del Irán

Dr. Sergio Aguilar-Gaxiola, Estados Unidos de América

Dr. Atalay Alem, Etiopía

Dr. Lorna Bailie, Canadá

Dr. Russell Blamey, Australia

Dr. Carlos Gómez-Restrepo, Colombia

Dr. Oye Gureje, Nigeria

Dr. Holub Jiri, República Checa

Sr. Mark Isserow, Sudáfrica

Dr. Feng Jiang, China

Sr. Jean-Louis Lanoe, Francia

Profesor Howard Meltzer, Reino Unido de Gran Bretaña e Irlanda del Norte

Sr. Steve Motlatla, Sudáfrica

Sra. Lipika Nanda, India

Dr. Kültegin Őgel, Turquía

Dr. Gustavo Olaiz Fernández, México

Dr. Mhamed Ouakrim, Marruecos

Dr. Jorun Ramm, Noruega

Dr. Wafa Salloum, República Árabe Siria

Dr. Shen Mingming, China

Dr. Benjamín Vicente, Chile

Consultores de muestreo

Steve Heeringa, Universidad de Michigan, Instituto de Investigación Social, Estados Unidos de América

Nanjamma Chinnappa, India, ex-presidente de la Asociación Internacional de Estadísticas de Encuesta

Asesores regionales de la OMS

Sra. M. Mohale M., Asesora regional para la Oficina Regional de la OMS para África

Dr. Siddiqi Sameen, Asesor regional para la Oficina Regional de la OMS para el Mediterráneo oriental

Dra. Amina Elghamry, Asesora regional para la Oficina Regional de la OMS para el Mediterráneo oriental

Dr. Lars Moller, Asesor regional para la Oficina Regional de la OMS para Europa

Dr. Myint Htwe, Asesor regional para la Oficina Regional de la OMS para Asia sudoriental

Dr. Soe Nyunt-U, Asesor regional para la Oficina Regional de la OMS para el Pacífico occidental

REFERENCIAS BIBLIOGRÁFICAS

Biemer, P. P. y otros, comps. (1991). *Measurement Errors in Surveys*. Nueva York: Wiley.

Bryant, B. E. (1975). Respondent selection in a time of changing household composition. *Journal of Marketing Research*, vol. 12, págs. 129-135.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* vol. 20, págs. 37-46.

De Lepper, M. H., H. Scholten y R. Stern, comps. (1995). *The Added Value of Geographical Information Systems in Public and Environmental Health*. Dordrecht, Países Bajos: Kluwer Academic Publishers.

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*, 2a. ed. Nueva York: John Wiley and Sons.

Kish, L. (1995a). *Survey Sampling*. Nueva York: John Wiley and Sons.

_____ (1995b) Methods for design effects. *Journal of Official Statistics*, vol. 11, págs. 55-77.

Lyberg, L. E. y otros, comps. (1997). *Survey Measurement and Process Quality*. Nueva York: Wiley.

Organización Mundial de la Salud (2000). *Informe Mundial sobre la Salud*. Ginebra: OMS.

_____ (2002). *World Health Survey: Quality Assurance and Guidelines: Procedures for Quality Assurance Implementation by Country Survey Teams and Quality Assurance Advise.* Ginebra: OMS.

Statistics Canada (1998). *Quality Guidelines,* 3a. ed. Ottawa.

Üstün, T. B. y otros (2001). *Disability and Culture; Universalism and Diversity*. Gotinga, Alemania: Hogrefe Huber.

_____ (2003a). WHO Survey Study on Health and Responsiveness 2000-2001. En *Health System Performance Assessment: Debates, Methods and Empiricim* (C.J.L. Murray y D.B. Evans, comps.). Ginebra: OMS.

_____ (2003b). The World Health Surveys. En *Health System Performance Assessment: Debates, Methods and Empiricism* (C. J. L. Murray y D. B. Evans, comps.). Ginebra: OMS.

_____ (2003c). *World Health Organization Disability Assessment Schedule II (WHO DAS II): Development and Psychometric Testing*. Ginebra: OMS. En colaboración con WHO/ National Institute of Health Joint Project Collaborators.

Valentine, N. B., A. de Silva y C. J. L. Murray (2000). *Estimating Responsiveness Level and Distribution for 191 Countries: Methods and Results*. Programa Mundial sobre Pruebas Científicas para las Políticas de Salud, No. 22. Ginebra: OMS.

Capítulo XI

Presentación de informes y compensación de errores no muestrales para encuestas realizadas en el Brasil: práctica actual y desafíos futuros

Pedro Luis do Nascimento Silva
Escola Nacional de Ciências Estadísticas/
Instituto Brasileiro de Geografia e Estatística
(ENCE/IBGE)
Río de Janeiro, Brasil

RESUMEN

En este capítulo se examinan algunas de las prácticas actuales para la presentación de informes y la compensación de errores no muestrales en el Brasil, teniendo en cuenta tres clases de error: errores de cobertura, falta de respuesta y errores de medición y procesamiento. Se identifican también algunos factores que impiden que se preste mayor atención a la medición y control de los errores no muestrales. Además, se señalan algunas iniciativas recientes que podrían ayudar a mejorar la situación.

Términos clave: Proceso de la encuesta, cobertura, falta de respuesta, errores de medición, informe de la encuesta, calidad de los datos.

A. Introducción

1. Antes de nada, conviene definir el concepto de error cuando se aplica a una estadística o estimación de una cantidad (o parámetro) que se desconoce. En ese contexto, error es la diferencia entre la estimación ($\hat{Y}$) y el "verdadero valor del parámetro" teórico (Y) que se obtendría o registraría si se eliminaran todas las fuentes de error. Quizá, como algunos argumentan, sería mejor hablar de desviación (véase Platek y Särndal (2001, sección 5). No obstante, el término error está tan arraigado que seguiremos utilizándolo en estas páginas. Lo que aquí nos interesa son los errores de las encuestas, es decir, los errores de las estimaciones basadas en datos de encuestas. Según Lyberg y otros (1997, pág. xiii), "los errores de las encuestas pueden desglosarse en dos grandes categorías: errores muestrales y errores no muestrales". El examen de los errores de las encuestas, según la terminología moderna, forma parte del debate más amplio sobre la calidad de los datos.

2. Para ilustrar este concepto, supongamos que en una encuesta el ingreso mensual medio de una población determinada se estima en 900 dólares de los EE. UU. y que el verdadero ingreso mensual medio de los miembros de esa población, obtenido a partir de una enumeración completa libre de errores de notificación y elaboración, es de 850 dólares. En este ejemplo, el error de la estimación sería de +50 dólares. En general, los errores de la

encuesta pasan inadvertidos, ya que tampoco se observan los verdaderos valores del parámetro, o incluso son inobservables. Un caso en el que pueden observarse al menos los errores muestrales de las estimaciones estadísticas es el ofrecido por un muestreo basado en registros informáticos en que, si es necesario, se pueden calcular las diferencias entre las estimaciones y los valores obtenidos utilizando la totalidad de los conjuntos de datos. Un ejemplo serían las muestras de uso público de los registros procedentes de un censo de población. En el Brasil las muestras de este tipo se han seleccionado a partir de registros de censos de población desde 1970. No obstante, situaciones como ésta son la excepción, no la norma.

3. Errores de muestreo son las diferencias entre las estimaciones basadas en una encuesta por muestreo y los valores de la población correspondiente que se obtendrían si se llevara a cabo un censo utilizando los mismos métodos de medición, y "se deben al hecho de que se observa una muestra en vez de toda una población" (Särndal, Swensson y Wretman, 1992, pág. 16). "Los errores no muestrales comprenden todos los demás errores" (ibíd.) de una encuesta. Los errores no muestrales pueden ocurrir y ocurren de hecho en encuestas de todo tipo, incluidos los censos. En éstos y en las encuestas que emplean grandes muestras, los errores no muestrales son la principal causa de error que se debe tener en cuenta.

4. Las estimaciones de la encuesta pueden estar sometidas a dos tipos de error: el sesgo y el error variable. Sesgo es un error que afecta al valor previsto de la estimación de la encuesta, y hace que se aleje del verdadero valor del parámetro. Los errores variables afectan a la dispersión de la distribución de las estimaciones de la encuesta en las posibles repeticiones del proceso. En cuanto a los errores muestrales, el sesgo se evita normalmente, o se reduce a niveles insignificantes, utilizando procedimientos de muestreo, tamaños de muestra y métodos de estimación adecuados. Por eso la dispersión es el principal aspecto de la distribución del error muestral que se debe considerar. Un parámetro fundamental que describe esta dispersión es el error estándar, a saber, la desviación estándar de la distribución del error muestral.

5. Los errores no muestrales incluyen dos clases amplias de error (Särndal, Swensson y Wretman, 1992, pág. 16): "errores debidos a la falta de observación" y "errores en las observaciones". Los errores debidos a la falta de observación se producen cuando no llegan a obtenerse los datos necesarios de algunas partes de la población objetivo (errores de cobertura) o de parte de la muestra seleccionada (error de falta de respuesta). Los errores de cobertura o de marco hacen referencia a la inclusión, omisión y duplicación errónea de unidades de la encuesta en el marco muestral, lo que daría lugar a una cobertura inadecuada, por exceso o por defecto, de la población objetivo. Los errores de falta de respuesta se producen cuando no se llegan a obtener datos sobre algunas unidades seleccionadas para la encuesta. Los errores en las observaciones pueden ser de tres tipos: errores de especificación, errores de medición y errores de procesamiento. Según Biemer y Fecso (1995, cap. 15) se producen errores de especificación cuando "1) los conceptos de la encuesta son incuantificables o están mal definidos; 2) los objetivos de la encuesta no están debidamente especificados, o 3) los datos recopilados no corresponden a los conceptos especificados o a las variables deseadas". Los errores de medición se producen cuando los valores observados para las preguntas y variables de la encuesta después de la recopilación de datos difieren de los verdaderos valores correspondientes que se obtendrían si se utilizara un método de medición ideal o "patrón oro". Los errores de procesamiento son los que se introducen al procesar los datos recopilados, es decir, durante la codificación, traspaso, edición, ponderación y tabulación de los datos de la encuesta. Estos tipos de errores se tratan en los subapartados del apartado B, con excepción de los errores de especificación. Su exclusión no significa que no sean importantes, sino únicamente que el examen y el tratamiento de estos errores no están muy adelantados en el Brasil.

6. Otros planteamientos para clasificar los errores no muestrales se examinan en el manual de las Naciones Unidas (véase, Naciones Unidas, 1982). En algunos casos no hay

una línea divisoria clara entre los errores por falta de respuesta, de cobertura y de medición, como en el caso de una encuesta de hogares por muestreo en varias etapas cuando un miembro del hogar no aparece en el empadronamiento: ¿Se trata de un error de medición, de falta de respuesta o de cobertura?

7. Los errores no muestrales pueden dividirse también en varianza no muestral y sesgo no muestral. La varianza no muestral mide la divergencia en las estimaciones de la encuesta si la misma muestra se sometiera a repeticiones hipotéticas del proceso de la encuesta en las mismas condiciones fundamentales (Naciones Unidas, 1982, pág. 20). Los sesgos no muestrales son los errores resultantes del proceso de la encuesta y de sus condiciones, y darían lugar a estimaciones con un valor previsto diferente del verdadero valor del parámetro. Como ejemplo de sesgo no muestral, supongamos que los individuos de una población tienden a infravalorar sus ingresos un 30%, en promedio. Luego, independientemente de los procedimientos de diseño de la muestra y de estimación, sin ninguna información externa las estimaciones del ingreso medio en la encuesta serían, en promedio, un 30% menores que el valor verdadero del ingreso medio de los miembros de la población. En la mayor parte del presente capítulo se estudiará cómo evitar o compensar el sesgo no muestral.

8. Las cuestiones relativas a la calidad de los datos en las encuestas por muestreo han recibido cada vez mayor atención en los últimos años, y varias iniciativas y publicaciones se ocupan de este tema, incluidas algunas conferencias internacionales (véase el apartado D). Por desgracia, el examen se limita todavía predominantemente a los países desarrollados, con poca participación y contribución de los países en desarrollo y en transición. Ésta es la principal conclusión a que se llega después de examinar las actas y publicaciones aparecidas después de estas diferentes conferencias e iniciativas. No obstante, se han publicado recientemente varios documentos sobre este tema en relación con las encuestas en los países en transición en la revista *Statistics in Transition* (Kordos, 2002), pero al parecer su circulación no está muy extendida entre las bibliotecas del mundo en desarrollo.

9. En cuanto a los errores muestrales, existe una teoría unificada de medición y estimación [véase, por ejemplo, Särndal, Swensson y Wretman (1992)], respaldada por la amplia difusión de los métodos de muestreo probabilístico y técnicas como la norma para el muestreo en la práctica de las encuestas (Kalton, 2002), así como por los programas informáticos estándar que permiten la aplicación práctica de esta teoría a las encuestas reales. Si las muestras se toman y recopilan adecuadamente, las estimaciones de la variabilidad muestral de las estimaciones de la encuesta son relativamente fáciles de calcular. Así se está haciendo ya en muchas encuestas en países en desarrollo y en transición, aunque esta práctica dista todavía mucho de ser la norma obligatoria.

10. La divulgación y el análisis de estas medidas de la variabilidad, en cambio, están más retrasados. En muchas encuestas las estimaciones del error muestral ni se calculan ni se publican, o se calculan y publican únicamente con respecto a una pequeña selección de variables/estimaciones. En general, no están disponibles en la mayoría de las estimaciones de las encuestas debido a que suponen un gigantesco esfuerzo de cálculo. En consecuencia, un usuario externo puede tener dificultades para evaluar el grado de variabilidad muestral en una variable concreta, aunque no obstante es posible calibrar su orden de magnitud comparándola con una variable semejante cuyo error estándar se ha estimado. En los comentarios acerca de las estimaciones de la encuesta muchas veces no se tiene en cuenta el grado de variabilidad de las estimaciones. Por ejemplo, la encuesta mensual de mano de obra del Brasil (Instituto Brasileiro de Geografia e Estatística, 2002b), iniciada en 1980, calcula y publica cada mes estimaciones de los coeficientes de variación de los principales indicadores estimados a partir de la encuesta. No obstante, no se estiman los errores estándar con respecto a las diferencias de dichos indicadores entre meses sucesivos, ni con respecto al mismo mes del año anterior. Por otro lado, la mayor parte de los comentarios publicados cada mes junto con las estimaciones hacen referencia al cambio (variaciones en los indicadores mensuales). Estas

estimaciones de los errores estándar para las estimaciones del cambio sólo han comenzado a calcularse recientemente para el análisis interno [véase Correa, do Nascimento Silva y Freitas (2002)], y todavía no se presentan periódicamente a los usuarios externos de los resultados de la encuesta. Lo mismo ocurre cuando las estimaciones son "complejas", como ocurre con las series estacionalmente ajustadas de indicadores del mercado de trabajo.

11. Si la situación dista mucho de ser la ideal en el caso de los errores muestrales, en que tanto la teoría como los programas informáticos están fácilmente disponibles y ha tenido lugar una amplia difusión de la cultura del muestreo, el tratamiento de los errores no muestrales en las encuestas de hogares y de otro tipo en los países en desarrollo está mucho menos avanzado. La falta de una teoría unificadora ampliamente aceptada [véase Lyberg y otros (1997, pág. xiii); Platek y Särndal (2001), y las páginas siguientes)], la falta de métodos estándar para compilar información y estimar los parámetros de los componentes del error no muestral, así como la ausencia de una cultura que reconozca la importancia de medir, evaluar e informar sobre estos errores, hacen que los errores no muestrales y su medición y evaluación reciban menos atención en las encuestas realizadas en los países en desarrollo o en transición. Con ello no se quiere decir que la mayoría de las encuestas realizadas en dichos países sean de baja calidad, sino que sabemos poco acerca de su nivel de calidad.

12. Con esta información básica sobre la situación de la medición y control del error no muestral en las encuestas realizadas en los países en desarrollo y en transición, podemos pasar ahora a examinar qué es lo ocurre de hecho (apartado B) en el Brasil. El examen se limita a lo que podemos observar en un solo país (Brasil), pero creemos que es de interés para los estadísticos de otros países en desarrollo, dado que las publicaciones sobre el tema son escasas. Luego indicamos qué desafíos presenta el futuro para mejorar la práctica de las encuestas en los países en desarrollo y en transición (sección C), de nuevo desde la perspectiva de lo que ocurre en el Brasil.

B. Práctica actual de registro y compensación de los errores no muestrales en las encuestas de hogares en el Brasil

13. En el Brasil, las principales encuestas periódicas de hogares por muestreo con amplia cobertura las realiza el Instituto Brasileiro de Geografia e Estatística (IBGE). Para ayudar al lector a compenetrarse con la referencias a estas encuestas, en el cuadro XI.1 presentamos sus principales características, cobertura y período.

1. Errores de cobertura

14. Se produce un error de cobertura insuficiente cuando se omiten del marco unidades de la población objetivo, que por lo tanto no quedarán abarcadas por la encuesta. La sobrecobertura tiene lugar cuando se incluyen en el marco unidades no pertenecientes a la población objetivo y no hay forma de separarlas de las unidades pertinentes antes del muestreo, o cuando algunas unidades aparecen repetidas en el marco. Los errores de cobertura pueden hacer también referencia a la clasificación errónea de las unidades de la encuesta en los estratos debido a una información muestral inadecuada o desfasada (por ejemplo, cuando se excluye un hogar del proceso de muestreo por considerar que está desocupado, cuando de hecho estaba ocupado en el momento de realización de la encuesta). En las estimaciones de una encuesta, la cobertura insuficiente tiene en general efectos más nocivos que la sobrecobertura. No hay forma de recuperar las unidades omitidas, pero muchas veces, durante el trabajo sobre el terreno o el procesamiento de los datos, es posible localizar las unidades no pertenecientes al universo de la encuesta, lo que permitirá su corrección y ajuste; no obstante, estas unidades ajenas al universo de la muestra provocan un aumento del costo de la encuesta por unidad válida.

Cuadro XI.1
Algunas características de las principales encuestas de hogares por muestreo del Brasil

Nombre de la encuesta	Periodicidad	Cobertura de la población	Tema/materia
Censo de población	Cada 10 años (la última, en 2000)	Residentes en hogares privados y colectivos en el país	Enseres del hogar, estado conyugal fecundidad, mortalidad, religión, raza, educación, mano de obra, ingresos
Encuesta nacional de hogares por muestreo (PNAD)	Todos los años, excepto los del censo	Residentes en hogares privados y colectivos en el país, excepto en las zonas rurales de la región septentrional	Enseres del hogar, religión, raza, educación, mano de obra, ingresos y suplementos especial sobre temas diversos
Encuesta mensual de mano de obra (PME)	Mensual	Residentes en hogares privados y colectivos en seis grandes zonas metropolitanas	Educación, mano de obra, ingresos
Encuesta sobre gastos de los hogares (POF)	1974-1975, 1986-1987, 1995-1996, 2002-2003	Nacional en la edición de 2002-2003; 11 grandes zonas metropolitanas en dos ediciones anteriores; nacional en la edición de 1974-1975	Enseres del hogar, gastos e ingresos familiares
Estudio de medición de los niveles de vida (PPV)	1996-1997	Residentes en hogares privados en las regiones nororiental y sudoriental	Amplia cobertura de temas relacionados con la medición de los niveles de vida
Encuesta sobre la economía informal urbana (ECINF)	1997	Residentes que participan en la economía informal pertenecientes a hogares privados de zonas urbanas	Mano de obra, ingresos y características de la empresa en la economía informal

15. Los problemas de cobertura se consideran muchas veces más importantes en el caso de los censos que en las encuestas por muestreo, ya que, en un censo no hay que preocuparse de los errores de muestreo. No obstante, se trata de una idea equivocada. En algunas encuestas por muestreo la cobertura puede a veces ser un problema tan importante como un error de muestreo, si no mayor. Por ejemplo, las encuestas por muestreo pueden a veces excluir del proceso de muestreo las unidades de algunas zonas de difícil acceso o de categorías que son difíciles de investigar (lo que significaría para ellas una probabilidad cero de inclusión). La razón puede ser por problemas de seguridad para el entrevistador (por ejemplo, cuando hay zonas en conflicto o con alto nivel de violencia) o de costo (por ejemplo, cuando el desplazamiento a algunas partes del territorio para realizar las entrevistas resulta prohibitivamente caro o lleva demasiado tiempo). Si en la definición de la población objetivo no se describen con precisión esas exclusiones, la encuesta resultante dará lugar a problemas de cobertura insuficiente. Estos problemas repercutirán probablemente en las estimaciones en forma de sesgo, ya que las unidades excluidas de la población de la encuesta serán en general diferentes de las incluidas. Cuando la encuesta trata de abarcar estas poblaciones de difícil acceso se necesita una planificación especial para comprobar que la cobertura llega a incluir a esos grupos de la población objetivo, o la población sobre la cual deberán hacer inferencias.

16. Un problema afín es el asociado con algunas encuestas repetidas en países con escasa cobertura telefónica y quizá altas tasas de analfabetismo, donde la recopilación de datos debe basarse en las entrevistas directas. Cuando estas encuestas tienen un período de entrevistas breve, su cobertura puede reducirse en muchos casos a las zonas de fácil acceso. En el Brasil, por ejemplo, la Encuesta mensual de mano de obra (PME) se realiza sólo en seis zonas metropolitanas (Instituto Brasileiro de Geografia e Estatística, 2002b). Su limitada definición de la población objetivo es una de las principales objeciones sobre la pertinencia de esta encuesta: con una población objetivo que es demasiado restringida para muchos usos, no ofrece información sobre la evolución del empleo y el desempleo en otros lugares del país. Aunque en la encuesta se señala correctamente que las cifras hacen referencia a la "población de la encuesta" que vive en las seis zonas metropolitanas, muchos usuarios interpretan erróneamente que las cifras correspondientes a la suma de seis zonas representa la población de todo el Brasil. Está previsto modificar el diseño de la encuesta para abordar este problema en 2003-2004. Se plantean cuestiones semejantes en otras encuestas, como las encuestas sobre ingresos y gastos del Brasil de 1987-1988 y 1995-1996 (cobertura limitada a 11 zonas

metropolitanas) y el Estudio de medición de los niveles de vida (EMNV) de 1996-1997 (cobertura restringida a las regiones nororiental y sudoriental). En menor medida, ocurre lo mismo en la importante encuesta "nacional" anual de hogares por muestreo realizada en el Brasil (Instituto Brasileiro de Geografia e Estatística, 2002a). En esta encuesta no se incluyen las zonas rurales de la región septentrional del Brasil debido a los prohibitivos costos de acceso. Bianchini y Albieri (1998) examinan con más detalle la metodología y cobertura de varias encuestas de hogares realizadas en el Brasil.

17. Se encuentran problemas semejantes en muchas encuestas de otros países en desarrollo y en transición, donde la cobertura frecuente de algunas zonas de difícil acceso del país puede resultar demasiado costosa. Una norma importante a este respecto es que toda publicación basada en una encuesta debe incluir una aclaración explícita sobre la población efectivamente incluida en ella, seguida de una descripción de subgrupos potencialmente relevantes que se han excluido de la misma, en su caso.

18. Las medidas del error de cobertura no se publican periódicamente junto con las estimaciones de la encuesta para permitir a los usuarios externos una evaluación independiente del efecto de los problemas de cobertura en sus análisis. Estas medidas pueden obtenerse sólo cuando las cifras del censo de población se publican cada diez años aproximadamente, e incluso entonces no están directamente vinculadas con el problema de cobertura de las encuestas de hogares realizadas en el decenio precedente.

19. En el Brasil, la única "encuesta" en la que se lleva a cabo un análisis más completo de la cobertura es el censo de población. Éste se realiza normalmente mediante una combinación de encuestas por muestreo posteriores al empadronamiento y un análisis demográfico. La encuesta por muestreo posterior al empadronamiento (EMPE) es un tipo de encuesta realizada sobre todo para evaluar la cobertura de un censo o encuesta semejante, aunque en muchos casos se emplea también para evaluar el contenido de la encuesta. En el Brasil, la EMPE que siguió al censo de población de 2000 realizó una muestra de unas 1.000 zonas de empadronamiento y las investigó utilizando un equipo distinto independiente de empadronadores que tuvo que seguir los mismos procedimientos utilizados por los del censo ordinario. Después de recopilados los datos de la EMPE se procedió a su comparación para ubicar las unidades correspondientes en los datos del censo ordinario. Los resultados de esta armonización se utilizan para aplicar el método de estimación de doble registro [véase, por ejemplo, Marks (1973)], que arroja estimaciones de cobertura insuficiente como las mencionadas en el cuadro XI.2 *infra*. Puede utilizarse también el análisis demográfico de los totales y flujos de la población tomando como base los registros administrativos de nacimientos y fallecimientos, con el fin de comprobar los recuentos de población del censo y evaluar su grado de cobertura. En el Brasil esta práctica consigue resultados positivos en algunos estados del sur y del sureste, donde los registros de nacimientos y fallecimientos son lo bastante precisos como para aportar información útil con este fin.

20. Un impedimento grave para la aplicación generalizada de las EMPE en la estimación y análisis de la cobertura del censo es su costo elevado. Las encuestas deben planificarse y realizarse con gran atención, para que los resultados sean fiables. Asimismo, es importante que ofrezcan resultados desglosados, al menos hasta cierto punto, pues de lo contrario su utilidad será muy limitada. En algunos casos los recursos que se necesitarían para este tipo de encuesta son inexistentes; en otros, los encargados de la planificación del censo quizá consideren más oportuno destinar esos recursos a mejorar la aplicación del censo. No obstante, es difícil, por no decir imposible, introducir mejoras sin haber medido y localizado antes dónde se encuentran los problemas. La EMPE ayuda a sacar a la luz las principales causas o problemas de cobertura y puede ofrecer información acerca de los aspectos de la recopilación de datos que deben mejorarse en el futuro, así como estimaciones de falta de cobertura que podrían emplearse para compensar la cobertura perdida. Por ello, recomendamos encarecidamente que al preparar el presupuesto y al planificar el censo se asignen los

recursos necesarios para una EMPE de tamaño razonable, que se realizaría poco después de la recopilación de datos del censo. La evaluación de la cobertura mediante el análisis demográfico suele ser más barata que una EMPE, pero requiere acceso a fuentes de datos externos y conocimiento de los métodos demográficos. No obstante, siempre que sea posible debería presupuestarse la realización de este tipo de análisis y asignarse tiempo para ello dentro del proceso de evaluación del censo principal.

21. En la mayor parte de los países, desarrollados o no, las cifras del censo no se ajustan en función de la falta de cobertura. La razón es que no hay una teoría o método ampliamente aceptado para corregir los errores de cobertura, o que las estimaciones sobre la cobertura insuficiente resultantes de la EMPE no son lo bastante fiables, o que existen factores políticos que impiden el cambio de las estimaciones del censo. La causa puede ser también una combinación de todos estos y otros factores. Por ello, las estimaciones de población publicadas sobre la base de los datos del censo de población no suelen compensar la falta de cobertura. En algunos casos la información sobre la falta de cobertura del censo, si es que existe, puede considerarse como "reservada" y quizá no sea de libre acceso para los usuarios en general, debido a la opinión de que este tipo de información puede perjudicar la credibilidad de los resultados del censo si se interpreta inadecuadamente. Recomendamos que se acabe con esta práctica y que los resultados de la EMPE se publiquen o se pongan a disposición de las comunidades pertinentes de usuarios del censo.

22. Nos hemos referido hasta ahora a la cobertura general de las poblaciones de la encuesta. Evaluar la cobertura adecuada es aún más grave en el caso de las subpoblaciones de interés especial, como las minorías étnicas o de otro tipo, porque el tamaño de la muestra necesario en una EMPE supera en general el presupuesto disponible. Se sabe muy poco sobre el grado de cobertura de las subpoblaciones en los censos y en otras encuestas de hogares en los países en desarrollo. En el Brasil, ninguna EMPE realizada desde el censo de 1970 ofreció estimaciones sobre los grupos étnicos u otras subpoblaciones pertinentes que podrían ser de interés. Sus estimaciones se han limitado a las omisiones generales de hogares y personas, desglosadas por grandes zonas geográficas (estados). Los resultados de las estimaciones de la cobertura insuficiente para el censo de población de 2000 han aparecido recientemente (Oliveira y otros, 2003). Aquí presentamos únicamente los resultados relativos al país, incluidas las estimaciones para las tasas de omisión de hogares y personas en los censos de 1991 y 2000. Las tasas de falta de cobertura fueron semejantes en 1991 y en 2000, pero ligeramente más bajas en el último año. Una recomendación para mejorar la EMPE realizada en los centros de población del Brasil ha sido ampliar la estimación de la falta de cobertura para incluir las subpoblaciones pertinentes, como las definidas por los grupos étnicos o de edad.

23. Las cifras del cuadro XI.2 son más elevadas que las mencionadas en censos semejantes en algunos países desarrollados. Las tasas de omisión revelan un nivel de falta de cobertura que no es insignificante. Hasta la fecha, los resultados del censo se publican en el Brasil, como ocurre en la gran mayoría de los países, sin ningún ajuste en función de la falta de cobertura estimada. Estos ajustes se introducen posteriormente en las proyecciones de población publicadas después del censo. Es preciso realizar investigaciones para determinar

Cuadro XI.2
Estimaciones de la tasas de omisión en los censos de población del Brasil obtenidas a partir de las encuestas posteriores al empadronamiento de 1991 y 2000 (*porcentaje*)

Categoría de la cobertura	Censo de 1991	Censo de 2000
Hogares privados ocupados	4,5	4,4
Personas que viven en hogares privados ocupados no omitidos	4,0	2,6
Personas omitidas de los hogares privados ocupados	8,3	7,9

Fuente: Oliveira y otros (2003).

el posible efecto del ajuste de las estimaciones del censo en función de la falta de cobertura así como analizar, planificar y decidir la fiabilidad necesaria de las estimaciones de la EMPE, para que puedan utilizarse para tal finalidad.

2. Falta de respuesta

24. Se habla de "falta de respuesta" cuando faltan datos relativos a algunas unidades de la encuesta (falta total de respuesta), a algunas unidades de la encuesta en una o más rondas de una encuesta repetida o por panel (falta de respuesta por oleada) o incluso a algunas variables dentro de las unidades de la encuesta (falta de respuesta sobre una partida). La falta de respuesta afecta a todas las encuestas, sean censos o por muestreo. Puede afectar también a los datos de fuentes administrativas que se utilizan para la elaboración de estadísticas. En la mayoría de las encuestas se emplean algunos procedimientos operacionales para evitar o reducir la incidencia de la falta de respuesta. Ésta constituye un problema mayor cuando la respuesta no es "en forma aleatoria" (falta de respuesta diferenciada entre importantes grupos de subpoblación) y las tasas de respuesta son bajas. Si la falta de respuesta es aleatoria, su principal efecto es una mayor varianza de las estimaciones de la encuesta debida a la reducción del tamaño de la muestra. No obstante, si la participación en la encuesta (respuesta) depende de algunas circunstancias y características de los informantes y/o los entrevistadores, el principal motivo de preocupación es entonces el sesgo, sobre todo en los casos de tasas considerables de falta de respuesta.

25. Särndal, Swensson y Wretman (1992, pág. 575) afirman lo siguiente: "Las principales técnicas para hacer frente a la falta de respuesta son el ajuste de la ponderación y la imputación. El ajuste de la ponderación significa aumentar las ponderaciones aplicadas en la estimación a los valores *y* de los informantes para compensar los valores perdidos debido a la falta de respuesta... Por imputación se entiende la introducción de valores artificiales 'válidos' en sustitución de los valores omitidos".

26. Entre los tres tipos de respuesta, la relativa a las unidades es la más difícil de compensar, ya que usualmente hay muy poca información en los marcos muestrales y en los registros que pueda usarse para ese fin. El método de compensación más frecuente para contrarrestar los efectos negativos de esta falta de respuesta es el ajuste de la ponderación, cuando las unidades informantes ven incrementada su ponderación para tener en cuenta la pérdida de unidades muestrales debido a la falta de respuesta; pero no siempre se aplica ni siquiera este tipo sencillo de compensación. En el caso de falta de respuesta relacionada con oleadas y partidas la compensación se realiza con frecuencia por imputación, ya que en tales casos las unidades que no responden habrán facilitado cierta información que puede utilizarse para orientar la imputación y de esa manera reducir el sesgo (véase Kalton, 1983; 1986).

27. La falta de respuesta tiene varias causas. Puede atribuirse a la falta de contacto con las unidades seleccionadas de la encuesta, debido a factores como la necesidad de presentar las encuestas en forma oportuna, la existencia de hogares que son difíciles de empadronar y el hecho de que los informantes no se encuentren en casa. Puede deberse también a la negativa a cooperar o a la incapacidad de responder o participar en la encuesta. Los casos de negativa a participar son pocos en las encuestas de hogares realizadas en países en desarrollo, debido sobre todo a que como la potenciación ciudadana a través de la educación está menos avanzada, los posibles informantes no tienen mucha disposición ni capacidad de resistirse a cooperar con las encuestas; el nivel más alto de analfabetismo significa que la mayor parte de la recopilación de datos se realice todavía mediante entrevistas directas, a diferencia de las entrevistas telefónicas o de los cuestionarios enviados por correo. Ambos factores contribuyen a reducir las tasas de rechazo o de no cooperación y pueden dar lugar también a una falta de respuesta diferenciada dentro de las encuestas, en las que la población con mayor nivel de instrucción y de riqueza tendría mayor propensión a no responder.

Al mismo tiempo, la respuesta o la participación en la encuesta no significa necesariamente que los informes sean más precisos; en muchos casos, la tasa de respuesta más elevada puede ocultar de hecho informaciones deliberadamente erróneas de algunos datos, sobre todo en las variables relacionadas con los ingresos o el patrimonio, debido a la desconfianza hacia los funcionarios públicos.

28. Los censos de población de los países en desarrollo se ven afectados por la falta de respuesta. En el Brasil, el censo de población utiliza dos tipos de cuestionarios: uno breve, con sólo algunas preguntas sobre cuestiones demográficas (sexo, edad, relación con el jefe del hogar y alfabetización), y otro más amplio y detallado, con contenidos socioeconómicos (raza, religión, educación, trabajo, ingreso, fecundidad, mortalidad, etcétera), en el que se incluyen también todas las preguntas del modelo resumido. El formulario largo se utiliza para los hogares seleccionados mediante una muestra probabilística de hogares en todas las zonas de empadronamiento. La tasa de muestreo es mayor (1 de 5) en los municipios pequeños, y menor (1 de 10) en los municipios con una población estimada de 15.000 o más habitantes en el año del censo. En conjunto, la falta de respuesta de unidades enteras es muy baja en el censo (aproximadamente el 0,8 % en el censo brasileño de 2000). No obstante, en lo que respecta a las variables del formulario breve (las que requieren respuestas de todos los hogares participantes, o "universo") no se introduce ninguna compensación por la falta de respuesta. Las razones son tres: primero, la falta de respuesta se considera bastante baja; segundo, la información disponible sobre los hogares que no responden es insuficiente para adoptar medios de compensación eficaces; tercero, no hay un marco natural para realizar el ajuste de la ponderación en el contexto de un censo. La alternativa a la imputación de los formularios del censo ausentes mediante una especie de donante no está tampoco muy extendida por las dos primeras razones indicadas antes y por el prejuicio añadido contra la imputación cuando se lleva a cabo en casos como éste. En cuanto a las estimaciones obtenidas a partir de la muestra dentro del censo, se introducen ajustes de ponderación basados en los métodos de calibración que compensan parcialmente la falta de respuesta de la unidad.

29. Se ha adoptado un planteamiento semejante en algunas encuestas por muestreo. Dos de las principales encuestas de hogares en el Brasil, la Encuesta nacional de hogares por muestreo (PNAD) (de carácter anual), y la Encuesta de mano de obra (PME) (mensual), no utilizan ningún método específico de compensación por falta de respuesta (véase Bianchini y Albieri, 1998). Los únicos ajustes en las ponderaciones de las unidades informantes se realizan mediante calibración del total de la población en la zona metropolitana o en el estado respectivo, lo que significa que no pueden compensar la falta de respuesta diferencial dentro de los grupos de población definidos por sexo y edad, por ejemplo. Las razones están relacionadas fundamentalmente con consideraciones de tipo práctico, como el mantenimiento de programas informáticos adaptados para fines específicos utilizados para la estimación y que se elaboraron mucho antes y la aparente facilidad de hacer caso omiso de la falta de respuesta. Ambas encuestas registran sus niveles de falta de respuesta, pero ésta no la incluyen en las publicaciones que luego recogen los principales resultados de las dos encuestas. De todas formas, se puede acceder a archivos de microdatos a partir de los cuales pueden realizarse estimaciones sobre la falta de respuesta, ya que en esos archivos se incluyen también los registros de las unidades que no han respondido, con códigos adecuados que especifican los porqués de esa omisión. La PME se remodeló recientemente (Instituto Brasileiro de Geografia e Estatística, 2002b) y comenzó a utilizar al menos un método de reponderación simple para compensar la falta de respuesta observada de las unidades. Otros procedimientos pueden ser la introducción de estimadores de calibración, que tratarán de corregir la falta de respuesta diferencial por edad y sexo. No obstante, los estudios sobre este tema, motivados por la observación de que la falta de respuesta es una de las causas probables del sesgo de los grupos de rotación (Pfeffermann, do Nascimento Silva y Freitas, 2000) en las estimaciones mensuales de la tasa de desempleo, se encuentran todavía en sus comienzos.

30. Una encuesta brasileña que utiliza métodos más avanzados de ajuste es la Encuesta sobre gastos de los hogares (POF) (la última ronda fue en 1995-1996, y actualmente se está llevando a cabo la de 2002-2003). En esta encuesta se utiliza una combinación de métodos de reponderación e imputación para compensar la falta de respuesta (Bianchini y Albieri, 1998). Se realizan ajustes de ponderación para compensar la falta de respuesta de las unidades y se recurre a métodos de imputación de donantes para cubrir las variables o bloques de variables cuya respuesta falta todavía después de la recopilación de datos y del proceso de edición. Se ha prestado más atención al tratamiento de la falta de respuesta por el incremento de respuesta observado en esta encuesta al compararla con las encuestas generales de hogares. Se trata de un resultado previsible, dada la carga mucho mayor impuesta por el tipo de encuesta (los hogares son visitados al menos dos veces y se les pide que anoten detalladamente los gastos a lo largo de dos semanas). En los informes de metodología de la encuesta se ha incluido un análisis de la falta de respuesta, pero no ocurre lo mismo en las publicaciones que presentan los resultados principales.

31. Otra encuesta realizada en el Brasil, la Encuesta de medición de los niveles de vida (PPV), que formaba parte del programa de encuestas del Estudio de medición de los niveles de vida del Banco Mundial, recurrió a sustituir los hogares para compensar la falta de respuesta de las unidades. En el Brasil esta práctica es poco frecuente, y no se ha adoptado en ninguna de las otras grandes encuestas de hogares.

32. Después de examinar estas encuestas llevadas a cabo dentro del mismo país se observa que en general no hay un planteamiento uniforme para compensar la falta de respuesta de las unidades e informar al respecto. Los métodos y el tratamiento de la falta de respuesta varían según cada encuesta, en función de los niveles de falta de respuesta, del seguimiento de las normas internacionales y de la necesidad y capacidad percibidas de aplicar métodos y procedimientos de compensación. Un planteamiento que podría adoptarse para mejorar esta situación es la preparación periódica de informes sobre el "perfil de calidad" para las encuestas de hogares. Esto podría ser más práctico e útil que tratar de incluir toda la información disponible sobre los métodos utilizados y las limitaciones de los datos en el censo básico o en las publicaciones sobre las encuestas.

33. En cuanto a la falta de respuesta sobre una partida, la situación no es muy diferente. En los censos de población del Brasil, a partir de 1980 se han empleado métodos de imputación para cubrir las lagunas y también para sustituir los valores contradictorios localizados mediante normas de edición especificadas por especialistas en las materias respectivas. En 1991 y 2000 se utilizó una combinación de procedimientos de imputación con registro de donante y métodos de Fellegi-Holt, aplicados en programas informáticos como DIA (Detección e Imputación Automática de datos) (García Rubio y Criado, 1990) y NIM (Nueva Metodología de Imputación) (Poirier, Bankier y Lachance, 2001), para realizar una edición e imputación integradas de los formularios breves y largos del censo. En 2000, además de la imputación de las variables categóricas, se realizó también la imputación de las variables de ingreso, mediante métodos del árbol de regresión utilizados para encontrar registros donantes cuyos valores observados en cuanto al nivel de ingreso se utilizaron luego para rellenar las partidas sin respuesta de los registros incompletos. Éste fue el primer censo de población del Brasil en que ninguno de los registros del censo de los archivos de microdatos tenía valores ausentes al final del procesamiento. La estrategia de edición e imputación del censo de población está bien documentada, aunque la mayor parte de la información referente al nivel de edición y de imputación sólo puede obtenerse en informes especializados. Una recomendación para facilitar el acceso a estos informes es difundirlos a través de Internet.

34. El tratamiento de los datos ausentes y sospechosos en otras encuestas de hogares no está tan bien desarrollado. Tanto en la PNAD como en la PME se utilizan programas informáticos para descubrir los errores, pero todavía hay un proceso considerable de "edición manual", y son pocos los casos en los que se utilizan métodos de imputación con ayuda de

computadora para compensar la falta de respuesta sobre partidas concretas. Si hay todavía partidas incompletas al finalizar la fase de edición, se considera que la información es "desconocida". Los progresos realizados en los últimos años se han centrado en la integración de las fases de edición con la entrada de datos, con el fin de reducir el costo y el tiempo de procesamiento. La aparición de computadoras portátiles mejores y menos costosas ha permitido al IBGE avanzar hacia una mayor integración. La PME revisada para el decenio de 2000 comenzó en octubre de 2001 a recopilar una muestra paralela, del mismo tamaño que la empleada en la encuesta ordinaria, en la que los datos se obtienen utilizando entrevistas directas con ayuda de computadora (computadoras de bolsillo). No hay todavía informes finales sobre los resultados conseguidos con estas computadoras de bolsillo, pero después de los primeros meses la recopilación de datos procedía sin problemas. Esta tecnología ha permitido a los responsables de las encuestas mejorar la calidad en el origen, incorporando todas las instrucciones de salto y comprobaciones de validez dentro del instrumento de recopilación de datos, con lo que se evitan errores de traspaso de datos y de otro tipo en la fuente. La falta de respuesta en el caso del ingreso se compensará utilizando métodos del árbol de regresión para encontrar donantes, como en el censo de población. No obstante, los resultados de esta nueva encuesta sólo se han conocido recientemente y la recopilación de datos se realizó en paralelo con la antigua serie durante todo un año antes de que se divulgaran y la nueva serie sustituyera a la anterior. Todavía se está llevando a cabo una evaluación más amplia y más detallada de los resultados de este nuevo planteamiento para la recopilación y el procesamiento de datos.

35. En la PME, cada hogar se mantiene en la muestra durante dos períodos, cada uno de ellos de cuatro meses, separados entre sí por ocho meses. Por ello, en principio, los datos de las entrevistas completas anteriores podrían utilizarse para compensar la falta de respuesta de una oleada siempre que un hogar o un miembro de un hogar quede omitido en una de las rondas de la encuesta posteriores a la primera. Este uso de los datos no se da en la serie antigua ni está previsto para la nueva serie, aunque representa un avance que debería ser tenido en cuenta por los responsables de las encuestas.

36. La pauta emergente de un análisis transversal de encuestas sobre las prácticas de edición e imputación para la falta de respuesta sobre una partida y los datos contradictorios o sospechosos no presenta signos de estandarización, y las diferentes encuestas siguen distintos rumbos metodológicos. Los censos han servido sin duda de ocasión para aplicar en gran escala métodos automáticos de edición e imputación, mientras que las encuestas de menor magnitud no adoptan con tanta frecuencia métodos semejantes. Quizá haya un efecto de escala de la encuesta en el sentido de que la inversión en métodos y procedimientos aceptables para la imputación automática es justificable en el caso de los censos pero no en el de las encuestas menores, que tienen además menos tiempo para conseguir sus resultados. En el caso de una encuesta repetida, como la PME del Brasil, aunque el tiempo para la presentación de resultados es breve, probablemente podrían conseguirse mejoras si se invirtiera más en métodos para la edición e imputación de datos, por la posibilidad que habría de rentabilizar esta inversión en numerosas rondas de encuestas sucesivas.

3. Errores de medición y procesamiento

37. Los errores de medición y de procesamiento implican valores observados en las preguntas y variables de la encuesta después de la recopilación y procesamiento de datos que son distintos de los verdaderos valores correspondientes que se obtendrían si se utilizaran métodos de medición y procesamiento ideales ("patrón oro").

38. Este tema es probablemente el que recibe menos atención en lo que respecta a su medición, compensación y notificación en las encuestas de hogares realizadas en los países en desarrollo y en transición. Varias novedades recientes han permitido mejorar la práctica

de las encuestas y reducir así el error de medición. En primer lugar, el uso de métodos con ayuda de computadora para recopilar los datos ha hecho posible la reducción del error de transcripción, en el sentido de que las respuestas del informante se incorporan directamente a la computadora y están disponibles inmediatamente para su edición y análisis. Asimismo, el flujo de las preguntas se controla mediante la computadora y puede condicionarse a las respuestas, lo que impediría la introducción de errores por el entrevistador. Las respuestas pueden comprobarse con respecto a los intervalos previstos e incluso cotejarse con las respuestas anteriores del mismo informante. Pueden señalarse los datos sospechosos o sorprendentes y se puede pedir al entrevistador que sondee al informante al respecto. Por todo ello, pueden obtenerse datos de mejor calidad y menos sujetos a errores de medición. No obstante, hay pocas pruebas de mejora de la calidad en las entrevistas con ayuda de computadora con respecto a las entrevistas "con lápiz y papel", salvo la ventaja de reducir las tasas de valores ausentes o que quedan fuera de los intervalos previstos.

39. Otro progreso ha sido el desarrollo y aplicación de programas informáticos generalizados para la edición e imputación de datos (Criado y Cabria, 1990). Como se ha mencionado en el apartado B, los censos de población han adoptado programas automatizados de edición e imputación para descubrir y compensar el error de medición y algunos tipos de errores de procesamiento (por ejemplo, errores de codificación y perforación) y, al mismo tiempo, la falta de respuesta sobre partidas concretas. Así ha ocurrido en algunas encuestas por muestreo. No obstante, el tipo de compensación que se aplica dentro de este enfoque sólo puede resolver los llamados errores aleatorios. Los errores sistemáticos rara vez afloran o compensan para utilizar programas informáticos de edición.

40. Otro tipo de progreso que puede dar lugar a la reducción de errores de procesamiento en las encuestas ha sido el desarrollo de programas de codificación con ayuda de computadora, así como los programas y equipo de captación de datos.

41. La prevención de los errores de medición y de procesamiento puede que haya experimentado algunos progresos, pero no ocurre lo mismo con la aplicación de los métodos para medir, compensar en su caso y notificar los errores de medición. La práctica relativa a los errores de medición muchas veces se centra en la prevención, y una vez realizado lo que se considera importante en este sentido ya no se presta mucha atención a la evaluación de la eficacia con que se ha planificado y ejecutado la encuesta. La falta de una teoría rectora estándar sobre la medición dificulta la tarea de establecer objetivos de calidad y evaluar el logro de los mismos. Por ejemplo, aun cuando existen algunos planes de muestreo en los que el tamaño de la muestra se ha definido con el fin de conseguir coeficientes de variación (errores estándar relativos) de determinadas estimaciones clave por debajo de un valor determinado previamente, pocas veces se encuentran planes de recopilación y procesamiento que traten de mantener los niveles de imputación de partidas por debajo de un nivel especificado, o que se propongan observar las medidas dentro de una tolerancia especificada (es decir, desviación máxima) con respecto a los "verdaderos valores", con un alto nivel de probabilidad. Quizá sea imposible en la práctica establecer por adelantado resultados cuantitativos realistas para todos los tipos de error no muestral; no obstante, proponemos que las organizaciones encargadas de las encuestas hagan un esfuerzo para medir los errores no muestrales y utilicen esas mediciones para fijar objetivos de mejora en el futuro y supervisar la consecución de dichos objetivos.

C. Desafíos y perspectivas

42. Después de más de 50 años de difusión generalizada de las encuestas (por muestreo) como instrumento clave de observación en las ciencias sociales, el concepto de error de muestreo y su control, medición e interpretación han alcanzado cierto nivel de madurez a pesar de que, como hemos observado, los resultados de muchas encuestas en todo el mundo

se publican sin incluir estimaciones sobre el error de muestreo. Son muchos menos los progresos conseguidos en lo que respecta a los errores no muestrales, al menos en las encuestas llevadas a cabo en los países en desarrollo. Ello no ha sido fruto de la casualidad. El problema de los errores no muestrales en las encuestas es de difícil solución. En primer lugar, proceden de numerosas causas. Los esfuerzos por combatir un determinado tipo de error muchas veces sólo consiguen incrementar errores de otro tipo. Los métodos de prevención dependen de la tecnología, pero también de la cultura y del entorno, lo que hace muy difícil generalizar y divulgar las experiencias positivas. Los métodos de compensación suelen ser complejos y su aplicación adecuada resulta costosa. La medición y la evaluación son tareas difíciles en un contexto de encuestas realizadas con presupuestos muy limitados y con plazos de publicación cada vez más imperiosos, para poder atender la creciente demanda de nuestras sociedades hambrientas de información. En un contexto semejante, es correcto que se dé siempre prioridad a la prevención, más que a la medición y a la compensación, pero eso deja poco espacio para evaluar hasta qué punto fueron eficaces los esfuerzos de prevención y reduce, por tanto, las perspectivas de mejoras futuras.

43. Algunos usuarios que pueden tener escaso conocimiento de las cuestiones estadísticas quizá interpreten erróneamente los informes acerca de los errores no muestrales en las encuestas. Por ello, la publicación de los informes de este tipo algunas veces se considera como poco deseable en algunos contextos, sobre todo por la falta general de cultura y conocimientos estadísticos, cuyo desarrollo puede resultar particularmente difícil entre las poblaciones con bajos niveles de alfabetización y de conocimientos numéricos, como es el caso en muchos países en desarrollo. Ocurre también con frecuencia que en los organismos encargados de la preparación de las encuestas no se dispone de suficiente personal especializado en estadística, lo que dificulta el reconocimiento de los problemas y la adopción de medidas decididas para contrarrestarlos, y además no permite cuantificar hasta qué punto han sido eficaces las medidas adoptadas. En cualquier caso, alentamos a que se preparen y publiquen los informes sobre errores no muestrales, y a que los organismos de estadística se esfuercen por hacerlos lo más claros y accesibles que sea posible para adultos con un nivel mínimo de instrucción.

44. Aun cuando el panorama no sea muy favorable, algunas novedades son alentadoras. La atención reciente al tema de la calidad de los datos en algunos importantes organismos de estadística, asociaciones académicas de estadística y encuestas e incluso organizaciones gubernamentales multilaterales es un gran avance. Las principales iniciativas a que nos referiremos aquí son el Sistema General de Divulgación de Datos (GDDS) y la Norma Especial de Difusión de Datos (SDDS) del Fondo Monetario Internacional (FMI), que están tratando de promover la estandarización de los informes sobre la calidad de los datos estadísticos mediante la adhesión voluntaria de los países a una de esas dos iniciativas. Según Fondo Monetario Internacional (2001): "El GDDS es un proceso estructurado a través del cual los países miembros del Fondo se comprometen voluntariamente a mejorar la calidad de los datos elaborados y divulgados por los sistemas de estadística a largo plazo con el fin de atender las necesidades de análisis macroeconómico". Igualmente, según la misma fuente, "el GDDS fomenta prácticas estadísticas acertadas con respecto a la compilación y la divulgación de estadísticas económicas, financieras y sociodemográficas. Identifica conjuntos de datos que son de especial interés para el análisis económico y la evolución social y demográfica y establece objetivos y recomendaciones referentes a su elaboración, producción y divulgación. Se presta particular atención a las necesidades de los usuarios, que se abordan mediante directrices relativas a la calidad e integridad de los datos y al acceso del público a éstos" (Ibíd.).

45. La principal contribución de estas iniciativas es ofrecer a los países: *a*) un marco para la calidad de los datos (www.dsbb.imf.org/dqrsindex.htm) que ayude a identificar las principales zonas problemáticas y los objetivos para mejorar la calidad de los datos; *b*) el in-

centivo económico para considerar la mejora de la calidad de los datos dentro de una amplia gama de encuestas y productos estadísticos (en forma de renovación o consecución de acceso a los mercados internacionales de capital); *c*) una comunidad que comparta una motivación común a través de la cual pueda promover el debate sobre la calidad de los datos sin temor a interpretaciones erróneas, y *d*) apoyo técnico para los programas de evaluación y mejora, en caso necesario. Ello no representa una iniciativa universal, ya que no todos los países son miembros del FMI. No obstante, se establecieron contactos con 131 países y, según las últimas cifras, 46 han decidido adherirse al GDDS y otros 50 han conseguido un considerable nivel de adhesión a la SDDS, tras haber superado un conjunto de controles y criterios rigurosos de evaluación de la calidad de sus productos estadísticos.

46. No entra dentro de los objetivos de este capítulo examinar detalladamente las normas de calidad de los datos promovidas por el FMI u otras organizaciones. De todas maneras, alentamos al lector a que examine este tema utilizando las referencias aquí mencionadas. Los países en desarrollo deberían sumarse al debate sobre las normas actualmente en curso, decidir si van a intentar o no adherirse a una de las iniciativas mencionadas y, en su caso, contribuir a la definición y revisión de las normas. La mayor parte de los organismos estadísticos importantes de los países en desarrollo pueden utilizar estas normas como puntos de partida (si no existe nada semejante a nivel local) para promover una mayor conciencia sobre la calidad tanto entre sus miembros y funcionarios como dentro de las comunidades de usuarios.

47. La otra iniciativa que mencionaremos aquí, en particular porque afecta al Brasil y a otros países de América Latina, es el Proyecto de Cooperación Estadística de la Unión Europea (EU) y el Mercado Común del Sur (MERCOSUR)[1]. De acuerdo con el objetivo del proyecto, "la Unión Europea y los países del MERCOSUR han firmado un acuerdo sobre 'Cooperación estadística con los países del MERCOSUR', cuyo principal objetivo es una aproximación[2] de los métodos estadísticos con el fin de hacer posible el uso de los diferentes datos estadísticos basados en términos mutuamente aceptados, en particular los referentes a los bienes y servicios objeto de comercio y, en general, cualquier esfera que pueda ser objeto de medición estadística". Se prevé que el proyecto "consiga al mismo tiempo la estandarización de los modelos estadísticos entre los países del MERCOSUR y entre ellos y la Unión Europea" (véase información más detallada en www.ibge.gov.br/mercosur/englishindex.html). Este proyecto ha promovido ya varios cursos y seminarios de capacitación y de esa manera está contribuyendo a mejorar la práctica de las encuestas y a lograr una mayor sensibilización sobre los errores de las mismas y su medición.

[1] MERCOSUR es el Mercado Común del Sur, grupo de países con un tratado de libre comercio integrado por Brasil, Argentina, Paraguay y Uruguay.

[2] Este término se utiliza aquí con el significado de armonización.

48. Iniciativas como ésta son fundamentales para ayudar a los organismos de estadística de los países en desarrollo a mejorar su situación: sus estadísticas quizá sean de buena calidad, pero muchas veces no saben hasta qué punto lo son. La cooperación internacional de los países desarrollados con los países en desarrollo y también entre estos últimos es fundamental para progresar hacia una mejor medición y notificación sobre los errores no muestrales y otros aspectos de la calidad de los datos de las encuestas.

D. Recomendaciones sobre lecturas complementarias

Entre las reuniones recomendadas cuyas actas y publicaciones de su trabajo pueden servir como lectura complementaria se incluyen las siguientes:

- Conferencia Internacional sobre errores de medición en las encuestas, celebrada en Tucson, Arizona, en 1990 (véase Biemer y otros, 1991).
- Conferencia internacional sobre la medición de las encuestas y el proceso de calidad, celebrada en Bristol (Reino Unido) en 1995 (véase Lyberg y otros, 1997).

- Conferencia internacional sobre la falta de respuesta en las encuestas, celebrada en Portland, Oregón, en 1999 (véase Groves y otros, 2001).
- Conferencia internacional sobre la calidad en las estadísticas oficiales, celebrada en Estocolmo (Suecia) en 2001 (véase www.q2001.scb.se).
- Statistics Canada Symposium 2001, celebrado en Ottawa (Canadá), cuyo tema principal fue el logro de la calidad de los datos en un organismo de estadística desde una perspectiva metodológica (véase www.statcan.caenglishconferences/symposium2001/session21/s21c.pdf).
- Quincuagésimotercera reunión del Instituto Internacional de Estadística (ISI), celebrada en Seúl (República de Corea) en 2001, donde se presentó un documento sobre "Programas de calidad en los organismos de estadística", en la que se examinaba cómo se ocupaban las oficinas nacionales e internacionales de estadística de la calidad de los datos (véase www.nso.go.kr/isi2001).
- Seminario sobre calidad de las estadísticas, 2000, patrocinado por el FMI, celebrado en la isla de Jeju (República de Corea) en 2000 (véase www.nso.go.kr/sqs2000/sqs12.htm).
- Conferencia internacional para mejorar las encuestas, celebrada en Copenhague (Dinamarca) en 2002 (www.icis.dk/).

REFERENCIAS BIBLIOGRÁFICAS

Bianchini, Z. M. y S. Albieri (1998). A review of major household sample survey designs used in Brasil. En *Proceedings of the International Conference on Statistics for Economic and Social Development*. Aguascalientes, México, 1998: Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Biemer, P. P. y R. S. Fecso (1995). Evaluating and controlling measurement error in business surveys, Cox y otros, comps. En *Business Survey Methods*, Nueva York: John Wiley and Sons.

Biemer, P. P. y otros (1991). *Measurement Errors in Surveys*. Nueva York: John Wiley and Sons.

Correa, S. T., P. L. do Nascimento Silva y M. P. S. Freitas (2002). Estimação de variância para o estimador da diferença entre duas taxas na pesquisa mensal de emprego. En *15º Simpósio Nacional de Probabilidade e Estatística*. Aguas de Lindóia (Brasil), São Paulo (Brasil): Associação Brasileira de Estatística.

Criado, I. V. y M. S. B. Cabria (1990). *Procedimiento de depuración de datos estadísticos,* cuaderno 20. Vitoria-Gasteiz, España: Instituto Vasco de Estadística (EUSTAT).

Fondo Monetario Internacional (2001). Guía del Sistema General de Divulgación de Datos (GDDS). Washington, D.C.: Departamento de Estadística del FMI. Disponible en www.dsbb.imf.org/vgn/images/pdfs/spaguide.pdf).

García Rubio, E. e I. V. Criado (1990). DIA System: software for the automatic imputation of qualitative data. En *Proceedings of the United States Census Bureau Sixth Annual Research Conference* (Arlington, Virginia). Washington, D.C.: Oficina del Censo de los Estados Unidos.

Groves, R. M. y otros (2001). *Survey Non-response*. Nueva York: John Wiley and Sons.

Instituto Brasileiro de Geografía e Estatística (2002a). www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad99/metodologia99.shtm.

_____(2002b).http://www.ibge.net/home/estatistica/indicadores/trabalhoerendimento/pme/default.shtm.

Kalton, G. (1983). *Compensating for Missing Survey Data*. Research Report Series. Ann Arbor, Michigan: Instituto de Investigación Social, Universidad de Michigan.

_____ (1986). Handling wave non-response in panel surveys. *Journal of Official Statistics*, vol. 2, No. 3, págs. 303-314.

_____ (2002). Models in the practice of survey sampling (revisited). *Journal of Official Statistics*, vol.18, No. 2, págs. 129-154.

Kordos, J. (2002). Comunicaciones personales.

Lyberg, L. y otros, comps. (1997). *Survey Measurement and Process Quality*. Nueva York: John Wiley and Sons.

Marks, E. S. (1973). *The role of dual system estimation in census evaluation*. Informe interno. Washington, D.C.: Oficina del Censo de los Estados Unidos.

Naciones Unidas (1982). *National Household Survey Capability Programme: Non-sampling errors in household surveys: sources, assessment and control*: *Preliminary Version*. DP/UN/INT-81-041/2. Nueva York: Departamento de Cooperación Técnica para el Desarrollo.

Oliveira, L. C. y otros (2003). *Censo Demográfico 2000: Resultados da Pesquisa de Avaliação da Cobertura da Coleta*. Textos para Discussão, No. 9. Río de Janeiro: IBGE, Directoria de Pesquisas.

Pfeffermann, D., P. L. do Nascimento Silva y M. P. S. Freitas (2000). *Implications of the Brasilian Labour Force rotation scheme on the quality of published estimates*. Informe interno. Río de Janeiro: IBGE, Departamento de Metodología.

Platek, R. y C. E. Särndal (2001). Can a statistician deliver? *Journal of Official Statistics*, vol. 17, No. 1, págs. 1-20.

Poirier, P., M. Bankier y M. Lachance (2001). Efficient methodology within the Canadian Census Edit and Imputation System (CANCEIS). Paper presented at the Joint Statistical Meetings, American Statistical Association.

Särndal, C. E., B. Swensson y J. Wretman (1992). *Model Assisted Survey Sampling*. Nueva York: Springer-Verlag.

Sección D

Costo de las encuestas

Introducción

James Lepkowski
Universidad de Michigan
Ann Arbor, Michigan
Estados Unidos de América

1. En las secciones anteriores se han examinado los errores muestrales y no muestrales resultantes de las encuestas de hogares con el fin de comprender mejor la calidad de las estimaciones de las encuestas. En casi todos los tipos de estos errores hay métodos que pueden utilizarse para reducir su magnitud. No obstante, la aplicación de esos métodos supone costos adicionales. Como las encuestas tienen presupuestos fijos para cubrir los gastos, la asignación de recursos adicionales para reducir una fuente de error significa transferir recursos de una esfera a otra. El diseño de las encuestas supone una constante transacción entre costos y errores.

2. Por ejemplo, supongamos que en una determinada encuesta de hogares hay un subgrupo de la población que habla un idioma al que no se ha traducido el cuestionario de la encuesta. Los encargados de su diseño pueden decidir inicialmente excluir a este grupo, con lo que se crearía un problema de cobertura. Otra posibilidad alternativa es decidir reducir el tamaño de la muestra para contener los costos y luego utilizar los ahorros para traducir el cuestionario a un nuevo idioma, contratar entrevistadores que lo hablen e introducir de nuevo a esos hogares en la encuesta.

3. Dado que el diseño de la encuesta supone con frecuencia la búsqueda de soluciones de compromiso, con el fin de tomar decisiones acertadas, es necesario disponer de información válida sobre el carácter y el tamaño de los errores resultantes de las diferentes causas (como la varianza de muestreo y sesgo por falta de cobertura, en el ejemplo del párrafo anterior) y sobre los costos asociados con los diferentes procedimientos de la encuesta. En las secciones anteriores se han examinado las fuentes de error y su magnitud. En ésta se examinará el carácter de los costos de la encuesta.

4. En una encuesta el problema de los costos se plantea en tres niveles. El primero es el de la fase de planificación, cuando los costos deben estimarse por adelantado. Es difícil estimar el costo en la fase de planificación o de presupuestación si no se tiene experiencia previa. Las operaciones habituales de la encuesta pueden ser una interesante fuente de datos sobre los costos para planificar nuevas rondas de la encuesta, aunque muchas veces deben tenerse también en cuenta consideraciones sobre los costos en el próximo nivel: el seguimiento de los costos de la encuesta.

5. Las organizaciones que se dedican a realizar encuestas, e incluso las que lo hacen de forma ocasional, no suelen tener sistemas desarrollados de seguimiento de los costos hasta el punto de utilizar los datos sobre los costos para la planificación. Los costos se integran en un sistema contable, pero estos sistemas no clasifican los costos de acuerdo con las categorías que el encargado de la encuesta necesitaría para la planificación. En los casos en los que se

intenta el seguimiento de los costos, ello puede representar un gasto adicional si los nuevos sistemas deben agregarse a las operaciones.

6. Si los costos se están supervisando en una operación continuada, es posible considerar de manera más sistemática las modificaciones del diseño de la encuesta durante la recopilación de datos. La información sobre los costos puede utilizarse para hacer proyecciones sobre la magnitud de los ahorros en una operación y sobre el efecto de la reasignación de recursos a otra esfera.

7. La reasignación de recursos en la planificación de una encuesta se determina considerando las soluciones de compromiso entre el nivel de costos y el error entre los diferentes motivos de error. La elaboración del diseño de la muestra es una esfera en la que pueden realizarse, y se realizan de hecho, compromisos formales para buscar una solución óptima al problema de la asignación de recursos.

8. Por ejemplo, como se ha examinado en el capítulo II, las encuestas que están basadas en conglomerados procedentes de una muestra probabilística de la zona a partir de una población muy dispersa deben considerar la posibilidad de limitar el número de conglomerados con el fin de reducir los costos de la recopilación de datos. La limitación del número de conglomerados significa, no obstante, que hay que aumentar también el número de observaciones realizadas en cada conglomerado de la muestra a fin de mantener el tamaño global de la misma. De todas formas, este incremento del tamaño de la submuestra en cada conglomerado incrementa también la variabilidad de las estimaciones de la muestra. En otras palabras, a medida que se reducen los costos, limitando los conglomerados, aumenta la varianza del muestreo. Lo que se necesita es orientación sobre la forma de seleccionar muchos conglomerados de manera que se puedan reducir los costos, dado que es preciso alcanzar un nivel especificado de precisión, o que la varianza de muestreo debe mantenerse lo más pequeña posible con un costo dado. En el diseño de la muestra hay una solución matemática para este problema.

9. La solución de compromiso entre costo y error se plantea asimismo en otros aspectos del diseño de la encuesta. Por ejemplo, un método para reducir la falta de respuesta de los hogares en una encuesta es visitar más de una vez los hogares sobre los que no se ha podido obtener respuesta con una sola visita. Se pueden dar instrucciones al entrevistador para que durante el período de recopilación de datos de la encuesta visite al menos cuatro o cinco veces los hogares con el fin de obtener respuesta. Las visitas repetidas a algunos hogares de la muestra reduce el número de hogares de la muestra que se pueden incluir en ella. El costo de la repetición de las visitas para reducir la falta de respuesta de los hogares limita el tamaño de la muestra. El costo de intensificación de los esfuerzos por reducir la falta de respuesta con el fin de reducir también el sesgo de la falta de respuesta aumenta, por lo tanto, la varianza muestral. Una vez más, los esfuerzos para reducir los costos en una esfera suponen la reasignación de recursos e introducen la posibilidad de aumentar el error en otra esfera del diseño de la encuesta.

10. En los capítulos de esta sección se consideran varias cuestiones centradas en torno a la planificación, el seguimiento y la reasignación de los costos en el diseño de la encuesta. Se utilizan datos de las encuestas de hogares en países en desarrollo y en transición para ilustrar tanto los tipos de costos contraídos en la recopilación de datos de la encuesta como, hasta cierto punto su magnitud. Como las operaciones de las encuestas varían tanto de un país a otro, y más entre continentes, la información específica sobre los costos aquí ofrecida quizá no sea útil para planificar una encuesta en un país dado. No obstante, se espera que los motivos y niveles de los costos que se presentan en los capítulos siguientes ayuden a los encargados de diseñar las encuestas en diferentes contextos a comprender las soluciones de compromiso entre costos y error en sus propias encuestas.

Capítulo XII
Análisis de las cuestiones relacionadas con los costos de las encuestas en los países en desarrollo y en transición

Ibrahim S. Yansaneh*
Comisión de Administración Pública Internacional
Naciones Unidas, Nueva York

* Ex jefe de la Unidad de Metodología y Análisis, División de Estadítica de las Naciones Unidas.

RESUMEN

En este capítulo se examinan, en términos generales, las cuestiones clave relacionadas con el costo del diseño y realización de encuestas de hogares en los países en desarrollo y en transición. El costo global de una encuesta se desglosa en componentes más detallados asociados con los diversos aspectos de su diseño y realización. Se consideran por separado los factores relativos a los costos en los países con una amplia infraestructura para las encuestas y los que tienen poca o ninguna infraestructura en ese terreno. Se examina también la cuestión de la comparabilidad de los costos entre diferentes países.

Términos clave: Infraestructura para las encuestas, costo incremental por entrevista, eficiencia, comparabilidad de los costos, factores de costo.

A. Introducción

1. Criterios para diseños muestrales eficientes

1. En general, un diseño muestral eficiente debe cumplir uno de estos dos requisitos: ofrecer estimaciones razonablemente precisas en un contexto de restricciones presupuestarias o reducir el costo de ejecución con un determinado nivel de precisión. En el presente capítulo se tiene en cuenta el primer criterio, es decir, la elaboración del diseño más eficiente que pueda realizarse con costos que estén en consonancia con los presupuestos disponibles y permitan un uso razonablemente eficiente de los recursos. En los países en desarrollo y en transición el costo de las encuestas es uno de los mayores obstáculos para la formulación de decisiones críticas sobre el diseño y la realización. En esos países, lo mismo que en los desarrollados, el diseño de una muestra supone las habituales transacciones entre precisión de las estimaciones y costo de la realización. La precisión suele medirse en función de las varianzas de los estimadores de las cantidades de la población seleccionada que se consideran de especial interés. Otras medidas relacionadas de la precisión son el error cuadrático medio o error total de la encuesta, que incorpora también el componente de sesgo del error.

2. El desarrollo matemático formal de las soluciones de compromiso entre precisión y costo normalmente supone la optimización de las funciones adecuadas de costo o de va-

rianza sujetas a restricciones relativamente sencillas. No obstante, debido a limitaciones de la información disponible sobre el costo y la varianza, este planteamiento de optimización muchas veces debe interpretarse como una mera aproximación remota al diseño preferido, y a los valores de precisión y costo que se conseguirán de hecho en la realización. Estas cuestiones se han examinado en profundidad en relación con las encuestas realizadas en los países desarrollados. Véanse, por ejemplo, Andersen, Kasper y Frankel (1979), Cochran (1977), Groves (1989), Kish (1965; 1976) y Linacre y Trewin (1993) y las referencias citadas en sus bibliografías. Además, en de Vries (1999, pág. 70) y en su bibliografía puede verse un debate más amplio sobre el costo y precisión en cuanto dos de los principales criterios para la evaluación de los sistemas estadísticos nacionales. Por lo que respecta a los análisis empíricos de los costos de determinadas encuestas en países en desarrollo y en transición y a un examen más detallado de las soluciones de compromiso entre costo y error en el diseño de las encuestas en los países en desarrollo y en transición, véanse los capítulos XIII y XIV, *infra*, y la introducción a toda esta sección.

3. Una importante limitación en el diseño de encuestas en países en desarrollo y en transición es la falta o insuficiencia de información sobre los costos asociados con los diversos aspectos de la realización de la encuesta. A pesar de esa limitación, muchas veces se encuentra en los costos de las diferentes encuestas cierta estructura común que puede ser útil para el diseño de una nueva encuesta. En algunos casos esta estructura común se limita a indicaciones cualitativas de las magnitudes relativas de varios componentes o motivos de costo. En otros se conocen los costos efectivos, que pueden considerarse bastante homogéneos en un conjunto de países, en particular si presentan semejanzas en cuanto a las distribuciones de la población y los niveles de infraestructura para las encuestas.

4. En este capítulo se realiza un análisis de las cuestiones relativas al costo en el contexto de las encuestas en los países en desarrollo y en transición y se investiga hasta qué punto pueden utilizarse los componentes de costos y otros afines de un país para mejorar el diseño de una encuesta semejante de otro país. En otras palabras, se trata de abordar la cuestión de la transferibilidad de los costos de las encuestas en diferentes países. La utilidad de este análisis es doble. En primer lugar puede ofrecer una solución parcial al problema de la escasez de información sobre el costo de las encuestas en los países en desarrollo y en transición. En segundo lugar, en la medida en que hay semejanzas entre los países en cuanto al diseño de la muestra, a la infraestructura para las encuestas y a las distribuciones de población, serían previsibles semejanzas al menos en algunos de los componentes del costo de las encuestas entre los distintos países. Esta información sobre el costo puede extraerse de una encuesta en un país y utilizarse para diseñar una nueva encuesta en otro, o para mejorar la eficiencia del diseño de la misma encuesta en el mismo país. En tal caso el responsable de diseñar la encuesta debe reconocer la gran variabilidad de las estructuras de costo de las encuestas en los diferentes países. Los componentes de costo variables suelen ser específicos de cada país, pero es probable que algunos costos fijos sean comparables entre los países.

2. Componentes de las estructuras de costo en las encuestas de los países en desarrollo y en transición

5. Nos fijaremos de entrada en el primer criterio para conseguir un diseño muestral eficiente; es decir, un diseño que genere estimaciones razonablemente precisas con una determinada asignación presupuestaria. Muchas encuestas realizadas en los países en desarrollo y en transición las encargan organismos internacionales financieros y de desarrollo que necesitan datos para tomar decisiones sobre los proyectos de asistencia al desarrollo o para ayudar a los responsables de la toma de decisiones y a las autoridades de los países beneficiarios. Tres ejemplos notables de encuestas de países en desarrollo son las encuestas demográficas y de salud (DHS) realizadas por Macro para la Agencia de los Estados Unidos

para el Desarrollo Internacional; las encuestas del Estudio de medición de los niveles de vida (EMNV), realizadas por el Banco Mundial, y las encuestas agrupadas de indicadores múltiples (MICS), del Fondo de las Naciones Unidas para la Infancia (UNICEF). Muchas otras encuestas son realizadas en forma periódica por las oficinas nacionales de estadística y otros organismos pertenecientes a los sistemas nacionales de estadística. Hay también un gran número de encuestas en menor escala encargadas por los donantes y realizadas por pequeñas organizaciones locales (por ejemplo, organizaciones no gubernamentales). Obviamente, el tema del costo es igualmente fundamental en el diseño de estas encuestas menores.

6. En lo que respecta a los costos, es importante reconocer que los diseños de las encuestas en los países en desarrollo tienen muchas características comunes. Por ejemplo, la mayoría de ellas están basadas en un diseño probabilístico estratificado por áreas en varias etapas. Las unidades primarias de muestreo se construyen con frecuencia a partir de las zonas de empadronamiento identificadas y utilizadas en un censo nacional de población precedente. Las unidades secundarias de muestreo suelen ser unidades de habitación u hogares, y las unidades últimas de muestreo son en general hogares o personas. Los estratos y campos analíticos se forman normalmente a partir de la intersección de regiones administrativas y subcampos urbanos/rurales de estas regiones. Debido a estas semejanzas, y habida cuenta de los estudios mencionados en el párrafo 2, es interesante considerar hasta qué punto se pueden identificar estructuras de costos comunes dentro de los grupos de encuestas de países en desarrollo. En la sección A de la parte I (Diseño de la encuesta) y en los estudios de casos incluidos en la parte II de esta publicación se encuentra información básica sobre el diseño y realización de encuestas en los países en desarrollo y en transición. En el capítulo XIII se examinan con más detalle los componentes de costo de una encuesta específica. En el capítulo XIV pueden verse comparaciones empíricas de los componentes de costo de las encuestas realizadas en algunos países en desarrollo y en transición.

7. En este capítulo nos limitaremos a considerar las grandes encuestas de hogares nacionales realizadas por las oficinas regionales de estadística u otros organismos gubernamentales del sistema nacional de estadística. Entre ellas se incluyen las encuestas sobre el presupuesto de los hogares, las encuestas sobre ingresos y gastos y las encuestas demográficas y de salud. Aun cuando las encuestas de mercado y otras encuestas de hogares en menor escala realizadas por varias organizaciones con carácter ocasional representan una fuente útil de información y se tienen en cuenta en las políticas y los planes de desarrollo nacionales, no se examinan en estas páginas. De todas formas, las cuestiones clave planteadas en el debate se aplican también a ese tipo de encuestas. La mayor parte de los ejemplos se basa en las encuestas DHS y EMNV, pero los aspectos fundamentales son aplicables en general a todas las encuestas de hogares.

3. Panorama general de este capítulo

8. El capítulo está organizado como sigue: En el apartado B se examina el desglose clásico del costo general de una encuesta en componentes más detallados. En los tres apartados que le siguen se hace una descripción cualitativa de algunos factores que influyen en los costos globales de las encuestas realizadas en los países en desarrollo y en transición. En el apartado C se analizan los factores de costo que pueden ser importantes para los casos en que se ha establecido ya una importante infraestructura para encuestas; en el D se consideran los casos en que la infraestructura es limitada o nula, y en el E se analizan los cambios en la estructura de costos que pueden resultar de las modificaciones introducidas en los objetivos de las encuestas. El apartado F contiene observaciones cautelares acerca de la interpretación de los costos registrados. El apartado G recoge algunas observaciones finales y un resumen de los puntos más destacados que no se han desarrollado plenamente en el debate. El capítulo se cierra con un anexo que presenta un ejemplo de marco utilizado en la presupuestación de

las encuestas por conglomerados de indicadores múltiples (MICS) del UNICEF hechas en países en desarrollo y en transición, tal como puede verse en Ajayi (2002).

B. Componentes de costo de una encuesta

9. Las bases matemáticas de los costos de una encuesta presuponen en general un costo global, *C*, en cuanto función lineal de los números de algunas unidades primarias de muestreo y de algunos elementos seleccionados. Un ejemplo de esta función sería el siguiente:

$$C = c_0 + \sum_{h=1}^{L} n_h \; c_h + \sum_{h=1}^{L} \sum_{i=1}^{n_h} n_{hi} \; c_{hi} \qquad (1)$$

donde c_0 representa los costos fijos de iniciar una encuesta; c_h es igual al costo incremental de recopilar información de una unidad primaria de muestreo (UPM) adicional dentro del estrato h; n_h es el número de UPM incluidas en la muestra; c_{hi} es igual al costo incremental de entrevistar un hogar adicional dentro de la UPM i en el estrato h; y n_{hi} es el número de hogares de la muestra en la UPM i. Véanse, por ejemplo, Cochran (1977, apartados 5.5 y 11.13-11.14) y Groves (1989, cap. 2). En general, los coeficientes de costo c_0, c_h y c_{hi} dependerán de un gran número de factores, que pueden variar según los países y las diferentes encuestas dentro de cada país. Estos factores se examinan con detalle en los apartados siguientes.

10. Téngase en cuenta que la expresión (1) es una de las muchas funciones de costo que podrían considerarse. Por ejemplo, Cochran (1977, pág. 313) examina la inclusión de un componente de costo independiente asociado con la enumeración de las unidades secundarias de muestreo (como etapa intermedia antes del submuestreo de hogares para la entrevista) dentro de unidades primarias seleccionadas, en que dicho componente depende del número de unidades secundarias de cada unidad primaria. Asimismo, en el caso de un diseño en varias etapas, es decir, en que se seleccionan aleatoriamente las personas de los hogares a las que se hará la entrevista, habrá un término adicional en la expresión (1) anterior, que denotaría el costo incremental asociado con el hecho de entrevistar a una persona adicional dentro de un hogar seleccionado.

11. Además, muchas veces una función de costo realista suele ser más una función escalonada que lineal. Por ejemplo si pueden realizarse 10 entrevistas en un solo día, la adición de una undécima entrevista requiere un día adicional de trabajo y, por lo tanto, un costo añadido considerable, mientras que la adición de una duodécima entrevista quizá represente sólo un pequeño aumento del costo total. Asimismo, es importante observar que las decisiones sobre cuestiones como el número de UPM de la muestra dependen a veces de consideraciones prácticas, además de las relacionadas con el costo y la precisión. Por ejemplo, puede ocurrir que alguien desee dedicar toda una semana a realizar entrevistas en una UPM. En tal caso no sería viable establecer un volumen de trabajo de menos de una semana, pero sería posible un volumen de trabajo doble que equivaliera a dos semanas. En tal situación, el número de UPM de la muestra no dependería directamente de los costos y de los efectos del diseño, sino de motivos prácticos relacionados con la ejecución.

12. En el siguiente apartado examinamos los costos de las encuestas atendiendo al nivel de infraestructura del país en cuestión. El mensaje central es que hay una enorme disparidad en los costos globales de las encuestas entre los países que tienen una considerable infraestructura y los que carecen de ella. No obstante, conviene recordar que en los países en desarrollo y en transición habría que evaluar el grado de infraestructura en la fase de planificación de una encuesta en vez de fiarse de los registros históricos. No es raro que un país con magnífica infraestructura para las encuestas en un determinado momento experimente un deterioro constante a lo largo del tiempo hasta el punto de pasar del primer grupo de países (considerados en el apartado C) al segundo (examinados en el apartado D).

C. Costos de las encuestas cuando se dispone de una infraestructura considerable

1. Factores relacionados con los preparativos

13. Gran parte del costo de una encuesta que se realiza una sola vez corresponde al financiamiento de las actividades de preparación [véase, por ejemplo, Grosh y Muñoz (1996, pág. 199)], cuyos fondos se desembolsan en una fase inicial del proceso. Los preparativos con costos relativamente fijos son la coordinación de la planificación de la encuesta por diferentes organismos gubernamentales, la elaboración del marco, el diseño de la muestra, el diseño del cuestionario, la impresión de cuestionarios y otros materiales de la encuesta y la publicidad destinada a los posibles informantes. Los costos de las actividades preparatorias que dependen del tamaño de la muestra (sea en relación con la unidad primaria o con el hogar) comprenden la contratación y capacitación del personal sobre el terreno (por ejemplo, los encargados de la confección de listas, entrevistadores, supervisores y traductores).

14. Los costos de las actividades preparatorias dependen de factores locales como la magnitud del personal de la encuesta y las tasas de remuneración, el tipo y volumen de equipo, los precios de partidas como el material de oficina y otros suministros y las formas de transporte y comunicación. Dependen también mucho de si la encuesta es un estudio transversal realizado por primera vez —en cuyo caso los costos unitarios son relativamente más altos— o forma parte de una encuesta continuada, cuyos costos unitarios son más bajos.

2. Factores relacionados con la recopilación y procesamiento de datos

15. Los costos de la recopilación y procesamiento de los datos implican también componentes fijos y variables, pero los costos de la recopilación de datos son en su mayor parte variables, es decir, dependen del número de unidades primarias de muestreo y de hogares seleccionados. Entre estos costos figuran los de elaboración de listas de hogares dentro de las unidades primarias seleccionadas o de personas dentro de los hogares elegidos, la realización de las entrevistas y la supervisión sobre el terreno. El costo de la recopilación de datos incluye también el de los gastos de viaje, tanto entre diferentes UPM como dentro de ellas. Estos costos de recopilación de datos dependen de la organización de las operaciones de entrevistas, la longitud del cuestionario, la necesidad de utilizar intérpretes y el número de unidades que se deben entrevistar.

16. Una opción para reducir los gastos de viaje es crear equipos nacionales integrados por supervisores y entrevistadores y trasladar a los equipos de una región a otra, en vez de establecer equipos regionales. Es importante observar que esta opción mejora también la calidad de los datos. Este planteamiento puede ser asimismo útil en situaciones en que la recopilación de datos se realiza de forma continua, o cuando las operaciones de la encuesta suponen el uso de equipo costoso. El modelo de equipos de encuesta múltiples se ha utilizado en muchas encuestas de los países en desarrollo y en transición, como la serie de encuestas del EMNV (Grosh y Muñoz, 1996, cap. 5). Cuando en esos países hay diferencias lingüísticas regionales, quizá sea más eficiente organizar los equipos teniendo en cuenta su dominio de la lengua hablada en cada región.

17. Una parte significativa de los costos de la recopilación y procesamiento de los datos está relacionada con la coordinación de las actividades sobre el terreno y los materiales de la encuesta. En un sistema centralizado de recopilación y procesamiento de datos, los costos asociados con la recuperación de los cuestionarios terminados y su transmisión a la sede pueden ser considerables. Además, el presupuesto debe tener en cuenta los costos potencialmente significativos asociados con el seguimiento de las actividades y resultados de la encuesta, como los procedimientos de elaboración de listas y de submuestreo llevados a

cabo sobre el terreno, las tasas de respuestas sobre los principales campos de interés con respecto a niveles previamente especificados, etcétera. El seguimiento eficaz de estas actividades permite a los encargados de realizar las encuestas adoptar las medidas correctivas oportunas durante la recopilación de datos en vez de descubrir las deficiencias una vez terminada esa recopilación, cuando su corrección podría resultar prohibitivamente costosa.

18. Dentro del procesamiento de los datos, la labor de entrada, edición e imputación puede suponer una combinación de costos fijos y variables, según el grado de automatización del proceso. Los otros costos principales del procesamiento de datos son probablemente fijos, e incluyen el equipo y programas de computación, y la formulación de ponderaciones y estimadores de la varianza y otras labores de análisis de datos. Por ejemplo, las ponderaciones podrían calcularse independientemente del número de UPM o de hogares incluidos en la muestra, y una vez que se haya establecido y programado un procedimiento de ponderación, el costo incremental de calcular una ponderación para un hogar adicional sería insignificante.

19. El costo del procesamiento de los datos depende de cuántos niveles de análisis se incluyen en el presupuesto. En el caso de algunas encuestas sólo se lleva a cabo un análisis preliminar sobre los datos recogidos en forma de cuadros en otras encuestas. En otras, como las DHS y las EMNV, se realizan análisis estadísticos más detallados como base para las recomendaciones sobre políticas de los gobiernos beneficiarios y organismos donantes. Por ejemplo, tanto en el contexto de las DHL como de las del EMNV se realizan varios tipos de análisis detallados sobre sus microdatos, y se publican sus conclusiones en una serie de informes analíticos y metodológicos (en el caso de las DHS) y documentos de trabajo (en el de las EMNV). En varias de las referencias que se citan más adelante se incluyen algunos ejemplos. Se contraen también costos considerables en la preparación de informes y divulgación de los resultados y en relación con varios servicios para otros analistas, en los que se puede incluir la preparación de metadatos y la organización de seminarios de capacitación.

D. Costos de las encuestas con poca o nula infraestructura previa

20. En un país con relativamente poca infraestructura previa para la realización de encuestas es probable que el organismo patrocinador tenga que destinar una cantidad considerable de recursos a actividades de fortalecimiento de la capacidad, que no serían necesarias en un país con mejor infraestructura (Grosh y Muñoz, 1996, cap. 8). Los costos de las actividades preparatorias, de operaciones sobre el terreno y de elaboración de datos pueden aumentar de forma sustancial si no se dispone de la infraestructura necesaria.

21. El fortalecimiento de la capacidad supone en general una importante labor de capacitación inicial del personal. En un país con poca o nula infraestructura para las encuestas, a diferencia de otro con una infraestructura bien desarrollada, se incurre normalmente en considerables costos asociados con el uso de personal especializado externo necesario para la realización de la encuesta. Además, el tiempo del personal sobre el terreno suele utilizarse con mayor eficiencia a medida que la organización gana experiencia. Asimismo, en los países con considerable historial en este terreno, la necesidad de desplazamiento es mucho menor, ya que los organismos de estadística tendrán probablemente equipos regionales habituados a recopilar datos o dotarán al personal sobre el terreno de los medios de transporte necesarios. Estas ventajas dan lugar a ahorros en los costos de transporte, capacitación y otros gastos de personal. Los países sin historial de realización de encuestas normalmente incluyen los vehículos en el presupuesto de la encuesta, partida que puede representar una parte considerable del costo general (Grosh y Muñoz, 1996, cap. 8). Otros casos de partidas presupuestarias en las que repercute notablemente la existencia de una infraestructura o de un historial previo de realización de encuestas son el equipo informático y los mapas para la localización de los hogares.

E. Factores relacionados con las modificaciones en los objetivos de la encuesta

22. Como se ha visto anteriormente, muchos factores de costo están vinculados con características del diseño de la encuesta, como el tamaño de la muestra, la longitud del cuestionario, el número de módulos y los métodos específicos empleados en la selección de la muestra y la confección de listas, las comprobaciones piloto y el diseño y traducción del cuestionario. Con un diseño dado, algunos de los costos resultantes son aproximadamente constantes en los diferentes países.

23. No obstante, los diseños de la encuesta en los países en desarrollo y en transición muchas veces tienen que modificarse para tener en cuenta algunas especificaciones especiales de los gobiernos beneficiarios o de otras partes interesadas. Por ejemplo, es posible que un gobierno decida ampliar los objetivos de la encuesta para incluir ciertas prioridades nacionales. Eso puede generar la inclusión de módulos adicionales en el cuestionario o un aumento del número de campos de información si las variables fundamentales de los grupos nacionales se desea que tengan el mismo nivel de precisión que las estimaciones de alcance nacional.

24. Esas modificaciones pueden repercutir de varias maneras en las soluciones de compromiso entre costos y calidad de los datos. En primer lugar, pueden ocasionar directamente aumentos significativos del tiempo total que requiere el entrevistador para la recopilación de datos debido a una longitud media mayor de las entrevistas por la inclusión de módulos adicionales en el cuestionario [párrafo 23 *a*)] o por un incremento, por órdenes de magnitud, del número de entrevistas como consecuencia del mayor número de campos de información [párrafo 23 *b*)]. En segundo lugar, si una organización dispone de un número relativamente fijo de entrevistadores y supervisores bien capacitados, las modificaciones pueden dar lugar a mayores costos debido a la necesidad de capacitar entrevistadores adicionales y al mayor tiempo de supervisión necesario por minuto de tiempo de entrevista. La otra posibilidad es que el número de personal bien capacitado pueda mantenerse constante con una secuencia dual de un período prolongado de recopilación de datos y, por lo tanto, mayores costos. En tercer lugar, los aumentos mencionados pueden generar un crecimiento de la magnitud del error no muestral con respecto al error muestral. Por ejemplo, la inclusión de módulos adicionales en un escenario puede multiplicar el error no muestral debido a la falta de comprobación de las cuestiones o a la fatiga de los informantes. El error no muestral puede aumentar también debido al uso de un mayor número de entrevistadores relativamente poco experimentados, por un incremento del número de entrevistas o por la longitud media de éstas.

F. Algunas advertencias sobre los informes relativos a los costos de las encuestas

25. Es preciso tener en cuenta varios factores para garantizar que las comparaciones de costos entre las diferentes encuestas y países se realicen sobre una base razonablemente común. En primer lugar, las encuestas de los países en desarrollo y en transición están patrocinadas por organizaciones diferentes, que muchas veces tienen políticas y procedimientos contables distintos. En el caso de algunos organismos patrocinadores, puede ser importante distinguir entre el costo para el organismo patrocinador y el costo global de realización de la encuesta.

26. En segundo lugar, puede ser importante contabilizar en forma comparativa el apoyo ofrecido en especie; por ejemplo, vehículos para el transporte del personal sobre el terreno. En algunos casos la ayuda en especie puede provenir de la oficina nacional de estadística, por medio de una asignación de su personal permanente sobre el terreno a una encuesta

patrocinada por instituciones internacionales. Puede considerarse que estos costos son en especie y no deben incluirse en el presupuesto detallado pero que representan un costo de oportunidad en la medida en que la realización de la encuesta es una actividad adicional que quita tiempo a otras posibles actividades que podrían ser realizadas por la oficina nacional de estadística.

27. Podría decirse algo semejante acerca de la prestación de asistencia técnica externa. Esta partida puede ser especialmente importante en los países sin infraestructura o historial de realización de encuestas. Esta asistencia técnica es ofrecida en muchos casos en especie por organismos internacionales que realizan o patrocinan las encuestas, y por eso no se incluye directamente en el presupuesto de la encuesta. De todas formas, esta asistencia técnica algunas veces se subcontrata. En ese caso tampoco se incluye en el presupuesto. Por ejemplo: la encuesta de Turkmenistán de 1998, inspirada en el modelo EMNV, se realizó con asistencia técnica del Research Triangle Institute (RTI) mediante contrato con el Banco Mundial.

28. En tercer lugar, debido a la estructura jerárquica de los costos (expresión 1) como se observa en el apartado B, es importante distinguir entre el costo total de una encuesta y el costo por entrevista terminada. Por ejemplo, debido a la disponibilidad de mayores recursos y al creciente interés por contar con estimaciones fiables registradas en el plano subnacional, los grandes países en desarrollo y en transición suelen utilizar muestras de mayor tamaño en sus respuestas (Fondo de las Naciones Unidas para la Infancia, 2000, cap. 4). Debido a los altos costos asociados con el transporte y los salarios de un número mayor de empleados, las encuestas de los grandes países suelen tener costos totales más elevados que las que se realizan en los de menor tamaño. En cambio, los grandes países con mayores costos generales pueden a veces tener costos más bajos por entrevista terminada, debido a las economías de escala y a la distribución de los costos fijos entre una muestra más amplia.

29. En cuarto lugar, la evaluación de los costos globales y por entrevista puede complicarse debido a las características especiales del diseño de la muestra. Por ejemplo, los costos pueden sobrevalorarse debido al sobremuestreo o a la selección de muestras para garantizar el logro de los objetivos de precisión en determinadas subpoblaciones que son reducidas o difíciles de identificar a partir de la información del marco (por ejemplo, los hogares con niños de menos de cinco años). Finalmente, en el caso de las encuestas de poblaciones con tamaños de hogares muy variables puede ser también importante distinguir entre los costos por hogar contactado y los costos por entrevista realizada.

G. Resumen y observaciones finales

30. La mayor parte de las encuestas de los países en desarrollo y en transición se realizan en un contexto de fuertes restricciones presupuestarias y de incertidumbre acerca de la ejecución del presupuesto aprobado. Por ello, el análisis de los factores que influyen en el costo de las encuestas es uno de los aspectos más importantes del proceso de diseño y planificación para los países en desarrollo y en transición. En este capítulo se ha presentado un marco para este análisis y se ha examinado también hasta qué punto las cifras sobre el costo de la encuesta y los componentes afines pueden transferirse de un país a otro cuando haya semejanzas con respecto al diseño de la encuesta y a la distribución de la población de los hogares, o por otros factores.

31. Para ilustrar las cuestiones clave se han utilizado las encuestas nacionales en gran escala, pero este examen podría aplicarse a muchos otros tipos de encuestas de menor tamaño realizadas dentro de los sistemas nacionales de estadística en los países en desarrollo y en transición. En la medida en que es posible identificar estructuras de costos comunes en estas encuestas, se puede utilizar la información sobre los componentes del costo de una encuesta en un país con el fin de obtener orientaciones útiles para el diseño de una encuesta

semejante en otro país, o para aumentar la eficiencia del diseño de una nueva encuesta en el mismo país. Se ha señalado que hay una gran discrepancia en los costos de las encuestas entre los países con una amplia infraestructura para las encuestas en el momento de su planificación y los que tienen poca o ninguna infraestructura. También se ha hecho hincapié en algunas advertencias que deberían tenerse en cuenta al hacer comparaciones entre los costos globales de las encuestas en diferentes países.

32. Concluimos reiterando algunos puntos importantes relacionados con el costo de las encuestas en los países en desarrollo y en transición:

a) Aun cuando un análisis atento de los componentes de costos pueda revelar estructuras de costos comunes en los diferentes grupos de países o de encuestas, es preciso reconocer que los presupuestos de las encuestas muchas veces son sólo válidos para un país concreto y para un momento dado. Por ello es importante compilar datos sobre los costos y preparar un informe administrativo en que se documenten los distintos componentes del costo de cada etapa de la encuesta para cada una de las encuestas de hogares. El mismo tipo de información debería documentarse en el caso de las varianzas y sus componentes. Esta información sobre los costos y las varianzas puede ser útil en dos sentidos: primero, para adoptar importantes decisiones sobre los presupuestos y la gestión, y segundo, para demostrar cómo las distintas decisiones sobre el diseño de la muestra han acusado los efectos de los diferentes componentes de los costos y de las varianzas. En general, la documentación de los costos y las varianzas y sus componentes en cada una de las etapas del proceso de la encuesta debería formar parte integrante de los procedimientos operativos habituales de las oficinas nacionales de estadística en los países en desarrollo y en transición;

b) Aun cuando el costo global de la encuesta incorpore aspectos fijos y variables, son los variables los que deben controlarse y manipularse con mayor atención en el proceso de diseño de una encuesta. Algunos costos fijos, como los de coordinación de la planificación de la encuesta por múltiples organismos gubernamentales o los de la publicidad orientada hacia posibles informantes, escapan muchas veces del control de quien diseñó la encuesta, y en cualquier caso están demasiado vinculados con el país, el tiempo y la encuesta especifica en cuestión;

c) Como se examina en el capítulo XIV, hay una diferencia en las consideraciones sobre el presupuesto entre las encuestas pagadas por el usuario y las presupuestadas por el país. Mientras que las primeras están bien diseñadas y suelen llevarse a cabo sin grandes sobresaltos y con todos los componentes críticos pagados por adelantado, las segundas suelen estar sometidas a restricciones y asignaciones presupuestarias de un país. En estas últimas muchas veces hay una gran disparidad entre el presupuesto planificado y el presupuesto efectivo, que no depende de consideraciones relativas a la precisión, sino de la disponibilidad de fondos para la encuesta, a veces desviados hacia otras prioridades presupuestarias del país;

d) Debido a las fuertes restricciones presupuestarias con que se llevan a cabo la mayor parte de las encuestas de los países en desarrollo y en transición, para quienes diseñan las encuestas es importante explorar procedimientos no monetarios de presupuestación para una encuesta o poder llevar a cabo algunos aspectos de la misma sin necesidad de presupuesto. Por ejemplo, quizá fuera posible compartir la infraestructura con una encuesta ya existente, utilizar una submuestra de unidades ya seleccionadas para otra encuesta o hacer que un entrevistador recopile datos para varias encuestas. Debería tenerse también en cuenta la posibilidad de presupuestar determinados aspectos de una encuesta teniendo en cuenta el tiempo necesario para llevarlos a efecto.

e) En las páginas anteriores hemos mantenido que el costo de una encuesta aumenta significativamente si un país no cuenta con la debida infraestructura y la capacidad general para la realización de estadísticas. Por ello, la creación y fortalecimiento de esa infraestructura es una inversión valiosa que podría dar lugar a largo plazo a presupuestos más bajos para las encuestas en los países en desarrollo y en transición. Uno de los planteamientos más eficaces para desarrollar esa infraestructura y para promover el desarrollo general de las estadísticas es la cooperación técnica entre las oficinas nacionales de estadística de los países en desarrollo y en transición y las de los sistemas estadísticos más avanzados, en colaboración con los organismos internacionales de estadísticas y financiamiento y otras partes interesadas. No obstante, con el fin de obtener resultados positivos para los países beneficiarios, estos esfuerzos de cooperación técnica deben estar bien orientados y aplicados. Las Naciones Unidas (1998, anexo) elaboraron orientaciones prácticas para la cooperación técnica en el terreno de las estadísticas, que fueron ratificadas por la Comisión de Estadística de las Naciones Unidas en su trigésimo período de sesiones, el 4 de marzo de 1999.

Reconocimientos

Agradezco los positivos comentarios de tres relatores y de participantes en la reunión del Grupo de Expertos sobre el análisis de las características operativas de las encuestas en los países en desarrollo y en transición, que tuvo lugar en la Sede de las Naciones Unidas en Nueva York en octubre de 2002, que ayudaron a mejorar notablemente el primer borrador de este capítulo. No obstante, las opiniones aquí expresadas son mías y no reflejan necesariamente las de las Naciones Unidas.

REFERENCIAS BIBLIOGRÁFICAS

Ajayi, O. O. (2002). *Budgeting framework for surveys.* Comunicación personal.

Andersen, R., J. Kasper y M. R. Frankel (1979). *Total Survey Error.* San Francisco, California: Jossey-Bass.

Cochran, W. G. (1977). *Sampling Techniques,* 3a. ed. Nueva York: Wiley.

de Vries, W. (1999). Are we measuring up? Questions on the performance of national statistical systems. *International Statistical Review,* vol. 67, págs. 63-77.

Fondo de las Naciones Unidas para la Infancia (2000). *End-Decade Multiple Indicator Cluster Survey Manual*. Nueva York: Fondo de las Naciones Unidas para la Infancia.

Grosh, M. E. y J. Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No. 126. Washington, D.C.: Banco Internacional de Reconstrucción y Fomento, Banco Mundial.

Groves, R. M. (1989). *Survey Errors and Survey Costs.* Nueva York: Wiley.

Kish, L. (1965). *Survey Sampling.* Nueva York: Wiley.

_____ (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society, Series A,* vol. 139, págs. 80-95.

Linacre, S. J. y D. J. Trewin (1993). Total survey design: application to a collection of the construction industry. *Journal of Official Statistics,* vol. 9, págs. 611-621.

Naciones Unidas (1998). "Algunos principios rectores sobre buenas prácticas relativas a la cooperación técnica en estadística". Nota de la Secretaría. E/CN.3/1999/19. 15 de octubre.

Yansaneh, I. S. y J. L Eltinge (2000). Design effect and cost issues for surveys in developing countries. *Proceedings of the Section on Survey Research Methods.* Alexandria, Virginia: American Statistical Association, págs. 770-775.

ANEXO

Marco de presupuestación para las encuestas por conglomerados de indicadores múltiples (MICS) del Fondo de las Naciones Unidas para la Infancia (UNICEF)

Categorías de costo	Costos totales	Categorías de actividad								
		Preparación/ sensibilización	Iniciativa piloto	Diseño de la encuesta y preparación de la muestra	Capacitación	Realización de la encuesta principal	Incorporación de los datos	Procesamiento y análisis de los datos	Redacción del informe	Divulgación y ulteriores análisis
Personal										
Viáticos										
Transporte										
Bienes fungibles										
Equipos										
Otros gastos										
Costo total										

Organismos de ejecución (nombres)

Detalles suplementarios

1. Tamaño de la muestra: Número de hogares: _____ Número de conglomerados:_____
2. Duración del empadronamiento: Número de días:_____
3. Duración de la capacitación de los empadronadores: Número de días:_____
4. Número de empadronadores/Supervisores sobre el terreno: Empadronadores: _____ Supervisores: _____
5. Entrada de datos: Pulsaciones por cuestionario: Cantidad:_____
6. UNICEF: Contribución (en dólares de los EE. UU.)_____

Cálculo del costo del marco
Partidas incluidas en las categorías de costo y de actividad

Categorías de costo	Categorías de actividad
Personal (salarios) Honorarios de consultores Supervisores sobre el terreno Entrevistadores/empadronadores Conductores Traductores Guías locales Encargados de la entrada de datos Programadores informáticos Horas extraordinarias Prima de incentivo Comité de coordinación	**Preparación/sensibilización** Preparación del cuestionario Preparación de modelos de cuadros Traducción directa e inversa Comprobación previa del cuestionario Publicidad antes y después del empadronamiento
Viáticos (alojamiento y comida) Supervisores sobre el terreno Entrevistadores/empadronadores Conductores Traductores Gruías locales (subsidio para la comida) Consultores/observadores	**Encuesta piloto** Capacitación Recopilación de datos Análisis de datos Informe sobre la encuesta piloto
Transporte Alquiler de vehículos Subsidio para transporte público Combustible Costos de mantenimiento Visitas de consultores	**Diseño de la encuesta y preparación de la muestra** Planificación Preparación de la muestra
Bienes fungibles Material de escritorio (papel, lápices, bolígrafos, etcétera) Tarjetas de identificación Sobres Computación; suministros (papel, disquetes, cintas, cartuchos)	**Capacitación** Preparación de materiales de capacitación Traducción a la lengua de capacitación Actividades de capacitación
Equipo Equipo antropométrico (balanzas, metros, etcétera)	**Realización de la encuesta principal** Realización Seguimiento y supervisión Recuperación de datos
Otros gastos Impresión (cuestionario, etcétera) Fotocopias de mapas, listas y manuales de instrucción Mantenimiento del equipo Comunicaciones (teléfono, fax, correo, etcétera) Contratos (procesamiento de datos, redacción de informes)	**Incorporación de los datos** Entrada de datos Comprobación de errores
	Procesamiento y análisis de los datos Procesamiento de los datos Depuración de datos Producción de indicadores Cuadros de análisis
	Redacción del informe
	Divulgación y posteriores análisis Impresión del informe Distribución Reuniones para presentación de observaciones Ulteriores análisis

Capítulo XIII
Modelo de costos para una encuesta de ingresos y gastos

Hans Pettersson
Statistics Sweden
Estocolmo, Suecia

Bounthavy Sisouphanthong
Centro Nacional de Estadística
Vientiane, República Democrática Popular Lao

RESUMEN

En el presente capítulo se describe el establecimiento de un modelo de costos para una encuesta sobre gastos y consumo en la República Democrática Popular Lao. Se comienza con una breve descripción de los modelos de costos y los problemas existentes para estimar los componentes del modelo, y luego se describe el diseño de la Encuesta sobre gastos y consumo de la República Democrática Popular Lao de 2002. El modelo de costos, basado en las estimaciones presupuestarias para la encuesta, se utiliza para los cálculos del tamaño ideal de los conglomerados con diferentes supuestos sobre las tasas de homogeneidad de los mismos. El capítulo concluye con un análisis de la eficiencia del diseño de la muestra elegido, en comparación con la eficiencia en condiciones óptimas.

Términos clave: Diseño de la encuesta, costos de la encuesta, eficiencia, modelo de costos, tamaño óptimo de la muestra.

A. Introducción

1. El diseño de una muestra por conglomerados en varias etapas supone una serie de decisiones. Una muy importante es cómo distribuir la muestra entre las diferentes etapas de la mejor manera posible. La aglomeración de la muestra en general tiene influencias contradictorias en los costos y en las varianzas: reduce los costos y aumenta las varianzas. El diseño económico de una muestra en varias etapas presupone que un estadístico de muestreo estime y pondere esas influencias. Para esta tarea, necesita buena información sobre las varianzas atribuibles a las diferentes etapas de muestreo e información sobre los costos variables que dependen del tamaño de la muestra en cada etapa.

2. Si bien se han elaborado modelos de varianza para muchos diseños comunes en varias etapas, la formulación de modelos de costos ha recibido menos atención por los estadísticos. Actualmente, los efectos del diseño y las varianzas se compilan al menos en las estimaciones más importantes de muchas encuestas en los países en desarrollo. El uso de modelos de costos para diseñar la muestra es menos común. Parte del problema es la escasez de información detallada sobre los costos de las encuestas en muchos institutos nacionales de estadística, lo que dificulta la preparación de un presupuesto adecuado para una encuesta y la creación de un modelo de costos realista.

3. En este capítulo examinamos brevemente los modelos de costos y describimos la manera de utilizarlos junto con los modelos de la varianza para encontrar el tamaño óptimo

de la muestra dentro de las unidades primarias de muestreo (UPM) en un diseño en dos etapas. Elaboramos un modelo de costos para una encuesta sobre gastos y consumo en la República Democrática Popular Lao y utilizamos el modelo para calcular el tamaño óptimo de la muestras dentro de las UPM.

B. Modelos de costos y estimaciones de costos

1. Modelos de costos

4. Un sencillo modelo de costos para una muestra en dos etapas puede representarse como sigue:

$$C = C_0 + C_1 \cdot n + C_2 \cdot n \cdot m \quad (1)$$

donde n = número de unidades primarias de muestreo (UPM) de la muestra; m = número de unidades secundarias de muestreo (USM) (por ejemplo, hogares) en la muestra de cada UPM; C_0 = costos fijos de realizar la encuesta, independientemente de las UPM de la muestra y de las USM por UPM, incluidos los costos para la planificación de la encuesta, la elaboración del diseño, los preparativos, la gestión de la encuesta y el procesamiento y análisis de los datos y la presentación de los resultados (parte de los costos para el procesamiento de los datos dependen del tamaño de la muestra; como no se trata de costos fijos, no se tiene en cuenta aquí); C_1 = costos medios de agregar una UPM a la muestra, es decir, los costos de viaje de los entrevistadores y supervisores entre las UPM y la base o entre las UPM (gastos de combustible, salarios de los conductores) y los salarios de los entrevistadores, incluido el costo de la adquisición de mapas y otros materiales para las UPM, el de realización de la encuesta en la zona local, que se incluiría, por ejemplo, reuniones con las autoridades locales y la obtención de los pertinentes permisos, y costo de elaboración de registros y muestreo de las unidades de habitación/hogares dentro de la UPM; C_2 = costo medio de incluir un hogar adicional en la muestra; este valor acogería los gastos para localizar, contactar y entrevistar a un hogar, más los salarios y viáticos de los entrevistadores y supervisores, así como los costos de desplazamiento de los entrevistadores y supervisores dentro de las UPM.

5. Este modelo de costos es sencillo en comparación con otros más complejos que se han elaborado hasta ahora. Hansen, Hurwitz y Madow (1953) formularon un modelo que aislaba los costos de viaje entre UPM, en el cual

$$C = C_0 + C_1 \cdot n + C_2 \cdot n \cdot m + C_3 \cdot \sqrt{n} \quad (2)$$

El costo de agregar una UPM (C_1) incluye el costo de viaje de instalación (desplazamiento hasta la primera UPM visitada desde la base del entrevistador y regreso a la base desde la última UPM visitada durante el viaje de recopilación de datos) pero no el costo de viaje entre las UPM, que se incluye en el término $C_3 \cdot \sqrt{n}$. Se han propuesto también modelos para aislar los viajes entre las UPM y el de instalación (Kalsbeek, Mendoza y Budescu, 1983). Groves (1989) debate ampliamente los modelos de costos, en particular varias formas complejas, como la expresión de costos en una función no lineal, discontinua y escalonada. No obstante, la complejidad de la fórmula matemática de los modelos de costos muchas veces dificulta la búsqueda de la solución óptima. Además, la ausencia de datos precisos obstaculiza también en muchas ocasiones el uso de modelos complejos. En este capítulo se utilizará el modelo sencillo (1), y se supone que las unidades de la segunda etapa son hogares.

2. Estimaciones de costos

6. El administrador de la encuesta muchas veces tiene una idea adecuada del tiempo necesario para las operaciones de una encuesta específica tomando como base la informa-

ción de encuestas anteriores de carácter semejante. La experiencia de encuestas previas (o de encuestas piloto) puede utilizarse muchas veces para formular estimaciones razonables del tiempo por hogar necesario para su localización y la realización de las entrevistas. En esos casos podrían compilarse estimaciones razonables de C_2. Normalmente es más problemático estimar C_0, que supone la asignación de costos indirectos y de los costos del personal que trabaja en varios proyectos/actividades. Muchas veces es difícil formular estimaciones del tiempo necesario para el personal administrativo, profesional y de supervisión. En general, no hay registros de costos de buena calidad de encuestas anteriores que indiquen los costos de ese tipo de personal. Asimismo, muchas encuestas emplean apoyo técnico de donantes extranjeros. En muchos casos puede ser difícil desglosar el tiempo que dedican los consultores de asistencia técnica a una encuesta específica.

7. Muchas veces es difícil calcular una estimación razonable de C_1 porque hay que determinar el efecto del desplazamiento adicional del entrevistador cuando se agrega una UPM a la muestra. El viaje depende del tamaño de la zona incluida, el número de UPM asignadas a cada entrevistador y el sistema de viaje de los entrevistadores. En este concepto se incluyen el viaje entre UPM durante una misión de recopilación de datos y el viaje de instalación.

8. No hay una forma fácil de superar las dificultades para realizar estimaciones de costos adecuadas. La contabilidad precisa y detallada de encuestas anteriores o de una encuesta piloto es muy valiosa. Además de la experiencia anterior y de las iniciativas piloto, podrían obtenerse también datos sobre los costos instituyendo capacidades especiales de supervisión de los costos en las encuestas en curso, como se hace, por ejemplo, en la encuesta nacional de salud por entrevistas en los Estados Unidos de América (Kalsbeek, Botman y Massey, 1994).

C. Modelos de costos para un diseño muestral eficiente

9. Los modelos de costos pueden utilizarse con un doble fin:

- Para fines presupuestarios, con el fin de establecer un presupuesto de la encuesta basado en los costos unitarios del modelo de costos y el tamaño previsto de la muestra en diferentes etapas;
- Para encontrar un diseño eficiente de la muestra combinando el modelo de costos con un modelo de error muestral.

10. En este capítulo nos interesa sobre todo el uso de modelos de costos para encontrar un diseño eficiente. Suponemos un diseño en dos etapas con hogares seleccionados de las UPM en la segunda etapa. El problema puede expresarse de esta manera: dada la estructura de costos representada en el modelo, ¿cómo debería distribuirse la muestra entre las dos etapas? Normalmente se preparan modelos de costos independientes para los estratos urbanos y rurales y, en algunos casos, para otros estratos. En ese caso el problema incluye también la distribución de la muestra entre los estratos urbano y rural (y otros).

11. No tenemos que considerar los costos fijos (C_0) al tratar de establecer un diseño eficiente; la parte importante es la de los costos del trabajo sobre el terreno: $C_1 \cdot n + C_2 \cdot n \cdot m$. El costo estimado del trabajo sobre el terreno por entrevista (C_f) se obtiene dividiendo el total de los costos sobre el terreno por el número de entrevistas ($n \cdot m$), con lo que se obtiene

$$C_f = C_2 + C_1/m \tag{3}$$

La varianza correspondiente al diseño puede expresarse así:

$$Var = V \cdot (1 + \rho\,(m - 1)) \tag{4}$$

donde V es la varianza en el marco del muestreo aleatorio simple de los hogares; ρ es la tasa de homogeneidad (Kish, 1965; véase también el capítulo VI *supra*); y m es el tamaño de la muestra dentro de las UPM.

De (3) se deduce claramente que los costos del trabajo sobre el terreno por entrevista (C_f) podrían reducirse haciendo que m fuera lo mayor posible. De (4) se deduce también claramente que la varianza crece cuando aumenta m y que la varianza se reduce cuando $m = 1$. El número óptimo de hogares, m_{opt}, es el valor de m que reduzca al mínimo posible $Var \cdot C_f$ donde

$$Var \cdot C_f = V \cdot (1 + \rho\,(m - 1)) \cdot (C_2 + C_1/m) \tag{5}$$

Se ha comprobado (Kish, 1965) que el tamaño óptimo de la muestra puede encontrarse mediante

$$m_{opt} = \sqrt{\frac{C_1}{C_2} \cdot \frac{(1-\rho)}{\rho}} \tag{6}$$

12. El primer factor en la ecuación (6), C_1/C_2, es el coeficiente de costos entre los costos unitarios en la primera y la segunda fase. El costo de incluir una nueva UPM en la muestra (C_1) será siempre mayor que el de incluir un nuevo hogar en una UPM seleccionada (C_2), donde el coeficiente de costos será siempre superior a 1,0. Cuanto mayor sea el coeficiente de costo, más costoso será seleccionar una nueva UPM en comparación con la selección de más hogares en las UPM seleccionadas; por consiguiente, deberíamos seleccionar más hogares en las UPM ya seleccionadas.

13. El valor de ρ mide la homogeneidad interna de la UPM. Cuando la homogeneidad interna es alta no conviene adoptar una muestra grande de hogares en la UPM, ya que la información adicional obtenida de cada nuevo hogar de la muestra será pequeña (pues los hogares son muy semejantes). Así se refleja en el segundo factor de (6). Cuando ρ tiene un valor elevado, este factor y m_{opt} son pequeños (para un determinado coeficiente de costos).

14. Los valores de ρ muchas veces se derivan de los efectos del diseño estimados a partir de encuestas anteriores, y suelen ser pequeños —con frecuencia, inferiores a 0,01— en muchas variables demográficas. En el caso de numerosas variables socioeconómicas, los valores ρ pueden ser superiores a 0,1 y, en algunos casos, llegar hasta 0,2 ó 0,3.

15. El coeficiente de costo debe basarse también en experiencias de encuestas anteriores. Cabe señalar que no es necesario expresar el coeficiente en términos de costos. El tiempo (días de entrevistador necesarios) se utiliza con frecuencia como unidad, en vez de los costos: los resultados serán aproximadamente idénticos (quizá se pasen por alto algunos costos de viaje). El nivel del coeficiente de costos depende del diseño del trabajo sobre el terreno. En una encuesta en la que el tiempo dedicado a la entrevista sea muy breve, el coeficiente de costos puede ser 20-50. Si, por ejemplo, el tiempo necesario por UPM, independientemente de las entrevistas de los hogares, es de tres días y el entrevistador puede ocuparse de 10 hogares por día, el coeficiente de costos (calculado en forma de coeficiente de tiempo T_1/T_2) será 30 ($T_1 = 3$ días y $T_2 = 0{,}1$ días). En las encuestas con entrevistas muy largas, el coeficiente de costos puede ser inferior a 10.

16. Las operaciones matemáticas empleadas en los cálculos pueden dar la impresión de que es posible obtener una respuesta precisa y clara a la pregunta de cuántos hogares se deben elegir de cada UPM. Sin embargo, eso no ocurre casi nunca, debido a, entre otros, los siguientes factores:

- El modelo de costos es sólo una reproducción aproximada de la realidad. Se necesita cierta simplificación para que el modelo de costos sea manejable (como se examina en el apartado B).

- Las estimaciones de costos y de ρ están sujetas a incertidumbre.
- La situación óptima se aplica sólo a una variable de la encuesta de entre las muchas existentes. Si las variables importantes de la encuesta tienen diferentes niveles de ρ no habrá un único tamaño de conglomerado óptimo, sino más bien varios de ellos.

17. Los cálculos permiten hacerse una idea más bien aproximada de lo que debe ser el tamaño óptimo de la muestra para los diferentes valores de ρ. Esta información puede utilizarse para decidir el tamaño de la muestra dentro de las UPM que corresponda en forma razonable a todas las variables importantes de la encuesta. En cuanto a la decisión final, puede haber también otros factores que considerar, relacionados muchas veces con los problemas prácticos del trabajo sobre el terreno.

D. Estudio de un caso: Encuesta sobre gastos y consumo de la República Democrática Popular Lao de 2002 (LECS-3)

18. El Centro Nacional de Estadística de la República Democrática Popular Lao ha realizado dos encuestas de gasto y consumo en el último decenio. La primera (LECS-1) se llevó a cabo en 1992-1993; la segunda (LECS-2), en 1997-1998; la tercera (LECS-3), en 2002-2003. En la presente sección se describe la LECS-3.

19. Los datos de las encuestas se utilizan para diferentes fines, el más importante de los cuales sería elaborar estimaciones nacionales del consumo y producción de los hogares con destino a las cuentas nacionales. Ello supone la estimación de la producción en las actividades agrícolas y comerciales de los hogares.

1. Diseño de la muestra para la LECS-3

20. La muestra estuvo integrada por 8.100 hogares seleccionados a través de un diseño de la muestra en dos etapas. Las aldeas fueron las unidades primarias de muestreo (UPM) y se estratificaron en 18 provincias, desglosándose cada provincia en sector urbano o rural. Las aldeas rurales se estratificaron en aldeas "con acceso a carreteras" y "sin acceso a carreteras". El total de la muestra de la primera etapa consistía en 540 aldeas. La muestra se asignó a las provincias en forma proporcional a la raíz cuadrada del tamaño de la población, según el censo de población. Las UPM se seleccionaron con un procedimiento sistemático de probabilidad proporcional al tamaño (PPT) en cada provincia.

21. Los hogares de las aldeas seleccionadas se registraron antes de la encuesta. Se seleccionaron 15 hogares con muestreo sistemático en cada aldea, lo que permitió obtener una muestra de 8.100 hogares. La decisión de seleccionar 15 hogares por aldea estuvo basada fundamentalmente en consideraciones prácticas. En el apartado E comparamos la eficiencia de las muestras de 15 hogares con los tamaños óptimos de la muestra en diferentes supuestos sobre las tasas de homogeneidad.

2. Recopilación de datos en la LECS-3

22. Los datos se recopilaron por medio de un cuestionario de hogares, un cuestionario de aldeas y un formulario de recopilación de precios. Los dos últimos cuestionarios sirvieron sobre todo como instrumentos para recoger información suplementaria para la encuesta de hogares.

23. Una gran parte del cuestionario de los hogares se mantuvo sin cambios con respecto a encuestas anteriores, con excepción de algunas modificaciones en las preguntas que no habían funcionado bien en la encuesta anterior. Se recopilaron datos sobre el gasto y el

consumo durante todo un mes tomando como base el registro diario de todas las transacciones. Al final del mes se formularon a los hogares preguntas acerca de las compras de bienes duraderos durante los 12 meses anteriores. A lo largo del mes, cada miembro del hogar debería haber registrado la utilización del tiempo durante un período de 24 horas. El consumo de arroz de cada miembro del hogar se cuantificó con respecto a un "ayer" para obtener una medición más precisa de la ingestión en cada comida por cada persona.

24. El cuestionario que se administró a los jefes de las aldeas incluía aspectos como las carreteras y el transporte, el agua, la electricidad, los servicios de salud, los mercados locales, las escuelas, etcétera. El formulario de recopilación de precios fue utilizado por los entrevistadores para recoger datos sobre los precios locales de 121 productos básicos.

3. Trabajo sobre el terreno

25. La medición del consumo cotidiano mediante un diario llevado por los hogares representaba una fuerte carga tanto para ellos como para los entrevistadores. Muchos hogares, sobre todo en las zonas rurales, necesitaban con frecuencia ayuda para llevar el diario. Con el fin de garantizar una calidad aceptable de los datos se había considerado necesario mantener a los entrevistadores en la aldea durante todo un mes en vez de hacerles desplazarse repetidamente para las entrevistas y el seguimiento. Otro motivo en favor de esa decisión era el hecho de que muchas aldeas, sobre todo en las zonas montañosas, eran de difícil acceso (en algunos casos había que viajar a pie durante varios días).

26. En las encuestas anteriores el trabajo sobre el terreno se había realizado por equipos de dos entrevistadores en cada aldea. Para la LECS-3 se consideró la posibilidad de que lo llevara a cabo un solo entrevistador. No obstante, en último término, los factores relacionados con la seguridad de los entrevistadores y su bienestar aconsejaron mantener los equipos de dos entrevistadores por aldea. Los entrevistadores realizaron varias visitas a los hogares seleccionados durante el período de cuatro semanas y colaboraron también con los dirigentes locales para completar el cuestionario y actualizar los registros de la aldea. Durante ese mes, los entrevistadores recopilaron asimismo datos sobre los precios en el mercado local.

27. El personal sobre el terreno estuvo integrado por 180 entrevistadores organizados en 90 equipos, cada uno de ellos de dos miembros. Había también 36 supervisores de las oficinas provinciales de estadística y 10 supervisores centrales de la oficina principal.

E. Modelo de costos para el trabajo sobre el terreno en la Encuesta sobre gastos y consumo de la República Democrática Popular Lao

1. Estimaciones de los costos

28. La Encuesta sobre gastos y consumo de la República Democrática Popular Lao (LECS-3) fue en gran medida semejante a las dos anteriores LECS. Por ello, para estimar el costo del trabajo de campo se aprovecharon las informaciones sobre el tiempo necesario para esa labor de las dos encuestas anteriores.

29. El cuadro XIII.1 contiene estimaciones del tiempo necesario para trabajos sobre el terreno en las aldeas con ocasión de la LECS-3. Se han elaborado estimaciones independientes para las zonas urbanas y rurales.

Cuadro XIII.1
Tiempo estimado para el trabajo sobre el terreno en una aldea

	Viajes	Presentación de la encuesta, enumeración y selección de los hogares en las aldeas, recopilación de información sobre la aldea	Entrevistas de los hogares
	No. de días/aldea	No. de días/aldea	No. de días/aldea
Zonas urbanas (100 aldeas)			
Supervisores de provincia	1,5	0,5	3
Entrevistadores (equipos de dos)	3	7	47
Zonas rurales (440 aldeas)			
Supervisores de provincia	3	0,5	3
Entrevistadores (equipos de dos)	6	7	47

30. El cuadro XIII.2 contiene los costos estimados para el trabajo sobre el terreno calculado sobre la base de las estimaciones del cuadro XIII.1. Los costos incluyen los gastos de viaje (normalmente en coche o en autobús) y los viáticos correspondientes al tiempo de trabajo sobre el terreno. Todo el personal de la encuesta era personal permanente del Centro Nacional de Estadística asignado a la encuesta en el desempeño de sus funciones habituales. Por ello, en las partidas de costos no se incluían los salarios ordinarios.

Cuadro XIII.2
Costos estimados para la LECS-3 (viático en dólares de los EE. UU.)

	Gastos de viaje (viático correspondiente al tiempo de viaje y los gastos de viaje estimados)	Presentación de la encuesta, enumeración y selección de los hogares en las aldeas, recopilación de información sobre la aldea	Entrevistas de los hogares
	A	B	C
Zonas urbanas (100 aldeas)			
Supervisores de provincia	1 540	450	2 710
Entrevistadores (equipos de dos)	2 490	5 060	33 970
Zonas rurales (440 aldeas)			
Supervisores de provincia	15 850	1 990	11 950
Entrevistadores (equipos de dos)	25 560	22 260	149 460
TOTAL	**45 440**	**29 760**	**198 090**

2. Modelo de costos

31. En las columnas A y B del cuadro XIII.2 se presentan los costos relacionados con la selección y preparación de las aldeas para la encuesta. La suma de las partidas de estas columnas divididas por el número de aldeas constituye el costo medio (C_1) en dólares de los EE. UU. de la inclusión de una aldea en la encuesta:

- Zonas urbanas: $C_1 = (1.540 + 2.490 + 450 + 5.060)/100 = 95$;
- Zonas rurales: $C_1 = (15.850 + 25.560 + 1.990 + 22.260)/440 = 149$.

Por viaje se entiende siempre el desplazamiento entre aldeas; todos los gastos de viaje se incluyen, por lo tanto, en C_1.

32. En la columna C del cuadro XIII.2 se presentan los gastos de la encuesta relacionados con las entrevistas de los hogares. La principal partida es el tiempo del entrevistador. La suma de las partidas de esta columna dividida por el número de hogares constituye el costo medio (C_2) en dólares de los EE. UU. de incluir un hogar en la encuesta:

- Zonas urbanas: $C_2 = (2.710 + 33.970)/(100,15) = 24$;
- Zonas rurales: $C_2 = (11.950 + 149.460)/(440,15) = 24$.

Al insertar los valores estimados para C_1 y C_2, la función de costos es como sigue:

- Zonas urbanas: $C_{trabajo\ sobre\ el\ terreno} = 95 \cdot n + 24 \cdot n \cdot m$
- Zonas rurales: $C_{trabajo\ sobre\ el\ terreno} = 149 \cdot n + 24 \cdot n \cdot m$

33. El hecho de que los gastos de personal no incluyeran los salarios del personal permanente da lugar a una infravaloración de C_1 y C_2 y, por consiguiente, a una infravaloración de $C_{trabajo\ sobre\ el\ terreno}$. No obstante, lo más importante para la optimización del diseño es la correlación de costos C_1/C_2. Cabría esperar que el coeficiente de costos estuviera sólo ligeramente afectado por la omisión de los salarios, ya que ésta tendrá efectos más bien semejantes en C_1 y C_2.

34. El coeficiente de costos entre las muestras en la primera y la segunda etapa es $C_1/C_2 = 95/24 = 3,9$ en las zonas urbanas y $149/24 = 6,1$ en las zonas rurales. Estos coeficientes de costos son más bien bajos debido al hecho de que la encuesta necesitaba bastante tiempo para las entrevistas y seguimiento por hogar durante el mes en el que se utilizó el método de diario con ayuda del entrevistado. La LECS-3 era una encuesta anómala en ese sentido.

3. Tamaño óptimo de la muestra dentro de las aldeas

35. En las anteriores LECS los dos entrevistadores se habían ocupado de 20 hogares en cada aldea. En la LECS-3, ese número descendió a 15 hogares. La reducción del volumen de trabajo de 20 a 15 hogares se debió al hecho de que las entrevistas de hogares eran considerablemente más largas en la LECS-3 que en encuestas anteriores. Asimismo, la LECS-3 contenía un cuestionario de precios que no se había incluido en las encuestas anteriores.

36. ¿Hasta qué punto era eficiente el diseño con dos entrevistadores en la aldea para una muestra de 15 hogares? El modelo de costos, junto con un modelo de la varianza, podría utilizarse para evaluar la eficiencia relativa de las muestras de 15 hogares.

37. En el cuadro XIII.3 se presenta el valor óptimo de m para diferentes valores de ρ. La eficiencia relativa de nuestro diseño puede verse en las líneas tres y cuatro. Se calcula como coeficiente entre el mínimo de $Var \cdot C_f$ (véase (5)) y el valor efectivo de $Var \cdot C_f$ para un valor dado de ρ y un tamaño de muestra de 15. La eficiencia es razonablemente elevada para valores de ρ de hasta 0,10; la eficiencia es más bien baja y tiende a deteriorarse para valores de ρ iguales a 0,2 y superiores.

Cuadro XIII.3
Tamaño óptimo de la muestra en las aldeas (m_{opt}) y eficiencia relativa del diseño efectivo (m = 15) con diferentes valores de ρ

	$\rho = 0,01$	$\rho = 0,05$	$\rho = 0,10$	$\rho = 0,15$	$\rho = 0,2$	$\rho = 0,25$
m_{opt} Zonas urbanas	20	9	6	5	4	4
m_{opt} Zonas rurales	24	11	8	6	5	4
Eficiencia relativa (porcentaje). Zonas urbanas	99	94	82	73	66	61
Eficiencia relativa (porcentaje). Zonas rurales	96	98	89	81	75	70

38. Los cálculos de ρ en la anterior LECS revelaron que había una clara diferencia entre zonas urbanas y rurales de ρ en algunas variables importantes de la LECS. Los valores de ρ en las zonas urbanas son considerablemente más bajos que en las zonas rurales. Sería de prever que ρ estuviera en el intervalo de 0,04-0,08 en muchas estimaciones urbanas de

la LECS, en cuyo caso una muestra de ocho a nueve hogares sería la ideal. Nuestro diseño con una muestra de 15 hogares por UPM tendrá una eficiencia relativa de entre el 85% y el 95%. Los valores de ρ en las zonas rurales están en el intervalo de 0,11-0,20, en cuyo caso una muestra de cinco a siete hogares sería la solución óptima. Nuestra muestra tendrá una eficiencia relativa de entre el 75% y el 88%. Hay algo de incertidumbre, sobre todo en relación con los valores de ρ previsibles en relación con variables importantes de la LECS-3. No obstante, podemos concluir con confianza que nuestra muestra de 15 hogares está por encima del nivel óptimo.

39. ¿Qué conclusiones pueden extraerse para las futuras LECS? Como las pérdidas de eficiencia son pequeñas en las zonas urbanas, podemos decidir mantener la alternativa de 15 hogares. Desearíamos reducir la muestra por UPM en las zonas rurales. No obstante, el actual sistema de trabajo sobre el terreno, por el que los entrevistadores deben permanecer en la UPM durante todo un mes, dificulta la reducción del volumen de trabajo en forma considerable. Ello significa que los entrevistadores no estarán plenamente ocupados durante todo el mes. Es posible confiar a los entrevistadores otras tareas cuando dispongan de tiempo libre, como que durante el mes de permanencia realicen encuestas comunitarias en la zona. Habría que examinar hasta qué punto es viable esta opción.

F. Observaciones finales

40. Se ha elaborado y analizado un modelo de costos para el trabajo sobre el terreno en la LECS-3. En él se observa que el coeficiente de costos, C_1/C_2, de la encuesta era más bien bajo. La principal razón es el método del diario con ayuda del entrevistador utilizado para la LECS-3, que requería mucho tiempo y obligaba a los entrevistadores a quedarse en la aldea durante todo el mes para poder ayudar a los hogares a llevar el diario. En ese sentido, la LECS-3 fue una encuesta más bien anómala entre las encuestas de ingresos y gastos de hogares, en las que el tiempo de entrevista por hogar era normalmente más bajo.

41. Los cálculos sobre el tamaño óptimo de la muestra dentro de las UPM revelan que el actual tamaño de la muestra de 15 hogares es superior al óptimo, sobre todo en las zonas rurales. No obstante, hay consideraciones prácticas que pueden dificultar la reducción del tamaño de la muestra.

42. Conviene señalar que el modelo de costos constituye sólo una representación aproximada de la realidad, pues no es posible reflejar toda la complejidad del mundo real en un modelo sencillo. Podrían elaborarse modelos más complejos, por ejemplo, mediante fórmulas con varias funciones escalonadas. No obstante, la complejidad matemática de los modelos de costos muchas veces no permite determinar cuál es la opción ideal.

REFERENCIAS BIBLIOGRÁFICAS

Groves, R. M. (1989). *Survey Error and Survey Costs.* Nueva York: John Wiley and Sons.

Hansen, M. H., W. N. Hurwitz y W. G. Madow (1953). *Sample Survey Methods and Theory,* vol. I. Nueva York: John Wiley and Sons.

Kalsbeek, W., O. M. Mendoza y D. V. Budescu (1983). Cost models for optimum allocation in multi-stage sampling. *Survey Methodology*, vol. 9, No. 2, págs. 154-177.

Kalsbeek W. D., S. L. Botman y J. T. Massey (1994). Cost efficiency and the number of allowable call attempts in the National Health Interview Survey. *Journal of Official Statistics,* vol. 10, No. 2, págs. 133-153.

Kish, L. (1965). *Survey Sampling.* Nueva York: John Wiley and Sons.

Capítulo XIV
Elaboración de un marco para la presupuestación de las encuestas de hogares en países en desarrollo

Erica Keogh
Departamento de Estadística
Universidad de Zimbabwe
Harare, Zimbabwe

RESUMEN

En el presente capítulo se proponen algunas recomendaciones sobre la presupuestación atenta y lógica de las encuestas. Se describen las dos maneras de considerar el presupuesto —en función de las categorías contables o de las actividades de la encuesta— y se aconseja un planteamiento que señale con detalle las categorías contables dentro de cada actividad de la encuesta. El producto final es una matriz de costos, que puede utilizarse también en toda la encuesta para registrar los gastos reales. Se alienta también encarecidamente la documentación y examen de los costos reales de la encuesta con el fin de aportar material para futuras iniciativas. Se insiste en todo momento en la relación crítica entre el diseño de la encuesta por muestreo y su presupuestación.

Términos clave: Diseño, presupuesto, realización de la encuesta.

A. Introducción

1. Las encuestas requieren una gran cantidad de recursos, tanto de tiempo como de dinero; por ello es imprescindible planificar con detalle los gastos que se van a contraer, desde el comienzo de la actividad hasta su conclusión. Además, hay que planificar los imprevistos, las situaciones de emergencia y los cambios económicos inesperados y garantizar que estos acontecimientos imprevisibles se tengan en cuenta en el presupuesto. Una forma de planificar los imprevistos es introducir en el proceso de la encuesta la capacidad de ajustar el volumen de trabajo de la encuesta, en particular el tamaño de la muestra, y de esa manera disfrutar de cierta flexibilidad para responder mejor a los cambios económicos imprevistos que puedan repercutir en la ejecución de la encuesta. Un presupuesto debe considerarse como una parte dinámica del proceso de la encuesta, ya que cambia de acuerdo con las necesidades reales durante su realización. Los instrumentos para supervisar los gastos se establecerán al mismo tiempo que el presupuesto, y se actualizarán constantemente a medida que vaya evolucionando el presupuesto.

2. Como la magnitud del presupuesto y su asignación a diversos componentes dentro de la encuesta repercute directamente en la calidad de los resultados, nunca se insistirá demasiado en la importancia de una planificación y presupuestación detalladas. En el capítulo XII se han examinado en forma pormenorizada las cuestiones relacionadas con los

costos en el diseño de las encuestas de hogares. En Naciones Unidas (1984) se destaca la importancia de lograr un equilibrio entre costos y calidad: "En principio, las prioridades deberían establecerse sobre la base del análisis de costos y beneficios de varias posibilidades alternativas de utilizar los escasos recursos" (párrafo 1.5). Muchas veces, el presupuesto para la encuesta es fijo y el encargado de establecer la muestra debe elaborar un diseño, con niveles de errores aceptables, ajustado a ese presupuesto.

3. El establecimiento de un presupuesto detallado para una encuesta suele ser una labor engorrosa, ya que implica una planificación y preparación meticulosas. Además, los planificadores de la encuesta se encuentran ante un dilema en el momento de la planificación, ya que el presupuesto no puede estimarse hasta que esté en marcha el plan final de la encuesta, pero la presupuestación debe realizarse antes de la fase final de la planificación y diseño. Por eso es importante la experiencia en presupuestación y estimación de costos en encuestas anteriores. Es preciso recordar también que la asignación óptima de la muestra no puede considerarse sin tener también en cuenta los costos: por ejemplo, en el muestreo estratificado no se puede elegir entre reducir el costo para un nivel fijo de precisión o la precisión óptima para unos costos fijos (Scheaffer, Mendenhall y Ott, 1990). No obstante, los modelos de costos muchas veces no son realistas, no prevén los cambios de circunstancias que pueden producirse durante la realización de la encuesta y normalmente consideran sólo los errores de una variable. Por consiguiente, es importante mantener registros detallados de la presupuestación y los gastos que pudieran producirse, para confirmar la tendencia de los profesionales de las encuestas que insisten en la disponibilidad de la información sobre los costos y ayudar a la planificación de encuestas en el futuro.

4. Los datos sobre las encuestas son necesarios para planificar y/o decidir sobre políticas, de ahí que los resultados se necesiten lo antes posible. Muchas veces la encuesta tendrá que realizarse dentro de un marco cronológico estricto, con plazos para la terminación de las diversas etapas previamente especificadas por los organismos que la financian. De todas formas, conviene recordar que si se dispone de un pequeño margen adicional los datos pueden ser de calidad mucho mayor; por ello, los profesionales dedicados a las encuestas deberían defender esta opción en la fase de presupuestación. Por ejemplo, si, como ocurre muchas veces, el tiempo y/o el presupuesto asignado para la gestión y análisis de los datos son insuficientes, puede correr peligro la calidad de los resultados de la encuesta. Por ello, en la etapa de presupuestación hay que buscar soluciones de compromiso entre el tiempo, los costos y los errores, a fin de llegar a la solución más adecuada.

5. En las páginas de este capítulo se intenta arrojar algo de luz sobre los aspectos siguientes:

- Cómo preparar un presupuesto.
- Inconvenientes previsibles a la hora de realizar la encuesta.
- Establecimiento de instrumentos para gestionar las finanzas de la encuesta e informar al respecto.

Se hace especial referencia a las encuestas de hogares mediante entrevistas personales en países en desarrollo.

B. Consideraciones preliminares

1. Fases de una encuesta

6. Como punto de partida, antes de examinar con cierto detalle los grandes componentes del presupuesto de una encuesta de hogares conviene tener en cuenta sus principales fases, ya que es preciso planificar los costos de cada una de ellas y cumplir los planes establecidos. Las fases de la encuesta pueden resumirse como sigue:

- Diseño y preparación.
- Ejecución.
- Presentación de informes.

Los componentes de estas fases se han explicado con cierto detalle en capítulos anteriores.

2. Calendario para la encuesta

7. Un segundo elemento esencial que se debe considerar al elaborar el presupuesto es el calendario general. Normalmente, cuando se planifica una encuesta los fondos deberán asignarse teniendo en cuenta una fase de terminación y posiblemente algunos otros plazos preestablecidos. Para que la encuesta avance satisfactoriamente es imprescindible establecer un calendario realista junto con el marco presupuestario, y luego cumplirlo durante la realización de la encuesta.

Ejemplo 1

8. Supongamos que se ha recibido el encargo de realizar una encuesta en una ciudad grande para disponer de información básica sobre las empresas del sector informal, sus actividades y logros. Varios donantes están interesados en los resultados porque desean cooperar en forma de capacitación empresarial y microfinanciamiento con empresarios en ciernes. En particular, los donantes desearían tener certeza de que se abordan las cuestiones de género, y en el futuro querrían supervisar los efectos de su cooperación. Los donantes están dispuestos a asignar fondos para una encuesta de tamaño reducido en la que se entrevistará a 500 hogares/propietarios de pequeñas empresas en la ciudad. Se establecerá un período de tres meses para la recopilación de datos y otro adicional para la elaboración de un proyecto de informe básico. Debe presentarse también un presupuesto para la encuesta.

9. A continuación puede verse un primer borrador (cuadro de Gantt) del posible calendario. Cuando se considera el tiempo dedicable para tareas concretas hay que estimar el

Cuadro de Gantt aplicado a un proyecto de calendario para una encuesta sobre el sector informal

	Semana																
Tarea	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
Consultas con los donantes/publicidad	●	●				●	●						●	●	●	●	●
Diseño y comprobación del cuestionario	●	●	●														
Diseño del muestreo y selección de la muestra		●	●	●	●	●											
Diseño de la entrada de datos			●	●													
Planificación del análisis de datos			●	●	●												
Contratación del personal sobre el terreno		●	●	●													
Capacitación de empadronadores y encuesta piloto				●	●	●											
Impresión de los cuestionarios					●	●	●										
Trabajo sobre el terreno y comprobación							●	●	●	●							
Entrada y validación de datos						●		●	●	●	●	●					
Depuración y análisis de datos						●					●	●	●	●	●	●	
Producción de gráficos y cuadros														●	●	●	
Preparación del informe				●	●			●	●	●		●		●	●	●	●
Archivado				●	●											●	●

personal necesario para realizar y terminar dichas tareas en el tiempo asignado. Por ejemplo, si se han reservado cuatro semanas para 500 entrevistas, incluidas las de repetición, habrá que prever unas 24 entrevistas por día. La longitud del cuestionario, el número de entrevistas diarias y las distancias entre los informantes determinarán el personal sobre el terreno necesario.

10. En el esquema de Gantt de la página anterior puede observarse:

- Cómo se superponen las fases de la encuesta. Por ejemplo, el diseño de la entrada de datos tendrá lugar al mismo tiempo que se ultima el cuestionario, la entrada de datos comienza muy poco después de que se disponga de los primeros cuestionarios respondidos, y la depuración de datos puede comenzar incluso antes de haber introducido todos los datos.
- Cómo algunas tareas se prolongan a lo largo de toda la encuesta. Por ejemplo, la preparación del informe debería ser una tarea continua para los coordinadores de la encuesta, ya que en el informe debe registrarse cada paso del estudio.
- Cómo en algunos casos no es posible comenzar una etapa antes de terminar otra. Por ejemplo, la impresión final de los cuestionarios no puede tener lugar hasta que la prueba piloto esté terminada; luego, el margen de tiempo para la impresión es breve y coincide con las actividades principales de capacitación (se recomienda siempre comenzar el proceso de entrevista lo antes posible una vez terminada la capacitación).

3. Tipo de encuesta

11. El proceso presupuestario puede depender del tipo de encuesta que se va a realizar. Con respecto a la presupuestación, hay dos tipos principales de encuestas que deben considerarse aquí: las encuestas incluidas en el presupuesto sobre países concretos y las encuestas pagadas por los usuarios.

Encuestas incluidas en el presupuesto de los países

12. Cada país cuenta con departamentos específicos (gubernamentales) que tienen la responsabilidad de realizar encuestas periódicas sobre salud y nutrición, sobre aspectos demográficos de los hogares, sobre el ingreso, el consumo y el gasto y sobre la agricultura y la ganadería, etcétera. La mayor parte de esos estudios y encuestas tendrán probablemente las siguientes características:

- Cuentan con una infraestructura común que está ya establecida y se utiliza una y otra vez en las iniciativas de este tipo, es decir, que forma parte de un programa "integrado";
- Están incluidos en el presupuesto del gobierno central, aunque se puede pedir a los donantes que aporten financiamiento adicional;
- Tienen personal permanente que participa en las encuestas;
- Disponen de equipo de tecnología de la información y servicios de transporte.

En otras palabras, esas encuestas forman parte del quehacer cotidiano que se desarrolla en algunas secciones del sector público. Precisamente por tratarse de una actividad cotidiana es por lo que en buena parte suelen basarse en estudios anteriores, que deberán considerarse al presupuestar el estudio en curso. Estas encuestas se realizan normalmente utilizando una muestra representativa nacional y suelen tener un calendario algo flexible, con plazos expresados no en días, sino en meses. Algunas de las partidas presupuestarias presentadas en el resto de este capítulo quizá no puedan aplicarse a estas encuestas.

Encuestas pagadas por los usuarios

13. Las encuestas pagadas por los usuarios no están vinculadas con un programa gubernamental central y son realizadas por una organización privada financiada por varias organizaciones no gubernamentales y por donantes, tanto nacionales como internacionales. Estas encuestas pueden ser iniciativas excepcionales, que deben conseguir resultados de calidad y a corto plazo. Por otro lado, pueden utilizarse para el seguimiento de los programas, y en algunos casos quizá se necesite un prolongado análisis de datos para la elaboración de modelos con el fin de planificar actividades futuras. En ciertas ocasiones los organismos que realizan estas encuestas pueden tener:

- Una infraestructura limitada como base de todo el proceso;
- Poco personal a quien recurrir para dichos estudios;
- Insuficiencias en el equipo de tecnología de la información y en los servicios de transporte;
- Recursos fijos limitados.

En otras ocasiones los organismos están bien equipados por haber realizado previamente varios de estos estudios en el pasado reciente. Los recursos fijos y los gastos generales deben excluirse del presupuesto, y si la organización es privada deberán tenerse en cuenta las posibilidades de obtener beneficios. El tamaño de la muestra de estas encuestas no suele ser demasiado grande. En muchos casos la encuesta se concentra únicamente en algunas zonas geográficas del país. Requerir rigurosos calendarios y plazos suele ser uno de los rasgos distintivos de estas encuestas, lo que hace que, por desgracia, la calidad de los datos muchas veces sufra los efectos de una planificación no demasiado realista.

4. Presupuestos frente a gastos

14. La presupuestación de una encuesta se realiza mucho antes de que comience su realización. El marco presupuestario debe formularse y presentarse al organismo de financiación antes de que comience la planificación real. En consecuencia, deben adoptarse algunos supuestos básicos sobre el diseño de la encuesta antes de que comience la preparación del presupuesto. Por otro lado, los gastos efectivos de la encuesta reflejan lo que ocurre realmente a lo largo del estudio. Los encargados de realizar la encuesta deben tener conciencia de esta distinción y comprender que el presupuesto debe tener en cuenta los costos eventuales. El gasto depende mucho del tiempo en relación con aspectos como la inflación o el tipo de cambio, y naturalmente diferirá de un país a otro, algunas veces en forma sustancial. Se recomienda que la presupuestación se haga en forma de días-hombre, distancias recorridas, etcétera, así como del costo previsto (utilizando una moneda internacional), con el fin de tener en cuenta los efectos de la inflación y otros cambios imprevistos en la situación macroeconómica de un país. Como se ha mencionado antes, el presupuesto de la encuesta es una entidad dinámica, que se presta a actualización constante una vez que dan comienzo la fase de realización y con ella el gasto real.

5. Estudios anteriores

15. Dice un conocido adagio que "la experiencia es madre de la ciencia". No obstante, en el caso de los presupuestos para una encuesta la tarea es mucho más difícil de lo que cabría prever. Parece que en todo el mundo existe cierta tendencia a presentar informes incompletos y de poca calidad sobre los costos de las encuestas, lo que significa que es más bien difícil extraer de ellos información para planificar la encuesta siguiente. Al solicitar información sobre los costos a organizaciones que han realizado recientemente encuestas, el autor descubrió que muchas veces sólo disponían de los presupuestos originales, aunque

se reconocía que las asignaciones reales habían sido diferentes de las que figuraban en el presupuesto, debido a una serie de factores ajenos, como la inflación. Al parecer, los costos efectivos no se registraron en ningún lugar. Y lo que es curioso es que se aceptaba que ésta fuera una práctica normal siempre que los gastos no fueran superiores al presupuesto. Otro problema al examinar los presupuestos de encuestas pasadas es la falta de información sobre los costos ocultos, como el uso gratuito de vehículos o el salario del director. El hecho de que estos costos muchas veces se consideren como costos generales y no se incluyan en el presupuesto de la encuesta casi seguro llevará a engaño a los investigadores en el futuro.

16. Es de esperar que la lectura de este capítulo aliente a los encargados de realizar las encuestas a registrar los costos cotidianos y a documentarlos plenamente para que los investigadores del futuro puedan aprender de la experiencia pasada. Una documentación completa que permita conocer el costo por entrevista es sumamente valiosa para quienes deseen presupuestar iniciativas semejantes en el futuro. El costo por entrevista recoge todos los aspectos del costo efectivo de la encuesta, con inclusión del diseño, el trabajo sobre el terreno, el procesamiento de los datos y la presentación de informes, y constituye un útil resumen general de los costos reales.

C. Principales categorías contables dentro del marco presupuestario

17. Hay dos maneras de considerar el presupuesto y, en definitiva, el gasto de la encuesta: en función de las actividades o en función de los procedimientos contables habituales. Se recomienda que al preparar el marco presupuestario se consideren por separado las categorías contables dentro de cada actividad de la encuesta. Luego es posible resumir las categorías contables en general, utilizando la información de cada actividad, y agruparlas para su presentación a los organismos de financiamiento. Al mismo tiempo, sería útil presentar al organismo de financiación la presupuestación detallada de cada una de las actividades de la encuesta, con el fin de destacar las necesidades concretas para cada actividad. En el cuadro XIV.1 se presenta un ejemplo de este planteamiento. En él se utiliza una matriz para ilustrar la necesidad de presupuestación teniendo en cuenta los dos puntos de vista.

18. Si se compara el cuadro XIV.1 con el calendario presentado en el esquema de Gantt, se observa que en ambos casos se trata de las mismas actividades. Aunque por razones de espacio se han realizado aquí algunas agrupaciones, habría que detallar en cada casilla del cuadro XIV.1 cuál es el aumento exacto de los costos.

19. Más adelante trataremos de determinar las categorías contables de interés para la presupuestación de la encuesta. En el apartado D nos ocuparemos del presupuesto de las actividades de la encuesta; en el apartado E presentaremos una sinopsis. Las categorías mencionadas a continuación no son exhaustivas y quizá sea necesario revisar categorías adicionales específicas de la encuesta.

Cuadro XIV.1
Matriz de categorías contables frente a actividades de la encuesta

	Consultas	Diseño	Muestreo	Trabajo sobre el terreno	Procesamiento de los datos	Presentación de informes	Total
Personal							
Transporte							
Equipo							
Bienes fungibles							
Otros							
Total							

1. Personal

20. Los sueldos y salarios de todo el personal deben calcularse cuidadosamente e incorporarse al presupuesto. Entre los costos adicionales que deberán considerarse aquí figuran los que se producen cuando la encuesta dura un largo período de tiempo: la subida de la inflación, por ejemplo, podría obligar a aumentar los salarios. Hay que planificar también los problemas de salud y la movilidad del personal.

21. Los salarios que se paguen al personal deben estar en consonancia con las condiciones locales, pero sin olvidar que el personal de la encuesta tiene largas jornadas de trabajo, que a veces se prolongan incluso por la noche y durante los fines de semana, y que muchas veces se trata de trabajo por contrata. Quizá haya que prever prestaciones suplementarias e incluirlas en el proceso presupuestario. No hay que olvidar que los trabajadores que consideran que no se les paga lo suficiente suelen cometer imprecisiones, con lo que aumentarían los errores no muestrales. Según cuál sea la duración de la encuesta, el sistema de pagos podría ser diario, semanal o mensual. Es fundamental disponer desde el comienzo mismo de los fondos necesarios para el pago oportuno y completo de los sueldos y salarios. Deberán preverse subsidios para los empadronadores y jefes de equipo que trabajen fuera de su ciudad. Algunos encargados de la realización de la encuesta tienden a pagar al personal "por entrevista terminada", pero esta práctica puede dar lugar a distorsiones, por lo que no es muy recomendable.

22. Es preciso tener en cuenta todas las categorías de personal, desde la más baja a la más alta, con inclusión de quienes quizá trabajen sólo a tiempo parcial en el proyecto. El calendario de la encuesta servirá de guía para evaluar el tiempo trabajado por cada posible miembro del personal.

23. Un cuadro de carga de personal permite elaborar la sección de salarios y sueldos del presupuesto. En este caso se utiliza también el concepto de matriz para presentar un panorama del posible uso del tiempo por parte de cada miembro del equipo de la encuesta. Puede verse un ejemplo en el cuadro XIV.2. Como antes se indica, durante el proceso de planificación se requerirá información más detallada dentro de cada casilla.

Cuadro XIV.2
Matriz del tiempo de personal previsto (días) con respecto a las actividades de la encuesta

	Número de miembros del personal	Número de días-hombre en cada actividad						
		Consultas	Diseño	Muestreo	Trabajo sobre el terreno	Procesamiento de los datos	Presentación de informes	Total de días
Director								
Supervisor								
Jefe de equipo								
Empadronador								
Auxiliar de proceso de datos								
Analista								
Personal de secretaría								
Conductores								
Otros								
Total de días								

24. Los trabajadores sobre el terreno deberían recibir un viático diario para el pago de sus comidas, bebidas y otras necesidades básicas mientras realizan su trabajo. La magnitud de este viático debería estar en consonancia con las circunstancias locales, pero ser quizá algo mayor de lo habitual, para prevenir situaciones en que la alimentación escasee y haya que garantizar que el personal dispone de fondos para casos de emergencia como ése.

25. Los gastos de alojamiento del personal que trabaje fuera de casa deben incluirse en el presupuesto y pagarse de forma oportuna. En muchos casos, los propios trabajadores prefieren buscar alojamiento por ellos mismos mientras se desplazan de una zona a otra, pero hay circunstancias en que será más conveniente establecer un mecanismo centralizado.

2. Transporte

26. Los costos de transporte pueden estimarse bastante bien si se conoce la ubicación de los informantes, es decir, una vez que se haya establecido el diseño muestral básico. Según las circunstancias, podría aconsejarse a los empadronadores que se ocuparan de su propio transporte, registrando los costos para su reembolso en el futuro, o se pueden ofrecer medios de transporte a cada uno de los trabajadores sobre el terreno. Esta última opción es más recomendable, ya que entonces el equipo funcionará como tal y el jefe podrá supervisar mucho más fácilmente el calendario de entrevistas. Entre los costos adicionales no fácilmente previsibles se incluyen la subida de los precios del combustible, imprevistos atmosféricos que impidan la circulación por algunas carreteras, etcétera, circunstancias que deberían incluirse en los costos imprevistos.

27. En el presupuesto deberían preverse asimismo los costos de transporte para las reuniones habituales de los jefes de equipo con los directores de la encuesta; también en este caso con el fin de mantener métodos coherentes de recopilación de datos.

28. Quizá sea necesario comprar o alquilar vehículos/motos/bicicletas para el trabajo sobre el terreno. Su inclusión en el presupuesto puede resultar difícil en un contexto de inflación reciente.

3. Equipo

29. Normalmente pueden ofrecerse estimaciones válidas de los gastos probables en equipo mucho antes de que comience la encuesta. Los problemas que pueden presentarse con estos aspectos del presupuesto normalmente se centran en la subida de precios y en la disponibilidad de los productos necesarios. Si es probable que ocurra algo semejante, se aconseja adquirir los productos con bastante antelación y en cantidad suficiente para toda la encuesta. En este contexto debería considerarse el equipo de tecnología de la información y de comunicaciones, fotocopia e impresión.

4. Bienes fungibles

30. Entre los artículos que deben considerarse en esta parte del presupuesto figuran el material de oficina, los programas informáticos y elementos para el trabajo sobre el terreno como bolsas, mapas, documentos de identificación y portapapeles y otro material de oficina posible. Los productos fungibles para impresión y reproducción de documentos constituirán una parte importante de esta sección del presupuesto, ya que es imprescindible tener acceso durante las 24 horas del día a los servicios de reproducción mientras dure la encuesta.

5. Otros costos

31. Habrá que pensar siempre en dedicar un módico apartado a gastos de publicidad e información durante la encuesta. El alcance de estas actividades dependerá por completo del carácter y del tamaño de la encuesta, y puede distribuirse en varios momentos de la

misma. Como ejemplo de dichas actividades cabe citar las reuniones o talleres con las partes interesadas, incluidos los líderes comunitarios y usuarios finales; el contacto previo con los informantes; los anuncios, etcétera. La publicidad debería ser una actividad constante mientras dure la encuesta, ya que la información se hace llegar a las partes interesadas como preparación para la divulgación final de los resultados.

32. Durante algunas fases de la encuesta el personal empleado es muy numeroso. Es fundamental contar con espacio suficiente para reuniones prolongadas (por ejemplo, durante la capacitación) y para almacenamiento de cuestionarios y con locales donde los administrativos y supervisores de la entrada de datos puedan trabajar en condiciones confortables. Algunas veces habrá que alquilar emplazamientos alternativos más próximos a donde se realizan las actividades sobre el terreno, mientras que en otras ocasiones se tendrá acceso a ese tipo de instalaciones.

33. Los costos de capacitación pueden aumentar en forma alarmante si no se realiza la preparación adecuada. Entre esos costos se incluyen los gastos de alojamiento asociados con las actividades de capacitación y los de transporte para las entrevistas, más los viáticos para todos los afectados. Todos estos costos deben tenerse en cuenta por adelantado.

34. Es fácil olvidar algunas de las comunicaciones que serán necesarias para realizar una encuesta. Entre ellas se incluyen el uso de teléfonos, el correo electrónico, el fax y el correo ordinario. Muchas veces es difícil presupuestar esas partidas, ya que nunca se sabe cuál será la cantidad necesaria. En general, se llega a un monto global, muchas veces como porcentaje del total, con lo que se espera sufragar los gastos reales. La comunicación constante con los equipos sobre el terreno es fundamental para poder responder con rapidez en caso de acontecimientos imprevistos y para mantener métodos coherentes de recopilación de datos. En los países donde la red de teléfonos celulares/móviles es fiable, estos instrumentos constituyen un medio sumamente útil de comunicación inmediata.

35. Los costos "ocultos" se refieren a las partidas/servicios de infraestructura ya existentes, como las computadoras o el espacio de oficina. Otras facetas ocultas del presupuesto que quizá no sean demasiado obvias son los costos de operación del personal empleado para realizar tareas en más de un proyecto y los correspondientes al transporte y a los productos fungibles que se utilizarán en varios proyectos diferentes, cada uno de ellos con su propio presupuesto. Normalmente, conviene tratar de estimar el tiempo/cantidad que se gastará/utilizará realmente en la actividad que se está planificando, aunque en ocasiones es posible estimar a grandes rasgos estos costos generales adicionales en forma de porcentaje del total. Es importante que todos estos costos ocultos se identifiquen y contabilicen para que al planificar encuestas futuras se cuente con ello y se puedan planificar en consecuencia, aun cuando es posible que la situación haya cambiado entre tanto.

6. Ejemplos de presupuestación por categorías contables

36. Como ya se ha dicho, es sumamente difícil conseguir información sobre los costos efectivos de una encuesta. El ejemplo siguiente, recogido de Ajayi (2002), y hace referencia a cálculos de costos de varios países africanos en relación con las encuestas sobre las metas para el final del decenio realizadas en la fase precedente a la petición de las Naciones Unidas de formular indicadores sobre la salud y bienestar materno-infantil.

Ejemplo 2

37. Se contó con información de 12 países acerca de los costos de la encuesta de acuerdo con las categorías contables. En el cuadro XIV. 3 pueden verse ejemplos de las categorías utilizadas en las encuestas incluidas en el presupuesto de los países, y se indica la proporción del presupuesto total asignado a cada una de ellas.

Cuadro XIV.3
Costos desglosados por categorías contables, en proporción del presupuesto total. Encuesta sobre las metas para el final del decenio (1999-2000), algunos países africanos (*porcentaje*)

País	Personal [a]	Transporte	Equipo	Bienes fungibles	Otros	Tamaño de la muestra
Angola	62,7	22,2	9,6	1,3	4,2	6 000
Botswana	79,2	0 [b]	10,1	3,5	7,2	7 000
Eritrea	64,0	0 [b]	28,0	4,8	3,2	4 000
Kenya	62,3	22,8	3,3	4,7	6,9	7 000
Lesotho	75,1	5,2	5,8	2,3	11,6	7 500
Madagascar	31,2	6,5	33,3	12,8	16,1	6 500
Malawi	32,0	17,3	23,9	21,6	5,2	6 000
Somalia	43,8	17,7	5,0	1,0	32,5	2 200
Sudáfrica	69,3	24,0	1,5	3,7	1,5	30 000
Swazilandia	29,8	4,3	1,9	1,0	63,0	4 500
República Unida de Tanzanía	77,9	12,8	1,6	1,2	6,5	3 000
Zambia	81,8	5,2	2,0	5,6	5,4	8 000
Total	62,9	14,9	7,4	6,3	8,5	7 054

Fuente: Ajayi (2002).
[a] Incluidos los viáticos.
[b] Imposibilidad de obtener esta información por separado.

38. Del cuadro XIV.3 se desprende que hay una notable divergencia en el presupuesto por categorías contables usado en encuestas semejantes en otros países. Sería previsible un aumento del tamaño de la muestra acompañado de un incremento del presupuesto para costos de personal, pero en la realidad no es así. Por ejemplo, si se compara Sudáfrica con la República Unida de Tanzanía. No obstante, es posible que la mayor parte de las encuestas dediquen hasta dos tercios del presupuesto para costos de personal y viáticos durante el trabajo sobre el terreno. La siguiente partida más costosa sería el transporte, cuyo monto variará dependiendo de la zona incluida pero que absorberá entre el 15% y el 20% del presupuesto. El financiamiento de estas encuestas corrió por cuenta del Fondo de las Naciones Unidas para la Infancia (UNICEF) y del gobierno en cuestión; la proporción sufragada por el UNICEF presentaba considerables diferencias de un país a otro.

Ejemplo 3

39. Este ejemplo hace referencia a la presupuestación de una encuesta de hogares realizada en 1999 en el contexto de los estudios del Proyecto de Evaluación del impacto de los servicios a microempresas (Assessing the Impact of Microenterprise Services, AIMS) (Barnes y Keogh, 1999; Barnes, 2001) en donde se investigan las operaciones de microfinanciamiento en Zimbabwe y, por lo tanto, se hace referencia a una encuesta pagada por los usuarios [financiada por Management Systems International (MSI) a través de la Agencia de los Estados Unidos para el Desarrollo Internacional (USAID)].

40. En el cuadro XIV.4 se ve que gran parte del presupuesto de la encuesta (el 75%) se asignó a gastos de personal y viáticos. Ello se debió en parte al tipo de encuesta, que era una actividad de seguimiento del estudio de referencia de 1997 y que requería la ubicación y/o identificación de los mismos informantes, para lo que se requiere mucho tiempo.

Cuadro XIV.4
Proporción del presupuesto asignado a categorías contables: Evaluación del impacto de los servicios a microempresas (AIMS), Zimbabwe (1999) *(porcentaje)*

Personal	Transporte	Bienes fungibles	Otros	Tamaño de la muestra*
75	8	912	5	691

* El tamaño final de la muestra fue de 599, por la imposibilidad de localizar a 92 de los 691 informantes, por varias causas.

D. Actividades fundamentales de la encuesta en el marco del presupuesto

41. Una vez que se toma conciencia de todos los aspectos de la encuesta que deberán presupuestarse, se puede proceder a definir y establecer las categorías contables que se van a utilizar. A continuación se consideran las fases de la encuesta y se prepara un presupuesto completo, utilizando las categorías contables definidas, especificando por separado cada una de las etapas. Ello permitirá crear un marco presupuestario basado en el concepto de matriz que se esboza en el apartado C.

42. Para facilitar la documentación de los costos en el futuro, los costos reales se hacen patentes a medida que se avanza fase por fase a lo largo de la encuesta; de igual modo, la elaboración del presupuesto fase por fase facilitará mucho las comparaciones y favorecerá que salten a primer plano aspectos y diferencias notables entre presupuesto y costos.

43. Además, este planteamiento nos ayudará a mantenernos conscientes de los estrechos vínculos que existen entre la calidad de los datos, el calendario de la encuesta y el presupuesto.

1. Presupuestación para la preparación de la encuesta

44. En esta fase se decide y se presupuestan todos los preparativos necesarios para poner en marcha la encuesta. Deben considerarse todas las categorías contables en forma sucesiva y estimar exactamente qué se necesitará para cada una de ellas. Quizá sea aconsejable formular de entrada algunos pedidos de productos fungibles, material de oficina, equipo, uso de vehículos, etcétera, si se está trabajando en un contexto de alta inflación. La contratación y selección de personal y la publicidad serán también actividades importantes, lo mismo que la preparación y conclusión del diseño de la muestra y los cuestionarios, con sus correspondientes manuales, además de los preparativos iniciales para la entrada y gestión de los datos.

45. Una parte considerable del proceso de diseño de la encuesta es la preparación del marco muestral. El tipo de encuesta impondrá la naturaleza del marco, pero algunas veces se necesitará bastante tiempo o numerosos viajes, o ambas cosas, para actualizar un marco existente o para generar uno nuevo. Ello puede suponer entre otras cosas la necesidad de decidir sobre las enumeraciones, sean de hogares, aldeas o de alguna unidad de nivel superior. Estas enumeraciones requieren asignaciones presupuestarias independientes.

46. Otras actividades que pueden exigir bastante tiempo son la preparación de los cuestionarios y de los manuales para la capacitación y el trabajo sobre el terreno.

2. Presupuestación de la realización de la encuesta

47. Es probable que la realización de la encuesta sea el aspecto más costoso de todo el proceso, lo que significa que la presupuestación dentro de cada una de las categorías contables, con todas las hipótesis posibles, reviste la máxima importancia. El tiempo y el presupuesto asignado a la impresión final de los cuestionarios debe analizarse atentamente y planificarse con suficiente antelación y recurriendo a fuentes fiables. Es importante recordar que en el momento en que comienza el trabajo sobre el terreno las actividades de la oficina central deberían orientarse mayormente a la entrada de datos.

48. Como ya se ha dicho, el afán de mantenerse dentro de los límites del presupuesto no debería mover a que haya recortes en el tiempo dedicado al trabajo sobre el terreno, pues si así fuere se podría poner en peligro la calidad de los datos, al aumentar los errores no muestrales.

3. Presupuestación del procesamiento de datos de la encuesta

49. Los presupuestos correspondientes a la entrada, validación, depuración y análisis de datos deberían planificarse teniendo en cuenta todos los posibles escenarios, para evitar que esas actividades se hagan en forma precipitada, lo que daría lugar a informes deficientes e incompletos. Gran parte del trabajo de impresión se realizará durante esta etapa, y si se escatima en material de oficina se pondrá en peligro la calidad global de los resultados. Se necesitarán también servicios adecuados de tecnología de la información, con inclusión de copias de seguridad (CD, discos).

4. Presupuestación de los informes sobre la encuesta

50. Cuando se ha terminado el trabajo sobre el terreno y ha comenzado la entrada de datos habrá que ocuparse de la siguiente fase del presupuesto: la relativa a los informes y a la terminación de la encuesta. También en este caso es importante que la encuesta haya estado bien diseñada, ya que esta parte determina hasta qué punto se necesita un análisis de datos y cuál va a ser el nivel de información. Se recomienda encarecidamente la documentación continuada a lo largo de toda la encuesta, ya que un diario de actividades, con detalles sobre decisiones, problemas y costos, podrá utilizarse con provecho en las secciones descriptivas del informe. Deberán considerarse con atención las categorías contables y asignar cantidades suficientes a cada una de ellas en esta fase final de la encuesta.

5. Ejemplos de presupuestación de actividades de la encuesta

51. A continuación se presenta la información contenida en los ejemplos 2 y 3 del apartado C.6 pero desde una perspectiva sobre la actividad de la encuesta.

Ejemplo 4

52. Volviendo al ejemplo 2 (encuestas sobre las metas de final del decenio), se dispone de información para calcular el costo de actividades concretas de la encuesta en 10 países. El cuadro XIV.5 es un resumen de ello.

53. En todos los países, con excepción de Swazilandia, se observa la gran proporción de costos que deben asignarse a la realización de la encuesta. Parece razonable estimar que entre el 70% y el 90% del presupuesto se destinará a esta fase de la encuesta. Dado que (como

Cuadro XIV.5
Costos de las actividades de la encuesta en proporción del presupuesto total. Encuesta sobre las metas de final del decenio (1999-2000), algunos países africanos (*porcentaje*)

País	Preparación	Realización[a]	Procesamiento de datos[b]	Presentación de informes[c]	Tamaño de la muestra
Angola	0 [d]	83,6	6,1	10,3	6 000
Botswana	10,4	59,1	21,7	8,8	7 000
Kenya	0 [d]	93,9	2,6	3,5	7 000
Lesotho	0 [d]	73,2	18,6	8,8	7 500
Madagascar	0,3	78,6	3,0	18,1	6 500
Malawi	5,0	62,7	16,4	15,9	6 000
Sudáfrica	1,3	93,1	2,9	2,7	30 000
Swazilandia	63,0	23,4	7,5	6,1	4 500
República Unida de Tanzanía	22,7	72,4	3,6	1,3	3 000
Zambia	0,4	92,0	6,4	1,2	8 000
Total	7,0	81,0	6,0	6,0	7 054

Fuente: Ajayi (2002).
[a] Se incluye: capacitación, diseño, comprobación piloto y recopilación de datos.
[b] Se incluye: entrada, gestión y análisis de datos.
[c] Se incluye: preparación y divulgación del informe.
[d] Imposibilidad de obtener esta información por separado.

se ha observado en el cuadro XIV.3) Malawi presentaba un alto nivel de costos correspondientes al equipo, ello podría explicar la mayor proporción destinada a los costos de procesamiento de los datos y presentación de informes que se observa en el cuadro XIV.5. Lo que no se ha encontrado ha sido una explicación a las proporciones relativamente altas asignadas por Botswana y Lesotho a los costos de procesamiento de los datos. En este caso se pidió a los países que presentaran una "matriz" de los costos en la que se indicaran las categorías contables dentro de las actividades de la encuesta. Por desgracia sólo la República Unida de Tanzanía y Eritrea presentaron este resumen.

Ejemplo 5

54. Volviendo al ejemplo 3, en el cuadro XIV.6 puede verse la información sobre los costos por actividad para la encuesta AIMS de Zimbabwe, de 1999.

55. La proporción bastante elevada de costos de realización de la encuesta en el presupuesto total, como ilustra este ejemplo de encuesta pagada por el usuario, se debe probablemente al hecho de que la muestra para esta encuesta AIMS estaba integrada por 691 informantes que habían participado ya en la encuesta anterior (1997), lo que supuso costos bastante elevados para su localización (22% del presupuesto total).

Cuadro XIV.6
Costos de las actividades de la encuesta en proporción del presupuesto total: AIMS, Zimbabwe (1999) (*porcentaje*)

Preparación	Realización [a]	Procesamiento de datos [b]	Presentación de informes [c]	Tamaño de la muestra
4	85	8	3	599

[a] Se incluye: localización de los informantes, diseño, capacitación, comprobación piloto y recopilación de datos.
[b] Se incluye: entrada, gestión y depuración de datos.
[c] Se hace referencia únicamente a la presentación de informes localizados hasta la producción de conjuntos de datos depurados; el análisis detallado de los datos y el informe final se realizaron en el marco de contratos independientes.

E. Sinopsis

56. Una vez que se han preparado los costos dentro de cada categoría contable para cada tipo de actividad de la encuesta, puede prepararse una matriz de categorías contables por actividad para facilitar la consideración final del presupuesto. Esta matriz ayuda a los encargados de la planificación a considerar todo el conjunto en un plano global, con lo que se corregirían las contradicciones y superposiciones y se haría hincapié en los grandes costos previstos; asimismo, y ayudaría a los organismos de financiamiento a comparar los costos entre varias encuestas, lo que cooperaría a evaluar mejor la validez del presupuesto propuesto.

57. Como se ha mencionado antes en el ejemplo 4, sólo 2 de los 21 países que participaron en las encuestas sobre las metas de final del decenio elaboraron la matriz de costos solicitada en categorías contables por actividades de la encuesta. Ésa es la razón por la cual para este ejemplo no podemos compilar una matriz de categorías contables por actividades de la encuesta.

58. En cambio, sí se cuenta con la información obtenida para la encuesta AIMS. El cuadro XIV.7 (véase la página siguiente) contiene la clasificación cruzada de los cuadros XIV.4 y XIV.6.

59. Una matriz como la presentada en el cuadro XIV.7, en la que se observaren claramente las necesidades presupuestarias para una encuesta, alentaría a los organismos de financiamiento a considerar favorablemente cualquier solicitud. Además, si se dispone de estos detalles es posible ajustar fácilmente el presupuesto para atender las necesidades imprevistas cuando suba la inflación. Finalmente, el registro continuado de los gastos que debe llevarse a cabo durante todo el proceso de la encuesta se adapta fácilmente a una matriz semejante de los costos efectivos. Obviamente, se necesitará también una matriz como la aquí presentada pero que además contenga las cifras reales en dólares.

Cuadro XIV.7
Costos de las categorías contables por actividad de la encuesta en cuanto proporción prevista del presupuesto. AIMS, Zimbabwe (1999) (*porcentaje*)

	Preparación	Realización	Datos	Presentación de informes	Total
Personal	3,0	65	5	2,0	75
Transporte	0,0	8	0	0,0	8
Bienes fungibles	0,9	9	2	0,1	12
Otros	0,1	3	1	0.9	5
Total	4,0	85	8	3,0	100

60. El resumen final que un organismo de financiamiento desearía ver cuando se le presente un presupuesto es una estimación del costo por hogar o por otra unidad de muestreo. Una vez más, esta cifra puede servir como línea de delimitación para una consideración realista del presupuesto por comparación con iniciativas semejantes.

61. Esta matriz se presta fácilmente a cambios dinámicos durante la realización de la encuesta, ya que presenta un panorama global, que nos permite ver cómo es posible reducir los gastos en una esfera al tiempo que se aumentan en otra más necesitada. De esta manera pueden registrarse los cambios en el diseño de la encuesta, el financiamiento recibido y la experiencia real de la ejecución. Cuando la encuesta AIMS (1999) se puso realmente en práctica hubo que introducir cambios en el presupuesto formulado, sobre todo en la esfera de los costos personales, debido a una inflación imprevista y galopante. Los encargados de la realización de la encuesta pudieron también transferir fondos de los bienes fungibles, del transporte y de otras categorías incluidas en las actividades sobre el terreno de la encuesta (realización), para absorber los costos adicionales en concepto de personal que fueran necesarios. En el cuadro XIV.8 puede verse la matriz de gasto real efectivo de esta encuesta.

Cuadro XIV.8
Costos de las categorías contables por actividad de la encuesta en cuanto proporción realizada del presupuesto. AIMS, Zimbabwe (1999) (*porcentaje*)

	Preparación	Realización	Datos	Presentación de informes	Total
Personal	3,3	69,3	5,6	2,5	80,7
Transporte	0,0	6,6	0,0	0,0	6,6
Bienes fungibles	0,6	7,1	2,1	0,1	9,9
Otros	0,1	2,5	0,0	0,2	2,8
Total	4,0	85,5	7,7	2,8	100

F. Posibles limitaciones y dificultades presupuestarias

62. Por muy atentamente que se planifique la encuesta, la realidad sobre el terreno nunca corresponde a las expectativas. Es importante ser consciente de ello desde el comienzo, ya que de esa manera se puede incluir en la aplicación final del presupuesto lo que se conoce como costos imprevistos. Esta categoría normalmente se evalúa como porcentaje del costo total, de acuerdo con las recomendaciones ya adelantadas. Normalmente es aceptable entre un 5% y un 10% del costo total.

63. Aparte de contar y dedicar cierta cantidad a gastos imprevistos, cuando se planifica una encuesta hay que tener muy en cuenta las condiciones del país, sobre todo si su situación política y/o económica no es estable. Los organismos de financiamiento deben

tomar conciencia de estas posibilidades en el momento de presentación del presupuesto y no olvidarlas durante el tiempo que dure la realización de la encuesta. Ésa es la única manera de estar alertados y poder actuar con rapidez ante los acontecimientos que provocan desajustes en el presupuesto. Estos acontecimientos pueden producirse por causas humanas o por razones ambientales; por otro lado, en la discusión constante con quienes aportan los fondos y/o encargan la encuesta deben incluirse también cuestiones como la política local, la economía, la meteorología, los movimientos migratorios, etcétera.

64. Por ejemplo, en la encuesta AIMS de Zimbabwe de 1999 la inflación había aumentado de forma constante durante algunos meses. Los coordinadores de la encuesta consideraron que habían tenido ese factor suficientemente en cuenta al elaborar el presupuesto. No obstante, cuando el trabajo sobre el terreno estaba a punto de comenzar, las autoridades congelaron el tipo de cambio del dólar de los Estados Unidos en un nivel muy bajo y poco realista que no estaba en sintonía con el aumento constante de la tasa de inflación. En ese contexto, los costos previstos resultaron totalmente dispares de la realidad. Por fortuna, Management Systems International comprendió la situación y aceptó un aumento del costo para que finalizase la iniciativa.

65. En casos como éste quizá sería aconsejable que los encargados de realizar la encuesta redujeran el personal, reteniendo sólo a quienes demuestren ser más eficientes, o recortaran costos de otras partidas; por ejemplo, utilizando material de oficina más barato o viajando en el transporte público en vez de en el privado, fusionando actividades, para reducir gastos generales, etcétera. Otra posibilidad aconsejable, si la autorizan quienes financian la encuesta, sería incluir en la exposición del proceso de la encuesta una nota en el sentido de que el alcance del presupuesto previsto podría modificarse debido a circunstancias imprevistas, lo que permitiría un cambio en el tamaño de la muestra acorde con el aumento de los costos.

G. Sistema de registro y resúmenes

66. Ya se ha indicado que el registro diario de los acontecimientos durante la encuesta es un requisito fundamental para tener constancia de todas las decisiones adoptadas y de las opciones consideradas al tomar esas decisiones. Ello supone, entre otras cosas, el registro de los gastos.

67. Los coordinadores de la encuesta deberían, en la fase de preparación de ésta, concebir una serie de formularios para uso de todos los empleados, con el fin de registrar las actividades y los gastos diarios con todo detalle. En estos formularios deberían incluirse datos concretos sobre las horas trabajadas, las tareas realizadas, los pormenores de las entrevistas, el transporte, etcétera, que resumirían una vez por semana. De esta manera se podrá supervisar atentamente el presupuesto e identificar los posibles problemas desde el comienzo. Además, debería instituirse, cuando sea posible, un sistema de pago sólo previa presentación de recibos válidos.

68. El seguimiento y registro de las actividades realizadas diariamente para la encuesta y sus correspondientes costos es una de las principales responsabilidades de la gestión de la encuesta. Pueden utilizarse diferentes formas de registro para las diferentes fases de la encuesta.

1. Diseño de la encuesta

69. Durante esta fase el director de la encuesta estará en estrecho contacto con todas las actividades, lo que simplificará bastante el seguimiento de las tareas. Un diario es un sistema útil para registrar quién ha hecho cada cosa. Esa información podría resumirse en un informe semanal. Puede mantenerse un registro paralelo de los costos efectivos por con-

cepto de transporte, bienes fungibles, alojamiento, etcétera, que se complementará con un resumen semanal para poder disponer de un informe sobre los costos a lo largo de la semana. En el anexo de este capítulo se presentan algunos ejemplos de formularios para mantener los registros diarios y semanales.

2. Realización de la encuesta

70. Durante la realización de la encuesta el director deberá contactar frecuentemente con los jefes de equipo sobre el terreno para que le comuniquen las actividades incluidas en el diario más los costos y las entradas registrados. Hay que repetir de nuevo que el director debería realizar un resumen semanal en que se especifiquen todos los costos y días trabajados por los miembros del equipo, para poder comprobar fácilmente el porcentaje del presupuesto utilizado. En el anexo pueden verse ejemplos de formularios que pueden utilizarse con este fin.

3. Presentación de informes sobre la encuesta

71. También durante esta fase el director de la encuesta deberá estar en contacto estrecho con las actividades. Un sistema de registro diario de las actividades y de los costos permitirá preparar un resumen semanal. En el anexo pueden verse algunos formularios para tal fin.

4. Comparación entre gastos y presupuesto

72. Conviene confiar a una persona la responsabilidad de comprobar en todo momento la situación de los gastos en relación al presupuesto. Esa persona debería presentar un panorama semanal de los gastos realizados hasta esa fecha, junto con las asignaciones presupuestarias (véase un ejemplo en el anexo). Si este mecanismo está en marcha desde el comienzo de la encuesta será bastante sencillo prever los problemas y solicitar reasignaciones del presupuesto en caso necesario. Los profesionales de las encuestas deben tener presente que el aumento del presupuesto una vez que ha comenzado la encuesta es una práctica muy excepcional y que, por lo tanto, para el éxito del producto final es fundamental tener la capacidad de efectuar los ajustes pertinentes.

H. Conclusiones

73. En este capítulo se han ofrecido algunos consejos útiles para planificar el presupuesto de las encuestas valiéndose de una consideración detallada de todos los componentes de las mismas. Se ha recomendado y mostrado con ejemplos un planteamiento dinámico que incorpora un mecanismo presupuestario desde dos puntos de vista.

74. Conviene destacar que una planificación detallada es fundamental para poder realizar con éxito un estudio transparente, fiable y de alta calidad. Igualmente importante es el registro diario de todas las actividades y acciones, que luego podrá incorporarse sin problemas al proceso contable y mantenerse como registro fiable al que poder recurrir para planificar futuras encuestas.

ANEXO

Ejemplos de formularios para los registros diarios y semanales

Registro de actividades diarias del personal			
Nombre			
Fecha	**Actividad**	**Lugar**	**Tiempo dedicado**
		Número total de días	

Registro diario de entrevistas				
Nombre del empadronador				
Fecha	**Lugar**	**No. de código de la entrevista**	**Tiempo dedicado**	**Resultado de la entrevista/Comentarios**

Resumen semanal de las actividades del personal			
Nombre del jefe de equipo			
Fecha del informe			
Nombre	**Resumen de las actividades**	**Lugar**	**Número total de días**

Registro de gastos diarios del personal					
Nombre					
Fecha	**Lugar**	**Actividad**	**Detalles del gasto**	**Cantidad (en dólares)**	**No. de recibo**
			Total (en dólares)		

Registro de gastos semanales					
Nombre del jefe de equipo					
Fecha del informe					
Nombre	**Lugar**	**Actividad**	**Detalles del gasto**	**Cantidad (en dólares)**	**No. de recibo**
			Total (en dólares)		

Resumen de gastos semanales						
Nombre						
Partida	**Semana 1**	**Semana 2**	**Semana 3**	**Semana 4**	**Semana 5**	**Semana 6**
Personal						
Sueldos/Salarios						
Alojamiento						
Comidas						
Otros						
Transporte						
Bienes fungibles						
Otros						

Resumen de gastos semanales *								
Partida	**Presupuesto**	**Gastos SING. O PLU. acumulado**	**Semana 1**	**Semana 2**	**Semana 3**	**Semana 4**	**Semana 5**	**Semana 6**
Personal								
Sueldos/salarios								
Alojamiento								
Comidas								
Otros								
Transporte								
Combustible								
Vehículos								
Público								
Otros								
Equipo								
Bienes fungibles								
Otros								
Total hasta la fecha								

* Puede hacerse en forma de hoja de cálculo (por ejemplo, con EXCEL).

REFERENCIAS BIBLIOGRÁFICAS

Ajayi, D. (2002). Comunicación personal.

Barnes, C. (2001). *An Assessment of Zambuko Trust Zimbabwe. Assessing the Impact of Microenterprise Services (AIMS) Project.* Washington, D.C.: Management Systems International.

_____ y E. Keogh (1999). *An Assessment of the Impact of Zambuko's Microfinance Program in Zimbabwe: Baseline Findings. Assessing the Impact of Microenterprise Services (AIMS) Paper.* Washington, D.C.: Management Systems International.

Greenfield, T. (1996). *Research Methods: Guidance for Postgraduates,* pág. 306. Nueva York: Arnold, John Wiley and Sons, Inc.

Groves, R. M. (1989). *Survey Errors and Survey Costs.* Nueva York: J. Wiley and Sons, Inc.

_____ y J. M. Lepkowski (1985). Dual frame mixed mode survey designs. *Journal of Official Statistics*, vol. 1, No. 3, págs. 263-286.

Naciones Unidas (1984). *Manual de encuestas sobre hogares (versión revisada).* Estudios de métodos, No. 31. No. de venta: S.83.XVII.13.

Scheaffer, R. L., W. Mendenhall y L. Ott (1990). *Elementary Survey Sampling* (4a. ed.), pág. 97. Belmont, California: Wadsworth, Inc.

Sección E

Análisis de datos de encuestas

Introducción

Graham Kalton
Westat
Rockville, Maryland
Estados Unidos de América

1. Una vez que se han recopilado los datos para una encuesta, deben prepararse para su análisis. Este paso tiene tres componentes importantes. En primer lugar, como se examina en el capítulo XV, deben adoptarse decisiones sobre qué formato permitía un análisis más eficaz de los datos teniendo en cuenta los medios informáticos disponibles y el programa de análisis que se va a utilizar. Los análisis de la encuesta suponen con frecuencia dos o más unidades de análisis diferentes; en particular, los hogares y las personas son unidades de análisis independientes en muchas encuestas de hogares. Por ello, los archivos de datos deben poder manipular con eficiencia estructuras jerárquicas. Por ejemplo, es preciso tener en cuenta que las personas están incluidas dentro de los hogares y que el número de personas varía de unos hogares a otros.

2. El segundo componente de la preparación de los datos es su depuración o edición. Inevitablemente, las respuestas de la encuesta contienen errores identificables de varias formas; por ejemplo, respuestas que se contradicen con otras o que quedan fuera del intervalo de posibilidades. Estos errores deben resolverse antes de iniciar los análisis (véase en el apartado C información más detallada sobre depuración de datos en el capítulo XV).

3. A este respecto, es importante ultimar el estado analítico de cada unidad de la muestra. Todas las unidades seleccionadas para la muestra deben colocarse en una de las siguientes categorías: informante, no informante admisible, unidad no admisible o unidad no informante de admisibilidad desconocida (véase el capítulo VIII). La clasificación como informante generalmente presupone algo más que la mera presencia de un cuestionario para la unidad incluida en la muestra. En general, para que la unidad pueda clasificarse como tal debe recolectarse un mínimo de datos aceptables. Por ello, para aclarar esta cuestión se necesita un examen de los cuestionarios. No obstante, téngase en cuenta que aun cuando una unidad deba figurar en el análisis como informante, puede haber algunas partidas sobre las que no se hayan obtenido respuestas aceptables. Para hacer frente a este problema puede utilizarse un método de imputación para asignar valores a las respuestas omitidas.

4. La determinación del estado analítico de todas las unidades de la muestra es condición necesaria para el último paso de la preparación de datos: la computación de las ponderaciones de la encuesta. Éstas se calculan en relación con cada una de las unidades de análisis. Como el punto de partida en la construcción de las ponderaciones es determinar las probabilidades de selección de todas las unidades de la muestra, es de vital importancia mantener con esmero los registros de probabilidades de selección durante el proceso de selección de la muestra. Las ponderaciones iniciales, o básicas, para las unidades de la muestra se calculan en forma de inversos de las probabilidades de selección de las unidades. Luego, las ponderaciones básicas de los informantes se ajustan para compensar los no informantes

admisibles y para una proporción de los no informantes cuya admisibilidad se desconoce. Muchas veces se introduce otro ajuste para conseguir que las distribuciones ajustadas de la media ponderada de algunas variables clave correspondan a distribuciones conocidas de estas variables disponibles en una fuente externa. La elaboración de ponderaciones se describe en los capítulos XV y XIX.

5. Un elemento importante en la preparación de datos es consignar en cada uno de los registros de datos de los informantes la información sobre el muestreo necesaria para el análisis. Se necesitan ponderaciones para cada una de las unidades informantes de análisis a fin de poder elaborar estimaciones válidas de parámetros de la población de la encuesta. Se necesita información sobre la UPM y el estrato de cada unidad informante para poder calcular correctamente los errores de muestreo en las estimaciones de la encuesta. (véase el capítulo XXI).

6. Hay dos diferencias entre los análisis de los datos de la encuesta y los descritos en los textos estadísticos habituales. Una es la necesidad de utilizar ponderaciones en los análisis de la encuesta con el fin de compensar las probabilidades de selección desiguales, la falta de respuesta y la falta de cobertura. Si no se introducen ponderaciones en los análisis, el resultado podría ser una estimación distorsionada de los valores de la población.

7. La segunda peculiaridad de los análisis de la encuesta es la necesidad de calcular los errores de muestreo para las estimaciones de la encuesta de tal manera que se tenga en cuenta el diseño muestral complejo. La teoría presentada en los textos estadísticos habituales supone en efecto un muestreo irrestricto, mientras que la mayor parte de las encuestas de hogares emplean el muestreo estratificado en varias etapas. En general, los errores de muestreo para las estimaciones a partir de una muestra estratificada en varias etapas son mayores que los de una muestra irrestricta del mismo tamaño, por lo que la aplicación de las fórmulas de los textos estadísticos habituales sobrevalorará la precisión de las estimaciones (véanse los capítulos VI, VII y XXI). Ello significa que los programas informáticos estándar de estadística producen estimaciones inválidas del error estándar en las estimaciones de las encuestas. No obstante, existen, afortunadamente, bastantes programas informáticos de análisis de encuestas que se pueden utilizar para obtener estimaciones adecuadas del error muestral a partir de datos de encuestas obtenidos con diseños muestrales complejos. En el capítulo XXI se examinan algunos de esos programas.

8. Gran parte del análisis realizado con las encuestas gubernamentales es de carácter descriptivo. Con frecuencia, los resultados se presentan en forma de cuadros, cuyas casillas contienen medias, porcentajes o totales; algunas veces se presentan en forma de gráficos. En términos estadísticos estrictos, las estimaciones aquí implicadas suelen ser muy sencillas, y el único problema es la necesidad de comprobar que se han utilizado las ponderaciones de la encuesta. No obstante, hay algunos problemas importantes de definición y presentación que se deben considerar. Es muy importante definir qué es lo que se quiere medir (por ejemplo, la pobreza: véase el capítulo XVII) y especificar el conjunto de unidades con respecto a las cuales se va a efectuar la medición, y hacerlo de forma adecuada para conseguir el objetivo deseado. Asimismo, los resultados deben presentarse de manera que se comunique claramente qué es lo que se ha medido y en relación con qué unidades. En el capítulo XVI se facilita orientación sobre la presentación de estimaciones descriptivas sencillas.

9. Muchas veces la realidad que se debe medir puede definirse en forma lógica relativamente clara, a partir de las respuestas ofrecidas. Otras veces resulta más complicado, y quizá haya que crear un índice que utilice métodos estadísticos de múltiples variables, como el análisis por conglomerados y el análisis del componente principal. En el capítulo XVIII se presentan varios ejemplos, entre ellos uno en que se construyó un índice de "riqueza" utilizando información de variables como la disponibilidad de corriente eléctrica, el número de personas por dormitorio y el tipo principal de agua potable.

10. Finalmente, conviene señalar que si bien la producción de estimaciones descriptivas continúa siendo la principal forma de análisis de encuestas, cada vez se utilizan más las técnicas analíticas con los datos de las encuestas. Estas técnicas se aplican con frecuencia al examen de las relaciones entre variables y a la búsqueda de posibles relaciones de causa-efecto. En la forma más frecuente de este tipo de análisis se construye un modelo estadístico para predecir mejor una variable dependiente en términos de un conjunto de variables independientes (o predictoras). Si la variable dependiente es continua (por ejemplo, el ingreso de los hogares), pueden utilizarse varios métodos de regresión lineal. Si se trata de una variable categórica con una respuesta binaria (por ejemplo, si el hogar tiene o no tiene agua corriente), pueden utilizarse métodos de regresión logística. Estos métodos, y los efectos del diseño muestral complejo, se describen en los capítulo XIX y XX. En el capítulo XIX se describe también el uso de modelos en varios niveles en el contexto de una encuesta, y en el capítulo XX se examina también el efecto de los diseños muestrales complejos en las pruebas de chi^2 de las asociaciones entre variables categóricas.

Capítulo XV
Guía para la gestión de datos de las encuestas de hogares

Juan Muñoz
Sistemas Integrales
Santiago, Chile

RESUMEN

En el presente capítulo se describe la función de la gestión de datos en el diseño y realización de las encuestas nacionales de hogares. En primer lugar se examina la relación entre la gestión de datos y el diseño del cuestionario, y luego se analizan las opciones pasadas, presentes y futuras para la entrada y edición de datos de la encuesta y sus consecuencias para la gestión de la encuesta en general. Los apartados que siguen orientan sobre la definición de los criterios de control de calidad y la elaboración de programas de entrada de datos para encuestas nacionales complejas de hogares, incluida la difusión de los conjuntos de datos resultantes. En el último apartado se examina la función de la gestión de datos como apoyo para la realización de un diseño de la muestra.

Términos clave: Comprobación de la coherencia, depuración de datos, edición de datos, gestión de datos, encuesta de hogares, criterios de control de calidad.

A. Introducción

1. Aunque la importancia de la gestión de datos en las encuestas de hogares se ha resaltado ya en muchas ocasiones, en general se considera todavía como un conjunto de tareas relacionadas con la fase de tabulación de la encuesta, en otras palabras, actividades que se realizan hacia el final del proyecto, que utilizan computadoras en el contexto aséptico de las oficinas de la sede central y que generalmente quedan bajo el control de los analistas de datos y los programadores informáticos.

2. Esta visión restrictiva de la gestión de datos de encuestas está cambiando. La experiencia de los dos últimos decenios revela que esa actividad puede y debe desempeñar un papel fundamental ya desde las primeras fases de la iniciativa. Cada vez es más claro que no se termina con la publicación de los primeros informes estadísticos.

3. La muestra más clara de esfuerzos eficaces de gestión de datos antes de las fases de análisis se encuentra en el Estudio de medición de los niveles de vida del Banco Mundial, y otras encuestas que han conseguido integrar con éxito controles de calidad basados en el uso de computadoras con operaciones sobre el terreno. Aun cuando la entrada de datos no se realice como parte del trabajo sobre el terreno, los gestores de datos deberían participar en el diseño de los cuestionarios para garantizar que las unidades estadísticas observadas por encuesta se reconozcan e identifiquen debidamente, que las instrucciones de salto para los

entrevistadores sean explícitas y correctas y que en los cuestionarios se incorporen repeticiones deliberadas que puedan luego aplicarse para controlar la coherencia.

4. En el otro extremo del calendario del proyecto de encuesta, la idea de que el producto final esperado de la encuesta es una publicación impresa, con una colección de cuadros estadísticos, ha dejado paso al concepto de una base de datos que no sólo puede ser utilizada por el organismo de estadística para preparar los cuadros iniciales sino que será también accesible a los investigadores, las autoridades y el público en general. El resumen descriptivo de los resultados de la encuesta ya no se considera el paso final, sino más bien como punto de partida de varias actividades de análisis que pueden durar muchos años, una vez que el proyecto se haya clausurado oficialmente y que se haya disuelto el equipo de la encuesta.

5. El apartado que sigue aborda las relaciones entre gestión de datos de la encuesta y diseño del cuestionario. En el apartado C se hace una exploración de las opciones pasadas, presentes y futuras para la entrada y la edición de datos, y sus consecuencias para la gestión de la encuesta en general, y en los apartados siguientes se presentan orientaciones para la definición de los criterios de control de calidad y el desarrollo de programas de entrada de datos para encuestas nacionales complejas de hogares, y la divulgación de los conjuntos de datos de la encuesta. En el último apartado se analiza la función de la gestión de datos como apoyo a la realización del diseño de la muestra.

B. Gestión de datos y diseño del cuestionario

6. La gestión de los datos de una encuesta comienza en paralelo con el diseño del cuestionario y puede influir en éste de manera considerable. Debe consultarse al gestor de datos sobre cada uno de los proyectos importantes de cuestionario, ya que él podrá captar con especial nitidez las deficiencias en la definición de las unidades de observación, pautas de salto, etcétera. En este apartado se examinan algunos de los aspectos formales del cuestionario que merecen especial atención.

7. *Naturaleza e identificación de las unidades estadísticas observadas.* Todas las encuestas de hogares recopilan información sobre una gran unidad estadística —el hogar— así como acerca de una variedad de unidades subordinadas dentro del hogar: personas, partidas presupuestarias, parcelas, cosechas, etcétera. El cuestionario debe indicar de forma clara y explícita cuáles son esas unidades y velar por que cada una de ellas se denomine adecuadamente con un identificador que permita identificarla de forma inequívoca.

8. La identificación del hogar suele aparecer en la cubierta del cuestionario. Algunas veces consiste en una larga serie de cifras y letras que representan la ubicación geográfica y los procedimientos de muestreo utilizados para seleccionar el hogar. Aunque pueda parecer obvio, el uso de todos estos códigos como identificadores del hogar debería examinarse con suma atención, ya que puede ser engorroso, dar lugar a errores y resultar muy caro (en muchos casos se necesitan 20 o más dígitos para identificar los pocos centenares de hogares de la muestra); en ocasiones no se llega ni siquiera a garantizar una identificación inequívoca, como cuando los códigos geográficos de la cubierta identifican la vivienda pero no consideran el caso de múltiples hogares en una vivienda. Una alternativa más sencilla y segura es identificar los hogares mediante un número de serie sencillo que se pueda escribir a mano o con un sello en la cubierta del cuestionario, o incluso que venga ya impreso de imprenta. La ubicación geográfica, la condición urbana/rural, los códigos de muestreo y el resto de los datos de la cubierta se convierten en importantes atributos del hogar, que en cuanto tal deben incluirse en los conjuntos de datos de la encuesta, pero no necesariamente como medio de identificación. Un compromiso aceptable entre esos dos extremos (la lista de todos los códigos detallados y un simple número de serie del hogar) sería asignar un número de serie de tres o cuatro dígitos a las unidades primarias de muestreo (UPM) utilizadas en la encuesta y luego un número de serie de dos dígitos a los hogares dentro de cada UPM.

9. La naturaleza de las unidades estadísticas subordinadas es muchas veces obvia (por ejemplo, los miembros de los hogares son personas individuales), pero pueden presentarse ambigüedades cuando lo que parece una unidad individual es de hecho un conjunto de unidades de diferentes tipos. Así puede ocurrir, por ejemplo, cuando un hombre a quien se ha pedido que informe sobre la actividad principal de su empleo realiza simultáneamente varias actividades, igualmente importantes, o tiene más de un empleo en un determinado período de referencia. De la misma manera, puede haber ambigüedad cuando se pregunta el género o peso del último hijo a una mujer que ha tenido dos gemelos —un niño y una niña— con pesos diferentes. No obstante, aunque estas situaciones deben evitarse mediante un diseño adecuado del cuestionario y la realización de pruebas piloto, muchas veces se presentan por desapercibimiento. Ése es el momento en que la visión crítica de un gestor de datos experimentado puede ofrecer una asistencia inapreciable a los especialistas en la materia para hacer aflorar los problemas.

10. Cualquiera que sea su naturaleza, las unidades subordinadas dentro del hogar deberían estar siempre identificadas de forma inequívoca. Ello puede conseguirse mediante códigos numéricos asignados por el entrevistador, pero en general conviene que siempre que sea posible estos identificadores estén ya impresos en el cuestionario.

11. *Repeticiones deliberadas*. Al diseñar el cuestionario se puede considerar la posibilidad de incluir repeticiones deliberadas, con el fin de descubrir los errores cometidos por el entrevistador o en la entrada de datos. Los ejemplos más comunes son los siguientes:

- Agregar una última línea correspondiente a los "totales" debajo de las columnas que contienen valores monetarios. El cálculo de las cifras totales puede ser competencia del entrevistador, pero aun cuando así no fuera, su inclusión es conveniente, ya que constituye una manera muy eficaz (a veces la única) de detectar errores u omisiones en la entrada de datos. De hecho, los totales pueden agregarse como medio de control de calidad al final de cada columna numérica, aun cuando la suma no represente una medida significativa de magnitud; por ejemplo, puede agregarse un total al final de una columna que contenga la cantidad (no los valores monetarios) de varios productos alimenticios comprados, aun cuando ello signifique la suma de valores heterogéneos, como kilos de pan y de patatas o incluso litros de leche. Este punto se explica con mayor detalle al hablar de las comprobaciones tipográficas.
- Agregar un dígito de comprobación a los códigos de algunas variables importantes (como la ocupación o actividad de una persona, o el carácter del producto de consumo). Un dígito de comprobación es un número o una letra que puede quitarse del resto de los dígitos del código mediante operaciones matemáticas realizadas en el momento de la entrada de datos. Un algoritmo frecuente utilizado con este fin es el siguiente: multiplicar el último dígito del código por 2, el penúltimo por 3, etcétera (si el código tiene más de seis dígitos, se repite la secuencia de multiplicadores 2, 3, 4, 5, 6, 7), y sumar los resultados. El código de comprobación es la diferencia entre esta suma y el múltiplo superior más próximo de 11 (el número 10 está representado por la letra K). Los algoritmos del dígito de comprobación se introducen para que los errores más frecuentes de codificación, como la transposición u omisión de dígitos, produzcan el dígito de comprobación erróneo.

C. Estrategias operacionales para la entrada y la edición de los datos

12. Muchas encuestas de hogares consideran todavía la entrada y la edición de los datos como actividades que deben realizarse en un lugar centralizado una vez finalizadas las operaciones sobre el terreno, mientras que en otros casos se está adoptando ya el concepto de integración de la entrada de datos en las operaciones sobre el terreno. En el futuro próximo,

esta idea puede llevar a la aplicación de entrevistas con asistencia de computadora. En este apartado se examinan las consecuencias organizativas de varias estrategias y las características comunes y específicas de los programas de entrada y edición de datos elaborados en el marco de cada alternativa.

13. *Entrada de datos centralizada.* La entrada de datos centralizada era la única opción conocida antes de la aparición de las microcomputadoras, y todavía se utiliza en muchas encuestas. Considera la entrada de datos como un proceso industrial y centralizado que comenzaría una vez finalizadas las entrevistas. El objetivo de la operación es transformar la materia prima (la información recogida en los cuestionarios impresos) en un producto intermedio (archivo de lectura mecánica) que debe elaborarse ulteriormente (mediante programas de edición y procesos manuales) con el fin de obtener como producto final lo que se conoce como base de datos "limpia".

14. Durante la fase inicial de entrada de datos, las prioridades son agilizar el proceso y garantizar que la información de los archivos corresponda perfectamente con la recopilada en los cuestionarios. Lo que se pide a los encargados de la entrada de datos no es que "piensen" sobre lo que están haciendo, sino que se limiten a copiar los datos que hayan recibido. Algunas veces los cuestionarios se someten a un proceso de doble anonimato o "doble ciego" con el fin de comprobar que los resultados son correctos.

15. Hasta mediados del decenio de 1970, la entrada de datos se realizaba con máquinas especializadas que tenían capacidades muy limitadas. Aunque en la actualidad casi siempre se utilizan microcomputadoras que pueden programarse, con comprobaciones de control de calidad, esta capacidad no suele utilizarse en la práctica. Predomina la opinión de que en el proceso de entrada de datos deben introducirse pocas comprobaciones de ese tipo, ya que los operadores no están capacitados para tomar decisiones sobre qué hacer si aparece un error. Además, la detección de errores y sus soluciones frena el proceso de entrada de datos. Según esta opinión, las comprobaciones de control de calidad deberían reservarse exclusivamente para el proceso de edición.

16. *Entrada de datos sobre el terreno.* A partir de mediados del decenio de 1980, la integración de los controles de calidad mediante computadora en las operaciones ha sido una de las claves para mejorar la calidad y oportunidad de las encuestas de hogares. Estas ideas fueron propuestas inicialmente en las encuestas del Estudio de medición de los niveles de vida (EMNV), del Banco Mundial, y posteriormente se han aplicado a otras encuestas de hogares complejas. En esta estrategia se aplican controles de entrada de datos y de coherencia hogar por hogar en el marco de las operaciones sobre el terreno, para poder corregir las contradicciones mediante nuevas visitas a los hogares.

17. El beneficio más importante y directo de la integración es que mejora significativamente la calidad de la información, ya que permite corregir los errores y contradicciones mientras los entrevistadores se encuentran todavía sobre el terreno, en vez de "depurarlos" posteriormente desde la sede. Además de ser prolongados y tediosos, los procedimientos de depuración centralizada permiten obtener en el mejor de los casos bases de datos que son internamente coherentes pero que no reflejan necesariamente las realidades observadas sobre el terreno. La incertidumbre es resultado del número infinito de decisiones —en general no documentadas— que deben tomarse lejos de donde se han recopilado los datos y mucho después de esa recopilación.

18. La integración de controles de calidad mediante computadora puede generar también bases de datos que se presten fácilmente a tabulación y análisis en forma oportuna, generalmente pocas semanas después de finalizadas las operaciones sobre el terreno. De hecho, las bases de datos pueden prepararse incluso mientras se realiza la encuesta, lo que ofrecería a los directivos la posibilidad de supervisar eficazmente las operaciones sobre el terreno.

19. Otra ventaja indirecta de la integración es que fomenta la aplicación de criterios uniformes por todos los entrevistadores a lo largo de todo el período de recopilación de datos, objetivo difícil de alcanzar en la práctica con los métodos anteriores a la integración. La computadora se ha convertido para los supervisores de encuestas en un auxiliar incorruptible e incansable.

20. La integración de controles de calidad mediante computadora en las operaciones sobre el terreno tiene también varias consecuencias para la organización de la encuesta, la más importante de las cuales es que el personal sobre el terreno debe organizarse en equipos. Éstos normalmente son presididos por un supervisor y cuentan con un operador de entrada de datos, además de entre dos y cuatro entrevistadores.

21. La organización de operaciones sobre el terreno depende de las opciones tecnológicas disponibles. Las dos configuraciones más utilizadas suponen el uso de computadoras de escritorio y portátiles, y requieren los siguientes pasos:

- El operador encargado de la entrada de datos puede utilizar una computadora de escritorio en un lugar fijo (generalmente una oficina regional del organismo de estadística) y el trabajo sobre el terreno debe organizarse de manera que el resto del equipo visite cada lugar de la encuesta (en general, una unidad primaria de muestreo) al menos dos veces, para dar al operador tiempo de introducir y verificar la coherencia de los datos entre las visitas. En la segunda y posteriores visitas, los entrevistadores volverán a formular las preguntas en los casos en que el programa de entrada de datos haya detectado errores, omisiones o contradicciones.
- El operador encargado de la entrada de datos debe trabajar con una computadora portátil y sumarse al resto del equipo en sus visitas a los lugares de la encuesta. Todo el equipo permanece en el lugar hasta que se han introducido todos los datos y el programa de entrada de datos certifica que son completos y correctos.

22. Ambas opciones tienen requisitos externos que deben ser examinados atentamente por los encargados de la planificación y gestión de la encuesta. Uno de ellos es garantizar un suministro permanente de energía eléctrica para las computadoras, lo que puede representar un problema en algunos países donde el servicio eléctrico no está suficientemente desarrollado. Si se utilizan computadoras de escritorio en lugares fijos, quizá haya que instalar generadores y garantizar la disponibilidad constante de combustible para éstos. Si se utilizan computadoras portátiles, quizá haya que utilizar paneles solares portátiles.

23. Una diferencia obvia pero importante entre las dos estrategias es que si los controles de calidad mediante computadora se van a integrar en el trabajo sobre el terreno, el programa de entrada y edición de datos debe elaborarse y ponerse a punto antes de que comience la encuesta. En el caso de un sistema centralizado de entrada de datos, ello puede resultar conveniente (para poder avanzar en paralelo con las operaciones sobre el terreno), pero no es absolutamente necesario.

24. *Entrevistas sin cuestionarios en papel.* El uso de computadoras portátiles para prescindir por completo de los cuestionarios impresos resulta muy atractivo, ya que permite automatizar algunas partes de las entrevistas, como las instrucciones de salto. No obstante, aunque esta tecnología existe ya desde hace casi 20 años, se ha hecho muy poco para aplicarla seriamente a las encuestas de hogares complejas en los países en desarrollo. De hecho, incluso en los organismos nacionales de estadística más avanzados, los cuestionarios sin papel se han limitado hasta ahora a actividades relativamente sencillas, como las encuestas de empleo y la recopilación de precios para el índice de precios de consumo.

25. Una de las posibles razones es que, aunque las entrevistas sin papel pueden realizarse en forma lineal, con un principio y un fin, muchas encuestas de hogares en los países en desarrollo y en transición quizá necesiten múltiples visitas a cada hogar, entrevistas por

separado con cada miembro del hogar u otros procedimientos estructurados de forma menos estricta.

26. A pesar de la ausencia de experiencia empírica real, pueden formularse algunas observaciones sobre lo que debe tenerse en cuenta en el diseño y aplicación de un cuestionario sin papel:

- La interfaz del programa de entrada de datos consistirá en algunos casos en una serie de preguntas que aparecen una después de otra en la pantalla de la computadora, pero en otros habrá que reproducir la estructura y formato visual de los cuestionarios de papel, en los que se observan muchos campos de entrada de datos al mismo tiempo. Ello parece ser especialmente importante en los módulos sobre gasto y consumo, en que el entrevistador necesita "ver" simultáneamente muchas partidas de consumo. La interfaz debe ofrecer también la posibilidad de introducir preguntas en caso de dudas y de regresar al hogar para una segunda entrevista sin tener que repetir todas las preguntas.
- El proceso de diseño del cuestionario presupone en general muchos meses de trabajo y la participación de muchas personas diferentes (especialistas en el tema, profesionales de las encuestas, etcétera). Con un cuestionario de papel, el proceso se lleva a cabo preparando, distribuyendo, examinando y comprobando varias "generaciones" del cuestionario hasta que se llega a un acuerdo sobre la versión final. Hay que determinar todavía los pasos equivalentes en un proceso que nunca se reflejará de hecho sobre el papel.
- La capacitación de los entrevistadores deberá reajustarse en torno a la nueva tecnología. Sabemos cómo capacitar entrevistadores para que administren un cuestionario de papel (sesiones teóricas, simulaciones, entrevistas ficticias, manuales de capacitación, etcétera), pero se ha hecho poco para elaborar técnicas equivalentes con destino a una encuesta sin papel.
- Finalmente, hay que establecer métodos eficaces de supervisión. A lo largo de medio siglo se han acumulado procedimientos numerosos y diversos (inspección visual de los cuestionarios, entrevistas de comprobación, etcétera) para verificar el trabajo realizado por los entrevistadores sobre el terreno. Todos ellos están basados en el concepto de un cuestionario de papel, y sería entonces preciso adaptarlos al nuevo tipo de entrevistas. Es muy probable que las nuevas tecnologías ofrezcan opciones completamente diferentes —y quizá mucho más productivas— para una supervisión eficaz; por ejemplo, la mayor parte de las computadoras portátiles ofrecen posibilidades de grabar la voz, lo que podría utilizarse para registrar en forma automática y aleatoria algunas partes de la entrevista junto con los archivos de datos. La introducción del sistema de posicionamiento global (GPS) permite registrar automáticamente la hora y el lugar de las entrevistas. También en este caso hay que establecer y comprobar sobre el terreno los detalles e incorporarlos al plan general de trabajo sobre el terreno.

D. Criterios de control de calidad

27. Cualquiera que sea la estrategia elegida para el control de calidad, los datos de los cuestionarios deben someterse a cinco tipos de comprobación: de límites, de relación con los datos de referencia, de salto, de coherencia y de los errores tipográficos. Aquí examinaremos el carácter de estas comprobaciones y cómo pueden aplicarse en diferentes circunstancias operacionales.

28. Los *controles de límite* tratan de garantizar que todas las variables de la encuesta contengan únicamente datos dentro de un campo limitado de valores válidos. Las variables

categóricas pueden tener sólo uno de los valores predefinidos para ellas en el cuestionario (por ejemplo, el género puede codificarse únicamente como "1" en el caso de los varones o "2" en el de las mujeres); las variables cronológicas deberían contener fechas válidas, y las variables numéricas deberían situarse entre los valores mínimo y máximo prescritos (por ejemplo, entre 0 y 95 años de edad).

29. Un caso especial de control de límites se produce cuando los datos de dos o más campos estrechamente relacionados pueden compararse *con cuadros de referencia externos*. Son frecuentes las siguientes situaciones:

- *Coherencia de datos antropométricos.* En este caso, los valores registrados de altura, peso y edad se comparan con los cuadros de referencia estándar de la Organización Mundial de la Salud. Todo valor para los indicadores estándar (altura-edad, peso-edad y peso-altura) que se alejen más de tres desviaciones estándar de la norma deberán marcarse como posible error, para que se pueda repetir la medición.
- *Coherencia de los datos sobre el consumo de alimentos.* En este caso los valores registrados para el código de alimentos, la cantidad comprada y el total pagado se comparan con un cuadro por partidas específicas de posibles precios unitarios.

30. Aun cuando los datos se introduzcan en una oficina centralizada, normalmente conviene buscar y corregir los errores graves en la fase inicial, en vez de aplazar este control a la fase de edición, ya que los errores de límite muchas veces son consecuencia del funcionamiento mismo de la entrada de datos más que de errores del entrevistador. Una posibilidad es que aparezca una señal de alerta; por ejemplo, que suene un pitido o que el texto parpadee en la pantalla cuando se introduce un valor que supera los límites mínimo y máximo y que, por tanto, se encuentra fuera del intervalo. Si el error es meramente tipográfico, el operador de la entrada de datos puede corregirlo inmediatamente. No obstante, debería ser posible hacer caso omiso de la señal de alerta si el valor introducido representa lo que aparece en el cuestionario. En tal caso debe hacerse constar, para que el error pueda ser corregido por el personal administrativo al examinar el cuestionario (o por el entrevistador durante una segunda entrevista, si los datos se están introduciendo sobre el terreno). Mientras tanto, el dato sospechoso puede almacenarse con un formato especial que ponga de manifiesto su dudoso valor.

31. *Comprobaciones de salto.* Su objetivo es verificar si las pautas de salto se han seguido adecuadamente. Por ejemplo, una comprobación sencilla verifica que las cuestiones que deben preguntarse sólo a los niños de las escuelas no se registran en el caso de un niño que ha respondido negativamente a la pregunta inicial sobre si está o no matriculado. Una comprobación más complicada verificaría que se han rellenado los módulos adecuados del cuestionario para cada uno de los informantes. Según su edad y género, cada miembro del hogar se supone que debe responder (o pasar por alto) determinadas secciones del cuestionario. Por ejemplo, los niños de menos de 5 años de edad deberían medirse en la sección antropométrica, pero no debería hacérseles preguntas sobre su ocupación. En la sección sobre la fecundidad podrían incluirse las mujeres de 15 a 49 años, pero no los hombres.

32. Es posible que en el futuro las entrevistas con ayuda de computadora (sin papel) sean más frecuentes en las encuestas de los países en desarrollo, con lo que el mecanismo de salto será controlado posiblemente por el mismo programa de entrada de datos, al menos en algunos casos. No obstante, en otros contextos operacionales (entradas de datos en oficinas centrales y sobre el terreno) el programa de entrada de datos no debería seguir por sí solo las pautas de salto. Por ejemplo, si se responde negativamente a la pregunta "¿Está matriculado?", los campos donde deben introducirse los datos sobre el tipo de escuela, el nivel alcanzado, etcétera, deberían presentarse no obstante al operador de la entrada de datos. Si hay respuestas realmente registradas en el cuestionario, pueden introducirse, y el programa enviará una alerta de salto incorrecto. El supervisor o el entrevistador (o el personal admi-

nistrativo centralizado encargado de la edición) puede determinar el carácter del error posteriormente. Puede ocurrir que la respuesta negativa quizá debería haber sido afirmativa. Si el programa de entrada de datos hubiera saltado automáticamente los siguientes campos, el error no habría saltado a la vista ni corregirse.

33. *Comprobaciones de coherencia.* Estas comprobaciones verifican que los valores de una pregunta estén en consonancia con los de otras. Una comprobación sencilla es la que se produce cuando ambos valores son de la misma unidad estadística; por ejemplo, la fecha de nacimiento y la edad de un determinado individuo. Las comprobaciones más complicadas suponen una comparación de informaciones procedentes de dos o más diferentes unidades de observación.

34. No hay ningún límite natural impuesto con respecto a la cantidad de comprobaciones de coherencia que pueden definirse. Las versiones bien preparadas del programa de entrada de datos para una encuesta compleja de hogares pueden tener centenares de ellas. En general, cuantas más comprobaciones se definan, mayor será la calidad del conjunto de datos final. No obstante, dado que el tiempo disponible para desarrollar los programas de entrada y de edición de datos es siempre limitado (normalmente unos dos meses), se necesita experiencia y buen criterio para decidir exactamente qué es lo que se debe incluir. Algunas comprobaciones de coherencia que pueden aplicarse a casi todas las encuestas de hogares han demostrado ser especialmente eficaces, y por ello se han convertido de hecho en la norma. En ellas se integran los siguientes elementos:

- *Coherencia demográfica del hogar.* La coherencia entre las edades y géneros de todos los miembros del hogar se comprueba con el fin de determinar las relaciones de parentesco. Por ejemplo, los padres deberían ser al menos 15 años mayores que sus hijos, los cónyuges deberían ser de géneros diferentes, etcétera.
- *Coherencia de ocupaciones.* La presencia o ausencia de determinadas secciones deberá estar en consonancia con las ocupaciones declaradas individualmente por los miembros del hogar. Por ejemplo, la sección sobre agricultura debería estar presente si y sólo si algunos miembros del hogar aparecen registrados como agricultores en la sección de mano de obra.
- *Coherencia de la edad y otras características individuales.* Es posible comprobar que la edad de cada persona está de acuerdo con características personales como el estado civil, la relación con el jefe de hogar, curso en que se ha matriculado (en el caso de los niños actualmente escolarizados) o último curso aprobado (en el caso de quienes han abandonado la escuela). Por ejemplo, un niño de 8 años de edad no debería estar por encima del tercer curso elemental.
- *Gastos.* En este caso son posibles varias comprobaciones de coherencia. Sólo en un hogar donde uno o varios de los registros individuales indiquen que hay un niño escolarizado debería haber números positivos en el registro de consumo de artículos como libros de texto y gastos de matrícula. De la misma manera, sólo los hogares que tienen servicios eléctricos deberían registrar gastos de electricidad.
- *Totales de control.* Como ya se ha dicho, la adición de un total de control siempre que haya una lista de números es un principio sabio en el diseño de cuestionarios. El programa de entrada de datos debería comprobar que el total de control equivale a la suma de los números individuales.

35. *Comprobaciones tipográficas.* En el pasado, la comprobación de errores tipográficos era casi el único control de calidad realizado en el momento de la entrada de datos. Ello se conseguía generalmente por el simple procedimiento de introducir cada cuestionario dos veces, encargándose de ello un operador distinto cada vez. Estos procedimientos de "doble ciego" no suelen utilizarse actualmente, ya que ahora existen otros sistemas de control de

coherencia que los hacen redundantes. No obstante, es posible que en algunos casos ese optimismo sea más fruto de una ilusión que de una realidad documentada.

36. Un error tipográfico habitual consiste en la transposición de dígitos (por ejemplo introducir "14" en vez de "41"). Este error podría detectarse en el caso de la edad mediante comprobaciones de coherencia con el estado civil o las relaciones familiares. Por ejemplo, el cuestionario de un adulto casado o viudo de 41 años cuya edad figura erróneamente como 14 años enviará un mensaje de error cuando se constate su estado civil. No obstante, el mismo error en el gasto mensual en carne puede pasar fácilmente inadvertido, ya que nada impide que la cifra real sea 14 ó 41 dólares.

37. Ello demuestra la importancia de incorporar perspectivas de gestión de datos en la fase de diseño del cuestionario. Los totales de control, por ejemplo, pueden reducir significativamente los errores tipográficos, ya que pedir al entrevistador que sume las cifras con una calculadora de bolsillo es semejante a introducirlas con procedimientos de doble ciego. Con este fin pueden utilizarse igualmente dígitos de comprobación en algunas variables importantes. Es también posible aplicar métodos reales de doble ciego para introducir los datos de algunas partes del cuestionario, pero intentar hacerlo en todo el cuestionario es innecesario e inviable, entre otras razones porque las estrategias modernas de entrada de datos se basan generalmente en el trabajo de un único operador de entrada de datos, no en el de dos operadores diferentes.

E. Desarrollo del programa de entrada de datos

38. El desarrollo de un buen programa de entrada y de edición de datos de encuestas es una técnica y un oficio. En esta sección se examinan algunas de las plataformas de desarrollo disponibles en la actualidad para facilitar los aspectos técnicos del proceso y algunas de las cuestiones más sutiles relativas al diseño de interfaces para los operadores de la entrada de datos y los futuros usuarios de los datos de la encuesta.

39. *Plataformas de desarrollo.* Hay muchas plataformas de desarrollo de programas de entrada y de edición de datos en el mercado, pero pocos de ellos están adaptados expresamente a las necesidades de gestión de datos de las encuestas complejas de hogares. Un examen del Banco Mundial realizado a mediados del decenio de 1990 permitió comprobar que en esas fechas dos plataformas basadas en DOS eran adecuadas: el programa del Estudio de medición de los niveles de vida (EMNV) desarrollado internamente por el Banco Mundial y el Sistema integrado para el procesamiento mediante microcomputadoras (*Integrated Microcomputer Processing System*, IMPS) de la Oficina del Censo de los Estados Unidos. Ambas plataformas han progresado desde que se llevó a cabo ese examen, en respuesta a los avances del equipo informático y de los sistemas operativos. El IMPS ha dejado paso al Sistema de procesamiento de censos y encuestas (*Census and Survey Processing System*, CSPro), aplicación en Windows que aporta algunas capacidades de tabulación, además de realizar su función primaria de entorno para el desarrollo de programas de entrada y edición de datos. El programa del EMNV se ha transformado en el LSD-2000, aplicación basada en Excel que trata de desarrollar el cuestionario de las encuestas y el programa de entrada de datos de forma simultánea.

40. Tanto CSPro como LSD-2000 (o sus antecedentes) han demostrado su capacidad de contribuir al desarrollo de programas eficaces de entrada y de edición de datos para las encuestas nacionales complejas de hogares en muchos países. Estas plataformas son fáciles de obtener y utilizar. Casi todo programador —de hecho, casi cualquiera básicamente familiarizado con las computadoras— puede adquirir en un par de semanas la capacidad técnica necesaria para iniciar el desarrollo de un programa práctico de entrada de datos.

41. *Principios de diseño.* Por desgracia, esas plataformas no pueden aconsejar a los programadores qué programa de entrada de datos se debe desarrollar. Podría argumentarse incluso que la facilidad de manejo para los usuarios corre el riesgo de hacer demasiado fácil el desarrollo de programas inadecuados de entrada de datos. La confusión del dominio de los instrumentos con la capacidad de utilizar esas herramientas con buen fin es un error que los administradores de la encuesta deben evitar integrando a programadores experimentados y a especialistas en la materia en el desarrollo de los programas de entrada y edición de los datos de la encuesta. En este sentido pueden ser útiles las siguientes orientaciones prácticas:

- *Diseño de la pantalla de entrada de datos.* Las pantallas de entrada de datos deben tener una apariencia lo más semejante posible a las páginas correspondientes del cuestionario, pero esta norma tiene muchas excepciones. Por ejemplo, si el cuestionario presenta preguntas personales en forma de matriz (con preguntas en las filas y los miembros de los hogares en las columnas, o al revés), en general es mejor preparar una pantalla independiente de entrada de datos para cada persona, en vez de reproducir las cuadrículas del papel en la pantalla de la computadora. Una razón es que el número de informantes es variable. Otra razón más poderosa es que las unidades estadísticas observadas son personas, no hogares.
- *Distinguir entre situaciones imposibles e improbables.* El programa de entrada de datos debería marcar como error cualquier situación que represente una imposibilidad lógica o natural (como una niña que tuviera más años que su madre), pero debe reaccionar también ante situaciones que no son naturalmente imposibles pero que resultan muy poco probables (como el que una niña sea menos de 15 años más joven que su madre). En principio, el programa de entrada de datos debería evaluar la gravedad de los errores y reaccionar de distinta manera según los casos, lo mismo que un supervisor humano haría si estuviera inspeccionando visualmente el cuestionario. Este tipo de programación "inteligente" es especialmente importante cuando la entrada de datos se integra en las operaciones sobre el terreno. Por desgracia, algunos programadores no dedican el esfuerzo suficiente a esta cuestión. Una señal reveladora es la tendencia a definir siempre el tramo superior de las variables cuantitativas como "999..." (con tantos nueves como admite la longitud del campo de entrada de datos). Esta práctica es contraproducente por razones obvias: los campos de entrada de datos deberían ser, naturalmente, lo bastante largos como para poder introducir incluso los mayores valores posibles, pero los márgenes superiores deberían ser lo bastante pequeños como para presentar como anomalías los valores improbables.
- *Forma de señalar el error.* Algunos de los criterios de control de calidad que incluye el programa de entrada de datos pueden señalar los errores por medios relativamente sencillos autoevidentes o que requieren poca capacitación para entenderlos. Por ejemplo, el programa EMNV señala los errores de límites con flechas que parpadeantes, hacia arriba ↑↑↑ o hacia abajo ↓↓↓ junto con el valor sospechoso, según se considere que es demasiado bajo o demasiado alto. No obstante, los controles de coherencia más complejos requieren una información mucho más clara y explícita. Por ejemplo, la comprobación de la coherencia demográfica del hogar podría dar lugar a un texto como el siguiente: "Advertencia: Lucy (código de identificación 05, niña de 9 años de edad) difícilmente puede ser la hija de Mary (código de identificación 02, mujer de 21 años de edad)", y el texto debería aparecer impreso y no sólo en la pantalla de la computadora. Este tipo de programación "inteligente y explícito" puede exigir más tiempo que otras alternativas aparentemente más sencillas (como el uso de códigos de error), pero ahorrará muchas horas de trabajo sobre el terreno y de capacitación del personal y liberará a los propios programadores de la carga de escribir una lista con códigos de error.

- *Códigos variables.* Una encuesta compleja de hogares normalmente contiene centenares de variables. Los programadores encargados del programa de entrada de datos deberán hacer referencia a ellos mediante códigos, de acuerdo con convenciones específicas de la plataforma utilizada. Es importante seleccionar para este fin un sistema racional y sencillo de codificación desde el comienzo del proceso de desarrollo del programa de entrada de datos, ya que facilitará la comunicación entre los miembros del equipo de desarrollo y porque permitirá ahorrar tiempo en los casos posteriores de preparación y divulgación de los conjuntos de datos de la encuesta. No obstante, la búsqueda de un sistema de codificación válido puede ser más difícil de lo que parece. El proceso puede comenzar sin dificultad, asignando a las primeras variables códigos como "EDAD", "GÉNERO", etcétera, pero poco después puede resultar inmanejable, ya que resulta cada vez más difícil encontrar códigos adecuados que puedan recordarse fácilmente. Una opción válida es hacer referencia simplemente a los números de sección y de pregunta que aparecen en el cuestionario, sin tratar en absoluto de que los códigos sean autoexplicativos (por ejemplo, si "edad" y "género" son las variables 4 y 5 de la sección 1, podrían codificarse como "S1Q4" y "S1Q5", respectivamente).
- *Distribución del trabajo de entrada de datos.* Cuando la entrada de datos se integra en las operaciones sobre el terreno, la unidad de trabajo más natural para la entrada de datos es el hogar. La razón es que en esas condiciones el operador de la entrada de datos tiene siempre sólo un cuestionario o un reducido número de ellos, con que trabajar, y también porque los controles de coherencia y la comunicación de errores se realizan hogar por hogar. En una oficina centralizada de entrada de datos, la unidad de trabajo pueden ser bloques de 10-20 hogares (como los lugares de la encuesta o las UPM). Lo que se propone es: *a*) el bloque debería ser introducido por un único operador de entrada de datos en una única computadora en un plazo máximo de dos días, y *b*) el conjunto correspondiente de cuestionarios debería poder almacenarse y recuperarse fácilmente en cualquier momento.

F. Organización y divulgación de los conjuntos de datos de la encuesta

42. La estructura de los conjuntos de datos de la encuesta debe reflejar la naturaleza de las unidades estadísticas observadas por la encuesta. En otras palabras, los datos de una encuesta compleja de hogares **no pueden** almacenarse en la forma presentada en el cuadro XV.1, es decir, como un sencillo archivo rectangular con una fila para cada hogar y columnas para cada uno de los campos que aparecen en el cuestionario.

Cuadro XV.1
Datos de una encuesta de hogares almacenados en forma de sencillo archivo rectangular

	Variable 1	**Variable 2**		**…**		**Variable *j***		**…**		**Variable *m***
Hogar 1				…				…		
Hogar 2				…				…		
…	…	…		…		…		…		…
Hogar *i*				…		Dato *i,j*		…		
…	…	…		…		…		…		…
Hogar *n*				…				…		

43. Esta estructura (conocida también con el nombre de "archivo plano") sería adecuada si todas las preguntas hicieran referencia al hogar en cuanto unidad estadística, pero, como ya se ha explicado, no es ése el caso. Algunas de las preguntas se refieren a unidades estadísticas subordinadas que aparecen en número variable dentro de cada hogar, como personas, cosechas, artículos de consumo, etcétera. El almacenamiento de la edad y del género de cada miembro del hogar como variables diferentes del hogar sería un desperdicio de recursos (ya que el número de variables necesarias se definiría en virtud del tamaño del hogar más numeroso en vez del hogar de tamaño medio) y resultaría sumamente engorroso en la fase de análisis (ya que incluso las tareas sencillas, como obtener la distribución por edad-género, implicarían un laborioso examen de un número variable de pares edad-género en cada hogar).

44. Las plataformas CSPro y LSD-2000 utilizan una estructura de archivos que resuelve bien las complejidades que se plantean al tratar muchas unidades estadísticas diferentes, al mismo tiempo que se reducen las necesidades de almacenamiento y se mantiene una interfaz adecuada con los programas informáticos de estadística en la fase de análisis.

45. La estructura de datos continúa siendo una correspondencia unívoca entre cada unidad estadística observada y los registros de los archivos de la computadora, utilizando un tipo de registro diferente para cada tipo de unidad estadística. Por ejemplo, para manipular los datos recogidos en la lista de hogares se definiría un tipo de registro para las variables de la lista, y los datos correspondientes a cada individuo se almacenarían en un registro independiente del mismo tipo. Igualmente, en el módulo de consumo de alimentos, un tipo de registro correspondería a los productos alimenticios y los datos correspondientes a cada producto concreto se almacenarían en registros independientes del mismo tipo.

46. El número de registro en cada uno de los diferentes tipos puede variar. Con ello se economiza espacio de almacenamiento, ya que los archivos no tienen que prever que cada uno de los casos adquiera la mayor dimensión posible.

47. En principio, sólo se necesita un tipo de registro para cada unidad estadística, aunque algunas veces puede definirse más de uno para la misma unidad, por razones prácticas. Por ejemplo, las preguntas sobre educación y salud pueden almacenarse en dos tipos de registro independientes, aun cuando la unidad estadística sea la persona en ambos casos.

48. Cada registro se identifica de manera inequívoca mediante un código con tres o más partes. La primera parte es el "tipo de registro". En ella se indica si la información procede, por ejemplo, de la portada o del módulo de salud o corresponde a los gastos en alimentos. El tipo de registro es seguido —en todos los registros— por el número del hogar. En la mayoría de los casos, será necesario un identificador que distinga entre unidades estadísticas independientes del mismo tipo dentro del hogar, por ejemplo, el número identificación de la persona o el código de la partida de gastos. En algunos casos habrá sólo una unidad correspondiente al nivel de observación, por lo que el tercer identificador no será necesario. Por ejemplo, las características de la vivienda normalmente se recopilan en una sola casa por hogar. En algunos casos puede haber un cuarto código adicional. Por ejemplo, el tercer identificador podría ser la empresa familiar, y el cuarto código se aplicaría a cada parte del equipo que es propiedad de cada empresa.

49. Después de los identificadores, siguen los datos registrados por la encuesta en relación con cada unidad concreta, que se consignan en campos de longitud fija y en el mismo orden que las preguntas del cuestionario. Todos los datos se almacenan en formato ASCII.

50. Los conjuntos de datos de la encuesta deben organizarse únicamente como archivos planos independientes (uno para cada tipo de registro) en el momento de su divulgación, ya que el formato de campo de longitud fija de la estructura original es también adecuado para transferir los datos a los sistemas de administración de bases de datos (DBMS) para su ulterior procesamiento, o a un programa estadístico estándar para su tabulación y análisis.

La transferencia de los datos a los DBMS es muy sencilla, ya que la estructura original se traduce casi directamente al formato de la base de datos estándar (DBF), que se acepta en todos los casos como punto de entrada para cuadros individuales (en este caso, los identificadores de registro funcionan como vínculos naturales de relación entre los cuadros).

51. La divulgación requiere también que la estructura de cada tipo de registro se documente debidamente en la lista de códigos de la encuesta, que debe entregarse a todo usuario interesado en trabajar con los conjuntos de datos. En dicha lista deberían especificarse claramente la posición y la longitud de cada variable del registro. En las variables categóricas, debería especificarse también la codificación. En la figura XV.1 puede verse una página del manual de códigos de la encuesta sobre el nivel de vida de Nepal (las codificaciones de algunas variables aparecen abreviadas).

Figura XV.1
Segunda encuesta sobre el nivel de vida en Nepal

Tipo de registro 002						**Sección 1, Parte A1: Miembros del hogar**	
VARIABLE	**CÓDIGO**	**RT**	**DESDE**	**LONGITUD**	**TIPO**		
Hogar		2	4	5	QNT		
CÓDIGO ID	IDC	2	9	2	QNT		
1 Nombre	Q01a	2	11	24	TYP		
D Etnia	Q01A	2	35	3	QLN	Chhatri	001
						Brahmin Hill	002
						…	…
						Otros	102
2 Género	Q02	2	38	1		Hombre	1
						Mujer	2
3 Relación	Q03	2	39	2		Jefe	1
						Cónyuge	2
						Hijo	3
						…	…
						Otros familiares	11
						Empleado/Familiar del empleado	12
						Arrendatario/Familiar del arrendatario	13
						Otra persona no emparentada	14
4A Distrito de nacimiento	Q04A	2	41	2	QLN	Taplejung	01
						Panchthar	02
						…	…
						Otros país	93
4B Distrito de nacimiento: U/R	Q04B	2	43	1	QLN	Urbano	1
						Rural	2
5 Edad	Q05	2	44	2	QNI		
6 Estado civil	Q06	2	46	1	QLN	Casado	1
						Divorciado	2
						Separado	3
						Viudo	4
						Soltero	5
7 ¿Cónyuge en la lista?	Q07	2	47	1	QLN	Sí	1
						No	2
8 Código de ID del cónyuge	Q08	2	48	2	QNT		
9 Meses en casa	Q09	2	50	2	QNT		
10 ¿Miembro o no?	Q10	2	52	1	QLN	Sí	1
						No	2

52. Las plataformas CSPro y LSD-2000 permiten elaborar una lista de códigos como producto secundario del proceso de desarrollo del programa de entrada de datos. LSD-2000 ofrece también interfaces para convertir los archivos de entrada de datos en archivos DBF y transferir los datos al programa informático de estadística más utilizado (Ariel, CSPro, SAS, SPSS y Stata). Ello demuestra la importancia de definir atentamente el sistema de codificación variable en la fase de desarrollo del programa de entrada de datos: si se hace debidamente, los analistas de la encuesta podrán utilizar los conjuntos de datos desde el momento mismo en que estén disponibles.

G. Gestión de datos en el proceso de muestreo

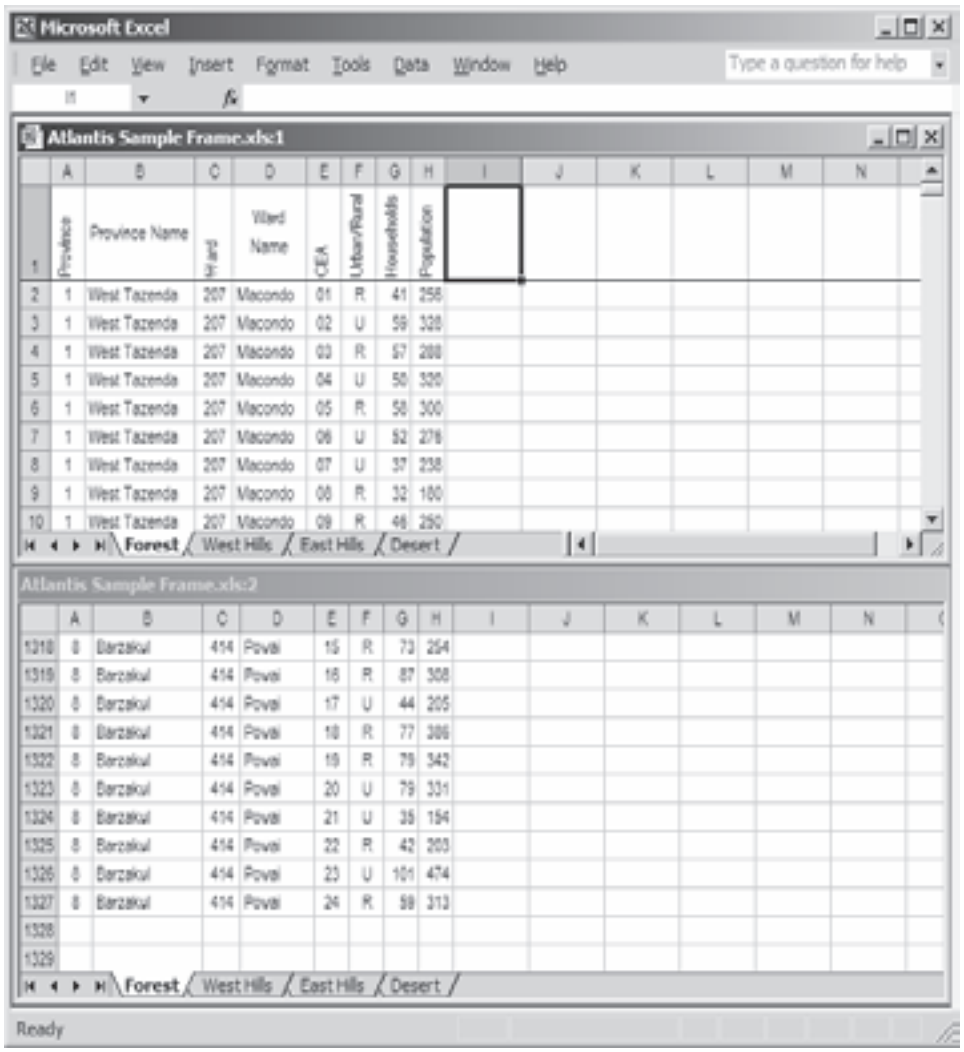
53. En este apartado se examina el papel de la gestión de datos en el diseño y realización de las muestras de encuestas de hogares. Se ofrecen recomendaciones para la informatización de los marcos muestrales y para la realización de las primeras etapas de selección de la muestra, con inclusión de métodos prácticos para la estratificación implícita y el muestreo de las UPM con probabilidad proporcional al tamaño (se examina la constitución de una base de datos con unidades penúltimas de muestreo como productos derivados de las fases anteriores de muestreo, haciéndose hincapié en su importancia como instrumento de gestión mientras se lleva a cabo la encuesta, y se explica cómo se puede actualizar su contenido con información generada sobre el terreno (como los resultados de la preparación de la lista de hogares y los datos sobre la falta de respuesta) con el fin de generar las ponderaciones que deberán utilizarse en la fase del análisis.

54. *Organización del marco muestral de la primera etapa.* Las unidades de muestreo de la primera etapa en muchas encuestas de hogares son las zonas de empadronamiento censal (ZE) definidas en el censo nacional más reciente disponible. La creación de un archivo de computadora con la lista de todas las ZE del país es un medio cómodo y eficiente de crear el marco muestral de la primera etapa. Con excepción de los países con un número ingente de ZE (como Bangladesh, donde son más de 80.000), el mejor método es un programa de hojas de cálculo, como Excel, con una fila para cada ZE, y columnas para toda la información que pueda ser necesaria. Debe incluirse la identificación geográfica completa de la ZE y un indicador de su tamaño (como el número de habitantes, de hogares o de viviendas). En general conviene crear una hoja de cálculo diferente para cada uno de los estratos de la muestra. En la figura XV.2 puede verse el aspecto de un marco muestral de la primera etapa en el estrato "bosques" de un país hipotético (la pantalla de Excel se ha dividido en dos ventanas para recoger simultáneamente la primera y la última ZE).

55. En ese ejemplo, las 1.327 ZE del estrato "bosques" se identifican mediante códigos geográficos y nombres de las divisiones administrativas del país (provincias y distritos) y mediante un número de serie dentro de cada distrito. El marco muestral contiene también el número de hogares y la población de cada ZE en el momento del censo, e indica si se trata de una ZE urbana o rural.

56. Antes de iniciar los pasos siguientes de la selección muestral, es fundamental verificar que el marco muestral es completo y correcto comprobando las cifras de población y comparándolas con los totales del censo publicados por el organismo de estadística. Es también importante verificar que el tamaño de todas las ZE es lo bastante grande como para que puedan utilizarse como unidades primarias de muestreo. Si el diseño de la muestra exige conglomerados de la penúltima etapa, por ejemplo, 25 hogares en cada caso, no será posible cumplir ese requisito. En tal caso, las ZE pequeñas deberían combinarse con ZE geográficamente adyacentes para constituir unidades primarias de muestreo. Este proceso puede resultar tedioso si la búsqueda de ZE adyacentes debe realizarse manualmente, mediante referencia constante a los mapas del censo. No obstante, como los organismos de estadística muchas veces asignan los números de serie de las ZE de acuerdo con algún criterio geográfico

Figura XV.2
Utilización de una hoja de cálculo como marco muestral de la primera etapa



[*Province* = provincia; *Province Name* = nombre de la provincia; *Ward* = distrito; *Ward Name* = nombre del distrito; *CEA* = ZE; *Urban/Rural* = urbano/rural; *Households* = hogares; *Population* = población]

(orden en forma de serpentina o espiral) para que las ZE próximas en las hojas de cálculo lo sean también en el territorio, en general es posible hacer las combinaciones automáticamente en la hoja de cálculo. En nuestro ejemplo, cada ZE tiene más de 30 hogares, por lo que no hace falta ninguna reagrupación. Conviene señalar, no obstante, que la ilustración anterior no es demasiado realista si se desea recurrir a ese procedimiento, ya que las ZE urbanas y rurales están mezcladas en la lista numérica, situación poco probable en un país real. En otras palabras, la agrupación informática de las ZE adyacentes no puede realizarse cuando las ZE urbanas y rurales no están agrupadas, sino dispersas, en la lista.

57. Otro paso previo a la primera etapa de muestreo es decidir si el marco debe clasificarse de acuerdo con determinados criterios de diseño para estratificar implícitamente la muestra dentro de cada uno de los estratos explícitos. Casi siempre se utilizan con este fin las divisiones administrativas pero en algunos casos podría considerarse más importante el criterio de la estratificación en zonas rurales y urbanas. Suponiendo que en nuestro ejemplo ocurra así con respecto a la clasificación en zonas urbanas/rurales, el marco muestral debe clasificarse por zonas urbanas/rurales, luego por provincia, después por distrito y finalmente por número de serie de ZE. Ello puede hacerse fácilmente con la orden de clasificación (*sort*) prevista en el programa de hojas de cálculo de la figura XV.3.

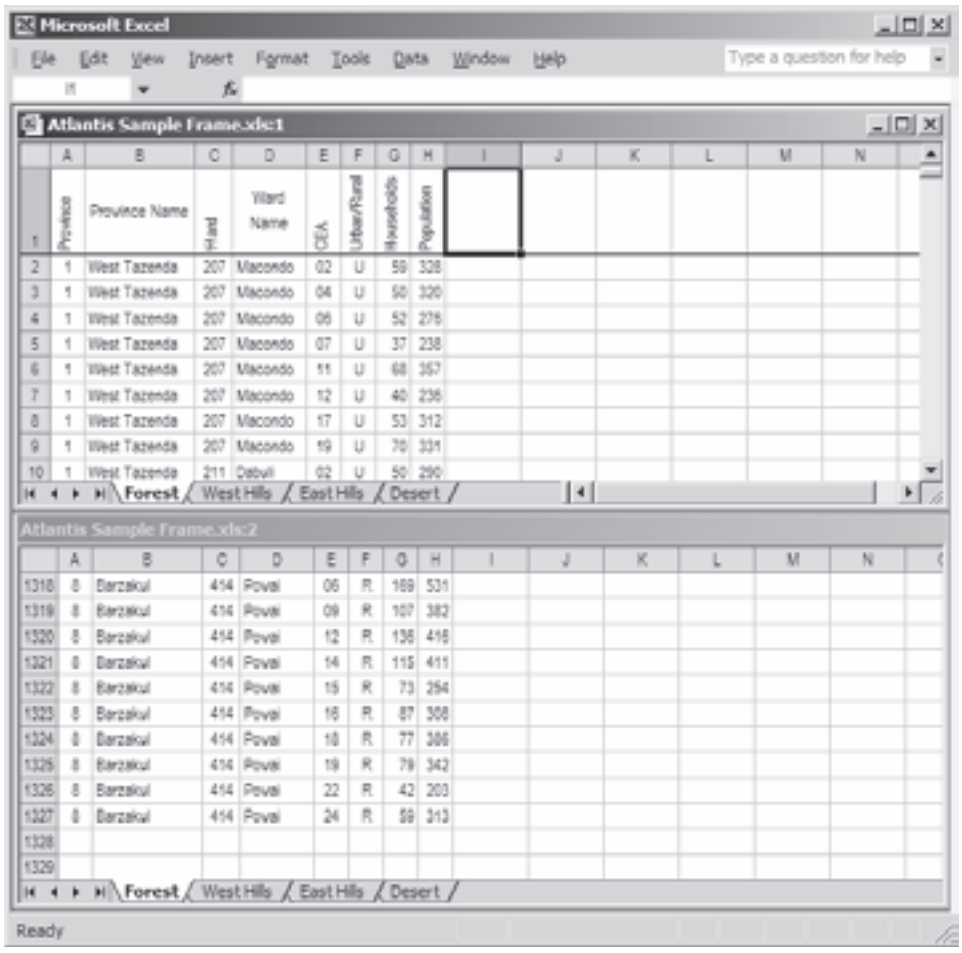
58. *Selección de unidades primarias de muestreo con probabilidad proporcional al tamaño*. La mayor parte de las encuestas de hogares seleccionan las unidades primarias de muestreo utilizando la probabilidad proporcional al tamaño (PPT). Cuando se encuentra en el marco muestral, suele utilizarse como medida de su tamaño el número de hogares de la ZE, pero en algunos casos puede emplearse la población o el número de viviendas. A continuación ilustraremos el procedimiento PPT, suponiendo que en el diseño se prevé la selección de 88 ZE con probabilidad proporcional al número de hogares (columna **G** de la hoja de cálculo) en el estrato "bosques" (*forest*) (véase la figura XV.4).

59. Lo primero que debe hacerse es crear una nueva columna en la hoja de cálculo, con el tamaño acumulado de las ZE. Introduzca la fórmula **=I1 + G2** en la casilla **I2** y cópiela hasta la última línea de la columna **I** (tenga en cuenta que esta última línea de la columna **I** contiene el número total de hogares en el estrato *forest,* es decir, 110.388).

60. En segundo lugar, cree otra columna con el tamaño acumulado ajustado de las ZE, multiplicando los valores de la columna **I** por el factor de ajuste 88/110.388 (de lo que se trata es de tener una columna que pase del cero al número de ZE que se debe seleccionar, en forma proporcional al tamaño de las ZE; véase la figura XV.5). Introduzca la fórmula **=I2*88/110388** en la casilla **J2** y cópiela hasta la última línea de la columna **J.**

61. En tercer lugar, introduzca un número aleatorio distribuido de manera uniforme entre 0 y 1 en la casilla superior de una nueva columna y súmelo a todas las líneas de

Figura XV.3
Puesta en práctica de la estratificación implícita



la columna **J**, para crear una nueva columna con el nuevo tamaño acumulado ajustado con una modificación aleatoria (véase la figura XV.6 en la página 289). Es posible seleccionar números aleatorios automáticamente dentro de la hoja de cálculo, pero es mejor seleccionar este cambio aleatorio desde el exterior (utilizando un cuadro de números aleatorios, por ejemplo) para evitar que el sistema seleccione una muestra diferente siempre que se vuelva a calcular todo el conjunto de hojas de datos. Introdúzcase, por ejemplo, el número aleatorio 0,73 en la casilla **K1**, luego introduzca la fórmula **=J2 + K$1** en la casilla **K2** y cópiela en toda la columna **K**.

62. La muestra se define por las columnas donde ha cambiado el número entero del tamaño acumulado ajustado y modificado. En este ejemplo, dicho número pasa de 0,97 a 1,02 en el caso de la ZE número 17 en el distrito número 207 (Macondo) de la provincia número 1 (West Tazenda), lo que significa que ésta es la primera ZE que debe seleccionarse en la muestra. El valor cambia de nuevo, de 1,99 a 2,09, en la ZE número 01 del distrito 226 (Balayan) de la misma provincia, por lo que ésta es la segunda ZE seleccionada. La muestra seleccionada puede resaltarse automáticamente introduciendo la fórmula **=INT(K2)-INT(K1)** en la casilla **L2** y copiándola en la columna **L**. La muestra se define por las líneas con valor distinto de cero en la columna **L** (véase la figura XV.7, en la página 290).

63. La lista de todas las unidades de muestreo seleccionadas en la primea etapa debe transferirse a una hoja de cálculo distinta, que será un instrumento fundamental para la ges-

Figura XV.4
Selección de una muestra PPT (primer paso)

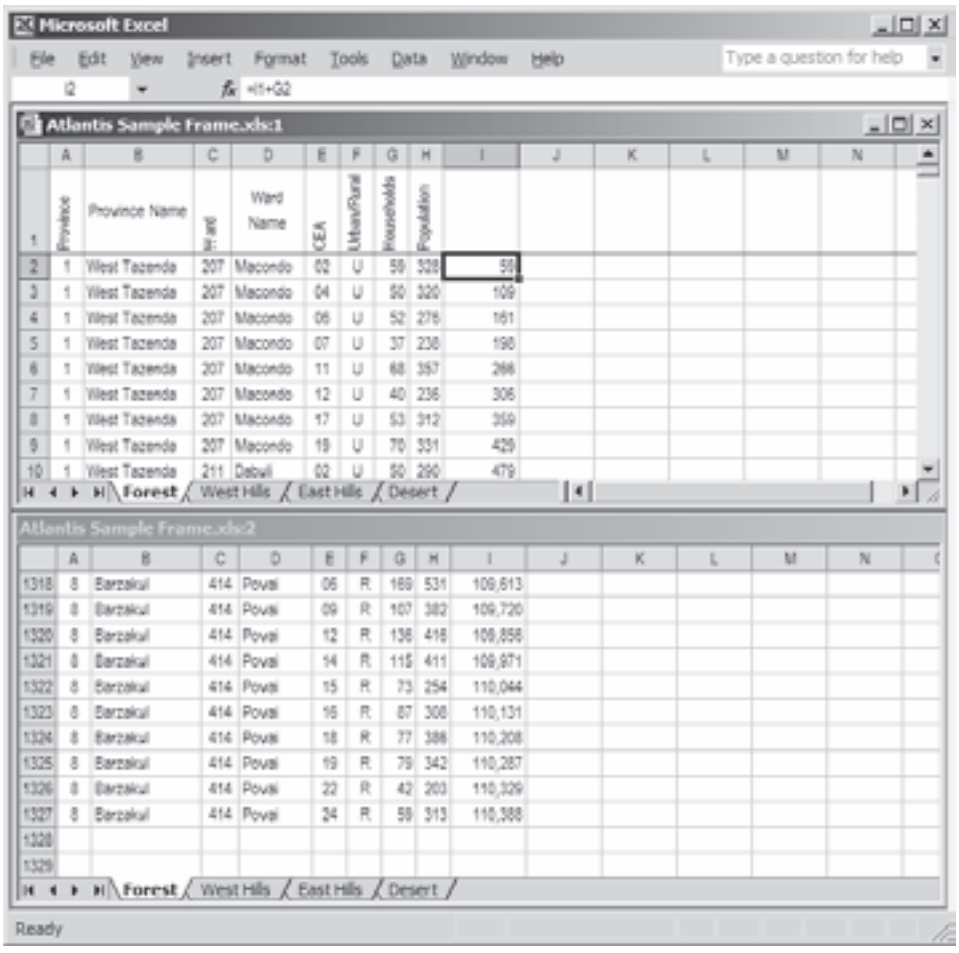
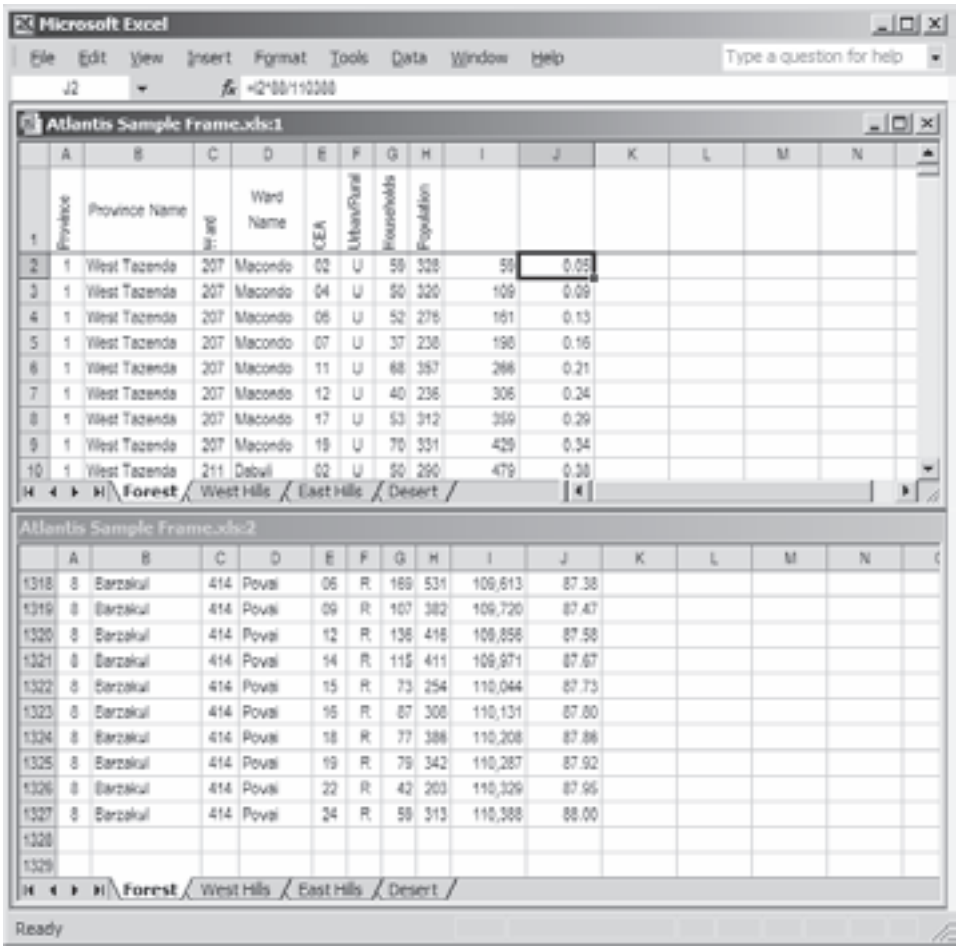


Figura XV.5
Selección de una muestra PPT (segundo paso)



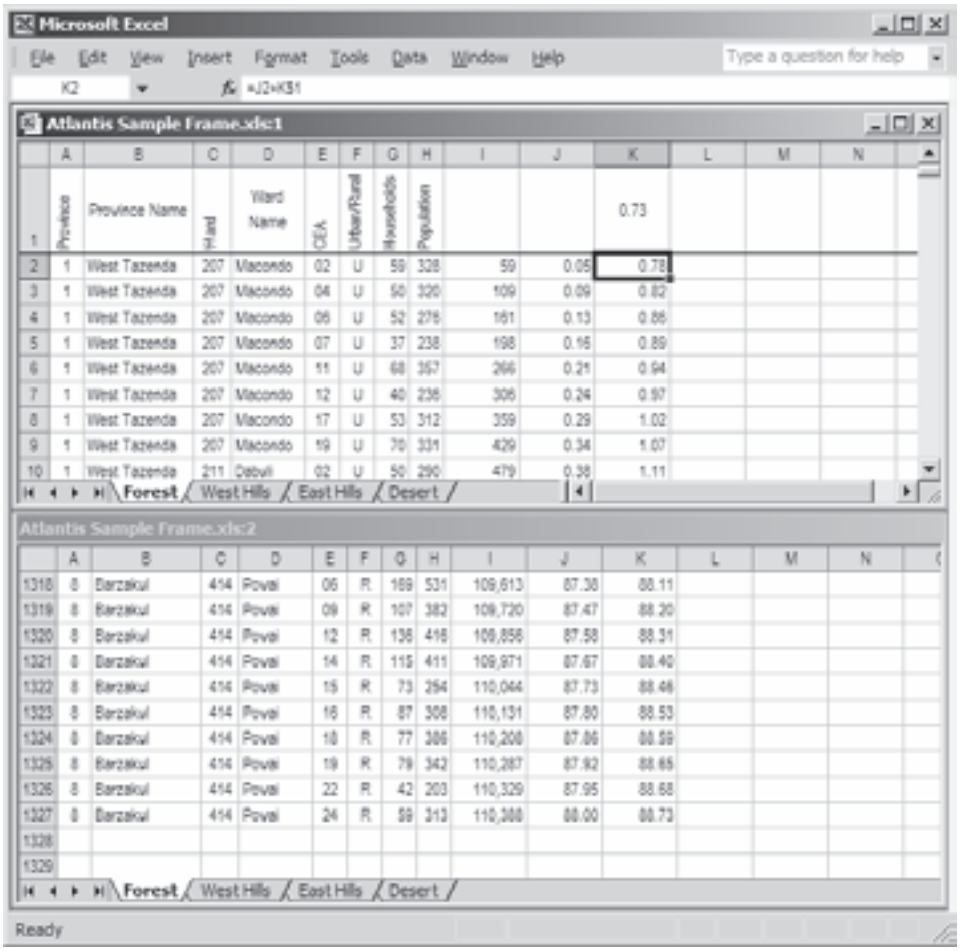
tión de la encuesta. Los responsables de la encuesta pueden, por ejemplo, añadir columnas para registrar los detalles de todas las grandes actividades de cada UPM (fechas previstas y reales del trabajo sobre el terreno y la entrada de datos, identificación del equipo responsable, etcétera).

64. La hoja de cálculo se utilizará, en particular, para calcular las probabilidades de selección y los correspondientes factores de incremento (o ponderaciones) necesarios para obtener estimaciones no sesgadas a partir de la muestra. Esta hoja de cálculo sinóptica no tiene que desglosarse por estrato. Es mejor colocar todas las UPM seleccionadas en una hoja de cálculo única, especificándose el estrato en sólo una de las columnas. En nuestro ejemplo, la "muestra" (*sample*) de la hoja de cálculo para las 19 primeras de las 88 ZE seleccionadas es la que puede verse en la figura XV.8, en la página 291.

65. *Probabilidades de selección y ponderaciones de muestreo.* Las probabilidades de selección en la primera etapa P(1) pueden calcularse fácilmente en la "muestra" de hoja de cálculo multiplicando el número de hogares de la UPM por el número de UPM seleccionadas en cada estrato (columnas **G** y **K** en la figura XV.9, en la página 291) y dividiendo el resultado por el número total de hogares del estrato (columna **J**). Esa operación se refleja en la fórmula **=K2*G2/J2** en la casilla **L2,** que se copia en toda la columna **L**.

66. Las probabilidades de selección en las etapas posteriores dependen, naturalmente, de los detalles del diseño de muestreo. Ilustraremos los cálculos para un diseño mues-

Figura XV.6
Selección de una muestra PPT (tercer paso)

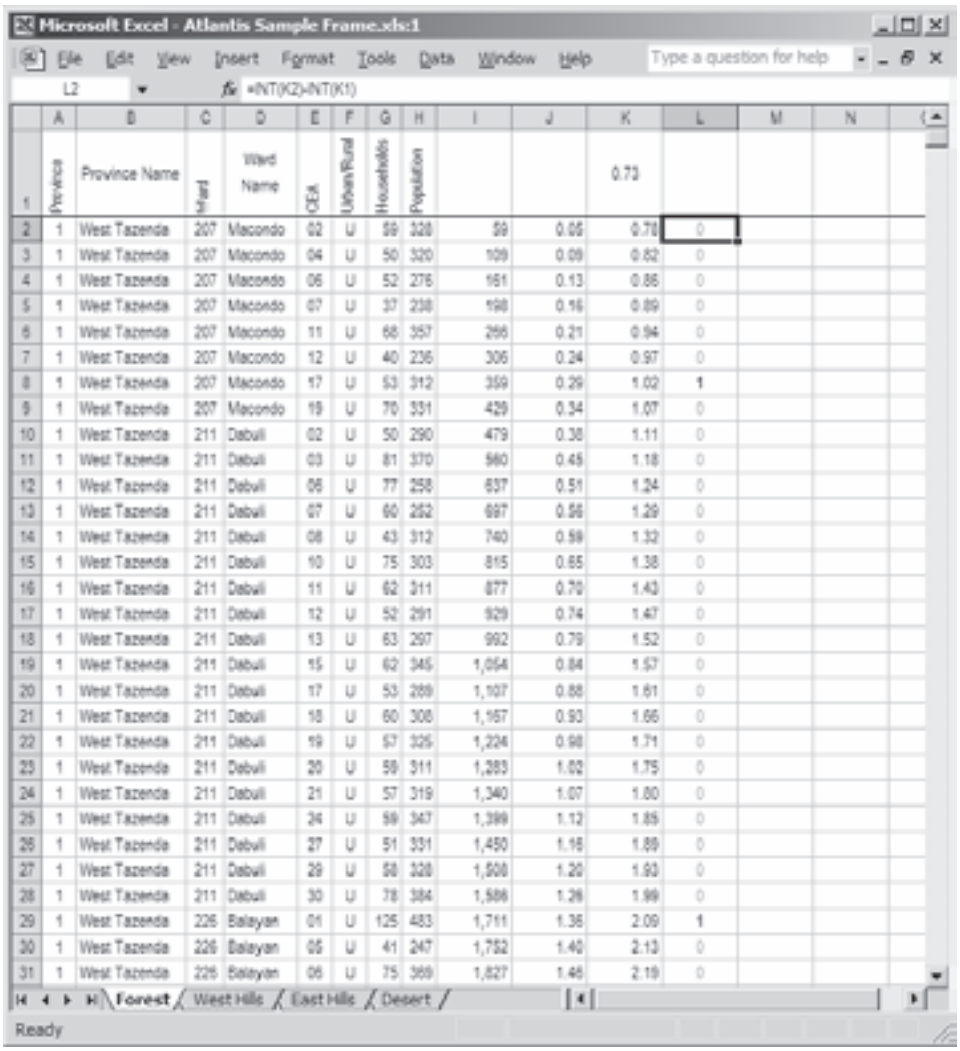


tral en dos etapas con un número fijo de hogares seleccionados con probabilidad igual en cada UPM en la segunda fase. Este diseño es de hecho uno de los más utilizados en la práctica. El número de hogares por UPM seleccionados en la segunda etapa puede variar según los estratos, pero en el país hipotético de nuestro ejemplo supondremos que hay 12 hogares por ZE en todos los estratos.

67. Esta etapa de muestreo requiere generalmente que se lleve a cabo una enumeración de hogares en cada una de las UPM seleccionadas. Esas listas no tienen que estar informatizadas, ya que la selección de los hogares que se deben visitar en la encuesta puede realizarse a mano a partir de las listas impresas. No obstante, la introducción de las listas en archivos de computadora presenta muchas ventajas (por ejemplo, si las UPM seleccionadas en la primera etapa constituyen lo que se conoce como muestra maestra, válida para varias encuestas, o para varias rondas de una encuesta por panel).

68. El número de hogares que se encuentran de hecho en cada una de las UPM de la muestra en el momento de la elaboración de la lista será en general diferente del "número de hogares" registrado originalmente por el censo en el marco de la primera etapa. **Deberá adjuntarse una columna a la "muestra" para registrar el número de hogares incluidos en la lista.** Si los formularios de listado están informatizados, esta columna puede rellenarse con un programa (utilizando macros de Excel, por ejemplo). De lo contrario, la cumplimentación de esta columna en la enumeración de los hogares debería revestir la máxima prioridad

Figura XV.7
Selección de una muestra PPT (cuarto paso)



Microsoft Excel - Atlantis Sample Frame.xls:1

File Edit View Insert Format Tools Data Window Help — Type a question for help

L2 =INT(K2)-INT(K1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population			0.73			
2	1	West Tazenda	207	Macondo	02	U	59	328	59	0.05	0.78	0		
3	1	West Tazenda	207	Macondo	04	U	50	320	109	0.09	0.82	0		
4	1	West Tazenda	207	Macondo	06	U	52	276	161	0.13	0.86	0		
5	1	West Tazenda	207	Macondo	07	U	37	238	198	0.16	0.89	0		
6	1	West Tazenda	207	Macondo	11	U	68	357	266	0.21	0.94	0		
7	1	West Tazenda	207	Macondo	12	U	40	236	306	0.24	0.97	0		
8	1	West Tazenda	207	Macondo	17	U	53	312	359	0.29	1.02	1		
9	1	West Tazenda	207	Macondo	19	U	76	331	429	0.34	1.07	0		
10	1	West Tazenda	211	Dabuli	02	U	50	290	479	0.38	1.11	0		
11	1	West Tazenda	211	Dabuli	03	U	81	370	560	0.45	1.18	0		
12	1	West Tazenda	211	Dabuli	06	U	77	258	637	0.51	1.24	0		
13	1	West Tazenda	211	Dabuli	07	U	60	282	697	0.56	1.29	0		
14	1	West Tazenda	211	Dabuli	08	U	43	312	740	0.59	1.32	0		
15	1	West Tazenda	211	Dabuli	10	U	75	303	815	0.65	1.38	0		
16	1	West Tazenda	211	Dabuli	11	U	62	311	877	0.70	1.43	0		
17	1	West Tazenda	211	Dabuli	12	U	52	291	929	0.74	1.47	0		
18	1	West Tazenda	211	Dabuli	13	U	63	297	992	0.79	1.52	0		
19	1	West Tazenda	211	Dabuli	15	U	62	345	1,054	0.84	1.57	0		
20	1	West Tazenda	211	Dabuli	17	U	53	289	1,107	0.88	1.61	0		
21	1	West Tazenda	211	Dabuli	18	U	60	308	1,167	0.93	1.66	0		
22	1	West Tazenda	211	Dabuli	19	U	57	325	1,224	0.98	1.71	0		
23	1	West Tazenda	211	Dabuli	20	U	59	311	1,283	1.02	1.75	0		
24	1	West Tazenda	211	Dabuli	21	U	57	319	1,340	1.07	1.80	0		
25	1	West Tazenda	211	Dabuli	24	U	59	347	1,399	1.12	1.85	0		
26	1	West Tazenda	211	Dabuli	27	U	51	331	1,450	1.16	1.89	0		
27	1	West Tazenda	211	Dabuli	29	U	58	328	1,508	1.20	1.93	0		
28	1	West Tazenda	211	Dabuli	30	U	78	384	1,586	1.26	1.99	0		
29	1	West Tazenda	226	Balayan	01	U	125	483	1,711	1.36	2.09	1		
30	1	West Tazenda	226	Balayan	05	U	41	247	1,752	1.40	2.13	0		
31	1	West Tazenda	226	Balayan	06	U	75	369	1,827	1.46	2.19	0		

Forest / West Hills / East Hills / Desert /

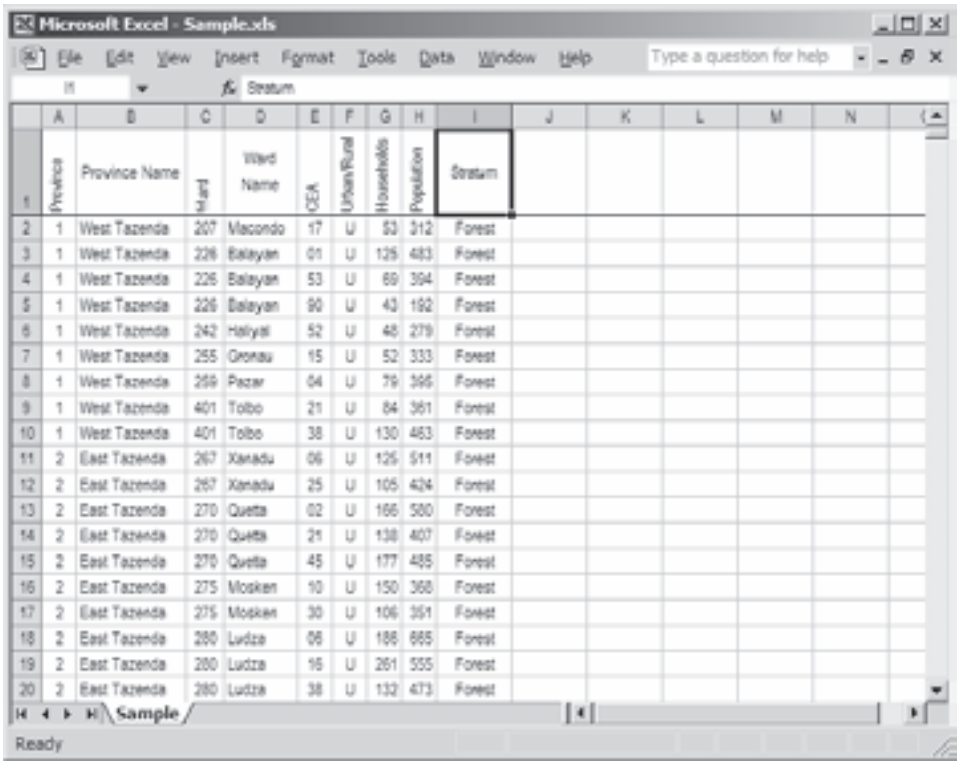
Ready

para los encargados de la encuesta. En la figura XV.10, el número de hogares (*number of households*) y el número hogares enumerados (*number of households listed*) aparecen, respectivamente, en las columnas **G** y **M.**

69. Según se vayan terminando las operaciones sobre el terreno y de gestión de datos, deberán agregarse nuevas columnas a la "muestra", para recoger, **por UPM**, los hogares sobre los que se ha registrado información útil en los conjuntos de datos de la encuesta, y aquellos otros sobre los que no se dispone información por diversas razones. Las razones habituales para incorporar la columna "cuestionario inservible" (*useless questionnaire*) se examinan con detalle en otros lugares de la presente publicación (véanse, por ejemplo, el capítulo VIII, y el apartado F del capítulo XXII, sobre la negativa a responder, las viviendas desocupadas, etcétera). Quizá haya que añadir una columna semejante cuando la encuesta no pueda integrar controles de calidad informatizados en las operaciones sobre el terreno. Éste es, por desgracia, un resultado común de las técnicas centralizadas de entrada de datos.

70. Continuando con el ejemplo presentado en la figura XV.11 (en la página 292), simplificaremos la situación suponiendo que en la "muestra" se agregan dos columnas adicionales, correspondientes al número de hogares de los conjuntos de datos (*number of households in the data sets*) y al total de falta de respuesta (*non-response*).

Figura XV.8
Hoja de cálculo con las unidades primarias de muestreo seleccionadas


Microsoft Excel - Sample.xls

File Edit View Insert Format Tools Data Window Help Type a question for help

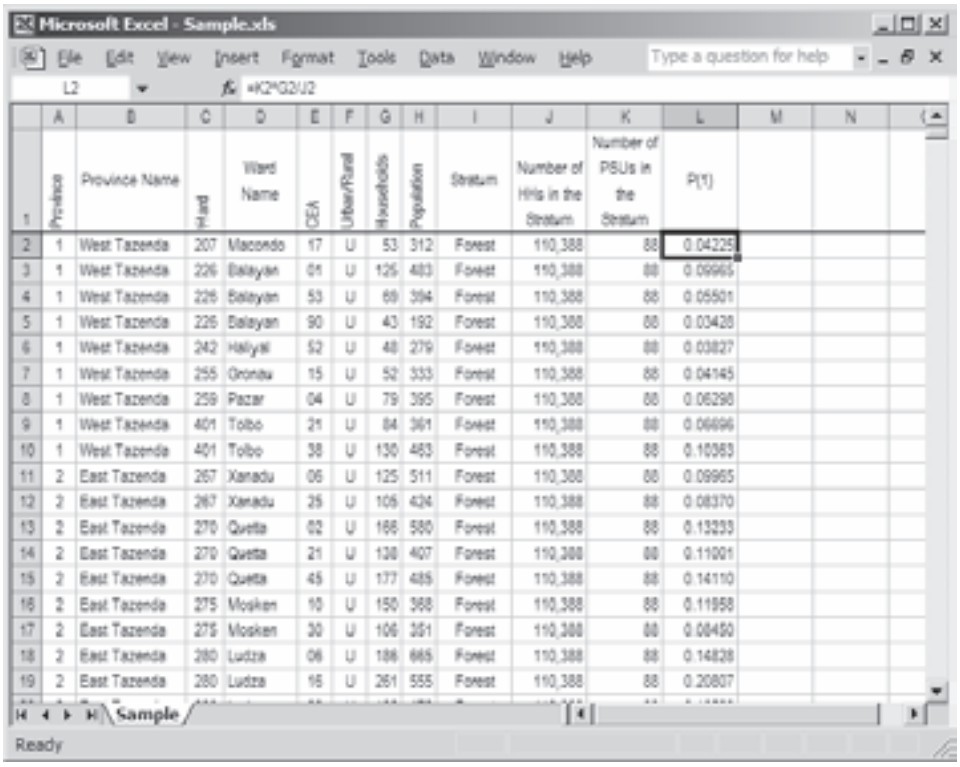
I1 Stratum

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population	Stratum					
2	1	West Tazenda	207	Macondo	17	U	53	312	Forest					
3	1	West Tazenda	226	Balayan	01	U	125	483	Forest					
4	1	West Tazenda	226	Balayan	53	U	69	394	Forest					
5	1	West Tazenda	226	Balayan	90	U	43	192	Forest					
6	1	West Tazenda	242	Haliyal	52	U	48	279	Forest					
7	1	West Tazenda	255	Gronau	15	U	52	333	Forest					
8	1	West Tazenda	259	Pazar	04	U	79	395	Forest					
9	1	West Tazenda	401	Tolbo	21	U	84	361	Forest					
10	1	West Tazenda	401	Tolbo	38	U	130	463	Forest					
11	2	East Tazenda	267	Xanadu	06	U	125	511	Forest					
12	2	East Tazenda	267	Xanadu	25	U	105	424	Forest					
13	2	East Tazenda	270	Quetta	02	U	166	580	Forest					
14	2	East Tazenda	270	Quetta	21	U	138	407	Forest					
15	2	East Tazenda	270	Quetta	45	U	177	485	Forest					
16	2	East Tazenda	275	Mosken	10	U	150	368	Forest					
17	2	East Tazenda	275	Mosken	30	U	106	351	Forest					
18	2	East Tazenda	280	Ludza	06	U	186	665	Forest					
19	2	East Tazenda	280	Ludza	16	U	261	555	Forest					
20	2	East Tazenda	280	Ludza	38	U	132	473	Forest					

Sample

Ready

Figura XV.9
Cálculo de las probabilidades de selección en la primera etapa


Microsoft Excel - Sample.xls

File Edit View Insert Format Tools Data Window Help Type a question for help

L2 =K2*G2/J2

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Province	Province Name	Ward	Ward Name	CEA	Urban/Rural	Households	Population	Stratum	Number of HHs in the Stratum	Number of PSUs in the Stratum	P(1)		
2	1	West Tazenda	207	Macondo	17	U	53	312	Forest	110,388	88	0.04225		
3	1	West Tazenda	226	Balayan	01	U	125	483	Forest	110,388	88	0.09965		
4	1	West Tazenda	226	Balayan	53	U	69	394	Forest	110,388	88	0.05501		
5	1	West Tazenda	226	Balayan	90	U	43	192	Forest	110,388	88	0.03428		
6	1	West Tazenda	242	Haliyal	52	U	48	279	Forest	110,388	88	0.03827		
7	1	West Tazenda	255	Gronau	15	U	52	333	Forest	110,388	88	0.04145		
8	1	West Tazenda	259	Pazar	04	U	79	395	Forest	110,388	88	0.06298		
9	1	West Tazenda	401	Tolbo	21	U	84	361	Forest	110,388	88	0.06696		
10	1	West Tazenda	401	Tolbo	38	U	130	463	Forest	110,388	88	0.10363		
11	2	East Tazenda	267	Xanadu	06	U	125	511	Forest	110,388	88	0.09965		
12	2	East Tazenda	267	Xanadu	25	U	105	424	Forest	110,388	88	0.08370		
13	2	East Tazenda	270	Quetta	02	U	166	580	Forest	110,388	88	0.13233		
14	2	East Tazenda	270	Quetta	21	U	138	407	Forest	110,388	88	0.11001		
15	2	East Tazenda	270	Quetta	45	U	177	485	Forest	110,388	88	0.14110		
16	2	East Tazenda	275	Mosken	10	U	150	368	Forest	110,388	88	0.11958		
17	2	East Tazenda	275	Mosken	30	U	106	351	Forest	110,388	88	0.08450		
18	2	East Tazenda	280	Ludza	06	U	186	665	Forest	110,388	88	0.14828		
19	2	East Tazenda	280	Ludza	16	U	261	555	Forest	110,388	88	0.20807		

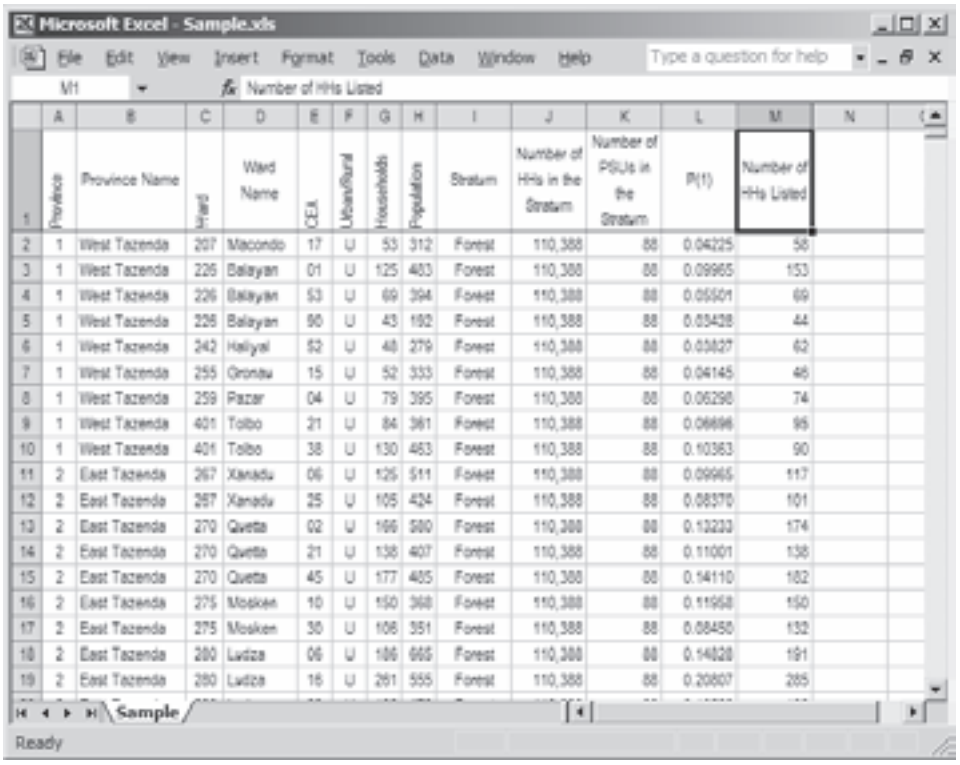
Sample

Ready

[*Province* = provincia; *Province Name* = nombre de la provincia; *Ward* = distrito; *Ward Name* = nombre del distrito; *CEA* = ZE; *Urban/Rural* = urbano/rural; *Households* = hogares; *Population* = población: *Stratum* = estrato; *Number of HHs in the Stratum* = número de hogares en el estrato; *Number of PSU in the Stratum* = número de UPM en el estrato]

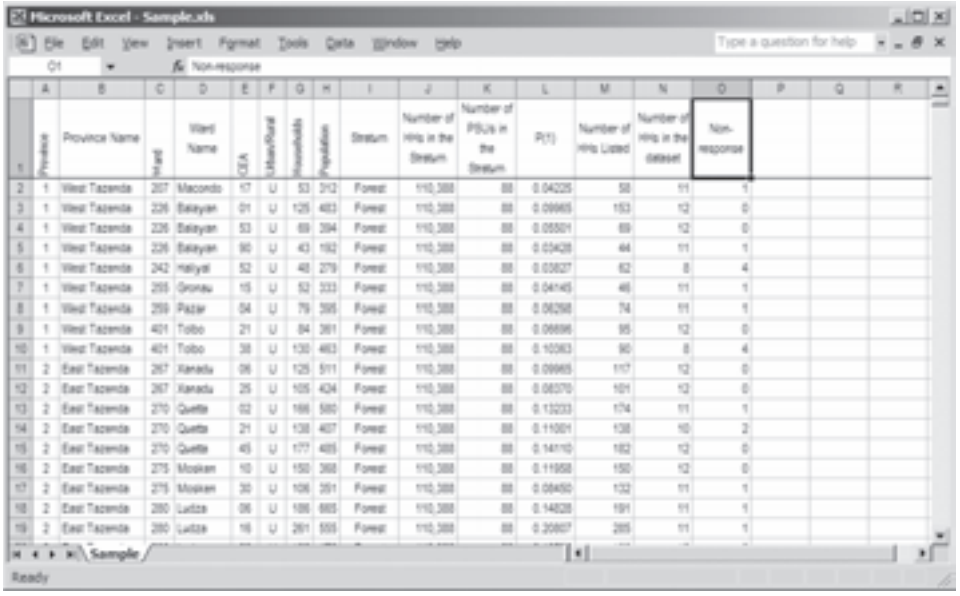
71. Aunque no hay modelos universalmente aceptados para la falta de respuesta, un supuesto muy común es que los hogares "útiles" de los conjuntos finales de datos son de hecho una muestra con probabilidad igual de todos los hogares enumerados en sus respectivas UPM (véanse los capítulos II y VIII, donde se explica el tema detenidamente). De

Figura XV.10
Documentación de los resultados de la enumeración de hogares



[*Province* = provincia; *Province Name* = nombre de la provincia; *Ward* = distrito; *Ward Name* = nombre del distrito; *CEA* = ZE; *Urban/Rural* = urbano/rural; *Households* = hogares; *Population* = población: *Stratum* = estrato; *Number of HHs in the Stratum* = número de hogares en el estrato; *Number of PSU in the Stratum* = número de UPM en el estrato; *Number of HHs Listed* = número de hogares incluidos]

Figura XV.11
Documentación de la falta de respuesta



[*Province* = provincia; *Province Name* = nombre de la provincia; *Ward* = distrito; *Ward Name* = nombre del distrito; *CEA* = ZE; *Urban/Rural* = urbano/rural; *Households* = hogares; *Population* = población: *Stratum* = estrato; *Number of HHs in the Stratum* = número de hogares en el estrato; *Number of PSU in the Stratum* = número de UPM en el estrato; *Number of HHs Listed* = número de hogares incluidos *Number of HHs in the dataset* = número de hogares en el conjunto de datos; *Non-response* = sin respuesta]

acuerdo con esta hipótesis, la probabilidad P(2) de seleccionar cada uno de estos hogares en la segunda etapa puede calcularse sencillamente dividiendo el número de hogares útiles por el número de hogares enumerados. La probabilidad de selección total de cada uno de los hogares de la UPM es el producto P(1)*P(2), y la ponderación de muestreo es el inverso de dicha probabilidad.

72. Estas fórmulas pueden aplicarse fácilmente en la hoja de cálculo (véase la figura XV.12). Escriba las fórmulas **=N2/M2** en la casilla **P2**; **=L2·P2** en la casilla **Q2,** y **=1/Q2** en la casilla **R2**; luego cópielas en las columnas **P**, **Q** y **R.**

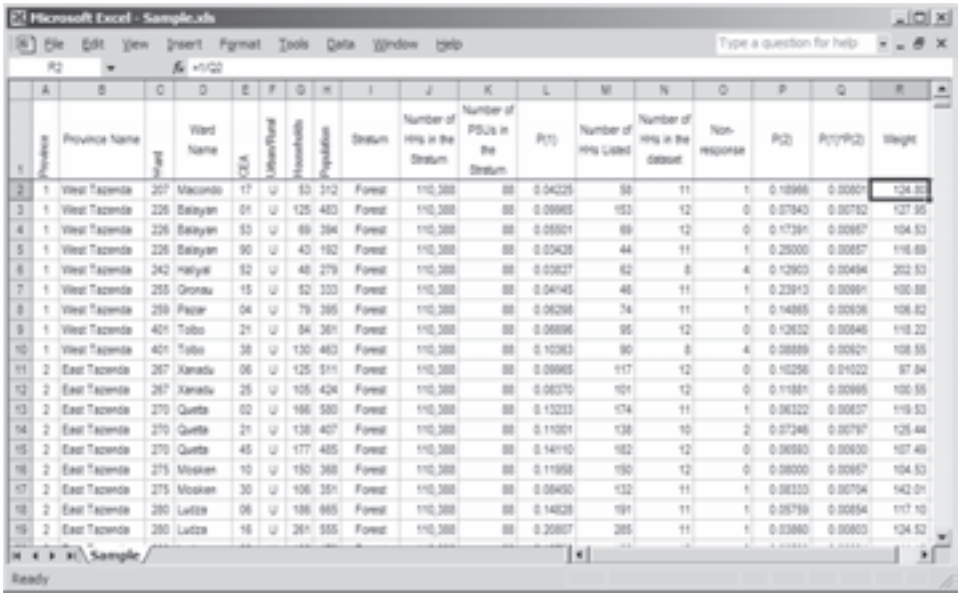
73. Las ponderaciones basadas en las probabilidades calculadas de esta manera se aplican a todos los hogares de UPM. Algunos profesionales pueden utilizar técnicas de "posestratificación" para ajustar mejor estas ponderaciones y garantizar que las estimaciones de la encuesta se correspondan con algunas distribuciones de población conocidas (como las distribuciones por edad y género, y las cifras sobre consumo total obtenidas de fuentes externas a la misma encuesta). Estos ajustes se introducen directamente, con programas informáticos especializados, en las hojas de cálculo de muestreo, y generalmente se aplican hogar por hogar o persona por persona, en vez de por UPM.

H. Resumen de las recomendaciones

74. En este capítulo se ha intentado arrojar algo de luz sobre la conveniencia de incorporar criterios sobre la gestión de datos en todas las fases de una encuesta, y no sólo en las últimas fases de análisis. Uno de los ejemplos más claros es el de las encuestas del Estudio de medición de los niveles de vida, que han tratado de diseñar sus cuestionarios, planificar y realizar las operaciones sobre el terreno y ocuparse de la entrada y procesamiento de datos de tal manera que éstos puedan gestionarse adecuadamente incluso antes de su recopilación. Los principios rectores de esa iniciativa constituyen la base del presente capítulo, y aun cuando consideran características que son diferentes según la aplicación específica de cada país, dichos principios pueden resumirse y codificarse como sigue:

a) La gestión de los datos de la encuesta comienza con el diseño del cuestionario; en ese punto los objetivos son los siguientes:

Figura XV.12
Cálculo de probabilidades de la segunda etapa y ponderaciones de muestreo



[*Province* = provincia; *Province Name* = nombre de la provincia; *Ward* = distrito; *Ward Name* = nombre del distrito; *CEA* = ZE; *Urban/Rural* = urbano/rural; *Households* = hogares; *Population* = población: *Stratum* = estrato; *Number of HHs in the Stratum* = número de hogares en el estrato; *Number of PSU in the Stratum* = número de UPM en el estrato; *Number of HHs Listed* = número de hogares incluidos; *Number of HHs in de dataset* = número de hogares en el conjunto de datos; *Non-response* = sin respuesta; *Weight* = ponderación]

i) *Identificación adecuada de las unidades estadísticas*. Lo que se recomienda es utilizar números de serie sencillos o mejorados de tres o cuatro dígitos para las UPM de la encuesta, y luego un número de serie de dos dígitos para cada uno de los hogares dentro de ella, además de un número de identificación de serie de cada unidad subordinada dentro del hogar;

ii) *Repeticiones deliberadas*. El diseño del cuestionario debería incluir repeticiones deliberadas, con el fin de detectar los errores del entrevistador o de la entrada de datos. Como ejemplo cabe citar las líneas finales donde se suma el total de las líneas anteriores o la agregación de un dígito de comprobación a los códigos de las variables importantes.

b) Durante las operaciones sobre el terreno, deberían tenerse en cuenta los siguientes aspectos:

i) *Estrategias operacionales para la entrada y la edición de los datos.* Se recomienda que los países consideren atentamente la opción de introducir todos los datos sobre el terreno. Ello puede realizarse utilizando operadores de entrada de datos que trabajen en un lugar fijo distinto de los hogares incluidos en la encuesta, haciendo que un operador se sume al resto del equipo de entrevistas e introduzca los datos directamente en una computadora portátil en cada hogar o, finalmente, recurriendo al método, todavía no bien comprobado, de las entrevistas sin papel mediante una computadora de bolsillo (aunque este procedimiento debe ser objeto de mayor investigación). La introducción de todos esos datos sobre el terreno en vez de en un lugar centralizado contribuirá enormemente a garantizar la calidad y la coherencia;

ii) *Criterios de control de calidad.* Los datos del cuestionario deben someterse a cinco mecanismos de control para comprobar los límites, la relación con los cuadros de referencia, las pautas de salto, la coherencia y los errores tipográficos.

iii) *Tecnología de entrada de datos.* Según un examen del Banco Mundial realizado en 1995, existen dos plataformas adecuadas para la entrada y edición de datos en las encuestas complejas de hogares: el programa LSMS, desarrollado internamente por el Banco Mundial, y el IMPS, de la Oficina del Censo de los Estados Unidos. Sus versiones actualizadas son, respectivamente, LSD-2000 y CSPro. Teniendo en cuenta la disponibilidad de personal especializado y otros factores que influyen en las condiciones de cada país, hay algunas orientaciones básicas que deberían tenerse en cuenta al diseñar los instrumentos de entrada y edición de datos: prescindiendo de las posibles excepciones, las pantallas de la computadora deberían parecerse a las secciones correspondientes del cuestionario; los programas de entrada de datos deberían distinguir las situaciones imposibles e improbables y marcar cada una de ellas de forma específica; la terminología y las expresiones para la comunicación del tipo de error deberían ser informales y fácilmente comprensibles;

iv) *Organización y divulgación de los conjuntos de datos de la encuesta*. Para estas actividades, los archivos planos no son los más indicados porque no permiten tratar adecuadamente las unidades estadísticas subordinadas (personas, cosechas, artículos de consumo, etcétera) dentro del hogar. Es preferible una estructura con un tipo de registro diferente para cada tipo de unidad estadística.

c) Finalmente, la gestión de datos puede resultar también importante para realizar el protocolo de muestreo, orientándolo a lo largo de sus fases principales: organi-

zación del marco muestral de la primera etapa, que se crea normalmente a partir del último conjunto disponible de zonas de empadronamiento (ZE) del censo; selección de las unidades primarias de muestreo con probabilidad proporcional al tamaño, medidas en función del número de hogares, viviendas o personas; cálculo de las probabilidades de selección y las correspondientes ponderaciones de muestreo.

REFERENCIAS BIBLIOGRÁFICAS

Ainsworth, M. y J. Muñoz (1986). *The Côte d'Ivoire Living Standards Survey: Design and Implementation*. Living Standards Measurement Study Working Paper, No. 26. Washington, D.C.: Banco Mundial.

Blaizeau, D. (1998). Seven expenditure surveys in the West African Economic and Monetary Union. En *Proceedings of the Joint Asociation of Survey Statisticians/International Association for Official Statistics (IASS/IAOS) Conference on Statistics for Economic and Social Development.* Aguascalientes, México: Instituto Internacional de Estadística.

_____ y J. L. Dubois (1990). *Connaître les conditions de vie des ménages dans les pays en développement.* París: Documentation française.

Blaizeau, D, y J. Muñoz (1998). *LSD-2000. Logiciel de saisie des données: pour saisir les données d'une enquête complexe.* París: Institut national de la statistique et des études économiques.

Grosh, M. y J. Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*, Living Standards Measurement Study Working Paper, No. 126. Washington, D.C.: Banco Mundial.

Muñoz, J. (1989). Data management of complex socioeconomic surveys: from questionnaire design to data analysis. En *Proceedings of the 47th Session of the Instituto Internacional de Estadística.* París: Instituto Internacional de Estadística.

_____ (1996). Cómo mejorar la calidad de la información: opciones para mejorar la organización del trabajo de campo, el sistema de entrada de datos, el análisis de consistencia y el manejo de la base de datos. En *Reunión de iniciación del Programa para el mejoramiento de las encuestas de condiciones de vida en América Latina y El Caribe*. Asunción: Banco Interamericano de Desarrollo.

_____ (1998). Budget-Consumption Surveys: New Challenges and Outlook. En *Proceedings of the Joint Association of Survey Statisticians/International Association for Official Statistics (IASS/IAOS) Conference on Statistics for Economic and Social Development*. Aguascalientes, México: Instituto Internacional de Estadística.

Oficina del Censo de los Estados Unidos. CSPro Census and Survey Processing System. Disponible en http://www.census.gov/ipcwww/cspro/.

Capítulo XVI
Presentación de estadísticas descriptivas sencillas a partir de los datos de las encuestas

Paul Glewwe
Departmento de Economía
Universidad de Minnesota
St. Paul, Minnesota
Estados Unidos de América

Michael Levin
Oficina del Censo de los Estados Unidos
Washington, D.C.
Estados Unidos de América

RESUMEN

En el presente capítulo se ofrecen orientaciones generales para calcular y presentar estadísticas descriptivas básicas con los datos de encuestas de hogares. El análisis es básico en el sentido de que consiste en la presentación de cuadros y gráficos relativamente sencillos que puedan ser entendidos con facilidad por un público numeroso. Se ofrece también asesoramiento sobre la manera de introducir cuadros y gráficos en el informe general, que será objeto de amplia divulgación.

Términos clave: Estadísticas descriptivas, cuadros, gráficos, resumen estadístico, divulgación.

A. Introducción

1. El verdadero valor de los datos de las encuestas de hogares sólo se hace realidad cuando se analizan los datos. Ese análisis puede ir desde la presentación de sencillas estadísticas sinópticas hasta el examen sumamente complejo de variables múltiples. Este capítulo es una presentación de los cuatro siguientes y, por lo tanto, en él se consideran las cuestiones básicas y los métodos relativamente sencillos. En los cuatro capítulos que siguen se tratarán cuestiones más complejas.

2. La mayor parte de los datos de las encuestas de hogares pueden utilizarse de formas muy distintas para arrojar algo de luz sobre los fenómenos que constituyen el centro principal de atención de la encuesta. En cierto sentido, el punto de partida para el análisis de datos es una estadística descriptiva básica, como cuadros de medias y frecuencias de las principales variables. No obstante, el punto de partida más fundamental para el análisis de datos son las preguntas a las que se intenta dar respuesta mediante los datos recopilados. Así pues, en casi todas las encuestas de hogares la primera tarea es establecer los objetivos de la encuesta y diseñar su cuestionario de manera que los datos recopilados permitan conseguir esos objetivos. Ello significa que el diseño y planificación de la encuesta para el análisis de datos deberán llevarse a cabo en forma simultánea, antes de que comience la recopilación de datos. Así se explica con mayor detalle en el capítulo III. En el presente capítulo se consideran los numerosos aspectos prácticos del análisis de datos, en el entendimiento de que se ha formulado previamente una estrategia razonable para ese análisis, siguiendo las pautas ofrecidas en el capítulo III.

3. La organización de este capítulo es como sigue: En el apartado B se examinan diferentes tipos de variables y estadísticas descriptivas sencillas; en el apartado C se da una orientación general sobre cómo preparar y presentar estadísticas descriptivas básicas a partir de los datos de encuestas de hogares; y en el apartado D se formulan recomendaciones para preparar un informe general (conocido muchas veces con el nombre de compendio estadístico) que haga llegar los resultados básicos de una encuesta de hogares a un público amplio. En el breve apartado final se presentan algunas observaciones a modo de conclusión.

B. Variables y estadísticas descriptivas

4. Muchas encuestas de hogares recopilan datos sobre un tema concreto; otras, abarcan temas muy diversos. En uno y otro caso, los datos recopilados pueden considerarse como una colección de variables, algunas de las cuales pueden interesar en forma aislada, mientras que otras son valiosas especialmente cuando se comparan con otras. Muchas de las variables presentan divergencias en cuanto al hogar —por ejemplo, en lo que respecta al tipo de vivienda—, mientras que otras difieren en el plano individual —por ejemplo, en cuanto a la edad y el estado civil—. Algunas encuestas recopilan datos que varían únicamente en el plano comunitario; un ejemplo serían los precios de varios productos vendidos en el mercado local[1].

5. El primer paso en todo análisis de datos es generar un conjunto de datos que contenga todas las variables de interés. Los analistas pueden luego calcular estadísticas descriptivas básicas que permitan a las variables "hablar por sí solas". Hay un número relativamente pequeño de métodos para ello. En este apartado se explica cómo hacerlo. Se comienza con un breve examen de los diferentes tipos de variables y estadísticas descriptivas y luego se analizan los métodos para presentar los datos sobre una, dos y tres o más variables.

1. Tipos de variables

6. Las encuestas de hogares recopilan datos sobre dos tipos de variables: "categóricas" y "numéricas". Las variables categóricas son características que no constituyen números propiamente dichos, sino categorías o tipos. Como ejemplos cabe citar las características de la vivienda (cubierta del suelo, material de las paredes, tipo de aseo, etcétera), y características individuales, como grupo étnico, estado civil y ocupación. En la práctica podrían asignarse números de código a estas características designando, por ejemplo, un grupo étnico como "código 1", otro como "código 2", y así sucesivamente, pero se trata de una convención arbitraria. Por el contrario, las variables numéricas son por su misma naturaleza números. Cabría citar, como ejemplos, el número de habitaciones de una vivienda, la superficie de tierra poseída o los ingresos de un determinado miembro del hogar. En este capítulo, los posibles resultados de las variables categóricas se denominarán "categorías", mientras que los resultados posibles de las variables numéricas se conocerán como "valores".

7. Al presentar los datos para uno u otro tipo de variable conviene establecer otra distinción, acerca del número de categorías o variables que puede presentar una variable. Si el número de categorías/variables es pequeño —menos de 10—, es conveniente (e ilustrativo) presentar la información completa sobre la distribución de la variable. En cambio, si el número de variables/categorías es elevado —más de 10—, es mejor presentar sólo estadísticas agregadas o sinópticas referentes a la distribución de la variable. Por ejemplo, en un país la población puede estar formada por un pequeño número de grupos étnicos, quizá sólo cuatro. En dicho país, es relativamente fácil presentar en un cuadro o gráfico sencillo el porcentaje de los hogares de la muestra que pertenecen a cada grupo. En cambio, en otro país puede haber centenares de grupos étnicos. Sería muy tedioso presentar el porcentaje de los hogares de la muestra que corresponden, por ejemplo, a 400 grupos diferentes. En la mayoría de los

[1] En la mayor parte de las encuestas de hogares se entiende por hogar un grupo de individuos que: *a*) viven en la misma vivienda; *b*) realizan juntos al menos una de las comidas cada día, y *c*) ponen en común los ingresos y otros recursos para la compra de bienes y servicios. Algunas encuestas de hogares modifican esta definición de acuerdo con las circunstancias locales, pero ese aspecto sobrepasa el contenido de este capítulo. La "comunidad" es más difícil de definir, pero a los efectos de este capítulo puede entenderse por tal una colección de hogares que viven en la misma aldea, ciudad o sección de una gran urbe. Véase un examen detallado de la definición de "comunidad", en Frankenberg (2000).

casos sería más sencillo y suficientemente ilustrativo agregar los numerosos grupos étnicos diferentes en un pequeño número de categorías amplias y presentar el porcentaje de hogares que se incluyen en cada una de estas categorías agregadas.

8. El ejemplo anterior contiene una variable categórica, el grupo étnico, pero puede aplicarse también a las variables numéricas. Algunas variables numéricas, como el número de días que una persona ha estado enferma la semana anterior, contiene sólo un pequeño número de valores, y por lo tanto toda la distribución puede reflejarse en un sencillo cuadro o gráfico. En cambio, muchas otras variables numéricas, como el número de animales poseídos, pueden presentar un gran número de valores, por lo que es mejor presentar sólo estadísticas sumarias de la distribución. La principal diferencia en el trato de las variables categóricas y numéricas se debe a la forma de agregación cuando el número de posibles valores/categorías es elevado. En el caso de las variables categóricas, una vez que se ha tomado la decisión de no presentar toda la distribución, no hay otra opción que la agregación en categorías amplias. En cuanto a las variables numéricas, es posible la agregación en categorías amplias, pero existe también la opción de presentar estadísticas sinópticas, como la media, la desviación estándar y, quizá, los valores mínimo y máximo. En el subapartado siguiente se presenta un breve examen de la mayor parte de las estadísticas descriptivas comunes.

2. Estadísticas descriptivas sencillas

9. Los cuadros y gráficos pueden ofrecer información básica sobre las variables de interés utilizando estadísticas descriptivas sencillas. En ellas se incluyen, entre otras informaciones, las distribuciones porcentuales, las medianas, las medias y las desviaciones estándar. En este subapartado se analizan esas estadísticas sencillas, y se presentan ejemplos basados en la utilización de datos de la encuesta de hogares de Saipán, que pertenece a la Comunidad de las Islas Marianas del Norte, y de Samoa Americana.

10. *Distribuciones porcentuales.* Las encuestas de hogares pocas veces recopilan datos en que se incluyan exactamente 100 ó 1.000 ó 10.000 personas u hogares. Supongamos que se tengan datos sobre las categorías de una variable categórica, como el número de personas de una población que son varones o que son hembras, o datos sobre una variable numérica, como la edad, expresada en años, de los miembros de la misma población. La presentación del número de observaciones que se incluyen en cada categoría normalmente no resulta tan útil como la presentación del porcentaje de las observaciones dentro de cada categoría. Así se observa cuando se consideran las tres primeras columnas de números del cuadro XVI.1. Para la mayoría de los usuarios sería difícil interpretar estos resultados si no se les diera ninguna distribución porcentual. Las tres últimas columnas del mismo cuadro se comprenden mucho más fácilmente si lo que se quiere saber es la proporción de varones y de hembras en los diferentes grupos de edad. Naturalmente, a veces lo que interesa son los porcentajes de la columna, es decir, el porcentaje de hombres y el porcentaje de mujeres que se incluyen en los diferentes grupos de edad. Así puede verse en el cuadro XVI.2 (una tercera posibilidad es presentar porcentajes que sumen un 100% en todas las edades, desglosados por categoría de sexo, pero ello suele presentar menor interés). En ambos cuadros se observa que las distribuciones porcentuales pueden corresponder a variables categóricas o numéricas.

11. Del cuadro XVI.1 se deduce claramente que la distribución por sexo difiere en los distintos grupos de edad. Ello se debe a algo que no puede observarse en los cuadros XVI.1 y XVI.2, a saber, que Saipán tiene muchos trabajadores inmigrantes —en particular mujeres trabajadoras— que encuentran empleo en sus fábricas de ropa. Si bien Saipan tiene un número ligeramente más alto de hombres que de mujeres entre las personas de pocos años, el siguiente grupo de edad, el de 15-29 años, tiene sólo 30 hombres por cada 70 mujeres. El grupo de edad de 30-44 años tiene también más mujeres que hombres. Ello está en consonancia con el hecho de que la mayor parte de los trabajadores de las fábricas

Cuadro XVI.1
Distribución de la población por edad y sexo, Saipán, Comunidad de las Islas Marianas del Norte, abril de 2002: porcentajes de las filas

Grupo de edad amplio, en años	Número de observaciones			Porcentajes de las filas		
	Total	Varones	Hembras	Total	Varones	Hembras
Total de personas	67 011	29 668	37 343	100,0	44,3	55,7
Menos de 15 años	16 915	8 703	8 212	100,0	51,5	48,5
De 15 a 29 años	18 950	5 765	13 184	100,0	30,4	69,6
De 30 a 44 años	20 803	9 654	11 149	100,0	46,4	53,6
De 45 a 59 años	8 105	4 458	3 648	100,0	55,0	45,0
60 años o más	2 239	1 088	1 150	100,0	48,6	51,4

Fuente: Ronda 10 de la encuesta permanente de mano de obra de las Islas Marianas del Norte.
Nota: Datos tomados de una muestra aleatoria del 10% de hogares y todas las personas que viven en centros colectivos.

Cuadro XVI.2
Distribución de la población por edad y sexo, Saipán, Comunidad de las Islas Marianas del Norte, abril de 2002: porcentajes de las columnas

Grupo de edad amplio, en años	Número de observaciones			Porcentajes de las columnas		
	Total	Varones	Hembras	Total	Varones	Hembras
Total de personas	67 011	29 668	37 343	100,0	100,0	100,0
Menos de 15 años	16 915	8 703	8 212	25,2	29,3	22,0
De 15 a 29 años	18 950	5 765	13 184	28,3	19,4	35,3
De 30 a 44 años	20 803	9 654	11 149	31,0	32,5	29,9
De 45 a 59 años	8 105	4 458	3 648	12,1	15,0	9,8
60 años o más	2 239	1 088	1 150	3,3	3,7	3,1

Fuente: Ronda 10 de la encuesta permanente de mano de obra de las Islas Marianas del Norte.
Nota: Datos tomados de una muestra aleatoria del 10% de hogares y todas las personas que viven en centros colectivos.

de ropa son mujeres con edades comprendidas entre 20 y 40 años. En el siguiente grupo, el de 45-59 años de edad, hay más hombres que mujeres. Los porcentajes de las columnas del cuadro XVI.2 revelan que el grupo de edad más numeroso de varones era el de 30-44 años, mientras que en el caso de las mujeres lo era el de 15-29 años, que es el grupo de mujeres con mayores probabilidades de trabajar en las fábricas de ropa.

12. *Medianas.* Las dos medidas estadísticas más frecuentes de las variables numéricas son las medias y las medianas (por definición, las variables categóricas no son numéricas y, por lo tanto, no permiten el cálculo de medias ni de medianas). La mediana es el punto medio de una distribución, mientras que la media es el promedio aritmético de los valores. La mediana se utiliza con frecuencia en variables como la edad y los ingresos, ya que acusa menos la influencia de los valores atípicos. Como ejemplo extremo, supongamos que en una encuesta hay 99 personas con ingresos comprendidos entre 8.000 y 12.000 dólares anuales, y distribuidos simétricamente en torno a 10.000 dólares. En ese caso la media y la mediana serían 10.000 dólares. Supongamos ahora que se añade durante el año una persona más con un ingreso de 500.000 dólares; la media sería de aproximadamente 15.000, mientras que la mediana sería todavía 10.000. En el caso de muchas variables sobre el ingreso, los informes publicados suelen indicar tanto la media como la mediana.

13. Volviendo a los datos de Saipán, la edad mediana de la población era de 28,5 años en abril de 2002, es decir, la mitad de la población tenía más de 28,5 años y la otra mitad menos de esa edad. La edad mediana femenina era inferior a la edad mediana masculina (27,6 años frente a 30,5), debido al mayor número de mujeres jóvenes inmigrantes que trabajan en las fábricas de ropa.

14. *Medias y desviación estándar.* Como ya se ha observado, la media es el promedio aritmético de una variable numérica. Son muy frecuentes las medias del número de niños nacidos (de mujeres), los ingresos y otras variables numéricas. La desviación estándar mide la distancia media de una variable numérica con respecto a la media de dicha variable y, por lo tanto, permite medir la dispersión de cualquier variable numérica dentro de la distribución.

Cuadro XVI.3
Estadísticas sinópticas de los ingresos de hogares por grupo étnico, Samoa Americana, 1994

Ingreso anual	Total	De Samoa	De Tonga	Otros grupos étnicos
Número de hogares encuestados	8 367	7 332	244	790
Mediana (dólares de los EE. UU.)	15 715	15 786	7 215	23 072
Media (dólares de los EE. UU.)	20 670	20 582	8 547	25 260

Fuente: Encuesta de hogares de Samoa Americana, 1995.
Nota: Datos no ponderados; muestra aleatoria del 20 % de los hogares.

15. En el cuadro XVI.3 pueden observarse las medianas y las medias de los ingresos anuales según la Encuesta de hogares de 1995 de Samoa Americana. La encuesta comprendía una muestra aleatoria del 20% de todos los hogares del territorio. Era previsible que el ingreso medio de los hogares fuera superior al ingreso mediano, ya que algunos hogares ganaban salarios significativamente más altos y obtenían ingresos superiores de otras fuentes. Los inmigrantes de Tonga son relativamente pobres, como se observa en su bajo nivel medio y mediano de ingreso, mientras que la alta media y mediana de los ingresos de "otros grupos étnicos" revela que se encuentran en situación relativamente desahogada.

3. Presentación de estadísticas descriptivas para una variable

16. El caso más sencillo al presentar estadísticas descriptivas de una encuesta de hogares es aquel en el que interviene una sola variable. En este subapartado se explica cómo puede procederse con variables tanto categóricas como numéricas.

17. *Presentación de toda la distribución.* Las variables categóricas o numéricas que representan un pequeño número de categorías o valores —10 o menos— son las más fáciles de representar. Puede utilizarse un cuadro para presentar toda la distribución (porcentual) de las variables indicando la frecuencia de cada una de las categorías o valores numéricos de la variable. Puede verse un ejemplo en el cuadro XVI.4, en el que se aprecia la distribución porcentual y de recuentos de frecuencia de la muestra (no ponderada) con respecto a las fuentes principales de iluminación en los hogares vietnamitas. Muchas encuestas de hogares exigen el uso de ponderaciones para estimar la distribución de una variable de la población, en cuyo caso la presentación de las frecuencias brutas de la muestra puede dar lugar a error, por lo que no son aconsejables. El uso de ponderaciones se examinará más adelante en el apartado C (la encuesta de Viet Nam estaba basada en una muestra autoponderada, por lo que no se necesitaban ponderaciones). Una observación final es que conviene también comunicar los errores estándar de las frecuencias porcentuales estimadas (véase en el capítulo XXI un examen detallado de este tema, que se complica por el uso de ponderaciones y por otras características del diseño de la muestra de la encuesta).

18. En algunos casos, el número de categorías o valores incluidos en una variable puede ser elevado, pero la mayor parte de la distribución se concentra en un reducido número de categorías o valores. En tales casos quizá sea necesario mostrar la frecuencia de cada categoría o valor. Una opción para impedir que el volumen de información acabe con la paciencia del lector de un cuadro es combinar los casos raros en la categoría general "Otros", donde podría introducirse toda categoría o valor con frecuencia inferior al 1%. Eso es lo que se ha hecho en el cuadro XVI.4, donde en "Otros" se incluyen casos raros como las antorchas y linternas. En ciertos casos puede haber otros grupos naturales. Por ejemplo, en muchos países los grupos étnicos y religiosos pueden dividirse en un gran número de categorías diferentes, pero esas categorías quizá puedan encajarse en un número mucho menor de grupos amplios. En ciertos casos bastará con presentar cifras de los grupos más generales. La principal excepción a esta norma se refiere a las categorías que pueden presentar especial interés aun cuando sean excepcionales. En general, estas categorías "de especial interés pero raras" podrían registrarse por separado, aunque es muy importante presentar los errores estándar en tales casos, porque la precisión de las estimaciones es más baja en las categorías raras.

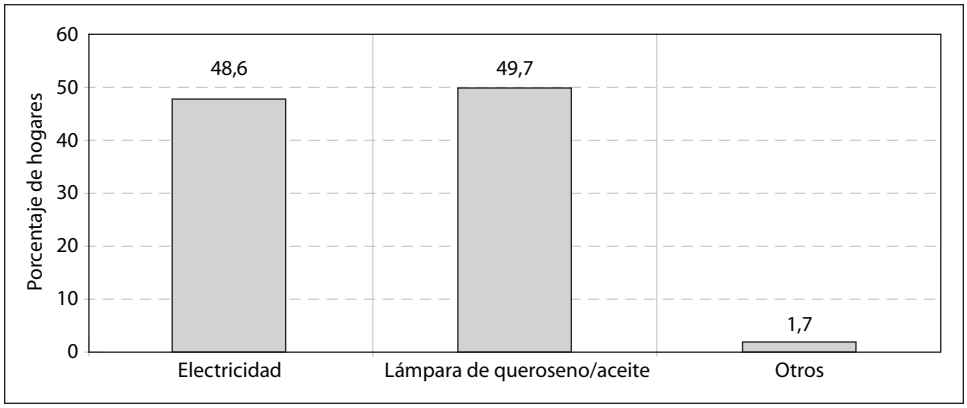
19. En muchos casos la presentación de datos puede resultar más interesante e intuitiva si se hace en forma de gráfico, en vez de en cuadro. Cuando se trata de una variable única que tiene sólo un pequeño número de categorías o valores, una forma normal de presentar gráficamente los datos es un gráfico de columnas o histograma, en que la frecuencia relativa de cada categoría o valor se indica por la altura de la columna. Puede verse un ejemplo en la figura XVI.1, en la que se utilizan los datos del cuadro XVI.4. Otra forma habitual de presentar la frecuencia relativa de las categorías o valores de una variable es el gráfico o diagrama de sectores, que es un círculo en que pueden verse las frecuencias relativas reflejadas en el tamaño de las secciones. Puede verse un ejemplo en la figura XVI.2, que presenta también la información recogida en el cuadro XVI.4. En Tufte (1983) y Wild y Seber (2000) se ofrece información sobre la manera de presentar gráficos que sea eficaz.

Cuadro XVI.4
Fuentes de iluminación en los hogares vietnamitas, 1992-1993

Método	Número de hogares	Porcentaje de hogares (error estándar)
Electricidad	2 333	48,6 (0,7)
Lámpara de queroseno/aceite	2 386	49,7 (0,7)
Otros	81	1,7 (0,2)
Total de hogares de la muestra	4 800	100,0

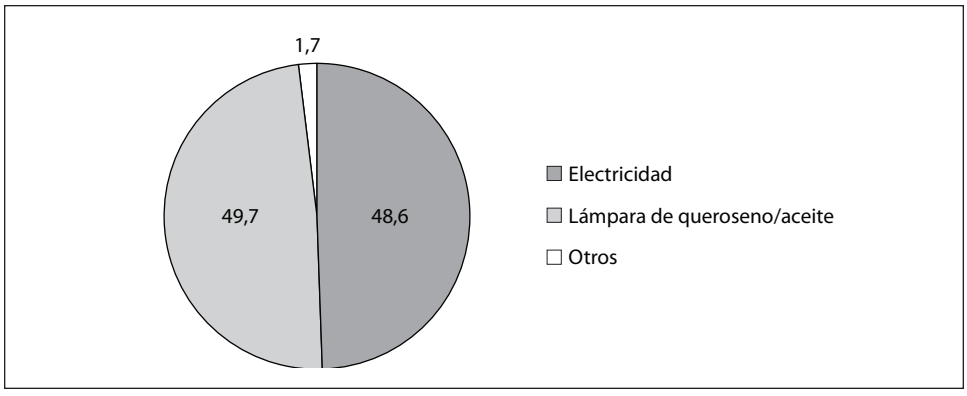
Fuente: Encuesta sobre el nivel de vida de Viet Nam, 1992-1993.
Nota: Los datos no están ponderados.

Figura XVI.1
Fuentes de iluminación en los hogares vietnamitas, 1992-1993 (*gráfico de columnas*)



Fuente: Encuesta sobre el nivel de vida de Viet Nam, 1992-1993.
Nota: Tamaño de la muestra: 4.800 hogares.

Figura XVI. 2
Fuentes de iluminación en los hogares vietnamitas, 1992-1993 (*diagrama de sectores*) (*porcentaje*)



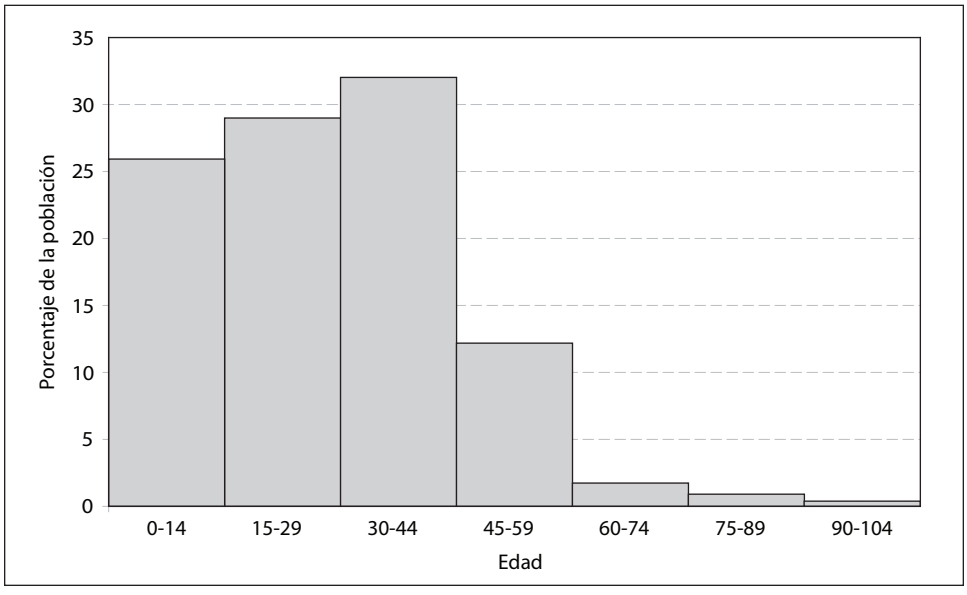
Fuente: Encuesta sobre el nivel de vida de Viet Nam, 1992-1993.
Nota: Tamaño de la muestra: 4.800 hogares.

20. *Presentación de variables con numerosas categorías o valores.* Las variables tanto categóricas como numéricas suelen tener muchas posibles categorías o valores. En el caso de las variables categóricas, la única manera de evitar la presentación de cuadros y gráficos sumamente detallados es agregar las categorías en grupos amplios y/o combinar todos los valores raros en la categoría "Otros", como ya se ha explicado. En cuanto a las variables numéricas, hay dos opciones diferentes.

21. La primera es dividir el intervalo de cualquier variable numérica con muchos valores en un pequeño número de intervalos y presentar la información en alguna de las formas descritas anteriormente para los casos en que una variable tiene sólo un pequeño número de categorías o valores. Así se ha hecho, por ejemplo, con la variable de edad en los cuadros XVI.1 y XVI.2. Esta opción puede utilizarse también en los gráficos. La información sobre la distribución de una variable numérica que tiene muchos valores puede presentarse utilizando un gráfico que indique la frecuencia de la variable en un pequeño número de categorías. Un ejemplo sería el histograma, que reproduce aproximadamente la función de densidad de la variable subyacente. Los histogramas dividen el intervalo de una variable numérica en un número relativamente pequeño de "subintervalos". Cada clase está representada por una columna que tiene una superficie proporcional al porcentaje de la muestra que se incluye en el subrango correspondiente a la columna. En la figura XVI.3 se aplica este proceso a los datos sobre la edad del cuadro XVI.2. La primera clase es el subintervalo de 0 a 14; la siguiente, de 15 a 29, y así sucesivamente[2]. Obsérvese que, a diferencia del gráfico de columnas de la figura XVII.1, no hay una distancia entre las "columnas" del histograma. Ello se debe a que el eje horizontal de un histograma representa el intervalo de la variable, y las variables normalmente no tienen "lagunas" en su intervalo.

22. La segunda opción, quizá la más frecuente, para representar una variable numérica con muchos valores es utilizar una estadística sinóptica de su distribución, como la media, la mediana y la desviación estándar. La única manera de conseguirlo es presentar esas estadísticas en un cuadro; no es posible presentar estadísticas sinópticas de una única variable numérica en un gráfico. Además de la media, la mediana y la desviación estándar, es útil presentar también los valores mínimo y máximo, los valores del cuartil superior e inferior[3], y quizá una indicación de la asimetría. Puede verse un ejemplo en el cuadro XVI.5.

Figura XVI.3
Distribución por edades de la población de Saipán, abril de 2002 (*histograma*)



Fuente: Ronda 10 de la Encuesta permanente de mano de obra de la Comunidad de las Islas Marianas del Norte.

[2] Este histograma divide la población de 60-99 años en tres grupos (60-74, 75-89 y 90-104), cada uno de los cuales abarca el número de años: 15, igual que los gupos de población de menos de 60 años. El objetivo es garantizar que la superficie de cada columna del histograma sea proporcional al porcentaje de la población de cada grupo.

[3] El cuartil inferior de una distribución es el valor por debajo del cual se encuentra el 25% de las observaciones, mientras que el 75% tendría un valor más alto, y el cuartil superior es el valor por debajo del cual se encuentra el 75% de las observaciones, y sólo sería superado por el 25% de éstas.

Cuadro XVI.5
Información sinóptica sobre el total de gastos de los hogares: Viet Nam, 1992-1993 (*miles de dong anuales*)

Media	6 531
Desviación estándar	5 375
Mediana	5 088
Cuartil inferior	3 364
Cuartil superior	7 900
Valor menor	235
Valor mayor	100 478

Fuente: Encuesta sobre el nivel de vida de Viet Nam, 1992-1993.
*Nota: T*amaño de la muestra: 4.799 hogares.

4. Presentación de estadísticas descriptivas para dos variables

23. El examen de las relaciones entre dos o más variables resulta con frecuencia mucho más revelador que el de una sola variable aislada. Al mismo tiempo, las posibilidades de presentación de los datos se multiplican. En este subapartado se describen algunos métodos frecuentes, en los que se distingue entre variables con un pequeño número de categorías o variables y las que corresponden a un gran número de valores.

24. *Dos variables con un reducido número de categorías o valores.* El caso más sencillo para representar la relación entre dos variables es aquel en que ambas variables tienen un pequeño número de categorías o valores. En una sencilla tabulación doble, las categorías o valores de una variable pueden ocupar las columnas, mientras que las categorías o valores de la otra serían las filas. Puede verse un ejemplo en el cuadro XVI.6, en el que se ilustra el uso de diferentes tipos de proveedores de servicios de salud en zonas urbanas y rurales de Viet Nam. En este ejemplo, las columnas suman hasta el 100%. Como se ha explicado antes, una alternativa sería que las columnas sumaran el 100%. En el ejemplo de Viet Nam, las figuras porcentuales que suman el 100% en cada fila indicarían cómo se distribuye el uso de cada tipo de servicio de salud en las diferentes zonas urbanas y rurales del Viet Nam. Una tercera alternativa sería que cada "casilla" de este cuadro indicara la frecuencia (en cifras porcentuales) de la probabilidad "conjunta" de visitar un centro de salud para la población de una región geográfica determinada (urbana o rural), en cuyo caso, la suma de los porcentajes en todas las filas y columnas sería el 100%. No obstante, este procedimiento se usa pocas veces, ya que suelen ser más interesantes las distribuciones condicionales. En cualquier caso, es buena práctica presentar datos suficientes para que cualquier lector pueda deducir los tres tipos de frecuencias a partir de los datos presentados en el cuadro.

Cuadro XVI.6
Uso de servicios de salud por la población (de todas las edades) que visitó un centro de salud en las cuatro últimas semanas, desglosado por zonas rurales y urbanas de Viet Nam, en 1992-1993

	Zonas urbanas		Zonas rurales	
Lugar de la consulta	**Frecuencia**	**Porcentaje (error estándar)**	**Frecuencia**	**Porcentaje (error estándar)**
Hospital o dispensario	251	45,0 (2,1)	430	25,0 (1,0)
Centro comunitario de salud	30	5,4 (1,0)	318	18,5 (0,9)
Hogar del proveedor	213	38,2 (2,1)	595	34,6 (1,1)
Hogar del paciente	50	9,0 (1,2)	376	20,1 (1,0)
Otros	14	2,5 (0,7)	29	1,7 (0,3)
Total	558	100,0	1718	100,0

Fuente: Encuesta sobre el nivel de vida de Viet Nam, 1992-1993.

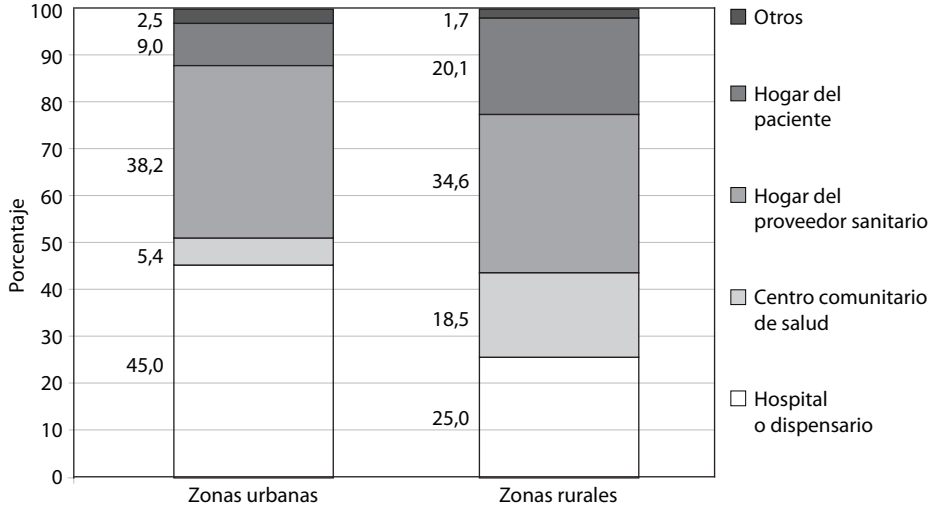
25. Hay varias maneras de utilizar los gráficos para presentar información sobre la relación entre dos variables que contienen un pequeño número de valores. Al reproducir los porcentajes de las columnas o las filas, un método adecuado es mostrar varias columnas verticales que sumen el 100%. Cada columna representa un valor particular de una de las variables, y la distribución de frecuencia de la otra variable se indica en forma de zonas sombreadas en cada columna. Así se ha hecho en la figura XVI.4 con relación a los datos sobre los servicios de salud en Viet Nam. Los programas informáticos de hojas de cálculo presentan muchas otras variaciones que podrían utilizarse.

26. *Una sola variable con un pequeño número de categorías/valores y una variable numérica con muchos valores.* Otra situación frecuente es la que contiene dos variables. Una representa un pequeño número de categorías o valores (quizá después de su agregación para reducir el número) y la otra es una variable numérica que contiene múltiples valores. En este caso la forma más frecuente de presentar los datos es la mediana del valor numérico, condicionada a cada valor de la variable que contiene un pequeño número de categorías o valores. Podría incorporarse también otra información, como la mediana y la desviación estándar. Puede verse un ejemplo en el cuadro XVI.7, en el que se observan los niveles medios de gasto total de los hogares en Viet Nam en 1992-1993, donde los hogares están clasificados atendiendo a las siete regiones del país. Para ello podría utilizarse un gráfico de columnas con un "diagrama de perfil", en el que cada columna (eje *x*) representaría una región y la longitud de las columnas (eje *y*) sería proporcional a los ingresos medios de cada región.

27. Otra opción es transformar la variable continua en una variable discreta dividiendo su intervalo en un pequeño número de categorías. Por ejemplo, a veces conviene dividir los hogares distinguiendo entre el 20% más pobre, el 20% más pobre siguiente, y así sucesivamente, teniendo en cuenta los ingresos o gastos de los hogares. Luego se pueden utilizar los mismos métodos para presentar los datos de dos variables discretas, como ya se ha descrito. Un ejemplo específico es modificar la figura XVI.4 de manera que tenga cinco columnas, una para cada quintil de ingreso.

28. *Dos variables numéricas con múltiples valores.* Los estadísticos presentan muchas veces información sinóptica sobre dos variables numéricas en función de su coeficiente de correlación (covarianza de las dos variables dividida por la raíz cuadrada del producto

Figura XVI.4
Uso de servicios de salud por la población (de todas las edades) que visitó un centro de salud en las cuatro últimas semanas, desglosado por zonas urbanas y rurales de Viet Nam, en 1992-1993 (*porcentaje*)



Fuente: Encuesta sobre el nivel de vida de Viet Nam, 1992-1993.
Nota: Tamaño de la muestra: 2.276.

Cuadro XVI.7
Gasto total de los hogares por región en Viet Nam, 1992-1993 (*miles de dong al año*)

Región	Gastos medios totales (*errores estándar entre paréntesis*)
Tierras altas del norte	4 792 (95,5)
Delta del Río Rojo	5 306 (110,4)
Norte (centro)	4 708 (107,7)
Costa (centro)	7 280 (234,8)
Tierras altas (centro)	6 173 (373,7)
Sudeste	10 786 (398,5)
Delta del Mekong	7 801 (167,4)
Todo Viet Nam	6 531 (77,6)

Fuente: Encuesta sobre el nivel de vida de Viet Nam, 1992-1993.
Nota: Tamaño de la muestra: 4.799 hogares.

de las varianzas). No obstante, estas estadísticas muchas veces son poco conocidas del gran público. Una alternativa es presentar gráficamente los datos en un diagrama de dispersión en que aparezca un punto por cada observación. De esa manera se podría indicar, por ejemplo, en qué medida está correlacionado el ingreso de los hogares durante dos períodos de tiempo, utilizando observaciones sobre los mismos hogares en dos encuestas diferentes (una para cada período de tiempo).

29. Un problema al utilizar los diagramas de dispersión es que cuando el tamaño de la muestra es grande el gráfico está demasiado "poblado" para que pueda interpretarse fácilmente. Para evitarlo, se puede trazar una submuestra aleatoria de las observaciones (por ejemplo, un décimo del total), para que el diagrama no esté tan cargado. Otro problema de los diagramas de dispersión es cómo ajustarlos teniendo en cuenta las ponderaciones de la muestra. Un método sencillo es crear observaciones repetidas, en las que la ponderación de la muestra sería el número de repeticiones de cada observación. Con ello se recargará sin duda el diagrama; por eso, después de crear los duplicados, sólo debería incluirse en el diagrama una submuestra aleatoria de las observaciones.

5. Presentación de estadísticas descriptivas para tres o más variables

30. En principio, es posible presentar relaciones entre tres o más variables utilizando cuadros y gráficos. No obstante, esto debería hacerse sólo en casos aislados, ya que se incorporan dimensiones adicionales que complican tanto la comprensión de las relaciones subyacentes como los métodos para reproducirlas en cuadros o gráficos sencillos. En la práctica, algunas veces es posible presentar las relaciones descriptivas entre tres variables, pero casi nunca es viable indicar las relaciones descriptivas entre cuatro variables o más.

31. Cuando son tres las variables, el procedimiento más sencillo es designar una de ellas como variable "condicionante". Ésta tendrá un pequeño número de valores discretos o, si es continua, deberá "discretizarse" calculando su distribución con respecto a un pequeño número de intervalos en toda su extensión. Luego pueden crearse cuadros o gráficos para cada categoría o valor de esta variable condicionante. Por ejemplo, supongamos que lo que interesa es presentar la relación entre tres variables: el nivel de instrucción del jefe del hogar, el nivel de ingresos del hogar y la incidencia de la malnutrición infantil. Ello podría realizarse generando un cuadro o gráfico independiente de la relación entre ingresos y un indicador del estado de nutrición de los niños (como la incidencia del retraso del crecimiento) para cada nivel de instrucción. De esa manera se puede observar, por ejemplo, que la asociación entre ingresos y nutrición infantil es más débil en los hogares con niveles de instrucción más elevados.

C. Orientaciones generales para presentar estadísticas descriptivas

1. Preparación de los datos

32. Antes de que se genere ninguna cifra que deba introducirse en los cuadros y gráficos, habrá que preparar los datos para el análisis. Ello implica tres tareas distintas: comprobar los datos para eliminar las observaciones que pueden ser muy imprecisas, generar variables complejas (derivadas) y documentar a fondo la preparación del conjunto de datos "oficiales" que se utilizarán para todo el análisis. En estas tres tareas la atención al detalle desde los primeros momentos puede ahorrar mucho tiempo y muchos recursos en el futuro. En este apartado se describen brevemente esas tareas; puede verse una exposición mucho más detallada en el capítulo XV.

33. Prácticamente todas las encuestas de hogares, por muy atentamente que se planifiquen y ejecuten, contienen observaciones sobre algunas variables que no parecen creíbles. Estos problemas van desde la falta de respuesta sobre partidas concretas (véase el capítulo XI) y otros errores obvios —por ejemplo, un niño de tres años que aparece como jefe de hogar— a otros mucho menos claros, como un hogar con ingreso muy alto pero con un nivel medio de gastos. En muchos casos los errores se deben a la entada de datos inexactos a partir de los cuestionarios impresos. En ese caso habría que comprobar primero ese cuestionario. Estos errores de entrada de datos pueden resolverse fácilmente. Si los datos sospechosos aparecen también en el cuestionario, hay varias opciones. En primer lugar, se podría cambiar el valor de la variable por un espacio en blanco. Si hay sólo un pequeño número de estos casos, podrían excluirse esas observaciones al calcular un cuadro o gráfico en que se utilice tal variable[4]. Si los casos son muy numerosos, los valores en blanco pueden calcularse como una categoría distinta de la variable categórica, que recibiría el nombre de "no registrados" o de "no declarados". En segundo lugar, si la mayor parte de los casos se concentran en un pequeño número de hogares, podría prescindirse de éstos. En tercer lugar, si hay muchas observaciones cuestionables en muchos hogares de algunas variables, quizá haya que tomar la decisión de no presentar los resultados de dicha variable.

[4] Esta opción tiene la desventaja de que el tamaño de la muestra diferirá ligeramente en cada cuadro. Si bien ello podría dar lugar a confusión, una nota al pie de cada cuadro en que se explique que se han eliminado algunas observaciones debería ser una aclaración suficiente.

34. Una manera de proceder cuando faltan datos es "imputar" los valores ausentes utilizando uno de los varios métodos disponibles. Los métodos de imputación asignan valores a casos desconocidos o "no registrados", así como a los casos con valores poco plausibles. Entre esos planteamientos se incluyen el método de imputación conocido como *hot deck* o el del vecino más próximo, en los que se intenta "adivinar" la respuesta cuando no hay ninguna disponible. La idea en que se basan estos métodos es sencilla: los hogares o las personas que tienen algunas características similares tendrán probablemente otras características semejantes. Por ejemplo, es probable que los hogares de una determinada aldea rural tengan paredes y tejados semejantes a los de las casas de otras zonas rurales, a diferencia de las casas de las zonas urbanas. De la misma manera, la mayor parte de las personas de un hogar suelen tener la misma religión y pertenecer a la misma etnia. El equipo de la encuesta debe decidir las normas específicas que deben seguirse teniendo en cuenta las condiciones demográficas, sociales, económicas y de vivienda del país.

35. Si bien los métodos de imputación son útiles, pueden presentar también serios problemas. Los miembros del equipo encargados del análisis de datos deben decidir si cambiar los datos omitidos caso por caso o utilizar alguna especie de método de imputación. Deben tenerse en cuenta los efectos en las tabulaciones finales. La imputación de un 1% o un 2% de los casos debería tener poco o ningún efecto en los resultados finales. Si aproximadamente el 5% de los casos son omisiones o representan contradicciones con otras partidas, probablemente debería considerarse todavía la posibilidad de imputación. En cambio, si hubiera que imputar una proporción de valores mucho mayor, digamos el 10% o más, quizá la variable no podría utilizarse en la presentación ni en el análisis, por lo que no debería

facilitarse ningún resultado sobre esa variable. El lector debe consultar los capítulos VIII y XI y las referencias bibliográficas de los mismos, donde se presenta información adicional sobre la imputación y la manera de proceder cuando se han omitido algunos valores.

36. Otro aspecto de la preparación de los datos es el cálculo de variables complejas (derivadas). En muchas encuestas de hogares, el total de los ingresos o gastos del hogar, o ambos, se calculan teniendo en cuenta los valores de un gran número de variables. Por ejemplo, el gasto total suele calcularse sumando los gastos en 100 o más alimentos y otros productos concretos. Si bien en teoría el cálculo de estas variables es sencillo, en la práctica pueden presentarse muchos problemas. Por ejemplo, al calcular los ingresos y gastos agrícolas de los hogares rurales, algunas veces ocurre que los beneficios de la explotación son negativos. Cuando se producen resultados extraños en determinados hogares, puede ser útil considerar cada uno de los componentes introducidos en el cálculo global. Puede ocurrir que uno o dos sean la causa del problema. Continuando con el ejemplo de los beneficios agrícolas, es posible que el precio de algunos insumos comprados sea excepcionalmente alto. En tal caso, podría calcularse de nuevo el beneficio utilizando un precio medio.

37. Por desgracia, la preparación de conjuntos de datos cuando se presentan problemas tiene más de arte que de ciencia. Habrá que tomar decisiones cuando no sea claro cuál es la mejor opción. Finalmente, es importante documentar las opciones adoptadas y, en términos más generales, todo el proceso a través del cual los "datos brutos" se transforman en cuadros y gráficos. En la documentación debería incluirse una breve descripción del proceso así como de todos los programas informáticos empleados para procesar y transformar los datos.

2. Presentación de los resultados

38. La mejor manera de presentar resultados estadísticos básicos depende del tipo de encuesta y del público destinatario. No obstante, hay algunas orientaciones que deberían aplicarse en casi todos los casos.

39. El consejo general más importante es el de presentar los resultados con claridad. Para ello deben tenerse en cuenta otras recomendaciones más concretas. En primer lugar, todas las variables deben definirse con precisión y claridad. Por ejemplo, al presentar cuadros y gráficos sobre el "ingreso" de los hogares, la variable del ingreso debería ser "ingreso per cápita" o "ingreso total del hogar", pero nunca sólo "ingreso". Las variables complejas, como el ingreso y el gasto, deberían definirse claramente en el texto y en las notas de los cuadros y gráficos. ¿El ingreso se entiende antes o después de pagar los impuestos? ¿Se incluye el valor de la vivienda ocupada por el propietario? ¿El ingreso hace referencia al ingreso por semana, por mes o por año? Esta información debe constar con toda claridad. En el caso de muchas variables, conviene presentar en el texto la fórmula utilizada en el cuestionario de hogares a partir de la cual se ha obtenido la variable. Por ejemplo, cuando se trata de datos sobre alfabetización de adultos, debe quedar muy claro cómo se ha decidido dicha variable. Puede definirse en función del número de años que la persona ha ido a la escuela o de si esas personas son capaces de firmar su propio nombre, o de la declaración del informante de que es o no capaz de leer un periódico, o puede estar basada en alguna prueba aplicada al informante. Las diferentes definiciones puedan arrojar resultados distintos.

40. Una segunda recomendación concreta es que las distribuciones porcentuales de las variables discretas deben especificar de forma inequívoca si se trata de porcentajes de hogares o porcentajes de personas (es decir, de la población). En muchos casos los resultados serán diferentes. En numerosos países, las personas con mayor nivel de instrucción tienen familias relativamente menos numerosas. Ello significa que la proporción de la población que vive en hogares con jefes de hogar de alto nivel de instrucción es menor que la proporción de hogares con un jefe del hogar poco o nada instruido. Una tercera recomendación es que

los gráficos indiquen los números en que se basan las formas de dichos gráficos. Por ejemplo, en el gráfico por columnas de la figura XVI.1 se observan los porcentajes de cada una de las fuentes de iluminación en los hogares vietnamitas, y lo mismo ocurre con el gráfico por sectores de la figura XVI.2.

41. Finalmente, conviene tener en cuenta algunas otras orientaciones concretas. En primer lugar, los informes no deben presentar un número enorme de cuadros ni una gran diversidad de números en cada cuadro. Los organismos de estadística a veces presentan centenares de cuadros con información minuciosa que probablemente no reviste demasiado interés para la mayoría de los destinatarios, y lo mismo cabría decir en cuanto al grado de detalle de cada uno de los cuadros. Los encargados de preparar los informes deberían examinar el objetivo de los distintos cuadros, y si se considera que un cuadro concreto o un detalle determinado de un cuadro presentan poco interés, debería prescindirse de ellos. En segundo lugar, deberían presentarse estimaciones de los errores de muestreo en lo que respecta a una selección de las variables más importantes recopiladas en la encuesta; además, es muy útil indicar los intervalos de confianza de las principales variables o indicadores. Esta observación es obvia, pero muchas veces no se tiene en cuenta. Lo que se trata de destacar es la importancia de comunicar al lector el grado de exactitud de la información ofrecida por la encuesta de hogares. En tercer lugar, debe indicarse el tamaño de la muestra en cada uno de los cuadros.

3. ¿Qué es lo que hace que un cuadro tenga o no interés?

42. En este subapartado se presentan algunas indicaciones concretas sobre la preparación de cuadros con información procedente de una encuesta de hogares. Al preparar los cuadros y gráficos debe tenerse cuenta el siguiente principio general: la información incluida en los cuadros debe ser suficiente para que el usuario pueda interpretarlos correctamente sin tener que consultar el texto del informe. Ello es muy importante, ya que muchos usuarios de los informes fotocopian los cuadros y luego los utilizan sin tener en cuenta el texto correspondiente.

43. Las orientaciones que se formulan a continuación son de carácter general. En toda encuesta el equipo debe decidir qué convenciones son las más indicadas. Una vez elegidas éstas, deben aplicarse estrictamente. No obstante, en algunos casos quizás sea necesario alejarse de esas convenciones para ilustrar puntos específicos o poner de manifiesto tipos concretos de análisis estadísticos. Un último punto referente a este subapartado es que casi todas estas orientaciones relativas a los cuadros son también válidas para los gráficos.

44. En el cuadro XVI.6 pueden verse las distintas partes de un cuadro bien concebido. Cada cuadro debe contener un título claro, indicadores geográficos (cuando convenga), títulos de las columnas, títulos de línea, fuente de datos y las posibles notas de interés.

45. *Título*. El título debería ser una descripción sucinta del cuadro. Ésta debería incluir los siguientes aspectos: *a*) número del cuadro; *b*) población u otro universo considerado (incluida la unidad de análisis, como los hogares o individuos); *c*) una indicación de lo que aparece en las líneas; *d*) una indicación de lo que aparece en las columnas; *e*) el país o región incluido en la encuesta, y *f*) el año o años de la encuesta.

46. En lo que respecta al número del cuadro, la mayor parte de los informes estadísticos numeran sus cuadros en forma consecutiva, comenzando, por ejemplo, con el cuadro XVI.1 y continuando hasta el último cuadro. Algunas veces los países utilizan letras y números para diferentes conjuntos de cuadros, por ejemplo, H01, H02, etcétera, para los cuadros de vivienda, y P01, P02, etcétera, para los cuadros de población. Si bien este procedimiento es sencillo y claro, presenta el inconveniente de que la numeración es demasiado rígida, lo que hace muy difícil las adiciones o supresiones.

47. El universo es la población o vivienda básicas incluidas en el cuadro. Si en el cuadro se incluye toda la población, el universo puede omitirse del título: se supone que es

el total de la población. Por el contrario, si un cuadro comprende una subpoblación, como la población activa, y se entiende por población activa potencial el número de personas con edades de 10 o más años, el título podría contener las siguientes palabras: "población de 10 años o más".

48. En el título del cuadro XVI.6 se incluye también una indicación de lo que aparece en las líneas y en las columnas del cuadro. En particular, se indica que el cuadro contiene información sobre los tipos de servicios de salud utilizados (las filas) y que esa información se presenta por separado en las zonas rurales y urbanas (las columnas). La inclusión del país o región en el título hace inmediatamente visible el universo geográfico. Esta característica es especialmente importante para los investigadores que comparan los resultados de diferentes países. Obviamente, la oficina nacional de estadística que recopila los datos sabrá de qué país se trata, pero las personas que utilicen cuadros de diferentes países quizá necesiten esta información para distinguir entre unos y otros.

49. Finalmente, en el título debe figurar el año o años de la encuesta, para poder determinar de inmediato el marco cronológico. Algunas veces el organismo nacional de estadística del país desea presentar datos de dos o más encuestas diferentes en el mismo cuadro. En ese caso pueden aparecer dos fechas, por ejemplo, "1900 y 2000" o "desde 1980 hasta 2000". El equipo encargado de la encuesta debe decidir si desea especificar una serie de fechas (por ejemplo, "1980, 1990 y 2000", en vez de la indicación más sencilla pero menos completa "desde 1980 hasta 2000"); no obstante, una vez tomada la decisión, el país deberá mantenerse siempre fiel a ella.

50. *Designadores geográficos.* Siempre que se repita el mismo cuadro para niveles geográficos inferiores, cada cuadro debe tener un designador geográfico que permita aclarar qué cuadro se aplica a cada región geográfica. Por ejemplo, si se repitiera el cuadro XVI.6 para cada una de las siete regiones de Viet Nam, el nombre de la región podría aparecer entre paréntesis en una segunda línea inmediatamente por debajo del título del cuadro. Podrían utilizarse también designadores "no geográficos". Por ejemplo, podría repetirse un cuadro para los principales grupos étnicos o nacionalidades.

51. *Títulos de las columnas.* Cada columna de un cuadro debe llevar un "título". Estos títulos pueden tener más de un "nivel", por ejemplo, en el cuadro XVI.6 el título de las dos primeras columnas es "Zonas urbanas" y el de las dos últimas es "Zonas rurales"; dentro de las zonas rurales y urbanas, hay títulos distintos para la frecuencia de observaciones y para la distribución porcentual de dichas observaciones. Cabe hacer otra observación sobre las columnas de "totales" o "sumas", como la primera columna del cuadro XVI.3. El equipo de la encuesta debería elegir una convención con respecto al lugar donde se introducirán estas columnas. Tradicionalmente, el total se coloca al final, y todos los atributos aparecen antes, en las diferentes columnas. No obstante, si un cuadro es muy largo y ocupa muchas páginas, con muchas columnas de información, el equipo de la encuesta quizá prefiera colocar primero el total (a la izquierda) de la serie de columnas. Cuando el total aparece en primer lugar, todo usuario sabrá inmediatamente el total de la serie de columnas, sin tener que ir pasando las páginas hasta llegar al final del cuadro.

52. Los títulos de las columnas y las correspondientes columnas de datos deben distribuirse de tal manera que se reduzcan en lo posible los espacios en blanco de la página. La distribución de las columnas debe tener en cuenta el número de dígitos de las cifras mayores que aparecerán en las columnas, el número de letras de los nombres de los atributos que aparezcan en las columnas y el número total de "espacios" permitidos por el tipo de letra concreto utilizado. El tipo de letra es muy importante y debería elegirse en los primeros momentos del proceso de tabulación.

53. *Títulos de fila*. El equipo de la encuesta deberá determinar también las convenciones que se van a utilizar para los encabezamientos y entradas de las filas. El texto debería

estar justificado y hacer referencia a una sola variable. En él se indican los nombres de las variables reflejadas en la fila correspondiente. Puede incluir también subcategorías (variables anidadas). Por ejemplo, un título general puede tener dos renglones distintos, uno para los hombres y otro para las mujeres. Es preciso introducir algunas convenciones para distinguir entre los diferentes grupos, lo que en la mayoría de los casos supondrá una indentación diferente para cada "nivel" de las variables.

54. *Precisión de los números*. Muchos cuadros tienen el problema de que presentan demasiados dígitos significativos. Cuando se señalan los porcentajes, casi siempre basta con incluir un solo dígito más allá de la coma decimal; la presentación de dos o más dígitos pocas veces aporta información útil y tiene tres inconvenientes: distrae al lector, representa un desperdicio de espacio y da una falsa sensación de precisión. Los números con cuatro o más dígitos casi nunca necesitan un decimal. Cuando se presentan grandes cifras, deben aparecer en "miles" o "millones" para que no aparezca ningún numero de más de cuatro o cinco dígitos.

55. *Fuente*. La fuente de los datos debe aparecer como el nombre completo de la encuesta, normalmente al final del cuadro (como puede verse en el cuadro XVI.6). No obstante, algunas veces las tabulaciones presentan varias encuestas de un país, o encuestas de más de un país. Cuando ello ocurre, la información de esas fuentes resulta más importante. Debería incluirse la fecha junto con el nombre de la encuesta. Si la fuente es un informe publicado, conviene distinguir entre la fecha de publicación del informe y el año de recopilación de los datos. Por ejemplo, es posible que un país haya recogido los datos en 1990 pero los haya publicado en 1992. En la fuente podría indicarse "Encuesta de fecundidad de 1990 (1992)", en que 1992 indicaría la fecha de publicación.

56. *Notas.* Las notas contienen información inmediata para poder interpretar debidamente los resultados que aparecen en el cuadro. Por ejemplo, las notas de los cuadros XVI.1 y XVI.2 indican que la población de la muestra incluye todas las personas residentes en viviendas individuales o colectivas. Además de las notas situadas al final de un cuadro, pueden aparecer en el texto, junto a los cuadros, definiciones y explicaciones pertinentes. En el texto se incluirían las definiciones de las características; por ejemplo, que el lugar de nacimiento hace referencia a la residencia de la madre inmediatamente antes de ir al hospital para dar a luz, en vez de al lugar donde se encuentra el hospital. En el texto podrían incluirse también explicaciones acerca de cómo se obtuvieron o se van a utilizar los datos. Por ejemplo, si la fecha de nacimiento y la edad se recogieron debidamente, pero la fecha de nacimiento prevalece sobre la edad cuando se observa una contradicción, esta información podría ayudar a algunos estudios, como los demográficos, a evaluar el mejor método de interpretar los datos.

4. Uso de ponderaciones

57. En este subapartado se presenta un breve panorama sobre el uso de las ponderaciones cuando se preparan cuadros y gráficos con datos de las encuestas de hogares. Puede encontrarse información mucho más detallada en los capítulos II, VI, XIX, XX y XXI y en sus correspondientes referencias bibliográficas.

58. Con respecto a la ponderación de la encuesta, el tipo más sencillo de diseño de la muestra para una encuesta de hogares es el "autoponderado". En tal caso no es preciso utilizar ponderaciones en el análisis, ya que cada hogar de la población tiene la misma probabilidad de selección en la muestra. Un ejemplo ilustrativo es la Encuesta sobre el nivel de vida de Viet Nam de 1992-1993, utilizada en varios de los ejemplos de este capítulo. No obstante, la divergencia en las tasas de respuesta de los diferentes tipos de hogares normalmente obligará a calcular las ponderaciones para corregir dicha discrepancia. La mayoría de las encuestas de hogares sobre todo, no están autoponderadas porque se basan desproporcionadamente en muestras muy amplias de algunas partes de la población que revisten particular interés.

En estas encuestas deben utilizarse ponderaciones para reflejar las diferentes probabilidades de selección y calcular debidamente las estimaciones no sesgadas de las características que presentan mayor interés para la encuesta.

59. Las ponderaciones precisas deben contener tres componentes. El primero comprende las "ponderaciones básicas" o "ponderaciones del diseño". Éstas tienen en cuenta la divergencia en las probabilidades de selección entre los diferentes grupos de hogares (es decir, cuando la muestra no es autoponderada), tal como se estipula en el diseño inicial de la muestra. El segundo componente es el ajuste teniendo en cuenta la divergencia en las tasas de falta de respuesta. Por ejemplo, en muchos países en desarrollo los hogares en mejor situación económica tienen menos probabilidad de aceptar someterse a una entrevista que los hogares de ingreso medio y bajo. Las ponderaciones básicas deben "sobrevalorarse" mediante el inverso de la tasa de respuesta para todos los grupos de hogares. Finalmente, en algunos casos puede haber "ajustes de postestratificación", ya que una fuente de datos independientes, como un censo, puede ofrecer estimaciones más precisas de la distribución de la población por edad, sexo y grupo étnico. Si las estimaciones de la encuesta sobre estas distribuciones no se corresponden fielmente con las de la fuente independiente, es posible que los datos de la encuesta vuelvan a ponderarse para conseguir que las dos distribuciones coincidan. Puede verse una descripción más detallada de los dos últimos componentes en Lundström y Särndal (1999).

D. Preparación de un informe general (compendio o informe estadístico) para una encuesta de hogares

60. La mayor parte de las encuestas de hogares divulgan primero sus resultados publicando un informe general en el que se registra, sin excesivo detalle, el conjunto de la información recogida en la encuesta. Estos informes suelen tener una distribución mucho más amplia que otros de carácter más especializado, en los que se aprovechan plenamente algunos aspectos de esa información. Estos informes generales algunas veces se conocen con el nombre de "compendios estadísticos". En este apartado se presentan algunas recomendaciones específicas para la preparación de esos informes, basadas en Grosh y Muñoz (1996).

1. Contenido

61. El material principal de todo informe general de estadística es un gran número de cuadros y gráficos. Éstos deben reflejar los principales tipos de información recopilados en la encuesta; el análisis en profundidad de temas más concretos debería dejarse para informes especiales más pormenorizados. Los cuadros deben ir acompañados de una explicación en forma de texto que permita aclarar el tipo de información contenida en ellos. No es preciso formular conclusiones normativas concretas, aunque pueden sugerirse posibles interpretaciones como posibilidades útiles de investigación en el futuro.

62. La información más básica puede desglosarse por regiones geográficas, por sexo y, quizá, por edad. Si la encuesta contiene datos sobre ingresos o gastos, éstos pueden desglosarse también por grupos de ingreso o de gasto. En algunos países habrá grandes diferencias entre los distintos tipos, y el carácter de estas divergencias puede analizarse todavía más en cuadros adicionales. En otros países, algunas de estas diferencias no son muy considerables, por lo que no habrá necesidad de ofrecer información más detallada.

63. Además de los resultados de los datos de la encuesta de hogares, el informe general debería contener varias páginas en las que se describan la encuesta en cuanto tal, en particular el tamaño y el diseño de la muestra, la fecha de comienzo y terminación de la encuesta y algunos detalles sobre la forma de recopilación de los datos. El cuestionario o los cuestionarios utilizados deberían incluirse como anexo del informe principal.

2. Proceso

64. Un informe estadístico de calidad requiere la participación de un equipo de personas, varias de las cuales deberían tener experiencia en informes anteriores. Algunos miembros del equipo se centrarán en los aspectos técnicos de la generación de cuadros y gráficos, mientras que otros se responsabilizarán sobre todo del contenido y del texto que acompaña a los cuadros. Los miembros del equipo de orientación más técnica pueden elegir el programa informático de estadística con el que estén más familiarizados, ya que la mayor parte de esos programas pueden producir las cifras necesarias para los cuadros y gráficos. No obstante, es probable que la estimación de los errores estándar exija programas especialmente concebidos con ese fin, ya que los diseños de las muestras de encuestas de hogares son casi siempre demasiado complejos para que puedan manejarse adecuadamente con los programas estándar de estadística (véase en el capituló XXI un examen más detallado sobre estas cuestiones).

65. Los miembros del equipo encargados del contenido deberían reunirse con expertos de los organismos gubernamentales para tratar los temas que deberán incluirse en el informe. De esa manera se garantizará que los cuadros y gráficos presenten los datos en la forma más útil para dichos organismos. Quizá sea útil también consultar a organismos internacionales de ayuda, que podrían considerar útiles esos datos para planificar sus programas (véase en el capítulo III un examen más general sobre la manera de constituir un equipo eficaz).

E. Observaciones finales

66. Este capítulo puede servir como introducción para la presentación de estadísticas descriptivas sencillas basadas en los datos de las encuestas de hogares. La presentación ha sido muy general y de nivel muy básico. Nuestra exposición es más bien un resumen de afirmaciones de sentido común, por lo que los analistas de datos deberían utilizar su propio sentido común al enfrentarse a cuestiones particulares referentes al análisis de sus encuestas. Pueden utilizarse también métodos más complejos para analizar los datos de las encuestas de hogares, algunos de los cuales se examinan en capítulos posteriores. En definitiva, el análisis de los datos para cualquier encuesta de hogares deberá acomodarse a los temas y objetivos principales de la encuesta, y los investigadores deberán consultar publicaciones y revistas especializadas para asesorarse sobre las cuestiones específicas relacionadas con esos temas.

REFERENCIAS

Frankenberg, Elizabeth (2000). Community and price data. En *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study.* M. Grosh y P. Glewwe, comps. Nueva York: Oxford University Press, para el Banco Mundial.

Grosh, Margaret y Juan Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No. 126. Washington, D.C.: Banco Mundial.

Lundström, S. y C. E. Särndal (1999). Calibration as a standard method for the treatment of non-response in sample surveys. *Journal of Official Statistics,* vol. 13, No. 2, págs. 305-327.

Tufte, Edward (1983). *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press.

Wild, C. J. y G. A. F. Seber (2000). *Chance Encounters: A First Course in Data Analysis and Inference.* Nueva York: Wiley.

Capítulo XVII

Uso de encuestas de hogares sobre varios temas para mejorar las políticas de reducción de la pobreza en los países en desarrollo

Paul Glewwe
Departamento de Economía Aplicada
Universidad de Minnesota
St. Paul, Minnesota
Estados Unidos de América

RESUMEN

En este capítulo se explica la forma en que las entrevistas de hogares pueden ser utilizadas por los investigadores y funcionarios públicos de los países en desarrollo para formular políticas encaminadas a reducir la pobreza. Se comienza con análisis descriptivos relativamente sencillos, en que se destaca la contribución clave de los datos contenidos en las encuestas de hogares: facilitan información sobre quiénes son pobres y sobre sus características. Luego se realizan análisis más complejos con variables múltiples, basados en técnicas de regresión múltiple. En cada tipo de análisis se presentan ejemplos sobre la posibilidad de utilizar los datos de encuestas de hogares para formular políticas de reducción de la pobreza.

Términos clave: Pobreza, formulación de políticas, análisis descriptivos, análisis con múltiples variables.

A. Introducción

1. Casi todos los países en desarrollo aceptan el hecho de que un objetivo primario del desarrollo económico y social es la reducción y posible eliminación de la pobreza. Si bien todos los gobiernos pueden tener el mismo objetivo, las políticas que adoptan para reducir la pobreza no deben ser necesariamente las mismas. La naturaleza de la pobreza y las características de los pobres varían de un país a otro, lo que significa que las políticas adecuadas deben ser también diferentes.

2. Un capítulo de estas características no da para examinar con detalle las numerosas formas en que las políticas gubernamentales pueden repercutir en la pobreza de los países en desarrollo. (Véase un examen detallado reciente en Lipton y Ravallion (1995) y Banco Mundial (2001).) No obstante sí puede presentarse un panorama general. A nuestros efectos, conviene dividir las políticas gubernamentales en cuatro tipos generales. El primero comprende las políticas macroeconómicas, que afectan a toda la economía y tienen repercusiones en el crecimiento económico y la estabilidad. Las políticas macroeconómicas más importantes son el nivel general de impuestos y de gasto público, las políticas monetarias (que influyen en las tasas de interés y de inflación), las políticas económicas internacionales (que

repercuten en el tipo de cambio, el comercio exterior y los flujos de capital extranjero), y las políticas referentes a los bancos y otras instituciones financieras. El segundo tipo de políticas gubernamentales lo integran las que afectan a los precios, como los impuestos y las subvenciones a algunos bienes y servicios. La prestación gubernamental de servicios públicos y de infraestructura, como los centros de salud, las escuelas y las redes de transporte y comunicación, representa el tercer tipo general de política gubernamental. El último está integrado por programas gubernamentales cuyo objetivo es dar asistencia directa a los pobres. Como ejemplo cabría citar el Programa de Educación, Salud y Alimentación (PROGRESA) de México, que ofrece ayudas económicas a las familias pobres si sus hijos asisten regularmente a la escuela, y el programa de cupones de alimentos de Jamaica, que ofrece a las familias pobres cupones utilizables para comprar alimentos en las tiendas locales. Los cuatro tipos de políticas pueden tener importantes repercusiones en la pobreza.

3. Esas repercusiones dependen de las características y comportamiento de los pobres, y en algunos casos de las características y comportamiento de quienes no lo son. Por ejemplo, la repercusión de los subsidios gubernamentales para alimentos específicos, que debería contribuir a reducir el precio de éstos, depende de que los pobres opten o no por comprar esos productos. Eso significa que los gobiernos necesitan información sobre las características y comportamiento de los pobres en sus países a fin de elegir las políticas más eficaces para reducir la pobreza. Esta información crucial puede obtenerse a través de las encuestas de hogares.

4. Casi todos los países en desarrollo, hasta los más pobres, realizan algún tipo de encuestas de hogares, como las encuestas de ingresos y gastos, las encuestas de mano de obra y las encuestas demográficas y de salud. Gracias a ellas se obtiene información abundante que puede utilizarse para comprender mejor el carácter de la pobreza y los probables efectos de las políticas gubernamentales en los pobres. En este capítulo se explica cómo pueden utilizarse las encuestas de hogares de los países en desarrollo para formular políticas de lucha contra la pobreza. El apartado B comienza con una descripción de lo que puede aprenderse a partir de estadísticas descriptivas sencillas calculadas con los datos de las encuestas. En el apartado C se examinan métodos más complejos basados en el análisis con múltiples variables. El capítulo se termina con una breve conclusión.

B. Análisis descriptivo

5. Para garantizar que las políticas y programas gubernamentales destinados a ayudar a los pobres sean eficaces, se necesita información sobre si las políticas llegan de verdad a los pobres y sobre los efectos que están produciendo. Por desgracia, esta información muchas veces no se encuentra en los países en desarrollo. Por ejemplo, las políticas que pueden favorecer el crecimiento económico pueden elevar los ingresos de algunas ocupaciones más que los de otras. Por otro lado, se plantea la pregunta de qué ocupaciones son las más comunes entre los pobres. Algo semejante cabe decir en cuanto a las políticas de fijación de precios. La repercusión de los planes gubernamentales de elevar los impuestos de los productos del petróleo, por ejemplo, depende de si los hogares pobres consumen una cantidad significativa de esos productos. La misma cuestión se plantea al considerar la conveniencia de crear nuevas escuelas o dispensarios de salud en determinadas zonas del país: habría que saber si esas zonas tienen una concentración relativamente elevada de hogares pobres. Finalmente, en todo programa que ofrezca prestaciones directas a los pobres, sean servicios o transferencias en especie o monetarias, los administradores del programa desearían saber qué proporción de los beneficiarios son pobres y qué proporción de ellos se benefician del programa.

6. Por desgracia, muchos países en desarrollo disponen de poca información acerca de la localización y características de los pobres, y por lo tanto tienen muy poca idea acerca de los beneficios que los pobres obtienen de las políticas y programas gubernamentales, o de

los daños que éstos les ocasionan. Las encuestas de hogares pueden cubrir muchas de esas lagunas informativas. En este apartado se examina cómo conseguirlo, utilizando numerosos ejemplos de países en desarrollo. Aunque muchos de los usos de los datos de encuestas de hogares para entender la pobreza son muy sencillos, ya que representan sólo la elaboración de cuadros y gráficos elementales, este tipo de información suele ser mucho más útil que la que se obtiene con análisis más complejos.

1. Definición de la pobreza

7. Antes de investigar los efectos de las políticas gubernamentales en los pobres hay que tener claro quiénes son los pobres, lo que supone a su vez una definición de la pobreza. No siempre hay acuerdo a este respecto. No obstante, en general y como principio se reconoce que existe un nivel de vida "decente" mínimo que todos los individuos y hogares deben ser capaces de alcanzar para poder tener la oportunidad de tener una vida satisfactoria. La mayor parte de los debates sobre la pobreza se centran en las necesidades materiales, a diferencia de las libertades políticas, los derechos humanos y el bienestar psicológico. En este capítulo se hará otro tanto. Las necesidades materiales más obvias y, por lo tanto, sobre las que hay mayor grado de consenso, son: *a*) una alimentación adecuada; *b*) alojamiento/vivienda básicos, y *c*) agua potable y medios sanitarios de eliminación de los desechos. La mayor parte de los observadores agregarían también oportunidades de educación básica y servicios elementales de atención de salud preventiva. Algunos incluso desearían contar con un conjunto más amplio de bienes y servicios, agregando, por ejemplo, las actividades culturales o recreativas, pero en este punto hay menos consenso sobre lo que se debe incluir o incluso sobre si se deben incluir o no estos bienes y servicios.

8. Los filósofos, economistas y sociólogos pueden dedicar, y con frecuencia dedican, mucho tiempo a debatir cuál es el conjunto mínimo de bienes y servicios que un individuo u hogar debería tener para que no sea considerado como pobre. Una vez que se haya llegado a un acuerdo sobre ese conjunto de bienes y servicios, la falta de consumo de componentes concretos del mismo puede utilizarse como "indicador" de pobreza. Un planteamiento más práctico, adoptado por muchos economistas, es considerar que casi todos los componentes de ese conjunto cuestan dinero, por lo que el tema se reduce realmente no a la composición exacta del conjunto, sino a su costo monetario. Este planteamiento establece un "umbral de pobreza" expresado como una cifra monetaria y se considera como pobre todo hogar cuyo ingreso o gasto sea inferior a esa cantidad. De hecho, el punto de partida para muchos umbrales de pobreza de carácter monetario utilizados en países en desarrollo es un conjunto de bienes y servicios que cumple unos requisitos mínimos. Por ejemplo, un componente sería un conjunto de alimentos que permita atender las necesidades nutricionales mínimas y que refleje también las pautas nacionales de consumo de alimentos. El paso siguiente sería calcular el costo de ese conjunto. En el resto de este capítulo se supondrá que es ése el planteamiento adoptado. Puede verse un examen más detallado de cómo establecer el umbral de pobreza en Ravallion (1998).

2. Elaboración de un perfil de pobreza

9. Una vez que se ha encontrado una definición práctica de la pobreza en función de los ingresos o gastos de los hogares, puede formularse una descripción de los pobres utilizando datos de las encuestas de hogares. Esta descripción se conoce muchas veces como "perfil de pobreza". Se realiza utilizando los datos sobre ingresos y/o gastos de los hogares para calcular el poder adquisitivo total de cada hogar (total de ingresos o total de gastos). Se consideran pobres aquellos hogares cuya capacidad adquisitiva es inferior al umbral de pobreza.

10. El párrafo anterior contiene implícitamente una enseñanza y también una pregunta. La enseñanza es que el análisis de la pobreza requiere datos de encuestas de hogares que incluyan información razonablemente precisa sobre el ingreso total o el gasto total de los hogares. Sin esos datos el análisis de la pobreza resulta difícil, ya que habrá que encontrar otra fórmula para clasificar los hogares como pobres o no pobres. Si bien podría obtenerse cierta información útil a partir de una encuesta sin contar con esos datos, es mucho más lo que se puede aprender de las encuestas de hogares que recopilan datos sobre ingresos y/o gastos. La pregunta es la siguiente: si se tiene una encuesta con datos sobre los ingresos y sobre los gastos, ¿cuál de estas dos informaciones habría que utilizar? En general, se prefieren los datos de gastos, ya que suelen ser más precisos que los relativos a los ingresos y porque los gastos de consumo están, en teoría, más estrechamente vinculados con el bienestar de los hogares, ya que el ingreso algunas veces se gasta para pagar deudas o ahorrar para el consumo futuro y por lo tanto no refleja necesariamente el bienestar del momento.

11. La primera tarea al construir un perfil de pobreza es describir quiénes son los pobres. Sin los datos de las encuestas de hogares, las autoridades y otros observadores muchas veces tienen sólo una remota idea de quiénes son los pobres y cuáles son sus características. Y, lo que todavía es peor, algunas de las ideas que tienen son inexactas. Por ejemplo, muchos funcionarios públicos y otros observadores pasan la mayor parte del tiempo en grandes zonas urbanas, y cuando piensan en los pobres tienen en cuenta lo que ven en esas zonas, cuando en casi todos los países la incidencia de pobreza es mucho más alta en las zonas rurales. Por ello, la primera tarea al utilizar los datos de encuestas de hogares es estimar la incidencia de la pobreza, describir la ubicación de los pobres distinguiendo entre zonas urbanas y rurales y entre las diferentes regiones del país, y calcular algunas características básicas de los pobres. Es importante comprobar las tasas de pobreza desglosadas por grupos étnicos y religiosos, por nivel de instrucción y por ocupación. Es también útil examinar las condiciones de vivienda entre los pobres, así como la propiedad de activos productivos. Con estas y otras informaciones, es posible comenzar a ofrecer asesoramiento útil a los responsables de la formulación de políticas.

12. Un ejemplo de algunas características básicas de los pobres es el presentado en un informe reciente del Banco Mundial (1999) sobre la pobreza en Viet Nam, donde, según las estimaciones, el 37% de la población era pobre en 1998. En ese país, el 79% de los pobres tiene ocupaciones agrícolas, y casi todos ellos trabajan por cuenta propia. Otro hecho básico es que la pobreza suele ser mucho más elevada entre los grupos de minorías étnicas: aunque los grupos minoritarios constituyen sólo el 14% de la población general, representan el 29% de los pobres en Viet Nam.

13. Una de las características más importantes de los pobres es dónde viven. En principio, las autoridades desearían conocer la incidencia de la pobreza en cada ciudad, pueblo y distrito rural. Por desgracia, el tamaño de la muestra de una encuesta típica de hogares suele ser de entre 3.000 y 15.000 hogares, cifra demasiado pequeña para ofrecer estimaciones precisas de la pobreza en un nivel tan desglosado. No obstante, si se dispone todavía de los datos del censo, es posible combinar esos datos con los de las encuestas de hogares para obtener estimaciones de la pobreza sobre zonas geográficas mucho más pequeñas. La idea básica es estimar la relación entre algunas variables "predictivas" y el ingreso o gasto de los hogares, utilizando los datos de la encuesta de hogares. Las variables productivas utilizadas son variables que se encuentran también en el censo. El hecho de que, con una estimación de la relación predictiva, los datos del censo puedan utilizarse para simular la distribución de los gastos en zonas geográficas relativamente pequeñas permite estimar la incidencia de la pobreza en esas zonas. Un ejemplo de esta utilización de los datos, en el caso del Ecuador, se encuentra en Hentschel y otros (1998). Pueden verse análisis más detallados de estos métodos en Rao (2002) y Kalton (2002).

14. Una observación final importante respecto de la definición de la pobreza y la creación de perfiles de pobreza es que muchas veces lo que interesa es comparar la pobreza en diferentes momentos dentro del mismo país, o en el mismo punto cronológico pero en diferentes países. Entonces, es importante que los datos de la encuesta de hogares definan los gastos o ingresos que deberán recopilarse de la misma manera a lo largo del tiempo o en los distintos países. Las diferencias en el diseño del cuestionario u otros cambios en el método de recopilación de datos, por pequeñas que sean, pueden dar lugar a cambios significativos y totalmente engañosos en las estimaciones, muchas veces de modo imprevisto. Es más, quizá no sea posible efectuar esas comparaciones si los datos recopilados o la forma en que se analizan ambas cosas no son iguales en las encuestas que se comparan. Por ello, todo cambio en la manera de recopilar los datos para las variables que definen la pobreza debe considerarse con suma atención, a fin de limitar el potencial de cambios observados debidos meramente a procedimientos estadísticos más que a cambios reales. Por ello, normalmente lo mejor es no cambiar de forma significativa la manera como se recopilan los datos.

3. Utilización de los perfiles de pobreza para análisis básicos de las políticas

15. El conocimiento de la ubicación de los pobres y algunas de sus características básicas es el punto de partida para poder asesorar a las autoridades. Naturalmente, los programas específicos para ayudar a los pobres deben situarse precisamente donde mayor es su concentración, pero es mucho lo que se puede conseguir, desde el punto de vista programático, si las estadísticas elementales sobre los pobres se analizan debidamente. En este apartado se describen cuatro tipos de información básica sobre los pobres que pueden utilizarse para extraer enseñanzas sobre la repercusión de varias políticas en ese grupo de la población.

16. *Cómo ganan los pobres sus ingresos*. Como se ha explicado antes, una de las formas en que las políticas gubernamentales repercuten en los pobres es su influencia en los ingresos que perciben. Por ello, una pregunta importante es qué hacen los pobres para ganarse la vida. Quizá lo primero que habría que preguntarse es si los pobres trabajan por cuenta propia, a diferencia de quienes perciben salarios trabajando para un empleador. En muchos países, la inmensa mayoría de los pobres son agricultores, artesanos o comerciantes que trabajan por cuenta propia. Por definición, los pobres que trabajan por cuenta propia no se ven directamente afectados por las políticas destinadas a los asalariados, como el cambio de las leyes sobre salario mínimo o un plan de "seguridad social" o de seguro médico sólo para los asalariados.

17. Muchos pobres son agricultores que trabajan por cuenta propia, por lo que es importante hacerse la siguiente pregunta: ¿qué cultivos producen y qué parte de su producción se vende? Podría proponerse como ejemplo concreto el de Côte d'Ivoire. Glewwe y de Tray (1990) observaron que muchos agricultores pobres de ese país producen algodón, mientras que en el caso de los agricultores que no eran pobres la producción de algodón era más bien rara. Por ello, las políticas gubernamentales que influyen en el precio del algodón repercutirán sobre todo en los pobres.

18. *Pautas de consumo de los pobres*. El bienestar económico de los pobres está determinado también por los precios de los productos y servicios que consumen. Por ejemplo, en Ghana menos del 1% del 20% más pobre de la población posee su propia motocicleta o automóvil (Glewwe y Twum-Baah, 1991). Ello significa que el aumento del precio de la gasolina tendrá poco efecto directo en los pobres del país, aunque quizá tenga repercusiones indirectas debido al aumento de los costos del transporte público.

19. En términos más generales, los datos sobre el consumo de los alimentos y los productos no alimenticios así como sobre la disponibilidad de electricidad y agua corriente, aportan abundante información que las autoridades pueden tener en cuenta. Cuando se considere la posibilidad de adoptar una subvención o impuesto para un tipo determinado de

producto, los datos deberían examinarse para comprobar hasta qué punto se verán afectados los pobres. Téngase también en cuenta que las políticas cambiarias influyen en los precios, por lo que resulta interesante conocer cuántos productos importados consumen los pobres. Un ejemplo ilustrativo es el de Ghana, que se acaba de mencionar: todos los productos del petróleo de Ghana son importados.

20. *Servicios utilizados por los pobres*. Las subvenciones a la salud y a la educación muchas veces se justifican, al menos en parte, por los beneficios que aportan a los pobres. No obstante, hay muchos tipos de servicios de salud y muchas formas y niveles diferentes de educación. Los datos sobre quién utiliza esos servicios ofrecen la oportunidad de comprobar la situación económica de los beneficiarios de las políticas concretas.

21. Un ejemplo reciente es el de Viet Nam. Gertler y Litvack (1998) observaron que la persona típica del 20% más pobre de la población realizaba aproximadamente una visita al año como paciente externo a un hospital gubernamental y aproximadamente dos a un centro de salud comunitario. Por el contrario, una persona típica del 20% más adinerado de la población visitaba entre cuatro o cinco veces al año un hospital gubernamental y sólo una el centro comunitario de salud. La principal razón de esta divergencia es que los hospitales gubernamentales se encuentran sobre todo en las zonas urbanas, mientras que aproximadamente el 90% de los pobres de ese país vive en las zonas rurales. De esas cifras sencillas se deduce claramente que las subvenciones a los centros comunitarios de salud benefician a los pobres más que a quienes no lo son, mientras que con las subvenciones a los hospitales ocurre lo contrario.

22. *Participación en programas.* Un último uso sencillo de los datos de una encuesta de hogares es examinar quién participa en los distintos programas gubernamentales cuyo objetivo es ayudar a los pobres. Para ello se requiere una encuesta de hogares con una o más preguntas específicas sobre la participación de los hogares en programas, así como datos sobre ingresos o gastos que puedan utilizarse para clasificar a los hogares en pobres o no pobres. Si bien estos datos eran raros en el pasado, son más frecuentes a medida que los encargados de diseñar las encuestas reconocen su valor.

23. Como ejemplo del uso de una encuesta de hogares para evaluar la orientación de un programa puede citarse el de Jamaica (Grosh, 1991). Los datos de la encuesta de hogares revelaban que, como era previsible, lo cupones para alimentos eran mucho más utilizados por los hogares pobres. Paradójicamente, los beneficios de las subvenciones generales a los alimentos solían recaer fundamentalmente sobre los hogares en mejor situación económica. Esta información se presentó al Gobierno en los últimos años ochenta, y en los primeros noventa los beneficios del programa de cupones se duplicaron mientras que se puso fin a las subvenciones alimentarias.

24. Una última observación general sobre el análisis descriptivo básico es que casi todas las encuestas de hogares están basadas en diseños muestrales complejos más que en muestras aleatorias. En consecuencia, los grupos de subpoblación de especial interés, como los pobres, aparecen excesivamente representados en la muestra, lo que significa que es preciso utilizar ponderaciones para obtener estimaciones no sesgadas de las estadísticas descriptivas básicas. Además, el cálculo de los errores estándar debe tener en cuenta el diseño de la muestra. Como estos puntos se examinan con mayor detalle en el capítulo XVI y en otros capítulos de esta publicación, el lector debería consultar esos capítulos antes de proceder a un análisis descriptivo.

C. Análisis de regresión múltiple de los datos de encuestas de hogares

25. Los ejemplos anteriores sobre el uso de las encuestas de hogares están basados en estadísticas muy sencillas que podría calcular cualquiera que disponga de un sencillo pro-

grama de estadística. No obstante, la enseñanzas que se han extraído para la formulación de políticas quizá sean demasiado simplistas, es el sentido de que no tienen en cuenta las respuestas conductuales a esas políticas. Por ejemplo, si se elimina un impuesto de un determinado producto agrícola porque es cultivado sobre todo por los pobres, los hogares que no son pobres quizá comiencen también a producir ese cultivo cuando aumenten los precios, por lo que algunos de los beneficios de la política podrían ir a hogares que no son pobres[1]. De la misma manera, un impuesto sobre un tipo determinado de alimento puede parecer que tiene importantes efectos negativos en los pobres si éstos lo consumen en grandes cantidades; pero si hay otro producto semejante que no está sometido a impuestos, es posible que los pobres simplemente comiencen a consumir ese producto, con sólo una pequeña reducción de su bienestar. Otro ejemplo sería el de la educación. El hecho de que los niños pobres de un determinado país no cursen habitualmente estudios secundarios indica que la reducción de las tasas de matrícula para esos tipos de escuelas produce pocos beneficios para los pobres, aunque es posible que esa reducción incremente notablemente la tasa de matrícula de los niños pobres en tales escuelas. A su vez, esta posibilidad significa que la consideración de las pautas actuales de matrícula infravalorarían los beneficios de dicha política para los pobres.

[1] Cuando el impuesto está vigente, el precio recibido por los productores es más bajo que el pagado por los consumidores, y la diferencia equivaldría al impuesto. Cuando se elimina el impuesto, el precio del productor será igual al precio del consumidor, y en casi todos los casos ello significa que el precio recibido por los productores aumentará y el precio pagado por los consumidores disminuirá.

26. Las encuestas de hogares pueden servir para ver cómo se comportan los hogares en respuesta a los cambios en las políticas. La tarea no es fácil, ya que requiere tipos de análisis mucho más complejos. Los métodos más comunes utilizados para realizar estas estimaciones son los del análisis de regresión múltiple. Los métodos más avanzados suelen utilizar datos de encuestas de hogares concebidas de manera específica y que recopilan los datos precisos necesarios para la realización de dicho análisis. Ello se debe a que estos métodos requieren con frecuencia datos que no se encuentran en las encuestas de hogares habituales. En los tres subapartados que siguen se describen tres formas comunes de utilizar los datos de las encuestas de hogares para estimar qué influencia pueden tener las políticas en el comportamiento de los hogares. Puede verse un examen más detallado en Deaton (1997).

1. Análisis de la demanda

27. Los economistas estiman con frecuencia los efectos de los precios y los ingresos de los hogares en las compras de bienes y contratación de servicios. Esta investigación se conoce con el nombre de análisis de la demanda. El concepto general es que para cualquier bien (i), las compras de dicho bien (q_i) por un hogar están determinadas por el ingreso (y) del hogar, el precio (p_i) de ese bien y los precios de todos los demás bienes. Esta relación puede expresarse así:

$$q_i = f(y, p_1, p_2, \ldots p_i, \ldots p_n) + \varepsilon \approx \beta_0 + \beta_1 p_1 + \beta_2 p_2 + \ldots + \beta_i p_i + \ldots \beta_n p_n + \beta_{n+1} y + \varepsilon$$

La función $f(y, p_1, p_2, \ldots p_i, \ldots p_n)$ es una representación muy general de cómo afectan a la demanda de los hogares los ingresos y los precios, donde ε refleja los efectos de otros factores causales (y quizá una variación aleatoria de q_i que no tiene nada que ver con ningún factor causal). Una simplificación frecuente del análisis de la demanda es asumir una representación lineal, que se expresa aquí mediante el término que se encuentra a la derecha del símbolo ≈, que indica que esta simplificación tiene carácter aproximativo. Si ε no está correlacionado con las variables de los precios, pueden utilizarse cuadrados mínimos ordinarios para obtener estimaciones no sesgadas de los coeficientes (β) del ingreso (y) y los precios (p_i) en esta relación lineal. En las aplicaciones reales este supuesto no es válido, y muchas otras cuestiones podrían complicar el análisis. Puede encontrarse más información sobre la estimación del sistema de demanda en la obra clásica de Deaton y Muellbauer (1980). Véanse estudios más recientes en Pollack y Wales (1992) y Lewbel (1997).

28. Para entender cómo el análisis de la demanda aporta información que va más allá de la obtenida utilizando estadísticas descriptivas sencillas, basta considerar los efectos de un

impuesto en un alimento importado, como el trigo (los países en desarrollo muchas veces gravan con impuestos los productos importados porque esos impuestos son relativamente fáciles de administrar, y el clima tropical de muchos países en desarrollo permite pensar que las importaciones pueden ser la única fuente de adquisición de trigo). Supongamos que el precio corriente de un kilogramo de harina de trigo es 10, y que el hogar pobre típico consume 60 kilogramos de harina de trigo al año. Suponiendo que el precio del trigo sea el precio internacional, un impuesto del 50% sobre las importaciones de trigo elevaría el precio de la harina de trigo a 15, lo que significa que el hogar pobre típico pagaría 300 (5×60) en impuestos adicionales. Naturalmente, este análisis basado en estadísticas descriptivas simples supone que los hogares pobres continuarán comprando el mismo volumen de trigo después de establecido el impuesto. En la práctica, es probable que reduzcan ese consumo y aumenten el de otros cultivos básicos (como arroz, maíz o yuca) en respuesta a la subida del precio de la harina de trigo. Las estimaciones del análisis de la demanda permiten calcular la magnitud de esa respuesta. Supongamos que la ecuación del párrafo anterior corresponde a la demanda de trigo, por lo que q_i representa los kilogramos de harina de trigo comprados por los hogares al año y p_i representa el precio de un kilogramo de harina de trigo. Si $\beta_i = -3$, una subida de 5 del precio de la harina de trigo reducirá el consumo de trigo en 15, por lo que el consumo anual de un hogar pobre medio sería de 45 kilogramos. A su vez, ello significa que el hogar pobre medio pagaría 225 (5×45) en concepto de impuestos adicionales, en vez de 300. Este ejemplo es muy simple, pero revela la necesidad de tener en cuenta el comportamiento de los hogares al examinar los efectos de políticas concretas.

29. Un ejemplo del uso del análisis de la demanda para analizar los efectos de las políticas gubernamentales en los pobres es el de Deaton, Parikh y Subramanian (1994). Los autores estiman un sistema de ecuaciones de la demanda para más de 10 tipos diferentes de alimentos. Calculan el efecto global de las subidas de los precios de los alimentos en el bienestar social nacional, y la repercusión de los cambios en el bienestar de los pobres. Un resultado particularmente interesante es que las subidas de precios del arroz tienen menos efectos negativos en el bienestar de los pobres que las subidas del precio de los cereales secundarios, ya que los pobres dependen más de estos últimos. Por ello, los impuestos que gravan el arroz perjudican a los pobres menos que los de los cereales secundarios.

2. Uso de los servicios sociales

30. Los programas de salud y educación pueden aportar muchos beneficios a los hogares pobres, pero la participación en ellos no implica necesariamente que se hayan recibido beneficios considerables. Algunos de esos programas pueden ser ineficaces. En la esfera de la salud, las autoridades desearían saber si la participación mejoró de hecho el estado de salud de los individuos. En la educación, les interesaría saber cuánto aprendieron realmente los niños en la escuela. Se han realizado muchos estudios en los que se utilizan los datos de encuestas de hogares de países en desarrollo con el fin de investigar hasta qué punto los programas de salud y de educación consiguen sus objetivos.

31. Un ejemplo reciente que ilustra el uso de la regresión múltiple es un análisis del efecto de algunas características de las escuelas en el aprendizaje de los alumnos y, por lo tanto, en los salarios futuros. Glewwe (1999) examinó esta cuestión estimando los efectos de las características de las escuelas y los hogares en los resultados académicos de los niños, reflejados en las calificaciones de los exámenes. Para su investigación utilizó datos de encuestas de hogares de Ghana. La ecuación utilizada adoptó la forma siguiente:

$$T_i = \beta_0 + \beta_1 \text{MED}_i + \beta_2 \text{FED}_i + \beta_3 y_i + \beta_4 \text{IQ}_i + \beta_5 \text{SC}_{1i} + \beta_6 \text{SC}_{2i} + \ldots + \varepsilon,$$

donde T_i es la calificación obtenida por el niño i, MED_i y FED_i son los niveles de educación de la madre y el padre del niño, respectivamente, y_i es el ingreso del hogar del niño i, y las

variables SC representan un gran número de características de la escuela y del personal docente. La estimación de esta ecuación es bastante complicada (véase Glewwe, 2002), pero una vez que se estiman las β, se dispone de información sobre la influencia que las diferentes características de las escuelas y profesores tienen en el nivel de instrucción de los alumnos. Si se comparan esos efectos con los costos de las diversas características de las escuelas y profesores se tiene cierta idea sobre qué tipos de gastos en educación son los más eficaces en función de los costos.

32. El análisis de Glewwe sobre los datos de Ghana reveló que la reparación de las goteras y la compra de pizarras para las aulas mejoró significativamente los resultados escolares y elevó el número de años pasados en la escuela. Con sencillos cálculos de la rentabilidad financiera de esas "inversiones" en la calidad de las escuelas se comprobó que la rentabilidad era muy elevada; en algunos casos, del 25% o más.

3. Efectos de programas gubernamentales concretos

33. Si bien es fácil utilizar los datos de encuestas de hogares para examinar si un hogar o individuo concreto participa en un determinado tipo de programa destinado a ayudar a los pobres, es más fácil establecer hasta qué punto su participación mejora de hecho su bienestar. El problema es que la participación puede tener otros efectos que reduzcan el bienestar. Por ejemplo, un programa de "alimentos por trabajo" puede ofrecer empleo a personas pobres, pero los beneficios del aumento del ingreso deben sopesarse teniendo también en cuenta el costo del trabajo, incluida la repercusión del trabajo en la salud. De la misma manera, cuando se ofrece a los hogares cupones para alimentos con el fin de aumentar su consumo alimentario, no siempre ocurre que el uso de esos cupones para comprar alimentos incremente su consumo, ya que los beneficiarios pueden desviar hacia otros usos el dinero que en circunstancias normales tendrían que utilizar para comprar alimentos. La evaluación del impacto de los programas en el comportamiento de los hogares requiere un análisis econométrico atento y complejo para comprender todos los efectos de la participación en el programa y, en definitiva, el efecto global de la participación en el bienestar del hogar.

34. Puede encontrarse un ejemplo reciente de ello en un estudio de Jacoby (2002), en donde se examinan los efectos de los programas de alimentación escolar en Filipinas. En concreto se analiza si el ofrecimiento de comidas a los alumnos escolarizados tenía como consecuencia que los padres les dieran menos alimentos en casa. Si bien la mayor parte de los economistas habrían previsto una reasignación de alimentos en ese sentido, Jacoby no encontró ninguna prueba que lo confirmara. Lo que comprobó es que la participación en el programa de alimentación escolar no tenía ningún efecto en el consumo de alimentos de los niños en casa, lo que significa que el consumo general de alimentos entre los niños participantes registró un aumento equivalente al volumen de los alimentos ofrecidos en las escuelas.

D. Resumen y observaciones finales

35. Las encuestas de hogares constituyen una rica fuente de información que pueden utilizar las autoridades y los encargados de diseñar los programas para determinar si las medidas adoptadas benefician a los hogares pobres. Una encuesta sólo puede ser útil si contiene datos sobre los ingresos o gastos, a fin de clasificar los hogares como pobres o no pobres, y datos que indiquen cómo se verá afectado el hogar por una política o programa concretos. Hasta hace poco, las encuestas de hogares utilizadas se habían concebido con otras finalidades. No obstante, en los decenios de 1980 y 1990 se elaboraron nuevas encuestas de hogares con la intención explícita de aportar ese tipo de información. Entre las más destacadas cabe citar las encuestas de hogares del Estudio de medición de los niveles de

vida (EMNV), del Banco Mundial. Puede verse una breve presentación de esas encuestas en Grosh y Glewwe (1998) y un análisis sumamente detallado en Grosh y Glewwe (2000). No obstante, incluso las encuestas estándar concebidas con otros fines pueden aprovecharse en forma mucho más útil para el análisis de la pobreza si se agregan algunas preguntas. Por ejemplo, sería muy útil incorporar preguntas sobre la participación en programas nacionales contra la pobreza (como los programas de empleo rural o los de cupones para alimentos) a una encuesta estándar de ingresos y gastos de los hogares.

36. En este capítulo se ha presentado un panorama sobre la manera de utilizar las encuestas de hogares para formular políticas que reduzcan la pobreza en los países en desarrollo. Ciertamente, se trata de un análisis demasiado breve, debido a los límites de espacio de esta publicación. Los lectores que deseen un examen más detallado deberían consultar las publicaciones y los documentos citados a lo largo del capítulo.

REFERENCIAS BIBLIOGRÁFICAS

Banco Mundial (1999). *Viet Nam: Attacking Poverty*. Joint Report of the Government of Viet Nam —Donor—. NGO Poverty Working Group. Hanoi.

_____ (2001). *Informe sobre el desarrollo mundial, 2000/2001: Lucha contra la pobreza*. Nueva York: Oxford University Press.

Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore, Maryland: Johns Hopkins University Press.

_____ y John Muellbauer (1980). *Economics and Consumer Behaviour*. Cambridge (Reino Unido) y Nueva York: Cambridge University Press.

Deaton, Angus, Kirit Parikh y Shankar Subramanian (1994). Food demand patterns and pricing policy in Maharashtra: an analysis using household-level survey data. *Sarvekshana,* vol. 17, págs. 11-34.

Gertler, Paul y Jennie Litvack (1998). Access to health care during the transition: the role of the private sector in Viet Nam. En *Household Welfare and Viet Nam's Transition,* D. Dollar, P. Glewwe y J. Litvack, comps. Washington, D.C.: Banco Mundial.

Glewwe, Paul (1999). *The Economics of School Quality Investments in Developing Countries.* Londres: Macmillan.

_____ (2002). Schools and skills in developing countries: education policies and socioeconomic outcomes. *Journal of Economic Literature,* vol. 40, No. 2, págs. 436-482.

_____ y Dennis de Tray (1990). The poor during adjustment: a case study of Côte d'Ivoire. En *Macroeconomic Policy Reforms, Poverty, and Nutrition,* P. Pinstrup-Andersen, comp. Ithaca, Nueva York: Cornell Food and Nutrition Policy Program Monograph, No. 3.

Glewwe, Paul y Kwaku Twum-Baah (1991). *The Distribution of Welfare in Ghana, 1987-1988*. Living Standards Measurement Study Working Paper, No. 75. Washington, D.C.: Banco Mundial.

Grosh, Margaret (1991). *The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions*. Living Standards Measurement Study Working Paper, No. 80. Washington, D.C.: Banco Mundial.

_____ y Paul Glewwe (1998). Data watch: the World Bank's Living Standards Measurement Study Household Surveys. *Journal of Economic Perspectives,* vol. 12, No. 1, págs.187-196.

_____ y Paul Glewwe (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Development Study*. Nueva York: Oxford University Press (para el Banco Mundial).

Hentschel, Jesko y otros (1998*). Combining Census and Survey Data to Study the Spatial Dimensions of Poverty*. Policy Research Working Paper, No. 1928. Washington, D.C.: Banco Mundial.

Jacoby, Hanan (2002). Is there an intra-household "flypaper effect"? Evidence from a school feeding programme. *Economic Journal,* vol.112, No. 476 (enero), págs. 196-221.

Kalton, Graham (2002). Models in the practice of survey sampling (revisited). *Journal of Official Statistics,* vol. 18, págs. 129-154.

Lewbel, Arthur (1997). Consumer demand systems and household equivalence scales. En *Handbook of Applied Economics,* vol. II, *Microeconomics.* H. Pesaran y P. Schmidt, comps. Oxford (Reino Unido): Blackwell.

Lipton, Michael y Martin Ravallion (1995). Poverty and policy. En *Handbook of Development Economics,* vol. 3. J. Behrman y T. N. Srinivasan, comps. North Holland.

Pollak, Robert y Terence Wales (1992). *Demand System Specification and Estimation*. Oxford (Reino Unido): Oxford University Press.

Rao, J. N. K. (2002). *Small Area Estimation*. Nueva York: Wiley.

Ravallion, Martin (1998). *Poverty Lines in Theory and Practice*. Living Standards Measurement Study Working Paper, No. 133. Washington, D.C.: Banco Mundial.

Capítulo XVIII
Métodos con múltiples variables para la elaboración de índices

Savitri Abeyasekera
Centro de Servicios de Estadística
Universidad de Reading
Reino Unido de Gran Bretaña e Irlanda del Norte

RESUMEN

Las encuestas, por su misma naturaleza, producen estructuras de datos que contienen múltiples variables. Aun reconociendo el valor de los métodos sencillos de análisis de datos de las encuestas, en este capítulo se ilustran los beneficios de un análisis más en profundidad para determinados subgrupos de población, mediante la aplicación de técnicas con múltiples variables. Actualmente pueden utilizarse programas informáticos que permiten aplicar estos métodos más avanzados en la investigación de las encuestas.

En este capítulo se presentan diversas situaciones en que los métodos con múltiples variables pueden contribuir a la creación de índices y a la exploración inicial de los datos con subconjuntos específicos de datos, antes de realizar ulteriores análisis para abordar objetivos específicos de la encuesta. Se insiste sobre todo en los métodos que implican el estudio simultáneo de diversas variables fundamentales. En este contexto, los métodos con múltiples variables permiten una exploración más profunda sobre las posibles pautas existentes en los datos, hacen posibles interrelaciones complejas entre numerosas variables que deben representarse gráficamente y ofrecen los medios de reducir la dimensionalidad de los datos para los resúmenes y ulteriores análisis. Al considerar la creación de índices se utiliza la interpretación más amplia de los métodos con múltiples variables a fin de incluir los métodos basados en la regresión.

En el capítulo se procura presentar un panorama de los métodos con variables múltiples a fin de que se pueda comprender, desde un punto de vista práctico, su valor para la creación de índices. Los destinatarios son tanto quienes participan en encuestas de hogares en gran escala como los investigadores de encuestas que intervienen en proyectos de investigación y desarrollo y que quizá tengan poca experiencia en la aplicación de los conceptos de análisis aquí descritos. El uso de estos métodos se ilustra con ejemplos adecuados y una exposición sobre la manera de interpretar los resultados.

Términos clave: Creación de índices, métodos con variables múltiples, componentes principales, análisis por conglomerados.

A. Introducción

1. Al analizar los datos de las encuestas, la mayoría de los analistas suelen utilizar planteamientos estadísticos sencillos. El más frecuente es el uso de cuadros de una, dos o más ejes de lectura, y de gráficos de barras, de líneas, etcétera. Puede encontrarse una exposición

sobre estos planteamientos y de algunos aspectos que necesiten atención durante el proceso de análisis de datos en Wilson y Stern (2001) y en los capítulos XV y XVI de la presente publicación. No obstante, en algunos casos conviene recurrir a procedimientos de análisis que vayan más allá de los simples resúmenes. En este capítulo se analizan precisamente algunos de esos procedimientos.

2. Los métodos con variables múltiples permiten el tratamiento simultáneo de distintas variables (Krzanowski y Marriott, 1994a y 1994b; Sharma, 1996). En estricto sentido estadístico, se refieren al estudio colectivo de un grupo de variables de resultados, por lo que tienen en cuenta la estructura de correlación de las variables dentro del grupo. No obstante, muchos investigadores utilizan también el término "múltiples variables" en la aplicación de técnicas de regresión múltiple porque intervienen distintas variables explicativas (predictivas) junto con la variable resultante principal (por ejemplo, Ruel, 1999). También en este caso la ventaja de explorar diversas variables juntas es la posibilidad de establecer intercorrelaciones. La regresión, que implica fundamentalmente la elaboración de un modelo de una variable de respuesta clave, se examina con más detalle en el capítulo XIX. Aquí nos concentraremos sobre todo en el estudio conjunto de diversas variables de medición como paso preliminar para nuestra interpretación más amplia de los métodos con múltiples variables en el examen de la creación de índices.

3. Las técnicas de múltiples variables se consideran con frecuencia como técnicas "avanzadas" que requieren un alto nivel de conocimientos estadísticos. Si bien es cierto que los aspectos teóricos de muchos procedimientos con variables múltiples y de su aplicación pueden inspirar temor incluso a los estadísticos, pueden ser de utilidad para el análisis de datos de las encuestas de países en desarrollo. Examinamos en primer lugar el uso eficaz de estos métodos para investigar las pautas de los datos, para identificar agrupaciones naturales de la población en orden a ulteriores análisis y para reducir la dimensionalidad del número de variables implicadas. Consideramos estos pasos como preliminares para la creación de índices a partir de variables de hogares, por ejemplo, con el fin de elaborar indicadores de la pobreza (véase, por ejemplo, Sahn y Stifel (2000)).

4. En el apartado B se presenta un panorama general de las técnicas con múltiples variables en cuanto estudio colectivo de un grupo de variables resultantes. En los cinco apartados siguientes se consideran las esferas de aplicación, con varios ejemplos ilustrativos. En el apartado final se presentan algunas conclusiones sobre el valor y las limitaciones de estas técnicas. Los detalles técnicos se han reducido al mínimo y se insiste en la comprensión de los conceptos implicados y en la interpretación. El lector que desee profundizar respecto a estas técnicas puede consultar Everitt y Dunn (2001); y Chatfield y Collins (1980).

B. Algunas restricciones al uso de los métodos con múltiples variables

5. En este capítulo insistimos en la utilización de planteamientos con múltiples variables como procedimientos descriptivos valiosos durante las fases iniciales de la exploración de datos y la creación de índices. No obstante, en la aplicación de estos métodos es importante insistir desde el comienzo en que un análisis aplicado a todo el conjunto de datos de una encuesta nacional de hogares no producirá probablemente conclusiones útiles, debido a la inevitable diversidad de hogares en cualquier país. Es posible que se pierda valiosa información si un análisis combina las poblaciones urbanas y las rurales y mezcla diferentes zonas agroecológicas, ya que los medios de subsistencia de los hogares dentro de esos diferentes estratos pueden presentar grandes diferencias. Las técnicas descritas en este capítulo deberían utilizarse, por lo tanto, únicamente después de un atento examen de la estructura de datos, para identificar los diferentes sectores o sustratos de la población a que pueden aplicarse los métodos, teniendo siempre presentes los principales objetivos de la encuesta.

6. Incluso dentro de estos sustratos, o en los casos en los que se necesita el análisis de toda la muestra, habrá que prestar atención a las ponderaciones de la muestra asociadas con las unidades incluidas. Si éstas varían sustancialmente en relación con los datos que son objeto de análisis, el uso de un programa informático que no cuente con recursos para tener en cuenta las ponderaciones de la muestra puede dar lugar a conclusiones erróneas. En ese caso la ponderación de las unidades de la muestra mediante las ponderaciones muestrales, utilizando por ejemplo la fórmula *WEIGHT* de SAS (2001) o la orden *aweight* de STATA (2003) permitirá superar esa dificultad con respecto a los métodos incluidos en los apartados C, D, E y F. Muchos más programas informáticos tendrán en cuenta las ponderaciones de muestreo con respecto a los métodos descritos en el apartado G. Cuando no se utilizan ponderaciones de muestreo hay que proceder con cuidado al interpretar los resultados, ya que pueden estar expuestos a sesgos.

C. Descripción general de los métodos con múltiples variables

7. El tema fundamental en el que se basa el uso de los métodos con variables múltiples en las investigaciones sobre encuestas es la simplificación. Por ejemplo, reducir un conjunto de datos voluminoso y posiblemente complejo a un número reducido de medidas sinópticas significativas, o la identificación de las características clave y de las posibles pautas de interés en los datos. El objetivo es muchas veces de carácter exploratorio: dichos métodos pueden ayudar al investigador a generar hipótesis interesantes, más que a comprobarlas. Muchos de los planteamientos utilizan métodos de libre distribución que no suponen una distribución estadística subyacente para ninguna de las variables. No obstante, como se necesita cierta cautela con respecto a los tipos de datos utilizados (por ejemplo, escala del intervalo, recuentos, binarios), haremos referencia a este tema cuando resulte pertinente según se desarrolle el capítulo.

8. El punto de partida es una matriz de datos en que las filas representan casos (las unidades de la muestra) y las columnas corresponden a las variables. Algunas veces las filas son de mayor interés; por ejemplo, si representan a los hogares agrícolas, puede ser conveniente agrupar los hogares en diferentes categorías de riqueza sobre la base de un número de criterios socioeconómicos representados por algunas columnas de la matriz de datos. En otros casos el máximo interés se centra en las columnas; por ejemplo, cuando un conjunto de variables que corresponden a un tema concreto deben combinarse en forma de índice compuesto para ulteriores análisis.

9. En los apartados siguientes nos centramos en cuatro planteamientos principales para manipular los datos con variables múltiples en las encuestas de países en desarrollo. Los tres primeros pueden considerarse como técnicas de exploración que dan lugar a la creación de índices. En primer lugar consideramos los procedimientos gráficos y las medidas sinópticas que contribuirán a una comprensión de los datos. Luego analizamos dos procedimientos muy conocidos con variables múltiples —el análisis por conglomerados y el análisis de componentes principales (ACP)—, por tratarse de dos de los procedimientos principales que pueden desempeñar una función preliminar útil en la creación de índices. Este último procedimiento se examina con más detalle en el apartado G, junto con otras maneras de creación de índices, adoptando la interpretación más amplia de los métodos "con múltiples variables", utilizados por muchos investigadores. En todo momento suponemos que se ha seleccionado un subconjunto adecuado de datos de la encuesta para el análisis y que el objetivo de someter esos datos a un procedimiento con múltiples variables es integrar un elemento de exploración en un análisis que trata de alcanzar un objetivo más amplio de la encuesta.

10. Naturalmente, hay muchos otros métodos con variables múltiples que podrían considerarse en situaciones concretas. En el cuadro XVIII.1 puede observarse una gran variedad de esos métodos, junto con una breve descripción de cada uno. En este capítulo se

Cuadro XVIII.1
Algunas técnicas con variables múltiples y su objetivo

	Tipo de técnica	Objetivo de la técnica
1.	Métodos descriptivos con variables múltiples	Exploración de datos; identificación de pautas y relaciones
2.	Análisis de componentes principales	Reducción de la dimensión formando nuevas variables (los componentes principales) como combinaciones lineales de las variables del conjunto con múltiples variables
3.	Análisis por conglomerados	Identificación de agrupaciones naturales entre los casos o variables
4.	Análisis factorial	Elaboración de modelos de la estructura de correlación entre variables en el conjunto de respuestas con variables múltiples relacionándolas con un conjunto de factores comunes
5.	Análisis multivariado de la varianza (MANOVA)	Ampliación del análisis multivariado de la varianza al estudio simultáneo de distintas variables. El objetivo es dividir la suma total de cuadrados y matriz de productos transversales entre un conjunto de variables de acuerdo con la estructura del diseño experimental
6.	Análisis discriminante	Determinar una función que permita la separación de dos o más grupos de individuos
7	Análisis de la correlación canónica	Estudiar la relación entre dos grupos. Supone la formación de parejas de combinaciones lineales de las variables en el conjunto de variables múltiples para que cada pareja produzca a su vez la máxima correlación entre individuos de los dos grupos
8.	Escalonamiento multidimensional	Creación de un "mapa" en el que se presente una relación espacial entre un número de objetos, comenzando con un cuadro de distancias entre los objetos

consideran sólo los tres primeros, ya que el objetivo es considerar la exploración de datos como primer paso necesario para la creación de índices. Es probable también que estos tres métodos sean los de mayor relevancia en el análisis de datos de encuesta. Junto con la aplicación más general del término "variables múltiples" en nuestro examen sobre la creación de índices, constituyen valiosos instrumentos metodológicos adicionales en el análisis de datos de encuestas. Los otros métodos del cuadro XVIII.1 pueden ser útiles en ocasiones específicas cuando están relacionados con los objetivos de la encuesta. No obstante, desbordan el contenido del presente capítulo, en el que se intenta ofrecer únicamente una presentación general de algunos de los métodos más sencillos.

D. Gráficos y medidas sinópticas

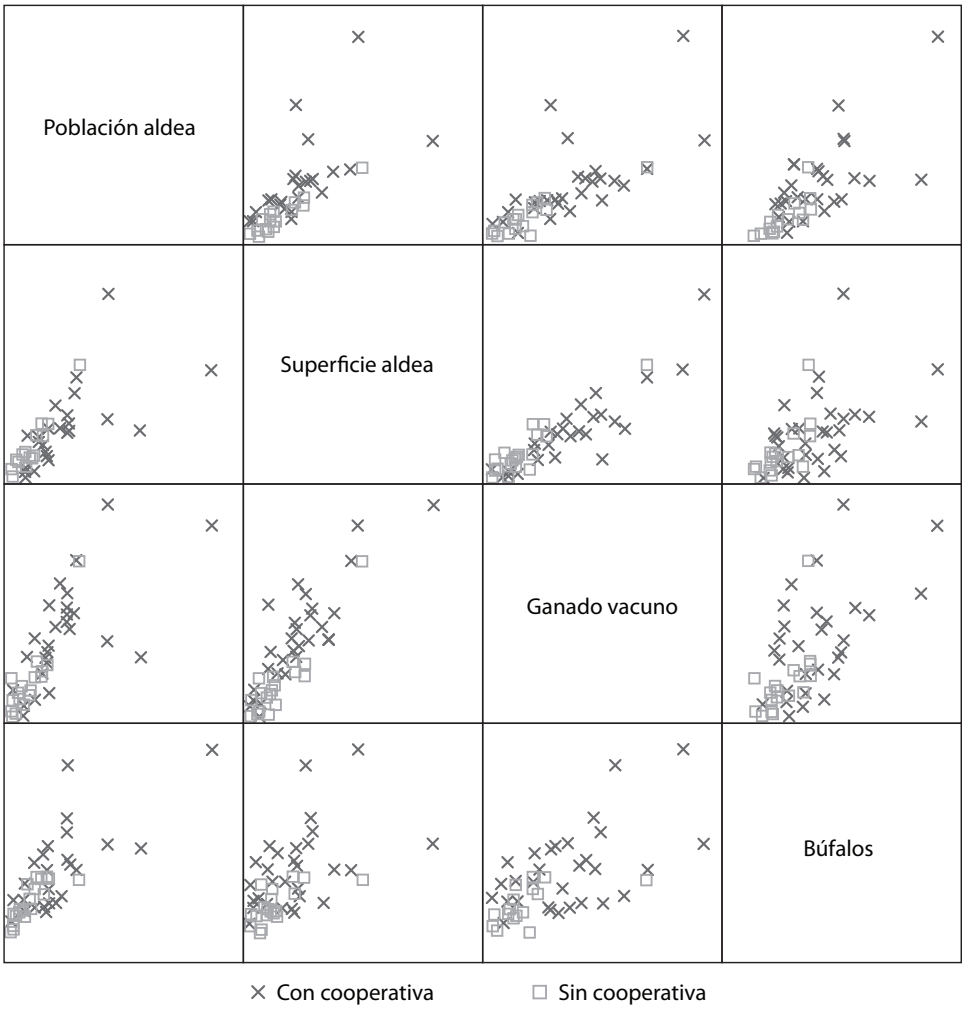
11. La comprensión preliminar de los datos es una fase inicial fundamental siempre que se procede a un análisis de datos. Su examen atento permite hacerse una idea del significado y pautas de distribución de los datos, identificar los posibles valores atípicos (observaciones en contradicción con la pauta de los demás datos), poner de manifiesto la configuración de los datos e indicar al usuario si algunas variables tienen mayor variabilidad que otras (véanse, por ejemplo, Tufte (1983) y Everitt y Dunn (2001)).

12. Igual que en un conjunto de análisis con una sola variante, en la medición de los cuadros de datos y de frecuencia de los datos binarios y categóricos son aconsejables las medidas sinópticas, como las medias y desviaciones estándar. Luego pueden considerarse pares de variables con el fin de identificar asociaciones entre las variables. En esta fase preliminar sería razonable considerar los datos en "bloques", quizá dos: uno que comprenda los datos cuantitativos (continuos o discretos) y otro con datos cualitativos (categóricos y binarios). En el primer caso serían significativos los gráficos de dispersión (en pares), mientras que para el segundo serían aconsejables los cuadros de doble entrada, también en pares, quizá combinados con algunas medidas de asociación y el uso del estadístico de prueba ji-cuadrado. Cuando sea pertinente, los diagramas de dispersión pueden presentarse también utilizando diferentes símbolos para indicar subconjuntos de datos identificados por una variable categórica.

13. La mayoría de los programas informáticos de estadística, como el procedimiento PLOT de SAS (2001), el menú Graph/Graphics de SPSS para Windows (SPSS, 2001) y GenStat para Windows (GenStat, 2002) permiten hacer gráficos matriciales. Se trata de presentaciones gráficas en las que se pueden reflejar juntos los diagramas de dispersión entre todos los pares de variables. Así se podría apreciar rápidamente la relación entre cada variable y el resto de variables del conjunto de datos con múltiples variables sometido a análisis.

14. Un ejemplo es la figura XVIII.1. Se trata de un diagrama matricial tomado de SPSS (2001) que recoge las relaciones entre cuatro variables correspondientes a 50 aldeas del estado de Gujarat (India), de las que se considera si tienen o no una cooperativa lechera. Las cuatro variables —población y superficie de la aldea y número de cabezas de ganado vacuno y de búfalos— representarían sólo un número reducido de variables de un grupo mayor. Los datos están tomados de un estudio de referencia hecho antes de la introducción de un plan para promover la capacitación en salud animal. Los ejes horizontal y vertical de cada gráfico se determinan mediante el eje que va paralelo a las casillas diagonales. Por ejemplo, las tres casillas con valores gráficos de la primera fila tienen la población de la aldea como eje vertical mientras que la superficie y el número de cabezas de ganado y de búfalos serían los ejes horizontales. Esos tres mismos gráficos aparecen en la primera columna pero con sus ejes invertidos. Hay posiblemente un valor atípico en el conjunto de datos, que se observa clara-

Figura XVIII.1
Ejemplo de un diagrama matricial entre cuatro variables



mente en las casillas de la primera fila correspondientes a una aldea con una población muy elevada. Se observa cierta asociación entre todos los pares de variables. Se aprecia también que los valores elevados de todas las variables consideradas tienen mayor probabilidad de corresponder a aldeas que tienen una cooperativa lechera que a las que no la tienen.

15. Si el gráfico matricial identifica pares particulares de variables que revelan pautas interesantes o valores atípicos, convendría repetirlas como un sencillo gráfico de dispersión de doble sentido, pero teniendo en cuenta las ponderaciones de muestreo asociadas con cada punto de los datos. Los gráficos de burbujas, en los que cada punto se representa con una burbuja de una superficie proporcional a la ponderación de la muestra (Korn y Graubard, 1998), son especialmente útiles y permiten una interpretación más significativa. Por ejemplo, un valor atípico con una gran ponderación de muestreo tendrá evidentemente una repercusión mayor que otro con una ponderación pequeña. Hay muchos otros procedimientos para tener en cuenta el diseño de la muestra en los gráficos de dispersión; por ejemplo, haciendo una submuestra de los datos con probabilidad proporcional a las ponderaciones de la muestra y luego representándolos sin tener en cuenta esas ponderaciones, o aplicando métodos de suavizado *kernel*. Puede verse información más detallada en Korn y Graubard (1998).

16. Existen muchos otros medios gráficos para presentar los datos con múltiples variables. Manly (1994) explica cómo varios objetos, descritos por distintas variables, pueden presentarse de tres maneras para indicar el perfil de los valores de las variables. En Everitt y Dunn (2001) hay un capítulo excelente que trata acerca de muchas presentaciones gráficas, entre ellas los gráficos *boxplots, coplots* y *trellis,* y en Jongman, Ter Braak y Van Tongeren (1995) se explica el uso de los *biplots*. No es posible presentar aquí información pormenorizada, por lo que animamos al lector a que examine las referencias citadas más arriba para mayor aclaración. No obstante, es importante señalar que estos procedimientos gráficos son especialmente valiosos cuando se utilizan con subgrupos específicos de la población.

E. Análisis por conglomerados

17. El análisis por conglomerados (Everitt, Landau y Leese, 2001) es una técnica basada en los datos, cuyo objetivo en general es identificar agrupaciones naturales entre las unidades de muestreo (por ejemplo, informantes, explotaciones agrícolas, hogares) para que las unidades de cada grupo (conglomerado) sean semejantes entre sí y que las unidades diferentes se encuentren en grupos distintos. Se presentan también situaciones en que la conglomeración de variables tiene sentido; por ejemplo, cuando se selecciona sólo una o dos variables de cada conglomerado para poder basar el análisis ulterior en un número menor de variables. Por ello constituye un instrumento útil en la exploración y/o reducción de datos. Puede utilizarse también para la generación de hipótesis y en otras situaciones específicas.

Ejemplo 1

18. Como ilustración, consideremos un estudio cuyo objetivo es investigar la eficacia de una serie de estrategias de manejo de plagas de bajo costo para su adopción por hogares agrícolas con pocos recursos en una determinada región. Supongamos que se lleva a cabo una encuesta entre los agricultores que puedan participar en futuros ensayos en las explotaciones con el fin de lograr un perfil socioeconómico de los hogares agrícolas, determinar cómo enfrentan las plagas los agricultores y conocer sus opiniones acerca de las plagas que sufren sus cultivos. Nos concentraremos aquí en el primero de esos objetivos para ver cómo puede aplicarse el análisis por conglomerados con el fin de elegir mejor los diferentes grupos de agricultores del estudio principal que se llevará a cabo en las explotaciones.

19. Durante la encuesta previa se midieron numerosas variables socioeconómicas para estratificar los hogares agrícolas en función de ellas. Un enfoque consiste en elegir dos variables clave y formar estratos definidos por combinaciones de categorías asociadas con

esas dos variables. Si las variables elegidas fueran el género del jefe de hogar (hombre/mujer) y el nivel de seguridad alimentaria del hogar (bajo, medio, alto), resultarían seis estratos.

20. La desventaja de este planteamiento es que no tiene en cuenta otras características socioeconómicas de los hogares. Un planteamiento con múltiples variables permite considerar simultáneamente muchas variables. El análisis por conglomerados, aplicado a los hogares agrícolas teniendo en cuenta todas las variables socioeconómicas pertinentes, es una manera más eficaz de estratificar los hogares en un número de conglomerados para que cada uno de éstos represente un grupo socioeconómico distinto de la población agrícola. Ello es importante en la medida en que las recomendaciones sobre las estrategias de manejo de las plagas no serán necesariamente adecuadas para todos los hogares agrícolas. Una clasificación inicial de los agricultores en conglomerados ayuda a establecer una base para elegir diferentes tipos de agricultores que participen en la exploración de una serie de estrategias de manejo de plagas y permite también centrarse en las características específicas de los conglomerados para poder investigar las interacciones entre esas características y las estrategias recomendadas. Puede verse una ilustración en Orr y Jere (1999).

21. Para hacer un análisis por conglomerados es preciso tomar dos decisiones. Primero hay que determinar si tomar una medida de semejanza, o de distancia, entre las unidades. La primera utiliza la información de distintas variables para ofrecer un valor numérico que refleje el grado de "proximidad" entre cada par de unidades. La de distancia es lo contrario, y refleja el grado de alejamiento entre los distintos pares de unidades. Cuando todas las variables son cuantitativas, o incluyen como mucho un reducido número de variables categóricas *ordenadas*, puede ser adecuado usar una matriz de distancia euclidiana[1]. No obstante, los datos de las encuestas suelen incluir variables categóricas binarias y no ordenadas. Para esos datos se han propuesto varias medidas de semejanza. Por ejemplo, si debe obtenerse una medida de semejanza entre dos variables binarias, los datos podrían someterse a un proceso de tabulación cruzada mediante esas dos variables para obtener el cuadro 2 × 2 siguiente:

[1] La distancia euclidiana puede concebirse sencillamente como reflejo del significado normal de "distancia" aplicado a un espacio multidimensional.

	0	1
0	*a*	*b*
1	*c*	*d*

22. Así pues, una posible medida de semejanza es $(a + d)/(a + b + c + d)$, que se conoce como coeficiente de concordancia simple. Otra es el coeficiente de Jaccard $d/(b + c + d)$. Pueden encontrarse más medidas en Krzanowski y Marriott (1994b). Véase en Gower (1971) una medida de similaridad válida cuando se utilizan tipos de datos mixtos. En la práctica, si deben utilizarse en el conglomerado numerosas variables de diferentes tipos, quizá sea mejor realizar diferentes análisis por conglomerados considerando variables del mismo tipo cada vez, y luego determinar si los diferentes tipos de conglomerados resultantes son similares. De esa manera se obtiene una validación cruzada de los componentes del conglomerado.

23. Una vez que se ha determinado una medida de distancia o de semejanza, hay que decidir el método de conglomeración. Los programas informáticos presentan numerosas opciones. SPSS (2001) ofrece siete (entre ellas, vinculación entre grupos, vinculación dentro del grupo, vecino más próximo, etcétera). En algunos casos se trata de procedimientos de conglomeración donde inicialmente las n unidades que se conglomeran forman n conglomerados con un miembro por conglomerado, y luego éstos se combinan de forma secuencial teniendo en cuenta su semejanza con los miembros de otros conglomerados. Otra posibilidad es un proceso de división en virtud del cual todas las n unidades comienzan como un único conglomerado, que luego se divide de forma secuencial hasta encontrar una solución satisfactoria. En uno y otro caso hay que proceder con cuidado al tomar la decisión sobre la manera de formar los conglomerados. Puede encontrarse un examen detenido de estos temas en Everitt, Landau y Leese (2001).

Ejemplo 2

24. Un caso especial es el que se plantea cuando todas las variables son binarias. El procedimiento puede ser bastante sencillo si se utiliza la conglomeración jerárquica. Como ejemplo usaremos sólo algunas observaciones de una pequeña encuesta sobre 74 agricultores en un programa de investigación realizado en explotaciones agrícolas. En el cuadro XVIII.2 pueden encontrarse algunas variables referentes a sólo ocho agricultores registradas durante las visitas. Las variables corresponden a las respuesta sí (+) y no (–). Uno de los objetivos era investigar si las explotaciones pueden agruparse en un reducido número de conglomerados teniendo en cuenta esas características.

25. Una vez más, por razones de claridad y simplificación, consideremos la formación de la matriz de semejanza que utiliza el número de + que cualquiera de las dos variables tienen en común. Los resultados pueden verse en el cuadro XVIII.3. Pueden formarse varios conglomerados considerando inicialmente las ocho explotaciones como si constituyeran ocho conglomerados, y luego agrupando los conglomerados más próximos sucesivamente hasta que todas las explotaciones se incluyan dentro de un único conglomerado.

Cuadro XVIII.2
Datos sobre las explotaciones en que se aprecia la presencia o ausencia de diversas características

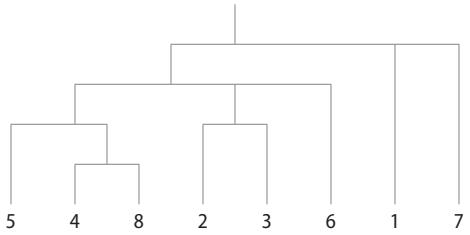
	Explotación (agricultor)							
Características	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
¿Tierras altas (+)/tierras bajas (–)?	–	+	+	+	+	+	–	+
¿Precipitaciones abundantes?	–	+	+	+	+	–	–	+
¿Ingresos elevados?	–	+	+	–	–	+	–	–
¿Hogar numeroso (>10 miembros)?	–	+	+	+	–	+	–	+
¿Acceso a leña a menos de 2 km?	+	–	–	+	+	–	+	+
¿Servicios de salud a menos de 10 km?	+	–	–	–	–	–	–	–
¿Presidido por una mujer?	+	–	–	–	–	–	–	–
¿Agua corriente?	–	–	–	–	–	–	+	–
¿Hay letrinas en la explotación?	+	–	–	–	–	+	–	–
¿Se cultiva maíz?	+	–	–	+	+	–	+	+
¿Se cultivan guisantes?	–	+	+	+	–	+	–	–
¿Se cultivan frijoles?	–	–	–	+	+	–	–	+
¿Se cultiva maní?	–	–	–	–	–	–	–	+
¿Se cultiva sorgo?	+	–	–	–	–	–	–	–
¿Tiene ganado?	+	+	–	–	+	–	+	–

Cuadro XVIII.3
Matriz de semejanzas entre ocho explotaciones agrícolas

		Explotación							
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Explotación	1	–	1	0	2	3	1	3	2
	2		–	5	4	3	4	1	3
	3			–	4	2	4	0	3
	4				–	5	3	2	6
	5					–	1	3	5
	6						–	0	2
	7							–	2
	8								–

26. La matriz de semejanza del ejemplo anterior se representa en forma gráfica en la figura XVIII.2. En este diagrama, conocido como dendograma, se observa cómo puede seleccionarse un número especificado de conglomerados cortando el árbol horizontalmente en cualquier punto. Por ejemplo, un corte horizontal cerca de la parte superior del árbol dará lugar a tres conglomerados, formados con los conjuntos (1), (7) y (2, 3, 4, 5, 6, 8). En la mayor parte de las situaciones reales se realizan juicios subjetivos para determinar el número de conglomerados que deben formarse a partir de una clasificación jerárquica. En Everitt, Landau y Leese (2001) se describen métodos formales para considerar este tema.

Figura XVIII.2
Dendograma formado por la matriz de semejanza entre explotaciones



27. Si se dispone de programas informáticos adecuados, el análisis de conglomerados puede realizarse fácilmente, aunque antes deberían tenerse en cuenta los tipos de datos utilizados, la medida de semejanza o distancia y el método empleado para conseguir los conglomerados. Deberá ponerse especial esmero si el programa permite utilizar sólo un tipo de datos para formar conglomerados. Por ejemplo, el SPSS (2001) exige que todas las variables utilizadas sean continuas, categóricas o binarias. Si hay una combinación de tipos de datos, sería mejor convertir todas las variables a puntuaciones binarias y utilizar una medida de semejanza adecuada a las variables binarias, aunque ello suponga pérdida de información.

28. Deben tenerse en cuenta otras dos cuestiones. La primera se refiere a la necesidad de considerar que (por lo que podemos saber) se desconoce la repercusión de los diseños muestrales complejos en el análisis de conglomerados. Si el diseño de la encuesta implicaba un procedimiento de muestreo de conglomerados, y había considerables diferencias entre los conglomerados incluidos en la muestra, el análisis de conglomerados aplicado a todo el conjunto de datos de la muestra sin tener en cuenta las ponderaciones de muestreo podría generar los conglomerados de diseño de la muestra. Por ello sería conveniente considerar la utilización de análisis de conglomerados con cada uno de los conglomerados de diseño de la muestra y estudiar la compatibilidad de los resultados. Una vez más, debe prestarse atención a las diferentes ponderaciones de muestreo dentro de los conglomerados de la encuesta, y los resultados deberían interpretarse con prudencia si el programa informático no puede tener en cuenta las ponderaciones.

29. La segunda cuestión se refiere a la posibilidad de que surjan dificultades de cálculo debido a limitaciones de la memoria del sistema. Esta circunstancia puede darse si el análisis de conglomerados se realiza utilizando toda la muestra de la encuesta. Si está en consonancia con los objetivos de realizar un análisis de conglomerados, el análisis puede limitarse a grupos más pequeños de la muestra con el fin de paliar ese problema.

F. Análisis de componentes principales (ACP)

30. Supongamos que hay cierto número de variables, pongamos por caso 12, que miden los distintos elementos de un tema importante de la encuesta. Por ejemplo, en una encuesta de nutrición, el estado de nutrición de los niños puede medirse en función de varios indicadores antropométricos, así como de variables que describan las características so-

cioeconómicas de sus familias. Es probable que estas variables estén correlacionadas, y entonces se plantea la cuestión de si podrían reducirse de alguna manera a un número menor de variables que reproduzcan en la medida de lo posible la variación existente en el conjunto original de datos. Eso es lo que intenta hacer el análisis de componentes principales (ACP). Esta técnica es estrictamente aplicable a un conjunto de mediciones que o son cuantitativas o tienen una escala ordinal. No obstante, como se trata en buena medida de una técnica descriptiva, no es probable que la inclusión de variables binarias y/o un número menor de variables categóricas nominales tenga consecuencias prácticas.

31. En el análisis de componentes principales se crea un nuevo conjunto de variables como combinaciones lineales[2] del conjunto original. La combinación lineal que explica la parte más considerable de la variación se conoce como primer componente principal. Luego se crea un segundo componente principal (otra combinación lineal), independiente del primero, que explica, en la medida de los posible, la variabilidad restante. Luego se van creando sucesivamente otros componentes, cada uno de ellos independiente de los anteriores. Si los tres primeros componentes explican una parte sustancial —el 90%, por ejemplo— de la variabilidad entre el conjunto original de 12 variables, el número de variables que en definitiva se deben analizar se ha reducido de 12 a 3.

[2] Si $X_1, X_2, \ldots, X_p$ son el conjunto original de variables, una variable Y formada a partir de una combinación lineal de éstas adopta la forma $Y = a_1X_1 + a_2X_2 + \ldots + a_pX_p$ donde las a_i ($i = 1, 2, \ldots, p$) son números, es decir, los coeficientes del componente principal.

32. Es importante observar que los estimadores del componente principal pueden sufrir un grave sesgo si el ACP se aplica a toda la muestra de la encuesta cuando no es autoponderada (Skinner, Holmes y Smith, 1986). Como se insiste en el apartado B, el ACP se recomienda en general en el análisis de datos de encuestas sólo para subconjuntos menores de la muestra que tengan (al menos aproximadamente) las mismas ponderaciones de muestreo. Si el subconjunto de datos pertinente tiene ponderaciones de muestreo muy diferentes, habría que proceder con cautela al interpretar los resultados.

Ejemplo 3

33. Pomeroy y otros (1997) aplicaron el ACP a los datos de una encuesta de 200 hogares en la que se había pedido a los informantes que calificaran en una escala de 1 a 15 los 10 indicadores que se les presentaron en forma de peldaños de escalera, como modo de reflejar su opinión sobre los cambios que habían tenido lugar en su área como consecuencia de varios proyectos de ordenación de recursos costeros de base comunitaria. Los indicadores se enumeran más adelante. Los resultados del ACP pueden verse en el cuadro XVIII.4.

El primer componente principal sería, por lo tanto:

CP1 = 0,24 (hogar) + 0,39 (recursos)... + 0,82 (observancia) + 0,38 (capturas)

Cuadro XVIII.4
Resultados de un análisis de componentes principales

	Componente		
Variable	**CP1**	**CP2**	**CP3**
1. Bienestar general del hogar	0,24	0,11	0,90
2. Situación general de los recursos pesqueros	0,39	0,63	0,02
3. Ingresos locales	0,34	0,51	0,55
4. Acceso a los recursos pesqueros	–0,25	0,72	0,17
5. Control de los recursos	0,57	0,40	0,12
6. Capacidad de participar en los asuntos comunitarios	0,77	0,13	0,29
7. Capacidad de influir en los asuntos comunitarios	0,75	0,22	0,34
8. Conflictos comunitarios	0,78	0,03	0,18
9. Cumplimiento de los compromisos por la comunidad y ordenación de recursos	0,82	0,12	0,07
10. Volumen de recursos capturados tradicionalmente en el agua	0,38	0,66	0,12
Porcentaje de la varianza explicado	33	19	14

34. Este primer componente se describe en Pomeroy y otros (1997) como indicador relacionado con el comportamiento de los miembros de la comunidad; el segundo componente, como asociado con los recursos pesqueros, y el tercero, como indicador de bienestar del hogar. Luego utilizan estos componentes como variables dependientes en análisis de regresión múltiple para investigar la validez de una serie de factores para explicar la variabilidad de cada indicador.

35. Aunque la interpretación de las variables es razonable en este caso, cabe preguntarse qué valor tiene, por ejemplo, utilizar el primer componente principal en la forma calculada más arriba para ulteriores análisis. Sólo las variables 5, 6, 7, 8 y 9 describen el comportamiento de los miembros de la comunidad y esas son las variables que merecen mejores calificaciones con respecto al componente principal 1 (CP1). Más que incluir las 10 variables en el cálculo del primer componente principal, sería mejor volver a calcular una variable como resumen simple de las variables de comportamiento en el conjunto de datos original adoptando, por ejemplo, una media aritmética simple de las variables 5, 6, 7, 8 y 9 o un promedio ponderado de éstas en el que el control de los recursos (variable 5) reciba una ponderación ligeramente inferior con respecto a las otras. De la misma manera, las variables relativas a los recursos (variables 2, 3, 4 y 10) podrían combinarse en una sola, mientras que la variable 1 se mantendría por cuenta propia. Si se utiliza de esta manera, el ACP indica cómo pueden resumirse los 10 indicadores de forma sencilla para obtener un nuevo conjunto de medidas significativas para ulteriores análisis, como han hecho, por ejemplo, Pomeroy y otros (1997) mediante el análisis de regresión para explorar factores que influyen en cada uno de sus tres componentes principales.

Ejemplo 4

36. Otro ejemplo práctico es el marco de medios de subsistencia sostenibles adoptado por el Departamento para el Desarrollo Internacional (DFID) del Gobierno del Reino Unido de Gran Bretaña e Irlanda del Norte. En él se consideran cinco activos: el capital social, el capital humano, el capital natural, el capital físico y el capital financiero. Una encuesta realizada para estudiar los medios de vida de los hogares supondría una medición de cada uno de estos activos de acuerdo con algunas variables subsidiarias. Por ejemplo, el capital social puede medirse en función de la dependencia de redes de apoyo, el porcentaje de ingresos del hogar obtenido a través de remesas, la confianza en el grupo, el grado de participación en la toma de decisiones, etcétera; el capital humano puede medirse en función del nivel de instrucción, el estado de salud, etcétera, y el capital físico, teniendo en cuenta la propiedad de una bicicleta o radio, la disponibilidad de agua corriente, electricidad, etcétera.

37. El objetivo es determinar una sola variable; una para cada uno de los cinco activos. Ello puede hacerse de forma sencilla en el caso de los activos físicos, por ejemplo, obteniendo una media ponderada simple de las respuestas binarias correspondientes a si un hogar posee o no los distintos artículos de una lista determinada, utilizando los precios como ponderaciones. Por el contrario, el capital social no puede combinarse de forma tan sencilla ya que es mucho más difícil asignar ponderaciones a variables que describan los activos sociales. Quizá tengamos que aceptar ponderaciones derivadas de datos a través de un ACP aplicado a un conjunto de variables sociales. Los resultados pueden utilizarse para obtener una medida global adecuada del capital social, también en este caso proponiendo una media ponderada simple una vez que se conozcan las ponderaciones relativas de cada variable en el primero o los dos primeros componentes principales.

G. Métodos con múltiples variables en la creación de índices

38. La creación de índices puede tener varios significados diferentes. En un estudio de salud, por ejemplo, el estado nutricional de los niños suele medirse elaborando índices a

partir de indicadores antropométricos, como el peso por edad, la altura por edad y el peso por altura, que representan la insuficiencia ponderal, el retraso del crecimiento y la emaciación, respectivamente.

39. En un examen más complejo, puede obtenerse un índice de alimentación infantil a partir de las respuestas proporcionadas sobre la lactancia materna, el uso de biberones, la diversidad dietética, el número de veces en que el niño ha recibido determinados grupos de alimentos en los siete últimos días o la frecuencia de la alimentación (Ruel y Menon, 2002). Éste es un segundo tipo de índice en que el investigador decide sobre las puntuaciones específicas que deberán asignarse, procurando que la escala ordinal de cada variable sea tal que los valores representen siempre o "bueno" o "malo". Cuando en la propiedad de distintos activos, por ejemplo, intervienen variables binarias, el precio del activo podría utilizarse para ofrecer ponderaciones diferentes de cada partida, como se ve en el ejemplo 4 (apartado F), *supra*.

40. Otro tipo de índice es el que se establece en el caso en que una encuesta intente determinar actitudes u opiniones sobre la calidad del acceso a los servicios de salud, por ejemplo. En este caso pueden formularse varias preguntas, que exigen respuestas en una escala de puntuación del 1 al 5, en el que 1 sería "muy malo" y 5 "muy bueno". También en este caso las puntuaciones resultantes podrían resumirse entre las diferentes preguntas pertinentes para obtener un índice que reflejara las opiniones de los hogares sobre el valor de los servicios de salud.

41. Nuestro examen va más allá, para considerar las situaciones en que los datos determinan la forma del índice mediante el uso de un procedimiento con múltiples variables. Con ello se conserva todavía la interpretación común de un índice como valor único que refleja la información de distintas variables en una medida compuesta, que normalmente adopta la forma siguiente:

$$\text{Índice} = a_1X_1 + a_2X_2 + a_3X_3 + \ldots\ldots\ldots + a_pX_p$$

donde los términos a_i son ponderaciones que deben determinarse a partir de los datos, y los términos X_1 son un conjunto adecuado de variables p medidas en la encuesta. Presentamos dos formas de determinar las ponderaciones a_i a partir de los datos (véase *infra*). Los objetivos de la creación del índice son, por lo general, los que determinan cuál de ellas es la más adecuada.

42. La primera está basada en un planteamiento de modelos de regresión; la segunda, en una aplicación del ACP. Aquí se presentan en relación con los índices utilizados para medir los indicadores sustitutivos de la riqueza del hogar o el estado socioeconómico en los países en desarrollo. Son muchas las publicaciones sobre este tema, y puede encontrarse una exposición general en Davis (2002). Véase también el capítulo XVII, en el que se realiza un análisis interesante sobre el uso de los datos de encuestas de hogares para comprender la pobreza.

1. Elaboración de modelos de gastos de consumo para la creación de una variable sustitutiva del ingreso

43. Un procedimiento para elaborar modelos del gasto de consumo como variable sustitutiva del ingreso es el presentado por Henstchel y otros (2000) y Elbers, Lanjouw y Lanjouw (2001). Supone la utilización de datos de una detallada encuesta sobre el presupuesto de los hogares para identificar variables indicativas de la pobreza. Para ello se utiliza el gasto de consumo como variable dependiente en un modelo de regresión lineal múltiple y una serie de variables relativas a los hogares (por ejemplo, activos poseídos por el hogar, calidad de la vivienda, acceso a servicios, etcétera) como posibles variables explicativas (predictivas) del modelo. El mejor subconjunto pequeño de las variables explicativas que expli-

que la variación máxima de la variable (dependiente) de la respuesta se utiliza para prever el gasto de consumo. Si las variables explicativas se han recopilado en un censo de población, la ecuación del modelo resultante puede aplicarse luego a los datos del censo para prever los gastos de consumo de cada hogar del censo. A partir de ahí podrán construirse luego mapas de pobreza de alcance nacional. Si la encuesta de presupuesto de hogares se realiza mucho antes de la fecha prevista del censo, el conjunto adecuado de valores predictivos puede identificarse a partir de los datos de la encuesta sobre el presupuesto. Presentamos a continuación un ejemplo para ilustrar este planteamiento.

Ejemplo 5

44. La Oficina Nacional de Estadística de la República Unida de Tanzanía realizó en 2000-2001 una encuesta nacional sobre el presupuesto de los hogares (HBS) en la que se incluyeron unos 22.000 hogares. Tomando como base los datos recopilados durante 28 días, se calculó en cada hogar el gasto total de consumo por equivalente de adulto durante esos 28 días. La elaboración de modelos de regresión con los datos preliminares disponibles de la HBS permitió identificar una serie de posibles variables de los hogares (conjuntos independientes para las zonas urbanas y rurales) que explicaban una buena parte de la variabilidad del gasto de consumo. Estas variables se incluyeron en un cuestionario administrado a un censo de hogares en tres puntos de vigilancia supervisados por el equipo del proyecto de morbilidad y mortalidad de adultos (AMMP), con sede en Dar es Salaam. El objetivo era establecer un índice que reflejara el gasto de consumo utilizando datos de la HBS para cada sitio del AMMP y aplicar el índice a los hogares considerados por el AMMP en cada uno de los sitios.

45. Puede verse toda la información sobre la elaboración de modelos y una evaluación de la eficacia de éstos en Abeyasekera y Ward (2002). Aquí presentamos un resumen de los resultados en una de las regiones rurales (véase el cuadro XVIII.5), para poner de manifiesto las variables que se introdujeron en la ecuación del modelo y las ponderaciones (coeficientes de regresión) utilizadas para calcular un índice del gasto de consumo.

Cuadro XVIII.5
Variables utilizadas y sus ponderaciones correspondientes en la creación de un índice predictivo del gasto de consumo para la región del Kilimanjaro en la República Unida de Tanzanía

Variable predictiva	Probabilidad de significación estadística	Ponderación (coeficiente del modelo) (estimación de STATA)
Tamaño del hogar	0,000	–0,239
Cuadrado del tamaño del hogar	0,000	0,0104
Edad del jefe del hogar (años)	0,038	–0,00226
Nivel de instrucción del jefe del hogar[a]	0,000	0, –0,00920, 0,159, 0,280
Fuente principal de ingresos[b]	0,017	0, ,0983, 0,1345, 0,0761
Días en que se comió carne en la última semana	0,000	0,0888
Superficie de tierra perteneciente al hogar	0,000	0,0258
Fertilizante[c]	0,000	0,1625
Semillas[c]	0,004	0,1186
Posesión de bicicleta	0,000	0,1287
Posesión de sofá	0,000	0,1995
Posesión de lámpara	0,001	0,1104
Constante en la ecuación del modelo	0,000	9,794

Tamaño de la muestra = 1.026.

R^2 = 0,651; R^2 ajustado = 0,646

[a] Ninguna; primaria; secundaria; terciaria y superior.

[b] Venta de cultivos, venta de ganado; negocios/sueldos/salarios; otras fuentes.

[c] Si se compraron en los 12 últimos meses.

46. De acuerdo con los datos del cuadro XVIII.5, el índice que prevé el gasto de consumo para los hogares en la región del Kilimanjaro de la República Unida de Tanzanía sería:

Índice de gasto de consumo =

= 9,79388 + (0,11043 * lámpara) + (0,19950 * sofá) + (0,12870 * bicicleta)+

+ (0,11858 * semillas) +(0,16254 * fertilizante) + (0,025824 * superficie tierra) +

+ (0,088769 * carne) + (0,076132 * ingreso4) + (0,13451 * ingreso3) +

+ (0,098303 * ingreso2) + (0,27985 * edu4) + (0,15878 * edu3) –(0,0091977 * edu2) –

– (0,0022552 ingreso2 * edad) + (0,010456 * tamaño hogar2) – (0,23902 * tamaño hogar)

47. El modelo explicaba el 65% de la variabilidad del gasto de consumo. Se trata de una cifra bastante elevada, dada la complejidad de los medios de subsistencia entre los hogares rurales. La calidad de este índice en su fase de elaboración se juzgaba con dos métodos distintos: *a*) comparándolo con valores verdaderos del gasto de consumo, y *b*) considerando su capacidad de identificar la verdadera proporción de hogares por debajo del umbral de pobreza de la República Unida de Tanzanía. El método *a*), en el que el índice se comparaba gráficamente con los valores verdaderos, revelaba una correspondencia muy aceptable. Los resultados no eran tan buenos cuando la población de los valores verdaderos y la población de los valores previstos se clasificaban en cinco quintiles de riqueza y se comparaban mutuamente en un cuadro. Sólo el 46% de los hogares se clasificaban en el quintil correcto. La clasificación en virtud del umbral de pobreza era más acertada: el 87% se clasificaba correctamente por encima o por debajo de ese umbral.

48. En los apartados finales del capítulo XIX se presentan otros ejemplos del enfoque de elaboración de modelos.

2. Análisis de componentes principales (ACP) utilizado para crear un índice "de riqueza"

49. La metodología examinada en el subapartado G.1 puede aplicarse sólo si se tienen datos fiables sobre el gasto de consumo —la variable dependiente— procedentes de una encuesta anterior. La dificultad de recopilar información fiable sobre el gasto de consumo, junto con los altos costos de la recopilación de datos, ha llevado a algunos investigadores a recomendar el uso de un índice de pobreza basado en los activos, derivado de la realización de un ACP. El primer componente principal sirve de indicio de la condición socioeconómica a raíz de investigaciones anteriores que han puesto de manifiesto la estrecha relación entre activos y consumo (Filmer y Pritchett, 1998). No obstante, debe procederse con cautela al interpretar el índice de activos como indicador de la pobreza, pues su eficacia dependerá de la elección de los activos utilizados y del conjunto concreto de datos a que se aplique el ACP. Un ejemplo de este planteamiento es el de Gwatking y otros (2000), que ilustra la metodología del ADP para determinar los quintiles de riqueza en la República Unida de Tanzania, utilizando las siguientes variables basadas en los activos y relacionadas con la salud:

- Disponibilidad de electricidad, radio, televisión, refrigerador, bicicleta, motocicleta, automóvil (en todos los casos, 1 = Sí, 0 = No);
- El número de personas por dormitorio (respuesta cuantitativa);
- Principales fuentes de agua potable para el hogar (siete categorías);
- Principal tipo de aseo utilizado por los miembros del hogar (cinco categorías);
- Principal tipo de material de recubrimiento de suelos del hogar (seis categorías).

Cuadro XVIII.6
Límites para dividir a la población en cinco quintiles de riqueza

Quintil de riqueza	Dar es Salaam[a] (HBS)	Kilimanjaro[a] (HBS)	Morogoro[a] (HBS)	Toda la República Unida de Tanzanía[a] (HBS)	Toda la República Unida de Tanzanía[b] (DHS)
20º percentil	–1,2993	–0,8452	–0,9190	–1,0317	–0,5854
40º percentil	–0,7709	–0,6289	–0,6180	–0,5704	–0,5043
60º percentil	–0,1054	–0,2459	–0,3645	–0,3051	–0,3329
80º percentil	1,1603	0,3239	0,4586	0,4609	0,3761

[a] Encuesta sobre el presupuesto de hogares, 2000-2001.
[b] Encuesta demográfica y de salud, 1996.

50. Los datos que ellos utilizaron proceden de la información recopilada a través del cuestionario de la encuesta demográfica y de salud (DHS). En el análisis se utilizaron las debidas ponderaciones de muestreo.

51. Los autores remarcaron que se trataba sólo de un esfuerzo inicial aplicado a la muestra de todo el país, pero que los intentos futuros de examinar las diferencias de la población en función de la clase socioeconómica arrojarían resultados diferentes y que ello podría deberse a la utilización de una base diferente de los activos para definir la condición socioeconómica, errores de muestreo u otras razones. Una razón más obvia serían las diferencias de riqueza en cada lugar. De hecho, había pruebas de diferencias en la línea de separación de los quintiles de riqueza cuando se aplicaba su metodología a tres subpoblaciones de la República Unida de Tanzania —las tres regiones mencionadas en el subapartado G.2— utilizando datos de la encuesta sobre el presupuesto de hogares (véase el cuadro XVIII.6 arriba). Por ello, los resultados del ACP no son transferibles a otros contextos ni siquiera dentro de un mismo país a lo largo del tiempo o cuando se aplican a diferentes estratos de la población.

52. Los investigadores han utilizado también el primer componente principal de un análisis de componentes principales como índice sinóptico para ulteriores análisis de los datos. Ruel y Menon (2002), por ejemplo, construyeron un índice socioeconómico y de los conjuntos de datos de la DHS con el fin de clasificar los hogares en terciles y controlar así la situación socioeconómica en un análisis de regresión múltiple realizado para determinar los factores que influyen en el estado nutricional de los niños. Trataron de separar los análisis correspondientes a las poblaciones rurales y urbanas utilizando siete conjuntos de datos de cinco países de América Latina. Las variables utilizadas eran la fuente de abastecimiento de agua, el saneamiento, los materiales de la vivienda (suelos, paredes, tejados) y la propiedad de una serie de activos. Los valores de estas variables se clasificaron por orden ascendente (de peor a mejor) antes de someter las variables a un análisis de componentes principales. Sólo se conservaron las variables con coeficientes de componentes principales superiores a 0,5. Este planteamiento resultó razonable, ya que el objetivo primario era la creación de un índice para corregir las diferencias socioeconómicas en un análisis posterior final.

H. Conclusiones

53. Nuestro objetivo en este capítulo ha sido explicar el uso de los métodos con múltiples variables en la creación de índices, con particular insistencia en la necesidad de instrumentos de exploración con variables múltiples como primera etapa del análisis. No obstante, la aplicación de estos métodos requiere cuidadosa atención, en particular, a su significado y sus limitaciones. La eficacia del ACP en la reducción de variables, por ejemplo, depende de que se pueda resumir una proporción considerable de la variación de los datos mediante un reducido número de índices parciales, y hacer una interpretación significativa de cada uno de ellos. Conviene también reflexionar atentamente acerca de la eficacia del procedimiento ACP si el primer componente principal explica sólo una pequeña parte de la variación en el conjunto completo de variables. Asimismo, debería tenerse en cuenta la idoneidad de las variables incluidas en el cálculo del índice en relación con los objetivos del análisis.

54. El análisis de conglomerados debe superar las dificultades asociadas con la identificación de una medida idónea de semejanza o de distancia y con las decisiones referentes al método de conglomeración que se debe utilizar. En este contexto deben tenerse en cuenta varios factores, entre ellos los tipos de datos utilizados, algunos aspectos relativos al cálculo y la solidez del procedimiento cuando se producen pequeños cambios en los datos.

55. Hay que insistir también una vez más en que los métodos descritos en este capítulo se aplican mejor a los conjuntos adecuados de la población cuando hay una estructura clara para dividir esa población. Así ocurre especialmente si los datos para el análisis proceden de una encuesta nacional. Es entonces cuando deben adoptarse, y justificarse debidamente, las decisiones referentes a la elección de los subconjuntos. Una consecuencia es que pueden producirse índices diferentes para los distintos subconjuntos. De todas formas, ello representa en cuanto tal un hallazgo útil, lo que indica que los análisis ulteriores serían más significativos dentro de los subconjuntos de población sometidos a consideración.

56. En este capítulo se ha realizado una evaluación de las técnicas con variables múltiples, en cuanto instrumento de exploración y, más en concreto, para la creación de índices. Actualmente se dispone de programas informáticos de estadística general —por ejemplo, SPSS (2001), STATA (2003) y otros— que permiten la realización de dichos análisis con relativa facilidad. Por ello, se alienta a los investigadores a que consideren la posibilidad de utilizarlos durante el análisis de los datos de encuestas con el fin de extraer la mayor información posible de los datos y contribuir provechosamente a los objetivos de la encuesta.

Reconocimientos

Deseo expresar mi sincero agradecimiento a mi colega Ian Wilson y a dos árbitros científicos por sus valiosas observaciones sobre los borradores iniciales de este capítulo. Deseo también agradecer a la Oficina Nacional de Estadística de la República Unida de Tanzanía su autorización para consultar los datos de algunos de los ejemplos utilizados en este capítulo, y al Departamento de Desarrollo Internacional (DFID) del Gobierno del Reino Unido de Gran Bretaña e Irlanda del Norte, que aportó ideas para este capítulo mediante el financiamiento de muchos proyectos interesantes relacionados con encuestas en el mundo en desarrollo. No obstante, el material de este capítulo es responsabilidad exclusiva del autor y no representa en absoluto la opinión del DFID.

REFERENCIAS BIBLIOGRÁFICAS

Abeyasekera, S. y P. Ward (2002). *Models for Predicting Expenditure per Adult Equivalent for AMMP Sentinel Surveillance Sites*. Dar es Salaam: Adult Morbidity and Mortality, Ministry of Health of the United Republic of Tanzania. Disponible en www.ncl.ac.uk/ammp/tools_methods/socio.html.

Chatfield, C. y A. J. Collins (1980). *Introduction to Multivariate Analysis*. Londres: Chapman and Hall.

Davis, B. (2002). Is it possible to avoid a lemon? Reflections on choosing a poverty mapping method. Disponible en www.povertymap.net/pub/Pov_mapping_methods_18-9-02.pdf.

Elbers, C., J. Lanjouw y P. Lanjouw (2001). Welfare in aldeas and towns: micro-level estimation of poverty and inequality. Mimeografiado. Universidades de Vrije y Yale y Banco Mundial.

Everitt, B. S. y G. Dunn (2001). *Applied Multivariate Data Analysis*. Londres: Arnold.

Everitt, B. S., S., Landau and M. Leese (2001). *Cluster Analysis*. Londres: Arnold.

Filmer, D. y L. Pritchett (1998). *Estimating Wealth Effects without Expenditure Data–or Tears: An Application to Educational Enrolments in States of India*. Washington, D.C.: World Bank Policy Research Working Paper, No. 1994.

GenStat (2002). *GenStat for Windows,* 6a. ed. Oxford, Reino Unido: VSN International, Ltd.

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics,* vol. 27, págs. 857-872.

Gwatkin, D. R. y otros (2000). *Socio-economic Differences in Health, Nutrition and Population in Tanzania*. Washington, D.C.: Thematic Group on Health, Population, Nutrition and Poverty of the World Bank. Disponible en ww.worldbank.orgpoverty/healthdatatranzaniatanzania.pdf (consultado el 30 de junio de 2004).

Hentschel, J. y otros (2000). Combining census and survey data to trace spatial dimensions of poverty: a case study of Ecuador. *The World Bank Economic Review*, vol. 14, No. 1, págs. 147-165.

Jongman, R. H. G., C. J. F. Ter Braak y O. F. R. Van Tongeren (1995). *Data Analysis in Community and Landscape Ecology*. Cambridge, Reino Unido: Cambridge University Press.

Korn, E.L. y B.I. Graubard (1998). Scatterplots with survey data. *The American Statistician*, vol. 52, No. 1.

Krzanowski, W.J. y F.H.C. Marriott (1994a). *Multivariate Analysis,* Part 1. *Distributions, Ordination and Inference*. Londres: Arnold.

_____ (1994b). *Multivariate Analysis,* Part 2. *Classification, covariance structures and repeated measurements*. Londres: Arnold.

Manly, B. F. J. (1994). *Multivariate Statistical Methods*: *A Primer*. 2a. ed. Londres: Chapman and Hall.

Orr, A. y P. Jere (1999). Identifying smallholder target groups for IPM in southern Malawi. *International Journal of Pest Management*, vol. 45, No. 3, págs. 179-187.

Pomeroy, R. S. y otros (1997). Evaluating factors contributing to the success of community-based coastal resource management: the Central Visayas Region Project-1, Filipinas. *Ocean and Coastal Management,* vol. 36, Nos. 1-3, pág. 24.

Ruel, M. T. y otros (1999). *Good Care Practices Can Mitigate the Negative Effects of Poverty and Low Maternal Schooling on Children's Nutritional Status: Evidence from Accra*. Food Consumption and Nutrition Division Discussion Paper, No. 62, Washington, D.C.: Instituto Internacional de Investigaciones sobre Política Alimentaria.

Ruel, M. T. y P. Menon (2002). *Creating a Child Feeding Index Using the Demographic and Health Surveys: an Example from Latin America*. Food Consumption and Nutrition Division Discussion Paper, No. 130, Washington, D.C.: Instituto Internacional de Investigaciones sobre Política Alimentaria.

Sahn, D. E. y D. Stifel (2000). Assets as a measure of household welfare in developing countries. Working Paper 00-11. St. Louis, Missouri: Universidad de Washington, Centro para el Desarrollo Social.

SAS (2001). *SAS Release 8.2.* Cary, North Carolina: SAS Institute, Inc., SAS Publishing.

Sharma, S. (1996). *Applied Multivariate Techniques*. Nueva York: Wiley and Sons, Inc.

Skinner, C. J., D. J. Holmes y T. M. F. Smith (1986). The effect of sample design on principal component analysis. *Journal of the American Statistical Association*, vol. 81, No. 395, págs. 789-798.

SPSS (2001). *SPSS for Windows. Release 11.0.* Chicago, Illinois: LEAD Technologies, Inc.

STATA (2003). *Intercooled Stata 8.0 for Windows*. College Station, Texas: Stata Corporation.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Chesire, Connecticut: Graphics Press.

Wilson, I. M. y R. D. Stern (2001). *Approaches to the Analysis of Survey Data*. Statistical Guideline Series supporting DFID Natural Resources Projects. Reading, Reino Unido: Statistical Services Centre, Universidad de Reading. Disponible en www.reading.ac.uk/ssc (consultado el 25 de junio de 2004).

Capítulo XIX
Análisis estadístico de los datos de las encuestas

James R. Chromy
Research Triangle Institute
Research Triangle Park, North Carolina
Estados Unidos de América

Savitri Abeyasekera
Centro de Servicios de Estadística
Universidad de Reading
Reino Unido de Gran Bretaña
e Irlanda del Norte

RESUMEN

El hecho de que los datos de las encuestas se obtengan a partir de unidades seleccionadas con diseños muestrales complejos es algo que debe tenerse en cuenta a la hora de los análisis: es preciso utilizar ponderaciones para analizar los datos de las encuestas, y las varianzas de las estimaciones de las encuestas deben calcularse de tal manera que correspondan al diseño de una muestra compleja. En este capítulo se describe la elaboración de ponderaciones y su uso para calcular las estimaciones de las encuestas y se presenta un panorama general de la estimación de la varianza para los datos de encuesta. En primer lugar se consideran lo que se denominan estimaciones "descriptivas", como los totales, las medias y las proporciones que se utilizan ampliamente en los informes de encuestas. Luego se examinan tres formas de planteamientos "analíticos" que pueden utilizarse para examinar las relaciones entre las variables de la encuesta: modelos de regresión lineal múltiple, modelos de regresión logística y modelos en varios niveles. Estos modelos constituyen un conjunto de instrumentos valiosos para analizar las relaciones entre una variable de respuesta clave y algunos otros factores. En este capítulo se presentan varios ejemplos para ilustrar el uso de estas técnicas de elaboración de modelos y se dan algunas orientaciones sobre la interpretación de los resultados.

Términos clave: Diseño de encuestas complejas, estadísticas analíticas, regresión, regresión logística, estructuras jerárquicas, elaboración de modelos en varios niveles.

A. Introducción

1. Las encuestas de hogares utilizan diseños muestrales complejos para controlar los costos de las encuestas. Los marcos muestrales completos que enumeran todos los individuos o todos los hogares son muy poco frecuentes. Incluso cuando se dispone de registros de la población, el costo de realización de una encuesta mediante entrevistas de hogares sobre la base de un simple diseño por muestreo aleatorio simple resulta prohibitivo. Las encuestas del Estudio de medición de los niveles de vida (EMNV) examinadas en el capítulo XXIII son un buen ejemplo de muchas de las características complejas de los diseños de encuestas de hogares.

2. En el cuadro XIX.1 puede verse la estructura típica de un diseño de encuestas de hogares. La mayor parte de los diseños para su utilización en encuestas de hogares son de carácter complejo e implican cierto grado de estratificación, muestreo en varias etapas y

Cuadro XIX.1
Estructura típica del diseño de las encuestas de hogares

Características	Definiciones posibles	Repercusiones
Estratos	Regiones Tipo de comunidad (urbana/rural)	Pueden reducir los errores estándar de las estimaciones. La distribución de control de la muestra puede dar lugar a muestreos desproporcionados
Unidades de muestreo de la primera etapa	Zonas de empadronamiento o áreas geográficas semejantes Aldeas en los estratos rurales	Facilitan la conglomeración de la muestra para controlar los costos Facilitan la elaboración de marcos completos de direcciones de unidades de vivienda únicamente en las zonas incluidas en la muestra Seleccionadas con probabilidad proporcional al tamaño
Unidades de muestreo de la segunda etapa	Direcciones de las unidades de vivienda	Pueden no contener un hogar o persona no relacionados, o contener uno o más Seleccionadas con igual probabilidad dentro de las unidades de muestreo de la primera etapa
Unidades de muestreo de la tercera etapa (cuando no todos los miembros del hogar se incluyen automáticamente en la muestra)	Miembros del hogar	Muestra seleccionada a partir de la lista de miembros del hogar obtenida de un miembro adulto responsable del hogar. Pueden dar lugar a ponderación desigual con el fin de tener en cuenta el número de componentes del hogar
Unidades de observación	Hogares Miembros del hogar Empresas agrícolas o comerciales administradas por los miembros del hogar Archivos especiales para algunos subgrupos; por ejemplo, adultos integrantes de la mano de obra Acontecimientos o episodios relacionados con los miembros del hogar Medidas repetidas a lo largo del tiempo (encuestas por panel)	Pueden requerir más de un archivo analítico para los análisis con fines específicos

tasas de muestreo desiguales, como se ha indicado ya. En el análisis se necesitan ponderaciones para compensar las tasas de muestreo desiguales, y los ajustes por falta de respuesta dan lugar a más ponderaciones desiguales. El diseño muestral complejo debe tenerse en cuenta al estimar la precisión de las estimaciones de las encuestas.

3. En el apartado B del presente capítulo se describe la elaboración de ponderaciones para su uso en los análisis de encuestas y el empleo de ponderaciones para la elaboración de estimaciones "descriptivas" simples, como los totales, las medias y las proporciones/porcentajes que se utilizan ampliamente en los informes de las encuestas. Se presenta también un panorama general del cálculo de la varianza para esas estimaciones sobre la base de diseños muestrales complejos.

4. Los apartados restantes se centran en tres formas de usos "analíticos" de los datos de encuestas que tratan de determinar cómo una respuesta clave o una variable dependiente —por ejemplo los resultados académicos de un niño escolarizado, el nivel de pobreza de un hogar— se ven afectadas por una serie de factores, conocidos muchas veces como variables explicativas o variables independientes. Los modelos de regresión lineal múltiple son útiles cuando la respuesta clave es una variable de medición cuantitativa, mientras que los modelos de regresión logística pueden aplicarse cuando la variable de la respuesta clave es binaria, es decir, cuando la respuesta admite sólo dos posibles valores (por ejemplo, sí/no, presente/ausente). Estos métodos de regresión pueden aplicarse a un conjunto de datos de encuesta no anidados, o a unidades de muestreo en un solo nivel de la jerarquía de un diseño en varias etapas. Por otro lado, el análisis quizá deba tener en cuenta las diferentes fuentes de variabilidad que ocurren en distintos niveles jerárquicos, y luego entran en acción los modelos en varios niveles. Este planteamiento tiene en cuenta la estructura de correlación entre las unidades de muestreo en un nivel, ya que ocurren dentro de unidades en diferentes niveles.

B. Estadísticas descriptivas: ponderaciones y estimación de la varianza

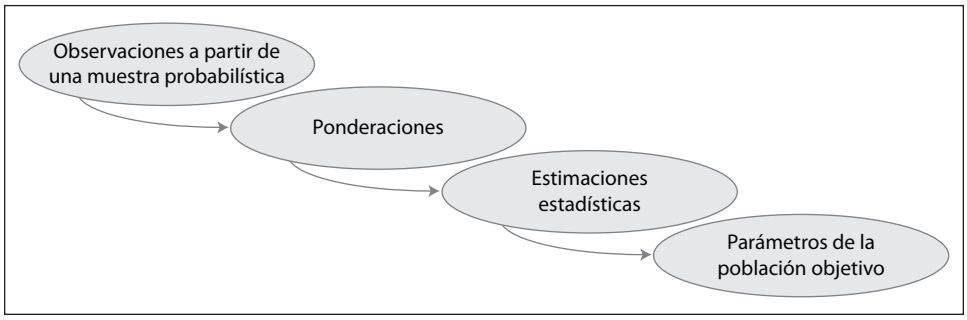
5. El objetivo de las encuestas de hogares suele ser conseguir estimaciones de los totales de la población, medias de población o coeficientes simples de los totales o las medias. Como ejemplos de totales cabría citar la población total, el número de hombres o de mujeres dentro de la población activa o el número de niños de cinco años de edad o menos. Ejemplos de medias serían el ingreso medio de la fuerza de trabajo o el ingreso medio de las mujeres o de los hombres integrantes de fuerza de trabajo. Los coeficientes podrían ser necesarios para estimar la proporción de hogares con un ingreso total inferior al umbral de pobreza o el ingreso medio de los hogares cuya principal fuente de ingresos salariales es una mujer.

6. Las encuestas de hogares producen estimaciones nacionales, pero pueden orientarse también de manera que presenten estimaciones sobre regiones geopolíticas u otros ámbitos intersectoriales. Además, pueden repetirse para obtener estimaciones periódicas (por ejemplo, anuales o quinquenales), que podrían considerarse como campos temporales. En la medida en que las estadísticas conseguidas consisten simplemente en estimaciones de totales, medias o tasas, aun cuando se elaboren para campos relacionados con la población (intersectoriales o temporales) consideramos el análisis necesario para obtener estas estimaciones como "descriptivo". Entre las estadísticas descriptivas se incluyen las propias estimaciones y cierta medida de la precisión de esas estimaciones. Entre los informes descriptivos cabe incluir los errores estándar de las estimaciones o las estimaciones de intervalo basadas en esos errores estándar. La estimación de los errores estándar requiere un análisis que tenga en cuenta el diseño muestral de la encuesta de hogares. Las estimaciones de intervalo exigen no sólo estimaciones de los errores estándar basadas en diseños adecuados sino también el conocimiento de los grados de libertad utilizados al calcular las estimaciones de los errores estándar. Estos tipos de estadísticas descriptivas bastante sencillas constituyen la mayoría de las estadísticas oficiales publicadas para describir los resultados de las encuestas de hogares.

7. Las ponderaciones de las encuestas[1] y las estimaciones estadísticas basadas en esas ponderaciones constituyen el nexo entre las observaciones a partir de una muestra probabilística de los hogares y las medidas sinópticas o parámetros de población en lo que respecta a la población de los hogares. En la figura XIX.1 se reproduce ese nexo. La población de todos los hogares algunas veces se conoce con el nombre de población objetivo o universo. Sin la aplicación de la ponderación y el muestreo probabilístico no se dispone de una teoría estadística que justifique el establecimiento de un nexo entre las observaciones de la muestra y los parámetros de la población objetivo.

8. Todo análisis que no tenga en cuenta el diseño de la muestra y las ponderaciones debe basarse en suposiciones. Si la muestra se ha concebido de manera que genere una muestra con probabilidad igual, se puede prescindir de las ponderaciones para estimar las medias,

Figura XIX.1
Aplicación de ponderaciones y estimación estadística



[1] Las ponderaciones basadas en el diseño se obtienen generalmente calculando el inverso de la probablidad de selección de las unidades de observación seleccionadas. Las ponderaciones de la encuesta ofrecidas en los archivos de análisis para las encuestas de análisis suelen ser ponderaciones basadas en el diseño que se han ajustado para tener en cuenta la falta de respuesta. Muchas veces se aplican ajustes adicionales para conseguir la posestratificación o la calibración adecuada para coincidir con totales marginales conocidos o mucho más precisos. Además, puede aplicarse cierta forma de compensación de la ponderación para limitar el efecto de ponderación desigual cuando las grandes ponderaciones se deben a acontecimientos imprevistos relacionados con el muestreo o la recopilación de datos sobre el terreno. El término "ponderaciones de la encuesta" se utiliza para distinguirlo de lo que son estrictamente "ponderaciones basadas en el diseño".

tasas o relaciones entre las variables. Kish (1965, págs. 20-21) acuñó el término "método de selección de probabilidad igual" (mspi) para referirse a estos diseños y observó que incluso las muestras complejas en varias etapas pueden concebirse como mspi para las unidades de muestreo en la etapa final del diseño, o cuando se está ya cerca de esa etapa. Como se señala más adelante, los ajustes para tener en cuenta la falta de respuesta podrían crear ponderaciones desiguales aun cuando el diseño fuera inicialmente mspi. Si la posestratificación o la calibración multidimensional se aplica a los datos mediante ajustes en las ponderaciones, estos procesos crearán casi siempre ajustes desiguales en las ponderaciones y por ello ponderaciones desiguales.

9. No obstante, algunos analistas están dispuestos a formular suposiciones que permitirían el análisis de los datos de las encuestas de hogares sin ponderaciones o con ponderaciones iguales. Esos supuestos son especialmente sostenibles cuando aplican modelos a los datos para estudiar las relaciones entre una variable dependiente y una serie de variables explicativas independientes.

10. Para el caso teórico de las encuestas con respuesta completa de todos los miembros de la muestra, el uso de ponderaciones basadas en el diseño calculadas como el inverso de cada probabilidad de selección de la unidad de observación ofrece estimaciones más sesgadas de los totales de la población y otras estadísticas lineales (Horvitz y Thompson, 1952). En la práctica, las encuestas de hogares siempre encuentran casos de falta de respuesta, lo que puede dar lugar a sesgos en las estimaciones si estas observaciones se eliminan del análisis sin tomar ninguna otra medida (véase el capítulo VIII). Se han elaborado también técnicas que tratan de reducir el sesgo debido a la falta de respuesta. El planteamiento más sencillo consiste en dividir la muestra en clases de ponderación para que dentro de estas clases las diferencias entre los parámetros de población para los informantes y los no informantes puedan considerarse mucho menores o insignificantes (Rubin, 1987). Luego se realizan ajustes del coeficiente en las ponderaciones dentro de las clases de ponderación para que cada una de éstas esté representada en las estimaciones ajustadas en la misma proporción en que lo habría estado en la muestra seleccionada.

11. El proceso de muestro probabilístico no garantiza necesariamente que la distribución de la muestra seleccionada con respecto a las características conocidas sea idéntica a la de la población total. La estratificación antes de la selección de la muestra puede garantizar que esta condición se mantenga en el caso de algunas características, pero quizá no sea posible en otras si la variable de clasificación no está disponible en el marco utilizado para seleccionar la muestra. En vez de realizar complejos ajustes del coeficiente para cada estimación producida a partir de los datos de la encuesta de hogares, muchas veces lo que se hace es incorporar la posestratificación como ajuste de la ponderación por una sola vez, que luego se aplica automáticamente a todas las estimaciones obtenidas utilizando las ponderaciones ajustadas. El planteamiento más sencillo para el ajuste de posestratificación utiliza una división de la muestra semejante a la empleada para ponderar el ajuste por falta de respuesta antes mencionado.

12. Los métodos de calibración que controlan la distribución muestral ponderada en varias dimensiones simultáneamente se utilizan algunas veces para el ajuste de la ponderación en caso de falta de respuesta, para la posestratificación o con ambos fines (Deville y Särndal, 1992; Folsom y Singh, 2000).

13. Las ponderaciones excesivamente grandes pueden sobrevalorar la varianza de las estimaciones de encuestas de hogares mediante un efecto del diseño (véanse los capítulos VI y VII). Algunas veces estas ponderaciones se reducen o compensan arbitrariamente, sobre todo si esa gran ponderación no es consecuencia de un diseño muestral planificado.

14. Las ponderaciones finales adjuntas a un archivo de análisis obtenido a partir de una encuesta de hogares pueden contener los siguientes factores:

- La ponderación basada en el diseño calculada como el recíproco de la probabilidad de selección;
- Un factor de ajuste por la falta de respuesta;
- Un factor de ajuste de postestratificación;
- Un factor de compensación de la ponderación.

15. Estos factores deberían documentarse para que pudieran ser examinados por los analistas. Los factores de ajuste aplicados a las ponderaciones iniciales basadas en el diseño suponen la formulación de juicios subjetivos y, en ocasiones, arbitrarios en la definición de las clases de ponderación, en la selección de los totales de control para el ajuste de posestratificación y en el alcance de la compensación de la ponderación aplicada para controlar el efecto de diseño. Cuando en las estimaciones de la encuesta aparecen resultados imprevistos o anomalías aparentes, no es extraño someter a detenido examen tanto el proceso de ponderación como todos los demás aspectos del diseño de la encuesta y de su ejecución.

16. En general, los usos analíticos de los datos de las encuestas de hogares representan especiales desafíos debido a los diseños de encuestas complejas que incluyen el uso de ponderaciones y una estructura de diseño. Los efectos del diseño debidos al diseño de encuestas complejas se examinan en varios de los capítulos de este manual. En el capítulo XX, en particular, se analizan los efectos del diseño de encuestas complejas en los resultados del análisis. Puede verse un análisis más a fondo del análisis de encuestas complejas o sobre algunos temas concretos en Skinner, Holt y Smith (1989); Korn y Graubard (1999); y Chambers y Skinner (2003). En el capítulo XX de la presente publicación se hace un examen más técnico acerca del análisis de las encuestas complejas, mientras que en el capítulo XXI se describen los programas informáticos y se presentan algunos ejemplos de planteamientos para analizar los datos de encuestas con ejemplos procedentes de datos reales.

17. **Estadísticas no lineales.** Incluso las estadísticas sencillas como las medias se convierten en no lineales en las encuestas complejas. Para estimar una media de población a partir de una encuesta compleja es necesario estimar una población total de la variable en cuestión (por ejemplo, el ingreso familiar) y el tamaño de la población (por ejemplo, el número total de familias). Luego la media se estima como coeficiente de las dos estimaciones. El ingreso familiar medio se estimaría como sigue:

$$\text{Estimación del ingreso familiar medio} = \frac{\text{Estimación del ingreso familiar total}}{\text{Estimación del número total de familias}}$$

Esta media estimada resulta ser una función no lineal (coeficiente) de dos estadísticas lineales. En las encuestas complejas el tamaño de la muestra (número de observaciones de un tipo determinado) es en sí mismo una variable aleatoria. Este tipo de estimaciones no lineales no están exentos de sesgo en las muestras pequeñas, pero son coherentes en el sentido superficial de que si el tamaño de la muestra aumentara hasta alcanzar el tamaño de la población finita, la estimación no lineal sería exactamente igual al valor de la población finita comparable (Cochran, 1977, págs. 21, 153 y 190). Si decidimos considerar la población finita como resultado de una población infinita hipotética, podemos considerar la posibilidad de dejar que el tamaño de la muestra aumente sin límites. En este caso podemos afirmar la coherencia del modelo cuando la estimación no lineal converge en probabilidad con el parámetro de la superpoblación a medida que aumenta el tamaño de la muestra (véase, por ejemplo, Skinner, Holt y Smith, 1989, págs. 17-18).

18. Los errores estándar de las estadísticas no lineales pueden expresarse sólo de manera indicativa utilizando aproximaciones de la serie de Taylor de primer orden. Las estimaciones de los errores estándar de las estadísticas no lineales pueden obtenerse utilizando aproximaciones de las serie de Taylor de primer orden o métodos de reproducción, como la reproducción repetida equilibrada o la repetición *jackknife*.

19. El mismo argumento aplican al análisis que utiliza modelos "lineales" cuando las funciones lineales requeridas de la variable tanto dependiente como independiente se estiman primero con respecto a toda la población.

20. En resumen, el uso de ponderaciones da lugar a estimaciones lineales no sesgadas y a estimaciones no lineales coherentes. En la práctica, el uso de estimaciones coherentes se considera satisfactorio para controlar el sesgo de estimación. Otros tipos de sesgos y de errores no muestrales, como los resultantes de la falta de respuesta, los errores de los entrevistadores o el error del informante, suelen tener mucha mayor importancia práctica, sobre todo cuando el tamaño de la muestra es grande.

21. **Estructura del diseño de la muestra en las encuestas de hogares.** En general, tanto la población como el diseño de la muestra pueden tener cierta estructura. En los diseños muestrales de encuestas de hogares suele imponerse una estructura anidada al marco muestral, como se ha visto en el apartado anterior e ilustra el cuadro XIX.1. Si bien la estructura no influye en la construcción de estimaciones estadísticas de primer orden, como totales, medias, relaciones o coeficientes del modelo, sí afecta a las estimaciones estadísticas de segundo orden (estimaciones de la varianza), que permiten a los analistas averiguar los errores estándar de las estadísticas de primer orden para elaborar pruebas de significación estadística referentes a hipótesis especificadas.

22. Toda la expresión de la varianza de las estimaciones basadas en muestras estratificadas en varias etapas tiene componentes en cada una de las etapas del diseño de la muestra. Por ejemplo, si la estratificación se emplea únicamente en la primera etapa, una estimación $\hat{T}$ de una población total T basada en un diseño en tres etapas con segmentos de área, los hogares y los miembros de los hogares, podría tener una varianza con la forma siguiente:

$$Var(\hat{T}) = \sum_h \left(fpc_{h1} \frac{S_{h1}^2}{n_{h1}} + fpc_{h2} \frac{S_{h2}^2}{n_{h2}} + fpc_{h3} \frac{S_{h3}^2}{n_{h3}} \right)$$

donde los términos dentro del estrato h se definen como sigue. Los términos fpc_{h1} son factores de corrección de la población finita en las fases de selección del segmento de área (i=1), selección de la unidad de habitación (i=2) y selección de las personas (i=3). Los términos S_{hi}^2 son componentes de la varianza basados en los datos ponderados en las tres etapas de muestreo. Los términos n_{h1} son los tamaños muestrales de los segmentos (i=1), los hogares (i=2) y las personas (i=3) dentro del estrato h. En la práctica no es raro que algunos de estos componentes de la varianza sean difíciles de estimar o que no sean estimables; ello puede ocurrir debido a submuestras de tamaño 1 o a otras razones. Cochran (1977, pág. 279) observa que si el factor de corrección de la población finita en la primera etapa (que se supone es 1) puede ignorarse, las estimaciones de la varianza pueden basarse en un análogo de esta fórmula mucho más sencillo que implica sólo la primera fase de muestreo. La suposición de un factor de población finita de la primera etapa de 1 se describe muchas veces como una estimación de la varianza del diseño de la muestra "con sustitución" para realizar un cálculo aproximado de la varianza para un diseño muestral "sin sustitución".

23. Para conseguir que esto funcione en las estimaciones lineales de los totales de población para el diseño en tres etapas antes examinado, cuando las unidades de observación son personas, podemos definir una nueva variable

$$Z_{hi} = n_{h1} \sum_j \sum_k w_{hijk} Y_{hijk}$$

donde w_{hijk} e Y_{hijk} son la ponderación y la variable observada para la persona k del hogar j del segmento de área i dentro del estrato h, respectivamente. Según ello, es posible obtener una estimación razonable de la varianza en la forma siguiente:

$$var(\hat{T}) = \sum_h \frac{\sum_i (Z_{hi} - \bar{Z}_h)^2}{n_{h1}(n_{h1} - 1)}$$

Ello funciona porque con esta formulación la estimación de la población total puede expresarse así:

$$\hat{T} = \sum_h \bar{Z}_h$$

Con una elección adecuada de Z_{hi}, las varianzas de las estadísticas no lineales y lineales puede estimarse utilizando aproximaciones de la serie de Taylor de primer orden[2]. Ello se amplía a las estimaciones del parámetro en regresión o regresión logística. Para que así ocurra no es necesario que las contribuciones de la varianza de las etapas posteriores sean estimables.

[2] Woodruff (1971) explica cómo pueden obtenerse variables linearizadas para facilitar el cálculo de aproximaciones de la varianza de la serie de Taylor compleja.

24. Si la corrección de la población finita de la primera etapa es notablemente inferior a 1, esta formulación sobrevalorará la varianza y dará lugar a una sobrestimación del error estándar de las estimaciones de la encuesta. Una pequeña sobrestimación daría lugar a intervalos de confianza moderadamente amplios o podría dar lugar a un número menor de declaraciones de significación estadística cuando se realicen las comprobaciones de la hipótesis. En ese sentido, el supuesto de una corrección de la población finita de la primera etapa de 1 se considera estadísticamente prudente, ya que servirá como protección frente a las declaraciones falsas de significación estadísticas. Conviene señalar que la aplicación de programas informáticos basados en la serie de Taylor y en la reproducción se simplifica con el supuesto de un factor de corrección de la población finita de 1 en la primera etapa de muestreo. (véase el capítulo XXI).

C. Estadísticas analíticas

25. En este apartado pasamos de la consideración de estimaciones descriptivas simples a lo que se conoce como "estadísticas analíticas", es decir, estadísticas que examinan las relaciones entre variables. De hecho, en cuanto los usuarios de los datos desean comparar estimaciones entre los campos, necesitan estadísticas "analíticas". Las estadísticas analíticas sencillas pueden estar basadas en diferencias entre campos, como ocurriría en una comparación de la proporción de hogares con ingreso total inferior al umbral de pobreza en dos subdivisiones geopolíticas o una comparación de la producción agrícola durante los dos últimos años. Algunas veces las estimaciones de una comparación simple son independientes entre sí, por lo que el error estándar de la diferencia puede determinarse estrictamente a partir del error estándar de las estimaciones individuales. En esas circunstancias el error estándar (*se*) de la diferencia estimada entre dos medias de campos puede ser el siguiente:

$$se(\bar{y}_1 - \bar{y}_2) = \sqrt{\{se(\bar{y}_1)\}^2 + \{se(\bar{y}_2)\}^2}$$

Esta fórmula correspondiente al error estándar de una diferencia supone que las dos estimaciones son independientes y que, en consecuencia, sus estimaciones no están correlacionadas. Esta forma de error estándar de las diferencias es práctica para los usuarios de datos, ya que puede deducir el error estándar de una diferencia a partir de los errores estándar de las estimaciones individuales publicados. En cambio, con diseños muestrales complejos, las estimaciones de los campos muchas veces están correlacionadas. La varianza de la diferencia de las dos estimaciones de los campos incluye entonces un término de covarianza:

$$se(\bar{y}_1 - \bar{y}_2) = \sqrt{\{se(\bar{y}_1)\}^2 + \{se(\bar{y}_2)\}^2 - 2\mathrm{cov}(\bar{y}_1, \bar{y}_2)}$$

26. El término de covarianza es en general positivo, por lo que da lugar a un error estándar de la estimación de la diferencia que es inferior al del caso independiente antes exa-

minado. Las encuestas de hogares pueden tratar de utilizar el término de la covarianza en los errores estándar de las estimaciones de las diferencias; las encuestas longitudinales por panel consiguen una alta covarianza positiva entre las estimaciones anuales utilizando una muestra común y continuada de individuos u hogares. Como el error estándar de la diferencia no puede deducirse de los errores estándar publicados de las estimaciones individuales, resulta necesario prever qué comparaciones son de mayor interés y publicar también sus errores estándar.

27. En las estadísticas estrictamente descriptivas sobre poblaciones finitas, el error estándar de las estimaciones descriptivas se reduce correctamente mediante la aplicación de un factor de corrección de la población finita (*fpc*). En el caso más sencillo de muestreo aleatorio simple, el factor de corrección de la población finita es el siguiente:

$$fpc = 1 - \frac{n}{N}$$

donde n es el tamaño de la muestra y N es el tamaño de la población. Si el objetivo es analítico, incluso en el caso más sencillo de significación estadística de la diferencia observada entre dos medias de campo, el uso del factor de corrección de la población finita es inadecuado (Cochran, 1977, págs. 34-35). Ello se debe a que la forma de la prueba de significación estadística obliga a formular una hipótesis sobre si las poblaciones de ambos campos podrían haber resultado de una población hipotética infinita común (una única superpoblación)[3]. El uso de factores de corrección de la población finita en un diseño complejo estructurado se examina más adelante.

[3] Cochran (1977, pág. 39) mantiene que el uso del factor de corrección de la población finita no es adecuado para la comprobación estadística de las diferencias entre las medias de campos. La interpretación de esta orientación resulta más ambigua cuando se aplica a diseños complejos que implican tanto estratificación como conglomeración; Chromy (1998) examina el problema con respecto al muestreo de alumnos dentro de las escuelas cuando éstas presentan altas tasas de estratificación y muestreo. Graubard y Korn (2002) han examinado recientemente este tema.

D. Observaciones generales sobre los modelos de regresión

28. Los métodos incluidos en los apartados siguientes implican una técnica de elaboración de modelos que establece un modelo de la variación de una variable de respuesta clave o variable dependiente e identifican los subconjuntos de un conjunto de posibles variables explicativas que contribuyen de manera más significativa a esa variación. La elección de este subconjunto "óptimo" puede realizarse mediante la aplicación de procedimientos adecuados de selección de variables o utilizando un procedimiento secuencial adecuado para explorar diferentes modelos teniendo muy en cuenta la idoneidad, desde un punto de vista práctico de las variables que se introducen o se eliminan del modelo en cada paso del procedimiento analítico.

29. Desearíamos insistir en que las técnicas examinadas en este capítulo deben considerarse suplementarias, no sustitutivas, de otros métodos de análisis más sencillos. La exploración inicial de los datos utilizando resúmenes descriptivos sencillos (medias, desviaciones estándar, etcétera), procedimientos gráficos (diagramas de dispersión, gráficos de barras, diagramas de recuadros, etcétera) y las tabulaciones de datos pertinentes es muy valiosa y debería constituir la primera etapa del análisis de datos. Algunas veces quizá sea lo único que haga falta. En cambio, en muchas ocasiones los objetivos de la encuesta exigen ulteriores análisis de los datos, en cuyo caso es probable que adquieran importancia las técnicas de elaboración de modelos.

30. Los métodos de elaboración de modelos aquí examinados son especialmente pertinentes en los casos en que se adopta un planteamiento de conjunto; por ejemplo, cuando el objetivo analítico es comprender la justificación de los sistemas agrícolas existentes y la forma en que los hogares gestionan sus limitados recursos para atender las necesidades tanto de producción como de consumo. En todo momento se insiste sobre todo en la aplicación práctica de una técnica adecuada de elaboración de modelos en la que se valoren las posibles dificultades que se presentarán sobre el terreno en los países en desarrollo. Se ponen de ma-

nifiesto las limitaciones del análisis para conseguir que los planteamientos aquí examinados se apliquen sólo después de atenta reflexión sobre la idoneidad del método aplicado para el contexto de investigación previsto.

31. Los modelos de regresión se utilizan para llegar a una mejor comprensión de la relación entre una variable dependiente y un conjunto de variables independientes o explicativas. No obstante, hay que tener en cuenta que normalmente no es posible asignar una relación de causa-efecto a ninguna conexión observada entre una variable dependiente y una variable explicativa, salvo en el caso de experimentos controlados y al azar bien concebidos[4]. Teniendo esto presente, es mucho lo que se puede aprender aplicando modelos de regresión a los datos observados procedentes de encuestas de hogares.

[4] En las encuestas pueden incluirse experimentos aleatorios. Muchas veces se trata de experimentos metodológicos en una muestra previa o en muestras complementarias de una encuesta en curso. Pueden realizarse también experimentos sociales contratando sujetos para un experimento social en el que se utilice una muestra de encuestas de hogares.

32. A diferencia de los datos procedentes de experimentos controlados que utilizan la aleatorización y el control de las variables auxiliares, los datos de encuestas de hogares se basan normalmente en la observación, con poco o ningún control sobre otros factores que puedan influir en las relaciones entre variables. Los métodos de regresión pueden eliminar algunas veces los efectos de estas variables de confusión incontroladas, por lo que quizá puedan obtenerse estimaciones menos sesgadas de la verdadera relación.

33. Los modelos basados en la regresión son muchas veces de carácter exploratorio. Pueden elaborarse distintos modelos para explicar el comportamiento de una variable dependiente. Las variables explicativas utilizadas en el modelo se limitan a las que se encuentran disponibles en el archivo de datos de la encuesta; en consecuencia, las variables seleccionadas para explicar la variación de una variable dependiente pueden ser únicamente correlaciones sólidas del factor causal verdadero. Puede haber correlaciones contradictorias del factor explicativo, ninguna de las cuales parece estar relacionada lógicamente con la variable dependiente. Los analistas de encuestas de hogares deberían aplicar la teoría sustantiva (por ejemplo, social o económica) al elegir variables explicativas y al determinar la forma de la relación (por ejemplo, lineal o no lineal).

34. Cuando la teoría sustantiva no revela fuertes relaciones teóricas o cuando parece indicar diversas variables explicativas enfrentadas, pueden aplicarse planteamientos de selección de variables a partir de programas estándar (no de encuestas) para identificar posibles variables explicativas. En muchos programas informáticos no relacionados con encuestas se ofrecen planteamientos de selección de variables de avance y retroceso que ayudan a identificar variables explicativas que tienen relaciones lineales con la variable dependiente. Si el programa lo permite, es muy recomendable el uso de ponderaciones de encuestas incluso para este análisis exploratorio. Las ponderaciones de las encuestas pueden normalizarse para que den una suma equivalente al total de la muestra a fin de disponer de mejores estimaciones del error y comprobaciones más correctas del significado estadístico (véanse ejemplos de este planteamiento en el capítulo XXI). Después de utilizar programas estadísticos no concebidos para encuestas con el fin de realizar una selección de variables, es conveniente evaluar el modelo utilizando un programa informático que emplee las ponderaciones de la encuesta y reconozca el diseño de la encuesta de hogares.

35. Las variables del modelo pueden ser variables categóricas, variables de recuento o variables de medición continua. Los modelos de regresión lineal se utilizan cuando las variables dependientes son recuentos o mediciones continuas; para los datos de recuento se proponen las transformaciones logarítmicas. Cuando la variable de recuento dependiente incluye valores de cero, la transformación logarítmica no vale, pero pueden utilizarse procedimientos como el PROC LOGLINK (SUDAAN 2001) para acomodarse al valor previsto del logaritmo de una variable de recuento. La regresión logística se utiliza cuando la variable dependiente es una variable categórica definida en dos niveles; los modelos de regresión multinomial pueden aplicarse también a las variables dependientes categóricas con más de dos niveles. A este respecto, clasificamos las variables explicativas como categóricas o continuas,

ya que las variables de recuento y continuas se tratan fundamentalmente de la misma manera en el contexto de la elaboración de modelos. Los datos de las encuestas pueden analizarse también utilizando modelos de supervivencia y otras técnicas con múltiples variables que no se examinan en este capítulo.

36. El uso de variables explicativas categóricas, que definen los campos del estudio, es análogo a la elaboración de comparaciones de campos simples sin utilizar modelos. El uso de modelos permite al analista introducir ajustes simultáneos en función de otras posibles variables explicativas. Es lo que se conoce con frecuencia con el nombre de ajuste en función de las covariables. Cuando no hay un ajuste de este tipo, los coeficientes del modelo de regresión reproducen comparaciones de campos simples y estiman las diferencias existentes en la población. Cuando se incluyen otras variables en el modelo en calidad de covariables, los coeficientes del modelo de regresión estiman las diferencias de campo que existirían hipotéticamente si las covariables se mantuvieran en los mismos niveles en todos los campos.

37. Los coeficientes del modelo de regresión para las variables explicativas continuas pueden obtenerse también con o sin ajuste en función de otras covariables. La decisión de introducir o no un ajuste en función de las covariables debe estar determinada por el objetivo del análisis. Las estimaciones no ajustadas describen una relación empírica entre variables dependientes y explicativas tal como existen en la población. Las estimaciones ajustadas describen la misma relación si las otras variables se mantienen hipotéticamente constantes. Si las otras variables incluidas en el modelo son también variables predictivas válidas de la variable dependiente, puede mejorar la precisión de los valores previstos para determinados niveles de las variables predictivas clave sometidas a estudio. La elección de los métodos de análisis debería depender del objetivo del análisis.

38. En los ejemplos siguientes sólo se consideran modelos sencillos de variables explicativas continuas. Cuando las variables explicativas son continuas, el analista debería investigar la relación de la variable dependiente con las posibles variables explicativas. Gráficos sencillos pueden poner de manifiesto que una relación lineal es inadecuada para establecer una debida relación entre las variables. Según sea el gráfico observado, pueden incorporarse términos adicionales (cuadráticos o cúbicos) para reflejar mejor la relación. Entonces la variable dependiente puede tener relaciones lineales con una variable explicativa, con su cuadrado y con sus términos cúbicos o superiores. Pueden utilizarse gráficos residuales, después de haber incluido algunas de las variables explicativas posibles, para determinar si otras variables u órdenes superiores (términos cuadrados o cúbicos) de las variables incluidas pueden estar influyendo en el encaje del modelo. En cuanto a las variables explicativas con una gran diversidad de valores y diferentes efectos en la variable dependiente dentro de ese intervalo, muchas veces son útiles los modelos *spline* que permiten que la relación cambie dentro de los subconjuntos del intervalo. Cuando la muestra de una encuesta incluye personas jóvenes, de mediana edad y ancianos, los efectos de la edad muchas veces pueden manifestarse mediante el uso de modelos de regresión de ese tipo.

39. Otros procedimientos de diagnóstico son el examen de la bondad del ajuste de los modelos propuestos y el examen de la significación estadística de los parámetros de regresión para las variables agregadas. Los procesos basados en procedimientos estándar (no de encuestas) pueden adaptarse a los datos ponderados de las encuestas. El concepto de variación explicada puede utilizarse en los datos de encuestas ponderados y en la regresión lineal. Pueden utilizarse conceptos como el de cuadro de contingencia para evaluar el ajuste de los modelos de regresión logística. En Korn y Graubard (1999, cap. 3) puede verse un útil examen de la adaptación de los procedimientos diagnósticos al análisis de los datos de encuestas generales.

40. La elaboración de modelos de regresión basados en datos observados implica obviamente el concepto de análisis de datos exploratorio (Tukey, 1977). Este tipo de análisis

puede llevar a planteamientos útiles sobre los datos y la relación entre las variables observadas, pero la significación estadística de las conclusiones de este análisis "no planificado" deberían ser objeto de confirmación o de validación en el futuro mediante el estudio de otros datos de encuestas.

E. Modelos de regresión lineal

41. Al examinar los modelos de regresión lineal (que se abordan en este apartado) y los de regresión logística (véase el apartado F, páginas 359 a 361), conviene suponer que el muestreo es "con sustitución" en la primera etapa. Suponemos también que el archivo analítico de los datos de observación incluye variables de índices para los estratos, designadas mediante h, y para la unidades primarias de muestreo (UPM), designadas por i. No es necesario identificar variables de estructura adicionales cuando se va a utilizar el supuesto de diseño con sustitución en la primera etapa de selección de la muestra, como se ha indicado en el apartado B. Las consecuencias de la utilización de un diseño muestral complejo de hogares se incorporan en las estimaciones de los coeficientes del modelo y sus errores estándar únicamente si utilizamos un programa de estadística que tenga en cuenta debidamente el diseño de la encuesta de hogares, incluidas las ponderaciones analíticas y la estructura del diseño (estratos y UPM). Cuando examinamos los modelos en varios niveles, lo que se tendrá en cuenta será más bien la incorporación de la estructura del diseño en el modelo, y el análisis permitirá la estimación de los efectos relacionados con las variables de la estructura.

42. Un modelo de regresión lineal con una variable explicativa continua y una variable explicativa categórica puede expresarse como sigue:

Modelo 1

$$y_{hij} = \alpha x_0 + \beta_1 x_{1hij} + \sum_{d=1}^{D} \gamma_d x_{2dhij} + \varepsilon_{hij}$$

43. En el modelo 1 las observaciones se representan mediante la variable dependiente observada y_{hij}, una variable de intersección x_0 que sería siempre 1; una variable explicativa continua observada x_{1hij} y un conjunto de variables indicativas x_{2dhij} que definirían D niveles de una variable categórica. Los parámetros del modelo de regresión α, β_1 y $\gamma_d (d=1, 2, ..., D)$ se conocen como coeficientes de regresión y se estiman mediante el análisis. El término final del modelo es el término de error y mide la desviación con respecto al modelo asociada con la j^a observación asociada con la i^a UPM del h^o estrato. Este es un modelo de los efectos principales, ya que no contiene ningún efecto de interacción.

44. Según cuál sea el programa informático aplicado, el conjunto de variables indicativas puede especificarse como una sola variable en una declaración del modelo; quizá sea necesario definir la variable como categórica y especificar el número de niveles con declaraciones u órdenes del programa. Luego, el programa define un vector de variables indicativas. Una variable indicativa, por ejemplo, x_{2dhij} se considerará como 1 si la observación hij pertenece a la categoría d; en caso contrario, sería 0. Para evitar la dependencia lineal entre las variables explicativas, el programa de análisis vuelve a establecer los parámetros de los indicadores para la variable categórica. Normalmente ello se hace abandonando la categoría final de la variable categórica; luego esta categoría se convierte en categoría de referencia[5]. El cuadro XIX.2 muestra algunos de los efectos que se pueden estimar para el modelo 1 cuando la variable dependiente es el ingreso de los hogares procedente de los salarios, la variable explicativa continua es el número de asalariados del hogar y la variable categórica define cuatro ámbitos regionales del país (norte, sur, este y oeste).

[5] Es posible también estimar los coeficientes de las variables categóricas agregando una limitación lineal; por ejemplo, exigiendo que la suma de los efectos sea cero o que los efectos ponderados sean cero.

45. Los coeficientes de regresión estimados para las variables se definen con respecto a la diferencia entre un campo y el campo de referencia. La prueba de significación estadís-

Cuadro XIX.2
Interpretación de las estimaciones del parámetro de regresión lineal cuando la variable dependiente son los ingresos del hogar procedentes de salarios (modelo 1)

Efecto (normalmente identificado en el resultado del programa)	Coeficiente de	Estimación de	Interpretación
Intersección	x_0=1	α	Ingreso de los hogares asalariados en la casilla de referencia o niveles cero: 0 asalariados en la región occidental
Asalariados en el hogar	x_{1hij}	β_1	Cambio en el ingreso asalariado de los hogares por asalariado adicional (ajustado en función de la región)
Región			Diferencias regionales en los ingresos de los hogares procedentes de salarios (ajustadas en función de los asalariados del hogar)
Norte (d=1)	$x_{21hij} - x_{24hij}$	$B_2 = \gamma_1 - \gamma_4$	Norte frente a oeste
Sur (d=2)	$x_{22hij} - x_{24hij}$	$B_3 = \gamma_2 - \gamma_4$	Sur frente a oeste
Este (d=3)	$x_{23hij} - x_{24hij}$	$B_4 = \gamma_3 - \gamma_4$	Este frente a oeste
Oeste (campo de referencia, d=4)	$x_{24hij} - x_{24hij} = 0$	$\gamma_4 - \gamma_4 = 0$	Ninguna estimación

tica de un coeficiente estimado para el campo norte comprueba realmente si norte y oeste podrían ser muestras aleatorias de la misma superpoblación común. Si el coeficiente de la región norte es significativamente distinto de 0 (sobre la base de una prueba hipotética con nivel de significación de 0,05), el analista puede concluir que es muy poco probable (5% o menos) que los salarios de los hogares en las regiones norte y oeste sean muestras de la misma superpoblación después de los ajustes necesarios en función del número de asalariados del hogar. Los programas de estadística permiten a los usuarios especificar diferentes conjuntos de referencia, bien ordenando las categorías (de manera que la categoría de referencia deseada sea la última) o mediante especificación explícita. Ello puede representar un instrumento útil para obtener estimaciones significativas de parámetros de regresión. Otras comparaciones pueden estimarse también mediante las funciones de los coeficientes estimados.

46. En el cuadro XIX.3 pueden verse algunas funciones estimables del modelo 1 basadas en estimaciones de los parámetros recogidos en el cuadro XIX.2. En el cuadro XIX.3 se observan las estimaciones del modelo 1 sobre los ingresos de los hogares resultantes de los salarios, clasificadas por región y número de asalariados del hogar. Sería fácil ampliar esto a tres o más asalariados por hogar.

47. Examinemos los supuestos que el analista debe formular al utilizar el modelo 1 para estudiar los ingresos del hogar procedentes de los salarios. Quizá el supuesto más importante es que los ingresos del hogar resultantes de los salarios están linealmente relacionados con el número de asalariados. Según el supuesto de linealidad, el cambio en los ingresos de los hogares resultantes de los salarios aumenta en la misma cantidad cuando se pasa de 0 a un asalariado, de uno a dos asalariados, de dos a tres asalariados, etcétera. Este supuesto

Cuadro XIX.3
Ingresos estimables de los hogares resultantes de los salarios (modelo 1)

Región	Hogares con un asalariado	Hogares con dos asalariado
Norte	$\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2$	$\hat{\alpha} + 2\hat{\beta}_1 + \hat{\beta}_2$
Sur	$\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_3$	$\hat{\alpha} + 2\hat{\beta}_1 + \hat{\beta}_3$
Este	$\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_4$	$\hat{\alpha} + 2\hat{\beta}_1 + \hat{\beta}_4$
Oeste	$\hat{\alpha} + \hat{\beta}_1$	$\hat{\alpha} + 2\hat{\beta}_1$

parece dudoso. Como las variables categóricas requieren menos supuestos sobre la forma de la relación entre la variable explicativa y la variable dependiente, el analista quizá prefiera convertir el número de asalariados en una variable categórica y, por lo tanto, utilizar un modelo que contenga únicamente variables categóricas[6]. Una variante del modelo 1 podría expresarse como sigue:

Modelo 2

$$y_{hij} = \alpha x_0 + \sum_{d=1}^{D_1} \gamma_{1d} x_{1dhij} + \sum_{d=1}^{D_2} \gamma_{2d} x_{2dhij} + \varepsilon_{hij}$$

[6] Puede verse una exposición adicional de la metodología para evaluar la bondad de un ajuste de un modelo de regresión lineal y algunas otras alternativas de relaciones no lineales en Korn y Graubard (1999, págs. 95-100).

48. En cuanto al modelo 2, el analista podría definir sólo dos categorías de asalariados o muchos menos, según la distribución del número de asalariados en los hogares. Para limitar el número de parámetros que se deben estimar quizá basten cuatro categorías:

- Categoría 1: ningún asalariado;
- Categoría 2: un asalariado;
- Categoría 3: dos asalariados;
- Categoría 4: tres o más asalariados.

49. Este modelo es todavía un modelo de efectos principales, pero el número de parámetros de regresión ha pasado de cinco a siete. En el cuadro XIX.4 puede verse la interpretación de los coeficientes de regresión estimados en el marco del modelo 2. Este modelo ya no requiere que el analista suponga una relación lineal de los ingresos salariales del hogar con el número de asalariados del hogar. No obstante, como en el modelo no hay términos de interacción, se supone lo siguiente:

- El efecto "asalariados del hogar" es el mismo en las cuatro regiones.
- El "efecto región" es el mismo en todos los niveles de "asalariados del hogar".

50. La mayor parte de los programas de regresión permiten comprobar las interacciones entre variables categóricas. En este caso habrá nueve grados de libertad de interacción. Si bien es posible interpretar los efectos de los modelos de regresión con dos efectos categóricos principales y una interacción, recomendaríamos un planteamiento diferente. En primer

Cuadro XIX.4
Interpretación de las estimaciones del parámetro de regresión lineal cuando la variable dependiente son los ingresos del hogar procedentes de los salarios (modelo 2)

Efecto (normalmente identificado en el resultado del programa)	Coeficiente de	Estimación de	Interpretación
Intersección	$x_0 = 1$	α	Ingresos de los hogares procedentes de los salarios en el nivel de referencia (ningún asalariado y región occidental)
Asalariados en el hogar			Cambios en los ingresos de los hogares procedentes de ingresos salariales por asalariado adicional (ajustados en función de la región)
Uno (*d*=1)	$x_{11hij} - x_{14hij}$	$B_1 = \gamma_{11} - \gamma_{14}$	Uno frente a ninguno
Dos (*d*=2)	$x_{12hij} - x_{14hij}$	$B_2 = \gamma_{12} - \gamma_{14}$	Dos frente a ninguno
Tres o más (*d*=3)	$x_{13hij} - x_{14hij}$	$B_3 = \gamma_{13} - \gamma_{14}$	Tres frente a ninguno
Ninguno (campo de referencia, *d*=4)	$x_{14hij} - x_{14hij} = 0$	$\gamma_{14} - \gamma_{14} = 0$	Ninguna estimación
Región			Diferencias regionales en los ingresos de los hogares procedentes de los salarios (ajustadas en función del número de asalariados en el hogar)
Norte (*d*=1)	$x_{21hij} - x_{24hij}$	$B_4 = \gamma_{21} - \gamma_{24}$	Norte frente a oeste
Sur (*d*=2)	$x_{22hij} - x_{24hij}$	$B_5 = \gamma_{22} - \gamma_{24}$	Sur frente a oeste
Este (*d*=3)	$x_{23hij} - x_{24hij}$	$B_6 = \gamma_{23} - \gamma_{24}$	Este frente a oeste
Oeste (campo de referencia, *d*=4)	$x_{24hij} - x_{24hij} = 0$	$\gamma_{24} - \gamma_{24} = 0$	Ninguna estimación

lugar compruébese la interacción: en este caso el modelo 2 podría incrementarse para incluir la interacción entre "asalariados del hogar" y "región". Si la prueba estadística indica que hay interacciones, puede incorporarse todo el modelo con 16 parámetros estimables aplicando un modelo más sencillo con una única variable categórica definida en 16 niveles. Este sería el modelo 3, que podría formularse así

Modelo 3

$$y_{hij} = \alpha x_0 + \sum_{d=1}^{16} \beta_{1d} x_{1dhij} + \varepsilon_{hij}$$

51. Los 16 niveles de la nueva variable categórica y sus estimaciones (entre paréntesis) son los siguientes:

- Norte, un asalariado ($\hat{\alpha} + \hat{\beta}_1$)
- Norte, dos asalariados ($\hat{\alpha} + \hat{\beta}_2$)
- Norte, tres o más asalariados ($\hat{\alpha} + \hat{\beta}_3$)
- Norte, ningún asalariado ($\hat{\alpha} + \hat{\beta}_4$)
- Sur, un asalariado ($\hat{\alpha} + \hat{\beta}_5$)
- Sur, dos asalariados ($\hat{\alpha} + \hat{\beta}_6$)
- Sur, tres o más asalariados ($\hat{\alpha} + \hat{\beta}_7$)
- Sur, ningún asalariado ($\hat{\alpha} + \hat{\beta}_8$)
- Este, un asalariado ($\hat{\alpha} + \hat{\beta}_9$)
- Este, dos asalariados ($\hat{\alpha} + \hat{\beta}_{10}$)
- Este, tres o más asalariados ($\hat{\alpha} + \hat{\beta}_{11}$)
- Este, ningún asalariado ($\hat{\alpha} + \hat{\beta}_{12}$)
- Oeste, un asalariado ($\hat{\alpha} + \hat{\beta}_{13}$)
- Oeste, dos asalariados ($\hat{\alpha} + \hat{\beta}_{14}$)
- Oeste, tres o más asalariados ($\hat{\alpha} + \hat{\beta}_{15}$)
- Oeste, ningún asalariado ($\hat{\alpha}$)

52. Con la decimosexta categoría definida como casilla de referencia, la estimación de intersección del modelo 3 $\hat{\alpha}$ corresponde a los ingresos del hogar resultantes de los salarios para dicha casilla (oeste, ningún asalariado). La estimación de los ingresos del hogar resultantes de los salarios en cada una de las otras 15 casillas se estima como en el caso de la decimosexta casilla pero agregando el coeficiente de regresión estimado para la casilla en cuestión. Estas 16 estimaciones podrían obtenerse también a partir de estimaciones directas. Si las ponderaciones de la encuesta y la estructura de diseño se aplican en programas adecuados para encuestas, las estimaciones y sus errores estándar estimados deberían ser idénticos en los dos planteamientos (modelo 3 o estimación directa). La aplicación del modelo 3 no da mejores resultados que la elaboración de 16 estimaciones directas.

53. Si el tamaño de la muestra de alguna de las 16 casillas es pequeña, las estimaciones correspondientes a esas casillas "con muestras pequeñas" serán poco precisas. La utilización de un modelo de efectos principales (modelos 1 ó 2) permite obtener estimaciones más precisas en las casillas con muestras pequeñas "tomando en préstamo" el tamaño de la muestra de las estimaciones marginales y formulando algunos supuestos adicionales (como se ha explicado antes) sobre la forma en que la población finita se deriva de la superpoblación hipotética.

54. Los analistas utilizan en general modelos para introducir ajustes en función de cierto número de variables explicativas. Supongamos que un analista desea introducir un ajuste de acuerdo con las características de la ciudad o la comunidad, como el nivel de ur-

banización (porcentaje urbano). El análisis quizá revele que el efecto región se reduce después de tener en cuenta y estandarizar el porcentaje urbano. En un modelo lineal de efectos principales el ajuste en función del porcentaje urbano (como variable explicativa continua o categórica) permite obtener estimaciones del efecto región suponiendo la misma distribución del porcentaje urbano (estándar) dentro de cada región. Sin el ajuste con arreglo a las covariables, el modelo (o las estimaciones directas) representa los parámetros regionales tal como existen; con el ajuste, el modelo representa los parámetros regionales tal como serían si se eliminaran los efectos de las covariables. En Korn y Graubard (1999, págs. 126-140) se examina el uso de los márgenes predictivos como método de estandarización.

F. Modelos de regresión logística

55. Cuando la variable dependiente es categórica no se aplican los planteamientos de regresión lineal. Aunque existen procedimientos para la elaboración de modelos multinomiales, examinaremos únicamente las variables categóricas binarias (dos niveles) que pueden analizarse utilizando modelos de regresión logística. En este sentido, la regresión logística es un caso especial y más sencillo de regresión multinomial.

56. En una variable dependiente binaria o de dos categorías codificadas como 0 ó 1, los planteamientos de la regresión lineal funcionan, pero pueden producir valores previstos que queden fuera del intervalo de 0 a 1. La regresión lineal podría utilizarse cono paso preliminar con una variable dependiente binaria para identificar variables explicativas que sean predictoras válidas de la variable dependiente, sobre todo si los programas informáticos de que dispone el analista incorporan procedimientos de selección de variables en el programa de regresión lineal pero no en el de regresión logística.

57. Se utilizan métodos numéricos para acomodarse a los parámetros de los modelos de regresión numérica; por ello, algunas veces puede haber dificultades para llegar a una solución. Los usuarios deberían tener en cuenta las advertencias que acompañan a los programas informáticos cuando se produzcan problemas de convergencia. En general, estos casos pueden resolverse simplificando el modelo.

58. Un modelo de regresión logística que suponga una variable explicativa continua y una variable explicativa categórica puede expresarse como sigue:

Modelo 4

$$\log\left(\frac{p(\underline{x}_{hij})}{1-p(\underline{x}_{hij})}\right)=\alpha x_0+\beta_1 x_{1hij}+\sum_{d=1}^{D}\gamma_d x_{2hij}+\varepsilon_{hij}$$

59. Excepto en el caso de la variable dependiente, los términos del modelo 4 se definen de la misma manera que en el modelo 1. Para comprender la transformación logística, consideremos un ejemplo en que $p(x_{hij})$ es una función de las variables explicativas. Designémoslo como p para mayor comodidad. Supongamos también que p es la probabilidad de que un hogar con un conjunto dado de valores para las variables explicativas tenga un nivel de ingreso inferior al umbral de pobreza establecido. En ese caso, $p/(1-p)$ es la probabilidad de encontrarse en la pobreza, y $\log(p/1-p)$ es el logaritmo de posibilidades (*log odds*) de p, denominado en algunas ocasiones *logit(p)*. El modelo 4 trata de relacionar el *logit* de p con las x. Las observaciones son hogares individuales donde observamos no su probabilidad de encontrarse en la pobreza, sino la situación real presente: pobreza o no pobreza. Asimismo, como la variable dependiente es un *logit* de p, cada parámetro $[\alpha, \beta_1$ y $\gamma_d (d = 1, ..., D)]$ está también en el *logit* de la escala p; además, se supone que la relación entre el *logit* de p y las x es lineal (compárese con el modelo 3, *supra*).

60. La reparametrización de las variables explicativas categóricas y la definición de las categorías de referencia es la misma que en la regresión lineal antes examinada. Los pará-

metros del modelo de regresión en el resultado del programa de regresión logística se parecen a los de la regresión lineal, pero tienen interpretaciones diferentes. En el cuadro XIX.5 se resume la interpretación de las estimaciones de los parámetros habituales para el modelo 4. Téngase en cuenta que hay cinco parámetros estimados (una intersección, α, y cuatro β).

61. Téngase también en cuenta que los parámetros del modelo logístico prevén el *probit* de encontrarse en situación de pobreza y no prevén directamente la probabilidad de encontrarse en la pobreza. Véase β_2 en el cuadro XIX.5. Se expresa como sigue, en forma de diferencia del logaritmo de posibilidades:

$$\beta_2 = \log\left(\frac{p(\text{norte})}{1-p(\text{norte})}\right) - \log\left(\frac{p(\text{oeste})}{1-p(\text{oeste})}\right)$$

Por las propiedades de los logaritmos, puede expresarse también como el logaritmo de una razón de posibilidades:

$$\beta_2 = \log\left(\frac{\left(\frac{p(\text{norte})}{1-p(\text{norte})}\right)}{\left(\frac{p(\text{oeste})}{1-p(\text{oeste})}\right)}\right)$$

Los resultados estándar de los procedimientos de regresión logística suelen ofrecer también la razón de posibilidades, ya que pueden calcularse fácilmente como sigue:

$$e^{\beta 2} = \left(\frac{\left(\frac{p(\text{norte})}{1-p(\text{norte})}\right)}{\left(\frac{p(\text{oeste})}{1-p(\text{oeste})}\right)}\right)$$

Además, las probabilidades de los hogares individuales de encontrarse en la pobreza pueden determinarse a partir del modelo en la forma siguiente:

$$p(\underline{x}_{hij}) = \frac{1}{1+e^{-\log it[p(\underline{x}_{hij})}}$$

Cuadro XIX.5
Interpretación de las estimaciones del parámetro de regresión logística cuando la variable dependiente es un indicador de los hogares por debajo del umbral de pobreza (modelo 4)

Efecto (normalmente identificado en el resultado del programa)	Coeficiente de	Estimación de	Interpretación
Intersección	x_0=1	α	*Logit* de encontrarse en la pobreza en el nivel de la casilla de referencia o cero: 0 asalariados en la región occidental
Asalariados en el hogar	x_{1hij}	β_1	Cambio en el *logit* de encontrarse en la pobreza por asalariado adicional (ajustado en función de la región)
Región			Diferencias regionales en los *logit* de encontrarse en la pobreza (ajustado en función de los asalariados del hogar)
Norte (d = 1)	$x_{21hij} - x_{24hij}$	$B_2 = \gamma_1 - \gamma_4$	Norte frente a oeste
Sur (d = 2)	$x_{22hij} - x_{24hij}$	$B_3 = \gamma_2 - \gamma_4$	Sur frente a oeste
Este (d = 3)	$x_{23hij} - x_{24hij}$	$B_4 = \gamma_3 - \gamma_4$	Este frente a oeste
Oeste (campo de referencia, d = 4)	$x_{24hij} - x_{24hij} = 0$	$\gamma_4 - \gamma_4 = 0$	Ninguna estimación

62. Al citar los resultados del ajuste de un modelo logístico, los autores interpretan algunas veces una razón de posibilidades de 2 como indicación de que la probabilidad de un evento (pobreza) en un campo (por ejemplo, norte) es dos veces superior a la posibilidad del evento (pobreza) en el otro campo (por ejemplo, oeste). Si bien este tipo de afirmación es aproximadamente exacto en eventos raros (p próxima a 0), dista mucho de ser verdad en eventos más frecuentes.

G. Uso de modelos en varios niveles

63. A continuación examinaremos la elaboración de modelos en varios niveles, y comenzaremos insistiendo en la necesidad de reconocer la estructura de datos de la encuesta. En este contexto es importante la estructura impuesta por encuestas concebidas en varias etapas. Por ejemplo, las regiones agroecológicas de un país pueden formar estratos, y de cada uno de ellos se pueden elegir varias unidades administrativas. Estas últimas formarán las unidades primarias de muestreo. Luego se seleccionan las unidades secundarias a partir de cada unidad primaria, a partir de las unidades secundarias se forman otras de rango inferior, y así sucesivamente. De esa manera se forma una estructura de datos jerárquica. Puede implicar el uso de variables de estratificación en uno o varios niveles.

64. Por ejemplo, una encuesta referente a los hogares agrícolas en una región puede suponer primeramente el uso de las divisiones administrativas de la región como unidades primarias, luego la elección de aldeas de cada división y después la de hogares de cada aldea, quizá garantizando que se incluyan las diferentes categorías de riqueza de los hogares. En este caso debe prestarse atención a las diferentes fuentes de variabilidad de los datos recopilados en los hogares. La variación general incorpora la variación entre las divisiones administrativas, entre aldeas y entre hogares dentro de las aldeas. Muchas veces los datos se recopilan también en cada nivel de la jerarquía, es decir, en este caso en relación con los hogares, las aldeas y las divisiones administrativas. Luego, es importante reconocer y observar qué variables se miden en las aldeas (disponiblidad de un agente de extensión agrícola; subvenciones gubernamentales para los fertilizantes, etcétera) y cuáles se miden en los hogares (por ejemplo, características socioeconómicas del hogar).

65. Para el análisis de datos pueden prepararse archivos "planos" independientes con hojas de cálculo para recopilar la información referente a las aldeas y a los hogares, utilizando algún identificador para relacionar dichos archivos. Así convendría en el caso de que los objetivos del análisis exijan que los datos referentes a las aldeas se analicen con independencia de los referentes a los hogares. En cambio, no sería aconsejable si el análisis debe combinar la información por aldeas con la información por hogares. Es mucho más conveniente una base de datos relacionados, es decir, una base de datos que permita almacenar datos en diferentes niveles en un único archivo, junto con enlaces que permitan relacionar los datos de un nivel con los de otro. El análisis debe integrar la información procedente de múltiples niveles, a fin de que las interrelaciones entre los niveles puedan explorarse para poder llegar a una interpretación global.

66. La elaboración de modelos en múltiples niveles es una técnica estadística clave de gran interés en este contexto. Este planteamiento (Goldstein, 2003; Snijders y Bosker, 1999; Kreft y de Leeuw, 1998) es aconsejable porque permite explorar las relaciones entre diferentes niveles o dentro de cada nivel de un diseño en varias etapas teniendo en cuenta la variabilidad en los distintos niveles (se consideran también las interrelaciones entre variables del mismo nivel) y ofrece, mediante el uso de programas informáticos adecuados, como *MLwiN* (Rashbash y otros, 2001) y SAS (2001), errores estándar basados en modelos para estimaciones procedentes de diseños de muestras complejas. Estos errores estándar pueden servir como aproximaciones razonables sobre errores estándar más exactos que tengan en cuenta la estratificación y la conglomeración. Conviene señalar que *MLwiN* podría tener

también en cuenta las ponderaciones de muestreo. Ello es importante porque la desigualdad de probabilidades de selección en un diseño muestral en varias etapas puede inducir a sesgos en los estimadores de los parámetros clave. Pfeffermann y otros (1998) y Korn y Graubard (2003) examinan estos temas más a fondo.

67. Conviene destacar brevemente las consecuencias que tendría el olvido de la estructura jerárquica en este punto, como puede ocurrir cuando los datos se agregan en un nivel superior o se desglosan hasta un nivel inferior. Si el análisis es pertinente y se necesita en un solo nivel, no hay problema. En cambio, hay que tener cuidado de que las inferencias se realicen sólo en ese nivel. No será posible hacer inferencias sobre un nivel concreto de la jerarquía a partir de datos analizados en otro nivel. Así pues, un análisis que pase por alto la jerarquía no permitirá explorar los efectos de diferentes niveles. Otra dificultad es la que se plantea si los datos se analizan en el nivel más bajo considerando las unidades de nivel superior como un factor del análisis. Esta práctica resulta ineficiente, ya que no permite generalizar las conclusiones a las unidades de nivel superior de la población; se aplicarán únicamente a las unidades incluidas en la muestra.

68. Presentamos a continuación un caso que ilustra los beneficios que puede aportar el uso de los modelos en distintos niveles al explorar las relaciones. Pueden encontrarse otros ejemplos en Congdon (1998), Langford, Bentham y McDonald (1998) y Goldstein y otros (1993).

Ejemplo 1

69. En un estudio de los factores que contribuyen a una ordenación comunitaria conjunta eficaz de los recursos pesqueros en los países insulares del Pacífico se eligieron 30 sitios de cinco países diferentes y se realizaron 133 entrevistas con pequeños grupos de debate integrados por entre dos y seis informantes de diferentes hogares (Banco Mundial, 2000). Fiji, Palau, Samoa, Islas Salomón y Tonga fueron los países elegidos para representar distintas condiciones de ordenación de las zonas costeras. Los 31 sitios se seleccionaron de manera que se abarcara una amplia gama de condiciones que parecían influir en el éxito de la ordenación. El estudio recopiló los "supuestos éxitos", es decir, las tendencias de las capturas percibidas por unidad de esfuerzo (CPUE), la situación de los hábitat, las amenazas para el lugar y una evaluación del cumplimiento de las normas. Los tres primeros indicadores se midieron en una escala de cinco puntos (5 = gran mejora; 1 = gran deterioro), mientras que el cumplimiento de las normas se midió según una escala de cuatro puntos.

70. Se recopilaron también datos de alcance nacional procedentes de las pesquerías y ministerios de medio ambiente de cada país y en los diferentes lugares. Además se pidió a cada grupo de debate, integrado por miembros de varios hogares, que expresara su opinión sobre hasta tres recursos pesqueros (en el caso de las CPUE), tres hábitat, tres amenazas y cinco normas de ordenación (en lo tocante al cumplimiento). Así pues, la información recopilada durante el estudio pertenecía a cuatro niveles: país, sitio, grupo de debate y recurso, hábitat y amenaza o norma específicos.

71. No obstante, es importante señalar que en esta encuesta se utilizó el muestreo no probabilístico, por lo que cabría argumentar que las conclusiones analíticas no pueden generalizarse a una población objetivo claramente definida. A los efectos del presente debate podemos suponer que el muestreo se había realizado con carácter probabilístico y que los datos referentes a los grupos de debate se analizarían utilizando un modelo en varios niveles, en cuyo caso la variable pertinente sería la percepción de la tendencia de la CPUE, obtenida calculando el promedio de las opiniones manifestadas con las puntuaciones de los tres recursos. El efecto país (en el nivel más alto de la jerarquía) pudo incluirse en el modelo como factor (efecto fijo), ya que se trata fundamentalmente de una variable de estratificación. No obstante, para que los resultados pudieran generalizarse a todos los sitios administrados

conjuntamente, habría que incluir los sitios como variable aleatoria más que como un efecto fijo. Los grupos de debate dentro de los diferentes sitios formarían parte también del modelo como efecto aleatorio. La esencia de los modelos en múltiples niveles está en la inclusión de una combinación de variables de efecto fijo y de variables de efecto aleatorio. Estos modelos hacen también posible la interacción entre las variables del sitio y las de los grupos de debate que se deben explorar.

72. Para ilustrar la manera como puede formularse un modelo en varios niveles para dar respuesta a las preguntas de una encuesta específica utilizamos un ejemplo tomado de una encuesta de producción y seguridad alimentaria realizada en Malawi en 2000-2001 (Levy y Barahona, 2001). El objetivo de la encuesta era evaluar un programa que trataba de aumentar la seguridad alimentaria entre los pequeños propietarios rurales mediante la distribución de recursos iniciales, en particular, fertilizantes y semillas de maíz y legumbres.

Ejemplo 2

73. La Encuesta de producción y seguridad alimentaria fue una encuesta nacional que utilizó un plan de muestreo estratificado en dos etapas y en la que los estratos estaban representados por los distritos. Se seleccionaron cuatro aldeas de cada uno de los 27 distritos de Malawi y unos 30 hogares por cada aldea. La selección de aldeas estuvo limitada a las que tenían más de 40 hogares (de forma que hubiera un número suficiente de hogares para garantizar que pudiera entrevistarse a los receptores de los recursos por distribuir) y menos de 250 hogares (para que el equipo pudiera realizar su trabajo dentro del tiempo previsto, de acuerdo con la disponibilidad de los recursos[7]). Respetando esta restricción, el muestreo en cada etapa se realizó en forma aleatoria. Durante la encuesta se visitaron 108 aldeas y 3.030 hogares en total.

74. Los datos que consideramos en el contexto de los modelos en varios niveles proceden de un cuestionario de hogares realizado durante la encuesta. El subconjunto de variables que consideraremos en nuestro ejemplo son el distrito, la aldea, el número de identificación del hogar, el sexo y la edad del jefe de hogar, el tamaño del hogar, el hecho de que el hogar hubiera recibido o no los recursos iniciales y dos índices que reflejaban los activos[8] y los ingresos[9] del hogar.

75. Hay varios modelos en distintos niveles que pueden encajarse con estos datos. Al formular el modelo, el primer paso es decidir qué variables son aleatorias y cuáles son los efectos fijos.

76. En el ejemplo 2, el distrito es una variable de estratificación y se consideraría como un efecto fijo. En general, todo efecto se califica como fijo si la repetición del proceso de muestreo da lugar al mismo conjunto de selecciones. Por el contrario, las aldeas de hogares se han seleccionado de forma aleatoria, por lo que constituyen los efectos aleatorios del modelo.

77. El modelo básico para analizar (por ejemplo) el índice de activos (IA) es

Modelo 5

$$y_{ijk} = \mu + d_k + U_{jk} + \varepsilon_{ijk}$$

donde d_k es el efecto de distrito (k = 1, 2, ..., 27), y los índices i y j corresponden al i^{o} hogar y a la j^{a} aldea, respectivamente. Algunas veces conviene considerar el parámetro de distrito como reflejo de la desviación del valor medio de IA para el distrito k a partir del valor medio global de los IA en todos los distritos. No obstante, los programas informáticos para la elaboración de modelos utilizan una parametrización diferente y consideran que uno de los efectos de distrito es cero. Los restantes efectos permiten comparar los IA de cada distrito con los IA del distrito cuyo efecto se ha fijado en cero.

[7] Esta restricción de la población objetivo limitó la posibilidad de inferencias a la población residente en las aldeas cuyo tamaño se incluyera dentro de ese intervalo.

[8] El índice de activos era un promedio ponderado basado en diferentes números de cabezas de ganado y activos de los hogares, como radio, bicicleta, carro de bueyes, etcétera.

[9] El índice de los ingresos estaba basado en los ingresos de diferentes procedencias.

78. En este modelo, U_{jk} y ε_{ijk} denotan variables aleatorias que representan, respectivamente, la variación entre todas las aldeas dentro del destrito k (que se supone es igual en todos los distritos), y la variación entre todos los hogares de la aldea j del distrito k (supuestamente igual para todas las combinaciones de aldeas y distritos). U_{jk} y ε_{ijk} son variables aleatorias y se supone que en el modelo son variables normalmente distribuidas con una media de cero y varianzas constantes σ_u^2 y σ_e^2, respectivamente. Se supone también que son independientes entre sí. Por ello, puede decirse que $U_{jk} \sim N(0, \sigma_u^2)$ y $\varepsilon_{ijk} \sim N(0, \sigma_e^2)$.

79. El ajuste de este modelo permite formular estimaciones de σ_u^2 y σ_e^2, y estimaciones de d_k, junto con los pertinentes errores estándar. Las estimaciones del parámetro para d_k ($k = 1, 2, ..., 27$) permiten una comparación de las medias de los IA en los 27 distritos.

80. Supongamos ahora que interesara investigar cómo se ha visto afectada la variación de los IA por el tamaño del hogar (variable cuantitativa) y el hecho de que el hogar hubiera recibido o no los recursos iniciales (una variable binaria). Estos aspectos se incluirían en el modelo de la misma manera que se haría en un modelo lineal general estándar. El modelo vendría dado por

Modelo 6

$$y_{ijk} = \mu + d_k + U_{jk} + t_{p(ijk)} + \beta\, x_{ijk} + \varepsilon_{ijk}$$

donde $t_{p(ijk)}$ representa el efecto correspondiente a la recepción de los recursos iniciales; x_{ijk} el tamaño del hogar, y β, la pendiente de la curva que describe la relación entre x_{ijk} e y_{ijk}; es decir, la relación del tamaño del hogar (HHSIZE) con el índice de activos (IA).

81. En este caso, tanto $t_{p(ijk)}$ como β se consideran como efectos fijos, mientras que U_{jk} y ε_{ijk} son (igual que antes) efectos aleatorios. La forma de este modelo supone que la relación de HHSIZE con los IA es la misma en todas las aldeas y distritos.

82. La inclusión de ambos componentes de la variación (aldea y hogar) en el modelo significa que éste tiene en cuenta la variabilidad en los dos niveles de la jerarquía. En otras palabras, los errores estándar asociados con $t_{p(ijk)}$ y β se calculan correctamente, como ocurriría con los resultados de las pruebas de significación asociadas con estos parámetros. No habría ocurrido así si el modelo lineal general no se hubiera ajustado considerando las aldeas como efectos fijos. Aun cuando se utilizara un modelo informático para encuestas que tuviera en cuenta las ponderaciones de muestreo, los procedimientos de la regresión estándar no tendrían en cuenta la estructura de correlación entre los hogares dentro de una aldea determinada.

83. El trato de las aldeas como efectos aleatorios tiene también otra ventaja importante. Si las aldeas se hubieran considerado como valores fijos, las conclusiones del análisis tendrían que aplicarse únicamente al conjunto de aldeas visitadas durante la encuesta. El considerar las aldeas como efectos aleatorios significa que las conclusiones referentes a la relación del tamaño del hogar con los IA, la comparación de los IA entre los hogares que reciben o no los recursos iniciales y la comparación entre los diferentes distritos pueden generalizarse para englobar la población general de todas las aldeas. La interacción entre el efecto fijo en el distrito d_k y el efecto de la recepción de los recursos iniciales $t_{p(ijk)}$ puede examinarse también incluyendo en el modelo dicho término de interacción.

84. Otro modelo útil se obtiene considerando el término de la pendiente β en el modelo 6 como variable aleatoria que varía en las distintas aldeas. Es lo que se conoce muchas veces como modelo de regresión de coeficiente aleatorio. El modelo se transforma en

Modelo 7

$$y_{ijk} = \mu + d_k + U_{jk} + t_{p(ijk)} + \beta_j\, x_{ijk} + \varepsilon_{ijk}$$

donde se supone que β_j es $N(\beta, \sigma_\beta^2)$. Además, como β_j es de carácter aleatorio en las diferentes aldeas, puede considerarse también que tiene una covarianza con U_{jk}, por ejemplo, $\sigma_{\beta u}$.

85. Así pues, en el análisis aquí presentado la comprobación de la hipótesis de que σ_β^2 es cero nos dice efectivamente si hay variabilidad en la pendiente de la relación IA-HH-SIZE en las aldeas. Si no se puede rechazar esta hipótesis, cabría concluir que la forma de la relación es la misma en todas las aldeas.

86. Es posible ampliar todavía más este modelo para incluir las variables relativas a las aldeas, por ejemplo, el acceso a abastecimiento de agua limpia o el grado de disponibilidad de asesoramiento de los agentes de extensión agrícola. En este caso se observan los beneficios reales de los modelos en múltiples niveles, pues sería posible explorar las relaciones entre dichas variables de las aldeas y las de los hogares. Por ello, el estudio de las relaciones entre variables en diferentes niveles de un plan de muestreo jerárquico resulta posible mediante la elaboración de modelos en múltiples niveles. La ventaja está en la posibilidad de tener en cuenta la estructura de correlación entre las unidades de nivel inferior cuando se analizan juntas variables de diferentes niveles. En el ejemplo anterior podrían considerarse otros modelos —por ejemplo modelos que incluyan el género y la edad del jefe de hogar— y las interacciones entre éstos y los términos antes incluidos en el modelo.

87. Hay, naturalmente, limitaciones asociadas con el ajuste de modelos en varios niveles. Lo mismo que con otros procedimientos semejantes, el modelo en varios niveles de nuestra hipótesis se supone que es "correcto" hasta un grado razonable y que se atiene al diseño de la muestra. Lógicamente, podría debatirse si estos supuestos son verdaderos o no.

H. Modelos en apoyo de los procesos de las encuestas

88. Aun cuando una encuesta de hogares se utilice estrictamente para obtener estadísticas descriptivas, es posible que se necesiten modelos para respaldar otros procesos de la encuesta. Los ajustes en caso de falta de respuesta muchas veces se basan directa o indirectamente en modelos estadísticos. En Groves y otros (2002, págs. 197-443) se examinan varios métodos para explicar la falta de respuesta, todos los cuales suponen cierto modelo estadístico. Los modelos de regresión logística pueden utilizarse para obtener propensiones de respuesta previsible, que permitirían ajustes en el caso de falta de respuesta, o para identificar clases de ponderación basadas en propensiones de respuestas semejantes (véanse, por ejemplo, Folsom (1991); Folsom y Witt (1994); o Folsom y Singh (2000)). Pueden utilizarse también modelos estadísticos predictivos en el marco del procedimiento para imputar los datos ausentes (véase, por ejemplo, Singh, Grau y Folsom (2002)). Finalmente, pueden utilizarse modelos estadísticos para evaluar los experimentos metodológicos insertos en las encuestas (véase, por ejemplo, Hughes y otros (2002)).

I. Conclusiones

89. Nuestro objetivo ha sido examinar las cuestiones implicadas en el análisis de los datos de encuestas. Entre ellas se incluyen el uso de ponderaciones y de métodos adecuados de estimación de la varianza, con planteamientos tanto descriptivos como analíticos. Se presenta también un panorama general de las situaciones prácticas en que las técnicas de elaboración de modelos pueden ayudar al análisis de datos. Se trata de instrumentos útiles, pero cuya aplicación requiere detenida reflexión y atención a sus supuestos subyacentes.

90. Hemos examinado el papel de las ponderaciones de la encuesta y el reconocimiento de la estructura de la muestra al elaborar estadísticas descriptivas y analíticas a partir de los datos. Los programas de análisis que utilizan ponderaciones de la encuesta y tienen en cuenta la estructura de la muestra pueden ser útiles para estimar los parámetros de los modelos de regresión tanto lineal como logística basados en los datos de encuestas. Las estimaciones basadas en la muestra son estimaciones de lo que se habría obtenido encajando los modelos con la población finita total. Además, pueden obtenerse errores estándar de las

estimaciones. Las variables explicativas en los modelos de regresión aplicados a los datos de las encuestas se observan casi siempre tal como existen en la población, más que mediante asignación aleatoria de acuerdo con un diseño experimental. Los encargados del análisis deben tener muy en cuenta que los coeficientes de regresión basados en los datos de encuestas simplemente reflejan la relación que existe entre la variable dependiente y las variables explicativas en la población y que no implican necesariamente una relación de causalidad. Hemos examinado cómo se relacionan los parámetros de regresión y los modelos de regresión logística con las estadísticas descriptivas simples y cómo pueden interpretarse en el caso de algunos modelos relativamente sencillos.

91. La elaboración de modelos en varios niveles, en particular, debería considerarse en general como una técnica más bien "avanzada", y es mejor aplicarla previa consulta con un estadístico familiarizado con su uso y sus limitaciones. En la actualidad, los modelos en varios niveles parecen utilizarse sólo en casos aislados para analizar las encuestas de países en desarrollo; no obstante, su uso sería muy conveniente, porque permiten conocer mejor las interrelaciones entre las variables en diferentes niveles y por su capacidad de tener en cuenta la variabilidad entre unidades de muestreo en diferentes niveles en un diseño por etapas.

92. Hemos comprobado que la formulación de modelos en varios niveles no es demasiado difícil para quien está familiarizado con la aplicación de modelos lineales generales (MLG), pero, como en otras ocasiones, hay algunos supuestos asociados con los modelos que deben comprobarse realizando análisis residuales, como ocurriría con los MLG. La elaboración de modelos en varios niveles puede realizarse también cuando la principal respuesta de interés es binaria, aunque no hayamos presentado un ejemplo de tal caso. Hay que proceder también con cuidado al decidir qué efectos son aleatorios y cuáles son fijos y cómo la especificación del modelo ayudará a responder a los objetivos específicos de la encuesta.

93. Como ocurre con todas las técnicas de estadística, los métodos de elaboración de modelos examinados en este capítulo presentan varias limitaciones que es preciso reconocer en su aplicación. Hemos aconsejado el uso de ponderaciones de las encuestas y programas de análisis que reconozcan la estructura de diseño de la muestra. Hay que reconocer la dificultad de contar con programas informáticos adecuados que tengan en cuenta el diseño muestral. En el capítulo XXI se describen varios programas que tienen en cuenta los problemas del diseño con respecto a los procedimientos de regresión múltiple y logística. No obstante, por desgracia estos programas no tienen opciones para el ajuste de modelos en varios niveles. Para ello, el usuario debe recurrir a programas estadísticos de alcance más general, como SAS (2001), GenStat (2002) y SPSS (2001), o a un programa especializado como *MLwiN* (Rahbash y otros, 2001).

94. También se han presentado algunas técnicas de elaboración de modelos que pueden servir como instrumentos útiles para el análisis de los datos de encuestas. Recomendamos que los analistas e investigadores consideren seriamente estos métodos, cuando estén en consonancia con los objetivos de la encuesta, durante el análisis de los datos, a fin de extraer la mayor información posible de unos datos cuya recopilación es por fuerza costosa.

Reconocimientos

Los autores desean manifestar su especial agradecimiento a los revisores y editores por sus útiles comentarios y sugerencias, y en particular al Dr. Graham Kalton por las mejoras introducidas en la exposición sobre la ponderación en las encuestas, así como al Departamento de Desarrollo Internacional (DFID) del Gobierno del Reino Unido de Gran Bretaña e Irlanda del Norte, cuya financiación de muchos proyectos interesantes relacionados con encuestas en el mundo en desarrollo ha servido para que de ellos extraigamos ideas interesantes que se han incluido en este capítulo. En cualquier caso, el material del presente capítulo es responsabilidad exclusiva de los autores y no expresa de ninguna manera las opiniones del DFID.

REFERENCIAS BIBLIOGRÁFICAS

Banco Mundial (2000). *Voices from the Village: A Comparative Study of Coastal Management in the Pacific Islands*. Pacific Islands Discussion Paper Series, No. 9. Washington, D.C.: World Bank. Papua New Guinea and Pacific Islands Country Management Unit.

Chambers, R. L. y C. J. Skinner (2003). *Analysis of Survey Data*. Chichester, Reino Unido: Wiley.

Chromy, James R. (1998). *The Effects of Finite Sampling on State Assessment Sample Requirements*. Palo Alto, California: NAEP Validity Studies, American Institutes for Research.

Cochran, W. G. (1977). *Sampling Techniques,* 3a. ed. Nueva York: John Wiley and Sons.

Congdon, P. (1998). A multi-level model for infant health outcomes: maternal risk factors and geographic variation. *The Statistician*, vol. 47, Parte 1, págs. 159-182.

Deville, J. C. y C. E. Särndal (1992). Calibration estimating in survey sampling. *Journal of the American Statistical Association,* vol. 87, págs. 376-382.

Folsom, Ralph E., Jr. (1991). Exponential and logistic weight adjustments for sampling and non-response error reduction. En *Proceedings of the Social Statistics Section, American Statistical Association*. Alexandria, Virginia: American Statistical Association, págs. 376-382.

Folsom, Ralph E. y Michael B. Witt (1994). Testing a new attrition non-response adjustment method for SIPP. En *Proceedings of the Survey Research Methods Section, American Statistical Association*, págs. 428-433.

Folsom, R. E. y A. C. Singh (2000). The general exponential model for sampling weight calibration for extreme values, non-response and and post-stratification. En *Proceedings of the Survey Research Methods Section, American Statistical Association*. Indianapolis, Indiana.

GenStat (2002). *GenStat for Windows,* 6a. ed. Oxford, Reino Unido: VSN International, Ltd.

Goldstein, H. (2003). *Multi-level Statistical Models,* 3a. ed. Londres: Arnold.

_____ y otros (1993). A multi-level analysis of school examination results. *Oxford Review of Education*, vol. 19, págs. 425-433.

Graubard, B. I. y E. L. Korn (2002). Inferences for super-population. *Statistical Science,* vol. 17, págs. 73-96.

Groves, Robert M. y otros (2002). *Survey non-response.* Nueva York: John Wiley and Sons, Inc.

Horvitz, D. G. y D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. En *Journal of the American Statistical Association,* vol. 47, págs. 663-685.

Hughes, Arthur y otros (2002). Impact of inteviewer experience on respondent reports of substance use. En *Redesigning an Ongoing National Household Survey: Methodological Issues,* J. Gfroerer, J. Eyerman y J. Chromy, comps. DHHS Publication, No. SMA 03-3768. Rockville, Maryland: Substance Abuse and Mental Health Services Administration, Office of Applied Studies, págs. 161-184.

Kish, Leslie (1965). *Survey Sampling*. Nueva York: John Wiley and Sons, Inc.

Korn, E. L. y B. I. Graubard (1999). *Analysis of Health Surveys*. Nueva York: John Wiley and Sons.

_____ (2003). Estimating variance components by using survey data. En *Journal of the Royal Statistical Society B,* vol. 66, págs. 175-190.

Kreft, I. y J. de Leeuw (1998). *Introducing Multi-level Modeling*. Londres.

Langford, I. H., G. Bentham y A. McDonald (1998). Multi-level modeling of geographically aggregated health data: a case study on malignant melanoma mortality and UV exposure in the European Community. *Statistics in Medicine*, vol. 17, págs. 41-58.

Levy, S. y C. I. Barahona (2001). The targeted inputs programme, 2000-2001: Main report. Inédito.

Pfeffermann, D. y otros (1998). Weighting for unequal selection probabilities in multi-level models. En *Journal of the Royal Statistical Society B*, vol. 60, págs. 23-40.

Rasbash, J. y otros (2001). *MLwiN Version 1.10.0007*. Multi-level Models Project. Londres: Institute of Education, University of London.

Research Triangle Institute (2002). *SUDAAN User's Manual*, *Release 8.0*. Research Triangle Park, North Carolina: Research Triangle Institute.

Rubin, Donald B. (1987). *Multiple Imputation for Non-response in Surveys*. Nueva York: John Wiley and Sons.

SAS (2001). *SAS Release 8.2*. Cary, North Carolina: SAS Institute, Inc., SAS Publishing.

Singh, Avinash, Eric Grau y Ralph Folsom Jr. (2002). Predictive mean neighborhood imputation for NHSDA substance use data. En *Redesigning an Ongoing National Household Survey: Methodological Issues,* J. Gfroerer, J. Eyerman y J. Chromy, comps. DHHS publication, No. SMA 03-3768. Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Skinner, C. J., D. Holt y T. M. F. Smith, comps. (1989). *Analysis of Complex Surveys*. Nueva York: Wiley.

Snijders, T. A. B. y R. J. Bosker, R. J. (1999). *Multi-level Analysis: An Introduction to Basic and Advanced Multi-level Modelling*. Londres: Sage.

SPSS (2001). *SPSS for Windows. Release 11.0.* Chicago, Illinois: LEAD Technologies, Inc.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley.

Woodruff, R. S. (1971). A simple method for approximating the variance of a complicated sample. En *Journal of the American Statistical Association,* vol. 66, págs. 411-414.

Capítulo XX
Planteamientos más avanzados para el análisis de los datos de encuestas

Gad Nathan
Universidad Hebrea
Jerusalén
Israel

RESUMEN

En este capítulo consideramos los efectos del diseño de muestras complejas utilizado en la práctica en la mayoría de las encuestas por muestreo para el análisis de datos de encuestas. También se especifican los casos en que el diseño puede influir o no en el análisis y se definen los conceptos básicos. Una vez que se ha establecido el modelo para el análisis, consideramos las posibles relaciones entre el modelo y el diseño de la muestra. Cuando el diseño puede influir en el análisis y no se puede agregar al modelo analítico una variable explicativa adicional relativa al diseño, pueden utilizarse dos metodologías básicas: el análisis clásico, que podría modificarse para tener en cuenta el diseño, o un nuevo instrumento analítico, que podría elaborarse para cada diseño. Los distintos planteamientos se acompañan con aplicaciones de datos reales a la regresión lineal, los modelos lineales y el análisis de datos categóricos.

Términos clave: Diseño muestral complejo, análisis de datos de encuestas, regresión lineal, modelos lineales, análisis de datos categóricos, análisis basado en modelos.

A. Introducción

1. Diseño muestral y análisis de datos

1. El objetivo principal de la inmensa mayoría de las encuestas por muestreo, tanto en los países en desarrollo como en los desarrollados, es de carácter descriptivo; es decir, ofrecer estimaciones puntuales y de intervalo de las medidas descriptivas de una población finita, como medias, medianas, distribuciones de frecuencia y tabulaciones cruzadas de variables cualitativas. No obstante, como se demuestra en los capítulos XV-XIX y se observará de nuevo en el capítulo XXI, hay cada vez mayor interés en realizar inferencias acerca de las relaciones entre las variables investigadas, en vez de limitarse simplemente a describir los fenómenos.

2. En este capítulo intentaremos evaluar los efectos que producen en el análisis de los datos de encuesta los diseños muestrales complejos habitualmente utilizados. Trataremos de identificar los casos en que el diseño puede influir en el análisis. En general, el diseño muestral no influye en el análisis cuando las variables en que está basado el diseño se incluyen en el modelo analítico. No obstante, con frecuencia algunas variables del diseño no se inclu-

yen en el modelo, debido a errores de especificación o a la falta de interés en esas variables de diseño en cuanto factores explicativos. Ello puede producir graves sesgos.

3. Hay dos metodologías básicas para manejar datos a partir de una muestra compleja cuando no se agregan al análisis variables adicionales relacionadas con el diseño. La primera modifica un instrumento analítico básico elaborado para la manipulación de datos de una muestra aleatoria simple. La segunda metodología establece un nuevo instrumento analítico para diseños complejos específicos.

4. A continuación presentaremos algunos ejemplos de los posibles efectos del diseño muestral en el análisis, definiremos algunos conceptos básicos y examinaremos el papel de los efectos del diseño en el análisis de datos muestrales complejos. En el apartado B se describen los dos planteamientos básicos del análisis de datos muestrales complejos. En los apartados C y D consideramos algunos ejemplos relativos al análisis de datos continuos y categóricos, respectivamente. El apartado final contiene un resumen y algunas conclusiones. En el anexo se presentan algunas definiciones formales y resultados técnicos.

2. Ejemplos de los efectos (y de la ausencia de efectos) del diseño muestral en el análisis

5. Con el fin de poner de manifiesto los posibles efectos del diseño muestral en el análisis, consideraremos un ejemplo sencillo pero ilustrativo (puede verse información más detallada en Nathan y Smith, 1989). Supongamos que Y es la variable de interés y que X es una variable auxiliar. Supongamos también que se puede aplicar a la población el modelo de regresión lineal $Y_i = \alpha + \beta X_i + \varepsilon_i$, donde

$$\varepsilon_i | X_i \underset{ind}{\sim} \mathrm{N}(0, \sigma_i^2)$$

El modelo es válido también para cualquier muestra aleatoria simple que se seleccione de la población. Algunas veces el supuesto de independencia entre $\varepsilon_i | X_i$ se adapta mejor a la muestra aleatoria simple que a la población de la que se ha tomado. En una población humana, por ejemplo, los valores ε_i pueden estar correlacionados con diferentes miembros del mismo hogar, mientras que en una muestra aleatoria simple de individuos, con una probabilidad muy pequeña de que se seleccione más de un individuo por hogar, la correlación sería insignificante.

6. En el contexto del muestreo aleatorio simple la estimación estándar del coeficiente de regresión no presenta sesgos. En la figura XX.1 se representa Y con respecto a X para la población total; la presentación sería la misma para una muestra aleatoria simple. En las cinco presentaciones de las figuras XX.2 a XX.6 las muestras se seleccionan a partir de la población utilizando métodos muy diferentes del muestreo aleatorio simple. Consideremos la selección de la muestra basada totalmente en el valor de X, por ejemplo, truncando los puntos de los datos con valores X más allá (o dentro) de los límites fijos, como en las figuras XX.2 y XX.3. Se deduce claramente de estas figuras que la selección no tiene efecto alguno en la estimación de los parámetros de regresión de la intersección (α) y la pendiente (β) (aunque podría influir en la varianza de los estimadores).

7. Consideremos ahora la selección muestral basada en los valores de la variable objetivo, Y, por ejemplo, truncando los puntos de los datos con valores Y más allá (o dentro) de los límites fijos, como en las figuras XX.4 a XX.6. En estos casos, es claro que las estimaciones de las pendientes de la regresión resultan sesgadas. En el último caso (figura XX.6), la truncación no es simétrica, y la estimación de la intercepción está también sesgada. Se trata de ejemplos extremos, ya que la selección basada en la truncación de la variable dependiente es rara en las encuestas por muestreo. En cambio, es muy frecuente en los estudios experimentales o de observación, como los estudios de control de casos en epidemiología o los

Figura XX.1
Ninguna selección

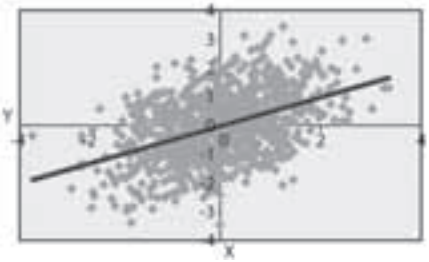


Figura XX.2
Selección en X: XL<X<XU

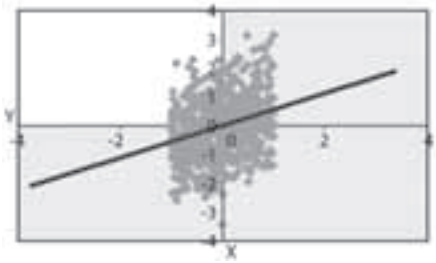


Figura XX.3
Selección en X: X<XL; X>XU

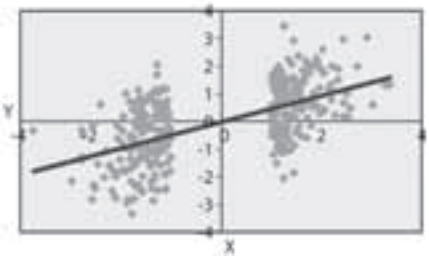


Figura XX.4
Selección en Y: YL<Y<YU

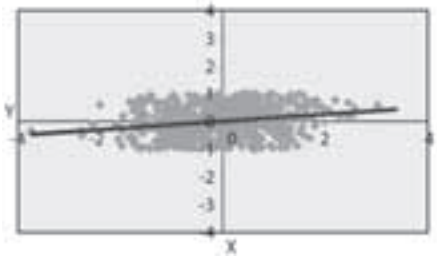


Figura XX.5
Selección en Y: Y<YL; Y>YU

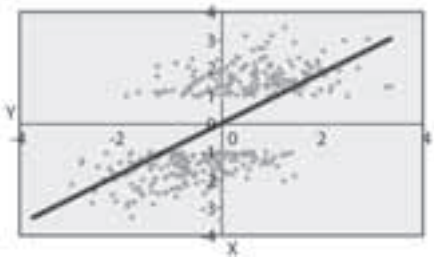
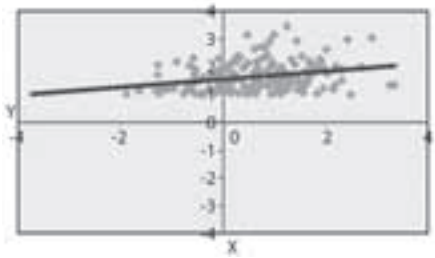


Figura XX.6
Selección en Y: Y>U



estudios basados en opciones en la economía (véanse, por ejemplo, Scott y Wild (1986) y Manski y Lerman (1977)). De todas formas, en muchos casos las muestras se seleccionan en función de variables de diseño que pueden estar estrechamente relacionadas con la variable dependiente. Así, un método de muestreo común muy utilizado en las encuestas de establecimientos y de explotaciones agrícolas consiste en seleccionar unidades con probabilidad proporcional al tamaño. La medida del tamaño —por ejemplo, la producción del año anterior— estará obviamente relacionada con la variable pertinente cuando dicha variable sea la producción del año en curso. Las estimaciones estándar de los parámetros del modelo, como los coeficientes de regresión, pueden estar sesgadas cuando no se tiene en cuenta el diseño de la muestra.

8. Los ejemplos precedentes ilustran los peligros que implica la realización de un análisis basado en datos de muestras complejas como si procedieran de un muestreo aleatorio simple revelan la necesidad de determinar cuándo es probable que el diseño repercuta en el análisis y de tener en cuenta el diseño, cuando así ocurra.

3. Conceptos básicos

9. La mayor parte de las encuestas por muestreo se conciben fundamentalmente con fines descriptivos (o enumerativos). Tratan de estimar los valores de los parámetros de la po-

blación finita, como el ingreso mediano de los hogares o la proporción de todos los adultos que han contraído el SIDA. Se trata de estadísticas que en principio pueden medirse exactamente si se incluye en la encuesta a toda la población, es decir, si se realizara el empadronamiento de la población y no de una muestra aleatoria de la misma. La teoría comúnmente aceptada sobre el muestreo da por supuesto que los datos de una muestra aleatoria pueden utilizarse para ofrecer estimaciones no sesgadas de los parámetros de la población finita y de sus errores de muestreo, por muy complejo que sea el diseño de la muestra. Se presupone en tal caso que se trata de un diseño muestral probabilístico, es decir, que cada unidad de la población tiene una probabilidad positiva conocida de figurar en la muestra. Estos modelos básicos de estimación de los parámetros de población finita se conocen como métodos basados en el diseño (o basados en la aleatorización), ya que toda inferencia está basada en las propiedades del diseño de la muestra, a través de la distribución probabilística de la muestra. No obstante, conviene señalar que la eficiencia de las diferentes estrategias de estimación (un diseño muestral junto con una fórmula de estimación) sólo puede evaluarse normalmente cuando se dispone de amplia información sobre la población. Este requisito no suele darse en la práctica. Así, incluso los textos de muestreo clásicos (por ejemplo, Cochran, 1977) se suelen basar en modelos para justificar métodos específicos de muestreo o estimación. Si los valores de la población se deducen de un modelo de regresión simple, por ejemplo, el estimador del coeficiente será más eficiente que el estimador de expansión simple, con determinados supuestos. Así pues, los métodos basados en el diseño muchas veces cuentan con la ayuda de modelos, pero no están basados ni dependen de ellos: en los métodos de diseño muestral y estimación de estadísticas descriptivas basados en modelos no se necesitan supuestos sobre el modelo para la estimación prácticamente no sesgada de parámetros de población finita.

10. El planteamiento del diseño y estimación basado en modelos (o teoría predictiva) supone que los valores de la población finita son plasmaciones de una distribución de la superpoblación basada en un modelo hipotético con parámetros de la superpoblación (modelo). En Brewer y Mellor (1973), Hansen, Madow y Tepping (1983), Särndal, Swensson y Wretman (1992) y Valliant, Dorfman y Royall (2000) puede verse un examen más detallado. A diferencia de los métodos de estimación con ayuda de modelos, la mayor eficiencia conseguida por los métodos de estimación basados en modelos depende de la validez del modelo supuesto. Así pues, si no hay ninguna idea de la validez de los supuestos acerca del modelo, la reducción aparente del cuadrado medio quizá no justifique el uso de un análisis puramente basado en modelos. Puede verse un ejemplo ilustrativo en Hansen, Madow y Tepping (1983), donde el supuesto de que no hay ninguna intercepción en un modelo de regresión, aun cuando dicha intercepción sea de hecho muy próxima a cero, da lugar a una inferencia inválida basada en el modelo.

11. Cada vez más, las encuestas se están utilizando con fines analíticos, además de descriptivos. Muchas veces, al formular las encuestas se tiene en cuenta su utilización para futuros análisis. Ello se debe a que las autoridades y los investigadores tienen interés en los procesos en que se basan los datos brutos al elaborar los modelos sobre las relaciones entre las variables investigadas. Dichos análisis requieren la formulación de supuestos sobre los modelos. El objetivo de un análisis es confirmar la validez de un modelo supuesto y estimar los parámetros del modelo más que los parámetros de la población finita. Así pues, el análisis está intrínsecamente basado en modelos. La inferencia sobre los parámetros del modelo, casi por definición, debe estar basada en modelos pertinentes.

12. Conviene señalar que cuando la población es muy numerosa y los modelos de la hipótesis son realmente válidos, en la práctica hay muy poca diferencia entre los parámetros del modelo y sus equivalentes de la población finita. Por ejemplo, si es válido el modelo de regresión lineal $Y_i = \alpha + \beta X_i + \varepsilon_i$, donde

$$\varepsilon_i | X_i \underset{ind}{\sim} \mathrm{N}(0, \sigma_i^2)$$

y el tamaño de la población es muy grande, el valor del coeficiente de regresión de la población estándar, B (véase el anexo), será muy próximo al valor del parámetro del modelo, β, debido al teorema del límite central. Así pues, aunque a continuación nos centraremos en la estimación de los parámetros del modelo, estos parámetros se sustituirán algunas veces por equivalentes de la población finita. Para mayor claridad en la presentación, la mayor parte de los ejemplos que siguen se formularán en términos de distribuciones de una sola variable (es decir, una sola variable dependiente y una sola variable explicativa). La ampliación de los resultados a un caso con variables múltiples suele ser bastante sencilla.

13. En resumen, los modelos hipotéticos forman parte integrante del análisis estadístico. Con el fin de analizar los datos de las encuestas por muestreo, la elección de un modelo adecuado para ajustar los datos forma parte decisiva del análisis. Antes de aplicar métodos analíticos, los investigadores y analistas deben comprender cabalmente los modelos en que se basan los procesos que desean estudiar, . Como veremos en las secciones siguientes, la aplicación a los datos procedentes de diseños muestrales complejos requiere comprender tanto el modelo subyacente como de qué forma puede verse afectado el análisis por un diseño complejo.

4. Efectos del diseño y su papel en el análisis de datos muestrales complejos

14. Los temas de los efectos del diseño y su estimación se han tratado con detenimiento en los capítulos VI y VII, sobre todo en lo que respecta a su función en el diseño y estimación de las encuestas enumerativas. En este capítulo veremos que son también importantes para el análisis de datos procedentes de encuestas con muestras complejas. La idea básica reside en el hecho de que, suponiendo que sean válidos los supuestos del modelo, en un contexto de muestreo aleatorio simple pueden obtenerse fácilmente estimaciones no sesgadas de los parámetros del modelo y estimaciones de las varianzas de dichas estimaciones. Esas estimaciones y las estimaciones de la varianza constituyen la base para comprobar las hipótesis relativas a los parámetros del modelo. Por ejemplo, con un muestreo aleatorio simple y suponiendo el modelo de regresión simple $Y_i = \alpha + \beta X_i + \varepsilon_i$ donde

$$\varepsilon_i \Big| X_i \underset{ind}{\sim} \mathrm{N}(0, \sigma_i^2)$$

el estimador de la muestra de mínimos cuadrados ordinarios, b, es un estimador no sesgado de β, y se dispone de un estimador no sesgado de su varianza, $v(b)$ (véase el anexo). Luego la prueba estándar de la hipótesis nula de que β = 0 se basa en el estadístico de prueba $b/\sqrt{v(b)}$ invocando el teorema del límite central. Cuando el diseño de la muestra es complejo —por ejemplo, una muestra conglomerada y estratificada—, el estimador b no está sesgado por el modelo, si el modelo de regresión es válido y el diseño de la muestra no depende de los valores Y_i (por ejemplo, en contraste con la situación de las figuras XX.4 a XX.6). Con ello queremos decir que con un modelo de regresión especificado el valor previsto de b es β, donde la expectativa es con respecto a la distribución de la superpoblación de los valores Y_i. Como veremos en el subapartado C.1, ello quizá no sea cierto si el modelo no es válido. No obstante, aun cuando se confirmen los supuestos del modelo, $v(b)$ ya no es un estimador válido de la varianza del modelo de b y debe modificarse. Muchas veces puede calcularse un estimador directo de la varianza del modelo —por ejemplo, utilizando uno de los programas informáticos descritos en el capítulo XXI— y puede utilizarse para sustituir $v(b)$. Si no se dispone de un estimador directo, frecuentemente se obtiene un estimador válido del efecto del diseño, denotado por $d^2(b)$. Éste puede utilizarse para la modificación del estadístico de prueba para sustituir $v(b)$ por $d^2(b)$ x $v(b)$. Más adelante presentaremos otros usos específicos de los efectos del diseño para modificar los estadísticos de pruebas estándar para otras aplicaciones.

B. Planteamientos básicos del análisis de datos muestrales complejos

1. La especificación del modelo, como base del análisis

15. La especificación correcta del modelo subyacente es un paso fundamental en todo análisis. Las consecuencias de una especificación errónea del modelo —tanto de la exclusión de las variables explicativas pertinentes (o la inclusión de superfluas) como de la utilización de una forma funcional errónea (por ejemplo, lineal en vez de cuadrática)— son bien conocidas y están ampliamente documentadas en los textos clásicos. Pueden adoptar la forma de sesgos en la estimación de los parámetros del modelo (principalmente, la exclusión de las variables pertinentes), pérdidas de eficiencia (vinculadas sobre todo con la inclusión errónea de variables explicativas) y alteración de los tamaños y potencia en las pruebas de las hipótesis. Estos efectos pueden exacerbarse cuando la especificación errónea está directamente relacionada con las variables del diseño muestral o con variables correlacionadas con las variables del diseño. No obstante, es importante caer en la cuenta de que las variables del diseño de la muestra quizá no sean pertinentes para los objetivos de la investigación. Además, quizá no haya una justificación de fondo para su inclusión en el modelo analítico.

16. Hay dos planteamientos básicos para incorporar las variables del diseño de la encuesta en el modelo. El planteamiento agregado considera que el modelo de interés se encuentra en el nivel de la población y es conceptualmente independiente del diseño muestral empleado para obtener los datos. Con este planteamiento, las variables del diseño se incluirán en el modelo sólo cuando sean pertinentes para el análisis del contenido en cuestión. Por ejemplo, supongamos que deseamos explicar la variable binaria empleada/no empleada por la variable explicativa años de educación, independientemente de la ubicación geográfica. La muestra está estratificada, por ejemplo, por regiones geográficas, en cuyo caso podrían ser pertinentes modelos distintos. Así ocurriría aun cuando se hubiera utilizado el muestreo aleatorio simple. En consecuencia, la estratificación puede incluirse en el modelo (véase el planteamiento desglosado que se describe más adelante) para reflejar las variaciones regionales en las relaciones entre las variables del modelo. Si, por el contrario, la estratificación y la asignación muestral a los estratos se llevaran a cabo simplemente por razones prácticas (comodidad o costo), las ponderaciones de la muestra probablemente no serían pertinentes para el modelo de población. La incorporación de ponderaciones de muestreo en un análisis por lo demás libre de los efectos del estrato dará lugar a cierta pérdida de eficiencia. No obstante, podría someterse a la fácil interpretación de un modelo libre de efectos del estrato, mientras que resistiría a la quiebra del modelo si existen realmente efectos de estrato ignorados.

17. El planteamiento desglosado amplía el modelo del analista para incluir no sólo las variables de la encuesta pertinentes sino también variables utilizadas en el diseño de la encuesta y las relativas a la estructura de la población que se refleja en el diseño. Las variables de diseño relativas a la estratificación y conglomeración se incluyen en el modelo para reflejar la estructura compleja de la población. Por ejemplo, en el caso anterior, el modelo contendría un conjunto diferente de coeficientes (tanto una intersección como una pendiente) para cada estrato geográfico. La inferencia en el marco del planteamiento desglosado tiene plenamente en cuenta el diseño de la muestra, suponiendo que todas las variables del diseño se incluyan correctamente en el modelo. El gran número de parámetros que deben estimarse en este planteamiento puede provocar dificultades y dar lugar a estimaciones menos precisas en comparación con las de modelos agregados más limitados. El planteamiento desglosado es adecuado únicamente cuando el analista considera que el modelo de la hipótesis puede ayudarle a conseguir sus objetivos.

18. El planteamiento adecuado que deberá utilizarse —agregado o desglosado— dependerá de los objetivos del analista. El planteamiento agregado es más adecuado para

estudiar factores que afectan al conjunto de la población, y por lo tanto puede ser más útil para evaluar medidas normativas de alcance nacional. El planteamiento desglosado es más aconsejable para estudiar los microefectos y los efectos de las decisiones de alcance local y relativas a sectores específicos. Pueden verse más ejemplos y un examen más detallado en Skinner, Holt y Smith (1989) y Chambers y Skinner (2003).

2. Posibles relaciones entre el modelo y el diseño muestral: diseños ilustrativos y no ilustrativos

19. Es importante establecer una distinción entre diseños muestrales ilustrativos y no ilustrativos al analizar los datos de encuestas complejas. Una vez que se ha hipotetizado un modelo, el analista debe considerar si, después de establecer las condiciones sobre las covariantes del modelo, las probabilidades de selección de la muestra están relacionadas con los valores de variable de respuesta. Un proceso de muestreo es ilustrativo si la distribución conjunta del modelo condicional de las observaciones para la muestra, dados los valores de las covariantes del modelo, difiere de la distribución condicional de la población. Únicamente cuando estas distribuciones son idénticas el diseño de la muestra resulta no ilustrativo (es decir, se podría ignorar), en cuyo caso los métodos analíticos estándar pueden emplearse como si las observaciones procedieran de una muestra aleatoria simple. Cuando el diseño de la muestra es ilustrativo, el modelo que se mantiene para los datos de la muestra es diferente del modelo de la población. Si no se tiene en cuenta el proceso de muestreo en tal caso pueden obtenerse estimadores puntuales sesgados, con lo que se distorsionaría el análisis, lo mismo que cuando se excluyen las variables del modelo en una análisis convencional. Téngase en cuenta que la inclusión en el modelo de variables relacionadas con el diseño dará lugar a que éste no sea ilustrativo.

20. Existen dos grandes problemas cuando se incluyen en un modelo todas las variables relacionadas con el diseño. En primer lugar, quizá no se sabe cuáles son exactamente las variables utilizadas en el diseño o, si se saben, quizá no se dispone de sus valores. Aun cuando las variables del diseño estén identificadas y cuantificadas, el analista quizá no sepa la forma exacta de la relación (por ejemplo, lineal o exponencial) entre ellas y la variable pertinente. Por ejemplo, si el diseño es estratificado, deberá comprobarse la posibilidad de que una relación de regresión tenga diferentes pendientes e intersecciones para los diferentes estratos.

21. En segundo lugar, cuando las variables del diseño se incluyen correctamente en el modelo, las estimaciones resultantes pueden ser de poco valor para el analista, ya que las variables agregadas no son de interés intrínseco desde el punto de vista del contenido (recuérdese lo que se ha dicho sobre el análisis agregado y desglosado). Ello significa que el efecto del diseño muestral complejo en el análisis no siempre se puede corregir limitándose a modificar el modelo subyacente. A continuación consideraremos la manera de modificar los métodos analíticos estándar para tener en cuenta un diseño complejo y la forma de desarrollar métodos de estimación y análisis que sean específicos para diseños especiales.

3. Problemas en la utilización de los programas informáticos estándar para el análisis de muestras complejas

22. El uso prácticamente universal de programas informáticos para el análisis estadístico ha dado lugar a un abuso generalizado de una práctica estadística válida. Este abuso se agrava con frecuencia cuando se analizan datos de encuestas con muestras complejas.

23. Las ventajas de los programas informáticos de estadística como medio de facilitar el análisis van acompañadas, por desgracia, de posibilidades de realizar análisis sin una comprensión elemental de los principios básicos pertinentes. Este problema ha llegado a ser grave en los estudios cuantitativos, sobre todo en las ciencias sociales. La situación se com-

plica por el hecho de que los programas de más fácil acceso tratan los datos como si fueran resultado de un muestreo aleatorio simple. Como se ha señalado antes, ello puede dar lugar a inferencias gravemente sesgadas cuando el diseño de la muestra es de carácter informativo. No obstante, si se procede con el debido cuidado, muchas veces los programas estándar pueden adaptarse para reflejar aproximadamente o explicar el efecto de un diseño complejo. En particular, el procedimiento SURVEYREG de la versión más reciente de SAS© (versiones 8 y 9) permite un análisis de regresión que tiene en cuenta el diseño de la muestra en forma semejante a como se describe más adelante (véase An y Watts (2001)).

24. Por ejemplo, consideremos el modelo de regresión lineal (heteroscedástico) definido por: $Y_i = \alpha + \beta X_i + \varepsilon_i$, con

$$\varepsilon_i | X_i \underset{ind}{\sim} \mathrm{N}(0, \sigma_i^2)$$

Los programas informáticos estándar normalmente calculan b, el estimador de cuadrados mínimos ordinarios (CMO) de β, o el estimador de los cuadrados mínimos generalizados (CMG), b_G, donde las sumas y productos se ponderan mediante los recíprocos de σ_i^2, cuyos valores (o valores relativos) se supone que son conocidos (véase el anexo, página 386). Ambos son estimadores no sesgados para el parámetro β, si se mantiene el modelo, aunque b_G es un estimador más eficiente en el caso heteroscedástico. Los programas estándar prevén también estimaciones de las varianzas del estimador CMO, $v(b)$, y del estimador CMG, $v(b_G)$, que no presentan sesgos si se utiliza el modelo adecuado (el modelo homoscedástico en el caso de $v(b)$).

25. En muchos casos puede haber duda sobre la validez del modelo, por lo que en vez de estimar β quizá sea más adecuado estimar la contraparte de la población finita de β, que designamos como B (véase el anexo). Aunque b (el estimador CMO) es un estimador no sesgado por el modelo para $\beta \approx B$, no se puede decir en general que no esté sesgado por el diseño. El estimador ponderado por la muestra (Horvitz-Thompson), b_W, con productos cruzados y cuadrados ponderados por los recíprocos de las probabilidades de inclusión, es coherente con el diseño y no está sesgado por el modelo, si se dan las condiciones adecuadas. Además, b_W puede obtenerse a partir de las opciones de regresión ponderada de muchos programas estándar utilizando los valores de w_i como ponderaciones. Alternativamente, puede obtenerse b_W mediante regresión no ponderada de las variables transformadas $Y_i/\sqrt{\pi_i}$, $X_i/\sqrt{\pi_i}$ en que la intersección se sustituye por $1/\sqrt{\pi_i}$. No obstante, conviene insistir en que en estas dos alternativas las estimaciones de la matriz de varianza-covarianza registradas por la mayoría de los programas estándar son incorrectas —en calidad tanto de estimadores para el error cuadrático medio del diseño como de estimadores para la varianza del modelo—, salvo en circunstancias excepcionales.

26. En resumen, el uso de programas informáticos estándar que no tengan en cuenta el diseño de las encuestas complejas no es recomendable a no ser que pueda comprobarse que el diseño complejo no tiene efectos graves en la estimación. Ello puede conseguirse muchas veces mediante el uso adecuado de programas estándar. Véase el ejemplo del subapartado C.2. Se recomienda el uso de programas informáticos en que se tienen en cuenta expresamente diseños muestrales complejos (véase el capítulo XXI).

C. Análisis de regresión y modelos lineales

1. Efecto de las variables de diseño no incluidas en el modelo y estimadores ponderados de regresión

27. El análisis de regresión y la elaboración de modelos lineales son aplicaciones muy comunes cuando habitualmente se aplican modelos estándar concebidos para muestras aleatorias simples a datos precedentes de encuestas con muestras complejas. Como ya se ha

señalado, ello puede dar lugar a análisis y conclusiones erróneas frecuentes. Una forma de prevenir el error es la identificación de variables que determinen o condicionen el diseño de la muestra, de manera que puedan incluirse en el modelo. No obstante, como hemos visto, incluso cuando se identifican estas variables, su inclusión en el modelo quizá no esté justificada desde el punto de vista del contenido. En este subapartado estudiaremos los efectos producidos en los estimadores tradicionales por la no inclusión de variables de diseño en el modelo e investigaremos las posibilidades de modificar esos estimadores a fin de tener en cuenta el diseño complejo. Para facilitar la exposición, consideraremos el caso de una única variable dependiente Y (designada aquí X_1 por razones técnicas), en el que el modelo pertinente tiene una única variable explicativa (X_2) y hay una única variable del diseño (X_3). Así pues, el modelo de interés es $E(E_1) = \mu_1 + \beta_{12}(X_2 - \mu_2)$, donde β_{12} es el parámetro pertinente, más que el modelo completo, que incluye la variable de diseño, X_3. Véanse en el anexo las fórmulas utilizadas en este subapartado.

28. En las condiciones bastante generales especificadas en Nathan y Holt (1980), el estimador CMO estándar para β_{12}, $b_{12} = s_{12}/s_2^2$, puede estar sesgado (por el modelo) y condicionado a X_3 y a la muestra, S, o sesgado de forma incondicional. Las expresiones correspondientes a la expectativa del modelo condicional y su expectativa no condicional (modelo y diseño conjunto) revelan que, en general, b_{12} está sesgado asintóticamente, a no ser que $\rho_{23,}$ que es la correlación entre X_2 y $X_{3,}$ sea cero o si la varianza muestral simple de X_3 es un estimador no sesgado de la varianza verdadera. Puede observarse que esta segunda condición no se aplica asintóticamente a un gran número de diseños muestrales de probabilidad igual (mspi), pero sólo en casos excepcionales, con diseños de probabilidades iguales (por ejemplo, diseños muestrales estratatificados no proporcionales).

29. En vez del estimador CMO puede utilizarse un estimador corregido y asintóticamente no sesgado basado en el estimador de probabilidad máxima en condiciones normales, $\widehat{\beta}_{12}$. Las expresiones de las varianzas de b_{12} y de $\widehat{\beta}_{12}$ pueden verse en Nathan y Holt (1980). Conviene señalar que el estimador habitual para la varianza de b_{12} —$v(b_{12})$— quizá diste mucho de ser no sesgado, aun cuando b_{12} sea un estimador coherente de β_{12}. Así puede ocurrir cuando los valores de ε_i no son independientes e idénticamente distribuidos entre las observaciones de la muestra.

30. Ni el estimador b_{12} ni $\widehat{\beta}_{12}$ dependen del diseño de la muestra, pero sí sus propiedades. La información sobre el diseño de la muestra puede ser útil para mejorar esos estimadores, bien cuando la información sobre los valores de la variable del diseño X_3 no está disponible para toda la población (por lo que S_3^2 no puede usarse para la estimación) o bien cuando el analista desea garantizar la solidez frente a alejamientos del modelo. Ello puede realizarse utilizando estimadores ponderados por la muestra basados en la estimación de Horvitz-Thompson para cada uno de los componentes de los estimadores no ponderados. Se pueden sustituir los momentos de la muestra no ponderada por sus versiones ponderadas en las expresiones correspondientes a b_{12} y a $\widehat{\beta}_{12}$ para obtener los estimadores ponderados, b_{12}^* y $\widehat{\beta}_{12}^*$.

31. Téngase en cuenta que b_{12}^* puede utilizarse cuando se desconoce la varianza de la población de X_3, S_3^2, pero que no se puede utilizar $\widehat{\beta}_{12}^*$ en esa situación. Se observa fácilmente que en condiciones bastante generales ambos estimadores son estimadores coherentes con el diseño del parámetro de la población finita, B_{12}.

32. Nathan y Holt (1980) han realizado comparaciones empíricas del desempeño de estos cuatro estimadores para una población de $N = 3.850$ explotaciones agrícolas de las cuales se disponía de información relativa a la tierra de cultivo (X_1), superficie total (X_2) y valor total de la producción en años precedentes (X_3). Las explotaciones se estratificaron sobre la base de los valores X_3 resultantes en seis estratos de los siguientes tamaños: 563, 584,

Cuadro XX.1
Sesgo y media cuadrática del estimador de cuadráticos mínimos ordinarios y varianzas de estimadores no sesgados para una población de 3.850 explotaciones agrícolas utilizando varios diseños de encuestas

Diseño de la encuesta	$E(b_{12}) - \beta_{12}$	MSE (b_{12})	$V(\widehat{\beta}_{12})$	$V(b^*_{12})$	$V(\widehat{\beta}^*_{12})$
A	0–000	0–000214	0–000197	0–000226	0–000197
B	0–000	0–000200	0–000198	0–000222	0–000196
C	0–031	0–001102	0–000160	0–000222	0–000196
D	0–027	0–000879	0–000163	0–000220	0–000195
E	0–042	0–001877	0–000152	0–000225	0–000196

Fuente: Nathan y Holt (1980); cuadro 1.

854, 998, 696 y 155. Se utilizaron los cinco diseños muestrales siguientes para seleccionar muestras de tamaño $n = 400$ (véase el cuadro XX.1):

A) Muestreo aleatorio simple;
B) Muestreo aleatorio simple estratificado proporcional;
C) Muestreo aleatorio simple estratificado de tamaño fijo;
D) Muestreo aleatorio simple estratificado con asignación más que proporcional a los estratos con altos valores X_3 (25, 30, 60, 80, 130, 75);
E) Muestreo aleatorio simple estratificado con asignación en forma de U (100, 80, 20, 20, 80, 100).

33. Los resultados revelan el sesgo de b_{12} para los diseños no mspi (C, D, E), mientras que los otros estimadores son o coherentes con el diseño o coherentes con el modelo (o ambas cosas). Revelan también la ventaja de $\widehat{\beta}_{12}$ con respecto a los estimadores ponderados en todos los diseños considerados. Ello es cierto aun cuando los supuestos del modelo completo no parecen aplicarse a la población. No obstante, cuando se desconoce S_3^2, b^*_{12}, menos eficiente pero coherente, es un estimador aceptable.

34. En resumen, cuando los datos están basados en diseños de probabilidades desiguales, conviene considerar los estimadores de probabilidad máxima tanto ponderados como no ponderados, en vez de los estimadores simples CMO. El estimador no ponderado parece ser más eficiente. No obstante, en muchas aplicaciones el analista no dispondrá de la información necesaria para calcular los estimadores de probabilidad máxima; los estimadores ponderados por la muestra, menos eficientes pero coherentes, son adecuados y, de hecho, se utilizan habitualmente (véase Korn y Graubard (1999)).

2. Comprobación del efecto del diseño en el análisis de regresión

35. Muchos analistas prefieren utilizar estimadores simples ponderados o no ponderados de coeficientes de regresión, que pueden obtenerse con programas estándar, en vez de los estimadores modificados que se proponen en el subapartado C.1. Hemos observado que el estimador simple CMO es coherente cuando el diseño no es informativo, o el efecto del diseño es insignificante, pero que en otros casos es preferible su contrapartida ponderada. DuMouchel y Duncan (1983) propusieron una prueba sencilla, basada en programas informáticos estándar, para decidir si deberían utilizarse ponderaciones cuando la base fueran datos procedentes de una muestra no conglomerada. Consideremos el caso de una sola variable con una variable explicativa única. La ampliación al caso con variables múltiples es sencilla. Suponiendo que $\widehat{\Delta} = b_w - b$, el objetivo es comprobar la hipótesis: $\Delta = \mathrm{E}\left(\widehat{\Delta}\right) = 0$. DuMouchel y Duncan demostraron que la prueba para $\Delta = 0$ es la misma que para $\gamma = 0$, en el contexto del modelo $Y_i = \alpha + \beta X_i + \gamma Z_i + \varepsilon_i$, donde $Z_i = w_i X_i$ y

$$\varepsilon_i | X_i \underset{ind}{\sim} \mathrm{N}(0,\ \sigma^2)$$

Los autores presentaron un ejemplo numérico para el caso de múltiples variables con un subconjunto de datos procedentes del Panel Study of Income Dynamics (University of Michigan Survey Research Center). Se seleccionó la muestra de 658 individuos con diversas probabilidades, lo que dio lugar a ponderaciones que iban de 1 a 83. El modelo final utilizado para explicar el nivel de instrucción incluía una constante y 17 variables explicativas, como la formación de los padres, los ingresos, la edad, la raza, el empleo y las interacciones. El resultado es el siguiente cuadro de análisis de la varianza:

Cuadro XX.2
Cuadro de análisis de la varianza de comparación de regresiones ponderadas y no ponderadas

Fuente	Grado de libertad	Suma de los cuadrados	Media cuadrática	F	Significación
Regresión	17	730,6	43,0	17,35	<0,0001
Ponderaciones	18	43,3	2,5	0,97	0,494
Error	622	1542,0	2,5		
Total	**657**	**2315,9**			

36. En conjunto, las 18 variables correspondientes a Z_i (las 17 variables explicativas y la constante, cada una de ellas multiplicada por w_i) tienen un valor F de 0,97 y un nivel de significación de 0,494. Así pues, está justificada una regresión no ponderada, aun cuando pueda implicar cierta pérdida de capacidad.

37. En general, a los analistas les preocupa tanto la aceptación de la hipótesis nula de que $\Delta = \mathrm{E}\left(\hat{\Delta}\right) = 0$ cuando es falsa, como su rechazo cuando es verdadera. En consecuencia, quizá decidan realizar un análisis ponderado (con el programa informático adecuado) o establecer un modelo de alcance menos precavido cuando el nivel de significación es considerablemente superior al 0,05 estándar. En el ejemplo citado, el nivel de significación es muy próximo a 0,5, lo que indica que puede hacerse caso omiso de las ponderaciones. No obstante, en una versión anterior del modelo, el nivel de significación de Z_i era 0,56, y en ese punto DuMouchel y Duncan añadieron algunos términos de interacción. Los resultados finales son los que muestra el cuadro de arriba.

38. En la prueba de DuMouchel-Duncan citada se supone que los valores ε_i son independientes y están distribuidos en forma idéntica. Muchas veces los datos de las encuestas están tomados de datos muestrales en varias etapas. Cuando los valores ε_i de las observaciones del mismo conglomerado muestral están correlacionados o cuando las observaciones tienen una heteroscedasticidad desconocida, independiente del diseño, esta prueba es inadecuada. Puede utilizarse una prueba de Wald siguiendo las líneas propuestas por Fuller (1984). En la práctica ello supone la utilización de programas informáticos como SAS/SURVEYREG y la introducción de cada punto de datos dos veces, una con una ponderación de la muestra igual a 1 y otra con una ponderación que equivalga a la ponderación efectiva.

39. Pfeffermann y Sverchkov (1999) propusieron un conjunto alternativo de ponderaciones que podrían utilizarse cuando el modelo lineal sea correcto y los errores sean independientes y estén idénticamente distribuidos pero el diseño muestral no sea ilustrativo. La prueba antes descrita puede utilizarse para evaluar sus ponderaciones en relación con las de la muestra. En Pfeffermann (1993) y Korn y Graubard (1999) puede verse un análisis más detenido de la función de las ponderaciones de muestreo al elaborar modelos sobre los datos de las encuestas.

3. Modelos en varios niveles en el marco de un diseño muestral ilustrativo

40. Desde hace poco se han venido utilizando cada vez más los modelos en varios niveles para el análisis de datos de poblaciones con estructuras jerárquicas complejas. Por

ejemplo, en la mayor parte de las encuestas de hogares los individuos incluidos dentro de los hogares son las unidades de investigación, y hay interés en conocer las relaciones tanto entre los individuos como entre los hogares. Existen estructuras jerárquicas semejantes para las encuestas de alumnos dentro de las escuelas y de empleados en los establecimientos.

41. Los modelos lineales habituales de un solo nivel pueden ampliarse fácilmente para tener en cuenta una jerarquía utilizando modelos mixtos (aleatorios y de efectos fijos) con una estructura de error que refleje la configuración jerárquica. Por ejemplo, lo que se conoce como modelo de intersección aleatoria puede formularse (en el caso de una sola variable explicativa) de la forma siguiente:

$$y_{ij} = \beta_{oi} + \beta x_{ij} + \varepsilon_{ij};\ \ \varepsilon_{ij} | x_{ij} \sim \mathrm{N}(0,\ \sigma_{\varepsilon}^2);\ \ (i = 1,\ ...,\ N;\ j = 1,\ ...,\ M_i)$$

donde y_{ij} es la variable resultante para la unidad del primer nivel j (por ejemplo, un individuo) dentro de la unidad de segundo nivel i (por ejemplo, el hogar), x_{ij} es una variable explicativa conocida y β un parámetro desconocido. La intersección, β_{oi}, es en este caso una variable aleatoria que se puede transformar en un modelo en la forma siguiente:

$$\beta_{oi} = \alpha + \gamma z_i + u_i;\ \ \ u_i | z_i \sim \mathrm{N}(0,\ \sigma_u^2);\ \ (i = 1,\ ...,\ N)$$

donde z_i es una variable explicativa conocida de segundo nivel y α y γ son parámetros desconocidos.

42. En un muestreo aleatorio simple, los modelos de este tipo pueden analizarse utilizando ampliaciones sencillas de la teoría del modelo lineal en un solo nivel. Por desgracia, no se dispone de formas cerradas para las estimaciones de los parámetros del modelo (α, β, γ, σ_{ε}^2, σ_u^2 en el modelo anterior). Lo que se hace es utilizar un procedimiento iterativo: el de los cuadrados mínimos generalizados iterativos (CMGI), que produce estimaciones que convergen en soluciones de probabilidad máxima. Por ello, los métodos de forma cerrada para adaptar los cuadrados mínimos ponderados con el fin de tener en cuenta el diseño de la muestra no pueden emplearse en este caso. Se ha elaborado una versión ponderada por la muestra de CMGI, que pondera las ecuaciones de estimación del primero y segundo nivel con ponderaciones adecuadas basadas en las probabilidades de selección, para obtener estimadores coherentes de los parámetros (véase información más detallada en Pfeffermann y otros (1998)).

43. Más recientemente, Pfeffermann, Moura y Silva (2001) han propuesto un planteamiento dependiente del modelo (puramente basado en el modelo) para el análisis en varios niveles que explica el muestreo informativo. Lo que se pretende con ello es extraer el modelo jerárquico válido para los datos de la muestra como función del modelo de población y las probabilidades de inclusión de la muestra de primer orden y luego ajustar el modelo muestral utilizando técnicas clásicas de estimación. Las probabilidades de estimación se convierten en variables resultantes adicionales que deben transformarse en modelos y, por lo tanto, mejorar el desempeño de los estimadores. Puede encontrarse información más detallada en Pfeffermann, Moura y Silva (2001). Un experimento de simulación que sigue de cerca el diseño del Estudio de evaluación de la enseñanza básica de Río de Janeiro, de 1996, indica que los resultados de la aplicación del método propuesto son prometedores.

D. Análisis de datos categóricos

1. Modificaciones en las pruebas de ji-cuadrado para las pruebas de bondad del ajuste y de independencia

44. Los intentos iniciales de evaluar los efectos del diseño muestral complejo en el análisis de los datos categóricos (es decir, datos en que cada uno de los puntos cae dentro

Cuadro XX.3
Coeficientes de tres pruebas de ji-cuadrado iteradas y pruebas de muestreo aleatorio simple*

Conjunto de datos	No. de estratos	Línea *x* columnas	Tamaño de la muestra	Tres pruebas de Nathan					
				Primera iteración			Última iteración		
				X^2	χ_1^2	*G*	X^2	χ_1^2	*G*
1	4	3x3	845	1·028	0·992	1·017	1·004	1·004	1·005
2	4	3x3	821	1·088	0·963	1·043	0·999	1·003	1·001
3	4	3x3	491	1·740	0·707	1·406	1·011	1·001	1·009
4	4	3x3	2 528	1·095	0·959	1·049	1·003	1·005	1·003
5	6	2x4	500	1·079	0·967	1·040	1·004	1·003	1·003
6	3	2x2	120	1·013	0·967	1·009	1·008	0·969	1·007
7	5	2x2	269	1·076	0·989	1·043	1·011	1·015	1·011
8	2	2x4	81	1·368	0·889	1·186	1·029	1·037	1·029

Fuente: Adaptación de datos tomados de Nathan (1972).

* Ocho cuadros de contingencia basados en muestras estratificadas proporcionales de Israel: Nos. 1-4, de ahorros; No. 5, de actitudes; No. 6, de datos de hospitales; No. 7, de medicamentos para aves de corral, y No. 8, de experimentos de percepción.

de un número finito de categorías o casillas) se referían a las modificaciones en las pruebas de ji-cuadrado que se emplean habitualmente para evaluar la bondad del ajuste entre la distribución de una sola variable categórica y una distribución hipotética o para comprobar la independencia entre dos variables categóricas. A pesar de que en las investigaciones publicadas se han propuesto varias pruebas modificadas de ji-cuadrado para datos procedentes de muestreo aleatorio simple estratificado proporcional, el efecto de ese diseño suele ser apenas relevante en la práctica. Así pues, en un estudio de estadísticos ji-cuadrado modificados en ocho conjuntos de datos procedentes de muestras estratificadas proporcionales en Israel —que se presentan en el cuadro XX.3 (tomado de Kish y Frankel (1974))—, ninguna de las estadísticas finales de interacción difería en más del 4% de las que se habrían obtenido con los supuestos de muestreo aleatorio simple, y en la mayoría de los casos la diferencia era de menos del 1%.

45. Aunque el efecto del diseño en el análisis de datos categóricos en el marco de un muestreo aleatorio simple estratificado proporcional suele ser pequeño, muchas veces eso no ocurre si se trata de muestreo conglomerado, como se demostró en el revelador estudio de Rao y Scott (1981). Al probar la bondad del ajuste, Rao y Scott observaron que, según la hipótesis nula, el estadístico ji-cuadrado habitual, X^2, se distribuye asintóticamente como una suma ponderada de las variables aleatorias $k-1$ de χ_1^2 independientes (es decir, el normal cuadrado). Las ponderaciones son los eigenvalores de una matriz D (véase el anexo, páginas 386 y ss.). La matriz puede considerarse como una extensión natural con distintas variantes del efecto del diseño para los estadísticos de una sola varianza (véanse los capítulos VI y VII). Sus eigenvalores, λ_{0i}^2 se califican como efectos del diseño generalizados y puede comprobarse que son efectos del diseño para determinadas combinaciones lineales de los efectos de diseño, d_i^2, de $\hat{p}_i$ (= la proporción estimada de la población de categoría *i*). Un estadístico ji-cuadrado modificado X_c^2, puede obtenerse dividiendo los estadísticos X^2 estándar por el promedio de las estimaciones de estos efectos de diseño generalizados, designados por $\hat{\lambda}^2$. Esta modificación requiere únicamente conocimientos acerca de los efectos del diseño de las estimaciones de la casilla. Aunque X_c^2 no tiene una distribución asintótica X_{k-1}^2 en el marco de la hipótesis nula, tiene el mismo valor previsto asintótico que X_{k-1}^2 (es decir, $k-1$), pero con una varianza mayor. Resulta que X_c^2 puede utilizarse empíricamente para comprobar la bondad del ajuste comparando el valor de este estadístico con el valor crítico de X_{k-1}^2. Así puede observarse en el cuadro XX.4 (tomado de Rao y Scott (1981)), en el que pueden verse los tamaños verdaderos de las pruebas basadas en X^2 y en X_c^2, respectivamente, en relación

Cuadro XX.4
Tamaños asintóticos estimados de pruebas basadas en X^2 y en X_c^2 para algunas partidas seleccionadas de la Encuesta general de hogares de 1971 del Reino Unido de Gran Bretaña e Irlanda del Norte (el tamaño nominal es 0,05)

Variable		*k*	*m*	$\hat{\lambda}^2$	Tamaño (X^2)	Tamaño (X_c^2)
G1	Edad de la construcción	3	33,1	3,42	0,41	0,05
G2	Tipo de propiedad	3	33,4	2,54	0,37	0,06
G3	Tipo de alojamiento	4	27,7	2,17	0,30	0,06
G4	Número de habitaciones	10	34,6	1,19	0,14	0,06
G5	Ingreso semanal bruto del hogar	6	26,6	1,14	0,10	0,06
G6	Edad del jefe del hogar	3	34,6	1,26	0,10	0,05

con seis conjuntos de datos de la Encuesta general de hogares de 1971 del Reino Unido de Gran Bretaña e Irlanda del Norte. La encuesta tenía un diseño estratificado en tres etapas. Los resultados revelan que el uso del estadístico ji-cuadrado estándar —X^2— puede ser muy engañoso, mientras que el uso del estadístico modificado — X_C^2 — funciona de manera muy satisfactoria.

46. Se mantienen resultados semejantes cuando se prueba la independencia en un cuadro de contingencia de doble entrada. En un cuadro de contingencia con r columnas y c filas, la hipótesis nula pertinente es $H_0 : h_{ij} = p_{ij} - p_{i+}\,p_{+j} = 0$ ($i = 1, ..., r$; $j = 1,..., c$), donde p_{ij} es la proporción de población de la casilla (i,j), y p_{i+}, p_{+j} son los totales marginales. El estadístico ji-cuadrado habitual para los datos de una muestra aleatoria simple — X_I^2— se distribuye asintóticamente como ji-cuadrado con $b = (r-1)(c-1)$ grados de libertad en el marco de la hipótesis nula. Ello no es necesariamente cierto cuando se trata de un diseño de la muestra complejo. De hecho, la distribución asintótica de X_I^2 es una suma ponderada de b variables aleatorias χ_1^2 independientes, algo semejante a lo ocurrido al probar la bondad del ajuste.

47. Puede elaborarse un estadístico Wald generalizado tomando como base la estimación de la matriz de varianza-covarianza completa de las estimaciones, $\hat{h}_{ij} = \hat{p}_{ij} - \hat{p}_{i+}\,\hat{p}_{+j}$ (véase información más detallada en Rao y Scott (1981)). Afortunadamente, una corrección de primer orden, que requiere únicamente estimaciones de las varianzas de $\hat{h}_{ij}$, $\hat{v}(\hat{h}_{ij})$, parece ser una aproximación adecuada. El estadístico modificado se define como

$$X_{I(C)}^2 = X_I^2 \Big/ \hat{\delta}^2_{\cdot}$$

donde $\hat{\delta}^2_{\cdot}$ es un promedio ponderado de los efectos estimados del diseño de $\hat{h}_{ij}$. Cuando no se dispone de las estimaciones de estos efectos del diseño, como ocurre con frecuencia con el análisis secundario de los datos publicados, puede obtenerse una modificación alternativa sustituyendo $\hat{\delta}^2_{\cdot}$ por $\hat{\lambda}^2$, promedio ponderado de los efectos estimados del diseño de las proporciones de la casilla, $\hat{d}_{ij}^2$. La idoneidad de estas aproximaciones depende en gran medida de la varianza relativa de los efectos del diseño. Puede obtenerse una corrección de segundo orden que podría utilizarse cuando la varianza relativa sea grande.

48. En el cuadro XX.5 (tomado de Rao y Scott (1981)) pueden verse los resultados empíricos de 15 cuadros de contingencia de doble entrada basados en datos de la Encuesta general de hogares del Reino Unido de Gran Bretaña e Irlanda del Norte. En él se observa también que el estadístico ji-cuadrado no corregido — X_I^2 — se comporta en forma bastante deficiente en muchos casos; que el estadístico corregido — $X_{I(C)}^2$ — basado en $\hat{\delta}^2_{\cdot}$, consigue el tamaño nominal casi exactamente, y que el estadístico corregido basado en $\hat{\lambda}^2$ peca por demasiado conservador.

Cuadro XX.5
Tamaños asintóticos estimados de las pruebas basadas en X_I^2, $X_I^2/\hat{\delta}_{\cdot}^2$ y en $X_I^2/\hat{\lambda}_{\cdot}^2$ para la clasificación cruzada de variables seleccionadas de la Encuesta general de hogares del Reino Unido de Gran Bretaña e Irlanda del Norte (el tamaño nominal es 0,05)

Clasificación cruzada	$r+c$	$\hat{\delta}_{\cdot}^2$	$\hat{\lambda}_{\cdot}^2$	Tamaño (X_I^2)	Tamaño ($X_I^2/\hat{\delta}_{\cdot}^2$)	Tamaño ($X_I^2/\hat{\lambda}_{\cdot}^2$)
G1 X G2	2 X 2	1,99	3,18	0,16	0,05	0,01
G1 X G3	2 X 3	1,97	2,36	0,22	0,05	0,03
G1 X G4	2 X 3	1,24	1,98	0,09	0,05	0,01
G1 X G5	2 X 6	0,91	1,23	0,04	0,05	0,02
G1 X G6	2 X 3	0,97	1,75	0,05	0,05	0,01
G2 X G3	2 X 3	1,94	2,49	0,21	0,05	0,03
G2 X G4	2 X 3	1,41	1,86	0,12	0,05	0,02
G2 X G5	2 X 6	1,02	1,18	0,06	0,05	0,03
G2 X G6	2 X 3	1,13	1,61	0,08	0,05	0,02
G3 X G4	3 X 3	1,26	1,72	0,11	0,05	0,01
G3 X G5	3 X 6	0,93	1,14	0,03	0,05	0,02
G3 X G6	3 X 3	0,96	1,51	0,05	0,05	0,01
G4 X G5	3 X 6	0,94	1,05	0,05	0,05	0,03
G4 X G6	3 X 3	0,93	1,21	0,04	0,05	0,02
G5 X G6	6 X 3	0,85	0,94	0,03	0,05	0,04

2. Generalizaciones para los modelos log-lineales

49. En Rao y Scott (1984) los resultados mencionados en los cuadros de doble entrada se han generalizado al modelo log-lineal utilizado para analizar los cuadros de múltiples direcciones. Designemos con π el vector T de las proporciones de la casilla de población, π_i, en el cuadro de varias direcciones con $\sum_1^T \pi_t = 1$ (por ejemplo, $T = 4$ en un cuadro 2 x 2). Denotemos el modelo long-lineal saturado (que incluye todas las interacciones posibles) como M_1. Tratemos de probar la hipótesis de que un submodelo anidado reducido, M_2, es suficiente. Supongamos que $\hat{\pi}$ es el estimador de probabilidad pseudomáxima de π en el marco de M_1. Éste se define como solución de la estimación muestral de las ecuaciones de probabilidad del censo (que se habrían obtenido a partir de los datos de población) y está basado en un estimador coherente con el diseño de π en el marco del diseño de la encuesta (véase el anexo). De la misma manera, supongamos que $\hat{\pi}$ es el estimador de probabilidad pseudomáxima de π en el marco de M_2. El estadístico ji-cuadrado estándar de Pearson para comprobar H_0, basado en $\hat{\pi}$ y en $\hat{\hat{\pi}}$, normalmente no tiene una distribución de ji-cuadrado asintótica en el contexto de la hipótesis nula. El caso es semejante al del cuadro de doble entrada, en la medida en que la distribución asintótica del estadístico ji-cuadrado estándar de Pearson es una suma ponderada de u variables aleatorias χ_1^2 independientes con ponderaciones δ_i^2, que son los eigenvalores de una matriz de efecto del diseño generalizada (véase información más detallada en el anexo).

50. Con el fin de tener en cuenta el diseño complejo se proponen los estadísticos ji-cuadrado modificados $X^2/\hat{\delta}_{\cdot}^2$, $X^2/\hat{\lambda}_{\cdot}^2$ y $X^2/\hat{d}_{\cdot}^2$. En este caso $\hat{\delta}_{\cdot}^2$ es el promedio de los eigenvalores estimados, $\hat{\lambda}_{\cdot}^2$ es el promedio de los efectos del diseño estimados de $X'\hat{p}$, y $\hat{d}_{\cdot}^2$ es el promedio de los efectos estimados del diseño de la casilla (véase información más detallada en el anexo). Conviene señalar que $\hat{\lambda}_{\cdot}^2$ y $\hat{d}_{\cdot}^2$ no dependen de la hipótesis nula, H_0, mientras que $\hat{\hat{\delta}}_{\cdot}^2$ sí depende. Además, $\hat{d}_{\cdot}^2$ requiere conocimientos únicamente sobre los efectos del diseño de la casilla, lo mismo que $\hat{\lambda}_{\cdot}^2$ cuando M_1 es el modelo saturado.

Cuadro XX.6

Niveles de significación (NL) asintótica estimada de X^2 y estadísticos corregidos $X^2/\hat{\delta}_{\cdot}^2$, $X^2/\hat{\lambda}_{\cdot}^2$, $X^2/\hat{d}_{\cdot}^2$: cuadro 2 x 5 x 4 y nivel de significación nominal α = 0,05

	Hipótesis a)	Hipótesis b)			Hipótesis c)		
	$\bar{1}\otimes\bar{2}\otimes\bar{3}$	$\bar{1}\otimes\bar{2}\bar{3}$	$(\bar{2}\otimes\bar{1}\bar{3})$	$(\bar{3}\otimes\bar{1}\bar{2})$	$\bar{1}\otimes\bar{2}\vert\bar{3}$	$\bar{2}\otimes\bar{1}\vert\bar{3}$	$\bar{3}\otimes\bar{1}\vert\bar{2}$
NL (X2)	0,72	0,33	0,76	0,72	0,43	0,30	0,78
NL ($X^2/\hat{\delta}_{\cdot}^2$)	0,16	0,11	0,14	0,13	0,095	0,11	0,12
NL ($X^2/\hat{\lambda}_{\cdot}^2$)	0,34	0,056	0,39	0,32	0,098	0,06	0,39
NL ($X^2/\hat{d}_{\cdot}^2$)	0,34	0,054	0,39	0,32	0,097	0,06	0,39
$\hat{\delta}_{\cdot}^2$	2,09	1,40	2,25	2,09	1,63	1,39	2,31
CV ($\hat{\delta}_i^2$)	1,54	1,02	1,37	1,27	0,86	1,05	1,11

51. En el caso importante de los modelos que admiten soluciones explícitas para $\hat{\pi}$ y para $\hat{\pi}$, Rao y Scott (1984) demuestran que $\hat{\delta}_{\cdot}^2$ puede calcularse aunque se conozcan únicamente los efectos del diseño de la casilla y los de sus marginales. Por ejemplo, en el caso de la hipótesis de la independencia completa en un cuadro en tres direcciones $I\times J\times K$, $H_0: \pi_{ijk} = \pi_{i++}\pi_{+j+}\pi_{++k}$, donde $\pi_{i++}\pi_{+j+}\pi_{++k}$ son los marginales en tres direcciones, el valor de $\hat{\delta}_{\cdot}^2$ puede calcularse explícitamente como función de las estimaciones de los efectos del diseño de los marginales en tres direcciones y de las estimaciones de los efectos del diseño de la casilla.

52. Los resultados relativos de estos estadísticos modificados y del no modificado pueden verse arriba en el cuadro XX.6 (tomado de Rao y Scott (1984)) basado en un cuadro 2×5×4 de la Encuesta de salud del Canadá de 1978-1979. Las variables son el género (I=2), el consumo de medicamentos (J=5) y el grupo de edad (K=4). Las hipótesis comprobadas eran las siguientes: *a*) independencia completa (denotada por $\bar{1}\otimes\bar{2}\otimes\bar{3}$); *b*) independencia parcial (por ejemplo, $\pi_{ijk} = \pi_{i++}\pi_{+jk} \Leftrightarrow \bar{1}\otimes\bar{2}\bar{3}$) y, de la misma manera, $(\bar{2}\otimes\bar{1}\bar{3})$ y $(\bar{3}\otimes\bar{1}\bar{2})$; *c*) independencia condicional (por ejemplo, $\pi_{ijk} = \pi_{i+k}\pi_{+jk}/\pi_{++k} \Leftrightarrow \bar{1}\otimes\bar{2}\vert\bar{3}$) y, de la misma manera, $(\bar{2}\otimes\bar{1}\vert\bar{3})$ y $(\bar{3}\otimes\bar{1}\vert\bar{2})$. El diseño era complejo, e implicaba la estratificación y el muestreo en varias etapas. Además, para mejorar las estimaciones se recurrió a la posestratificación.

53. Las comparaciones hacen referencia a los niveles de significación (NL) efectiva con el nivel nominal deseado, $\alpha = 0{,}05$. Los resultados presentan de nuevo valores inaceptablemente elevados de niveles de significación para el estadístico no corregido. Los estadísticos modificados $X^2/\hat{\lambda}_{\cdot}^2$ y $X^2/\hat{d}_{\cdot}^2$ que no dependen de la hipótesis, se comportan de forma muy semejante con valores de NL que van desde 0,06 hasta 0,39, que son muy elevados. La modificación basada en los efectos del diseño de la casilla y marginales, $X^2/\hat{\delta}_{\cdot}^2$, tiene un comportamiento más estable, con valores de NL que oscilan entre 0,095 y 0,16, todos ellos por encima del nivel nominal, debido probablemente al considerable coeficiente de variación (CV) de $\hat{\delta}_i^2$.

54. En resumen, se dispone de métodos de corrección para los estadísticos de prueba de ji-cuadrado estándar en el análisis de datos categóricos. Estas correcciones son con frecuencia necesarias para un análisis válido, en el caso de una muestra conglomerada, y pueden aplicarse con relativa facilidad utilizando los efectos del diseño de la casilla y marginales estimados. En el capítulo XXI puede encontrarse información detallada sobre los programas informáticos disponibles para resolver los efectos del diseño muestral complejo en las pruebas de ji-cuadrado y de la regresión logística.

E. Resumen y conclusiones

55. En este capítulo hemos ilustrado los métodos para evaluar los efectos que producen en el análisis de datos de encuestas los diseños muestrales complejos utilizados habitualmente. Se trata fundamentalmente de una exposición introductoria más que de un análisis prescriptivo. La evaluación y trato de los efectos del diseño muestral en el análisis puede resultar difícil y no se presta a la formulación de normas prácticas de fácil aplicación. Como se ha comprobado, los diferentes problemas pueden tener distintos (o varios) métodos posibles de solución. Éstos dependen en gran medida del modelo hipotético y de la validez de sus puestos subyacentes, de varios aspectos del diseño muestral (por ejemplo, las probabilidades de selección desigual, la conglomeración, etcétera) y del tipo de análisis previsto. Es imprescindible conocer la relación entre el modelo y las variables del diseño muestral. Por desgracia, no siempre es fácil tener acceso a esta información, lo que significa que quizá haya que limitarse a utilizar supuestos y aproximaciones.

56. Un primer paso fundamental en todo análisis es la especificación correcta del modelo subyacente. Ello es competencia del analista especializado en la materia, aunque la identificación final del modelo puede y debe basarse en técnicas estadísticas adecuadas. El análisis exploratorio inicial necesario para identificar el modelo adecuado puede realizarse utilizando métodos gráficos y descriptivos estándar sin tener en cuenta los efectos del diseño muestral.

57. Una vez que se ha hipotetizado un modelo de trabajo inicial, es necesario determinar si el diseño puede provocar cierta confusión en el análisis. Para ello se puede recurrir, por ejemplo, a una prueba que compare las estimaciones ponderadas y no ponderadas de los coeficientes de regresión lineal (véase el subapartado C.2). Si el diseño complejo debe incorporarse al análisis, hay que elegir el método adecuado para ello. El planteamiento desglosado se limita a agregar variables al modelo relacionado con el diseño de la muestra.

58. No obstante, en muchas situaciones el modelo no puede modificarse de manera que refleje los efectos del diseño muestral de forma significativa. Cuando así ocurre y debe utilizarse el planteamiento agregado, se han propuesto dos metodologías básicas para tener en cuenta el posible efecto del diseño muestral. Una supone la modificación de los instrumentos analíticos clásicos para tener en cuenta el diseño. Éste es el método que mejor se adapta para el análisis de datos categóricos, en que los estadísticos ji-cuadrado estándar pueden modificarse sobre la base de los efectos de diseño generalizados. El segundo planteamiento es el desarrollo de instrumentos analíticos adecuadamente definidos, sobre todo para el diseño. Se han propuesto estimadores ponderados de la muestra y un estadístico Wald de una muestra de gran tamaño. Se necesita un estimador fiable de la matriz de covarianza antes de utilizar el estadístico Wald. En la práctica, ello no siempre es posible.

59. Las numerosas investigaciones sobre los problemas relacionados con los efectos del diseño muestral complejo en el análisis han dado lugar a diversos métodos prácticos, algunos de los cuales se han descrito en este capítulo. Se están realizando nuevas investigaciones, y muchos de los métodos existentes se han incorporado ya a programas informáticos más o menos recientes. Por desgracia, debido a la complejidad del problema, no es probable que se elabore en el futuro un método uniforme general. Los métodos y programas informáticos disponibles deben aplicarse con suma cautela. Su aplicación requiere no sólo unos conocimientos fundamentales de la teoría básica sino también una comprensión a fondo y abundante experiencia en la construcción de modelos prácticos.

ANEXO
Definiciones formales y resultados técnicos

Modelos de regresión (subapartados B.2 y B.3)

- Modelo de regresión lineal estándar: $Y_i = \alpha + \beta X_i + \varepsilon_i$, con $\varepsilon_i | X_i \underset{ind}{\sim} \mathrm{N}(0, \sigma^2)$

- Coeficiente de regresión de **población** estándar: $$B = \frac{\sum_{i=1}^{N} Y_i (X_i - \overline{X})}{\sum_{i=1}^{N} (X_i - \overline{X})^2}$$

- Estimador de cuadrados mínimos ordinarios (CMO) de β: $$b = \frac{\sum_{i=1}^{n} y_i (x_i - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

- Estimador no sesgado de la varianza de b: $$v(b) = \frac{s^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2},$$

  donde s^2 es el estimador no sesgado de σ^2, basado en la varianza de los residuales de regresión estimados.

- Modelo de regresión lineal general (heteroscedástica): $Y_i = \alpha + \beta X_i + \varepsilon_i$,

  donde $\varepsilon_i | X_i \underset{ind}{\sim} \mathrm{N}(0, \sigma_i^2)$

- Coeficiente ponderado de regresión de **población**: $$B^* = \frac{\sum_{i=1}^{N} Y_i (X_i - \overline{X}_\sigma) / \sigma_i^2}{\sum_{i=1}^{N} (X_i - \overline{X}_\sigma)^2 / \sigma_i^2},$$

  donde $$\overline{X}_\sigma = \frac{\sum_{i=1}^{N} X_i / \sigma_i^2}{\sum_{i=1}^{N} 1 / \sigma_i^2}$$

- Estimador de cuadrados mínimos generalizados (CMG) de β: $$b_G = \frac{\sum_{i=1}^{n} y_i (x_i - \overline{x}_\sigma) / \sigma_i^2}{\sum_{i=1}^{n} (x_i - \overline{x}_\sigma)^2 / \sigma_i^2},$$

  donde $$\overline{x}_\sigma = \frac{\sum_{i=1}^{n} x_i / \sigma_i^2}{\sum_{i=1}^{n} 1 / \sigma_i^2}$$

- Varianza del estimador de CMG: $$v(b_G) = \frac{1}{\sum_{i=1}^{n} (x_i - \overline{x}_\sigma)^2 / \sigma_i^2}$$

- Estimador ponderado por la muestra (Horvitz-Thompson): $$b^* = \frac{\sum_{i=1}^{n} w_i y_i (x_i - \overline{x}^*)}{\sum_{i=1}^{n} w_i (x_i - \overline{x}^*)^2},$$

  donde $$w_i = \frac{1/\pi_i}{\sum_{k=1}^{n} 1/\pi_k};$$

  π_i es la probabilidad de inclusión y $$\overline{x}^* = \sum_{i=1}^{n} w_i y_i$$

Efecto de exclusión de las variables de diseño (subapartado C.1)

- Modelo de interés: $E(X_1) = \mu_1 + \beta_{12}(X_2 - \mu_2)$, donde β_{12} es el parámetro de interés.
- Modelo completo con variable de diseño, X_3: $E(X_1) = \mu_1 + \beta_{12\cdot3}(X_2 - \mu_2) + \beta_{13\cdot2}(X_3 - \mu_3)$
- Notación:
  - Notación habitual para el análisis de múltiples variables; por ejemplo, $\beta_{12\cdot3}$ denota el coeficiente de regresión condicional de X_1 sobre X_2, dado X_3
  - Primero y segundo momentos de población de Xi: $\overline{X}_i = \frac{1}{N} = \sum_{j=1}^{N} X_{ij}$; $S_i^2 = \frac{1}{N-1}\sum_{j=1}^{N}(X_{ij} - \overline{X}_i)^2$;
  - $S_{ik} = \frac{1}{N-1}\sum_{j=1}^{N}(X_{ij} - \overline{X}_i)(X_{ij} - \overline{X}_i)$;
  - Momentos muestrales: $\bar{x}_i = \frac{1}{n} = \sum_{j=1}^{n} x_{ij}$; $s_{ik} = \frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)$; $s_i^2 = s_{ii}$, donde suponemos una muestra, **S**, de tamaño fijo *n,* seleccionada por cualquier diseño, quizá dependiente de X_3.
- Estimador CMO estándar de β_{12}: $b_{12} = \frac{s_{12}}{s_2^2}$
- Expectativa condicional del modelo asintótico de b_{12}:

$$E_M(b_{12} \mid X_3, \mathbf{S}) = \frac{\beta_{12} + \beta_{13}\beta_{23}(s_3^2/\sigma_3^2 - 1)}{1 + \rho_{23}^2(s_3^2/\sigma_3^2 - 1)} + \mathrm{O}(n^{-1})$$

- Expectativa incondicional (modelo y diseño conjunto):

$$E_M(b_{12}) = \beta_{12} + \frac{\sigma_1}{\sigma_2}\frac{\rho_{13\cdot2}\rho_{23}\left[\left(1-\rho_{12}^2\right)\left(1-\rho_{23}^2\right)\right]^{\frac{1}{2}}(Q-1)}{1+\rho_{23}^2(Q-1)} + \mathrm{O}(n^{-1}), \text{ donde } Q = \mathrm{E}(s_3^2)/\sigma_3^2$$

- El estimador CMO, $b_{12,}$ está asintóticamente sesgado, incluso en forma no condicional, a no ser que $\rho_{23} = 0$ o $\mathrm{E}(s_3^2)/\sigma_3^2$, es decir, $Q = 1$
- Estimador asintóticamente no sesgado corregido (estimador de máxima probabilidad, EMP) en condiciones de normalidad):

$$\widehat{\beta}_{12} = \frac{s_{12} + \left(s_{13}s_{23}/s_3^2\right)\left(S_3^2/s_3^2 - 1\right)}{s_3^2 + \left(s_{23}^2/s_3^2\right)\left(S_3^2/s_3^2 - 1\right)}$$

- Estimadores ponderados: $b_{12}^* = \frac{s_{12}^*}{s_2^{*2}}$; $\widehat{\beta}_{12}^* = \frac{s_{12}^* + \left(s_{13}^* s_{23}^*/s_3^{*2}\right)\left(S_3^2/s_3^{*2} - 1\right)}{s_3^{*2} + \left(s_{23}^{*2}/s_3^{*2}\right)\left(S_3^2/s_3^{*2} - 1\right)}$, donde

$$\bar{x}_i^* = \sum_{j=1}^{n}\frac{x_{ij}}{N\pi_j};\ s_{ik}^* = \sum_{j=1}^{n}\frac{x_{ij}x_{kj}}{N\pi_j} - \bar{x}_i^*\bar{x}_k^*;\ s_i^{*2} = s_{ii}^*,\ \text{y}\ \ \pi_j = p(j \in \mathbf{S} \mid X_{3j})$$

son las probabilidades de inclusión de la muestra. Téngase en cuenta que $\sum_{j=1}^{n}\frac{1}{N\pi_j} = 1$ en el caso de muestreo aleatorio simple estratificado, que es el diseño que estamos suponiendo aquí. En los diseños más generales, $N\pi_j$ puede sustituirse por $1/w_j$, donde

$$w_j = \frac{1/\pi_j}{\sum_{k=1}^{n} 1/\pi_k}$$

- Resultado: $E_p(b_{12}^*) = E_p(\widehat{\beta}_{12}^*) = B_{12} + \mathrm{O}(n^{-1})$, donde E_P denota la expectativa del diseño (es decir, la expectativa con respecto a la selección muestral repetida).

Análisis de datos categóricos (apartado D)

- Comprobación de la bondad del ajuste:
  - Supongamos una distribución multinomial conocida con probabilidades $\mathbf{p}_0 = (p_{0,1},..., p_{0,k-1})$, donde k es el número de categorías y $\sum_1^k p_{0,1} = 1$.
  - Con H_0, el estadístico ji-cuadrado $X^2 = n\sum_{i=1}^{k}\frac{(\hat{p}_i - p_{0i})^2}{p_{0i}}$ (donde $\hat{p}_i$ son estimaciones muestrales de p_{0i}) se distribuye asintóticamente así: $X^2 = \sum_{i=1}^{k-1}\lambda_{0i}^2 Z_i^2;\ Z_i \underset{ind.}{\sim} \mathbf{N}(0,1)$, donde λ_{0i}^2 son los eigenvalores de $\mathbf{D} = \mathrm{P}_0^{-1}\mathrm{V}_0$, P_0 es la matriz de varianza de las estimaciones muestrales, con la hipótesis de nulidad para MAS, y $\mathbf{V}_0$ es su matriz de varianza verdadera con H_0.
  - Estadístico ji-cuadrado modificado: $X_C^2 = X^2 / \hat{\lambda}^2;\ \hat{\lambda}^2 = \sum_{i=1}^{k-1}(1-\hat{p}_i)\hat{d}_i^2 \big/ (k-1)$, donde $\hat{d}_i^2$ son estimaciones de los efectos del diseño, d_i^2, de $\hat{p}_i$.
- Prueba de independencia en cuadros de contingencia de doble entrada:
  - Hipótesis de interés: $H_0 : h_{ij} = p_{ij} - p_{i+}p_{+j} = 0$ ($i = 1, ..., r; j = 1, ..., c$), donde p_{ij} es la proporción de la población en la casilla (i,j)[a] y $p_{i+} = \sum_1^c p_{ij},\ p_{+j} = \sum_1^r p_{ij}$ son los totales marginales.
  - Estadístico ji-cuadrado habitual: $X_I^2 = n\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{(\hat{p}_{ij} - p_{i+}p_{+j})^2}{p_{i+}p_{+j}}$, donde $\hat{p}_{ij}$ denota el estimador muestral de p_{ij}.
  - X_I^2 está asintóticamente distribuido como suma ponderada de b variables aleatorias χ_1^2 independientes.
  - Corrección de primer orden: $X_{I(C)}^2 = X_I^2 \big/ \hat{\delta}_{\cdot}^2$ donde $\hat{\delta}_{\cdot}^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}(1-\hat{p}_{i+})(1-\hat{p}_{+j})\hat{\delta}_{ji}^2 / b$ y $\hat{\delta}_{ji}^2 = n\frac{\hat{v}(\hat{h}_{ij})}{\hat{p}_{i+}\hat{p}_{+j}(1-\hat{p}_{i+})(1-\hat{p}_{+j})}$ es el efecto estimado del diseño de $\hat{h}_{ij}$.
  - Modificación alternativa obtenida sustituyendo $\hat{\delta}_{\cdot}^2$ por $\hat{\lambda}_{\cdot}^2 = \frac{1}{rc-1}\sum_{i=1}^{r}\sum_{j=1}^{c}(1-\hat{p}_{i+})\hat{d}_{ij}$
- Generalizaciones para los modelos log-lineales
  - Modelo log-lineal: $\boldsymbol{\mu} = \tilde{u}(\theta)\mathbf{1} + \mathbf{X}\theta$ donde π es el vector T de las proporciones de la casilla de población, π_i; μ es el vector T de logaritmo de probabilidades $\mu_t = 1n\pi_t$; X es una matriz $T \times r$ conocida de pleno rango y X'1 = 0; θ es un vector r- de parámetros, y $\tilde{u}(\theta) = 1n\{1/[1'\exp(X\theta)]\}$ es un factor de normalización.
  - Hipótesis de interés: H_0: $\theta_2 = 0$, donde X = (X_1, X_2) y $\theta = (\theta_1, \theta_2)$; X_1 es $T \times s$ y X_2 es $T \times u$; θ_1 es $s \times 1$, y θ_2 es $u \times 1$.
  - Supongamos que $\hat{\pi}$ es el estimador de probabilidad pseudomáxima de π, con M_1, que es la solución de las ecuaciones de pseudoprobabilidad: $X'\hat{\pi} = X'\hat{p}$, donde $\hat{p}$ es un estimador coherente (del diseño) de π, en el marco del diseño de la encuesta. De la misma manera, supongamos que $\hat{\hat{\pi}}$ es el estimador de probabilidad pseudomáxima de π, con M_2.
  - Estadístico ji-cuadrado estándar de Pearson para comprobar H_0: $X^2 = n\sum_t \frac{(\hat{\pi}_t - \hat{\hat{\pi}}_t)^2}{\hat{\hat{\pi}}_t}$.
  - Distribución asintótica de X^2: $X^2 \sum_{i=1}^{u}\delta_i^2 Z_i^2;\ Z_i \underset{ind.}{\sim} \mathrm{N}(0,1)$ donde δ_i^2 son los eigenvalores de una matriz de efecto del diseño generalizada.

- Estadísticos ji-cuadrado modificados: $X^2/\hat{\delta}_{\cdot}^2$, $X^2/\hat{\lambda}_{\cdot}^2$ y $X^2/\hat{d}_{\cdot}^2$, donde:

  $\hat{\delta}_{\cdot}^2$ es la estimación del promedio de los eigenvalores, $\hat{\delta}_{\cdot}^2 = \frac{1}{u}\sum_i \delta_i^2$

  $\hat{\lambda}_{\cdot}^2$ es la estimación del promedio de los efectos del diseño de $\mathrm{X'\hat{p}}$

  $\hat{d}_{\cdot}^2 = \frac{1}{T}\sum_t \hat{d}_t^2$, donde $\hat{d}_t^2 = n\frac{\hat{v}(\hat{p}_t)}{\hat{\pi}_t(1-\hat{\pi}_t)}$ es el efecto estimado del diseño de la casilla *t*.

- Ejemplo: para la hipótesis de independencia completa en un cuadro $I\times J\times K$ de triple entrada, $H_0 : \pi_{ijk} = \pi_{i++}\pi_{+j+}\pi_{++k}$, donde $\pi_{i++}\pi_{+j+}\pi_{++k}$ son los marginales de triple entrada, el valor de $\hat{\delta}_{\cdot}^2$ viene dado por:

$$\hat{\delta}_{\cdot}^2 = \frac{\sum_i\sum_j\sum_k(1-\hat{\pi}_{i++}\hat{\pi}_{+j+}\hat{\pi}_{++k})\hat{d}_{ijk}^2 - \sum_i(1-\hat{\pi}_{i++})\hat{d}_i^2(r) - \sum_j(1-\hat{\pi}_{+j+})\hat{d}_j^2(c) - \sum_k(1-\hat{\pi}_{++k})\hat{d}_k^2(l)}{IJK - I - J - K + 2}$$

  donde $\hat{d}_i^2(r)$, $\hat{d}_j^2(c)$ y $\hat{d}_k^2(l)$ son estimaciones de los efectos del diseño de los marginales de triple entrada y $\hat{d}_{ijk}^2$ es la estimación del efecto del diseño de la casilla.

REFERENCIAS BIBLIOGRÁFICAS

An, A. y D. Watts (2001). *New SAS Procedures for Analysis of Sample Survey Data*. SAS Users Group International (SUGI) paper, No. 23. Cary, North Carolina, SAS Institute, Inc. Disponible en http://www2.sas.com/proceedings/sugi23/stats/p247.pdf (consultado el 2 de julio de 2004).

Berthoud, R. y J. Gershuny, comps. (2000). *Seven Years in the Lives of British Families: Evidence on the Dynamics of Social Change from the British Household Panel Survey*. Bristol, Reino Unido: The Policy Press.

Brewer, K. R. W. y R. W. Mellor (1973). The effect of sample structure on analytical surveys. *Australian Journal of Statistics,* vol. 15, págs. 145-152.

Chambers, R. L. y C. J. Skinner, comps. (2003). *Analysis of Survey Data*. Nueva York: Wiley and Sons, Inc.

Cochran, W. G. (1977). *Sampling Techniques*, 3a. ed. Nueva York: Wiley and Sons, Inc.

DuMouchel, W. H. y G. J. Duncan (1983). Using sample survey weights in multiple regression analyses of stratified samples. *Journal of the American Statistical Association,* vol. 78, págs. 535-543.

Duncan, G. J. y G. Kalton (1987). Issues of design and analysis of surveys across time. *International Statistical Review,* vol. 55, págs. 97-l17.

Feder, M., G. Nathan y D. Pfeffermann (2000). Time series multi-level modelling of complex survey longitudinal data with time varying random effects. *Survey Methodology,* vol. 26, págs. 53-65.

Fuller, W. A. (1984). Least squares and related analyses for complex survey designs. *Survey Methodology*, vol. 10, págs. 97-118.

Hansen, M. H., W. G. Madow and B. J. Tepping (1983). An evaluation of model-dependent and probability-sampling inferences in sample surveys. *Journal of the American Statistical Association,* vol. 78, págs. 776-793.

Kish, L. y M. Frankel (1974). Inference from complex samples. *Journal of the Royal Statistical Society B,* vol. 36, págs. 1-37.

Korn, E. L. y B. I. Graubard (1999). *Analysis of Health Surveys*. Nueva York y Chichester, Reino Unido: Wiley and Sons, Inc.

Manski, C. F. y S. R. Lerman (1977). The estimation of choice probabilities from choice based samples. *Econometrica,* vol. 45, págs. 1977-1988.

Nathan, G. (1972). On the asymptotic power of tests of independence in contingency tables from stratified samples. *Journal of the American Statistical Association,* vol. 67, págs. 917- 920.

_____ y D. Holt (1980). The effect of survey design on regression analysis. *Journal of the Royal Statistical Society B,* vol. 43, págs. 377-386.

_____ y T. M. F. Smith (1989). The effect of selection on regression analysis. En *Analysis of Complex Surveys,* C. J. Skinner, D. Holt y T. M. F. Smith, comps. Chichester, Reino Unido: Wiley and Sons, Inc., págs. 227-250.

Pfeffermann, D. (1993). The role of sampling weights when modeling survey data. *International Statistical Review,* vol. 61, págs. 317-337.

_____ y Sverchkov, M. (1999). Parametric and semi-parametric estimation of regression models fitted to survey data. *Sankhya*, *Series B*, vol. 61, parte 1, págs. 166-186.

Pfeffermann, D., F. Moura y N. S. Silva (2001). *Multi-level modelling under informative probability sampling*. Exposición escrita invitada al 53º período de sesiones del Instituto Internacional de Estadística, Seúl.

Pfeffermann, D. y otros (1998). Weighting for unequal selection probabilities in multi-level models. *Journal of the Royal Statistical Society B,* vol. 60, págs. 23-40.

Rao, J. N. K. y A. J. Scott (1981). The analysis of categorical data from complex sample surveys: hi-squared tests for of fit and independence in two-way tables. *Journal of the American Statistical Association,* vol. 76, págs. 221-230.

_____ (1984). On chi-squared tests for multi-way contingency tables with cell proportions estimated from survey data. *Annals of Statistics,* vol.12, págs. 46-60.

Särndal, C-E., B. Swensson y J. Wretman (1992). *Model Assisted Survey Sampling*. Nueva York: Springer-Verlag.

Scott, A. J. y C. J. Wild (1986). Fitting logistic models under case control or choice-based sampling. *Journal of the Royal Statistical Society B,* vol. 48, págs. 170-182.

Skinner, C. J., D. Holt y T. M. F. Smith, comps. (1989). *Analysis of Complex Surveys.* Chichester, Reino Unido: Wiley and Sons, Inc.

Valliant, R., A. H. Dorfman y R. M. Royall (2000). *Finite Population Sampling and Inference: A Prediction Approach*. Chichester, Reino Unido, y Nueva York: Wiley and Sons, Inc.

Capítulo XXI
Estimación de errores muestrales en los datos de encuestas*

Donna Brogan
Universidad de Emory
Atlanta (Georgia)
Estados Unidos de América

* Este capítulo se acompaña de un anexo (en cederrón y sólo en inglés) que contiene los análisis ilustrativos y comparativos que de los datos obtenidos de la Encuesta sobre inmunización en Burundi se realizaron con cinco paquetes de *software* estadístico. El contenido del cederrón, incluidos los códigos de los programas y los resultados de cada uno de los paquetes de *software*, puede descargarse directamente del sitio web de la División de Estadística de las Naciones Unidas (http://unstats.un.orgunsdhhsurveys/). Si se desea, también puede solicitarse el cederrón a la División de Estadística de las Naciones Unidas (statistics@un.org).

RESUMEN

Los diseños de encuestas complejas difieren de los muestreos aleatorios simples en aspectos como la probabilidad desigual del muestreo, el muestreo polietápico y la estratificación. Es necesario recurrir a análisis ponderados para obtener estimaciones de los parámetros de población no sesgadas (o casi sin sesgo). Para los estimadores, la estimación de varianzas depende de los detalles específicos del plan de muestreo y requiere de métodos aproximados, por lo general la linealización mediante el desarrollo en serie de Taylor o las técnicas de replicación.

Por lo general, no es posible utilizar paquetes de *software* estadístico estándar para analizar los datos de las encuestas, ya que suelen presuponer un muestreo aleatorio simple de los elementos. Estos paquetes producen estimaciones por puntos sesgadas de los parámetros de población (en análisis no ponderados), así como una subestimación de los errores estándar de las estimaciones por puntos. El uso de la variable de ponderación de la muestra con paquetes estándar produce adecuadas estimaciones por puntos de los parámetros de población. No obstante, los errores estándar estimados siguen siendo incorrectos por lo general, porque el procedimiento de estimación de varianzas no suele tener en cuenta los estratos y conglomerados según el plan de muestreo.

En este capítulo se hace una descripción general de ocho paquetes de *software* con capacidad para realizar análisis de datos de las encuestas, incluido su costo aproximado, métodos de estimación de varianzas, opciones de análisis, interfaz de usuario y ventajas e inconvenientes. Cuatro de los paquetes son gratuitos, lo que resulta interesante para los países en desarrollo que tengan un presupuesto limitado para la adquisición de *software*.

El conjunto de datos de una encuesta compleja realizada en Burundi sirve para ilustrar el hecho de que con el *software* estadístico estándar se obtienen análisis incorrectos. Mediante análisis descriptivos comentados de la encuesta de Burundi realizados con cinco de los ocho paquetes examinados (STATA, SAS, SUDAAN, WesVar y Epi-Info) se muestra cómo se utilizan estos paquetes. Por último, se comparan los resultados numéricos obtenidos con los cinco paquetes de *software* para objetivos analíticos comunes utilizando los datos de la encuesta de Burundi. Los cinco paquetes brindan resultados equivalentes en cuanto a la estimación de varianzas, tanto si se emplea la linealización mediante el desarrollo en serie de Taylor como si se utiliza la replicación repetida y equilibrada (BRR).

Téerminos clave: Linealización mediante el desarrollo en serie de Taylor, métodos de replicación, conglomerado final, estimación de varianzas, encuestas complejas, paquetes de *software*.

A. Diseños muestrales

1. Como se pone de manifiesto en muchos capítulos de esta publicación, los diseños muestrales para encuestas de hogares son complejos y suelen requerir muestreo polietápico. Como consecuencia del empleo de un diseño muestral complejo, no es posible aplicar sin más los métodos estadísticos y el *software* estándar para el análisis de los datos de las encuestas de hogares. En particular, a las unidades entrevistadas en una encuesta se les asignan ponderaciones para compensar la desigual probabilidad de ser seleccionadas, las unidades que no pueden ser entrevistadas y los problemas de cobertura. Esas ponderaciones pueden utilizarse para generar distribuciones ponderadas de la encuesta de manera que determinadas variables se adapten a las distribuciones conocidas de esas variables. Además, deben utilizarse en el análisis de la encuesta. Asimismo, en la computación de los errores del muestreo en las estimaciones de la encuesta ha de tenerse en cuenta el hecho de que en la selección de muestras para la encuesta se ha empleado un diseño muestral complejo. Afortunadamente, en la actualidad existen varios paquetes de *software* especializados para el análisis de encuestas que computan correctamente los errores del muestreo para obtener estimaciones ponderadas de las encuestas elaboradas con diseños muestrales complejos. En este capítulo se describen y examinan algunos de estos paquetes.

2. Como preparación para el estudio de los paquetes de análisis de encuestas, en los dos apartados siguientes se examina la cuestión de los análisis ponderados y los métodos de estimación de varianzas con diseños muestrales complejos. En ellos se comparan las estimaciones de varianzas obtenidas mediante ocho paquetes de *software* a partir de los datos de encuestas complejas y se ilustra el empleo de cinco de ellos con los datos de una encuesta realizada en Burundi. El capítulo se cierra con algunas conclusiones y recomendaciones. En uno de los anexos del cederrón (véase la nota al comienzo del capítulo) se presentan análisis comentados de los datos correspondientes a tres análisis realizados con los cinco paquetes de *software* seleccionados.

B. Aspectos del análisis de datos de encuestas complejas

1. Análisis ponderados

3. En muchas encuestas de hogares, las unidades de análisis —hogares o personas— se seleccionan con probabilidades desiguales, por lo que es necesario efectuar ponderaciones para compensar en los análisis estas desiguales probabilidades de selección. Además, aunque las unidades se seleccionen con igual probabilidad, a menudo se necesitan ponderaciones para compensar las faltas de respuesta de las unidades y para establecer referencias comparativas, como en el caso de la posestratificación (véase el capítulo XIX). Estas ponderaciones deben utilizarse en los análisis para estimar los parámetros de población. Los estimadores no ponderados (no recomendables) pueden presentar un gran sesgo para los parámetros de población, dependiendo de la encuesta en concreto. El valor de la variable de ponderación de la muestra, designada como *WTVAR*, para un determinado elemento muestral entrevistado *R* del conjunto de datos puede interpretarse como el número de elementos de la población representados por *R*. La suma del valor de *WTVAR* para todas las *R* del conjunto de datos representa una estimación del número de elementos de la población.

4. A veces la variable de ponderación de la muestra *WTVAR* se "normaliza" multiplicándola por (número de *R*) / (suma del valor de *WTVAR* de todas las R). La suma del valor de la "variable de ponderación normal" *WTNORM* para todas las *R* es el tamaño de la muestra para análisis (número de R). No importa si se utiliza la variable de ponderación de la muestra *WTVAR* o la variable de ponderación normal *WTNORM* para obtener una estimación por puntos de un parámetro de población "promedio" como, por ejemplo, una media o proporción: ambas producen el mismo resultado. No obstante, la variable de pon-

deración normal *WTNORM* no puede utilizarse para estimar directamente los totales de los parámetros de población, como, por ejemplo, el número total de niños desnutridos dentro de la población.

2. Descripción general de la estimación de varianzas

5. La importancia de la estimación de varianzas estriba en que indica la precisión de los estimadores, lo que permite establecer intervalos de confianza de los parámetros de población y ensayar hipótesis sobre dichos parámetros. En la estimación de varianzas con estimadores basados en datos de encuestas complejas se deben tener en cuenta los siguientes factores: *a*) la mayoría de los estimadores no son lineales (normalmente hay una cierta proporción de estimadores lineales); *b*) los estimadores están ponderados; *c*) en general, en el plan de muestreo se habrá recurrido a la estratificación antes del muestreo en la primera etapa (y quizá también en las etapas de muestreo posteriores); y *d*) los elementos de la muestra no serán, por lo general, independientes, debido al muestreo polietápico por conglomerados. En casi todos los casos resulta imposible obtener una expresión algebraica cerrada de la varianza estimada. En la bibliografía referente a la investigación de la estimación de varianzas para datos de encuestas complejas, los analistas de datos de encuestas pueden encontrar varios métodos aproximados entre los que elegir.

6. Los enfoques más comunes para aproximarse al valor de la varianza estimada son la linealización mediante el desarrollo en serie de Taylor (TSL) (Wolter, 1985; Shah, 1998) y las técnicas de replicación (Wolter, 1985; Rust y Rao, 1996). Estos enfoques se estudian más detalladamente en el apartado C. La mayoría de los paquetes de *software* para el análisis de datos de encuestas sólo utilizan uno de estos dos métodos. Para los estimadores que son funciones suavizadas de los datos muestrales (por ejemplo, totales, medias, proporciones, diferencias entre medias o proporciones, etcétera), ambos métodos ofrecen estimaciones comparables de las varianzas, y ninguno goza de clara preferencia. Para los estimadores que son funciones no suavizadas de los datos muestrales (por ejemplo, medianas), un procedimiento específico de replicación —la replicación repetida y equilibrada (BRR)— parece gozar de preferencia sobre el de linealización mediante el desarrollo en serie de Taylor y el *jackknife*, otro método de replicación (Korn y Graubard, 1999). Existe abundante bibliografía sobre la comparación entre las distintas técnicas de estimación de varianzas, incluidos casos particulares en los que un método puede ser preferible a otro (por ejemplo, véase Korn y Graubard (1999) y sus numerosas referencias, y también Kish y Frankel (1974)).

3. Factores de corrección al muestreo de poblaciones finitas para muestreo sin reemplazamiento

7. Para simplificar, consideremos inicialmente la estimación de una media de población obtenida de una muestra de tamaño n seleccionada con probabilidad igual de entre una población de tamaño N, y comparemos dos diseños muestrales. En uno de los diseños los elementos se seleccionan mediante muestreo aleatorio simple, esto es, se seleccionan sin reemplazamiento. En el otro diseño se seleccionan mediante muestreo irrestricto, esto es, con reemplazamiento (denominado también muestreo aleatorio simple con reemplazamiento). La diferencia en la varianza de las medias muestrales con estos dos diseños consiste en que en el muestreo aleatorio simple se introduce en la varianza un término de corrección al muestreo de poblaciones finitas (*fpc*), cosa que no ocurre en el caso del muestreo irrestricto (véase el capítulo VI). El término *fpc* es $(1 - f)$, donde $f = n/N$ es la fracción del muestreo. El término *fpc* tiene un límite superior de 1,0 y refleja la reducción de la varianza resultante del muestreo sin reemplazamiento. Si la fracción del muestreo f es pequeña, el término *fpc* tiende a 1,0 y su efecto sobre la varianza es mínimo. Puede despreciarse sin problema en la estimación de

varianzas. Dicho de otro modo, el muestreo sin reemplazamiento puede tratarse como si fuera un muestreo con reemplazamiento. Por lo general, se considera que una fracción del muestreo es pequeña cuando no pasa del 5% ó del 10 %. Por otra parte, si *f* es grande, despreciar el término *fpc* cuando la selección muestral se hace sin reemplazamiento dará lugar a una sobreestimación de la varianza. En diseños muestrales aleatorios estratificados con distintas fracciones del muestreo en distintos estratos, el término *fpc* puede ser suficientemente pequeño como para despreciarlo en algunos estratos, aunque no en otros.

8. La mayoría de las encuestas de hogares se basan en diseños muestrales complejos aplicados a poblaciones muy extensas. Por lo general, las unidades primarias de muestreo (UPM) se seleccionan mediante muestreo sin reemplazamiento con probabilidad proporcional al tamaño (PPT), lo que hace más complejo el concepto de "fracción del muestreo". No obstante, el número de UPM es a menudo grande y la fracción del muestreo de las UPM en cada estrato es muy pequeña, dando un valor próximo a 1,0 para todos los términos *fpc* en la primera etapa. Así, una aproximación común en el análisis de datos de encuestas complejas es aquella en la que las UPM se han sometido a muestreo con reemplazamiento en cada estrato. Si se aplica esta aproximación cuando existen estratos con fracciones grandes del muestreo en la primera etapa, la varianza resultará sobreestimada en alguna medida. A menudo se admite esta sobrestimación por razón de la complejidad de la estimación de varianzas sin la aproximación. Obsérvese que si en algún estrato el muestreo se hace con reemplazamiento en la primera etapa de muestreo, no habrá aproximación para ese estrato.

4. Pseudoestratos y pseudo-UPM

9. A los efectos de estimación de varianzas, a veces no se identifican los estratos y las UPM como se utilizaron realmente en el plan de muestreo. Es posible hacer modificaciones en la definición de los estratos y las UPM para la estimación de varianzas, con objeto de que el plan de muestreo utilizado se ajuste a una de las opciones de planes de muestreo disponibles en un paquete de *software*. Cuando se llevan a cabo estas modificaciones, las variables redefinidas de estratos y UPM para la estimación de varianzas se denominan a veces pseudoestratos y pseudo-UPM.

10. Un ejemplo común se presenta cuando se define un número muy grande de estratos antes del muestreo en la primera etapa, con sólo una UPM seleccionada (muestreada) en cada estrato. La estimación de varianzas es imposible con sólo una UPM por estrato, puesto que no se puede estimar la variabilidad entre las UPM en el estrato. En esta situación, dos estratos se reúnen o combinan en un pseudoestrato, dando así dos UPM seleccionadas en el pseudoestrato. La reunión de estratos se realiza con un criterio estratégico, y se basa en el conocimiento de la(s) variable(s) de estratificación de las UPM y del método de muestreo de las UPM (Kish, 1965).

11. La estratificación implícita representa otro ejemplo. Un país puede, por ejemplo, estar estratificado por el norte y por el sur, y las UPM serían las distintas localidades. En cada estrato, las UPM se ordenan por proximidad geográfica, seguidas de una selección de una muestra probabilística de muchas UPM (30, pongamos por caso) en cada estrato mediante muestreo sistemático por probabilidad proporcional al tamaño estimado (PPTE) (Kish, 1965). La ordenación geográfica de las UPM de población en el estrato, combinada con el muestreo sistemático, da lugar a una estratificación geográfica implícita de las localidades (UPM) en cada uno de los estratos del norte y del sur. Para tomar en consideración la estratificación implícita en la estimación de varianzas, el plan de muestreo tendría que comprender 15 pseudoestratos del norte y 15 pseudoestratos del sur, cada uno de ellos con dos UPM o pseudo-UPM seleccionadas. Las dos primeras UPM seleccionadas del marco de muestreo del norte irían al primer pseudoestrato, las dos siguientes UPM obtenidas irían al segundo pseudoestrato, y así sucesivamente.

12. Korn y Graubard (1999) brindan varios ejemplos adicionales de formación de pseudoestratos y pseudo-UPM para la estimación de varianzas para, por ejemplo, reducir el número de réplicas y la carga de cálculo. En el apéndice D de la guía del usuario de WesVar (2002) se brinda también orientación y ejemplos, con la descripción de diversos planes de muestreo para el *software* de estimación de varianzas según distintos métodos de replicación.

5. Una aproximación común (*WR*) para la descripción de muchos planes de muestreo complejo

13. Normalmente, las encuestas complejas utilizan el muestreo polietápico por conglomerados. Además, es habitual la estratificación de las UPM de la población antes del muestreo en la primera etapa. Por otra parte, la estratificación de las unidades de la segunda y posteriores etapas (en una UPM seleccionada) puede tener lugar antes del muestreo en estas etapas. Sin embargo, con los métodos aproximados de estimación de varianzas comúnmente utilizados para estos diseños complejos no es necesario tomar en consideración todas las etapas de muestreo y estratificación. El muestreo complejo en las etapas posteriores queda automáticamente cubierto de manera adecuada con la aproximación "con reemplazamiento" para la primera etapa de muestreo que hemos visto más arriba. De hecho, pocos paquetes de *software* tienen la capacidad de incluir por separado todas las etapas de muestreo en la estimación de varianzas, en los casos en los que no se practica la aproximación con reemplazamiento en la primera etapa.

14. Es muy común emplear en los diseños complejos la estimación de varianzas por conglomerados finales (EVCF), propuesta por primera vez por Hansen, Hurwitz y Madow (1953) y analizada también en Wolter (1985). La estimación de varianzas por conglomerados finales puede efectuarse indistintamente por linealización mediante el desarrollo en serie de Taylor o con una técnica de replicación. Con el enfoque de EVCF, las UPM se tratan como si se hubieran seleccionado con reemplazamiento entre los estratos de la primera etapa. En tal caso, cada *R* (elemento muestral entrevistado del conjunto de datos) sólo tiene que identificarse por el estrato de la primera etapa y por la UPM (dentro del estrato) de la que se haya seleccionado. A los efectos de la estimación de varianzas no se necesita información sobre las etapas de muestreo por debajo del nivel de las UPM pero antes de la etapa del elemento. De este modo la descripción del plan de muestreo se simplifica de forma que se asemeja a un muestreo por conglomerados estratificados en una etapa, esto es, una muestra estratificada de conglomerados finales completamente enumerados. Este enfoque por conglomerados finales produce una buena aproximación para la estimación de la varianza, siempre que sea razonable partir de la hipótesis de una primera etapa con reemplazamiento. Esta aproximación común (EVCF) se denomina a veces *WR* (por las siglas inglesas de la expresión "con reemplazamiento") en la bibliografía sobre encuestas, por lo que en adelante emplearemos las siglas *WR* con este significado.

15. Así, cuando el plan de muestreo se describe como *WR* sólo se necesitan tres variables del diseño de la encuesta para la estimación de varianzas:

a) La variable de ponderación de la muestra *WTVAR* (necesaria también para las estimaciones por puntos);

b) La variable de estratificación (o variable de pseudoestratificación) *STRATVAR* utilizada antes de la selección (de UPM) en la primera etapa; y

c) La variable UPM (o pseudo-UPM), designada como *PSUVAR*.

16. Cada entrevistado de la muestra *R* debe tener un valor para cada una de estas tres variables en el archivo de datos básicos. Por ejemplo, una *R* en particular puede representar a 8.714 elementos de la población (el valor de *WTVAR* será 8.714) y puede haber sido se-

leccionada del estrato o pseudoestrato No. 6 (el valor de *STRATVAR* será 6) y de la UPM o pseudo-UPM No. 3 del estrato 6 (*PSUVAR* tendrá un valor de 3 para *STRATVAR* = 6).

17. *WR* es la descripción por defecto o única del plan de muestreo en la mayoría de los paquetes de *software* o procedimientos de análisis de encuestas. Por ejemplo, con la linealización mediante el desarrollo en serie de Taylor, *WR* es la descripción por defecto en las aplicaciones SUDAAN, SAS, STATA, Epi-Info, PC-CARP y CENVAR. Con BRR y *jackknife*, *WR* es la descripción por defecto en las aplicaciones WesVar y SUDAAN. Obsérvese que el muestreo monoetápico de elementos, como en el caso del muestreo aleatorio simple o el muestreo aleatorio estratificado, es un caso especial de muestreo polietápico en el que las UPM de la población que aparecen en el marco de muestreo son los elementos de población y cada UPM seleccionada sólo contiene un elemento (en otras palabras, no hay conglomerado de elementos muestrales). Los paquetes de *software* que sólo incluyen la descripción del plan de muestreo *WR* pueden ofrecer la opción de incorporar términos *fpc* en la estimación de varianzas cuando se emplea el muestreo monoetápico de elementos sin reemplazamiento (por ejemplo, SAS, STATA o WesVar).

18. El empleo de *WR* para la aproximación al plan de muestreo complejo puede dar lugar a una ligera sobreestimación de las varianzas. Sin embargo, los analistas de datos de encuestas no tienen, por lo general, problema en aceptar un cierto grado de sobreestimación por la relativa simplicidad de la aproximación mediante *WR*. No obstante, hay que tener en cuenta que la sobrestimación puede ser apreciable si hay varios estratos en los que el muestreo en la primera etapa se hace sin reemplazamiento y con grandes fracciones del muestreo. En este caso puede ser preferible optar por un *software* que cuente con la opción de incorporar los factores *fpc* para la primera etapa.

6. Técnicas de estimación de varianzas y variables de diseño de la encuesta

19. Los conjuntos de datos de las encuestas para uso público suelen venir ya preparados para la estimación de varianzas mediante alguno de los dos enfoques más conocidos, sea la linealización mediante el desarrollo en serie de Taylor o las técnicas de replicación. Ocasionalmente, un conjunto de datos para uso público puede estar preparado para ambos enfoques de la estimación de varianzas. En el conjunto de datos para uso público deben incluirse las variables necesarias para estimar las varianzas, además de la documentación correspondiente referente a la definición de estas variables y el modo de usarlas.

20. Si en el conjunto de datos se emplea la linealización mediante el desarrollo en serie de Taylor, en la documentación deberán aparecer tres variables de diseño de la encuesta: la variable de ponderación de la muestra *WTVAR*, la variable de estratificación de la primera etapa *STRATVAR* y la variable UPM *PSUVAR* (por supuesto, los nombres de las variables no serán los aquí utilizados). Si para el conjunto de datos se emplea un método de replicación, tendremos la variable de ponderación de la muestra *WTVAR* y varias variables de ponderación replicadas, a menudo con nombres como REPL01—REPL52 (para 52 variables de ponderación replicadas). No es necesario conocer las variables *STRATVAR* o *PSUVAR* si en el conjunto de datos se incluyen las variables de ponderación replicadas.

21. El encuestador que se ocupa de su propia encuesta y prepara su conjunto de datos para el análisis ha de incluir en el mismo las necesarias variables de diseño de la encuesta y asignarles un valor por cada elemento muestral entrevistado (*R*). El conjunto mínimo de variables necesario es el siguiente: variable de ponderación de la muestra *WTVAR* variable de estratificación (o pseudoestratificación) en la primera etapa *STRATVAR*; y variable UPM (o pseudo-UPM) *PSUVAR* en el estrato. Estas tres variables de diseño de la encuesta hacen que la aproximación del plan de muestreo sea *WR* y permiten el uso directo de la linealización mediante el desarrollo en serie de Taylor o el cálculo, manual o con ayuda de *software*, de

ponderaciones replicadas para aplicar las técnicas de replicación en la estimación de varianzas. Si se desea incorporar términos *fpc* o etapas adicionales de muestreo o estratificación a la estimación de varianzas, el conjunto de datos tendrá que incluir variables adicionales de diseño de la encuesta y se necesitará un *software* de análisis de encuestas con capacidad para ello (por ejemplo, SUDAAN).

22. Lamentablemente, al adquirir un conjunto de datos de encuesta es bastante corriente que éste no incluya ninguna variable de diseño de la encuesta ni variable de ponderación replicada. Suponiendo que se ha utilizado el muestreo probabilístico, es necesario construir las variables de diseño de la encuesta *WTVAR* para la estimación y *STRATVAR* y *PSUVAR* para la estimación de varianzas. Con suerte, a partir de la documentación escrita o mediante la consulta personal con el equipo de encuesta se puede conseguir información del plan de muestreo suficiente para construir las variables de diseño de la encuesta. Si sólo se dispone de información limitada es posible realizar algunas aproximaciones en bruto. Por ejemplo, si no se pueden reconstruir las probabilidades de selección, podría suponerse razonablemente que la muestra de elementos es de probabilidad igual, y realizar un ajuste mediante posestratificación para obtener valores de *WTVAR*. Si no se pueden identificar con exactitud las UPM, se podrían desarrollar UPM vicarias si se conocen determinados identificadores geográficos. En tales casos hay que tener presentes las limitaciones del análisis de datos cuando las variables del diseño muestral son imprecisas.

7. Análisis de datos de encuestas complejas

23. Existen muchas cuestiones teóricas y prácticas en torno al análisis de datos de encuestas complejas, que van más allá de la realización de un análisis ponderado y de la correcta estimación de las varianzas de los estimadores. Estas cuestiones están bien tratadas e ilustradas en el reciente y extenso libro de Korn y Graubard (1999), que incluye temas como el ajuste de modelos (por ejemplo, la regresión logística) a los datos de las encuestas, la bondad del ajuste de los modelos, la estimación de varianzas para subpoblaciones, la combinación de encuestas múltiples y la formación de pseudoestratos y pseudo-UPM. Véanse también otros capítulos de la presente publicación.

C. Métodos de estimación de varianzas

1. Linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas

24. Supongamos un plan de muestreo con estratificación de UPM, muestreo politápico y probabilidad desigual del muestreo de los elementos. El estimador lineal $\Sigma\, w_i\, y_i$, que es una suma ponderada, estima el total de población para la variable y, donde w_i es el valor de la variable de ponderación de la muestra *WTVAR* para el elemento muestral i; y_i es el valor de la variable y para el elemento muestral i, y la suma Σ abarca todos los elementos de la muestra, $i = 1, 2, \ldots, m$. Si y es una variable binaria codificada como 1 para los varones diabéticos y 0 para los demás casos, el total de población sometido a estimación será el número total de varones diabéticos. La varianza estimada de $\Sigma\, w_i y_i$ puede obtenerse directamente partiendo del supuesto de *WR* estudiado más arriba.

25. Sea ahora x_i una variable binaria codificada como 1 para los varones y 0 para las mujeres. En tal caso, la prevalencia estimada de la diabetes entre los varones vendrá dada por $[\Sigma\, w_i y_i\,] \,/\, [\Sigma\, w_i x_i\,]$, que es el cociente de dos estimadores lineales (o dos sumas ponderadas). Partiendo del supuesto de *WR*, no es posible obtener directamente la varianza estimada de este cociente de estimadores. Incluso si se utiliza el muestreo aleatorio simple en vez de métodos de muestreo complejo, la estimación de la varianza de esta función no lineal, un cociente, no es directa y requiere el empleo de algún método aproximado.

26. La expresión algebraica del estimador no lineal anterior puede desarrollarse en una serie de Taylor infinita centrada en el valor esperado (estimado) del numerador y el valor esperado (estimado) del denominador. Seguidamente, el estimador no lineal se somete a aproximación algebraica conservando sólo los primeros términos de la serie infinita de Taylor, lo que produce una expresión algebraica que ya no es una función no lineal, sino una función lineal de los datos muestrales; esto es, el estimador no lineal cociente se ha "linealizado". Ahora es posible obtener directamente la varianza estimada de la función linealizada (incluidos los términos de covarianza pertinentes) con la hipótesis de *WR*, del mismo modo que se obtuvo la varianza estimada de $\Sigma\, w_i\, y_i$. En este proceso la varianza de la función linealizada se estima en cada estrato por separado (puesto que el muestreo es independiente para cada estrato) y seguidamente se suman las varianzas estimadas propias de cada estrato para obtener la varianza del estimador.

27. Cuando se emplea el enfoque de la linealización mediante el desarrollo en serie de Taylor es necesario obtener y programar una fórmula singular para la estimación aproximada de la varianza, no sólo para cada diferente estimador no lineal sino también para cada posible plan de muestreo en el que pudiera utilizarse ese estimador (donde *WR* sería uno de esos planes de muestreo). Esta característica se considera un inconveniente del enfoque de la linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas. De hecho es posible que un paquete de *software* capaz de analizar datos de encuestas con la linealización mediante el desarrollo en serie de Taylor no incluya la combinación del estimador específico que queramos utilizar con el plan de muestreo real o aproximado que hayamos empleado.

28. En los programas de *software* que emplean la linealización mediante el desarrollo en serie de Taylor es necesario especificar las variables del diseño *WTVAR*, *STRATVAR* y *PSUVAR*, según se precise para la aproximación al plan de muestreo *WR*. Algunos paquetes de *software* permiten la linealización mediante el desarrollo en serie de Taylor con otros planes de muestreo, aunque el empleo de estos planes puede requerir variables de diseño adicionales.

2. El método de replicación para la estimación de varianzas

29. El método de replicación para la estimación de varianzas de los estimadores de las encuestas, aunque conocido teóricamente desde hace tiempo, ha suscitado un interés creciente con la aparición de las computadoras de alta velocidad de proceso. El método de replicación exige un uso intensivo de las computadoras, pero es más flexible que el método de linealización mediante el desarrollo en serie de Taylor en cuanto al número de estimadores diferentes para los que se pueden computar las varianzas estimadas.

30. La idea general de los métodos de replicación es como sigue. En primer lugar, se emplea la totalidad de la muestra, de la misma manera que sucede en el método del desarrollo en serie de Taylor, para obtener una estimación por puntos del parámetro de población que nos interese; esto es, la fórmula del estimador para el parámetro de población se aplica a la totalidad de la muestra. Para este cálculo sólo se necesita la variable de ponderación de la muestra *WTVAR*.

31. En segundo lugar, para estimar la varianza de este estimador se forman a partir de la muestra completa muchas submuestras diferentes de manera que cada réplica refleje el plan de muestreo y los ajustes y procedimientos de ponderación de la muestra completa. Cada réplica se define por el valor de una variable de ponderación replicada. Por ejemplo, *REPWTj* es la variable de ponderación replicada de la réplica #*j*, donde $j = 1, 2, 3, \ldots, G$ (número total de réplicas). Una observación en la muestra completa tiene valor cero para *REPWTj* si esa observación no se incluye en la réplica #*j*, y un valor positivo si se incluye

en la réplica #*j*. La suma de los valores de *REPWTj* en todas las observaciones de la muestra completa es una estimación del número de elementos de la población.

32. Tercero, la fórmula del estimador se aplica a cada réplica para obtener una estimación por puntos del parámetro de población que nos interese (la estimación replicada), lo que produce *G* estimaciones replicadas del mismo parámetro de población.

33. Cuarto, a partir de la variabilidad de las *G* estimaciones replicadas, se calcula una varianza estimada del estimador de la muestra completa.

34. Las réplicas se pueden formar de distintas maneras, lo que da lugar a diversas técnicas de replicación. Dos de los principales métodos de formación de réplicas, cada uno de ellos con variantes, son la replicación repetida y equilibrada (BRR) y el *jackknife* (ambos se estudian más adelante). Los conjuntos de datos de las encuestas para uso público preparados para la estimación de varianzas según un método concreto de replicación suelen incluir las variables de ponderación replicadas. En este caso, el analista de datos secundarios debe emplear un *software* de estimación de varianzas que incluya la técnica de replicación específica para la que se hubieran generado las ponderaciones replicadas incluidas en el conjunto de datos.

35. No obstante, es posible que queramos utilizar una técnica de replicación para la estimación de varianzas sin disponer de las ponderaciones replicadas en el conjunto de datos. Algunos paquetes de *software* que aplican métodos de estimación de varianzas por replicación calculan también las ponderaciones replicadas. Para formar ponderaciones replicadas, los paquetes de *software* necesitan como mínimo las siguientes tres variables de diseño de la encuesta: variable de ponderación de la muestra *WTVAR*, variable de estratificación *STRATVAR* y variable UPM *PSUVAR* del estrato. Si la muestra completa ha sido objeto de ajuste para las faltas de respuesta o ha sido estratificada *a posteriori*, esta información puede introducirse también en el paquete de *software* para el cálculo de las ponderaciones replicadas (por ejemplo, en WesVar). Siempre podemos calcular por nosotros mismos las ponderaciones replicadas (es decir, sin valernos de un paquete de *software*), pero esta estrategia sólo es recomendable si quien va a realizar el cálculo es una persona familiarizada con los detalles de las técnicas de replicación.

3. Replicación repetida y equilibrada (BRR)

36. La replicación repetida y equilibrada es una técnica de replicación específica que puede utilizarse para diseños muy generales, como el muestreo polietápico. Sin embargo, se desarrolló para el caso específico en el que hay exactamente dos UPM seleccionadas (muestreadas) por estrato, muestreadas por lo general con probabilidades desiguales con o sin reemplazamiento. Su uso está también generalizado con la aproximación *WR* al plan de muestreo complejo (el enfoque EVCF).

37. Con la BRR cada réplica contiene exactamente la mitad de las UPM seleccionadas, una de cada estrato; es frecuente que a cada réplica se la denomine "media muestra". El número total de réplicas distintas posibles es 2^L, donde L es el número de estratos. No obstante, no es necesario utilizar la totalidad de las 2^L réplicas, lo que requeriría un número desmesurado de horas de cálculo. Basta con un conjunto menor y "balanceado" de réplicas para obtener la misma estimación de varianzas que con todas las réplicas posibles. Se forman *G* réplicas "balanceadas" utilizando una matriz Hadamard (Wolter, 1985), de manera que cada UPM seleccionada aparezca en el mismo número de réplicas y cada par de UPM seleccionadas procedentes de dos estratos diferentes aparezca también en el mismo número de réplicas. El número mínimo de réplicas necesarias *G* es el menor número entero que sea a la vez mayor o igual que *L* y divisible por 4. Por ejemplo, para 49 estratos, cada uno de ellos con dos UPM seleccionadas, harían falta 52 réplicas mediante BRR. Las observaciones en las UPM seleccionadas que no se han incluido en la réplica *j* tienen valor cero para la

variable de ponderación replicada *REPWTj*, y las observaciones en las UPM seleccionadas que se han incluido en la réplica *j* tienen un valor del doble de su ponderación muestral en la muestra completa, aunque ésta puede ser sometida a ajuste para las faltas de respuesta o a posestratificación.

38. Fay (Judkins, 1990) desarrolló una variante común de la técnica BRR debido a que la BRR estándar puede ser problemática si se desea una estimación de un campo pequeño o de un cociente de población en el que el denominador presenta pocos casos en la muestra completa. Según el método de Fay, las observaciones en las UPM seleccionadas que no son elegidas para la réplica *j* no toman el valor de cero como ocurre en la BRR estándar. En vez de ello, se reduce su ponderación muestral en un factor multiplicador K ($0 \leq K < 1$), mientras que se aumenta la ponderación muestral de las observaciones en las UPM seleccionadas escogidas para la réplica en un factor multiplicador $(2 - K)$. Haciendo $K = 0$ se llega a la técnica BRR estándar. Un valor que se suele recomendar para el método de Fay es $K = 0{,}3$.

4. Técnicas de replicación *jackknife* (JK)

39. La idea general de las técnicas *jackknife* consiste en eliminar cada vez una UPM seleccionada para formar réplicas, y a continuación volver a ponderar cada réplica según se requiera de forma que haga inferencia a la población representada por la muestra completa. Una UPM seleccionada podría estar formada por un solo elemento, como en el caso del muestreo aleatorio simple o el muestreo aleatorio estratificado, o contener varios elementos, como en la aproximación al plan de muestreo *WR*.

40. Consideremos en primer lugar el caso en el que no hay estratificación previa a la selección de las UPM y cada una de las G UPM seleccionadas (con el mismo número de elementos aproximadamente) se asemeja a la muestra completa. Se forma un total de G réplicas eliminando una UPM seleccionada cada vez. Para la réplica *j* con la variable de ponderación replicada *REPWTj*, las observaciones en la UPM seleccionada eliminada *j* tienen valor cero para *REPWTj*. Cada observación en las UPM seleccionadas restantes (no eliminadas) tiene un valor para *REPWGT* igual a la ponderación muestral para esa observación multiplicada por el factor $(G / (G - 1))$.

41. Un segundo ejemplo lo tenemos en el caso de L estratos con exactamente dos UPM seleccionadas por estrato; es decir, el diseño analizado anteriormente para la BRR. Eliminando una UPM seleccionada cada vez obtendríamos $2L$ réplicas. Para cada una de las $2L$ réplicas, la UPM seleccionada restante en el estrato al que pertenece la UPM seleccionada eliminada tendría la ponderación muestral de cada observación multiplicada por 2 (y la UPM seleccionada eliminada tendría la ponderación muestral de sus observaciones multiplicada por cero). No obstante, esta técnica se aplica generalmente con sólo L réplicas en vez de $2L$, y sólo se elimina una UPM seleccionada, escogida al azar, en cada uno de los L estratos. Para los estimadores lineales, el estimador de varianzas utilizando sólo las L réplicas es algebraicamente equivalente al estimador de varianzas utilizando las $2L$ réplicas.

42. El plan de muestreo más general es el muestreo polietápico estratificado con L estratos (antes de la selección de las UPM) y una selección de dos o más UPM por estrato. Cada UPM seleccionada se elimina para formar una réplica; el número de réplicas G es igual al número total de UPM seleccionadas de la muestra completa (n). En el estrato h, el valor de la variable de ponderación replicada *REPWTj* para cada observación de la UPM seleccionada eliminada es la variable de ponderación de la muestra *WTVAR* multiplicada por cero. El valor de la variable *REPWTj* para cada observación restante en el estrato h del que se eliminó la UPM seleccionada es la variable de ponderación de la muestra *WTVAR* multiplicada por el factor $(n_h / (n_h - 1))$, donde n_h es el número de UPM seleccionadas en el estrato h de la muestra completa.

5. Algunos errores comunes de los usuarios de *software* de estimación de varianzas

43. Varios de los paquetes de *software* requieren que el usuario ordene el conjunto de datos de entrada por alguna de las variables de diseño de la encuesta; por ejemplo, por *STRATVAR* y, dentro de *STRATVAR*, por *PSUVAR* (como se explica en el párrafo 35). La omisión de la ordenación previa puede producir varianzas estimadas incorrectamente, aunque la mayoría de los programas de *software* darán un mensaje de error si el conjunto de datos no está correctamente ordenado.

44. Un examen inadecuado de la documentación de la encuesta puede dar lugar a que los usuarios de conjuntos de datos para uso público especifiquen variables incorrectas de diseño de la misma. Si se especifica incorrectamente la variable de ponderación de la muestra se obtendrán estimadores sesgados y estimaciones incorrectas de la varianza; es decir, todos los análisis serán erróneos. Si la variable de ponderación de la muestra es correcta pero la variable de estratificación, la variable UPM o ambas son incorrectas, las estimaciones por puntos serán correctas pero las varianzas estimadas serán incorrectas.

45. Algunos conjuntos de datos para uso público tienen múltiples archivos de datos con distintas variables de diseño de la encuesta para cada archivo. Los distintos archivos de datos pueden tener distintas unidades de análisis (por ejemplo, personas, hogares o familias), por lo que es necesario prestar una cuidadosa atención a la interpretación de los resultados. Algunas variables de la encuesta sólo pueden medirse en una submuestra probabilística de la muestra completa, lo que precisa de una variable de ponderación de la muestra distinta de las variables que se miden en la muestra completa. Una lectura cuidadosa y completa de la documentación es esencial en todas las encuestas, tanto si su plan de muestreo es simple como si es desmesuradamente complejo.

D. Comparación de la estimación de varianzas con distintos paquetes de *software*

46. El sitio web informativo www.fas.harvard.edu/~stats/survey-soft/survey-soft.html tiene enlaces a toda una gama de paquetes de *software* para encuestas, incluidos los ocho examinados en este artículo. Véase también en Carlson (1998) un examen de los paquetes de *software* para datos de encuestas complejas. El SPSS no se encuentra entre los paquetes de *software* examinados. A primeros de 2003, el SPSS no tenía capacidad de estimación de varianzas de encuestas complejas, aunque a finales del mismo año, cuando se preparaba la publicación de este capítulo, salió al mercado un módulo adicional del programa.

47. En las páginas restantes de este capítulo se examina y compara la estimación de varianzas a partir de los datos de encuestas complejas según ocho paquetes de *software*: SAS, SUDAAN, STATA, Epi-Info, WesVar, PC-CARP, CENVAR e IVEware. El examen de los cinco primeros paquetes se ilustra con análisis descriptivos para los que se han empleado los datos procedentes de una encuesta realizada en Burundi en 1989; se estiman las proporciones, medias y totales de población, y se comparan los campos en estos parámetros. Los resultados de los análisis de la encuesta de Burundi se resumen en el cuadro XXI.1 del capítulo, y en el anexo incluido en el cederrón se ofrecen cuadros detalladas y ejemplos y resultados anotados para cada paquete. Los ejemplos anotados que se incluyen en el anexo pueden ayudar a los usuarios a aprender a utilizar los primeros cinco paquetes de *software* de estimación de varianzas.

48. De los cinco paquetes que se ilustran con los datos de la encuesta de Burundi, tres (STATA, SAS y Epi-Info) incluyen procedimientos para encuestas dentro de un paquete de *software* de estadística general. Los tres emplean la linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas. Los otros dos paquetes que se ilustran

Cuadro XXI.1
Comparación de procedimientos entre cinco paquetes de software: porcentaje estimado y número de mujeres que son seropositivas, con error estándar estimado; mujeres que han dado a luz recientemente, Burundi, 1988-1989

Paquete de software y procedimiento	Seropositivas			Seropositivas		
	Porcentaje	EE	IC 95%	Cantidad	EE	IC 95%
SAS 8.2 MEANS [a] Sin ponderación	74,88 Mal	2,12% Mal	N. a.	N. a.	N. a.	N. a.
SAS 8.2 MEANS [b] Con ponderación	67,20	2,30% mal	N. a.	N. a.	N. a.	N. a.
SAS 8.2 SURVEYMEANS	67,20	3,83%	59,38, 75,02	142.485	8.848,10	124.415, 160.556
SUDAAN 8.0 CROSSTAB y DESCRIPT Taylor y BRR	67,20	3,83%	N. d.	142.485	8.848,10	N. d.
STATA 7.0 Svymean	67,20	3,83%	58,38, 75,02	N. d.	N. d.	N. d.
STATA 7.0 Svytotal	N. d.	N. d.	N. d.	142.485	8.848,10	124.415, 160.556
Epi-Info 6.04 [d] CSAMPLE [c]	67,20	3,83%	59,70, 74,71 [c]	N. d.	N. d.	N. d.
WesVar 4.2	67,20	3,83%	59,38, 75,02	142.485	8.848,10	124.415, 160.556

Abreviaturas utilizadas:
IC = Intervalo de confianza,
N. a. = no aplicable,
N. d. = no disponible,
EE = error estándar.

[a] Análisis incorrectamente especificado; no tiene en cuenta la ponderación muestral, la clasificación ni la estratificación.

[b] Análisis incorrectamente especificado; incorpora la ponderación muestral pero no la clasificación ni la estratificación.

[c] El intervalo de confianza que ofrece Epi-Info 6.04d es más estrecho que el de otros paquetes de *software*. Epi-Info 6.04d usaba $z = t = 1{,}96$ para construir el intervalo de confianza del 95%, mientras que los demás paquetes de software usaban $t = 2{,}042$ de la distribución *t* de Student, con 30 gld (grados de libertad del denominador para la encuesta, calculados como el número de UPM menos el número de pseudoestratos). Es preferible usar el valor gld de la propia encuesta.

(WesVar y SUDAAN) fueron desarrollados expresamente para la estimación de varianzas de encuestas. WesVar emplea métodos de replicación, y SUDAAN ofrece tanto la linealización mediante el desarrollo en serie de Taylor como métodos de replicación.

49. Se examinan también otros tres paquetes de *software* (PC-CARP, CENVAR e IVEware), pero no se ilustran con los datos de la encuesta de Burundi. Tanto PC-CARP como CENVAR emplean la linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas. IVEware emplea tanto la linealización mediante el desarrollo en serie de Taylor como métodos de replicación.

50. Los ocho paquetes examinados incluyen muchas de las opciones posibles de estimación de varianzas de encuestas, pero no la totalidad. Tres de ellos (Epi-Info, CENVAR y WesVar 2) se escogieron porque ofrecen análisis descriptivos básicos y pueden descargarse de Internet gratuitamente, lo que será de interés para los analistas con poco o ningún presupuesto para compras de *software*. Otros dos (PC-CARP y WesVar 4) se escogieron porque, aunque no son gratuitos, su costo es bajo en comparación con otras opciones y porque ofrecen análisis descriptivos y regresión lineal y logística basada en el diseño. Dos paquetes de precio moderado (SUDAAN y STATA) se escogieron porque ofrecen, además de análisis descriptivos, una amplia variedad de modelos de regresión basados en el diseño. SAS se escogió porque, aunque es un paquete caro, domina en el campo de la gestión y el análisis de datos, y por sus procedimientos relativamente nuevos de análisis de datos de las encuestas. Por último, la nueva aplicación IVEware (versión beta) se escogió porque ofrece amplios análisis descriptivos y modelos de regresión basados en el diseño, junto con múltiples procedimientos de imputación. IVEware es gratuito (se puede descargar de Internet), pero se ejecuta como aplicación de *software* invocada desde SAS, por lo que es necesario disponer de este paquete.

51. En el cuadro XXI.2 se resume una amplia variedad de características de los ocho paquetes de *software*, incluidos los planes de muestreo que cubren, métodos de estimación de varianzas y tipos de análisis.

Cuadro XXI.2

Atributos de ocho paquetes de software con capacidad de estimación de varianzas para datos de encuestas complejas

Atributo	SAS 8.2	SUDAAN 8.0	STATA 8.0	Epi-Info 6.04d	WesVar 4.2	PC-CARP	CENVAR	IVEware
Serie de Taylor	Sí	Sí	Sí	Sí	No	Sí	Sí	Sí Desc.
Métodos de replicación BRR y *jackknife*	No	BRR JK	No	No	BRR JK	No	No	Modelos JK
Formación de ponderaciones replicadas	No	No-BRR Sí-JK	No	No	Sí BRR/JK	No	No	Sí JK
Conjunto de datos de entrada	SAS	SAS, SPSS, ASCII	STATA	Epi-Info	SAS, SPSS, STATA, ASCII, ODBC	ASCII	ASCII	SAS
Estimación del total	Sí	Sí	Sí	No	Sí	Sí	Sí	No
IC sobre total	Sí	No	Sí	No	Sí	Sí	Sí	No
CL sobre totales	No	Sí	Sí	No	Sí	Sí	Sí	No
Estimación de la media	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
IC sobre media	Sí	No	Sí	Sí–estrecho	Sí	Sí	Sí	Sí
CL sobre medias	No	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Estimación de proporciones	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
IC sobre proporción	Sí	No	Sí	Sí–estrecho	Sí	Sí	Sí	Sí
CL sobre proporciones	No	Sí	Sí	Sí–error	Sí	Sí	Sí	Sí
Estimación de razón	Sí	Sí	Sí	No	Sí	Sí	Sí	No
IC sobre razón	Sí	No	Sí	No	Sí	Sí	Sí	No
CL sobre razones	No	Sí	Sí	No	Sí	Sí	Sí	No
Análisis de campos	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Comparación de campos	No–8.2 Sí–9.0	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Análisis de subpoblaciones	No	Sí	Sí	No	Sí	Sí	Sí	Sí
Razones/medias estándar	No–8.2 Sí–9.0	Sí	Sí	No	Sí	No	No	No
Pruebas de ji-cuadrado	No–8.2 Sí–9.0	Sí	Sí	No	Sí	Sí	No	No
Regresión logística	No	Sí	Sí	No	Sí	Sí	No	Sí
Cociente de disparidad	No	Sí	Sí	Sí	Sí	Sí	No	Sí
Riesgo relativo	No	Sí	Sí	Sí	Sí	No	No	No
Regresión lineal	Sí	Sí	Sí	No	Sí	Sí	No	Sí
Modelos de regresión adicionales	No	Sí	Sí	No	No	No	No	Sí
Describe muchas etapas muestrales	No	Sí	No	No	No	No	No	No
Efecto del diseño	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Software de prueba gratuita	No	No	No	N. d. Gratuito	Sí	No	N. d. Gratuito	N. d. Gratuito
Paquete de estadística general	Sí	No	Sí	Sí	No	No	No	No
Capacidad de gestión de datos	Sí	No	Sí	Sí	Sí	No	No	No
Se ejecuta mediante programas de entrada de datos	Sí	Sí	Sí	No–6.04d Sí–2002	No	No	No	Sí
Se ejecuta mediante mandatos abreviados	No	No	Sí	No	No	No	No	No
Se ejecuta mediante selecciones del menú	No	No	No	Sí	Sí	Sí	Sí	No
Ordena el conjunto de datos por estratos y UPM	No	Sí	No	Sí	No	Sí	Sí	No
Los creadores ofrecen formación	Sí	Sí	Sí	No	Sí	No	No	No
Manual escrito /en línea	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Guías de procedimientos paso a paso	No	No	No	No	Sí	No	Sí	No
Precio	Elevado	Medio	Medio	Gratuito	Bajo–V4 Gratis–V2	Bajo	Gratuito	Gratuito
Cuota anual renovable	Elevada	Medio	No tiene	No tiene	No tiene	No tiene	No tiene	No tiene
Imputación de datos	No	No	No	No	No	Sí	No	Sí

Abreviaturas utilizadas: ASCII = American Standard Code for Information Interchange, BRR = replicación repetida y equilibrada, IC = intervalo de confianza, JK = *jackknife*, CL = contraste lineal, N. d. = no disponible, ODBC = Open DataBase Connectivity, V = versión.

E. El conjunto de datos de la encuesta de Burundi

52. Para todos los ejemplos numéricos de este capítulo se ha utilizado un conjunto de datos de una encuesta sobre cobertura inmunitaria con toxoide tetánico (TT) realizada en 1989. A continuación se presenta un breve resumen del diseño de la encuesta de Burundi (en la sección I del anexo incluido en el cederrón se aportan más detalles). Para información adicional sobre la metodología de esta encuesta y sus resultados publicados, puede consultarse el informe del Programa Ampliado de Inmunización (PAI) (1996) de la Organización Mundial de la Salud (OMS).

1. Población de inferencia y parámetros de población

53. La población de inferencia para esta encuesta es el conjunto de mujeres de Burundi que dieron a luz entre la Pascua de Resurrección de 1988 y febrero–marzo de 1989. El parámetro de población que nos interesa es el porcentaje (o la proporción) de mujeres que resultaron seropositivas a la antitoxina tetánica, protegiendo así a sus bebés contra el tétanos neonatal.

2. Plan de muestreo y recogida de datos

54. El plan de muestreo era una modificación propuesta por Brogan y otros (1994) de la metodología de encuesta por conglomerados desarrollada por la OMS para su Programa Ampliado de Inmunización. Con la modificación se obtiene una muestra probabilística de viviendas o unidades habitacionales y, por consiguiente, una muestra probabilística de mujeres, lo que no sería posible con la metodología estándar de muestreo por conglomerados del PAI de la OMS (ibídem).

55. El país fue estratificado en dos zonas geográficas: la capital, Bujumbura (estrato urbano), y el resto del territorio (estrato rural). Las unidades primarias de muestreo (UPM) eran las zonas geográficas; *collines* en el estrato rural y *quartiers* o *avenues* en el estrato urbano. El marco de muestreo de las UPM para cada estrato se ordenó por proximidad geográfica. Mediante un muestreo sistemático con probabilidad proporcional al tamaño estimado (PPTE) se seleccionaron 30 UPM por estrato. Dado que el 96% de la población de inferencia reside en las zonas rurales de Burundi y que se asignó el mismo número de UPM seleccionadas a cada estrato, las mujeres urbanas estaban sustancialmente sobrerrepresentadas en la muestra. La ordenación específica de las UPM en el marco de muestreo, combinada con el muestreo sistemático con PPTE de las UPM, produce una estratificación geográfica implícita en cada estrato.

56. Para obtener una muestra de viviendas ocupadas se desarrollaron etapas adicionales de muestreo probabilístico en las UPM seleccionadas. Para la muestra se seleccionaron todas las mujeres que reunían los requisitos de la encuesta en cada vivienda seleccionada. La seropositividad según el título de anticuerpos de la antitoxina tetánica se determinó mediante una muestra de sangre obtenida por punción de la yema del dedo. La tasa de respuesta de la encuesta fue prácticamente del 100%, cifra inusualmente elevada. Se entrevistó a un total de 206 mujeres urbanas y 212 mujeres rurales.

3. Procedimientos de ponderación y preparación para la estimación de varianzas

57. La variable de ponderación de la muestra *W* incluida en el conjunto de datos de la encuesta de Burundi se reajustó a *W2* de manera que el valor de *W2* para una entrevistada *R* de la muestra fuera una estimación del número de mujeres de la población de inferencia representadas por esa *R*. El valor de *W2* es aproximado y sólo se emplea para ilustrar la estimación de los totales de población con los distintos paquetes de *software*. Del análisis in-

cluido en este capítulo no deberían extraerse conclusiones sustantivas referentes a los totales de población para las mujeres que reunían los requisitos de la encuesta de Burundi en 1989. Es importante señalar que las proporciones y medias estimadas que se mencionan en este capítulo coinciden con los resultados publicados previamente con este conjunto de datos (Programa Ampliado de Inmunización, 1996), puesto que el valor revisado de *W2* es un múltiplo escalar del valor de *W* utilizado en análisis previos. El valor de *W2* era de 959,3 para las mujeres rurales de la muestra y de 42,0 para las mujeres urbanas de la muestra, lo que reflejaba la sustancial sobrerrepresentación de las mujeres urbanas en la muestra. En el plan de muestreo de la encuesta de Burundi se utilizó la aproximación mediante la descripción común *WR* para la estimación de varianzas, es decir, el enfoque EVCF con fracciones bajas de muestreo en la primera etapa.

58. Dado que las UPM estaban estratificadas implícitamente, en el plan de muestreo en cada uno de los estratos, urbano y rural, se consideró que se seleccionaban dos UPM de cada uno de 15 pseudoestratos. La descripción del plan de muestreo como un total de 30 pseudoestratos, cada uno de ellos con dos UPM seleccionadas, es preferible a la de 2 estratos, cada uno de ellos con 30 UPM seleccionadas, porque la primera produce una estimación de varianzas menos sesgada debido a que tiene en cuenta la estratificación implícita. La variable de pseudoestratificación, *PSTRA,* se codificó con los valores 1 a 30 y la variable pseudo-UPM, *PPSU,* se codificó con los valores 1 ó 2 en cada pseudoestrato.

59. Cuando se emplea la linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas sólo se necesitan las variables *W2*, *PSTRA* y *PPSU*. Cuando se emplean técnicas de replicación, sin embargo, se necesitan ponderaciones replicadas. Se utilizó WesVar para calcular ponderaciones replicadas mediante BRR de las variables *W2*, *PSTRA* y *PPSU*. Estas ponderaciones replicadas calculadas por WesVar se utilizaron tanto en WesVar como en SUDAAN para la estimación de varianzas mediante BRR.

4. Tres ejemplos de análisis de datos de encuestas

60. Las secciones II a VI del anexo en cederrón contienen análisis comentados de datos de los tres ejemplos que se presentan a continuación, para los que se emplearon cinco paquetes de *software* para datos de encuestas. Los ejemplos siguientes ilustran los análisis descriptivos y analíticos comunes a los que se someten los datos de encuestas, a saber: *a*) estimación de proporciones, totales y medias, tanto de la totalidad de la población como de los campos o estratos, y *b*) comparación de las medias o proporciones de los distintos campos o estratos. La población de inferencia es el conjunto de mujeres que reunían los requisitos de la encuesta en Burundi a comienzos de 1989.

- Ejemplo 1: Estimar el número de mujeres (total de población) y porcentaje de mujeres (proporción/porcentaje de la población) que eran seropositivas (variable *IMMUNE*, 1 = seropositiva, 2 = seronegativa). La variable *BLOOD* (1 = seropositiva, 0 = seronegativa) es una recodificación de *IMMUNE*.
- Ejemplo 2: Estimar los parámetros de población del ejemplo 1 entre las mujeres urbanas y rurales (variable *RUR_URB*, codificada como 1 = rural, 2 = urbana). Determinar si el hecho de que la residencia sea rural o urbana es estadísticamente independiente de la seropositividad (*IMMUNE*).
- Ejemplo 3: Estimar las unidades internacionales medias de antitoxina por mililitro (ml) (*IUML*) para la población de mujeres de inferencia y por tipo de residencia (rural o urbana). Determinar si existe alguna relación entre la media *IUML* y el hecho de que la residencia sea rural o urbana.

  Nota: la estimación de las unidades internacionales medias de antitoxina por mililitro (ml) (*IUML*) puede ofrecer resultados engañosos debido a la distribución

asimétrica hacia la derecha de esta variable. Quizá sería mejor usar la mediana o transformar las *IUML* antes del análisis, por ejemplo, en el logaritmo natural de *IUML*. En este capítulo se estima la media *IUML* (sin transformación) para mostrar las posibilidades de los cinco paquetes de *software*, no para ilustrar resultados sustantivos referentes a las *IUML*.

F. Empleo de procedimientos ajenos a las muestras para analizar datos de encuestas

61. En este apartado se muestra cómo el uso incorrecto de fórmulas de muestras aleatorias simples para el análisis de datos de encuestas complejas puede dar lugar a estimaciones por puntos sesgadas y a sesgos (por lo general muy pequeños) en los errores estándar estimados. Véase otra ilustración al respecto en Brogan (1998 y en curso de publicación). Para ilustrar este hecho se podría haber usado cualquier paquete de *software* estadístico y se hubieran obtenido respuestas comparables a las que se obtuvieron con SAS (utilizado en este apartado).

62. El parámetro de población que ha de estimarse es la proporción de mujeres seropositivas de la población de inferencia. Se calcula la variable indicadora *BLOOD* y se codifica como 1 = seropositiva y 0 = seronegativa. Así, la media de *BLOOD* indicará la proporción de mujeres seropositivas. En SAS, el procedimiento PROC MEANS estima la media de *BLOOD* en 0,74880, con un error estándar estimado de 0,02124 (fila 1 del cuadro XXI.1). Estos dos cálculos están sesgados porque PROC MEANS, incorrectamente aplicado, no tiene en cuenta la variable de ponderación de la muestra al estimar la proporción de la población, y además pasa por alto las variables de ponderación muestral, UPM y estratificación, al calcular el error estándar estimado de la estimación por puntos.

63. A continuación se emplea el procedimiento PROC MEANS de SAS con la variable de ponderación de la muestra *W2* en una sentencia WEIGHT. La media de *BLOOD* se estima en 0,67203, con un error estándar estimado (ponderado) de 0,02299 (fila 2 del cuadro XXI.1). En este análisis, PROC MEANS ofrece una adecuada estimación por puntos de la proporción de la población. No obstante, la aplicación incorrecta de PROC MEANS da lugar a un sesgo en el error estándar estimado, debido a que pasa por alto las variables de estratificación y UPM.

64. Por último, se emplea el procedimiento PROC SURVEYMEANS de SAS para obtener un análisis adecuado para datos de encuestas complejas. (En el siguiente apartado se explica cómo usar SURVEYMEANS). La estimación por puntos de la proporción de la población es 0,67203, con un error estándar estimado de 0,03830 (fila 3 del cuadro XXI.1). PROC SURVEYMEANS tiene en cuenta la variable de ponderación de la muestra para calcular la estimación por puntos y las variables de ponderación muestral, UPM y estratificación y determinar el error estándar estimado.

65. La comparación entre estos tres análisis que se presenta en el cuadro XXI.1 muestra que la estimación por puntos no ponderada (incorrecta) de 0,7488 (74,88%) difiere bastante de la estimación por puntos ponderada (correcta) de 0,6720 (67,20%). La estimación por puntos no ponderada es demasiado alta porque existe una mayor proporción de seropositivas entre las mujeres urbanas que entre las mujeres rurales (como se ilustra más adelante en este capítulo), y las mujeres urbanas están sobrerrepresentadas en la muestra porque se incluyó en la muestra un número excesivo de ellas: constituyen alrededor de la mitad de la muestra, pero sólo el 4% de la población de inferencia. Así, en un análisis no ponderado a partir del cual se quieran hacer inferencias sobre el conjunto del país, se atribuye a las mujeres urbanas mucha más influencia de la debida y se produce un sesgo por exceso en la proporción de la población estimada.

66. La comparación entre los dos análisis que producen la estimación por puntos ponderada correcta ilustra que, incluso con una sentencia WEIGHT la aplicación incorrecta de PROC MEANS en SAS da lugar a una grave subestimación del error estándar, con un cálculo incorrecto de 0,02299 (2,30%) frente al cálculo correcto de PROC SURVEYMEANS de 0,03830 (3,83%). Esto ocurre, fundamentalmente, porque PROC MEANS, con o sin una sentencia WEIGHT, pasa por alto la clasificación de las mujeres en las UPM seleccionadas, mientras que SURVEYMEANS reconoce la clasificación a efectos de la estimación de varianzas. Como el coeficiente de correlación intraclase es positivo para la mayoría de las variables medidas de las encuestas complejas, los procedimientos correctos de estimación de varianzas que tienen en cuenta la clasificación suelen arrojar valores mayores para los errores estándar estimados.

67. En general, se obtienen estimaciones por puntos sesgadas de los parámetros de población si los datos de la encuesta no se analizan con la variable adecuada de ponderación de la muestra. Además, incluso si se incorpora al análisis la variable de ponderación de la muestra, obteniéndose estimaciones por puntos adecuadas de los parámetros de población, los errores estándar suelen subestimarse cuando los elementos muestrales aparecen por conglomerados en los datos de la encuesta y la clasificación no se reconoce en la estimación de varianzas. La subestimación de los errores estándar da lugar a unos intervalos de confianza demasiado estrechos y a pruebas estadísticas de significación con valores *p* demasiado pequeños. En otras palabras, se sobrevalora el nivel de significación estadística.

68. Una aproximación a la magnitud de la subestimación de la varianza al pasar por alto la clasificación de los datos de la encuesta viene dada por la expresión $[1 + \rho\ (b - 1)]$, donde ρ es el coeficiente de correlación intraclase entre los elementos de la población y b es el número promedio de elementos muestrales por conglomerado muestral (UPM) (véase el capítulo VI). Por ejemplo, si el valor de la expresión es 2, al tener en cuenta la clasificación prácticamente se duplica la varianza estimada que se obtendría sin tenerla en cuenta. Obsérvese que la variable UPM de la encuesta de Burundi denominada *PPSU* permite que el *software* identifique qué elementos muestrales forman un conglomerado dentro de una misma UPM seleccionada (para un estrato dado).

69. Además del efecto de la clasificación sobre la varianza estimada, una variación sustancial del valor de la variable de ponderación de la muestra entre los entrevistados da lugar a un aumento de la varianza estimada. Así, si la variable de ponderación de la muestra se pasa por alto en el análisis, se subestima el error estándar estimado (y el estimador del parámetro de población resulta sesgado).

G. Procedimientos para análisis de encuestas en SAS 8.2

1. Descripción general de SURVEYMEANS y SURVEYREG

70. La versión 8.2 de SAS contiene SURVEYMEANS y SURVEYREG, dos procedimientos para el análisis de datos de encuestas desarrollados recientemente (aparecieron por primera vez en la versión 8.0). SAS incluye la descripción del plan de muestreo común *WR*, para la que se requieren las tres variables básicas de diseño de la encuesta. Es posible aplicar factores de corrección al muestreo de poblaciones finitas a los diseños de muestreo monoetápico como el muestreo aleatorio estratificado y el muestreo aleatorio simple. Para la estimación de varianzas se usa la linealización mediante el desarrollo en serie de Taylor. La versión 9 de SAS contiene dos nuevos procedimientos para datos de encuestas complejas: SURVEYFREQ para análisis de variables categóricas y SURVEYLOGISTIC para regresión logística. Otros procedimientos adicionales para el análisis de datos de encuestas con SAS se encuentran en desarrollo.

71. La sintaxis para especificar las variables de diseño de la encuesta necesarias para *WR* es igual para SURVEYMEANS que para SURVEYREG. Se emplea la palabra clave STRATA para designar la variable de estratificación; la palabra clave CLUSTER, para designar la variable UPM, y la palabra clave WEIGHT, para especificar la variable de ponderación de la muestra (al igual que en otros procedimientos de SAS, como MEANS). Estas sentencias, expresadas de la manera adecuada para una encuesta dada, deben estar en cada procedimiento de SAS relacionado con la encuesta; por lo general se mantendrán invariables durante todo el análisis de un mismo conjunto de datos de encuesta. Para el conjunto de datos de la encuesta de Burundi, las sentencias SAS que se indican a continuación describen el diseño de la encuesta para los procedimientos SURVEYMEANS o SURVEYREG:

STRATA	*PSTRA*
CLUSTER	*PPSU*
WEIGHT	*W2*

72. Si se omite la sentencia STRATA, SAS supone que en el plan de muestreo no había estratificación de UPM previa al de la primera etapa; si se omite la sentencia CLUSTER, supone que los elementos muestrales no están clasificados, es decir, que cada conglomerado muestral contiene exactamente un elemento, esto es, que los elementos se seleccionaron en la primera (y única) etapa de muestreo, o sea, que se utilizó el muestreo aleatorio simple o el muestreo aleatorio estratificado, y si se omite la sentencia WEIGHT, supone que cada *R* tiene el mismo valor para la variable de ponderación y asigna el valor 1,0 a esa variable. Si se omiten las tres sentencias de diseño de la encuesta (STRATA, CLUSTER y WEIGHT), ello equivale a especificar un muestreo aleatorio simple de una población infinita, que es la premisa de la mayor parte de los procedimientos ajenos a las encuestas de SAS.

2. SURVEYMEANS

73. Este procedimiento permite estimar las medias y totales de población para variables continuas, y las proporciones y totales de población para las variables categóricas, utilizando los datos de la encuesta. Se obtienen los errores estándar estimados y los coeficientes de variación para todas las estimaciones por puntos, así como los intervalos de confianza de los parámetros de población. En la sentencia PROC se puede pedir información estadística concreta, o aceptar el listado de información estadística por defecto, o incluir ALL en la sentencia PROC para obtener toda la información estadística que se puede calcular con SURVEYMEANS.

74. Las variables que han de analizarse (tanto continuas como categóricas) aparecen en la sentencia VAR. La sentencia CLASS produce listados de las variables categóricas que aparecen en la sentencia VAR; SAS supone entonces que todas las demás variables de la sentencia VAR son continuas.

75. Para especificar campos para el análisis de todas las variables de la sentencia VAR se usa la sentencia DOMAIN con una o más variables categóricas. SAS ofrece automáticamente análisis de la población marginal (en otras palabras, de la totalidad de la población), además del análisis de campos. Un programa sin la sentencia DOMAIN produce estimaciones solo para la totalidad de la población. Aunque la sentencia BY en SURVEYMEANS se puede usar para obtener estimaciones para campos, esto no es recomendable para los datos de encuestas porque cuando se emplea la sentencia BY no se usan las fórmulas adecuadas para la estimación de varianzas. Para el análisis de campos debe utilizarse la sentencia DOMAIN.

76. En la versión 8.2 de SAS no hay ninguna sentencia que permita el análisis de una subpoblación, por ejemplo, solo de mujeres mayores. Sí es posible realizar análisis de subpoblaciones definiendo primero una variable indicadora como, por ejemplo, *OLDERFEM*, que indica si el elemento muestral pertenece a la subpoblación. A continuación se puede

utilizar la sentencia DOMAIN *OLDERFEM* para lograr los análisis deseados; sólo hay que descartar los resultados de SAS para los elementos muestrales que no sean mujeres mayores. En SAS no debe utilizarse la sentencia IF para extraer del conjunto de datos el subconjunto de mujeres mayores antes de ejecutar el procedimiento SURVEYMEANS, porque los errores estándar pueden calcularse incorrectamente en la medida en la que SURVEYMEANS desconozca el número total de estratos y de UPM seleccionadas de la encuesta.

3. SURVEYREG

77. Con este procedimiento se realiza una regresión lineal de los datos de encuestas según el enfoque basado en el diseño (Korn y Graubard, 1999); esto es, el análisis tiene en cuenta las variables de diseño de la encuesta. Como en el caso de la regresión lineal para datos ajenos a encuestas, la variable dependiente es continua (o se supone que lo es) y las variables independientes pueden ser una mezcla de variables continuas y categóricas. La sentencia MODEL incluye la variable dependiente y todas las variables independientes. Todas las variables categóricas que aparezcan en la sentencia MODEL deben aparecer también en la sentencia CLASS, y la sentencia CLASS debe preceder a la sentencia MODEL en el programa SAS. SURVEYREG forma variables indicadoras ficticias (con codificación 1 ó 0) para las variables categóricas independientes, definiéndose como grupo de referencia aquel con el mayor valor codificado de la variable. Las demás opciones de SURVEYREG, así como sus resultados, son similares a la regresión lineal (ajena a las encuestas) de SAS.

78. La versión 8.2 de SAS carece de procedimientos de encuestas para comparar medias o proporciones de campos, aunque estas capacidades se encuentran en desarrollo. Una pregunta que sirve de ejemplo de esta situación es: ¿en la población de inferencia de la encuesta de Burundi, las mujeres rurales y las urbanas difieren en las unidades medias *IUML*, o en la proporción de seropositivas? En la versión 9.0 se puede utilizar SURVEYFREQ para realizar una prueba de ji-cuadrado de las dos variables residencia (rural/urbana) y seropositividad (Sí/No). Hasta que se terminen de desarrollar en SAS los procedimientos de comparación de campos para datos de encuestas, SURVEYREG puede utilizarse para comparar campos como se indica a continuación.

79. Si se desea comparar la media *IUML* de las mujeres rurales y urbanas de la población de inferencia, úsese la sentencia MODEL en SURVEYREG con la variable continua *IUML* como variable dependiente, y la variable de campo que designa la residencia rural o urbana como variable categórica independiente. Parte del resultado estándar de SURVEYREG es una prueba de la hipótesis de nulidad de que el coeficiente de regresión de la población para la residencia rural o urbana (con un grado de libertad) es igual a cero. Esta hipótesis de nulidad referente al coeficiente de regresión equivale a la hipótesis de nulidad de que las mujeres rurales y las urbanas de la población de inferencia tienen la misma media *IUML*.

80. Si se desea comparar la proporción de seropositivas (variable binaria) entre las mujeres rurales y urbanas de la población de inferencia, utilícese la variable indicadora *BLOOD* (1 = seropositiva, 0 = seronegativa) como variable dependiente. (Obsérvese que *BLOOD* es simplemente una recodificación de la variable *IMMUNE*, donde 1 = seropositiva y 2 = seronegativa). En la sentencia MODEL de SURVEYREG, defínase *BLOOD* como la variable dependiente, y la variable de campo que designa la residencia (rural o urbana) como la variable categórica independiente. La hipótesis de nulidad de que el coeficiente de regresión es cero equivale a la hipótesis de nulidad de que la proporción de seropositivas es igual para las mujeres rurales que para las urbanas de la población de inferencia.

4. Ejemplos numéricos

81. En la sección II del anexo incluido en el cederrón se ilustra el uso de SURVEYMEANS y SURVEYREG para desarrollar los tres ejemplos que se indican en el párrafo 60

supra. El examen de los programas SAS comentados (escritos por los usuarios) y de los resultados comentados de SAS preparará al lector para escribir sus propios programas SAS para SURVEYMEANS y SURVEYREG e interpretar los resultados.

82. En la fila 3 del cuadro XXI.1 se resume el resultado de SURVEYMEANS de la sección II del anexo para la estimación del porcentaje y el número de mujeres seropositivas de la población de inferencia de la encuesta de Burundi, con el error estándar estimado y el intervalo de confianza; la mayoría de estos resultados se analizaron en el apartado F de este capítulo. En la fila 1 del cuadro XXI.3 (sección VII del anexo incluido en el cederrón) se resume el resultado de SURVEYMEANS para la estimación del porcentaje de seropositivas en cada uno de los dos campos —mujeres rurales y mujeres urbanas—, 66,51% y 83,50% respectivamente. En la fila 1 del cuadro XXI.4 (sección VII del anexo contenido en el cederrón) se resume el resultado de SURVEYMEANS al comparar las mujeres rurales con las urbanas, que arroja un valor de la t de –3,52 con un valor p de 0,0014 para probar la hipótesis de nulidad de que las mujeres rurales y las urbanas no difieren en el porcentaje de seropositivas. Así pues, las mujeres rurales y urbanas de la población de inferencia difieren en el porcentaje de seropositivas: las mujeres urbanas tienen una mayor tasa de prevalencia de la seropositividad.

5. Ventajas e inconvenientes y costo

83. Quien ya sea usuario de SAS/STAT puede obtener sin costo alguno los procedimientos para encuestas de SAS y utilizar la sintaxis con la que está familiarizado. También puede disponer de todas las capacidades de SAS para la gestión de datos y la formación de variables nuevas. El sistema general de asistencia de SAS incluye la asistencia técnica y documentación para los procedimientos para encuestas. En comparación con el costo de los demás paquetes para encuestas examinados, el costo de SAS es elevado.

84. SAS 8.2 no tiene capacidad para comparar campos entre sí, aunque se puede usar SURVEYREG como solución provisional para este tipo de análisis. Si se añade la versión 9.0 de SURVEYFREQ se facilitará la comparación de variables categóricas entre campos.

85. SAS sólo emplea la linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas. Para el muestreo polietápico por conglomerados sólo utiliza la descripción del plan de muestreo común *WR*. No obstante, puede incorporar términos *fpc* en muestreos monoetápicos aleatorios estratificados o muestreos aleatorios simples.

86. La capacidad de SAS 8.2 para el análisis de datos de encuestas es básica y descriptiva y puede ser suficiente para las necesidades de análisis de muchos usuarios. La adición de SURVEYFREQ en la versión 9.0 proporciona capacidad descriptiva y analítica para variables categóricas. Los procedimientos para encuestas que se encuentran en desarrollo, como, por ejemplo, la regresión logística, harán de SAS, en el futuro, un paquete más comparable con otros paquetes de *software* que permiten someter las encuestas a análisis exhaustivos.

H. SUDAAN 8.0

1. Descripción general de SUDAAN

87. SUDAAN (Research Triangle Institute, 2001) es un paquete de *software* especializado desarrollado originalmente para el análisis de datos de encuestas complejas, pero ahora de uso generalizado para el análisis de datos correlacionados, con técnicas como el análisis de datos longitudinal y las ecuaciones de estimación generalizadas (GEE). SUDAAN es un acrónimo de *SUrvey DAta ANalysis* (análisis de datos de encuestas). Los procedimientos de estadística descriptiva y analítica son DESCRIPT, CROSSTAB y RATIO. Los procedi-

mientos de modelado basado en el diseño comprenden la regresión lineal, regresión logística (incluida la multinomial), regresión log-lineal y análisis de supervivencia.

88. SUDAAN 8.0 está programado en lenguaje C, con sentencias de mandatos preparadas por el usuario similares a las de SAS. Los conjuntos de datos de entrada pueden ser archivos SAS, SPSS o ASCII. SUDAAN puede ejecutarse como aplicación independiente (SUDAAN autónomo) o invocarse desde SAS (SUDAAN bajo SAS). En general, para los usuarios de SAS resulta más interesante el SUDAAN bajo SAS.

89. SUDAAN es el único paquete de análisis de encuestas que incluye los dos enfoques más corrientes de la estimación de varianzas: la linealización mediante el desarrollo en serie de Taylor y los métodos de replicación. En SUDAAN, este último enfoque incluye la replicación repetida y equilibrada (BRR), con o sin el factor de ajuste de Fay, y los métodos *jackknife*. Todos los métodos de replicación de SUDAAN se basan en la descripción del plan de muestreo común al que nos hemos referido anteriormente como *WR*. Si se emplea la BRR para la estimación de varianzas, el conjunto de datos de entrada debe incluir las ponderaciones replicadas para la BRR; SUDAAN no genera ponderaciones replicadas para BRR. SUDAAN generará ponderaciones replicadas para el método *jackknife* de eliminación de una UPM, o aceptará las ponderaciones replicadas *jackknife* incluidas en el conjunto de datos de entrada para el método *jackknife* citado y las variantes del mismo.

90. Para SUDAAN, el diseño de la encuesta se describe mediante tres sentencias: *a*) la sentencia PROC, con la que se escoge una opción para la palabra clave DESIGN; *b*) la sentencia NEST, con la que se especifican las variables de estratificación y clasificación; y *c*) la sentencia WEIGHT, con la que se especifica la variable de ponderación de la muestra. El conjunto de datos de entrada de SUDAAN debe ordenarse por todas las variables que aparezcan en la sentencia NEST, por lo general la variable de estratificación de la primera etapa y a continuación por la variable UPM de cada estrato.

91. Al contrario que con la mayoría de los paquetes de *software* con capacidad de análisis de encuestas, en SUDAAN es posible describir la segunda y siguientes etapas de muestreo y estratificación en el muestreo polietápico para la estimación de varianzas, con lo que se evita tener que utilizar siempre la descripción del plan de muestreo común *WR*. Además, SUDAAN tiene amplia capacidad para incorporar a la estimación de varianzas los términos de corrección al muestreo de poblaciones finitas (*fpc*) en múltiples etapas del muestreo sin reemplazamiento. El manual de SUDAAN, disponible impreso o como archivo pdf, ofrece varios ejemplos de descripción de los planes de muestreo en SUDAAN (véase el capítulo III).

92. En SUDAAN, el plan de muestreo por defecto es *WR,* como se ha definido anteriormente, ya sea para linealización mediante el desarrollo en serie de Taylor, BRR o *jackknife*. En SUDAAN, el empleo de la sintaxis DESIGN = *WR* en la sentencia PROC invoca no sólo el enfoque EVCF y el muestreo en la primera etapa con reemplazamiento o sin reemplazamiento pero con fracciones pequeñas de muestreo en la primera etapa, sino también el uso de la linealización mediante el desarrollo en serie de Taylor. Con DESIGN = *WR*, la sentencia NEST contiene una o más variables de justificación (generalmente una sola) y una variable UPM. Si en la sentencia PROC se omite la opción DESIGN =, SUDAAN supone que DESIGN = *WR*.

93. En SUDAAN, la sintaxis DESIGN = BRR invoca la descripción del plan de muestreo común *WR* (como se ha comentado anteriormente) con replicación repetida y equilibrada para la estimación de varianzas. Las variables de ponderación replicadas mediante BRR deben estar en el conjunto de datos de entrada, y la sentencia REPWGT del programa SUDAAN da los nombres de variables de las variables de ponderación replicadas.

94. En SUDAAN, la sintaxis DESIGN = JACKKNIFE, en ausencia de las sentencias JACKWGTS y JACKMULT, invoca la descripción del plan de muestreo común

WR con estimación de varianzas mediante la técnica *jackknife* de eliminación de una UPM, donde SUDAAN genera las ponderaciones replicadas *jackknife*. La sintaxis de SUDAAN DESIGN = JACKKNIFE, con la sentencia JACKWGTS, invoca la descripción del plan de muestreo común *WR* con las ponderaciones *jackknife* aportadas a SUDAAN como variables en el conjunto de datos de entrada.

95. El diseño de la encuesta de Burundi y la especificación de la linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas se describen en SUDAAN de la siguiente forma:

PROC ..... DESIGN = *WR* .....
NEST *PSTRA PPSU*
WEIGHT *W2*

96. El diseño de la encuesta de Burundi y la especificación de BRR (replicación repetida y equilibrada) para la estimación de varianzas se describen en SUDAAN de la siguiente forma:

PROC ..... DESIGN = BRR .....
WEIGHT *W2*
REPWGT *REPLWT01–REPLWT32*

Obsérvese que esta sentencia REPWGT identifica las variables de ponderación replicadas incluidas en el conjunto de datos de entrada. Estas 32 variables de ponderación replicadas se basan en los 30 pseudoestratos, con 2 UPM por pseudoestrato, y se obtuvieron con WesVar. Obsérvese también la ausencia de la sentencia NEST cuando se emplea la BRR; SUDAAN no necesita conocer las variables de estratificación y UPM, puesto que sólo usa las variables de ponderación replicadas para la estimación de varianzas.

2. DESCRIPT

97. El procedimiento DESCRIPT estima los totales y medias de población para variables continuas, y los totales y porcentajes de población para las variables categóricas. La sentencia VAR produce un listado de las variables (dependientes) para analizar. Para un programa DESCRIPT dado, las variables de la sentencia VAR deben ser todas continuas o todas categóricas. Si las variables de la sentencia VAR son categóricas, deberá usarse también la sentencia CATLEVEL para indicar para qué niveles de cada variable categórica se desean estimaciones. Por ejemplo, las dos sentencias siguientes sirven para estimar los porcentajes de seropositivos y no seropositivos de la población de inferencia de la encuesta de Burundi [suponiendo que la codificación de *IMMUNE* sea 1, 2 o . (punto) para "sin datos"].

VAR *IMMUNE IMMUNE*
CATLEVEL 1 2

98. Se pueden obtener estimaciones para los campos utilizando una sentencia TABLES que contenga una o más variables categóricas. Los campos pueden compararse entre sí mediante contrastes lineales utilizando las sentencias CONTRAST, PAIRWISE o DIFFVAR. Se pueden estimar coeficientes y medias estandarizados como, por ejemplo, la prevalencia de una enfermedad ajustada por edades, usando las sentencias STDVAR y STDWGT. Las tendencias lineales y de nivel superior (cuadráticas, etcétera) de medias o porcentajes pueden evaluarse en los distintos niveles de una variable categórica utilizando la sentencia POLY (POLYNOMIAL); SUDAAN utiliza contrastes lineales de polinomios ortogonales para estos análisis.

99. Todas las variables de una sentencia TABLES, CONTRAST, PAIRWISE, DIFFVAR, STDVAR o POLY deben aparecer también en una sentencia SUBGROUP, y una sentencia LEVELS obligatoria indica el valor codificado máximo en el análisis de cada variable categórica de la sentencia SUBGROUP.

100. La sentencia SUBPOPN de SUDAAN, que puede utilizarse en todos los procedimientos, restringe los análisis a una subpoblación, como, por ejemplo, mujeres mayores solamente. La sentencia SUBPOPN se debe usar con la totalidad del conjunto de datos de la encuesta introducido en SUDAAN, en vez de obtener del conjunto de datos de entrada el subconjunto correspondiente a la subpoblación de interés antes de utilizar SUDAAN, ya que este último procedimiento puede dar lugar a estimaciones incorrectas de los errores estándar en la medida en que falten algunas UPM seleccionadas del subconjunto de datos.

3. CROSSTAB

101. El procedimiento CROSSTAB sólo es para variables categóricas. La sentencia TABLES en CROSSTAB indica los cuadros simples, de doble entrada o de entrada múltiple para las que se estiman los porcentajes y totales de población. Para todas las variables de la sentencia TABLES se requieren las correspondientes sentencias SUBGROUP y LEVELS.

102. La sentencia TEST en CROSSTAB sirve para solicitar las pruebas de ji-cuadrado, para probar la hipótesis de nulidad de que las dos variables categóricas son estadísticamente independientes. Una prueba de ji-cuadrado se basa en un test de tipo Pearson (CHISQ), utilizando los cálculos de distribución "observada menos esperada" referidos a los totales de población estimados. La otra prueba de ji-cuadrado se basa en las disparidades en la población estimada (LLCHISQ). Se estiman los cocientes de disparidad y los riesgos relativos (en realidad, índices de prevalencia) para los cuadros 2 x 2, con intervalos de confianza, utilizando RISK = ALL en la sentencia PRINT. Por último, se puede utilizar la prueba de Cochran-Mantel-Haenszel (usando CMH en la sentencia TEST) para evaluar la independencia estadística de dos variables mientras se controla (se "estratifica") una tercera variable.

4. Ejemplos numéricos

103. En la sección III del anexo incluido en el cederrón se ilustra el uso de CROSSTAB y DESCRIPT para desarrollar los tres ejemplos que se indican en el párrafo 60, utilizando SUDAAN bajo SAS (SAS versión 8.2 y SUDAAN versión 8.0). Para la estimación de varianzas se utiliza tanto la linealización mediante el desarrollo en serie de Taylor como la BRR (replicación repetida y equilibrada). El examen de los programas SUDAAN comentados (escritos por los usuarios) y de los resultados comentados de SUDAAN preparará al lector para escribir sus propios programas SUDAAN e interpretar los resultados. En el anexo (sección III) sólo se incluyen los análisis de SUDAAN estudiados en TABLES 1, 3, 4, 5 y 6.

104. En la fila 4 del cuadro XXI.1 se resume el resultado de CROSSTAB y DESCRIPT de la sección III del anexo para la estimación del porcentaje y el número de mujeres seropositivas de la población de inferencia de la encuesta de Burundi, con el error estándar estimado. Los resultados de CROSSTAB y DESCRIPT en SUDAAN son idénticos para un método dado de estimación de varianzas (como cabe esperar), y los resultados para la serie de Taylor son idénticos a los de la BRR (lo que no siempre es cierto). Los resultados de SUDAAN coinciden con los del procedimiento SURVEYMEANS de SAS. Obsérvese que CROSSTAB y DESCRIPT no calculan intervalos de confianza para porcentajes o totales de población estimados.

105. En la fila 2 del cuadro XXI.3 de la sección VII del anexo (cederrón) se observa que con CROSSTAB y DESCRIPT (ya sea con la serie de Taylor o con la BRR) se obtienen los mismos resultados de estimación del porcentaje de mujeres seropositivas, pero para cada uno de los dos ámbitos (rural y urbano). Los resultados de los procedimientos CROSSTAB y DESCRIPT de SUDAAN coinciden con los de SURVEYMEANS de SAS.

106. En la fila 2 del cuadro XXI.4 de la sección VII del anexo (cederrón) se resume el resultado de DESCRIPT (con la serie de Taylor y con la BRR), que emplea el contraste lineal para comparar los porcentajes de mujeres seropositivas rurales con los de urbanas. Hay una diferencia despreciable entre el error estándar estimado con la serie de Taylor y el estimado con la BRR. La conclusión es que las mujeres urbanas y rurales de la población de inferencia de la encuesta de Burundi difieren en la prevalencia de la seropositividad; la prevalencia es mayor en las mujeres urbanas. Obsérvese que los resultados de los contrastes lineales con DESCRIPT coinciden con los obtenidos usando SURVEYREG de SAS para comparar dos campos.

107. En las filas 1 y 2 del cuadro XXI.5 de la sección VII del anexo (cederrón) se muestran los resultados de las dos pruebas de ji-cuadrado distintas disponibles en CROSSTAB: el test Pearson (CHISQ) y el log-lineal (LLCHISQ). Los resultados con la serie de Taylor son idénticos a los de la BRR. Tanto la prevalencia estimada de la seropositividad (utilizando CHISQ) como la disparidad estimada de la seropositividad (utilizando LLCHISQ) son claramente mayores en las mujeres urbanas que en las rurales.

108. En la fila 1 del cuadro XXI.6 de la sección VII del anexo (cederrón) se muestra el cociente de disparidad (0,393) y el índice de prevalencia (0,797) estimados de la seropositividad (rurales con respecto a urbanas), ambos con un intervalo de confianza del 95%. La serie de Taylor y la BRR muestran diferencias despreciables del cociente de disparidad en el límite superior del intervalo de confianza del 95%. El cociente de disparidad y el índice de prevalencia estimados difieren en magnitud porque la prevalencia de la seropositividad no es baja.

5. Ventajas e inconvenientes y costo

109. SUDAAN es un completo paquete de *software* para encuestas (y datos correlacionados) con gran capacidad analítica para análisis tanto descriptivos como de modelado. Tiene amplia capacidad para la estimación y prueba de matrices de contraste de parámetros de población especificadas por el usuario, coeficientes de regresión inclusive. Se puede ejecutar tanto en macrocomputadoras como en PC. Los usuarios de SAS tendrán posiblemente más facilidad para aprender a manejar SUDAAN, ya que su sintaxis es similar a la de SAS. No obstante, algunos aspectos de la sintaxis de SUDAAN resultan esotéricos, por lo que puede requerir más tiempo de aprendizaje que otros paquetes.

110. En comparación con el costo de otros paquetes de *software* examinados en este capítulo, el de SUDAAN es elevado, especialmente si se usa como SUDAAN bajo SAS, porque en ese caso también hay que tener SAS. Hay un servicio de asistencia técnica para usuarios con licencia. El manual del usuario de SUDAAN para la versión 8.0, más como referencia que para el aprendizaje de SUDAAN, contiene varios ejemplos detallados y comentados de análisis elaborados con datos de la encuesta NHANES-III (National Health and Nutrition Examination Survey III), que pueden resultar de utilidad para aprender a utilizar SUDAAN.

111. SUDAAN es el único paquete de *software* ilustrado en este capítulo que incluye los dos enfoques principales de estimación de varianzas: la linealización mediante el desarrollo en serie de Taylor y los métodos de replicación. Sin embargo, SUDAAN no construye ponderaciones replicadas para la replicación repetida y equilibrada (BRR), por lo que es el usuario quien tiene que aportar estas ponderaciones. SUDAAN construye ponderaciones replicadas para el procedimiento *jackknife* de eliminación de una UPM, y también acepta ponderaciones replicadas *jackknife* si se incluyen en el conjunto de datos de entrada.

112. SUDAAN es también el único de los paquetes de *software* examinados que posee amplia capacidad para describir varias etapas de muestreo, estratificación y términos *fpc* para su incorporación a la estimación de varianzas. Cuenta además con varias definiciones

para los cálculos del efecto del diseño, de manera que es posible excluir del efecto del diseño los efectos del sobremuestreo o de una ponderación desigual.

113. La introducción de datos ASCII en SUDAAN resulta engorrosa y son preferibles las otras dos opciones de entrada de datos: como conjuntos de datos SAS o SPSS. Si la entrada de datos al SUDAAN autónomo se hace mediante un conjunto de datos SAS, éste deberá haberse preparado con la versión 6.04 de SAS o como archivo de transporte de SAS. El SUDAAN bajo SAS puede leer cualquier conjunto de datos que pueda leer SAS. Los resultados generados por SUDAAN pueden grabarse electrónicamente en un formato de archivo de datos SAS para su uso posterior con SAS o con un *software* de hoja de cálculo como Excel. SUDAAN posee una capacidad muy limitada para recodificar variables, y no tiene capacidad de gestión de datos, por lo que es aconsejable que todas las tareas necesarias de recodificación y de formación de nuevas variables se realicen en SAS o en SPSS (según el tipo de conjunto de datos de entrada) antes de utilizar SUDAAN.

I. Procedimientos para análisis de encuestas en STATA 7.0

1. Descripción general de STATA

114. STATA es un paquete de *software* de estadística general al que se dotó en 1995 de amplia capacidad de análisis de datos de encuestas. Aquí examinaremos la versión 7.0 de STATA; la versión 8.0 se publicó en 2003. Para la estimación de varianzas se usa solamente la linealización mediante el desarrollo en serie de Taylor. La descripción del plan de muestreo común es, por defecto, *WR*. Con STATA pueden incorporarse términos *fpc* en la estimación de varianzas con planes de muestreo monoetápico sin reemplazamiento (muestreo aleatorio simple y muestreo aleatorio estratificado) y para muestreo monoetápico por conglomerados sin reemplazamiento (estratificado o no), donde se utiliza el muestreo con probabilidad igual para los conglomerados (UPM) de cada estrato y todos los elementos de una UPM seleccionada se incluyen en la muestra.

115. La amplitud de los análisis de encuestas de STATA es comparativamente superior a la de SUDAAN, con capacidad de estadística matemática para matrices de contraste de parámetros de población especificadas por el usuario, coeficientes de regresión inclusive. STATA se ejecuta interactivamente con mandatos abreviados y simples, por lo que resulta relativamente fácil de aprender. No obstante, el usuario puede escribir sus propios programas y ejecutarlos por lotes si lo desea. STATA distingue mayúsculas de minúsculas, y sus mandatos se escriben con minúsculas. STATA asigna por defecto una cantidad de memoria en la que carga una copia del conjunto de datos de entrada. Si la memoria es insuficiente para grandes conjuntos de datos, puede aumentarse con el mandato *set memory*.

116. En STATA, los mandatos para encuestas empiezan por *svy* (de *survey*, encuesta en inglés). Existen mandatos descriptivos para estimar la media de población (*svymean*), el total de población (*svytotal*), una proporción de la población (*svyprop*) y porcentajes y totales en cuadros de doble entrada (*svytab*). Para los intervalos de confianza en las proporciones de la población a partir de *svytab* se utiliza una transformación logit de manera que los límites inferior y superior estimados queden constreñidos entre (0, 1). En *svytab* se dispone de ocho pruebas de ji-cuadrado diferentes para los datos de encuestas en cuadros de doble entrada. Los procedimientos de modelado disponibles incluyen la regresión lineal, la regresión logística (incluida la multinomial con una variable nominal u ordenada), la regresión de Poisson y modelos probit.

117. El mandato *svyset* se utiliza para indicar a STATA el plan de muestreo. Para describir el plan de muestreo común *WR* (opción por defecto) se introducen interactivamente en STATA tres claves para el mandato *svyset*. La palabra clave *strata* precede al nombre de la variable de estratificación, la palabra clave *psu* (UPM) precede al nombre de la variable

UPM y la palabra clave *pweight* precede al nombre de la variable de ponderación de la muestra. De este modo el plan de muestreo de la encuesta de Burundi se describe en la versión 7 de STATA de la siguiente forma:

```
svyset    strata      pstra
svyset    psu         ppsu
svyset    pweight     w2
```

118. Como ya se ha señalado para los procedimientos de SAS para encuestas, la omisión de la palabra clave *strata* en STATA significa que no hay estratificación de UPM previa al muestreo en la primera etapa. La omisión de la palabra clave *psu* significa muestreo monoetápico de los elementos y no clasificación de los elementos seleccionados. La omisión de la palabra clave *pweight* significa elementos muestrales de ponderación igual, con un valor por defecto de 1,0 para la variable de ponderación. En la versión 8 de STATA se ha revisado la sintaxis del mandato *svyset*.

119. El mandato *svydes* indica a STATA que liste las variables de diseño de la encuesta que ha añadido al conjunto de datos (con los mandatos *svyset*) y resuma el número de estratos, el número de UPM por estrato y el número promedio de observaciones por UPM en cada estrato. Este resumen de las características del diseño de la encuesta resulta de gran utilidad.

2. *Svymean, Svyprop, Svytotal, Svylc*

120. El mandato *svymean* permite estimar una media de población, ya sea para una variable continua o para una variable indicadora codificada 1 ó 0 (esto es, una proporción de la población estimada). Las opciones de salida incluyen el error estándar estimado, el coeficiente de variación estimado, el efecto del diseño y el intervalo de confianza del parámetro de población.

121. El mandato *svyprop* es para datos categóricos: permite estimar la proporción de la población que hay en cada nivel de la variable categórica, junto con el error estándar estimado. En comparación con *svymean*, *svyprop* presenta menos opciones de salida.

122. El mandato *svytotal* permite estimar un total de población tanto para una variable continua como para una variable indicadora (0, 1), con el error estándar estimado, el coeficiente de variación estimado, el efecto del diseño y el intervalo de confianza.

123. Cada uno de los tres mandatos anteriores se puede utilizar para estimar parámetros de población de los campos utilizando la opción *by* en la línea de mandatos; por ejemplo, *by (stra)* o *by (urb_rur)* para analizar los dos campos —rural y urbano— de las mujeres de Burundi. STATA usa fórmulas correctas de estimación de varianzas para los campos con la sentencia *by* en sus mandatos *svy*.

124. Además, cada uno de los tres mandatos anteriores se puede usar con la opción *subpop* en la línea de mandatos para efectuar una estimación de parámetros de población para una subpoblación; por ejemplo, mujeres mayores. La sentencia *if* de STATA no debe usarse para análisis de subpoblaciones, porque las varianzas estimadas pueden ser incorrectas; se utilizará la opción *subpop*.

125. El mandato *svylc* permite estimar combinaciones lineales especificadas por el usuario de medias, proporciones o totales de campos, junto con el error estándar estimado, la prueba de la *t*, el valor *p* y el intervalo de confianza. Este mandato se puede usar para comparar campos entre sí. En la versión 8.0, el mandato *svylc* se ha sustituido por *lincom*. El mandato *svylc* sigue funcionando en la versión 8.0, pero ya no está documentado en la misma.

3. *Svytab*

126. El mandato *svytab* de STATA es para cuadros de doble entrada. Permite estimar los porcentajes de población (fila, columna o total) con los errores estándar estimados, los totales de población de las celdas del cuadro con los errores estándar estimados, y los intervalos de confianza. Se emplea una transformación logit para obtener los intervalos de confianza de las proporciones de la población de manera que los límites estimados inferior y superior queden constreñidos en el intervalo (0, 1). Se dispone de ocho pruebas de ji-cuadrado diferentes para probar la hipótesis de nulidad de la independencia estadística de las dos variables categóricas del cuadro. El mandato *subpop* puede utilizarse con *svytab*.

4. Ejemplos numéricos

127. En la sección IV del anexo incluido en el cederrón se ilustra el uso de los mandatos de STATA para desarrollar los tres ejemplos que se indican en el párrafo 60. Cada uno de los ejemplos desarrollados es un archivo de registro (log) de la sesión interactiva con STATA. El examen del registro comentado de STATA (mandatos del usuario y resultados producidos por STATA) ayudará al lector a utilizar los mandatos para encuestas de STATA e interpretar los resultados.

128. Los mandatos *svymean* y *svytotal* se utilizaron con la variable indicadora *BLOOD* (1 = seropositiva, 0 = seronegativa). En el cuadro XXI.1 (filas 5 y 6) se muestra el número y porcentaje estimados de mujeres seropositivas, con los intervalos de confianza. Los cálculos de STATA coinciden con los de SURVEYMEANS de SAS y con los de DESCRIPT y CROSSTAB de SUDAAN.

129. En el cuadro XXI.3 (fila 3) de la sección VII del anexo (cederrón) se muestra el porcentaje estimado de mujeres seropositivas, por tipo de residencia (rural o urbana). Las estimaciones por puntos del mandato *svytab* de STATA y los errores estándar estimados coinciden con los de SURVEYMEANS de SAS y con los de DESCRIPT y CROSSTAB de SUDAAN. No obstante, existen ligeras diferencias entre los intervalos de confianza de los campos obtenidos con el mandato *svytab* de STATA y los obtenidos con SURVEYMEANS de SAS, porque el mandato *svytab* de STATA emplea una transformación logit para obtener los intervalos de confianza.

130. El cuadro XXI.4 (fila 3) de la sección VII del anexo (cederrón) presenta los resultados del mandato *svylc* de STATA para el contraste lineal que compara los porcentajes de mujeres rurales y urbanas seropositivas, los cuales indican una diferencia sustancial entre los dos campos. Los resultados de STATA coinciden con los de DESCRIPT de SUDAAN y con los obtenidos utilizando SURVEYREG de SAS para comparar campos.

131. En el cuadro XXI.5 (filas 3 a 5) de la sección VII del anexo (cederrón) se presentan los resultados del mandato *svytab* de STATA para tres pruebas de ji-cuadrado de la hipótesis de nulidad de que la seropositividad es estadísticamente independiente de que la residencia sea rural o urbana. Las tres pruebas de ji-cuadrado de *svytab* presentan valores *p* similares (y bajos). La prueba por defecto de ji-cuadrado para el mandato *svytab* de STATA (fila 3) es una prueba de ji-cuadrado de tipo Pearson propuesta por Rao y Scott (1981; 1984) con una corrección de segundo orden). Las otras dos pruebas de ji-cuadrado de *svytab* (filas 4 y 5) son las mismas pruebas de ji-cuadrado que en CROSSTAB de SUDAAN, y los resultados de los cálculos de STATA y SUDAAN son iguales para estas dos pruebas.

132. Como el mandato *svytab* de STATA no produce cocientes de disparidad ni índices de prevalencia, se empleó el mandato *svylogit* para estimar el cociente de disparidad (entre residencia urbana y rural) de la seropositividad. El cociente de disparidad de STATA, con el intervalo de confianza se indica en el cuadro XXI.6 (fila 2) de la sección VII del anexo

(cederrón). El mandato *svylogit* de STATA produce los mismos cálculos que CROSSTAB de SUDAAN para la estimación por puntos y el intervalo de confianza.

5. Ventajas e inconvenientes y costo

133. STATA es un paquete completo de análisis estadístico general y posee también amplia capacidad analítica de datos de encuestas, con procedimientos descriptivos y de modelado basado en el diseño. Ofrece muchos procedimientos de modelado para datos de encuestas. STATA ha obtenido muy buenas críticas como paquete estadístico, es relativamente fácil de aprender y cuenta con un activo grupo de usuarios. En comparación con otros paquetes de *software* examinados, su costo es moderado.

134. Para quienes deseen probar sus propias hipótesis específicas o combinaciones estimadas de parámetros de población, STATA acepta matrices de contraste definidas por el usuario de parámetros de población estimados, incluidos los coeficientes de regresión. En general, para los usuarios con los necesarios conocimientos de estadística matemática, el paquete ofrece gran flexibilidad para la realización de análisis estadísticos.

135. STATA sólo usa la linealización mediante el desarrollo en serie de Taylor y está limitado a la descripción del plan de muestreo común *WR*. Sin embargo, puede incluir en la estimación de varianzas para muestreo sin reemplazamiento los términos *fpc* para el muestreo monoetápico de elementos y para el muestreo monoetápico por conglomerados. Es un tanto difícil, aunque posible, extraer resultados analíticos de STATA (por ejemplo, tamaños de la muestra no ponderados, estimaciones por puntos, errores estándar) para exportarlos a otros formatos de datos.

J. Procedimientos para encuestas en Epi-Info 6.04d y Epi-Info 2002

1. Descripción general de Epi-Info

136. Epi-Info ha sido desarrollado durante muchos años por los Centros para el Control y la Prevención de Enfermedades (CDC) y la Organización Mundial de la Salud (OMS). Este *software* está disponible gratuitamente para su descarga en el sitio web de CDC: http://www.cdc.gov/epiinfo/.

137. Existen dos versiones de Epi-Info: la última versión de Epi-Info para DOS (la versión 6.04d) y la más reciente, Epi-Info 2002, para Windows.

138. Las capacidades de Epi-Info incluyen el desarrollo de un cuestionario o modelo de recogida de datos de investigación, entrada de datos personalizada, análisis de datos y tratamiento de textos. Sus capacidades analíticas y estadísticas están orientadas al servicio de los epidemiólogos en todo el mundo. El resultado (analítico) de Epi-Info puede enviarse a la pantalla, a una impresora o a un archivo electrónico.

139. Ambas versiones de Epi-Info (DOS o Windows) tienen capacidad para análisis descriptivos básicos de datos de encuestas complejas. Sólo admite la descripción del plan de muestreo común *WR*. El conjunto de datos de entrada debe ordenarse por dos de las tres variables de diseño de la encuesta: la variable de estratificación *STRATVAR* y la variable UPM *PSUVAR* en cada estrato. Epi-Info no incorpora términos *fpc* en la estimación de varianzas. Tampoco realiza estimaciones de totales de población. Para la estimación de varianzas se usa la linealización mediante el desarrollo en serie de Taylor.

140. La capacidad analítica de Epi-Info para datos de encuestas complejas fue desarrollada originalmente para el Sistema de Vigilancia de los Factores Conductuales de Riesgo (BRFSS), programa de encuestas anuales sobre la salud patrocinado por CDC para los estados de EE. UU. (Brogan, 1988 y en curso de publicación) y para la metodología de muestreo por conglomerados de la OMS utilizada en todo el mundo por el Programa Ampliado de Inmunización (PAI) para estimar la cobertura de las vacunaciones entre la población infantil

(Brogan y otros, 1994). No obstante, los procedimientos para encuestas de Epi-Info pueden utilizarse para cualquier encuesta compleja que pueda ser descrita mediante la descripción del plan de muestreo común *WR*.

2. Módulo CSAMPLE de Epi-Info Versión 6.04d (DOS)

141. Epi-Info para DOS fue desarrollado conjuntamente por CDC y la OMS. La entrada de datos de Epi-Info 6.04d se hace mediante un archivo dBase o ASCII, que Epi-Info convierte a continuación en archivo de datos Epi-Info (*.rec). Existen paquetes de *software* para convertir archivos de datos SAS, SPSS o de otros tipos en archivos *.rec de Epi-Info; por ejemplo, DBMS-COPY (http://www.dataflux.com/conceptual/). Epi-Info 6.04d se ejecuta como programa interactivo y no puede ejecutarse por lotes. La versión DOS puede ser preferible a la versión Windows para quienes tengan computadoras o sistemas operativos más antiguos o con poca capacidad de almacenamiento en disco duro.

142. Con el módulo CSAMPLE de Epi-Info versión 6.04d se pueden realizar análisis de datos de encuestas complejas. CSAMPLE permite estimar medias de población (para variables continuas o variables indicadoras codificadas con 1 ó 0) o porcentajes de población (para variables categóricas), junto con el error estándar estimado, intervalo(s) de confianza y efecto del diseño. Estas estimaciones también pueden calcularse para campos formados por niveles de variables categóricas. Además, CSAMPLE permite estimar la diferencia entre medias de campos o porcentajes de campos, con su correspondiente error estándar estimado de la diferencia estimada y un intervalo de confianza de la diferencia de población. CSAMPLE permite estimar el cociente de disparidad y el riesgo relativo en los cuadros 2 x 2. Obsérvese que CSAMPLE no permite estimar totales de población.

143. Cuando se abre el módulo CSAMPLE en Epi-Info 6.04d, aparece una pantalla de entrada de datos en la que el usuario especifica las variables que van a utilizarse para el análisis. El usuario selecciona una variable para cada uno de los tres recuadros de diseño de la encuesta: STRATA (la variable de estratificación), UPM (la variable UPM o variable de agrupamiento) y WEIGHT (la variable de ponderación de la muestra). Para la encuesta de Burundi, la especificación de Epi-Info fue la siguiente:

STRATA *PSTRA*
UPM *PPSU*
WEIGHT *W2*

144. El usuario especifica la variable de análisis (o variable dependiente) en el recuadro MAIN. Esta variable puede ser continua, como *IUML*, o categórica, como *IMMUNE*. Si se desea obtener la media de población estimada para la variable continua (o que se supone continua) especificada en MAIN, el usuario hará clic en la opción MEANS. Si se desea obtener los porcentajes de población estimados para la variable categórica especificada en MAIN, el usuario hará clic en la opción TABLE.

145. Si se desea obtener las medias o porcentajes estimados para los campos, en el recuadro CROSSTAB se especifica la variable que define los campos y la variable de análisis se especifica en el recuadro MAIN.

146. Además, CSAMPLE permite estimar la diferencia de medias de una variable de análisis entre dos campos. El usuario puede especificar los dos niveles de la variable CROSSTAB que definen los dos campos que se van a comparar entre sí.

3. Epi-Info 2002 (Windows)

147. Epi-Info 2002 es una aplicación para Windows desarrollada por CDC. La entrada de datos para el análisis de datos con Epi-Info 2002 se hace mediante un archivo de Mi-

crosoft Access 97 (*.mdb) o un archivo dBase. Epi-Info 2002 también puede leer los archivos *.rec preparados para las versiones DOS de Epi-Info. El *software* se ejecuta interactivamente, pero también puede ejecutarse por lotes.

148. Epi-Info 2002 tiene tres procedimientos para muestras complejas, localizados en Advanced Statistics (estadística avanzada) de la opción Analyze Data (análisis de datos) del menú principal: el mandato *Complex Sample Frequencies* (frecuencias de muestras complejas) permite estimar la distribución porcentual simple de una variable categórica, con el error estándar estimado y los intervalos de confianza; con el mandato *Complex Sample Tables* (cuadros de muestras complejas) se estiman los porcentajes por filas y por columnas de un cuadro de doble entrada de variables categóricas ("exposure" (fila) y "outcome" (columna)), con los errores estándar estimados y los intervalos de confianza para los porcentajes de la fila (si es un cuadro 2 x 2, el procedimiento también permite estimar el cociente de disparidad y el riesgo relativo, con los intervalos de confianza); con el mandato *Complex Sample Means* (medias de muestras complejas) se estiman la media de una variable continua, con el error estándar estimado y el intervalo de confianza, y la media de campos formados por una variable categórica. Si la variable de campo tiene dos niveles, también se estima la diferencia entre las medias de campo, con el error estándar estimado y el intervalo de confianza.

149. En los tres procedimientos para muestras complejas, las variables de diseño de la encuesta se identifican en tres recuadros designados como *Weight*, *UPM* y *Stratify By* (la variable de estratificación de la encuesta). Para obtener como resultado los errores estándar estimados y los intervalos de confianza, se hace doble clic en el mandato *Set* (en *Options*) y se escoge *Statistics* = *Advanced*.

4. Ejemplos numéricos

150. En la sección V del anexo (cederrón) se ilustra el uso de CSAMPLE en Epi-Info 6.04d para desarrollar los tres ejemplos del párrafo 60. En cada uno de los ejemplos desarrollados se incluyen los resultados de Epi-Info, anotados con comentarios. El examen de los resultados comentados ayudará al lector a interpretar los resultados de CSAMPLE.

151. En el cuadro XXI.1 (fila 7) se indica la estimación del porcentaje de mujeres seropositivas obtenida con Epi-Info 6.04d. La estimación por puntos y el error estándar estimado de Epi-Info coinciden con los de SURVEYMEANS de SAS, *svymean* de STATA y DESCRIPT y CROSSTAB de SUDAAN. El intervalo de confianza del 95% para la prevalencia de la seropositividad es más estrecho que los intervalos de confianza obtenidos con SURVEYMEANS de SAS y *svymean* de STATA. Esto se debe a que Epi-Info usa $z = 1{,}96$ en su cálculo del intervalo de confianza del 95%, en vez del valor t de Student de 2,042 con 30 gld, los grados de libertad del denominador para la encuesta de Burundi (número de UPM (60) menos número de pseudoestratos (30)).

152. En el cuadro XXI.3 (fila 4) de la sección VII del anexo (cederrón) se ofrecen las estimaciones de Epi-Info para la prevalencia de la seropositividad por tipo de residencia, rural o urbana. Las estimaciones por puntos y los errores estándar estimados de Epi-Info coinciden con los de SURVEYMEANS de SAS, el mandato *svytab* de STATA, y DESCRIPT y CROSSTAB de SUDAAN. Los intervalos de confianza de los campos para Epi-Info son más estrechos que los de SURVEYMEANS de SAS y el mandato *svytab* de STATA, porque Epi-Info usa $z = 1{,}96$.

153. En el cuadro XXI.4 (fila 4) de la sección VII del anexo (cederrón) se indica el resultado en Epi-Info del contraste lineal que compara la prevalencia de la seropositividad de las mujeres rurales con la de las urbanas. El valor del contraste estimado (–16,99%) coincide con el de SURVEYREG de SAS, DESCRIPT de SUDAAN y *svylc* de STATA. Epi-Info no ofrece el error estándar estimado de la diferencia estimada, y el intervalo de confianza del 95% que ofrece Epi-Info para el valor del contraste es erróneo.

154. En el cuadro XXI.6 (fila 3) de la sección VII del anexo (cederrón) se indican las estimaciones de Epi-Info para el cociente de disparidad (residencia urbana con respecto a rural) y el índice de prevalencia de la seropositividad, con un intervalo de confianza del 95%. Las estimaciones por puntos de Epi-Info coinciden exactamente con las de CROSSTAB de SUDAAN y *svylogit* de STATA, y los intervalos de confianza de Epi-Info coinciden prácticamente con los de SUDAAN y STATA.

5. Ventajas e inconvenientes y costo

155. Una importante ventaja de Epi-Info es su costo: se puede descargar gratis del sitio web de CDC. Además, hay versiones para DOS y para Windows, lo que ofrece más flexibilidad en cuanto a los requisitos de *hardware* y *software* para ejecutar Epi-Info. La capacidad de análisis de encuestas de Epi-Info atraerá seguramente a quienes ya sean usuarios de Epi-Info para otros tipos de análisis epidemiológico o estadístico.

156. Epi-Info sólo usa la linealización mediante el desarrollo en serie de Taylor y sólo dispone de la descripción del plan de muestreo común *WR*. El módulo CSAMPLE de la versión para DOS y el correspondiente de la versión para Windows (tres procedimientos dentro del apartado Advanced Statistics) son adecuados para la estadística descriptiva básica con datos de encuestas complejas. Esto incluye la estimación de medias o porcentajes de población para toda la población y por campos, así como la comparación de campos. Epi-Info no posee capacidad de análisis de encuestas para estimar totales de población, realizar pruebas de ji-cuadrado, incorporar los términos *fpc* (corrección al muestreo de poblaciones finitas) en la estimación de varianzas ni para los análisis de modelado basado en el diseño (por ejemplo, regresión logística o regresión lineal).

K. WesVar 4.2

1. Descripción general de WesVar

157. WesVar es un paquete de *software* para el análisis de datos de encuestas que para la estimación de varianzas utiliza métodos de replicación (Rust y Rao, 1996): BRR, incluido el factor de Fay opcional, y tres variantes de *jackknife*. WesVar no tiene capacidad para la linealización mediante el desarrollo en serie de Taylor. Los diseños de encuestas que se prestan bien a la BRR tienen varios estratos y exactamente dos UPM seleccionadas por estrato. Los métodos *jackknife*, como la linealización mediante el desarrollo en serie de Taylor, se pueden aplicar a un diseño con cualquier número (≥ 2) de UPM seleccionadas por estrato.

158. El plan de muestreo por defecto para WesVar es el plan de muestreo común *WR* mencionado anteriormente. WesVar tiene capacidad para incluir factores *fpc* en la estimación de varianzas, pero sólo para técnicas *jackknife* y sólo para el muestreo monoetápico de elementos.

159. WesVar 4.2 puede leer los siguientes tipos de conjunto de datos de entrada: PC-SAS para DOS, archivos de transporte de SAS, SAS (versiones 6-8), SPSS, STATA, ASCII y archivos que cumplan el estándar ODBC como los de Microsoft Excel o Access. Como corresponde al plan de muestreo común *WR,* que emplea por defecto, si es necesario, construir ponderaciones replicadas, WesVar requiere las variables de estratificación, UPM y ponderación para cada observación. No obstante, una vez que las ponderaciones replicadas se encuentran en el archivo, ya no se necesitan los identificadores de UPM y de los estratos: esto representa una ventaja en cuanto a la confidencialidad de los métodos de replicación en los archivos de uso público. WesVar es el único de los paquetes examinados que puede ajustar las ponderaciones básicas de las encuestas para las faltas de respuesta, posestratificación y *raking*. Una vez completada la preparación del conjunto de datos de entrada, éste se guarda como archivo WesVar (*.var) para análisis de datos y para futura gestión de los mismos.

160. El paquete dispone de toda una gama de funciones de estadística descriptiva: estimaciones de medias de población, porcentajes, percentiles y totales, junto con el error estándar estimado, coeficiente de variación, intervalo de confianza y efecto del diseño. Un particular punto fuerte de WesVar, y de los métodos de replicación en general, es la capacidad de obtener estimaciones por puntos (con el error estándar estimado) de funciones especificadas por el usuario de parámetros de población como, por ejemplo, índices de prevalencia. WesVar permite realizar análisis de regresión basados en el diseño: lineal, logística y logística multinomial.

161. Desde la página web de WESTAT se puede descargar la versión 4 de WesVar para su evaluación durante 30 días. La versión 2 de WesVar está también disponible para descarga desde la página web, y se puede usar sin límite de tiempo gratuitamente (véase http://www.westat.com/WesVar/about). En comparación con la versión 2, la versión 4 de WesVar acepta una mayor variedad de conjuntos de datos de entrada, tiene más capacidad de manejo de archivos y gestión de datos, permite ajustar las ponderaciones replicadas para tener en cuenta las faltas de respuesta e incluye otras muchas opciones analíticas. El usuario puede empezar con la versión 2 de WesVar y pasarse después a la versión 4 si lo necesita.

2. Uso de WesVar versión 4.2

162. La interacción entre el usuario y WesVar tiene lugar a través de menús desplegables en un entorno Windows. Cuando se inicia el *software* WesVar, el primer menú presenta cuatro opciones. La primera opción, *New WesVar data file* (nuevo archivo de datos de WesVar): 1) lee un conjunto de datos de entrada que no está en el formato de conjunto de datos de WesVar; 2) crea ponderaciones replicadas o acepta las ponderaciones replicadas que se encuentran ya en el conjunto de datos de entrada; 3) recodifica, transforma, etiqueta y formatea las variables; 4) realiza ajustes de posestratificación, *raking* y faltas de respuesta; 5) define subpoblaciones para el análisis; y 6) modifica los gld si se le pide, y después guarda el conjunto de datos como archivo WesVar. La segunda opción, *Open WesVar data file* (abrir archivo de datos WesVar) lee un archivo de datos WesVar y permite realizar las seis operaciones enumeradas anteriormente.

163. La tercera opción, *New WesVar Notebook* (nuevo cuaderno WesVar), acepta las peticiones de análisis de un archivo de datos WesVar, ejecuta las peticiones, muestra el resultado y guarda las peticiones y el resultado en un "cuaderno", que es el sistema del que se vale WesVar para organizar los análisis pedidos y el resultado de los mismos. Se puede pedir uno de los siguientes tipos de análisis: cuadros o regresión (lineal, logística o multinomial). Una vez que se elige entre cuadros o regresión, se dispone de muchas opciones para concretar el análisis. La navegación por las pantallas de menús de análisis y la lectura de los resultados no resultan sencillas, pero en la guía del usuario de WesVar hay varios ejemplos útiles que ilustran la navegación por los menús y la organización de los resultados.

164. Si las peticiones y resultados obtenidos de una sesión anterior con WesVar se guardaron en el cuaderno, se podría escoger la cuarta opción del primer menú: *Open WesVar Notebook*. Se pueden añadir nuevas peticiones de análisis al cuaderno existente y guardarlo a continuación. Todos los análisis relacionados con un determinado archivo de datos o con un determinado proyecto de WesVar pueden organizarse en uno o más cuadernos.

165. Es necesario especificar uno de los cinco métodos de replicación de WesVar para construir ponderaciones replicadas o para reconocer ponderaciones replicadas ya existentes en el archivo de datos de entrada. Los métodos de replicación son:

a) Replicación repetida y equilibrada (BRR): dos UPM seleccionadas por estrato;

b) Método de perturbaciones de Fay (FAY) con BRR;

c) Método *jackknife* de eliminación de una UPM sin estratificación explícita (JK1);

d) Método *jackknife* con dos UPM seleccionadas por estrato (JK2);

e) Método *jackknife* con dos o más UPM seleccionadas por estrato (JKn).

166. En los apéndices A y D de la guía del usuario de WesVar hay una excelente descripción general de estos cinco métodos de replicación y se ilustra con ejemplos sobre cómo traducir distintos planes de muestreo a alguno de estos cinco métodos.

3. Ejemplos numéricos

167. Como el conjunto de datos de entrada de la encuesta de Burundi no contenía ponderaciones replicadas, fue necesario escoger una de las cinco técnicas de replicación disponibles y pedir seguidamente a WesVar que calculara las ponderaciones replicadas. Las variables de diseño de la encuesta de Burundi que necesitó WesVar fueron las siguientes: *PSTRA*, *PPSU* y *W2.* Como el plan de muestreo de la encuesta de Burundi es *WR,* con 30 pseudoestratos y dos UPM seleccionadas por estrato, las mejores opciones son BRR y JK2. Escogimos BRR sin factor de perturbación de Fay. No se realizaron ajustes para tener en cuenta las faltas de respuesta o posestratificación o *raking*, ya que tampoco se hicieron estos ajustes en el conjunto de datos completo cuando se utilizó la linealización mediante el desarrollo en serie de Taylor.

168. En la sección VI (cederrón) se ilustra el uso de WesVar para desarrollar los tres ejemplos que se indican en el párrafo 60. En cada uno de los ejemplos desarrollados se presentan los resultados de WesVar 4.1 ó 4.2, aunque no se muestran las pantallas del menú de entrada de datos correspondientes a los análisis pedidos. El examen de los resultados comentados de WesVar ayudará al lector a interpretar los resultados de WesVar.

169. En la fila 8 del cuadro XXI.1 se puede observar que WesVar coincide con los demás paquetes de *software* de encuestas en el porcentaje y número de mujeres seropositivas estimado (con errores estándar). Los intervalos de confianza de WesVar coinciden con los de SAS y STATA, pero no con los de Epi-Info, que son demasiado estrechos.

170. En la fila 5 del cuadro XXI.3 del anexo (cederrón) se puede observar que WesVar coincide con los demás paquetes de *software* en las estimaciones por puntos de los campos y en los errores estándar estimados. Los intervalos de confianza de WesVar son muy similares a los de SURVEYMEANS de SAS, aunque difieren ligeramente de los del mandato *svytab* de STATA (que emplea transformación logit) y de Epi-Info (que usa $z = 1{,}96$ en vez del valor *t* de Student).

171. En la fila 5 del cuadro XXI.4 de la sección VII del anexo (cederrón) se muestran los resultados del contraste lineal de WesVar comparando la prevalencia de la seropositividad entre las mujeres rurales y las urbanas. WesVar coincide con SURVEYREG de SAS, DESCRIPT de SUDAAN y *svylc* de STATA en el error estándar estimado para el contraste lineal y en el valor estadístico de la *t* de Student. Los intervalos de confianza del contraste lineal arrojan diferencias despreciables entre SAS, STATA y WesVar.

172. En las filas 6 y 7 del cuadro XXI.5 de la sección VII del anexo (cederrón) se muestran las dos pruebas de ji-cuadrado de Rao/Scott para datos de encuestas complejas ejecutadas con WesVar. Estos cálculos no coinciden exactamente con ninguna de las otras pruebas de ji-cuadrado de los demás paquetes.

173. En la fila 4 del cuadro XXI.6 de la sección VII del anexo (cederrón) se puede observar que el procedimiento de regresión logística de WesVar produce el mismo cociente de disparidad estimado y sustancialmente el mismo intervalo de confianza que CROSSTAB de SUDAAN y *svylogit* de STATA. En la fila 5 del cuadro XXI.6 se puede observar que el índice de prevalencia estimado por WesVar (utilizando las funciones de celdas en TABLES) coincide con el de CROSSTAB de SUDAAN y el de Epi-Info, con diferencias despreciables en los intervalos de confianza entre SUDAAN y WesVar.

4. Ventajas e inconvenientes y costo

174. WesVar sólo usa técnicas de replicación para la estimación de varianzas. El analista de datos secundarios de conjuntos de datos para uso público que incluyen las ponderaciones replicadas no necesita conocer los detalles del diseños muestral (por ejemplo, las variables de diseño de la encuesta *STRATVAR* y *PSUVAR*), aunque no necesita especificar en WesVar el método empleado para obtener las ponderaciones replicadas (información obtenida de la documentación de la encuesta). Si el usuario necesita construir con WesVar ponderaciones replicadas del conjunto de datos de encuesta, deberá tener cierto conocimiento de los métodos de replicación y disponer de las tres variables de diseño de la encuesta asociadas al plan de muestreo común *WR* (variable de estratificación *STRATVAR*, variable UPM *PSUVAR* para cada estrato, variable de ponderación de la muestra *WTVAR*).

175. WesVar posee amplia capacidad para construir ponderaciones replicadas de conjuntos de datos de encuestas. El paquete ofrece cinco técnicas de replicación distintas y permite ajustes para las faltas de respuesta, así como posestratificación y *raking*. Además, WesVar ofrece la opción de incorporar un término de corrección al muestreo de poblaciones finitas para muestreo monoetápico utilizando las técnicas *jackknife* para la estimación de varianzas.

176. Para quienes se adentran por primera vez en las técnicas de replicación para la estimación de varianzas, el apéndice A de la guía del usuario de WesVar brinda una excelente descripción general de la teoría y práctica de las técnicas de replicación, aunque la lectura de este material requiere poseer cierta formación en estadística matemática. Por otra parte, el apéndice D de la guía del usuario ofrece orientaciones muy útiles y varios ejemplos sobre cómo escoger un método de replicación para un plan de muestreo determinado.

177. WesVar puede estimar funciones definidas por el usuario de los parámetros de población, algo que es más difícil de hacer si se emplea la linealización mediante el desarrollo en serie de Taylor para la estimación de varianzas. Por consiguiente, tiene mayor flexibilidad intrínseca que los demás paquetes de *software* examinados en este capítulo por lo que se refiere a los parámetros de población que puede estimar. Aunque SUDAAN permite utilizar los métodos de replicación BRR y *jackknife* para la estimación de varianzas, SUDAAN no permite, al contrario que WesVar, que el usuario especifique funciones de los parámetros de población que se desea estimar.

178. En comparación con otros paquetes de *software* de análisis de encuestas, trabajar con los resultados directos que se obtienen con WesVar presenta cierta dificultad. Los resultados de WesVar se muestran en una fila por cada celda del cuadro solicitada (como se ilustra en la sección VI del anexo). No obstante, existe un visor de cuadros que se puede descargar gratuitamente desde el sitio web de WesVar. Este programa auxiliar convierte los resultados de WesVar 4 en un formato de malla o tabular que se puede imprimir o visualizar en pantalla, o grabar como archivo electrónico que puede pegarse en aplicaciones como Microsoft Word o Excel.

179. En comparación con otros paquetes de *software* examinados en este capítulo, WesVar es un paquete de bajo costo. La versión 4 se puede descargar gratuitamente para su evaluación durante 30 días, y la versión 2 se puede descargar gratuitamente y utilizar por tiempo indefinido.

L. PC-CARP

180. PC-CARP es un programa autónomo para MS-DOS desarrollado y distribuido por el Departamento de Estadística de la Universidad del Estado de Iowa. El programa gestiona el plan de muestreo común *WR* ya comentado y en los diseños más simples permite

incorporar términos *fpc* hasta en dos etapas de muestreo. Para la estimación de varianzas se usa la linealización mediante el desarrollo en serie de Taylor.

181. Se pueden construir estimaciones por puntos, errores estándar estimados e intervalos de confianza para totales, medias, proporciones, cuantiles, funciones de distribución empíricas, relaciones y diferencias de relaciones (y diferencias de medias, proporciones y totales, pues) de población y subpoblación. También incluye regresión lineal basada en el diseño y una función de análisis de cuadro de contingencia de doble entrada, incluida una prueba de ji-cuadrado. Se puede calcular el efecto del diseño y el coeficiente de variación para las estimaciones por puntos. El programa dispone de tres módulos adicionales: PC-CARPL para regresión logística basada en el diseño, POSTCARP para posestratificación de datos de encuestas, y EV CARP para análisis de regresión con error de medición en las variables explicativas.

182. La interfaz de usuario consiste en pantallas de menús de texto con desplazamiento mediante el teclado; no permite utilizar el ratón. Los datos de entrada sólo se aceptan en formato ASCII, con los registros delimitados mediante espacios o de longitud fija con una sentencia de formato de soporte con sintaxis FORTRAN. No hay restricciones en cuanto al número de observaciones del conjunto de datos, y en la mayoría de los análisis se aceptan hasta 50 variables. PC-CARP puede ejecutarse con sistemas informáticos antiguos bajo DOS 5.0 o posterior, Windows 3.1 o Windows 95 o posterior. Sólo ocupa 3 megabytes (Mb) de espacio en el disco duro y sólo requiere 450 kilobytes (Kb) de memoria de acceso aleatorio (RAM). Para ejecutar PC-CARP en sistemas más modernos, éstos deben poder soportar programas en DOS.

183. El precio de compra de PC-CARP es bajo en comparación con el de otros paquetes de *software* examinados. Se paga una sola vez y no hay que abonar ninguna cuota anual de renovación de licencia. Los módulos adicionales se pueden adquirir individualmente por una pequeña cantidad.

184. En el capítulo no se han incluido ejemplos de análisis de la encuesta de Burundi realizados con PC-CARP.

M. CENVAR

185. CENVAR es un componente de un completo sistema de *software* estadístico llamado Integrated Microcomputer Processing System (IMPS) que fue diseñado por la Oficina del Censo de los Estados Unidos para tratamiento, gestión y análisis de datos de encuestas complejas. IMPS, incluido CENVAR, se puede obtener sin costo alguno descargándolo de http://www.census.gov/ipc/www/imps/download.htm. A comienzos de 2003, parte de IMPS se había adaptado a Windows y otra parte aún funcionaba bajo DOS. En el capítulo no se ha incluido ningún análisis de IMPS.

186. CENVAR es una adaptación de PC-CARP y tiene, por tanto, muchas de sus características. CENVAR soporta los mismos diseños muestrales que PC-CARP, es decir, el plan de muestreo común *WR* y la posibilidad de incorporar términos *fpc* en la estimación de varianzas para diseños sencillos de una y dos etapas con muestreo sin reemplazamiento. Para la estimación de varianzas se usa la linealización mediante el desarrollo en serie de Taylor. El *software* se controla por menús y no permite el uso del ratón.

187. Se pueden construir estimaciones por puntos, errores estándar estimados, intervalos de confianza, coeficientes de variación y efectos de diseño para totales, medias, proporciones, relaciones y diferencias de relaciones (y, por tanto, diferencias de medias, proporciones y totales) de población y subpoblación. No incluye ninguna de las demás funciones de PC-CARP: regresión lineal basada en el diseño, función de análisis de cuadro de contingencia de doble entrada o estimación de cuantiles. Los módulos adicionales de PC-CARP no están incluidos en CENVAR.

188. La guía del usuario de CENVAR (1995), de unas 100 páginas, puede descargarse de la web. Contiene ejemplos y ejercicios útiles con datos de tres encuestas realizadas por la Oficina del Censo de los Estados Unidos. CENVAR sólo admite datos de entrada en formato ASCII y requiere el *software* "Diccionario de datos" de IMPS. El diccionario debe crearse antes de ejecutar CENVAR. Por consiguiente, antes de utilizar CENVAR es necesario un cierto grado de conocimiento de IMPS. CENVAR se ejecuta en PC en entorno DOS 3.2 o superior. Necesita 10 Mb de espacio de almacenamiento en disco y 640 Kb de memoria disponible. En este capítulo no se han incluido ejemplos de análisis de la encuesta de Burundi realizados con CENVAR.

N. IVEware (versión beta)

189. IVEware (*Imputation and Variance Estimation Software*) es una aplicación bajo SAS para análisis de datos de encuestas desarrollada recientemente por el Programa de Metodología de Encuestas de la Universidad de Michigan. La aplicación gestiona el plan de muestreo común *WR* y usa la linealización mediante el desarrollo en serie de Taylor o los métodos de replicación, dependiendo del procedimiento.

190. El módulo IMPUTE emplea un enfoque de regresión secuencial con múltiples variables para imputar valores omitidos de los elementos, incluida la imputación a múltiples conjuntos de datos. El módulo DESCRIBE permite estimar medias y proporciones de población y subpoblación, diferencias de subgrupos y contrastes lineales de medias y proporciones; se emplea la linealización mediante el desarrollo en serie de Taylor. El módulo REGRESS contiene varios modelos de regresión basados en el diseño (lineal, logística, etcétera); se emplea la técnica de replicación *jackknife*. El módulo SASMOD permite que el usuario tenga en cuenta características de diseño muestral complejo al utilizar varios procedimientos SAS para el análisis de datos como CATMOD, GENMOD y MIXED. Es posible realizar un análisis de imputación múltiple para los tres módulos de análisis de datos (DESCRIBE, REGRESS, SASMOD).

191. IVEware se ejecuta con SAS 6.12 o superior en PC con sistema operativo Microsoft Windows o Linux; existen versiones para otras plataformas. Aunque el usuario no necesita estar familiarizado con los bloques de construcción de IVEware en lenguaje de macros de SAS, C y FORTRAN, debe poseer cierta experiencia con SAS y, por supuesto, el *software* SAS. El *software* IVEware y la documentación se pueden descargar gratuitamente desde http://www.isr.umich.edu/src/smp/ive. En este capítulo no se han incluido ejemplos de análisis de la encuesta de Burundi realizados con IVEware.

O. Conclusiones y recomendaciones

192. A algunos analistas de datos les sorprenderá que se necesite *software* especializado para la estimación de varianzas con datos de encuestas complejas. Aunque algunos analistas preferirán usar para la estimación de varianzas con datos de encuestas complejas un *software* desarrollado para muestras aleatorias simples, no recomendamos esta práctica. Actualmente existen varias opciones de *software* para la estimación de varianzas, algunas de las cuales son gratuitas. La elección se basará posiblemente en la familiaridad con el *software*, el costo, la facilidad de uso y en el tipo de análisis que nos interese: análisis descriptivo básico o análisis más completo.

193. Para quien ya utilice un paquete de estadística general con capacidad de estimación de varianzas para encuestas, ese paquete será una elección obvia, ya que el coste de adquisición está pagado y se conoce la sintaxis. Los usuarios de STATA, que dispone de amplia capacidad de estimación de varianzas para encuestas, no tendrán necesidad de buscar otra cosa salvo que el conjunto de datos que vayan a analizar exija el uso de métodos de replica-

ción. Los usuarios de SAS disponen con la versión 9.0 de una mayor capacidad de estimación de varianzas para encuestas, y cabe esperar que esta capacidad se incremente en el futuro. No obstante, si la versión 9.0 de SAS no resultara suficiente para la estimación de varianzas para encuestas en una encuesta concreta, el paquete gratuito IVEware bajo SAS puede satisfacer sus necesidades. Los usuarios de Epi-Info sólo disponen con este paquete de capacidad básica de estimación de varianzas en los datos de encuestas, pero esto será suficiente si no se necesita nada más. SPSS, paquete de análisis estadístico de amplia aceptación, publicó un módulo adicional para encuestas complejas a finales de 2003, por lo que actualmente representa una opción viable.

194. Si el paquete de *software* de estadística general que se esté utilizando carece de la necesaria capacidad de estimación de varianzas para encuestas, convendrá considerar la adquisición de un paquete de *software* de encuestas especializado (por ejemplo, WesVar, SUDAAN, PC-CARP o CENVAR) u otro paquete de estadística general (por ejemplo, STATA o SAS con o sin IVEware, o SPSS, o quizá Epi-Info). SUDAAN suele resultar atractivo para los usuarios de SAS debido a que su sintaxis es similar a la de SAS y a la opción de ejecutarlo bajo SAS, aunque en un entorno autónomo acepta también conjuntos de datos de entrada de SPSS. WesVar, PC-CARP y CENVAR son todos ellos programas autónomos con su propia organización singular, por lo que el conocimiento de algún otro paquete estadístico no influirá en la elección entre estos tres. PC-CARP y CENVAR pueden interesar a los que deban o prefieran trabajar en un entorno DOS, y puede que no interese a quienes prefieran un entorno Windows.

195. Si el costo representa un factor de importancia en la selección del *software*, algunos paquetes serán definitivamente más preferibles. Epi-Info, aunque es gratuito, presenta limitaciones en cuanto a las opciones de análisis para la estimación de varianzas para encuestas, pero puede ser perfecto para análisis básicos. CENVAR, también gratuito, dispone de más opciones de análisis que Epi-Info, pero no cuenta con procedimientos de regresión basados en el diseño. WesVar versión 2 es también gratuito. IVEware es gratuito, pero debe ejecutarse en conjunción con SAS. WesVar versión 4 y PC-CARP son completos paquetes de *software* para encuestas de precio económico. STATA y SUDAAN en versión autónoma son de costo moderado; SAS resulta más costoso.

196. Otro factor para la elección del *software* puede ser el método de estimación de varianzas utilizado. Por ejemplo, para analizar un conjunto de datos a disposición del público que incluye ponderaciones replicadas con BRR o *jackknife* pero no las variables de identificación de estratos o UPM, un paquete de *software* que utilice la linealización mediante el desarrollo en serie de Taylor no resultará de utilidad. Entre los paquetes examinados, tanto SUDAAN como IVEware ofrecen linealización mediante el desarrollo en serie de Taylor y, además, métodos de replicación; WesVar sólo ofrece procedimientos de replicación, y STATA, SAS, PC-CARP, Epi-Info y CENVAR sólo ofrecen linealización mediante el desarrollo en serie de Taylor.

197. Por último, la elección del *software* dependerá de los análisis que se quieran hacer. Los ocho paquetes analizados en los apartados anteriores permiten realizar análisis básicos y descriptivos. De los ocho paquetes, los que van más allá del análisis básico son STATA, SUDAAN, WesVar, PC-CARP y SAS (con o sin IVEware). En el cuadro XXI.2 *supra* se resumen y comparan muchos de los atributos de estos ocho paquetes de *software*.

198. Los cinco paquetes de *software* sometidos a una comparación empírica en este capítulo (SAS, SUDAAN, STATA, Epi-Info y WesVar) ofrecen, como cabía esperar, las mismas estimaciones por puntos en todos los ejemplos descriptivos y analíticos. Los cinco paquetes producen básicamente los mismos errores estándar estimados, tanto si usan BRR como si emplean la linealización mediante el desarrollo en serie de Taylor. Algunos de los cálculos de intervalos de confianza presentan ligeras diferencias entre los cinco paquetes;

las causas de estas diferencias se han comentado anteriormente. Por consiguiente, no existe ninguna razón de peso para decidir la elección entre estos cinco paquetes en función de los análisis comparativos presentados en este capítulo.

199. El mercado de paquetes de *software* especializados para encuestas (orientados a la estimación de varianzas) podría desaparecer en el futuro, por la tendencia a incluir estas capacidades en los paquetes estadísticos estándar (por ejemplo, STATA, SAS y SPSS). Puede, pues, que en el futuro a los analistas de datos les resulte más fácil obtener y utilizar un *software* adecuado para la estimación de varianzas con datos de encuestas complejas.

ANEXO

Véase la llamada que al respecto figura al principio del capítulo (página 391).

Reconocimientos

Deseamos expresar nuestro reconocimiento a las siguientes personas:

A Michael S. Deming, MD, MPH, por facilitarnos el conjunto de datos de la encuesta de Burundi y su documentación, por su atenta lectura de numerosos borradores y por sus valiosas propuestas de revisión.

A Kevin Sullivan, PhD, por enseñarnos a navegar por Epi-Info y por sus valiosas indicaciones al respecto, por su atenta lectura de numerosos borradores y por sus valiosas propuestas de revisión.

A Z. T. Daniels, MS, MBA, por dar formato a los resultados de salida y a los cuadros de texto de WesVar y por localizar los datos de población de Burundi en Internet.

A Graham Kalton, PhD e Ibrahim Yansaneh, PhD, por sus valiosas propuestas de organización y de revisión.

A James Chromy, PhD, por su atenta lectura de los borradores y por sus valiosas propuestas de revisión.

A diversos árbitros y revisores académicos, por su atenta lectura de los borradores y por sus valiosas propuestas de revisión.

A Paul Weiss, MS, por enseñarnos a navegar por WesVar y por sus valiosas indicaciones al respecto.

Los errores que pueda haber en este capítulo son responsabilidad exclusiva del autor.

REFERENCIAS BIBLIOGRÁFICAS

Brogan, Donna (1998 y en curso de publicación). *Software* for sample survey data: misuse of standard packages. Capítulo por invitación en *Encyclopedia of Biostatistics*, Peter Armitage y Theodore Colton, comps. Nueva York: John Wiley, vol. 5, págs. 4167-4174. Artículo revisado en curso de publicación con la 2a. edición de la *Encyclopedia of Biostatistics*.

Brogan, Donna, y otros (1994). Increasing the accuracy of the expanded programme on immunization's cluster survey design. *Annals of Epidemiology*, vol. 4, No. 4, págs. 302-311.

Carlson, Barbara L. (1998). *Software* for sample survey data. En *Encyclopedia of Biostatistics*, vol. 5, Peter Armitage y Theodore Colton, editores jefes. Nueva York: John Wiley and Sons, págs. 4160-4167.

Cochran, William G. *Sampling Techniques*, 3a. ed. Nueva York: John Wiley and Sons.

Expanded Programme on Immunization (EPI) (1996). Estimating tetanus protection of women by serosurvey. *Weekly Epidemiological Record*. (Organización Mundial de la Salud), vol. 71, págs. 17-124.

Hansen, Morris H., William N. Hurwitz y William G. Madow (1953). *Sample Survey Methods and Theory*, vol. I, *Methods and Applications*. Nueva York: John Wiley and Sons.

Judkins, D. (1990). Fay's method for variance estimation. *Journal of Official Statistics,* vol. 6, págs. 23-240.

Kish, Leslie (1965). *Survey Sampling*. Nueva York: John Wiley and Sons.

————, y M. R. Frankel (1974). Inference from complex samples. *Journal of the Royal Statistical Society, Series B,* vol. 36, págs. 1-37.

Korn, Edward L., y Barry I. Graubard (1999). *Analysis of Health Surveys*. Nueva York: John Wiley and Sons.

Krotki, Karol P. (1998). Sampling in developing countries. En *Encyclopedia of Biostatistics*, vol. 5. Peter Armitage y Theodore Colton, editores jefes. Nueva York: John Wiley and Sons, págs. 3939-3944.

Levy, Paul S. y Stanley Lemeshow (1999). *Sampling of Populations: Methods and Applications*, 3a. ed. Nueva York: John Wiley and Sons.

Lohr, Sharon L. (1999). *Sampling: Design and Analysis*. Pacific Grove, California: Duxbury Press, Brooks/Cole Publishing.

Rao, J. N. K., y A. J. Scott (1981). The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables. *Journal of the American Statistical Association*, vol. 76, págs. 221-230.

———— (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *Annals of Statistics,* vol. 12, págs. 46-60.

Rust, K. F., y J. N. K. Rao (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, vol. 5, págs. 283-310.

Shah, Babubhai V. (1998). Linearization methods of variance estimation. En *Encyclopedia of Biostatistics*, vol. 3. Peter Armitage y Theodore Colton, editores jefes. Nueva York: John Wiley and Sons, págs. 2276-2279.

Som, R. K. (1995). *Practical Sampling Techniques*, 2a. ed. Nueva York, Basilea y Hong Kong: Marcel Dekker.

Wolter, Kirk M. (1985). *Introduction to Variance Estimation*. Nueva York: Springer-Verlag.

REFERENCIAS DE *SOFTWARE*

CENVAR Variance Calculation System: IMPS Version 3.1: User's Guide, 1995. Bureau of the Census, United States Department of Commerce, Washington, D.C. Sitio de descarga: http://www.census.gov/ipc/www/imps/download.htm.

Epi-Info. Sitio de descarga: http://www.cdc.gov/epiinfo/, tanto para el *software* como para la documentación.

IVEware. Sitio de descarga: http://www.isr.umich.edu/srcs/mp/ive, tanto para el *software* como para la documentación.

PC CARP (1986, 1989). *User's Manual,* Wayne Fuller y otros, eds. Statistical Laboratory, Iowa State University, Ames, Iowa. Sitio de descarga: http://cssm.iastate.edu/*software*.

Research Triangle Institute (2001). *SUDAAN User's Manual, Release 8.0*. Research Triangle Park, North Carolina: Research Triangle Institute. Sitio de descarga: www.rti.org/sudaan.

SAS/STAT. Sitio de descarga: http://www.sas.com/technologies/analytics/statistics/stat/index.html, para información sobre el *software* SAS/STAT, que incluye procedimientos para procesar los datos de las encuestas.

STATA. Sitio de descarga: http://www.stata.com (STATA); http://www.stata.com/help.cgi?svy (examen de los mandatos svy de STATA), y http://www.stata.com/bookstore (información sobre el manual de referencia).

WesVar 4.2 User's Guide (2002). Rockville, Maryland: Westat. Véase también el sitio web http://www.westat.com/WesVar/about/.

Segunda parte

Estudios de casos

Introducción

Gad Nathan
Hebrew University
Jerusalén, Israel

1. En la primera parte de esta publicación se han presentado los "últimos avances" aparecidos en los aspectos más importantes del diseño y análisis de encuestas de hogares en los países en desarrollo y en transición. Se ha insistido en los principios y metodologías generales para el diseño, realización y análisis de encuestas, aplicables a las encuestas de hogares en países en desarrollo y en transición, con especial insistencia en las características practicas: efectos del diseño, costos de la encuesta y errores no muestrales. Son muchos y muy variados los métodos y técnicas elaborados y aplicados a las encuestas de hogares en países en desarrollo y en transición. Por eso, la presentación de los capítulos precedentes ha sido muy amplia para poder abarcar el mayor número posible de aspectos. En los capítulos precedentes se han incluido numerosos ejemplos de aplicación, y en algunos capítulos específicos se han considerado aplicaciones concretas a una serie de encuestas en países en desarrollo y en transición. Así, en el capítulo VII se han descrito los diseños muestrales y se han presentado datos sobre los efectos del diseño en 11 encuestas de 7 países. Igualmente, en el capítulo XI se ha presentado un estudio de caso con detalles sobre las prácticas de notificación, control, evaluación y compensación de los errores no muestrales en el Brasil.

2. No obstante, para los profesionales es de la máxima importancia observar cómo se combinan en la práctica, en una aplicación de la vida real, las distintas técnicas y métodos propuestos y observar ejemplos concretos de la integración de los métodos en una encuesta de hogares completa bien concebida y analizada. Las condiciones específicas en cada país y su infraestructura tienen una importante influencia en la forma en que se aplican de hecho los principios generales y, en particular, en su combinación con una encuesta completa. En general, los estudios de casos son un instrumento de aprendizaje fundamental para cualquier ciencia aplicada, y el examen de la aplicación de los conceptos estadísticos teóricos y de sus resultados al diseño y análisis de encuestas estadísticas mediante estudios de casos detallados es especialmente fecundo. Por esta razón hemos reservado esta segunda parte a estudios de casos. Con ello esperamos situar en el contexto de la vida real los métodos examinados en la primera parte. Ello demostraría no sólo la aplicación de aspectos específicos de las técnicas estudiadas sino, sobre todo, su integración en programas completos de diseño y análisis para las encuestas de hogares en países en desarrollo y en transición.

3. En los cuatro capítulos de esta parte de la publicación se consideran varios centenares de encuestas de hogares de todos los lugares del mundo sobre temas muy heterogéneos, en condiciones diversas y con distintos diseños, en forma más o menos detallada. En la mayoría de los casos, los estudios de casos describen los objetivos y ámbito de las encuestas, la definición de la población y el diseño de la muestra, los instrumentos de la encuesta, el diseño del trabajo sobre el terreno y su ejecución, los errores de falta de respuesta y la evaluación, el análisis, la ponderación y los efectos del diseño. En algunos casos las encuestas descritas se estandarizaron en relación con los parámetros de diseño en un gran número de

encuestas de organizaciones internacionales. En otros casos había semejanzas en el diseño de la encuesta debido a condiciones semejantes en los países vecinos (por ejemplo, en los países en transición).

4. En el capítulo XXII se describen las características y el diseño del programa de encuestas demográficas y de salud (DHS) en más de 100 encuestas de hogares y de individuos en más de 50 países. En el capítulo XXIII se describen las características operativas de la serie de más de 60 encuestas del Estudio de medición de los niveles de vida (EMNV) realizadas bajo el patrocinio del Banco Mundial en más de 40 países. En el capítulo XXIV se examinan distintos diseños muestrales y cuestiones relacionadas con la medición que son específicas de las encuestas sobre el presupuesto de hogares (HBS), basadas en experiencias con tales encuestas en varios países en desarrollo y en transición. Un estudio acerca de la Encuesta sobre gasto y consumo de la República Democrática Popular Lao, de 1997-1998, incluye descripciones detalladas de las condiciones generales para las actividades de la encuesta, los instrumentos pertinentes, los métodos de medición, el diseño muestral y las actividades sobre el terreno. Una evaluación de las experiencias en estas esferas ha dado lugar a conclusiones interesantes. Finalmente, en el capítulo XXV se examinan los principales aspectos del diseño y realización de encuestas de hogares en 14 países en transición de Europa oriental, con descripciones detalladas de estudios de casos de encuestas de hogares en una selección de seis de ellos.

5. Algunas de las características descritas tienen mucho en común. Por ejemplo, todas las encuestas eran encuestas de hogares o incluían algún elemento relacionado con los hogares. No obstante, en muchos casos la unidad de análisis era fundamentalmente el individuo; un solo individuo por hogar (por ejemplo, las mujeres en las encuestas demográficas y de salud) o todos los individuos del hogar (por ejemplo, en las encuestas de mano de obra), y en muchos casos la respuesta se obtenía a través de un intermediario. Los diseños muestrales básicos eran bastante semejantes en casi todas las encuestas descritas: muestreo por conglomerados en varias etapas con grandes unidades geográficas como unidades primarias de muestreo (UPM). Se intentó también cierta estratificación de las UPM. En la mayoría de los casos, los diseños eran autoponderados en lo que respecta a los hogares. No obstante, cuando se elegía un solo individuo por hogar, la muestra de individuos ya no era autoponderada. Prácticamente todos los diseños eran diseños de probabilidad completa, aunque las encuestas sobre el presupuesto de hogares de la República Checa y de Eslovaquia utilizaban todavía el muestreo por cuotas.

6. Las metas y objetivos de las encuestas presentan grandes diferencias, Por ejemplo, las encuestas demográficas y de salud tratan de "ofrecer a los países los datos necesarios para supervisar y evaluar los programas de población, salud y nutrición". El programa EMNV tiene como principal finalidad la comprensión, medición y seguimiento de las condiciones de vida. El programa de encuestas sobre el presupuesto de los hogares trata de medir los aspectos importantes del presupuesto cotidiano de los hogares: ingresos y gastos. La gran variedad de encuestas de hogares en los países en transición se ha concentrado en el análisis de las condiciones de vida, la elaboración de índices de precios de consumo y las estadísticas sobre la mano de obra necesarias para la transición de una economía estatal a una economía de mercado.

7. Los instrumentos utilizados en estas encuestas se basaban, en general, en entrevistas sobre el terreno con cuestionarios rellenados con lápiz y papel. No obstante, se señaló un primer intento de utilización de entrevistas por teléfono con ayuda de computadora (CATI) en la encuesta de población activa de Estonia (capítulo XXV). Se dio prioridad a la capacitación y al control de los entrevistadores en muchas de las encuestas señaladas y se realizaron varios intentos de reducir la falta de respuesta y los errores de respuesta. Se señalaron altas tasas de respuesta en el caso de las encuestas democráticas y de salud (DHS): del 88% al

99% de los hogares y entre el 87% y el 99% de las mujeres. En las encuestas del EMNV se observaron también tasas altas de respuesta global (74%-99,7%). No obstante, se registraron también altas tasas de ausencia de datos sobre los ingresos, en particular en el caso de los trabajadores por cuenta propia. La Encuesta sobre el presupuesto de los hogares de la República Democrática Popular Lao tenía una tasa de falta de respuesta de sólo el 3,1%. Por el contrario, las encuestas sobre presupuestos de los hogares en los países en transición registraron tasas de falta de respuesta que oscilaban entre el 8% y el 49%. La respuesta fue algo mejor en las encuestas de población activa de esos países, ya que las tasas de falta de respuesta fueron del orden del 4% al 29%, y en algunos países eran sistemáticamente inferiores al 10%.

8. En gran parte de los estudios de casos se insiste mucho en los esfuerzos de depuración, edición e imputación de los datos. La mayoría del procesamiento y análisis se llevó a cabo mediante programas informáticos estándar, muchas veces sin ponderación. Los países en transición utilizaron métodos de ponderación y calibración con gran frecuencia. Muchos de los estudios trataron de estimar los efectos del diseño utilizando métodos estándar. Estas estimaciones se emplearon tanto en el análisis como para mejorar el diseño en el futuro. Así, un examen de los efectos del diseño del EMNV ha indicado la necesidad de utilizarlos en el análisis, pero las grandes discrepancias en los efectos del diseño para las distintas variables importantes no ha permitido llegar a conclusiones útiles sobre el diseño muestral debido a que en muchas encuestas se consideraban a la vez temas diversos.

9. Además de ofrecer posibilidades de aprender de una gran variedad de experiencias aquí presentadas en relación con distintos tipos de encuestas en diferentes países, los informes llegan a conclusiones importantes para el tipo de encuestas considerado. Entre ellas se incluye la necesidad de actualizar constantemente los marcos muestrales, la insistencia constante en la capacitación sobre el terreno y el control de los entrevistadores, la importancia de la preparación de datos de calidad, la formulación y actualización de los requisitos y análisis en materia de datos, el uso de los efectos del diseño y muchas otras consideraciones. Junto con los métodos descritos en la primera parte, los estudios de casos constituyen una parte integrante y muy valiosa de lo que se puede aprender de esta publicación.

Capítulo XXII
Las encuestas demográficas y de salud

Martin Vaessen
ORC Macro
Calverton, Maryland
Estados Unidos de América

Mamadou Thiam*
UNESCO
Montreal
Canadá

Thanh Lê*
WESTAT
Rockville, Maryland
Estados Unidos de América

* Mamadou Thiam y Thanh Lê colaboraron en el pasado con el programa DHS (encuestas democráticas y de salud), ORC Macro.

RESUMEN

En este capítulo se presenta un panorama general de los principales procedimientos empleados en el programa de encuestas demográficas y de salud (DHS) para la realización de encuestas de hogares e individuos en gran escala. Se describen a grandes rasgos el contenido general de las encuestas, los procedimientos de muestreo, las tasas de respuesta y los efectos del diseño, así como los procedimientos y planteamientos seguidos para todos los componentes importantes de las encuestas, desde la capacitación a la elaboración de datos y la redacción de informes. El capítulo contiene también una recopilación de las principales enseñanzas aprendidas hasta ahora gracias a la ejecución de este programa de encuestas.

Términos clave: Encuestas de hogares, tasas de respuesta, muestreo de encuestas, error de muestreo, efectos del diseño, trabajo sobre el terreno.

A. Introducción

1. El programa de encuestas demográficas y de salud (DHS) viene realizando encuestas de hogares en países en desarrollo de todo el mundo desde 1984. El principal objetivo de estas encuestas es ofrecer a los países los datos necesarios para el seguimiento y evaluación de los programas de población, salud y nutrición en forma periódica. La creciente insistencia de los donantes y de los países en la utilización de indicadores objetivos para cuantificar dicho progreso ha incrementado el uso de datos de encuestas de hogares periódicas, dada la ausencia de información adecuada procedente de las estadísticas administrativas y otros sistemas habituales de recopilación de datos. En cualquier encuesta DHS se selecciona una muestra de hogares en todo el país y luego se realizan entrevistas utilizando un cuestionario de hogares para recopilar información sobre las características del hogar e identificar a todos los miembros del mismo y sus características básicas. Se entrevista también a las mujeres con edades comprendidas entre 15 y 49 años utilizando un cuestionario de mujeres para recopilar información sobre todo acerca de las características básicas, comportamiento reproductivo, conocimiento y uso de los métodos anticonceptivos, salud de los niños y de la mujer y otras cuestiones. La duración media de una entrevista oscila entre 35 y 40 minutos, con un margen general de 10 a 90 minutos, aunque algunas entrevistas son todavía más largas. Las muestras presentan considerables diferencias de tamaño: entre 5.000 y 30.000 mujeres. En algunos países se entrevista también a una muestra de hombres con edades comprendidas entre 15 y 59 años. Muchas veces se trata de una submuestra de la muestra utilizada para seleccionar a las mujeres. Las entrevistas a los hombres suelen durar unos 25 minutos. En los

apartados siguientes se presenta la historia del programa DHS junto con el contenido general de sus encuestas, un panorama de sus procedimientos de muestreo y un análisis de la falta de respuesta por unidades. Se presentan también los efectos del diseño muestral, así como las diferentes fases de la realización de la encuesta y las enseñanzas aprendidas gracias a la realización de las encuestas demográficas y de salud en los países en desarrollo.

B. Historia

2. Las encuestas demográficas y de salud son la prolongación de dos programas anteriores de encuestas de hogares: los estudios mundiales sobre fecundidad (WFS) y los estudios sobre la frecuencia del uso de anticonceptivos (CPS). Los primeros tuvieron lugar entre 1973 y 1984; las segundos, entre 1977 y 1985. El programa WFS realizó encuestas en 41 países en desarrollo y colaboró en encuestas en 20 países desarrollados. Los estudios mundiales sobre la fecundidad pretendían obtener sobre todo información acerca de la fecundidad, la planificación familiar y, en cierto modo, la salud infantil. El programa estaba financiado conjuntamente por la Agencia de los Estados Unidos para el Desarrollo Internacional (USAID) y el Fondo de Población de las Naciones Unidas (FNUAP), con ayuda de los gobiernos del Reino Unido de Gran Bretaña e Irlanda del Norte, los Países Bajos y el Japón.

3. El programa CPS realizó 43 encuestas en 33 países y se centró más concretamente en la planificación familiar. Estuvo financiado por USAID y las encuestas se limitaron a países que habían recibido asistencia para el desarrollo de USAID.

4. Las encuestas demográficas y de salud comenzaron en 1984. Para finales de 2003, se habrían realizado unas 150 encuestas de mujeres, 75 encuestas de hombres y 10 encuestas de servicios sanitarios en unos 70 países. Las encuestas normalmente se realizan una vez cada cinco años, aunque en algunos países se hacen con mayor frecuencia. En general, se concentran en los países que han recibido asistencia de USAID, aunque algunos países han participado con financiamiento del Banco Mundial o el FNUAP. En muchos países, las encuestas cuentan con apoyo de otros donantes además de USAID, como el Departamento para el Desarrollo Internacional (DFID) del Reino Unido, el Fondo de las Naciones Unidas para la Infancia (UNICEF), los gobiernos del Japón y de Suecia, entre otros. Las encuestas demográficas y de salud permiten obtener un panorama general de la población y de la salud materna e infantil en los países participantes, y los datos son de libre acceso para los organismos interesados en actividades de seguimiento y evaluación. El contenido de las encuestas ha cambiado a lo largo de los años debido a la evolución de las circunstancias y prioridades.

C. Contenido

5. El contenido básico de cada ronda de encuestas demográficas y de salud es el mismo en todos los países, a fin de lograr la máxima comparabilidad de la información. Además de este contenido básico, los países pueden decidir incorporar al cuestionario módulos de especial interés para cada uno de ellos. El contenido básico de los cuestionarios de los países del África subsahariana es algo distinto al de los demás países, sobre todo en lo que respecta a su complejidad.

6. Los cuestionarios básicos correspondientes al período 1997-2002 incluían los siguientes elementos:

Cuestionario de los hogares. En él se obtenían datos básicos sobre edad, sexo, supervivencia de los padres y escolarización de los miembros del hogar. Contenía también información sobre el abastecimiento de agua y los servicios disponibles en el hogar. En el cuestionario de hogares se recopilaba además información sobre la altura y peso de las mujeres de 15 a 49 años y de los niños de menos de 5 años, así como sobre sus niveles de hemoglobina, para la medición de la anemia.

Cuestionario de las mujeres. Este cuestionario, aplicado a las mujeres en edad fecunda, contenía las siguientes secciones:

- Características básicas del informante;
- Historial de reproducción;
- Uso de métodos anticonceptivos;
- Embarazo, atención post-natal y lactancia materna;
- Inmunización, salud y nutrición;
- Matrimonio y actividad sexual;
- Preferencias en materia de fecundidad;
- Antecedentes del esposo y trabajo de la mujer;
- VIH/SIDA y otras infecciones de transmisión sexual.

En algunas encuestas se incluían pruebas del VIH/SIDA o sífilis u otros biomarcadores.

7. Hay también un cuestionario para hombres. En él se incluyen algunos de los temas ya mencionados en el cuestionario de la mujer. No se aplica en todos los países. Además, existe un cuestionario para proveedores de servicios de salud y planificación familiar, pero es independiente de la encuesta de hogares y se administra a los proveedores de los servicios. Se conoce con el nombre de cuestionario para la evaluación de la prestación de servicios. En él se incluyen todos los aspectos relacionados con la prestación de servicios mediante preguntas a proveedores y clientes y observaciones pertinentes.

8. El programa DHS ha elaborado una serie de módulos que los países pueden agregar a su cuestionario. Los hay sobre las siguientes materias:

- Mutilación genital de la mujer;
- Mortalidad materna;
- Consumo de pastillas;
- Experiencia en materia de esterilización;
- Consanguinidad (matrimonio entre parientes consanguíneos);
- Autopsia verbal (preguntas detalladas sobre la causa del fallecimiento);
- VIH/SIDA;
- Educación de los niños;
- Condición de la mujer;
- Violencia doméstica;
- Paludismo;
- Gastos en salud de los hogares.

9. Debido a la longitud del instrumento básico, generalmente los países no pueden añadir más de dos o tres módulos, aunque ello puede variar según sea la longitud de los módulos elegidos (véanse los cuestionarios y otros materiales en www.measuredhs.com).

D. Marco muestral

10. El tema de la disponibilidad de un marco muestral adecuado se plantea obviamente en las primeras etapas de planificación de una encuesta demográfica y de salud. Ésta recopila datos sobre individuos que residen en hogares privados, pero generalmente no se dispone de una muestra actualizada de dichos individuos u hogares. El marco muestral utilizado en la mayor parte de las encuestas demográficas y de salud es, por definición, una lista de unidades de área no superpuestas que abarcan todo el territorio nacional. Las caracterís-

ticas esenciales de estas unidades, en lo que se refiere al marco, son la existencia de límites bien definidos y mapas claramente delineados. Cada unidad de área tiene también un código de identificación único. Debe tener asimismo una medida actual o estimada de su tamaño (población y/o número de hogares). En cada unidad existen otras características, como su condición de urbana o rural, que pueden utilizarse para la estratificación.

11. En la mayor parte de los países, las unidades de área deseadas corresponden a zonas de empadronamiento censal (ZE), que ofrecen un marco conveniente para la primera fase de muestreo. En algunos países, las ZE pueden tener una población muy numerosa; en otros no es así. Cualquiera que sea el tamaño, las ZE suelen ser las unidades primarias de muestreo (UPM). En algunas encuestas, son también las unidades de área últimas cuando son lo suficientemente pequeñas. Si se utilizan como UPM y se comprueba que el tamaño es demasiado grande (número de hogares o población), se introduce en el diseño de la muestra una fase intermedia de selección: la segmentación.

12. Como se ha mencionado antes, el marco, independientemente de que comprenda las ZE del censo u otras unidades, quizá no esté actualizado. Normalmente deben adoptarse medidas para actualizar todo el marco o para actualizarlo parcialmente compilando una lista actualizada de hogares en la fase penúltima de selección.

13. En algunas encuestas se utiliza una muestra maestra preexistente como marco muestral si se comprueba que su diseño puede servir para el objetivo de medición de la encuesta demográfica y de salud.

E. Etapas de muestreo

14. Como ocurre con todo diseño muestral, las características del marco de muestreo y los objetivos de la encuesta determinan el número de etapas de muestreo. Aunque no se haya llegado a un modelo estándar para todos los países, el diseño muestral de cada encuesta demográfica y de salud está inspirado en los mismos principios generales: sencillez, muestreo probabilístico (probabilidad de selección conocida distinta de cero), conglomeración y estratificación. En las encuestas demográficas y de salud se requieren normalmente dos o más etapas de selección, según la medida del tamaño de las unidades de área en el marco muestral.

15. El diseño básico supone seleccionar las unidades de área en la primera etapa con probabilidad proporcional al tamaño, tomando como tamaño los recuentos de población o el número de hogares de cada unidad de área. La primera fase de la selección marca el punto más allá del cual las operaciones de muestreo se trasladan de la oficina al exterior para el proceso cartográfico y, en caso necesario, la enumeración de los hogares en las unidades de área seleccionadas. La cartografía consiste en trazar un mapa esquemático con las fronteras de cada UPM seleccionada y la ubicación de las viviendas dentro de la UPM. En los países con mapas detallados y precisos de las UPM, se trata simplemente de actualizar la ubicación de las viviendas. Cuando se piensa que el marco no está completamente actualizado se incorporan listas actualizadas de hogares a las UPM seleccionadas en las que se enumeran todos los hogares de cada vivienda ocupada, incluidos los hogares que se encuentran ausentes en el momento de la visita del equipo encargado de la enumeración. Las listas obtenidas sirven como marco de muestreo para la selección sistemática de hogares en la segunda etapa.

16. El tamaño de los conglomerados en una encuesta de hogares (número de hogares/mujeres que deben seleccionarse por UPM o conglomerado) depende de la variable considerada. Cuando se trata de variables fuertemente conglomeradas que requieren frecuentes comparaciones entre zonas geográficas (como la tasa de uso de anticonceptivos y sus determinantes), se considera que el tamaño óptimo es entre 15 y 20 mujeres por conglomerado. Otras variables de la fecundidad están menos conglomeradas, y cuando las comparaciones de interés no son geográficas (por ejemplo, comparaciones entre grupos de edad o niveles

de instrucción), el tamaño óptimo del conglomerado puede ser mayor. Las DHS utilizan un conglomerado de entre 30 y 40 mujeres en el sector rural. En las zonas urbanas el ahorro de costos asociado a un conglomerado de mayor tamaño suele ser pequeño, y las DHS utilizan conglomerados de entre 20 y 25 mujeres. Cuando se dispone de una lista reciente de hogares, estas cifras se reducen, ya que el factor que favorece los conglomerados de gran tamaño es el ahorro con respecto a las operaciones de enumeración (ORC Macro, 1996). Como las DHS recopilan también datos sobre la salud de los niños y éstos son hijos de las madres incluidas en la muestra, el tamaño del conglomerado debe ser también lo bastante grande como para contar con el número de niños necesario para el análisis.

17. Todos los individuos pertenecientes a los hogares seleccionados y que reúnan los debidos requisitos se incluyen en la muestra final. Aunque en la mayor parte de las muestras de las DHS el número de hogares seleccionado por UPM varía de unas a otras, en algunas encuestas se ha utilizado un tamaño de muestra fijo.

18. Muchas veces las UPM son de tamaño demasiado grande para que puedan enumerarse directamente. En ese caso se introduce en el diseño un proceso de segmentación para reducir el proceso de enumeración y mantener una carga de trabajo equitativa entre las UPM. Cada UPM de gran tamaño se divide en segmentos, uno de los cuales se conserva en la muestra con probabilidad proporcional al tamaño (PPT).

19. La mayoría de los diseños muestrales de las DHS son conglomerados y estratificados. La estratificación explícita se basa normalmente en criterios geográficos, como el desglose en zonas urbanas/rurales, y se introduce sólo en la primera fase de muestreo. Las UPM se seleccionan independientemente en cada estrato. La estratificación implícita se consigue mediante el uso de la técnica de selección sistemática. Normalmente, el número de UPM es elevado; oscila entre unos 300 y 550 UPM en una muestra de 10.000 hogares.

20. En las DHS se utiliza un diseño lo más sencillo posible, con el fin de facilitar su realización precisa. No obstante, el diseño básico se modifica de acuerdo con las condiciones específicas del país. Entre esas modificaciones se incluye el uso del diseño de segmentos estándar con o sin conglomerados compactos; se entiende por conglomerados compactos aquellos en que cada hogar de la muestra es geográficamente contiguo de otro, mientras que los conglomerados no compactos serían los formados por hogares geográficamente dispersos. Se trata de una variación del diseño muestral en que un tamaño de segmento estándar predeterminado, es decir, la unidad de área última especificada, es lo más pequeño posible, siempre que sea manejable. A cada UPM o zona de empadronamiento i del país se asigna un número de segmentos s_i dividiendo su población del censo por el tamaño del segmento estándar. Luego las UPM se incluyen en la muestra con probabilidad proporcional al tamaño (PPT), en donde la medida del tamaño es igual al número de segmentos s_i. Dentro de cada UPM seleccionada se selecciona luego un segmento aleatoriamente. El caso del segmento estándar con conglomerado compacto es aquel en el que los segmentos son de tamaño medio T, siendo T el tamaño de conglomerado deseado. De esta manera podría evitarse un proceso de enumeración utilizando el planteamiento de "inclusión total" (ORC Macro, 1996).

21. Las estimaciones de las DHS se presentan tanto para el país en su totalidad como para cada campo geográfico concreto, como las zonas urbanas y rurales y las regiones. Como los campos muchas veces son variables en lo que respecta al tamaño de la población, la muestra suele concebirse de tal manera que las áreas pequeñas estén sobrerrepresentadas, con el fin de disponer de los tamaños muestrales adecuados para el análisis. Ello, naturalmente, introduce un posible sesgo en las estimaciones nacionales, que se corrige mediante la ponderación adecuada de los datos. El principal componente de las ponderaciones de la muestra es la ponderación del diseño basada en las probabilidades de selección. En la ponderación se tiene también en cuenta la falta de respuesta tanto por hogares como por individuos. Puede utilizarse una fase final de ponderación en la que se realice un ajuste de posestratificación

siempre que se utilice para la selección muestral un marco de área desfasado, utilizando para ello proyecciones de población de fuentes fiables.

F. Información sobre la falta de respuesta

22. La sustitución de las unidades (hogares o individuos) que no han respondido a las preguntas no está autorizada en la DHS, que en este sentido se distingue de muchas otras encuestas. Con el fin de lograr el número deseado de unidades muestrales, las tasas de falta de respuesta de las unidades de la muestra se estiman a partir de encuestas anteriores o semejantes en el momento del diseño de la muestra y, luego, se utilizan para determinar el número deseado de unidades que se deben seleccionar. Además, durante el trabajo sobre el terreno se realizan numerosos esfuerzos por conseguir tasas de respuesta elevadas. A continuación se presenta un examen de las tasas de respuesta de las DHS, incluida una comparación de dichas tasas a lo largo del tiempo y entre diferentes regiones.

23. Como ya se ha mencionado, los datos de las DHS se recopilan en dos niveles: hogares e individuos. Los individuos que reúnen las debidas condiciones son en su mayoría mujeres en edad de procrear, pero en algunos países se entrevista también a hombres con edades comprendidas entre 15 y 59 años. En las encuestas demográficas y de salud, la falta de respuesta se refiere a la ausencia de entrevistas de hogares o individuos seleccionados para la muestra. Las tasas de respuesta de los hogares e individuos se miden contabilizando con esmero todos los hogares e individuos aptos. La computación operacional de las tasas de respuesta utiliza códigos de respuesta que se introducen en los cuestionarios. El cuestionario de hogares identifica a todos los individuos aptos dentro de cada hogar. Sólo se aplican cuestionarios individuales a las personas que reúnen los requisitos para la encuesta.

24. Los códigos de respuesta relativos a los hogares son los siguientes:

1H Entrevista terminada
2H Ningún miembro del hogar en casa o ningún informante competente en casa
3H Todo el hogar ausente durante un período largo
4H Aplazado
5H Rechazado
6H Vivienda abandonada o dirección que no corresponde a una vivienda
7H Vivienda destruida
8H Vivienda no encontrada
9H Otros

La tasa de respuesta de los hogares es la siguiente:

$$R_H = \frac{1H}{1H + 2H + 4H + 5H + 8H}$$

25. En las DHS se considera que no son aptos los hogares con los códigos 3H, 6H, 7H y 9H, y por lo tanto no se incluyen en el denominador[1]. El código 9H normalmente es incluido por los supervisores en uno de los códigos específicos, y por lo tanto es prácticamente inexistente. Los pocos casos de hogares con el código 9H pueden clasificarse como no aptos. Conviene señalar que debido a la falta de un buen sistema de direcciones en muchos países, la operación de confección de listas de las DHS identifica primero las viviendas en función de los nombres de los hogares ocupados, nombres que luego se utilizan en lugar de las direcciones. Cuando un nuevo hogar se traslada de vivienda entre el período de enumeración y la entrevista, ello no significa que no se haya producido una sustitución de la unidad de muestreo, ya que la vivienda es la verdadera base para la selección. Asimismo, cuando un hogar abandona la vivienda después de la enumeración y otro viene a sustituirlo no constituye una situación de falta de respuesta.

26. Los códigos de respuesta relativos a los individuos son los siguientes:

[1] Se considera que los hogares con el código 3H (Todo el hogar ausente durante un período largo) no son aptos para las DHS, por lo que este método de calcular la tasa de respuesta de los hogares es comparable con el método de tasas de respuesta RR5 establecido por las normas de 2000 de la American Association for Public Opinion Research (AAPOR). Este método exagera ligeramente la verdadera tasa de respuesta en cuanto que un pequeño número de los hogares codificados como 3H son aptos pero no se incluyen en el cálculo.

1I Entrevista terminada
2I Fuera de casa
3I Aplazada
4I Rechazada
5I Terminada en parte
6I Incapacitados
7I Otros

La tasa de respuesta de los individuos es la siguiente:

$$R_I = \frac{1I}{1I + 2I + 3I + 4I + 5I + 6I + 7I}$$

27. Las tasas de respuesta no ponderadas de hogares e individuos se calculan independientemente en cada estrato o campo de información y se presentan en el informe sobre el país de la DHS junto con las tasas globales de respuesta. La tasa global de respuesta es el producto de las tasas de respuesta de hogares e individuos. En las encuestas demográficas y de salud, las tasas de respuesta son semejantes en los diferentes campos. Como la muestra suele ser aproximadamente autoponderada dentro de cada campo, las tasas de respuesta ponderadas y no ponderadas del conjunto del país son muy semejantes. Conviene señalar que los citados códigos de respuesta se han utilizado en la mayor parte de las encuestas demográficas y de salud pero que se han modificado en algunas de ellas para tener en cuenta la situación de un país determinado.

G. Comparación de las tasas de falta de respuesta

28. Utilizando las fórmulas antes mencionadas, se calcularon las tasas de respuesta tanto de los hogares como de las mujeres en 66 encuestas realizadas en 44 países entre 1990 y 2000. En el anexo pueden verse los resultados correspondientes a las siguientes regiones mundiales: Asia, Eurasia, América Latina, Cercano Oriente y África subsahariana.

29. Los datos revelan que en esas encuestas las tasas de respuesta de los hogares oscilaron entre el 87,9% y el 99,5%, con un promedio del 97,5%, lo que indica que la inmensa mayoría de los hogares identificados en las muestras de las DHS realizaron las entrevistas. En las mismas encuestas, la tasa de respuesta de las mujeres oscilaba entre el 86,5% y el 99,3%, con un promedio del 95%. Así pues, se consiguió terminar la entrevista con casi todas las mujeres que reunían los requisitos.

30. Con excepción de América Latina, donde la tasa de respuesta global de los hogares fue del 95%, las demás regiones tuvieron una tasa media de aproximadamente el 98%. En cuanto a los hogares, la tasa media de respuesta de la mujer fue más baja en América Latina que en otras regiones incluidas en el programa DHS: el 92% frente al 97%. Dentro de cada región, las tasas de respuesta tanto de los hogares como de las mujeres presentaban pocas diferencias en los distintos países, y el coeficiente de variación osciló ente el 0,4% y el 3,7%.

31. La tasa media de respuesta de los hogares se mantuvo en un nivel elevado, el 97%, durante las tres últimas fases del programa DHS (DH II, DHS III y MEASURE-DHS+)[2], mientras que la tasa media de respuesta de las mujeres aumentó ligeramente, del 94% al 96%, a lo largo del tiempo.

32. Las altas tasas de respuesta de los hogares e individuos en las encuestas DHS son resultado de la rigurosa capacitación del personal sobre el terreno y de la estrecha supervisión de las actividades. Además, en todas las encuestas se procura que la fecha de las actividades de enumeración y las de las entrevistas no estén demasiado alejadas. Así, a diferencia de las encuestas realizadas en los países desarrollados, las encuestas de hogares en los países en desarrollo normalmente cuentan con un alto nivel de cooperación por parte de los posibles

[2] MEASURE es un proyecto global de USAID del que forma parte MEASURE DHS+. MEASURE corresponde a las siglas en inglés de la Iniciativa de Seguimiento y Evaluación para Valorar y utilizar los resultados (*M*onitoring and *E*valuation to *A*ssess and *U*se *Re*sults).

informantes. A lo largo del tiempo, las tasas de respuesta media de los hogares e individuos han sido notablemente semejantes en cada región.

H. Efectos del diseño muestral observados en las DHS

33. En este apartado se presenta un breve resumen de algunos efectos del diseño y de valores del coeficiente de correlación intraclase ($\bar{\rho}$) que aparecen en las encuestas demográficas y de salud. (Véase información más detallada en Lê y Verma (1997). Kish, Groves y Krotki (1976) y Verma, Scott y O'Muircheartaigh (1980) contienen análisis semejantes de los análisis de muestreo de las encuestas mundiales sobre fecundidad (WFS).)

34. El efecto del diseño es el coeficiente entre la varianza de muestreo de cualquier estimación obtenida a partir de un diseño muestral complejo y la varianza de la misma estimación que se aplicaría con una muestra aleatoria simple o una muestra no restringida del mismo tamaño (Kish, 1965), es decir:

$$D^2(y) = \frac{Var_{complejo}(y)}{Var_{no\ restringido}(y)}$$

35. Los efectos del diseño son resultado de la estratificación, de las probabilidades desiguales de selección, de los ajustes en la ponderación de la muestra (para tener en cuenta la falta de respuesta), los ajustes de ponderación de la población (en caso de falta de cobertura y para mejorar la precisión) y la conglomeración de todos los elementos de un diseño muestral complejo.

36. El efecto del diseño estimado debido a la ponderación puede calcularse a partir de la muestra como sigue:

$$d^2(\hat{y}) = 1 + cv^2(w_j)$$

donde cv^2 es el cuadrado de la variación del coeficiente de las ponderaciones de muestreo w_j.

37. El efecto del diseño debido a la conglomeración puede calcularse como sigue:

$$D^2(\hat{y}) = 1 + (b-1)\rho$$

donde b es el tamaño medio del conglomerado y ρ es la correlación intraclase.

38. En el capítulo VI de la presente publicación puede verse un examen completo de los efectos del diseño y de los coeficientes de correlación intraclase —definiciones, componentes de efectos del diseño, uso de los efectos del diseño y coeficientes de correlación intraclase en la elaboración de encuestas por muestreo. Para comprender el efecto de un diseño muestral complejo en los errores estándar suele utilizarse la raíz cuadrada del efecto del diseño $d(\hat{y})$.

39. Como ya se ha mencionado, las DHS están basadas en muestras de hogares nacionalmente representativas con un diseño muestral probabilístico estratificado en varias etapas en el que se incluye un número bastante considerable de UPM. Normalmente se elaboran estimaciones de alcance nacional para las zonas urbanas y rurales y para regiones geográficas menores que en muchos países coinciden normalmente con regiones administrativas.

40. Lê y Verma (1997) estudiaron los errores de muestreo en 48 encuestas demográficas y de salud realizadas entre 1985 y 1993. En lo que respecta a las estimaciones nacionales globales, el efecto cuadrático medio del diseño $d(\bar{y})$, donde $\bar{y}$ muchas veces era una proporción promediada de 37 variables y 48 encuestas, era de aproximadamente 1,50, con promedios que iban desde 1,13 en Trinidad y Tabago hasta 2,07 en Nigeria. Ello significa que la conglomeración, la ponderación y otros aspectos de los diseños hacían que los errores estándar de las estimaciones fueran, por término medio, 1,5 veces mayores (o 2,25 veces

mayores en el caso de las varianzas de las estimaciones) que en una muestra sin restricciones del mismo tamaño.

41. Se utilizaron tamaños de conglomerados semejantes en las zonas urbanas y rurales de la mayor parte de los países (tamaño medio del conglomerado de 24 en las zonas urbanas y de 30 en las zonas rurales). En consecuencia, la diferencia del promedio entre los valores urbanos y rurales $d(\bar{y})$ fue pequeña: 1,4 en las zonas urbanas y 1,5 en las rurales. Esta pauta se dio también en los valores $d(\bar{y})$ correspondientes a las regiones geográficas. Dentro de cada país, los valores $d(\bar{y})$ fueron muy semejantes en las diferentes regiones, siendo sólo marginalmente menores que el correspondiente al total del país $d(\bar{y})$, debido también en este caso a la utilización del mismo diseño en todas las regiones del país. Por el contrario, los valores $d(\bar{y})$ fueron notablemente menores que los nacionales en los subgrupos definidos en función de las características demográficas y socioeconómicas de los informantes individuales. Como estos subgrupos corresponden a distintas UPM, los tamaños de los conglomerados pertinentes (b_d) fueron menores que los correspondientes a la muestra total (b), por lo que los efectos del diseño del subgrupo eran en general menores. Por ejemplo, en la DHS de Túnez los valores de $d(\bar{y})$ en la variable "tamaño de familia ideal" eran 1,56 y 1,70 en los subgrupos de mujeres trabajadoras y mujeres no trabajadoras, respectivamente, en comparación con un valor de $d(\bar{y})$ de 1,79 para toda la muestra.

42. Las tasas de muestreo diferenciales de las zonas urbanas y rurales o de las regiones geográficas en las encuestas demográficas y de salud requerían la ponderación de los datos de la muestra. La ponderación era también necesaria para compensar la falta de respuesta diferencial y otros inconvenientes relacionados con la muestra. Esta ponderación tendía a sobrevalorar los errores de muestreo. El efecto de diseño debido a las ponderaciones variables se calculó en las DHS para las estimaciones basadas en las muestras totales. En las primeras encuestas de 1985-1990, el promedio $d(\bar{y})$ debido a la ponderación era de 1,08 (que representaba un aumento del 17% de la varianza). Subió a 1,15% (que significaría un aumento del 32% de la varianza) en las últimas encuestas demográficas y de salud de 1990-1993, que se alejaban más de la costumbre de utilizar muestras mspi (de probabilidad igual) dentro de las zonas urbanas y rurales, con el fin de hacer posibles las estimaciones regionales.

43. Como puede verse en el cuadro XXII.1, los valores de $d(\bar{y})$ para el promedio total de la muestra en los diferentes países presenta diferencias notables en función de la variable: los valores $d(\bar{y})$ oscilan desde un mínimo de aproximadamente 1,1 ó 1,2 en el caso de las variables de la mortalidad infantil hasta un máximo de 2,5 en una estimación de si el parto tuvo lugar en condiciones médicas adecuadas. Ello refleja la mayor correlación dentro de los conglomerados geográficos de la atención médica disponible. Al examinar la variabilidad de estos valores $d(\bar{y})$ conviene señalar las diferencias en las bases muestrales de las diferentes partes del cuadro. Por ejemplo, el conjunto superior de estimaciones está basado en todas las mujeres de 15 a 49 años; el segundo, únicamente en las mujeres de ese grupo de edad que estaban casadas en el momento de la encuesta, y el siguiente, en todos los nacimientos de los cinco últimos años. Los cambios en la base de la muestra dan lugar a diferentes valores b en los efectos del diseño correspondientes a la conglomeración, y este factor contribuye a la variabilidad de los valores $d(\bar{y})$ en el cuadro XXII.1.

44. La medida de homogeneidad ρ es más útil que el efecto del diseño debido a la conglomeración para las encuestas futuras de planificación, ya que el efecto del diseño depende tanto de ρ como del tamaño del conglomerado b. El efecto del diseño de una encuesta anterior será aplicable a la nueva encuesta únicamente si estos parámetros son iguales. No obstante, debería considerarse la posibilidad de cambiar b, ya que el tamaño del conglomerado puede controlarse mediante el muestreador, cosa que no ocurre con la correlación intraclase. Si se dispone de una estimación de ρ, el efecto del cambio de b puede examinarse calculando los efectos del diseño resultantes de la aglomeración para diferentes valores de b. Así pues, ρ es el factor de mayor interés. Se calcularon las estimaciones del promedio de $\bar{\rho}$

Cuadro XXII.1
Valores medios $d(\bar{y})$ y $\hat{\rho}$ en 48 encuestas demográficas y de salud (DHS), 1984-1993

Proporción/media	$d(\bar{y})$	$\hat{\rho}$
Todas las mujeres de 15-49 años [a]		
Actualmente casadas	1,43	0,03
Número de hijos nacidos	1,35	0,02
Número de hijos en los cinco últimos años	1,44	0,03
Número de hijos vivos de menos de 5 años	1,41	0,02
Número de hijos nacidos de mujeres de 40-49 años	1,26	0,02
Mujeres actualmente casadas de 15-49 años		
Que no desean más hijos	1,32	0,02
Que desean aplazar el siguiente parto dos o más años	1,24	0,01
Que conocen un método anticonceptivo	2,01	0,14
Que conocen un método anticonceptivo moderno	2,08	0,15
Que conocen una fuente de suministro de anticonceptivos	1,94	0,12
Que utilizan actualmente un método anticonceptivo	1,50	0,05
Que utilizan actualmente un método anticonceptivo moderno	1,43	0,04
Que utilizan actualmente el dispositivo intrauterino (DIU)	1,42	0,04
Que utilizan actualmente la píldora	1,41	0,04
Que utilizan actualmente el preservativo	1,38	0,03
Que utilizan actualmente un proveedor público de suministro de anticonceptivos	1,36	0,03
Esterilizadas	1,36	0,03
Todos los nacimientos de los cinco último años		
La madre recibió atención médica durante el parto	2,54	0,22
La madre recibió el toxoide tetánico	2,02	0,12
Niños de menos de cinco años		
Tuvieron diarrea en las dos últimas semanas	1,34	0,03
De ellos: el niño recibió tratamiento con sales de rehidratación oral (SRO)	1,25	0,12
Niños de 6-35 meses		
Altura por edad, menos de 2 desviaciones estándar por debajo de la norma	1,33	0,05
Peso por edad, menos de 2 desviaciones estándar por debajo de la norma	1,29	0,04
Peso por altura, menos de 2 desviaciones estándar por debajo de la norma	1,19	0,02
Niños de 12-23 meses de edad		
Tenían ficha médica	1,33	0,15
De ellos: el niño estaba totalmente inmunizado	1,31	0,21
Niños nacidos hace 1-4 ó 5-9 años		
Tasa de mortalidad infantil 1-4 años antes de la encuesta	1,23	0,02

[a] En aproximadamente una cuarta parte de las encuestas, la muestra y, por lo tanto, todas las variables de este grupo se limitaban a las mujeres que habían contraído matrimonio.

a partir de las encuestas demográficas y de salud, y los resultados pueden verse también en el cuadro XXII.1. Como puede deducirse del cuadro, los valores $\bar{\rho}$ varían considerablemente, desde un mínimo de 0,01 a un máximo de 0,22. Como era de prever, las estimaciones que dependen de la disponibilidad de servicios de salud local suelen tener valores $\bar{\rho}$ elevados.

45. Una comprobación importante en los análisis de error muestral en el programa DHS es que las estimaciones de $\bar{\rho}$ para una estimación dada son bastante transferibles de un país a otro, siempre que los diseños de la muestra sean comparables. Por ello, al diseñar una nueva encuesta en un país pueden emplearse datos empíricos sobre los errores de muestreo de una encuesta semejante en un país vecino, si se considera necesario y se hace lo posible por comprobar la comparabilidad entre una y otra.

I. Realización de la encuesta[3]

[3] Gran parte del material de los apartados relativos a la organización de la encuesta y las características de las encuestas demográficas y de salud se ha tomado del borrador del *Manual de Organización de Encuestas DHS,* que ha preparado uno de los autores de este capítulo.

46. Si bien se presta gran atención al muestreo científico y al cálculo de los errores de muestreo, no debe olvidarse que hay múltiples factores de error en las encuestas. Los errores relacionados con la variabilidad de muestreo pueden cuantificarse normalmente, mientras que otros errores no se cuantifican con tanta facilidad. De todas formas, los errores no muestrales suelen ser mayores que los errores muestrales. Así ocurre sobre todo cuando no se presta la debida atención a la capacitación y al reclutamiento del personal encargado de las actividades sobre el terreno y del procesamiento de datos. Por ello, el control del error no muestral es un objetivo importante en toda encuesta demográfica y de salud.

47. Con respecto a la puesta en práctica, muchas encuestas demográficas y de salud se realizan en países donde es difícil contratar personal sobre el terreno altamente cualificado y donde las actividades sobre el terreno plantean importantes problemas de transporte, alojamiento, higiene, suministro de alimentos, etcétera. La necesidad del personal sobre el terreno de viajar por el país plantea también problemas de seguridad y supervisión. Estas y otras son las principales razones por la que el programa DHS presta gran atención a la capacitación del personal sobre el terreno y a la supervisión sobre el terreno y en las oficinas. No obstante, aun cuando se haya hecho hincapié en la supervisión, ha habido casos en que los sistemas no se aplicaron debidamente y la calidad de los datos fue deficiente. A continuación se describen los pasos habituales para la realización de una encuesta demográfica y de salud, con especial atención a la necesidad de preparación detallada, capacitación profunda y supervisión.

48. Otro aspecto importante de las encuestas es hasta qué punto los datos se pueden poner a disposición en forma oportuna y accesible para los responsables de la toma de decisiones, los directores de programas y los analistas. Hay demasiadas encuestas, sobre todo en los países en desarrollo, que no se han analizado o divulgado debidamente. El programa DHS trata de garantizar que todas las encuestas se analicen de forma oportuna, que los resultados se publiquen y divulguen y que se disponga de datos para ulteriores investigaciones. A continuación se describe el proceso que debe seguirse con ese fin.

J. Preparación y traducción de los documentos de la encuesta

49. En cada país participante, los documentos de la encuesta suelen consistir en un cuestionario de hogares, cuestionarios individuales para las mujeres y/o los hombres y los correspondientes manuales. En los cuestionarios se incluyen cuestiones básicas sobre el programa DHS, adaptaciones específicas a cada país y módulos optativos. El personal del DHS colabora con sus homólogos locales en la adaptación de los cuestionarios, teniendo en cuenta las necesidades del país. Los cuestionarios inspirados en el modelo DHS son largos, por lo que las adiciones deben examinarse atentamente, dada la habitual gran longitud de los instrumentos. La calidad de los datos sufrirá probablemente si los cuestionarios resultan poco manejables y necesitan demasiado tiempo. Los manuales básicos del entrevistador y el supervisor se adaptan en cada país para tener en cuenta el contenido específico del cuestionario.

50. El DHS tiene por norma traducir los cuestionarios e imprimirlos en los principales idiomas locales para garantizar que las entrevistas se realicen en el idioma de los informantes. Todo grupo lingüístico que constituya el 10% o más de la muestra debería tener su propio cuestionario traducido. La necesidad de traducción sobre el terreno por el entrevistador o por alguna otra persona muchas veces no se puede evitar completamente, ya que quizá no haya una versión lingüística adecuada de un cuestionario para algunos informantes incluidos en la muestra. La necesidad de traducciones improvisadas debería reducirse al mínimo.

51. La traducción no es una tarea fácil y requiere un gran conocimiento lingüístico y una comprensión de los términos y expresiones propios de las encuestas demográficas y de salud. Pocas veces se encuentran ambos conocimientos en una única persona, sobre todo cuando deben utilizarse varias lenguas en el mismo país.

52. El planteamiento de la traducción en el DHS significa que una persona debe traducir el cuestionario DHS al idioma local pertinente, utilizando la versión en español, francés o inglés del cuestionario básico. En el caso de que se haya traducido ya una encuesta anterior del DHS o semejante, dicha traducción debería tenerse en cuenta. Normalmente, si se van a hacer las mismas preguntas, sería de prever que la traducción fuera también la misma, excepto en los casos en que se considere que la traducción anterior era deficiente.

53. El cuestionario traducido es objeto de traducción inversa al idioma original por un traductor independiente. Es importante que esta segunda traducción se realice sin tener a la vista el cuestionario original, para garantizar la plena independencia de las dos versiones. El paso siguiente es hacer que los dos traductores y el personal superior de la encuesta se reúnan para estudiar el original y la traducción inversa a fin de resolver las posibles discrepancias. Este es un paso importante, sobre todo en el caso de idiomas que no se escriben habitualmente, ya que su traducción no es un proceso sencillo.

54. El proceso debería dar lugar a cuestionarios que puedan ser comprendidos fácilmente por los informantes entrevistados en su propio idioma. No obstante, las traducciones deben comprobarse también sobre el terreno antes de adoptarlas para la encuesta. No hace falta realizar un gran número de entrevistas sobre el terreno, pero deberían llevarse a cabo entre tres y cinco en cada idioma antes de dar por finalizadas las traducciones. Es importante recordar que el objetivo de la traducción es garantizar que se haga la misma pregunta a todos los informantes. No obstante, ello no significa que la traducción deba ser literal. Una buena traducción transmitirá el mismo significado, aunque quizá no sea una traducción palabra por palabra. Las encuestas demográficas y de salud muchas veces se repiten en los países, aunque los cuestionarios de las diferentes rondas quizá tengan un contenido algo distinto. Por ello, pueden utilizarse también las traducciones anteriores de la mayor parte de las preguntas y la experiencia obtenida durante los controles previos y la actividad sobre el terreno anteriores.

55. La documentación de la encuesta y los manuales para los entrevistadores y los supervisores deberían traducirse a un idioma que puedan entender todos los trabajadores sobre el terreno, si no es posible utilizar las versiones en español, francés o inglés.

K. Prueba preliminar

56. Una prueba preliminar es un medio fundamental de controlar la calidad de las traducciones, las pautas de salto en los cuestionarios, los manuales de entrevistadores y supervisores y otros procedimientos de la encuesta. Es también un mecanismo a través del cual el personal superior puede conseguir mayor experiencia en la capacitación del personal sobre el terreno antes del curso principal de capacitación. El director del DHS en el país normalmente participa en las entrevistas de la prueba preliminar.

57. Para realizar las comprobaciones previas se capacita a un pequeño número de trabajadores sobre el terreno, normalmente durante unas dos semanas. Se ofrece capacitación a través del personal local, con asistencia del director del DHS en el país. Es práctica habitual del DHS capacitar a los futuros supervisores como entrevistadores para las comprobaciones previas, que luego participarán en actividades de capacitación de entrevistadores como supervisores. De esa manera se garantiza que tengan una capacitación amplia, que su función se haya podido establecer durante la formación de los entrevistadores y que haya personal suficiente para corregir y orientar las sesiones de práctica y las pruebas que tengan lugar durante la formación de los entrevistadores.

58. Normalmente, la prueba preliminar abarca 100-200 hogares y las entrevistas duran aproximadamente una semana. Estas entrevistas se realizan en zonas urbanas y rurales que no se hayan seleccionado para la encuesta principal a fin de evitar la contaminación de los resultados de la encuesta. La experiencia acumulada en el DHS con este tipo de encuesta es muy considerable, por lo que la prueba preliminar puede ser breve y no tiene que incluir muchas zonas diferentes del país.

59. El trabajo previo sobre el terreno sigue los mismos procedimientos que se adoptarán durante la encuesta. Así, los hogares se enumeran de tal manera que los equipos puedan familiarizarse con la aplicación de los procedimientos y de sus formularios de control. El personal superior supervisará activamente todas las etapas de la prueba preliminar para poder familiarizarse con los problemas que surjan y recomendar soluciones.

60. La experiencia de esta prueba preliminar es la base para la revisión de los cuestionarios y manuales. Deben corregirse los errores e introducirse las mejoras pertinentes teniendo en cuenta las actividades observadas. Un elemento fundamental es mantener un diario de todos los problemas encontrados durante la capacitación, las prácticas y las entrevistas. Los problemas surgidos durante estas últimas se documentan mediante informes del personal de la encuesta que observan las entrevistas de la prueba preliminar y mediante una reunión informativa diaria de los entrevistadores. Es importante que todo el personal implicado en la prueba preliminar tome notas de lo que observa.

61. Hay que evitar que en las posibles revisiones posteriores a la prueba preliminar se introduzcan nuevos errores. De hecho, si es necesaria una revisión amplia de los cuestionarios, conviene realizar algunas entrevistas sobre el terreno con el nuevo instrumento para garantizar que las revisiones se han hecho correctamente y no se han provocado nuevos problemas.

L. Contratación del personal sobre el terreno

62. La calidad de una encuesta sobre el terreno depende en gran medida de la calidad del personal sobre el terreno. Por ello, para este trabajo debe contratarse a las personas mejor calificadas. En los países en desarrollo hay pocas organizaciones que cuenten con personal permanente de entrevistadores y supervisores. Incluso si los tienen, los entrevistadores suelen ser predominantemente hombres. En las entrevistas para las encuestas demográficas y de salud también se necesitan mujeres entrevistadoras, a no ser que se trate de una encuesta para hombres. Por ello, las DHS generalmente se llevan a cabo con personal contratado especialmente para ese fin. Como la fase de recopilación de datos o de actividades sobre el terreno suele durar entre tres y seis meses, los candidatos suelen ser personas que no tienen empleo en ese momento y que están dispuestas a pasar varios meses lejos de casa. En algunos países donde las entrevistas tienen mayor contenido sanitario, se ha enviado como entrevistadores y supervisores a personal médico del Ministerio de la Salud.

63. En la contratación debe tenerse en cuenta el número de personas necesarias para hablar en cada uno de los idiomas en que se realizará la encuesta. El número de personas que recibirán capacitación deberá ser al menos entre un 10% y un 15% superior al necesario para las actividades sobre el terreno, para afrontar los efectos del desgaste y del despido de candidatos que resulten inadecuados. La contratación se basa sobre todo en una prueba objetiva de las capacidades de los candidatos. Los candidatos deberán tener un aspecto presentable y ser capaces de recorrer largas distancias y establecer buenas relaciones con las personas a las que vayan a entrevistar. Asimismo, deben tener espíritu de equipo. El criterio para la selección de candidatos no será en ningún caso su relación con el personal de la encuesta, el favoritismo ni ninguna otra práctica inaceptable de contratación.

64. Los cargos de supervisor y de editor sobre el terreno requieren personas que puedan ser jefes de equipo. Deben dar muestras de autoconfianza, fuerte motivación y excelente

espíritu de equipo. Todas estas características son también deseables en los candidatos para entrevistadores. No obstante, las principales características de un buen entrevistador son la capacidad de hacer preguntas de forma fluida y natural, conseguir que el informante se sienta cómodo y registrar correctamente las respuestas recibidas.

M. Capacitación de los entrevistadores

65. La capacitación de los entrevistadores es semejante a la organizada para la prueba preliminar, con la excepción de que generalmente dura entre tres y cuatro semanas, debido en parte al mayor número de participantes. Los candidatos a entrevistadores realizan al menos entre 5 y 10 entrevistas prácticas sobre el terreno durante la capacitación. Ésta es ofrecida por personal local, que cuenta con la ayuda de trabajadores capacitados para la prueba preliminar y con el director del programa DHS en el país.

66. La selección final de los entrevistadores se basa en los resultados que obtengan en una serie de pruebas escritas, en cómo se comportan cuando hacen prácticas de entrevistas en la oficina y en la calidad de sus entrevistas durante la prueba preliminar. Es de suma importancia que los criterios de selección sean objetivos. En muchos lugares el personal de capacitación se ve sometido a fuerte presión de otros individuos para que den los puestos a las personas recomendadas por ellos. Hay que repetir que la única manera de seleccionar al personal es la comprobación de que cumplen los requisitos necesarios para el empleo y que han conseguido una buena y objetiva calificación durante la capacitación. De hecho, la organización de pruebas escritas objetivas durante la capacitación puede ayudar al personal de la encuesta a documentar las razones por las que no se pudo aceptar a determinados candidatos.

N. Trabajo sobre el terreno

67. El programa DHS concibe el trabajo sobre el terreno como una labor de equipo. Las razones para trabajar en equipos son muchas, pero la principal es que así se puede conseguir un mayor nivel de supervisión. Otra razón es la necesidad de un medio especial de transporte para la mayor parte de los entrevistadores. En muchos países, un motivo adicional importante es la necesidad de salvaguardar el bienestar del personal sobre el terreno.

68. Los equipos están formados generalmente por un supervisor (jefe de equipo), una editora sobre el terreno, un técnico de salud y entre tres o cuatro entrevistadoras. Si se incluye también una entrevista de hombres, suele incorporarse a un entrevistador. En casi todos los países se asigna a cada equipo un vehículo y un conductor. Algunas veces el tamaño del equipo está limitado por la capacidad de transporte de los vehículos utilizados.

69. El supervisor es el encargado general del equipo y de la organización y supervisión cotidiana de su labor. El editor sobre el terreno se ocupa principalmente de comprobar la calidad de las entrevistas. En la práctica, el supervisor y el editor deberán compartir mutuamente sus responsabilidades con el fin de crear y mantener un buen equipo de entrevistadores.

70. Las principales consideraciones al determinar el número de equipos son el número de UPM, el tamaño de los conglomerados y la duración prevista del trabajo sobre el terreno. No obstante, otras consideraciones importantes son el número de vehículos disponibles, el número de entrevistadores y supervisores que puedan reclutarse y el número de idiomas hablados en el país. El trabajo sobre el terreno debería durar entre tres y seis meses. A veces la duración puede ser menor. No obstante, para conseguir datos de buena calidad, el número de entrevistadores se mantiene relativamente bajo, debido a problemas de capacitación, disponibilidad de buenos candidatos, etcétera. Ello limita a su vez el número de equipos que pueden utilizarse y determina la duración del trabajo sobre el terreno.

71. Si es posible, todos los equipos comienzan el trabajo sobre el terreno en la misma ubicación geográfica general (por ejemplo, la misma provincia), con el fin de hacer posible la supervisión de todos los equipos por el personal superior de la encuesta en el momento en que más necesaria es la supervisión. Si los equipos están dispersos por todo el país desde el comienzo, es muy difícil visitar a todos los equipos inmediatamente.

72. Al asignar las tareas correspondientes a cada equipo se deben tener en cuenta las lenguas que se hablan y otros requisitos, así como la necesidad de reducir en lo posible el tiempo de viaje por equipo. En términos generales, los equipos trabajan seis días por semana y están lejos de casa varias semanas o meses seguidos.

73. Si una entrevista no se termina en la primera visita, se realizan nuevos intentos con el hogar o informante incluido en la muestra, hasta tres veces y en tres días diferentes, antes de clasificar el caso como de falta de respuesta. Los contactos posteriores se programan a las horas en que la probabilidad de que el informante se encuentre en casa es mayor. Cuando la mayoría de los miembros del equipo han terminado su trabajo pero quedan todavía una o dos segundas visitas para otro día, no es raro que el equipo se desplace a un nuevo conglomerado y deje detrás a un entrevistador para "atar los cabos sueltos". Ello es posible cuando el nuevo conglomerado no se encuentra demasiado distante y el vehículo del equipo puede recoger al entrevistador rezagado. En otras circunstancias, todo el equipo continúa en el lugar hasta que se haya terminado el trabajo en ese conglomerado. Como ya se ha mencionado, no hay ninguna sustitución en el caso de los hogares o individuos que se resisten a ser entrevistados o que por cualquier razón se incluyen entre los casos de falta de respuesta.

74. Los equipos deben contar con suficientes cuestionarios y materiales para garantizar que el trabajo se lleva a pleno ritmo en todo momento. Los cuestionarios terminados deben empaquetarse, protegerse frente a la acción de los elementos y salvaguardarse hasta que puedan enviarse a la oficina local, normalmente por medio de los supervisores itinerantes sobre el terreno que visitan periódicamente a cada equipo.

75. La fuerte insistencia en la supervisión es una de las características de las encuestas demográficas y de salud. La experiencia revela que sin una supervisión continua la calidad de los datos sufre considerablemente. Por ello se utilizan varios niveles de supervisión. El supervisor del equipo y el editor sobre el terreno deben observar a los entrevistadores de vez en cuando y comprobar cada cuestionario exhaustivamente para determinar si es completo y exacto. Cuando se encuentran problemas de consideración, los entrevistadores deben volver a ponerse en contacto con la persona entrevistada para obtener la información correcta. Además, el supervisor normalmente se encarga de volver a entrevistar una submuestra de aproximadamente un 10% de los hogares seleccionados para cerciorarse de que la entrevista inicial se llevó a cabo y de que se identificaron correctamente todas las mujeres que reunían los debidos requisitos.

76. El director de la encuesta y el personal del DHS realizan también tareas de supervisión durante el trabajo sobre el terreno. Se visita a los equipos periódicamente para comprobar el trabajo de los entrevistadores, los editores y los supervisores. Durante esa comprobación se examinan al menos uno o dos cuestionarios de cada entrevistador que se hayan revisado por el editor sobre el terreno. Eso sirve además para descubrir al tiempo los posibles errores del entrevistador o del editor. Las visitas de supervisión sobre el terreno son sumamente importantes. Si la labor de los supervisores y editores deja bastante que desear, merma la calidad de la labor de los entrevistadores. En tal caso deben adoptarse medidas correctivas lo antes posible. Las visitas sobre el terreno son el principal mecanismo de rectificación. Un instrumento útil durante esas visitas son los "cuadros de calidad de los datos" que se aplican periódicamente durante el trabajo sobre el terreno para descubrir problemas específicos y las dificultades existentes con determinados equipos y entrevistadores. Dichos cuadros contienen información sobre la edad de los informantes y de los niños pequeños,

que puede utilizarse para comprobar que los informantes fueron seleccionados debidamente por los entrevistadores, y ofrecen también información sobre la mortalidad infantil para calibrar el nivel de omisiones de niños fallecidos. Para comprobar la productividad de cada equipo y entrevistador y ver si los hogares y/o los informantes se omiten deliberadamente de la encuesta se incluyen también las tasas de respuesta de hogares e individuos. Los problemas encontrados durante el examen de estos cuadros de calidad de los datos se comunican a quienes se encuentran sobre el terreno, para que no vuelvan a repetirse (véase también el apartado O, *infra*).

77. La enumeración de hogares que forma parte de la etapa de muestreo no se describe en esta sección referente al trabajo sobre el terreno. Se trata de una operación distinta que tiene lugar entre dos y tres meses antes de las actividades sobre el terreno y es llevada a cabo por personal especializado, tal como se describe en la sección E. La separación entre las actividades de enumeración y el trabajo sobre el terreno propiamente dicho garantiza que la enumeración esté bien supervisada y que los hogares puedan ser incluidos en la muestra por personal calificado en la oficina antes del verdadero trabajo sobre el terreno. La selección de la muestra en la oficina ayuda a evitar los sesgos que pueden ocurrir si los hogares son seleccionados por el personal sobre el terreno, sobre todo cuando el encargado de la lista y de la muestra son la misma persona.

O. Procesamiento de datos

78. En las encuestas demográficas y de salud, el procesamiento de datos generalmente comienza entre una y dos semanas después del inicio del trabajo sobre el terreno y normalmente se termina antes de transcurrido un mes desde la conclusión de dicho trabajo. El personal encargado de la entrada de datos recibe capacitación sobre los cuestionarios, para lo cual asiste a parte del curso de capacitación de entrevistadores o a sesiones especiales de dos o tres días. El coordinador del procesamiento de datos normalmente asiste a todo el curso de capacitación de entrevistadores.

79. La entrada de datos tiene lugar en una sala independiente, donde el personal puede trabajar sin ser molestado y los cuestionarios se encuentran a buen recaudo. Dicha sala debería estar próxima al espacio donde se almacenan los cuestionarios terminados. Todos los cuestionarios se manipulan varias veces durante la entrada y edición de los datos, de manera que la cercanía entre las instalaciones de almacenamiento y de entrada de datos puede reducir considerablemente el volumen de trabajo y la presión. El personal encargado de la entrada de datos no trabaja más de seis horas diarias, debido al carácter tan mecánico de la actividad. Según el número de computadoras disponibles para la operación de entrada de datos, quizá sea necesario organizar más de un turno de personal para acabar la entrada y edición de datos poco después de finalizar el trabajo sobre el terreno. Los dobles turnos se evitan cuando es posible, ya que la presencia de múltiples supervisores y editores en la oficina puede dar lugar a resultados contradictorios.

80. El DHS tiene por norma introducir los datos de todos los cuestionarios dos veces ("doble entrada"), comparar los resultados y resolver las contradicciones. Esta verificación del 100% reduce considerablemente el volumen de edición secundaria necesario para resolver las contradicciones y da lugar a un conjunto de datos más claro y preciso. La doble entrada de datos es llevada a cabo por dos personas diferentes, para garantizar resultados óptimos. Durante la entrada de datos, en cada cuestionario se comprueban la coherencia, las pautas de salto y los intervalos.

81. Un aspecto de la entrada y edición de datos está directamente relacionado con el control de la calidad. El programa DHS prepara periódicamente un conjunto seleccionado de cuadros durante la entrada y edición de datos, con el fin de comprobar los problemas que no puedan identificarse fácilmente durante el proceso manual de entrada y edición de

los datos en los cuestionarios individuales. La finalidad de estos "cuadros de comprobación sobre el terreno" es descubrir, por ejemplo, si los entrevistadores están manipulando las edades de los informantes o de sus hijos como manera de aligerar su carga de trabajo, presentan cifras demasiado bajas sobre la mortalidad infantil o si registran correctamente la edad en el momento del fallecimiento. Estos cuadros se aplican una vez que se cuenta con un número suficiente de cuestionarios —alrededor de 300—, y posteriormente cada dos semanas, para que el entrevistador o el equipo de entrevistadores puedan identificar las pautas de respuesta o las características de los informantes que deben ser objeto de especial atención, y son examinados por el personal del organismo encargado de la encuesta y del DHS. Los problemas observados se comunican a los equipos pertinentes para que adopten medidas correctivas.

82. Las tabulaciones básicas preparadas para cada país son las que se diseñaron tomando como base los datos recopilados en el cuestionario básico. Las tabulaciones de datos resultantes de las preguntas que se incorporaron al cuestionario básico se formulan en colaboración con las personas/instituciones que solicitaron los cuadros adicionales. Esta labor debe realizarse en los primeros momentos para garantizar que el proceso de tabulación se lleva a cabo sin problemas. Todas las tabulaciones son comprobadas exhaustivamente, tanto por el personal del DHS como por sus homólogos de los países.

83. Debido a la complejidad de los programas de entrada, edición, imputación y tabulación de datos, su preparación se encarga al personal encargado del procesamiento de datos del DHS, que visita el país para instalar los programas e iniciar el proceso. Normalmente, el especialista en procesamiento de datos regresa al final del proceso para ayudar a examinar el conjunto final de datos, volver a codificar algunas variables, imputar las fechas omitidas, determinar los factores de ponderación de la muestra y aplicar el conjunto de cuadros previamente designado para los informes preliminares y finales. Al tabular los datos, en los informes se presentan los números ponderados y los sin ponderar, aunque en los cálculos siempre se utilizan las ponderaciones muestrales finales.

P. Análisis y redacción del informe

84. La base para el análisis se fija en las tabulaciones del modelo de DHS, modificada por el director del DHS en el país y el personal del país anfitrión para tener en cuenta los cuestionarios utilizados. Estas tabulaciones se complementan con cuadros específicos de cada país, en los que se presentan datos adicionales recopilados en los países. El análisis da lugar a un informe completo sobre los datos de la encuesta.

85. Se prepara también un pequeño informe sobre las conclusiones principales, con el fin de dar la máxima difusión posible de los datos. Éste se prepara inmediatamente después o en paralelo con el informe de la encuesta principal, y puede distribuirse cuando se celebre el seminario nacional (véase el apartado Q sobre la divulgación, *infra*).

86. El DHS ayuda además a los países a realizar "ulteriores análisis" de mayor profundidad sobre los datos de la encuesta. Estos análisis suelen dar lugar a un documento de investigación de entre 30 y 60 páginas en el que se consideran temas de especial interés para el país o los organismos que financian la encuesta. Ese documento puede consistir también en una serie de tabulaciones especiales con breves exposiciones analíticas que permitan al país responder a las cuestiones relacionadas con las políticas y otros temas.

Q. Divulgación

87. La divulgación de los resultados de la encuesta de modo que lleguen a todos los grupos de interesados es un objetivo fundamental del programa. Los informes se distribuyen por todo el país y se ponen también a disposición de los organismos cooperadores y otras instituciones que trabajan en los respectivos países. Pueden también consultarse y descargarse del sitio del DHS en la web. Asimismo, se preparan gráficos en forma de mural y

de libro, calendarios, carteles y otros materiales, en paralelo con el seminario nacional, para conseguir una mayor divulgación de los resultados de la encuesta.

88. Además, se celebra un seminario nacional para presentar las principales conclusiones a los responsables de la formulación de políticas, directores de programas, investigadores y representantes de organizaciones donantes. El seminario suele tener eco en los medios de comunicación, lo que permite una mayor utilización de estos datos en las políticas y programas. Algunos países organizan seminarios regionales para dar a conocer los resultados y que se utilicen en esferas distintas de las políticas y en programas de alcance nacional.

89. Todos los datos de las encuestas del DHS se introducen en el archivo de datos del programa. Casi todos los países que participan en él han autorizado el uso de sus datos por investigadores responsables de todo el mundo. El equipo del archivo de datos supervisa las solicitudes de información y facilita datos y documentación a quienes están autorizados a utilizarlos. Ahora los datos pueden obtenerse gratuitamente a través de Internet, después del correspondiente registro y autorización de cada uno de los usuarios. Al final de 2002, ORC Macro había registrado más de 80.000 accesos a los archivos y subarchivos de datos de DHS. La dirección del sitio web es www.measuredhs.com. Puede obtenerse también más información sobre el programa DHS en ese mismo sitio.

R. Utilización de los datos del DHS

90. Los datos del DHS suelen utilizarse para supervisar y evaluar los programas de salud materna e infantil y de población en los países participantes. La disponibilidad de información correspondiente a encuestas repetidas ofrece a los países datos sobre la tendencia, necesarios para calibrar el progreso. Algunas veces los datos se utilizan para programas de intervención inmediata, relacionados, por ejemplo, con el suministro de complementos de hierro en los lugares donde la anemia hace estragos. Más frecuentemente, se han utilizado para orientar las políticas y modificar los objetivos de los programas de intervención y para la planificación de las actividades de salud y población a largo plazo. Los datos del DHS han sido también un agente catalizador de los programas de planificación familiar en el África subsahariana y en otros lugares, ya que han podido demostrar que el cambio es posible y real incluso en algunos de los países más pobres.

S. Fortalecimiento de la capacidad

91. Uno de los objetivos del programa DHS es aumentar la capacidad de los países participantes para recopilar y analizar datos mediante encuestas nacionales de hogares en gran escala. El mecanismo principal para ello es el establecimiento de una documentación básica con los avances más recientes, como cuestionarios y manuales; la elaboración de programas informáticos que faciliten el procesamiento de la encuesta en el contexto de los países en desarrollo, y la capacitación sobre el terreno de los homólogos locales durante todas las fases de las encuestas en los países.

92. Una contribución importante al fortalecimiento de la capacidad es la elaboración de nuevos programas informáticos. Inicialmente, el DHS creó el Sistema integrado de análisis de encuestas (ISSA), para el procesamiento de las encuestas. Este programa contribuyó en forma decisiva a la disponibilidad temprana de archivos de datos depurados y de informes. Para adaptarse a las innovaciones de los nuevos equipos y programas informáticos, el DHS ha establecido un nuevo programa de procesamiento de datos de encuestas conocido con el nombre de Sistema de procesamiento de censos y encuestas (Census and Survey Processing, CSPro), en colaboración con la Oficina del Censo de los Estados Unidos y una empresa de programas informáticos. Es de prever que este programa se utilice ampliamente y suplante a otros muchos utilizados por diferentes instituciones para el procesamiento de

encuestas en gran escala. La Oficina del Censo de los Estados Unidos está respaldando ya amplios programas de capacitación sobre el uso del sistema y se prevé que se convierta en el procedimiento estándar en la mayor parte de los países. Ello contribuirá enormemente a los esfuerzos de fortalecimiento de la capacidad.

93. El programa DHS ha ofrecido siempre capacitación continua y comentarios a los homólogos locales por medio de documentación básica detallada para la realización de encuestas, visitas periódicas de asistencia técnica (10-14 por país) y colaboración en los preparativos de los informes de la encuesta. La documentación básica incluye manuales sobre las fases importantes de ejecución de la encuesta. Estos tres mecanismos continúan siendo los vehículos principales para el fortalecimiento de la capacidad de los países participantes.

T. Enseñanzas aprendidas

94. Son muchas y valiosas las enseñanzas aprendidas durante la ejecución del DHS y sus predecesores, entre ellas las siguientes:

- Los marcos muestrales necesitan en muchos países un costoso proceso de actualización sobre el terreno para que puedan utilizarse en encuestas que traten de recopilar datos de alta calidad. Las listas de hogares muchas veces no están actualizadas o son inexistentes. El control de calidad exige que la selección de los hogares se realice en la oficina en vez de confiarla al personal sobre el terreno, con lo que se garantizaría que todos los hogares tengan una probabilidad de selección conocida. La selección de hogares en la oficina elimina problemas provocados por la tendencia de los entrevistadores a visitar las casas más accesibles y dejar de lado las más remotas. La selección a partir de una lista de hogares en la oficina permite una selección no sesgada de los hogares y una fácil supervisión de la selección sobre el terreno.
- La actualización de la muestra, cuando se realiza en la penúltima etapa de muestreo, debe estar estrechamente supervisada con el fin de hacer posible una enumeración total de los hogares. Se ha observado también en varias encuestas que los encargados de las enumeraciones de hogares pueden tener la tentación de omitir las viviendas más remotas o que se encuentran en zonas difíciles o peligrosas. Sin una buena supervisión, las listas preparadas por los encargados de visitar los hogares pueden estar sesgadas.
- Las tasas de respuesta son en general muy buenas, tanto por lo que se refiere a los hogares como a los informantes individuales (véase el apartado F).
- Deben calcularse los errores de muestreo y los efectos del diseño para obtener un conjunto representativo de temas de cada encuesta a fin de evaluar la eficacia del diseño de la muestra y la precisión de las estimaciones de la encuesta.
- El tamaño de conglomerado óptimo es de 15-20 mujeres en las encuestas demográficas y de salud, en las que es preciso lograr un equilibrio entre la variedad de los elementos demográficos y de salud —algunos más conglomerados que otros, y algunos de los cuales se refieren a los niños pequeños de las mujeres de la muestra— y el costo de la recopilación de datos.
- El efecto del diseño debido a la conglomeración es una función creciente del tamaño de conglomerado b y el coeficiente de correlación intraclase ρ. Como ρ es bastante transferible a otros países con diseños muestrales comparables, ρ, b y los efectos del diseño de una encuesta pueden utilizarse para preparar una nueva encuesta comparable en otros países, como se describe en el capítulo VI.
- La capacitación de entrevistadores y supervisores sobre encuestas complejas requiere entre tres y cuatro semanas. La capacitación del DHS normalmente dura

tres semanas. No obstante, ha habido muchas ocasiones en que se ha prolongado una semana más para conseguir la debida preparación del personal sobre el terreno. La mayor parte de los problemas con las encuestas están relacionados con el personal sobre el terreno, no con los informantes. La capacitación y supervisión adecuada son los principales instrumentos para evitar esos problemas.

- Los entrevistadores y supervisores pueden causar graves problemas en una encuesta. Por ello, la supervisión y control de calidad continuos son necesarios a fin de evitar descuidos y/o la manipulación deliberada de la muestra o de la entrevista por parte de algunos entrevistadores y supervisores que intentan aligerar su carga de trabajo. Las DHS han presentado amplias pruebas de que los entrevistadores tienen tendencia a codificar mujeres y/o niños como si se encontraran fuera del intervalo de edad, para así evitar entrevistarlos. Si bien este problema no es generalizable a todo el personal sobre el terreno, existe de hecho. Lo normal es que se limite a un reducido número de los equipos de entrevistadores. La vigilancia continua durante todo el proceso es un requisito imprescindible.
- Debería aspirarse a una entrevista que, por término medio, dure no más de una hora. Esta propuesta está basada en los comentarios recibidos del personal sobre el terreno más que en una experimentación real sobre el terreno con encuestas de diferente duración. Las encuestas demográficas y de salud varían enormemente de longitud según las características de los informantes y la facilidad con que puedan recordar las fechas y acontecimientos. La duración puede ir desde sólo 10 minutos en el caso de una mujer soltera sin hijos ni actividad sexual, hasta más de una hora y media cuando se trata de mujeres con muchos hijos que no recuerdan fácilmente los episodios contemplados en la encuesta.
- Uno de los principales obstáculos en lo que respecta a la logística sobre el terreno está asociado con la disponibilidad de vehículos de transporte adecuados para los equipos de la encuesta. La adquisición y el mantenimiento de los vehículos es costoso, ya que se necesitan vehículos todo terreno para poder transportar a todo el equipo. La falta de vehículos adecuados es igualmente costosa en términos de tiempo, y repercute negativamente en la moral del equipo. Incluso cuando se dispone de vehículos adecuados, los entrevistadores y supervisores deberán recorrer a pie grandes distancias hasta llegar a algunas viviendas. Por ello, el transporte hasta el lugar general de la encuesta debería ser lo menos penoso posible.
- Uno de los aspectos más difíciles de la logística es compaginar el entrevistador adecuado con el informante y el cuestionario pertinentes en el caso de los países donde se utilizan numerosos idiomas para las entrevistas. La composición de los equipos según las capacidades lingüísticas, junto con un plan detallado de despliegue que tenga en cuenta las necesidades lingüísticas de los equipos, es condición necesaria para conseguir que la mayor parte de los informantes mantengan las entrevistas en su lengua materna y que el entrevistador hable esa lengua y utilice un cuestionario traducido a ésta.
- El personal encargado de la entrada de datos debe seguir el curso de capacitación de entrevistadores para poder ocuparse de la entrada y edición de datos. Los cuestionarios de las DHS son bastante complicados. La participación en la capacitación para entrevistadores ofrece al personal encargado de la entrada de datos una comprensión adecuada del flujo del cuestionario y de la forma como se relacionan entre sí las diferentes partes del mismo. Este conocimiento es necesario para introducir correcciones durante el proceso interactivo de entrada y edición de datos.
- La doble entrada de datos puede parecer costosa pero ahorrará tiempo de edición. En las primeras encuestas demográficas y de salud, los datos se introducían sólo

una vez. Las encuestas posteriores han utilizado el sistema de doble entrada para determinar los errores que no pueden descubrirse con los programas de comprobación de límites y coherencia y que durante la fase de edición mo haya que corregir demasiados cuestionarios. El personal encargado del procesamiento de datos del DHS ha llegado a la conclusión de que los efectos beneficiosos de la doble entrada de datos en la edición de los mismos supera con creces los costos.

- La comunicación constante con el personal sobre el terreno acerca de los problemas encontrados en los cuestionarios terminados durante la entrada de datos es condición necesaria para conseguir datos de alta calidad. El personal sobre el terreno debe recibir, sobre todo en las fases iniciales de una encuesta, información constante sobre los errores que están cometiendo, para evitarlos en el futuro. La entrada interactiva de datos constituye un mecanismo muy valioso para la identificación precoz de los problemas sobre el terreno.
- Es preciso utilizar cuadros para revelar las pautas de respuesta que no podrán observarse al editarse los cuestionarios individuales. Por ejemplo, existe el peligro de que los entrevistadores codifiquen deliberadamente a los posibles informantes como si tuvieran más o menos edad de la que tienen realmente para no tener que entrevistarlos. La única manera de saber si está ocurriendo esto es estudiar las pautas de edad de los informantes comprobando varios centenares de entrevistas.
- En muchos países, la preparación del informe de la encuesta es una de las tareas más arduas. El fortalecimiento de la capacidad de investigación es uno de los objetivos del programa DHS. La redacción de informes es una de las esferas en que se realizan grandes esfuerzos por desarrollar la capacidad mediante una labor interactiva con los autores locales. Más recientemente, los seminarios sobre redacción de informes, durante los cuales todos los autores trabajan en diferentes capítulos del informe en colaboración con personal del DHS, se considera uno de los medios más eficaces de transferir capacidad. No obstante, la preparación de informes es también un arte que no todos poseen en la misma medida, independientemente de la titulación que se tenga en demografía o salud.
- La asistencia técnica es especialmente necesaria en el muestreo, el procesamiento de datos y la elaboración de informes. En otras esferas esta asistencia consiste muchas veces en garantizar que se ejecuten de manera oportuna los diferentes pasos de la encuesta. Las áreas antes mencionadas son las que han presentado mayores dificultades para el personal local en muchas encuestas demográficas y de salud, por no decir en la mayoría. En comparación, muchos organismos locales son capaces de ejecutar en forma muy satisfactoria las actividades de capacitación y el trabajo sobre el terreno. Por ello es preciso ofrecer la asistencia técnica necesaria y que se puedan superar las deficiencias en una o más de las áreas más problemáticas.
- Los países están dispuestos a intercambiar los datos de las encuestas con investigadores responsables. Los planes en ese sentido deberían aprobarse antes de la realización de la encuesta. El programa de encuestas demográficas y de salud ha conseguido eficazmente garantizar la aprobación de los países participantes en lo que respecta al intercambio de datos con investigadores responsables para futuros proyectos de investigación. Ello ha creado una base de datos transnacional única, de valor incalculable para los países y donantes. Para alcanzar este objetivo es preciso llegar a acuerdos con las autoridades de los países participantes en el momento en que se apruebe la encuesta. Si no se llega a esos acuerdos en tal momento, muchas veces no es posible negociarlos después, ya que el gobierno puede haber cambiado y quizá las personas encargadas de los departamentos gubernamentales responsables no sean ya las mismas que cuando se planificó la entrevista.

ANEXO

Tasas de respuesta de hogares y mujeres en 66 encuestas de 44 países, 1990-2000, en algunas regiones

Región	País	Año de la encuesta	Fase	Número de hogares	Tasa de respuesta de los hogares (%)	Número de mujeres	Tasa de respuesta de las mujeres (%)
Asia	Bangladesh	1994	DHS III	9 255	99,1	9 900	97,4
	Bangladesh	1997	DHS III	8 762	99,1	9 335	97,8
	Filipinas	1993	DHS III	13 065	99,5	15 332	98,0
	Filipinas	1998	DHS III	12 567	98,7	14 390	97,2
	Indonesia	1991	DHS II	27 106	99,1	23 470	97,6
	Indonesia	1997	DHS III	34 656	98,8	29 317	98,3
	Pakistán	1991	DHS II	7 404	97,2	6 910	95,7
Eurasia	Kazajstán	1995	DHS III	4 232	98,7	3 899	96,7
	Kazajstán	1999	MEASURE	5 960	98,1	4 906	97,8
	Kirguistán	1997	DHS III	3 695	99,4	3 954	97,3
	Turquía	1993	DHS III	8 900	96,8	6 862	95,0
	Turquía	1998	MEASURE	8 596	93,8	9 468	90,6
	Uzbekistán	1996	DHS III	3 763	98,4	4 544	97,2
América Latina	Bolivia	1994	DHS III	9 335	97,6	9 316	92,3
	Bolivia	1997	DHS III	12 281	98,6	1 831	94,6
	Brasil	1991	DHS II	6 416	94,5	6 864	90,7
	Brasil	1996	DHS III	14 252	93,2	4 579	86,5
	Colombia	1990	DHS II	8 106	91,4	9 715	89,0
	Colombia	1995	DHS III	11 297	89,5	2 086	92,2
	Colombia	2000	MEASURE	11 747	92,8	2 531	92,5
	Guatemala	1995	DHS III	11 754	96,1	3 388	92,6
	Haití	1994	DHS III	4 944	97,5	5 709	93,8
	Nicaragua	1997	DHS III	11 726	98,3	4 807	92,1
	Paraguay	1990	DHS II	5 888	96,5	6 262	93,1
	Perú	1992	DHS II	13 711	98,3	17 149	92,6
	República Dominicana	1991	DHS II	8 131	87,9	8 200	89,3
	República Dominicana	1996	DHS III	9 026	97,8	9 034	93,2
Cercano Oriente	Egipto	1992	DHS II	10 950	98,3	9 978	98,9
	Egipto	1995	DHS III	15 689	99,2	14 879	99,3
	Marruecos	1992	DHS II	6 635	99,1	9 587	96,5
	Yemen	1991	DHS II	12 934	99,2	6 515	92,2

Región	País	Año de la encuesta	Fase	Número de hogares	Tasa de respuesta de los hogares (%)	Número de mujeres	Tasa de respuesta de las mujeres (%)
África subsahariana	Benin	1996	DHS III	4 562	98,6	5 719	96,0
	Burkina Faso	1992	DHS II	5 283	97,3	6 848	92,8
	Burkina Faso	1999	DHS III	4 871	98,8	6 740	95,6
	Camerún	1991	DHS II	3 647	97,0	4 147	93,3
	Camerún	1998	DHS III	4 791	98,0	5 760	95,5
	Chad	1997	DHS III	6 930	98,7	7 705	96,7
	Comoras	1996	DHS III	2 277	98,9	3 160	96,5
	Côte d'Ivoire	1994	DHS III	5 977	99,3	8 271	97,9
	Ghana	1993	DHS III	5 919	98,4	4 700	97,1
	Ghana	1999	MEASURE	6 055	99,1	4 970	97,4
	Guinea	1999	MEASURE	5 216	97,6	7 117	94,9
	Kenya	1993	DHS III	8 185	97,1	7 952	94,8
	Kenya	1998	DHS III	8 661	96,8	8 233	95,7
	Madagascar	1992	DHS II	6 027	98,6	6 520	96,0
	Madagascar	1997	DHS III	7 349	97,6	7 424	95,1
	Malawi	1992	DHS II	5 409	98,4	5 020	96,6
	Malí	1996	DHS III	8 833	98,7	10 096	96,1
	Mozambique	1997	DHS III	9 681	95,9	9 590	91,5
	Namibia	1992	DHS II	4 427	92,6	5 847	92,7
	Níger	1992	DHS II	5 310	98,7	6 750	96,3
	Níger	1997	DHS III	6 007	98,7	7 863	96,4
	Nigeria	1990	DHS II	9 173	98,1	9 200	95,4
	Nigeria	1999	MEASURE	7 736	98,8	10 529	93,2
	República Centroafricana	1994	DHS III	5 583	99,4	6 005	98,0
	República Unida de Tanzanía	1992	DHS II	8 560	97,3	9 647	95,8
	República Unida de Tanzanía	1996	DHS III	8 141	97,9	8 501	95,5
	Rwanda	1992	DHS II	6 292	99,4	6 947	94,3
	Senegal	1993	DHS II	3 563	99,0	6 639	95,0
	Senegal	1997	DHS III	4 855	98,3	9 186	93,5
	Togo	1998	DHS III	7 620	98,6	8 964	95,6
	Uganda	1995	DHS III	7 671	98,4	7 377	95,8
	Zambia	1992	DHS II	6 245	99,4	7 247	97,4
	Zambia	1996	DHS III	7 365	98,9	8 298	96,7
	Zimbabwe	1994	DHS III	6 075	98,5	6 408	95,6
	Zimbabwe	1999	MEASURE	6 512	97,8	6 208	95,2

Reconocimientos

Los autores agradecen las valiosas observaciones de los miembros de un equipo de expertos convocados por las Naciones Unidas para examinar el proyecto de publicación, las consideraciones profesionales de los revisores externos y las del Dr. Alfredo Aliaga, de ORC Macro.

REFERENCIAS BIBLIOGRÁFICAS

Cleland, J. y C. Scott, comps. (1987). *The World Fertility Survey. An Assessment.* Nueva York: Oxford University Press.

Institute for Resource Development/Macro Systems, Inc. (1990). *An Assessment of DHS-I Data Quality*. Methodological Report, No. 1. Columbia, Maryland.

_____ *(1994). An Assessment of the Quality of Health Data in DHS-I Surveys.* Methodological Report, No. 2. Columbia, Maryland. *Review,* 56(3): 259-278.

Kish, L. (1965). *Survey Sampling.* Nueva York: Wiley.

_____, R. Groves y Krotki (1976). *Sampling Errors in Fertility Surveys*. World Fertility Survey Occasional Paper, No. 17, La Haya: Instituto Internacional de Estadística.

Lê, T. y V. Verma (1997). *An Analysis of Sample Designs and Sampling Errors of the Demographic and Health Surveys*. DHS Analytical Reports, No. 3. Calverton, Maryland: Macro International, Inc.

ORC Macro (1996). *Sampling Manual*. DHS-III Basic Documentation, No. 6. Calverton, Maryland.

_____ (2001). Survey organizational manual (en borrador). Calverton, Maryland.

Scott, Christopher y otros (1988). Verbatim questionnaires versus field translation or schedules: an experimental study. *International Statistical Review,* vol. 56, No. 3, págs. 259-278.

Vaessen, Martin y otros (1987). Translation of questionnaires into local languages. En *The World Fertility Survey: An Assessment.* John Cleland y Chris Scott, comps. Nueva York: Oxford University Press.

Verma, V. y T. Lê (1996). An analysis of sampling errors for the Demographic and Health Surveys. *International Statistical Review*, vol. 64, págs. 265-294.

Verma, V., C. Scott y C. O'Muircheartaigh (1980). Sample designs and sampling errors for the World Fertility Survey. *Journal of the Royal Statistical Society A*, vol. 143, págs. 431-473.

Capítulo XXIII
Encuestas del Estudio de medición de los niveles de vida

Kinnon Scott, Diane Steele y Tilahun Temesgen
Banco Mundial
Washington, D.C.
Estados Unidos de América

RESUMEN

El Estudio de medición de los niveles de vida (EMNV) surgió como consecuencia de la necesidad de mejorar los datos estadísticos sobre los hogares con miras a formular, aplicar y evaluar políticas sociales y económicas en los países en desarrollo. Este programa ha tratado sobre todo de comprender, medir y supervisar las condiciones de vida, la interacción del gasto público y los programas gubernamentales con el comportamiento de los hogares, las evaluaciones previas y *a posteriori* de las políticas y las causas de los resultados sociales observados. Las encuestas resultantes del EMNV utilizan numerosos instrumentos para obtener los datos necesarios para alcanzar esos objetivos y están basadas en importantes mecanismos de control para garantizar la alta calidad de los datos pertinentes. Sobre todo en los últimos años, el programa EMNV ha insistido en el proceso de implicación de los usuarios de datos en el diseño de las encuestas y se ha ocupado de los problemas de la sostenibilidad. En este capítulo se presenta un panorama general de las encuestas del EMNV y de los principales métodos de diseño y realización utilizados en ellas, así como de los esfuerzos por promover la capacidad de análisis. Se incluye una evaluación de los costos de la encuesta y de la calidad de los datos obtenidos y ejemplos de los diferentes usos normativos de los datos contenidos en las encuestas EMNV. En el capítulo se examinan también los cálculos de los efectos medios del diseño muestral y los coeficientes de correlación e intraclase de algunas variables referidas a hogares e individuos utilizando algunas encuestas del EMNV.

Términos clave: Medición de la pobreza, niveles de vida, metodología de la encuesta, efecto del diseño, correlación intraclase, control de calidad.

A. Introducción

1. Los gastos del sector público para servicios sociales y de infraestructura representan una parte significativa de los recursos, tanto en cifras absolutas como relativas. No es raro que el gasto de salud y de educación representen, en uno y otro caso, entre el 3% y el 4% del producto interno bruto (PIB). Según el país, esa cifra puede oscilar entre varios millones y varios centenares de millones de dólares. Los cambios importantes en las políticas económicas referentes a los impuestos y precios modifican sustancialmente los niveles relativos y absolutos de bienestar. No obstante, muchas veces, debido a la falta de datos, las políticas se formulan, aplican y revisan sin tener demasiada información sobre su eficacia global para mejorar las condiciones de vida de la población del país. La ausencia de datos adecuados

sobre los hogares obliga a las autoridades a recurrir a los datos administrativos que si bien son adecuados para determinados fines, muchas veces limitan gravemente la capacidad de comprender el comportamiento de los hogares, la forma en que las políticas gubernamentales repercuten en los hogares e individuos y los determinantes de los resultados sociales observados. La función de las encuestas de hogares es precisamente cubrir esas lagunas.

2. Las encuestas del Estudio de medición de los niveles de vida son uno de los instrumentos que los gobiernos pueden utilizar, y utilizan, para entender mejor las causas de los resultados observados y los efectos de sus políticas. Estas encuestas van más allá de la simple medición de los resultados y permiten establecer conexiones entre los muchos factores que afectan o provocan esos resultados. Las encuestas de hogares sobre un solo tema dan información valiosa y en profundidad sobre una cuestión específica, pero no bastan para explicar el porqué de ciertos resultados ni los factores que intervienen. El objetivo de una encuesta del EMNV es explorar la vinculación entre los distintos activos y características del hogar, por un lado, y las acciones gubernamentales, por otro, para así comprender las fuerzas que influyen en cada sector, sobre ciertos comportamientos y los resultados. La mayor comprensión gubernamental de los factores que influyen en las condiciones de vida permite mejorar las políticas y programas. A su vez, ello puede dar lugar a un uso más eficiente y eficaz de los no siempre sobrados recursos gubernamentales y privados y mejorar los niveles de vida.

3. Las encuestas del EMNV son una muestra de los esfuerzos de colaboración entre los gobiernos de los países que administran las encuestas, los principales usuarios de los datos en los países y el Banco Mundial, además de organizaciones donantes bilaterales y multilaterales[1]. Aunque estén basadas en un conjunto básico de conceptos, cada encuesta requiere cierta adaptación, según las necesidades de cada gobierno en un determinado momento. El principal organismo de ejecución suele ser la oficina nacional de estadística (ONE) que toma la iniciativa de diseñar el cuestionario, la muestra y las metodologías del trabajo sobre el terreno utilizando la técnicas consideradas más eficaces por el EMNV.

4. En este capítulo se presenta un panorama general del estudio de medición de los niveles de vida. En primer lugar se describe la historia del programa y luego se presenta información sobre las características fundamentales de las encuestas del EMNV. Se incluye una sección en que se explica de qué forma han repercutido las características del diseño del EMNV en la calidad de los datos recopilados. El capítulo finaliza con algunos ejemplos de utilización de los datos de encuestas del EMNV.

[1] Entre las instituciones que han colaborado en encuestas del EMNV cabe citar el Banco Interamericano de Desarrollo, organizaciones de las Naciones Unidas como el Programa de las Naciones Unidas para el Desarrollo, el Fondo de las Naciones Unidas para la Infancia, el Fondo de Población de las Naciones Unidas y donantes bilaterales de Canadá, Dinamarca, los Estados Unidos de América, Japón, Noruega y el Reino Unido de Gran Bretaña e Irlanda del Norte.

B. ¿Por qué una encuesta del EMNV?

5. Los esfuerzos del EMNV por responder a la necesidad que tienen los responsables de la formulación de las políticas de datos de calidad comenzaron en 1980. Tras un período de actividad de cinco años, en el que se realizó un examen de las encuestas de hogares existentes y se entablaron amplias consultas con investigadores y autoridades para determinar los tipos de datos necesarios y con especialistas en metodología de encuestas para determinar los mejores procedimientos para el trabajo sobre el terreno, las primeras encuestas del EMNV tuvieron carácter experimental en Côte d'Ivoire y Perú en 1985. Estas dos primeras encuestas fueron proyectos de investigación para comprobar toda la metodología y determinar la utilidad y calidad de los datos que podrían obtenerse[2]. El éxito de las dos primeras encuestas explica por qué se han realizado más de 60 encuestas del EMNV en más de 40 países desde 1985 (véase una lista completa en el anexo I).

[2] Puede verse una exposición más detallada de la historia del EMNV en Grosh y Glewwe (1995).

C. Características fundamentales de las encuestas del EMNV

6. A continuación se presenta un resumen de las características fundamentales del EMNV. El lector puede consultar el manual del EMNV de 1996, que contiene información más detallada sobre las encuestas y la manera de realizarlas[3].

[3] En Grosh y Muñoz (1996).

1. Contenido e instrumentos utilizados

7. Las encuestas del EMNV tienen hasta seis instrumentos independientes; entre ellos: *a*) un cuestionario de hogares, para recopilar información sobre hogares e individuos y sobre las actividades económicas del hogar (agricultura y empresas familiares); *b*) un cuestionario sobre la comunidad[4], para la recopilación de datos sobre el entorno de los hogares, con especial atención a los servicios disponibles, actividades económicas, acceso a los mercados y, recientemente, el patrimonio social; *c*) un cuestionario sobre precios, administrado en todas las áreas de hogares, para poder realizar ajustes del costo de vida[5], y *d*) un cuestionario sobre servicios, administrado a los proveedores locales de servicios para conocer el tipo y calidad de servicios a disposición de los hogares. El esquema XXIII.1 muestra los instrumentos utilizados para lograr los fines de las encuestas del EMNV y las variables necesarias.

8. Los contenidos de los instrumentos de la encuesta reflejan las necesidades prioritarias de datos del país en un determinado momento. Como lo que interesa en general es medir los niveles de vida en todas sus diversas facetas, las encuestas de hogares, en particular, tratan de recopilar información sobre la gran variedad de temas que influyen en ese sentido. En el cuadro XXIII.1 puede verse el contenido de una encuesta característica del EMNV; en este caso la de Viet Nam en 1997-1998.

[4] Téngase en cuenta que no se trata de una 'comunidad' en el sentido sociológico del término, sino más bien de un mecanismo para recopilar información sobre las áreas donde se encuentran los hogares seleccionados para la encuesta.

[5] Los índices nacionales sobre el precio de consumo muchas veces son inadecuados para este fin porque suelen considerar sobre todo las zonas urbanas y porque cuando incluyen las zonas rurales los precios no se recogen con el debido nivel de desglose.

Esquema XXIII.1
Relación entre objetivos del EMNV e instrumentos de las encuestas

Objetivo	Indicadores	Instrumentos
Medición del bienestar de los individuos y los hogares		
• Niveles, distribución y correlaciones	• Consumo • Ingreso • Riqueza, ahorro • Capital humano • Antropometría	• Cuestionario de hogares • Cuestionario de precios
Analizar las políticas		
• Quién se beneficia de los programas/gasto público • Efectos de los programas/gasto público • Disponibilidad de los servicios • Calidad de los servicios • Precio de los servicios • Efecto de las políticas económicas	• Uso de los servicios • Quién recibe los servicios, transferencias • Costos de los servicios • Repercusión de las políticas • Distancia al servicio más próximo • Tipos de servicios ofrecidos • Personal, presupuesto, otros insumos • Transferencias netas entre sectores	• Cuestionario de hogares • Cuestionario de la comunidad • Cuestionario de servicios • Cuestionario de precios
Identificar los determinantes		
• Por qué se producen los resultados observados • Qué influye en el comportamiento de los hogares	• Composición de los hogares • Capital humano, bienestar, servicios disponibles, etcétera	• Cuestionario del hogar • Cuestionario de la comunidad • Cuestionario de servicios • Cuestionario de precios

Cuadro XXIII.1
Contenido del cuestionario de hogares de Viet Nam, 1997-1998

Primera visita	Segunda visita
• Lista de hogares • Educación • Salud • Trabajo • Migración • Vivienda y servicios públicos	• Fecundidad • Agricultura, silvicultura y pesca • Empleo por cuenta propia no agrícola • Gastos y producción alimentarios • Bienes no alimenticios y duraderos • Ingresos procedentes de remesas • Endeudamiento, préstamos y ahorros • Antropometría

9. Existe un alto grado de adaptación del cuestionario a cada país, lo que da lugar a variaciones en el contenido general de los instrumentos de la encuesta y a incluir nuevos módulos y temas con el paso del tiempo. Por ejemplo, en Bosnia y Herzegovina en 2001 el módulo de salud se amplió para incorporar cuestiones relativas a la depresión con el fin de medir su incidencia e identificar las vinculaciones entre ella y otros aspectos del bienestar y la participación en el mercado de trabajo. En Guatemala, en 2002, se incorporó un módulo sobre patrimonio social para recoger información sobre las dimensiones sociales de la pobreza, como la participación en programas comunitarios/gubernamentales e iniciativas colectivas, las causas de la exclusión de la sociedad, las opiniones sobre el bienestar y las opiniones sobre la justicia y el acceso a ella. En Albania, Brasil, Nepal, Jamaica, Sudáfrica y Tayikistán se añadieron preguntas sobre las medidas subjetivas de la pobreza con el fin de examinar la relación entre éstas y otras medidas[6]. En el cuadro XXIII.2 se presenta una muestra de módulos que se han incorporado en los últimos años. En resumen, si bien existe un conjunto estándar de módulos, la encuesta del EMNV de cada país refleja sus propias prioridades, sus necesidades de datos o preocupaciones en el momento de la encuesta. Un proyecto reciente del Banco Mundial cuyo objetivo era mejorar la relevancia normativa de las encuestas del EMNV ha dado lugar a un nuevo libro en el que se esbozan, desglosadas por tema, las cuestiones normativas que pueden resolverse con los datos del EMNV y en el que se ofrece asesoramiento sobre el diseño del cuestionario[7].

10. La fase de diseño del cuestionario es un proceso orientado a garantizar la identificación e incorporación de las cuestiones normativas pertinentes. En la mayor parte de los países se forma un grupo de usuarios de datos o un comité directivo integrado por miembros de diferentes ministerios, donantes y del mundo académico, junto con la oficina nacional de estadística (ONE). Este grupo se encarga de identificar las necesidades de datos para políticas específicas con el fin de garantizar la recopilación de los datos pertinentes. La fase de diseño del cuestionario dura por término medio unos ocho meses y en ella se intenta contar con el mayor número de participantes posibles. Este proceso más bien prolongado tiene la ventaja adicional de generar la demanda de los datos resultantes y la identificación con los mismos. Ello, a su vez, contribuye a una mayor utilización de los datos en la formulación de las políticas.

[6] Puede verse información más detallada sobre las actividades relacionadas con el patrimonio social en Guatemala, en Banco Mundial (2002b). En cuanto a las medidas subjetivas de la pobreza, puede consultarse Pradhan y Ravallion (2000), Ravallion y Lokshin (2001), Ravallion y Lokshin (2002). El análisis de los datos sobre Bosnia y Herzegovina es una labor en curso.

[7] En Grosh y Glewwe, comps. (2000).

Cuadro XXIII.2
Ejemplos de módulos adicionales

Temas	Países y año
Actividades de la vida cotidiana	Kosovo (2000), Kirguistán (1993, 1996, 1997, 1998), Jamaica (1995) Nicaragua (1993)
Discapacidad	Nicaragua, (1993)
Efectos de la mortalidad relacionada con el SIDA	República Unida de Tanzanía-Kagera (1991-1994)
Pruebas de alfabetización y aritmética elemental	Viet Nam (1997-1998), Jamaica (1990), Marruecos (1990-1991)
Salud mental	Bosnia y Herzegovina (2001)
Privatización	Bosnia y Herzegovina (2001), Kirguistán (1996, 1997)
Crisis/vulnerabilidad	Bolivia (1999, 2000), Guatemala (2000), Paraguay (2000-2001), Perú (1999)
Patrimonio social	Guatemala (2000), Kosovo (2000), Panamá (1997), Paraguay (2000-2001)
Medidas subjetivas de la pobreza	Albania (2002), Brasil (1996), Jamaica (1997), Nepal (1996), Sudáfrica (1993), Tayikistán (1999)
Uso del tiempo	Guatemala (2000), Nicaragua (1998), Jamaica (1993), Pakistán (1991) Marruecos (1990-1991), República Unida de Tanzanía-Kagera (1991-1994)

2. Cuestiones relacionadas con la muestra

11. Normalmente, las encuestas del EMNV son encuestas nacionales que utilizan muestras probabilísticas en varias etapas[8]. Las muestras globales son pequeñas (en relación con algunas otras encuestas), y normalmente oscilan entre 2.000 y 5.000 hogares. Hay dos razones principales para limitar el tamaño de la muestra. En primer lugar, existe la preocupación por la calidad y la necesidad de compaginar el error de muestreo con el error no muestral (véase el subapartado C.4. *infra,* donde se examina con mayor detalle este punto). En segundo lugar, la orientación de las encuestas del EMNV se centra en los determinantes o relaciones entre las características de los hogares y no en las estimaciones precisas de las tasas, coeficientes o promedios específicos. Por estas razones las muestras del EMNV son razonablemente pequeñas y, en general, no lo bastante grandes como para que los resultados de la encuesta puedan desglosarse en áreas geográficas pequeñas, como estados, municipios o departamentos.

12. En todas las encuestas del EMNV se utiliza el muestreo probabilístico, aunque el diseño efectivo utilizado varía según el país y la situación[9]. Se identifican los campos de estudio (zonas urbanas/rurales, regiones) y dentro de cada campo se utiliza un diseño por conglomerado en dos etapas[10]. Como ocurre en la mayoría de las encuestas de hogares, las del EMNV utilizan un diseño por conglomerados en lugar de una muestra aleatoria simple. Ello se debe a motivos relacionados con los costos, aun cuando los diseños por conglomerados reduzcan la precisión de las estimaciones (véase el subapartado E.4 *infra*, donde se explican más detenidamente los efectos del diseño muestral resultantes de la utilización del muestreo en varias etapas; véase también el anexo III). Las unidades primarias de muestreo (UPM) son unidades de área geográficamente definidas, seleccionadas con probabilidad proporcional al tamaño. El marco muestral suele ser el censo de población más reciente del país, pero se han utilizado otras soluciones alternativas cuando el censo era inexistente o irrelevante (véanse, por ejemplo, los documentos de información básica correspondientes al EMNV de Nicaragua de 1993, donde los registros de votación complementaron la desfasada información del censo, y el EMNV de Bosnia y Herzegovina de 2001, en el que se necesitaron amplias operaciones de enumeración debido a la guerra civil, por ejemplo).

13. Una vez seleccionadas las UPM, se lleva a cabo una enumeración de las mismas con el fin de disponer de una lista exacta y actualizada de todas las viviendas de hogares. Esta operación se realiza en la fecha más próxima posible al trabajo sobre el terreno. Para evitar posibles sesgos, la operación no la realizan los entrevistadores mismos, sino el departamento de cartografía de la ONE. Una vez que se dispone de una lista actualizada completa de todas las viviendas de la UPM, se seleccionan sistemáticamente las unidades secundarias de muestreo (hogares), normalmente un número fijo de hogares dentro de cada UPM, que suelen ser entre 12 y 18. Luego se recopilan los datos de todos los miembros del hogar. Si bien los diseños muestrales de las encuestas del EMNV tratan de englobar a todo el país, en algunos casos, debido a conflictos civiles o a catástrofes naturales, algunas zonas concretas quedan excluidas.

14. Las estimaciones de las encuestas del EMNV en general requieren el uso de ponderaciones de la muestra. Aun cuando el diseño original de la muestra requiera un sistema autoponderado, por ejemplo, como en Ghana, Nicaragua (1993) y Tayikistán, las distintas tasas de falta de respuesta obligan a introducir ponderaciones diferenciales para el análisis de los datos. De hecho, la mayoría de los diseños muestrales no son autoponderados. Muchas veces el diseño de la muestra en un determinado país se ve afectado por consideraciones analíticas nacionales. Por ejemplo, subgrupos de población que son pequeños pero de interés para el gobierno (minorías étnicas, regiones remotas, personas dedicadas a una actividad económica concreta o que se encuentran en la zona de un proyecto gubernamental importante) quizá tengan que estar sobrerrepresentados para garantizar que se disponga de casos

[8] De hecho, como ocurre en la mayoría de las encuestas de hogares, lo que se selecciona es la vivienda y luego se entrevista a todos los hogares que se encuentran en la vivienda seleccionada. Cuando se recurre al diseño por panel, la selección de la vivienda o el hogar dependerá del objetivo del panel y de cuestiones relacionadas con la logística.

[9] El documento de información básica de cada encuesta contiene los detalles del diseño de la muestra en relación con cada una de las encuestas. Esta información puede encontrarse en el sitio web del EMNV: http://www.worldbank.org/lsms/.

[10] No obstante, en algunos países ha habido que utilizar diseños en tres etapas.

suficientes para realizar un análisis independiente de los mismos. También en este caso sería necesario introducir ponderaciones de la muestra en el análisis de los datos. Una observación final, dados los diseños muestrales utilizados en las encuestas del EMNV, es que las pruebas estadísticas de importancia realizadas sobre los datos deben tener en cuenta que se trata de un diseño en varias etapas (véanse los capítulos de esta publicación sobre los efectos del diseño muestral, donde se explica con más detalle este tema).

3. Organización del trabajo sobre el terreno

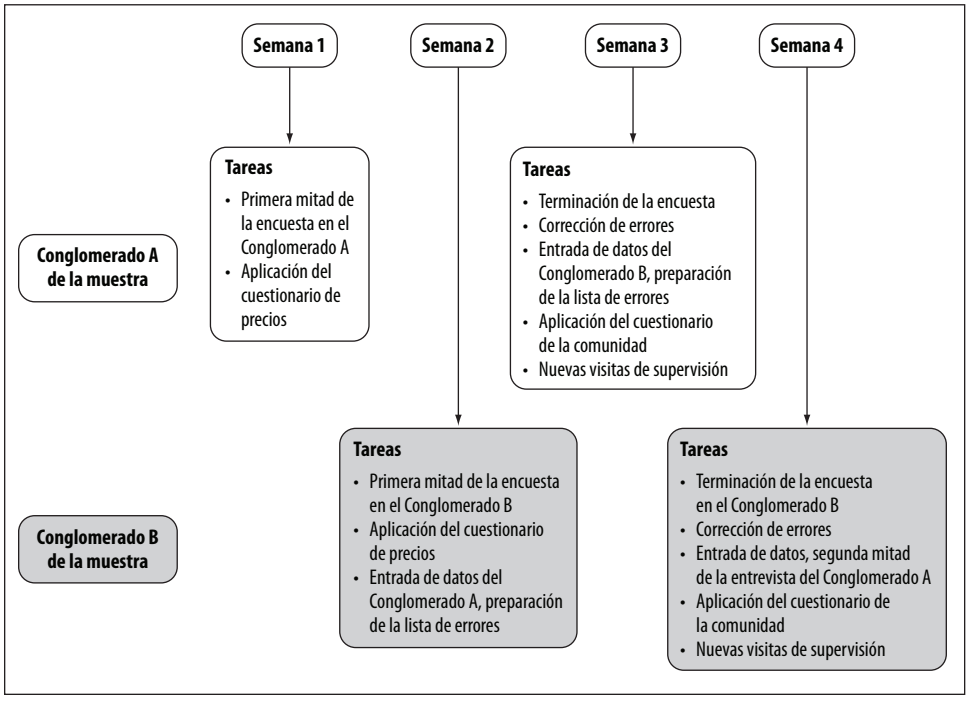
15. Como se ha visto antes, los objetivos de las encuestas del EMNV condicionan la estructura y contenido de las encuestas: se reflejan también en los métodos utilizados en el trabajo sobre el terreno. Éste se concibe de tal manera que los datos sean recopilados por equipos de entrevistadores móviles que incorporen las actividades de entrada de datos y una fuerte supervisión[11]. Cada hogar es visitado al menos dos veces, con un intervalo de dos semanas. En el esquema XXIII.2 se representa gráficamente cómo se realiza el trabajo sobre el terreno. Éste se organiza de tal manera que cada mes el equipo termine las entrevistas en dos comunidades seleccionadas (UPM). Los equipos trabajan en la primera comunidad durante la primera y la tercera semanas del mes y en la segunda comunidad durante la segunda y cuarta semanas. La primera mitad del cuestionario se termina en la primera visita, realizada en la semana 1 ó 2 según la comunidad. Entre las distintas visitas se procede a la entrada de datos de la primera visita y se comprueba la existencia de errores. La segunda visita se utiliza para corregir los errores de la primera, aplicar la segunda mitad de la encuesta y ofrecer un tiempo fijo para la información recopilada sobre los gastos en alimentos[12]. Los datos suelen recopilarse durante un período de 12 meses, con el fin de poder introducir los ajustes estacionales necesarios, aun cuando muchos países han preferido adoptar períodos más breves.

16. El supervisor se encarga de administrar los cuestionarios sobre la comunidad y sobre los precios en paralelo con su equipo de entrevistadores que recopilan los datos relati-

[11] Véase el anexo II, donde se describen con mayor detalle los equipos de entrevistadores.

[12] Si bien están previstas dos visitas, el uso de informantes directos en todas las secciones del cuestionario significa que de hecho los entrevistadores visitan cada hogar cuantas veces sea necesario para poder entrevistar a todos los miembros del mismo.

Esquema XXIII.2
Calendario de las actividades de cada equipo durante un mes



vos a los hogares en la UPM. Las encuestas de servicios pueden requerir personal adicional para su administración.

4. Calidad

17. Una preocupación fundamental y constante en las encuestas del EMNV es garantizar la alta calidad de los datos obtenidos. Debido a la complejidad de la encuesta, los mecanismos de control de calidad revisten particular importancia. Como puede verse en el cuadro XXIII.3, los controles de calidad adoptan diversas formas, desde los más sencillos —preguntas formuladas explícitamente, pautas expresas de salto, cuestionarios traducidos a los idiomas pertinentes en un país y cuestiones cerradas para reducir los errores del entrevistador— hasta el más complejo, consistente en la entrada simultánea de datos con visitas repetidas inmediatas a los hogares para corregir las contradicciones o incluir los datos omitidos. Evidentemente, no todos estos controles de calidad son privativos de las encuestas del EMNV, pero dada la complejidad de estas encuestas se ha hecho hincapié en incorporar un conjunto completo de controles de calidad. Además de los controles mencionados, el programa ha optado, en forma quizá más polémica, por una muestra de tamaño pequeño para reducir los errores no muestrales. La justificación es que, si bien los errores muestrales pueden ser considerables cuando se utilizan muestras de pequeño tamaño, esos errores pueden al menos cuantificarse. Por el contrario, los errores no muestrales proceden de muchas causas y su magnitud es prácticamente imposible de cuantificar. En todo caso, es un hecho bien conocido que la totalidad del error no muestral tiende a aumentar a medida que disminuye la muestra. Por ello se tomó la decisión de limitar estos errores no muestrales aun cuando ello limitara el nivel de desglose geográfico posible con los datos de la encuesta. La insistencia de las encuestas EMNV en explorar las relaciones entre aspectos de los niveles de vida, en vez de medir con gran precisión indicadores o tasas específicas, significa que ello no representa un obstáculo tan grande como podría ocurrir en otras encuestas[13]. Finalmente, los métodos recientes que vinculan los datos de las encuestas del EMNV (y otros) a los datos del censo que permiten una imputación de la pobreza dentro de los datos del censo, sirven para redu-

[13] Por ejemplo, una encuesta de población activa, que en principio debería reflejar cambios muy pequeños en las tasas de desempleo a lo largo del tiempo, requerirá una muestra mucho mayor de lo necesario para analizar los determinantes del desempleo, que deberían constituir el objetivo del análisis en las encuestas del EMNV.

Cuadro XXIII.3
Controles de calidad en las encuestas del EMNV

Área del control de calidad	Controles
Cuestionario	• Preguntas redactadas textualmente • Pautas de salto explícitas • Uso mínimo de preguntas abiertas • Traducción escrita a las lenguas pertinentes* • Temas delicados tratados al final • Presentación: un formulario para todos los datos de hogares e individuos
Fase piloto	• Prueba piloto formal del cuestionario y trabajo sobre el terreno
Informantes directos	• Individuos y personas mejor informadas
Entrada de datos paralela	• Comprobación de los errores de límite y coherencia • Nuevas visitas a los hogares para introducir correcciones
Formato en dos rondas	• Reduce la fatiga • Crea un período de recuerdo limitado • Permite la comprobación de la entrada de datos y su corrección con los hogares
Capacitación	• Capacitación intensiva de los entrevistadores (un mes), supervisores y personal encargado de la entrada de datos
Trabajo sobre el terreno descentralizado	• Equipos móviles formados por el supervisor, entre dos y tres entrevistadores y el encargado de la entrada de datos con computadora e impresora, y conductor con automóvil
Supervisión	• Un supervisor por dos o tres entrevistadores
Muestra de pequeño tamaño	• Limitación del error no muestral
Política de acceso a los datos	• Uso de los datos abiertos a todos los investigadores e instituciones

* En los países donde hay lenguas no escritas (lenguas indígenas de Panamá, por ejemplo) se utilizan entrevistadores bilingües. La solución no es perfecta, y sólo debería recurrirse a ella cuando sea absolutamente necesario.

cir, hasta cierto punto, el problema del pequeño tamaño, al menos en lo que respecta a las medidas de la pobreza y la desigualdad[14].

18. Otro mecanismo de control de calidad incorporado en las encuestas del EMNV es el uso de informantes directos, conocidos con el nombre de autoinformantes. Ello presenta dos grandes ventajas. Reduce la carga de los informantes y disminuye la fatiga. El cuestionario de hogares es en realidad una serie de breves entrevistas individuales (10-15 minutos); en él se somete a períodos de entrevistas más largos sólo a los informantes con mayores conocimientos sobre el consumo, la agricultura y las empresas domésticas[15]. El uso de informantes directos mejora también la calidad de los datos obtenidos, ya que garantiza que quien responda a las preguntas sea la persona mejor informada[16]. No es razonable pensar que una sola persona del hogar pueda presentar datos precisos y completos sobre la salud, la educación, el trabajo, la migración, el crédito, la fecundidad o las actividades de todos los demás miembros del hogar; es, sencillamente, demasiada información. Además, puede haber razones dentro de un hogar para no compartir entre todos los miembros cierta información (el crédito, el ahorro, los ingresos y el uso de métodos anticonceptivos son actividades acerca de las cuales quizá no se comparta toda la información). La utilización de informantes directos es, por lo tanto, la única manera de garantizar una información precisa sobre cada uno de los miembros del hogar. Se capacita a los entrevistadores para que en la medida de lo posible realicen las entrevistas individuales en privado.

19. La capacitación de todo el personal que participa en cada una de las encuestas del EMNV es otro mecanismo de control de calidad. Adopta la forma de capacitación "en el empleo" para el personal de la ONE, así como de cursos más formales cuando las necesidades lo requieren. En cuanto al personal sobre el terreno, los entrevistadores, los supervisores y los encargados de la entrada de datos, se invierten cuantiosos recursos en actividades de capacitación. Normalmente, la capacitación para el personal sobre el terreno dura cuatro semanas y tiene alcance teórico y práctico. Una vez terminada la capacitación, se selecciona al personal sobre el terreno de entre los que han superado el curso de capacitación. Los resultados satisfactorios normalmente se basan en una combinación de participación activa durante el curso y la superación de un examen al final de él.

20. Un último método para mejorar la calidad de los datos, muchas veces olvidado, es promover el libre acceso a los microdatos resultantes de la encuesta. El uso generalizado de los conjuntos de datos por una gran variedad de investigadores y autoridades da lugar a una atenta comprobación de los datos y se crea un circuito de intercambio de información con los productores de datos que permite mejorar la calidad de las encuestas futuras. Se han conseguido acuerdos de libre acceso a los datos para la mayor parte de los conjuntos de datos de las encuestas del EMNV y se está procurando ayudar a los gobiernos a divulgar dicha información. Aunque el Banco Mundial no es propietario de los conjuntos de datos de las encuestas del EMNV, está autorizado para divulgar directamente más de la mitad de ellos (de hecho, el 30% de todos los conjuntos de datos puede descargarse directamente del sitio web del EMNV)[17]. Del resto de conjuntos de datos, la mayoría pueden distribuirse si el gobierno aprueba la solicitud individual. Quienes han solicitado este tipo de permiso indican que la aprobación se concede aproximadamente en el 90% de los casos.

5. Entrada de datos

21. La entrada paralela de datos implica la utilización de programas informáticos avanzados que comprueben los errores de límite y las contradicciones entre los diferentes registros y dentro de cada uno de ellos incluso con cuadros externos de referencia (por ejemplo, los que contienen información antropométrica, sobre el rendimiento de las cosechas y sobre los precios). Cuando es posible, los datos se introducen en las computadoras móviles sobre el terreno durante la fase de recopilación de datos, y los encargados de la entrada de da-

[14] Véase el apartado F *infra* sobre los usos de los datos del EMNV, donde se expone con mayor detalle esta técnica.

[15] Incluso en el caso de los individuos "mejor informados" el tiempo real de la entrevista es de menos de una hora, tiempo que se considera el máximo que puede durar una entrevista personal. En algunos hogares ese límite a veces se supera y hay que hacer lo posible por evitar la fatiga del informante y el consiguiente descenso de la calidad de los datos.

[16] En el caso de nuños menores de 10 o 12 años, o de miembros del hogar incapaces de comunicarse, pueden utilizarse declarantes sustitutos. Cuando así ocurre, se indica el código de identificación de la persona que responde realmente.

[17] http://www.worldbank.org/lsms/.

tos forman parte integrante de los equipos móviles. Los datos se introducen inmediatamente después de terminada cada entrevista, y el proceso da lugar a una lista de errores, contradicciones y lagunas informativas. Luego el entrevistador visita de nuevo el hogar a o los hogares con información omitida, para aclarar con los miembros del mismo los posibles problemas y completar la información. De esa manera se evita limpiar grandes lotes de datos una vez terminada la encuesta. Esa limpieza es mejor evitarla porque aunque esa depuración tiende a crear conjuntos de datos internamente coherentes, no son los que mejor reflejan la situación de cada individuo. Además, requiere bastante tiempo, con lo que se retrasa el uso de los datos y, lo que es peor, cuando se dan a conocer ya están desfasados. Con la llegada de computadoras poco costosas pero de gran potencia y los nuevos avances de los programas informáticos es probable que algunas encuestas del EMNV se lleven a cabo en forma completamente electrónica utilizando métodos de entrevistas personales con ayuda de la computadora (CAPI). Éste es un camino que se está explorando actualmente, para reducir el tiempo entre el trabajo sobre el terreno y la publicación de los datos y para mejorar la calidad de los datos[18].

[18] El uso de los sistemas CAPI es uno de los factores que permiten a la Oficina del Censo de los Estados Unidos publicar los resultados de su encuesta mensual de población activa (Current Population Survey) a los 10 días de acabado el trabajo sobre el terreno. Está en proyecto un experimento para comparar los costos y beneficios del sistema CAPI con la entrada paralela de datos para las encuestas del EMNV en Albania.

6. Sostenibilidad

22. En su nivel más elemental, los tres mayores impedimentos para la sostenibilidad, para la aplicación a largo plazo de las encuestas del EMNV y para el uso de los datos resultantes en la formulación de políticas son las restricciones presupuestarias, la rotación de personal y la falta de capacidad de análisis. Si bien no existe un programa concreto para garantizar la sostenibilidad, la experiencia del EMNV ha dado varias pistas para aumentar la sostenibilidad. La primera es sobre la importancia de contar con los responsables de la formulación de políticas y los usuarios de los datos en la fase de diseño y análisis. Ésta comienza fundamentalmente con el proceso de crear una demanda de resultados del EMNV y el uso de los datos en las decisiones sobre políticas. Como son estos usuarios los que se benefician de los datos (no la ONE en cuanto tal), éste es el grupo más interesado en conseguir que se atiendan las necesidades presupuestarias de las futuras encuestas mientras se preparan los presupuestos del Estado. Muchas veces la creación o identificación de uno o varios "promotores" de la encuesta y de los datos fuera del sistema estadístico es un requisito fundamental para la sostenibilidad[19].

[19] Un buen ejemplo es el de Jamaica. La solicitud la hizo originalmente la Oficina del Primer Ministro. El Ministro de Planificación ha participado en todas las etapas del diseño y utilización de la encuesta en colaboración con la Oficina de Estadística, que es la que se encarga de realizar la encuesta. El EMNV se ha realizado anualmente desde los últimos años ochenta en Jamaica. Véase más información sobre este ejemplo en Grosh (1991).

23. La segunda enseñanza fundamental es que la sostenibilidad es un proceso a largo plazo: la inversión en encuestas aisladas tiene poca repercusión a largo plazo. Se necesita un esfuerzo sistemático a lo largo de varios años para capacitar un equipo de personal, demostrar la eficacia y utilización de los instrumentos, vincular a productores y a usuarios y adaptar la metodología a las necesidades y capacidades del país. Además, invertir en la documentación adecuada para los fines de la encuesta, en el almacenamiento de los datos y en la divulgación de las actividades es un modo de asegurar que el esfuerzo y el dinero empleados no están pendientes de algo tan contingente como el cese de un funcionario. Casi el 40% de los países que han realizado una encuesta del EMNV han organizado luego múltiples encuestas.

24. Finalmente, entre los objetivos explícitos debe figurar el fortalecimiento de la capacidad de análisis[20]. Con ello se consigue incrementar el uso de los datos, lo que ayuda a crear demanda de futuros conjuntos de datos. Además, la creciente capacitación del personal de la ONE y, por lo tanto, su mayor relevancia dentro de la Administración pueden animar al personal a continuar en la institución[21]. Finalmente, razones externas pueden ayudar también a que aumente la demanda de datos. Las estrategias de lucha contra la pobreza de los países que reciben financiamiento del Banco Mundial y del Fondo Monetario Internacional (FMI) en condiciones de préstamo, y las incluidas en los objetivos de desarrollo del Milenio requieren datos sobre la medición y seguimiento de la pobreza y de los indicadores sociales más importantes. El carácter a largo plazo de estos objetivos puede ayudar a fomentar los sistemas de seguimiento y evaluación que dependen fuertemente de las encuestas de hogares,

[20] Puede encontrarse un resumen de las enseñanzas aprendidas en las encuestas del EMNV, en lo que respecta al fortalecimiento de la capacidad analítica, en Blank y Grosh (1999).

[21] Hay siempre cierta preocupación por mantener la separación entre los encargados de recopilar los datos y de analizarlos. Cuando se es menos riguroso en ese sentido, pueden presentarse problemas de credibilidad.

como las del ENMV, junto con los datos administrativos y de proyectos[22]. Una evaluación reciente de un proyecto del Banco Interamericano de Desarrollo-Banco Mundial-Comisión Económica para América Latina y el Caribe (CEPAL) para mejorar las encuestas de hogares[23] subraya el carácter a largo plazo de la sostenibilidad y plantea la cuestión adicional de la transición del financiamiento de los donantes al financiamiento local, que también debe tenerse en cuenta[24].

[22] La creación del Consorcio de Estadísticas para el Desarrollo en el Siglo XXI (PARIS21) en apoyo de la mejora de los datos para ese fin subraya la importancia de la recopilación sostenible de datos, de su análisis y utilización.

[23] El proyecto del Banco Interamericano de Desarrollo-Banco Mundial-CEPAL se denomina "Programa para la mejora de las encuestas y la medición de las condiciones de vida en América Latina y el Caribe", pero se conoce por su sigla en español: MECOVI.

[24] Véase Ryten (2000).

D. Costos de realización de una encuesta del EMNV

25. La atención a la calidad repercute gravemente en los costos, tanto en lo que se refiere al tiempo como a los recursos de las encuestas. Los costos de una encuesta del EMNV oscilan entre 400.000 dólares y 1,5 millones, según el país y el año. En cifras anuales, el costo es comparable a los de otras encuestas complejas, como las de ingresos y gastos y las demográficas y de salud. Naturalmente, los costos varían en función de la capacidad de la ONE, el nivel de infraestructura estadística, los objetivos de la encuesta y la dificultad de desplazamiento por el país. Los costos son notablemente más bajos cuando el organismo encargado de la ejecución tiene ya una buena infraestructura y personal experimentado. Los fondos para cada encuesta suelen proceder de distintas fuentes: fondos públicos (para la ONE o de otros organismos), donaciones bilaterales y donaciones y créditos multilaterales. En algunos casos el sector privado ha financiado también parte de los costos de la encuesta[25].

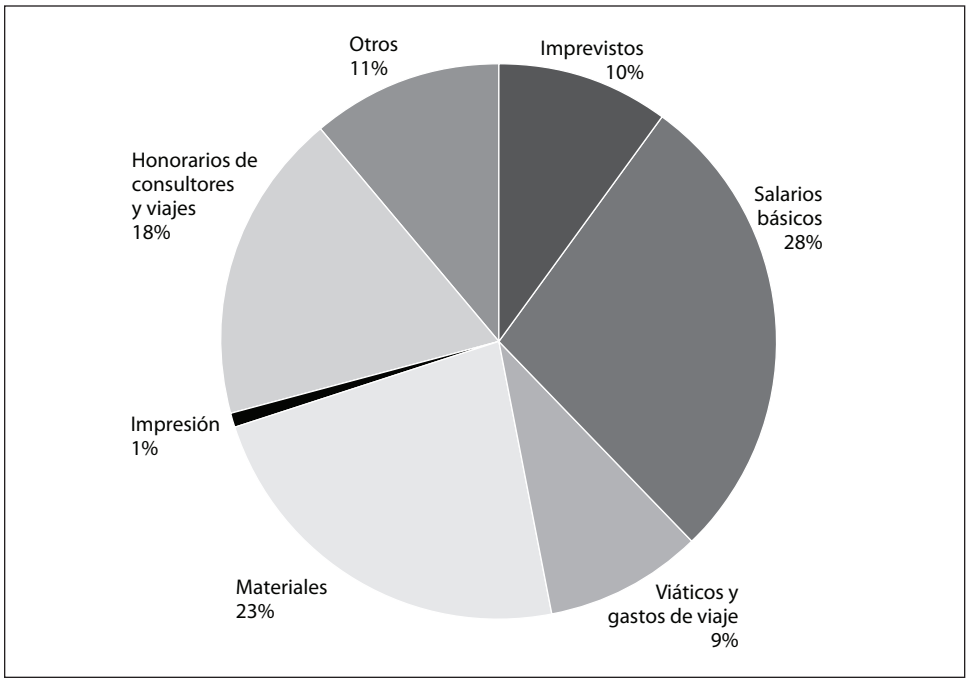
26. En general, el costo de una encuesta del EMNV depende de los métodos adoptados, el tamaño de la muestra y la complejidad del trabajo sobre el terreno. En la figura XXIII.1 pueden verse los componentes de costos de una encuesta del EMNV y la ponderación relativa de cada uno de ellos[26] (en el anexo II puede verse una práctica sencilla para ayudar al lector a preparar el presupuesto de una encuesta del EMNV).

27. La partida mayor de los costos corresponde a los salarios. Casi las tres cuartas partes se destina al personal sobre el terreno: entrevistadores, supervisores, encargados de la entrada de datos, antropometristas y conductores. El personal sobre el terreno de las en-

[25] Por ejemplo, en el Perú, una parte del cuestionario se reserva a empresas privadas o a investigadores que pagan para incluir preguntas específicas en algún lugar del cuestionario.

[26] Esta síntesis sobre los costos está basada en Grosh y Muñoz (1996) y en una presentación de Juan Muñoz en el curso del Banco Mundial sobre pobreza y desigualdad, 26-28 de febrero de 2002.

Figura XXIII.1
Componentes de costos de una encuesta del EMNV (parte del costo total)



Fuente: Basado en Grosh y Muñoz (1996), cuadro 8.2.

cuestas del EMNV es numeroso (en relación con el tamaño de la muestra) debido a la alta proporción supervisor-entrevistador (normalmente 1 sobre 3), el tamaño del cuestionario y el uso de informantes directos —tres elementos que limitan el número de hogares que pueden visitarse diariamente—, la necesidad de incluir la entrada de datos en los equipos en el terreno y el transporte para los miembros del equipo, para garantizar la movilidad y la integridad de sus miembros. Otros salarios corresponden al personal de la oficina: normalmente son funcionarios de la ONE, aunque a veces se contrata a un coordinador de proyecto externo.

28. La segunda mayor partida corresponde a materiales y equipo, lo que incluye las computadoras y los vehículos (compra o alquiler) y el mantenimiento, así como otro material de oficina. Ésta es la partida que más varía respecto a la infraestructura existente en la ONE o el organismo de ejecución. Las fuentes de financiamiento pueden incrementar los costos si están prohibidas las compras de vehículos: el alquiler de los vehículos necesarios puede ser algunas veces significativamente más costoso.

29. La asistencia técnica es el tercer componente principal. También en este caso hay grandes diferencias en función de la capacitación y experiencia disponibles en el organismo de ejecución. Los países que realizan sus segundas o terceras encuestas del EMNV requieren obviamente mucha menos asistencia técnica y equipo. Normalmente, los tipos de especialización que necesitan más asistencia técnica son el muestreo, el diseño de cuestionarios, la adaptación de la entrada de datos, la organización del trabajo sobre el terreno y las técnicas de análisis.

30. Naturalmente, los costos de una encuesta del EMNV están justificados si el resultado es realmente una mejor calidad de los datos, que podrían utilizarse para mejorar las políticas. Las encuestas del EMNV puede que sean costosas en términos absolutos, pero no si se considera la magnitud del gasto en políticas sociales, no son costosas. En el siguiente apartado se ofrecen muestras de encuestas recientes del EMNV sobre la calidad de los datos. Se presentan ejemplos en relación con la omisión de datos, su utilidad para los fines del EMNV, la coherencia interna y los efectos del diseño.

E. ¿Hasta qué punto ha influido en la calidad el diseño del EMNV?

1. Tasas de respuesta

31. Un primer indicador de la calidad es la tasa global de respuesta a la encuesta: ¿responden a la encuesta los hogares seleccionados en la muestra, son muchos los hogares no incluidos, con el consiguiente peligro de sesgo de los resultados finales?[27]. El examen de las tasas de respuesta es útil, ya que constituyen un indicador de la calidad de la capacitación, el diseño del cuestionario y los entrevistadores así como de los procedimientos de selección de la muestra (enumeración, actualización de mapas, etcétera). Las encuestas del EMNV de muchos países han conseguido tasas de respuesta notablemente elevadas. En el cuadro XXIII.4 pueden verse las tasas de respuesta de encuestas del EMNV recientemente terminadas. No obstante, estas encuestas no se libran de los posibles efectos de la situación específica de cada país. En los países que han atravesado un conflicto, donde los niveles previstos de confianza son bajos, las tasas de respuesta han sido también inferiores, como demuestran las de Bosnia y Herzegovina y Kosovo. En cambio, las bajas tasas de respuesta de Jamaica quizá ilustren mejor la importancia de los mecanismos de control de calidad. Jamaica ha utilizado menos técnicas estándar del EMNV para las actividades sobre el terreno y menos medidas del control de calidad: ello parece haberse traducido en tasas de respuesta más bajas[28]. En Guatemala, la baja tasa de respuesta se debió probablemente al prolongado período transcurrido entre la fecha de terminación de la enumeración de hogares y la de realización efectiva de la encuesta. En las últimas entrevistas ese lapso varió entre nueve a diez meses.

[27] Las tasas de participación hacen referencia al hogar en su conjunto, y no a los miembros individuales del mismo.

[28] Véase Banco Mundial (2001), donde se informa con más detalle acerca de las actividades sobre el terreno de la encuesta sobre condiciones de vida de Jamaica.

Cuadro XXIII.4
Tasas de respuesta en encuestas recientes del EMNV

País	Año	Número de viviendas elegidas	Tamaño efectivo de la muestra	Tasa de respuesta [a] (porcentaje)
Bosnia y Herzegovina	2001	5 400	5 402	82,6
Ghana [b]	1998-1999	6 000	5 998	97,4
Guatemala	2000	8 940	7 468	83,5
Jamaica	1999	2 540	1 879	74,0
Kosovo	2000	2 880	2 880	82,0
Kirguistán	1998	2 987	2 979	99,7
Nicaragua	1998-1999	4 370	4 209	96,3
Tayikistán	1999	2 000	2 000	..
Viet Nam	1997/98	5 994	5 999	93,9

[a] Bosnia y Herzegovina, Ghana, Kosovo, Tayikistán y Viet Nam utilizaron hogares de reemplazo. La tasa de respuesta estaba basada en el número de entrevistas terminadas menos el de hogares de reemplazo, dividido por el tamaño previsto de la muestra. En Bosnia y Herzegovina se utilizaron 938 hogares de reemplazo; en Ghana, 155; en Kosovo, 519; en Viet Nam, 372. Los autores no han podido determinar el número de hogares de reemplazo en Tayikistán.

[b] En la encuesta de Ghana se realizaron siete visitas a cada hogar. El tamaño final de la muestra es el número de hogares que participó en las siete visitas.

Nota: Dos puntos suspensivos (..) indican que faltan datos.

2. Falta de respuesta sobre partidas concretas

32. El cálculo del porcentaje de falta de respuesta sobre partidas concretas es otro indicador de calidad. Un examen del tema en las tres primeras encuestas del EMNV reveló que el problema ha sido casi irrelevante (faltaba menos del 1% de las respuestas en 10 variables clave)[29]. Es también interesante comparar las tasas de falta de respuesta sobre partidas concretas en las encuestas del EMNV con las obtenidas en otras encuestas que no cuentan con los mismos mecanismos de control de calidad. Aunque eso no siempre es posible; a continuación se presenta una pequeña comparación. Un examen de 1998 sobre las encuestas de población activa en América Latina compiló información sobre la frecuencia de los valores omitidos en relación con los ingresos de los trabajadores asalariados, los empleados por cuenta propia y los empleadores[30]. Tres países latinoamericanos realizaron también encuestas del EMNV antes de transcurrir un año de haberse realizado las encuestas de población activa. En la mayor parte de las comparaciones, como muestra el cuadro XXIII.5, las encuestas del EMNV consiguieron en esos países resultados notablemente mejores, o al menos iguales, que las encuestas de población activa. Aunque se trata sólo de un ejemplo limitado, parece demostrar el efecto positivo de la inversión del EMNV en los controles de calidad.

33. Para determinar la calidad global, sería preferible comprobar hasta qué punto se pueden utilizar los datos, en lugar de calcular el número de faltas de respuesta. En las encuestas del EMNV cuyo objetivo principal es medir el bienestar, lo más pertinente es determinar hasta qué punto los datos recogidos son adecuados para ese fin. El indicador monetario más frecuentemente utilizado para determinar el bienestar, por sus ventajas teóricas y prácticas, es el consumo total de los hogares. Se trata de una medida compleja que requiere datos de

[29] En Grosh y Glewwe (1995).

[30] En Feres (1998). Si bien los ingresos no constituyen el centro principal de las encuestas de población activa ni de las del EMNV, en ambas se recopila información sobre los ingresos de forma semejante.

Cuadro XXIII.5
Frecuencia de la ausencia de datos sobre los ingresos en el EMNV y en las encuestas de población activa

País	Encuesta	Porcentaje de ausencia de datos sobre ingresos de			Porcentaje de informantes directos
		Asalariados	Por cuenta propia	Empleadores	
Ecuador	LFS, 1997	6,3	6,7	13,2	..
	EMNV, 1998	3,6	8,5	6,5	96,5
Nicaragua	LFS (urbana) 1997	1,0	1,4	5,7	..
	EMNV, 1998	1,1	1,0	4,7	84,6
Panamá	LFS 1997	2,9	36,2	26,0	..
	EMNV, 1996	1,0	3,5	8,4	98,7

Fuente: La información acerca de las encuestas de población activa (LFS) está tomada de Feres (1998); en cuanto a las encuestas del EMNV, los cálculos han sido realizados por los autores.

Nota: En Nicaragua, en la encuesta EMNV de 1998, el porcentaje de falta de datos no incluyó los ceros, ya que según instrucciones recibidas por los entrevistadores el código debía ser cero si la persona no recibía ingresos en efectivo sino en especie. En esta categoría se incluían los agricultores de subsistencia, cuyos ingresos se calculaban en otro lugar del módulo sobre producción agrícola.

Dos puntos suspensivos (..) indican que faltan datos.

distintos módulos del cuestionario, tanto en el plano individual como en lo que se refiere a los hogares. En general, los datos de consumo se toman de los módulo de vivienda (valor de uso de la vivienda, servicios públicos y otros gastos de la vivienda), de bienes duraderos (para calcular el valor del flujo de servicios), de educación (enseñanza privada, gastos de bolsillo), de consumo de alimentos (comprados, de producción doméstica y recibidos en donación), de agricultura (para los alimentos de producción casera consumidos en el hogar si no se incluyen en el módulo de consumo de alimentos) y de gastos no alimentarios (partidas diversas, que van desde el jabón hasta los enseres domésticos).

34. En el cuadro XXIII.6 *infra* puede verse el porcentaje de hogares en los que se pudo establecer el total agregado de consumo. En la mayor parte de las encuestas muy pocos hogares tuvieron que excluirse del análisis por falta de datos. La excepción es Ghana. No está claro cuál fue el principal problema en su encuesta de 1998: la muestra era un poco mayor que otras, pero la diferencia no era tan grande como en el caso de Guatemala. Quizá influyera el hecho de que algunos datos sobre el consumo de alimentos se recopilaron a través de un diario (en vez de utilizar la metodología habitual del EMNV). Por desgracia, en la documentación de la encuesta no se considera este tema[30].

[30] Véase Ghana Statistical Service (2000).

3. Comprobaciones de coherencia interna

35. El logro de la coherencia interna de los datos es también de importancia crucial. El hecho de que la complejidad de los instrumentos de la encuesta dificulte su supervisión por los entrevistadores durante el proceso de la entrevista explica por qué hay tantos controles de calidad relativos a la coherencia. El cuadro XXIII.7 (en la página siguiente) muestra tres ejemplos de comprobación de la coherencia interna. El primero muestra la vinculación entre el cuestionario de la comunidad y los datos de los hogares. En el segundo, el porcentaje de niños de edad preescolar o escolar, tal como aparecen en la enumeración, con información completa sobre su escolarización/preescolarización. En el tercero se determina si las personas identificadas como trabajadores por cuenta propia en el módulo de población activa han facilitado detalles sobre sus actividades en el módulo de empresas domésticas no agrícolas.

36. Como se desprende del cuadro, las dos primeras comprobaciones revelan que la calidad de los datos ha sido muy elevada. En cambio, en la tercera se observan algunos problemas. Ello indica una falta de controles adecuados sobre el terreno entre las dos visitas a los hogares. Sólo en el caso de Viet Nam se incluyó en la segunda visita una pregunta expresa para el entrevistador, a fin de garantizar que el módulo se compilara. Se necesita una comprobación semejante en todas las encuestas.

Cuadro XXIII.6
Hogares con agregados completos de consumo. Ejemplos de encuestas recientes del EMNV

País	Año	Tamaño de la muestra final	Hogares con agregado completo de consumo (*porcentaje*)
Bosnia y Herzegovina	2001	5 402	99,9
Ghana	1998-1999	5 998	87,7
Guatemala	2000	7 468	97,4
Jamaica	1999	1 879	99,8
Kosovo	2000	2 880	100,0
Kirguistán	1998	2 979	99,4
Nicaragua	1998-1999	4 209	96,0
Tayikistán	1999	2 000	100,0
Viet Nam	1997-1998	5 999	100,0

Cuadro XXIII.7
Coherencia interna de los datos: relaciones correctas entre módulos (porcentaje) [a]

País	Vinculación correcta entre:			
	Encuesta de hogares y encuesta de la comunidad [a]	Enumeración y módulo de educación [b]		Módulo de empleo y módulo de empresas domésticas no agrícolas [c]
		Preescolar	Primaria	
Bosnia y Herzegovina	..	99,5	99,8	90,4
Ghana	99,9	..	96,5	70,2
Guatemala	100	100	100	93,0
Jamaica	..	..	96,4	..
Kosovo	100	..	100	58,6
Kirguistán	100	86,5	98,4	93,1
Nicaragua	..	97,9	97,5	62,0
Tayikistán	100	..	99,9	..
Viet Nam	100	..	99,6	98,1

Notas: Este cuadro hace referencia al porcentaje de vinculaciones correctas. Bosnia y Herzegovina, Jamaica y Tayikistán no incluyeron cuestionarios de la comunidad. Jamaica, Kosovo, Tayikistán y Viet Nam no incluyeron un módulo especial sobre la educación preescolar. Jamaica y Tayikistán no recopilaron información sobre las empresas domésticas no agrícolas.

Dos puntos suspensivos (..) indican que faltan datos.

[a] Comparación de los hogares con las comunidades en que estaban ubicados.

[b] Comparación de la variable de la edad tomada de la enumeración con la presencia de individuos en el módulo de educación.

[c] Comparación de quienes dijeron trabajar por cuenta propia en el módulo de empleo con la presencia de información en el módulo de empresas domésticas no agrícolas.

4. Efectos del diseño muestral

37. Un criterio final para juzgar las encuestas del EMNV se refiere al tamaño y al diseño de la muestra. Al utilizar los datos de una encuesta de hogares basada en un diseño complejo con múltiples etapas, estratificación y conglomeración, la verdadera varianza de las estimaciones se calcula teniendo en cuenta estas características del diseño de la muestra y la ponderación. El efecto del diseño es el coeficiente entre la varianza verdadera de una estimación, teniendo en cuenta la muestra en varias etapas, y la varianza de la estimación que se habría obtenido de haberse utilizado una muestra aleatoria simple del mismo tamaño[31]. Así pues, un efecto de diseño 1 indica que no ha habido ninguna pérdida de precisión en las estimaciones de la muestra como consecuencia del uso de un diseño en varias etapas, mientras que un efecto de diseño superior a 1 revela que el uso del diseño en varias etapas ha reducido la eficiencia de la muestra y la precisión de las estimaciones.

[31] Véase el anexo III del presente capítulo, u otros capítulos de esta publicación, donde hay información adicional sobre las cuestiones del diseño muestral.

38. En el marco de las actividades del EMNV se llevó a cabo un examen de los efectos del diseño en las principales variables e indicadores en algunas de las primeras encuestas del EMNV. El examen, realizado por Temesgen y Morganstein (2000), destacó varios puntos clave que deben tenerse en cuenta al utilizar los datos de las encuestas del EMNV (y los de otras encuestas de hogares que utilizan diseños muestrales en varias etapas) y al diseñar las muestras oportunas[32]. Lo más importante que cabe señalar es que la inclusión de varios temas en las encuestas del EMNV complica el proceso de diseño de una muestra eficiente. Los efectos del diseño varían enormemente en las variables relativas tanto a los individuos como a los hogares, como puede observarse en el cuadro XXIII.8, tomado del estudio de Temesgen y Morganstein (2000). En resumen, la reducción del efecto del diseño de una

[32] Varios de los cuadros del informe de Temesgen y Morganstein se incluyen en el anexo III de este capítulo.

Cuadro XXIII.8
Ejemplos de efectos del diseño en las encuestas del EMNV

País	Consumo per cápita			Acceso a servicios de salud			Tasa de desempleo		
	Total	Rural	Urbano	Total	Rural	Urbano	Total	Rural	Urbano
Côte d'Ivoire, 1988	6,7	3,6	5,5	6,3	5,7	2,2	7,0	4,4	5,7
Ghana, 1987	1,9	3,1	1,8	2,9	3,0	5,0	1,7	1,5	2,0
Ghana, 1988	3,2	2,9	2,9	2,2	2,5	3,6	1,3	1,1	1,4
Pakistán, 1991	1,6	1,1	2,6	5,0	4,0	5,2	4,6	4,7	2,5

Fuente: Temesgen y Morganstein (2000).

variable puede dar lugar a un aumento del mismo en otras variables. En segundo lugar, es claro que se debe buscar una solución de compromiso entre errores muestrales y no muestrales. Los efectos del diseño pueden ser elevados en las encuestas del EMNV. En el cuadro se observa que, en la medida en que las encuestas del EMNV se utilizan para producir medias, coeficientes y estimaciones puntuales, es trascendental que el diseño de la muestra se tenga en cuenta y que se preste atención al uso adecuado de los datos.

39. Como ocurre con otras encuestas, es interesante observar que el efecto del diseño varía no sólo entre variables sino también geográficamente dentro de un país en relación con la misma variable y con una variable determinada a lo largo del tiempo. Finalmente, los efectos del diseño pueden presentar enormes diferencias entre los distintos países. Un examen atento de las correlaciones intraclase y de los efectos del diseño en encuestas anteriores, cuando existen, ayudará a mejorar el diseño en futuras encuestas del EMNV. Debe procederse con cautela al presentar e interpretar los resultados del EMNV y otras encuestas que utilizan muestras en varias etapas, ya que el diseño muestral utilizado puede ser complejo.

F. Usos de los datos de las encuestas del EMNV

40. A lo largo de los años, los datos de las encuestas del EMNV se han utilizado para una gran variedad de actividades de formulación de políticas y de investigación. Algunos de ellos se han incluido en otros estudios[33]. En el sitio web del EMNV puede encontrarse una bibliografía amplia, aunque parcial, de estudios e informes basados en los datos de encuestas del EMNV. En esa bibliografía se observa el alcance de la utilización de los datos del EMNV para fines analíticos, pero su utilización es más amplia de lo que puede desprenderse de esa enumeración. La existencia de investigaciones en curso y de revisiones y enmiendas de los cuestionarios significa que el uso de los datos está cambiando constantemente. Como demostración de las distintas formas en que se han utilizado los datos del EMNV o se han combinado con otras informaciones, quizá valga la pena centrarse en un uso determinado —la orientación selectiva de los programas gubernamentales hacia los pobres— en vez de tratar de examinar a fondo los distintos usos que se les da.

[33] Véase Grosh (1997), por ejemplo.

41. En primer lugar, un ejemplo temprano de Jamaica demuestra como un análisis sencillo puede dar al gobierno información clara sobre los efectos de la orientación hacia los pobres utilizando programas alternativos. En el caso de Jamaica, como se observa en Grosh (1991), había tres importantes programas de nutrición: subvenciones generales a los alimentos, vales de alimentos y programas de alimentación escolar. La encuesta del EMNV en Jamaica permitió cuantificar el valor de los beneficios recibidos por los hogares pobres a través de los tres programas y demostró que la subvención alimentaria, a diferencia de los otros dos programas, era muy regresiva. Este análisis fue uno de los elementos que llevó a la decisión de eliminar las subvenciones y aumentar los recursos en favor de los otros dos programas.

42. Un segundo instrumento que puede crearse utilizando datos de encuestas del EMNV es la selectividad geográfica en favor de las zonas pobres. Aprovechando los datos del censo, las encuestas del EMNV pueden utilizarse para crear mapas de pobreza con el fin de asignar los recursos y los programas a las zonas pobres[34]. El método se basa en la existencia de una encuesta del EMNV y de datos del censo separados por pocos años[35]. La encuesta del EMNV es un indicador válido del bienestar (consumo total de los hogares) pero, debido al pequeño tamaño de la muestra, la capacidad de desglosar los datos resultantes acerca de la pobreza se limita sólo a las zonas urbanas y rurales, y a algunos otros grandes desgloses regionales del país. Obviamente, con ello no se cubren todas las necesidades de los gobiernos que tratan de orientar los recursos hacia las zonas pobres, ni ayuda en los sistemas descentralizados a asignar los recursos a los gobiernos locales. Además, en las grandes regiones hay muchas veces una gran heterogeneidad en cuanto a los niveles de pobreza de la población, aspecto que no aflora en una encuesta de hogares con una muestra pequeña.

[34] Puede verse más información sobre la metodología de creación de los mapas utilizando el indicador de bienestar de las encuestas y con los datos del censo, en Hentschel y otros (2000); Elbers, Lanjouw y Lanjouw (2002), Elbers y otros (2001) y Demombynes y otros (2001). Están en marcha nuevos estudios sobre el uso de esta técnica para vincular entre sí dos encuestas; aunque no es posible estimar los errores estándar correctos a partir de esa vinculación.

[35] Pueden utilizarse otras encuestas de hogares siempre que constituyan un indicador monetario del bienestar, como el consumo o el total de ingresos.

43. Para poder ofrecer información sobre la pobreza en forma más desglosada se requiere un conjunto de datos con una muestra que sea varias veces mayor que la de un EMNV. El conjunto más amplio de datos de un país es, naturalmente, el censo de población. No obstante, como abarca toda la población, el censo recopila información muy limitada de cada hogar y normalmente se realiza una vez cada diez años. Por ello no es posible establecer un indicador adecuado de la pobreza a partir del censo. Se está experimentando un planteamiento innovador que permite vincular los datos de las encuestas con los datos del censo. Esta técnica aprovecha el indicador del bienestar contenido en el EMNV y la cobertura del censo. El método supone una estimación de la pobreza en los datos de las encuestas del EMNV utilizando un vector de variables que se encuentra tanto en el censo como en la encuesta. Los parámetros estimados de esa manera se utilizan luego con datos del censo para prever la probabilidad de ser pobre de cada uno de los hogares y creando coeficientes de recuento para las pequeñas áreas a partir de los datos del censo. Los mapas de pobreza resultantes son un instrumento que ayuda al gobierno asignar los recursos con más equidad. Pueden encontrarse ejemplos de estos mapas de pobreza en el Ecuador, Guatemala, Madagascar, Nicaragua, Panamá y Sudáfrica.

44. Un tercer ejemplo del uso de datos de las encuestas del EMNV para mejorar la orientación de los programas sociales está basado en una evaluación del Fondo de Inversión Social de Emergencia (FISE) en Nicaragua. La evaluación examinó los problemas de orientación selectiva y los efectos de las inversiones del FISE en las comunidades en los sectores del abastecimiento de agua, saneamiento, educación, salud y alcantarillado[36]. En este caso se planificó una encuesta EMNV de alcance nacional. Se incluyó una sobremuestra de hogares integrada por hogares de zonas del proyecto FISE así como de comunidades semejantes no incluidas en el proyecto. La otra fuente de datos estuvo compuesta por registros administrativos y del proyecto utilizados para evaluar los costos administrativos del proyecto.

[36] Véase Banco Mundial (2000), donde se presenta información detallada sobre los objetivos de la evaluación, los métodos empleados y los resultados.

45. La sobremuestra de hogares de FISE y comunidades semejantes no incluidas en el FISE permitió la creación de grupos de control y de tratamiento para medir los efectos de las inversiones del FISE y la eficacia de su orientación selectiva. Además, la muestra nacional de la encuesta del EMNV permitió crear un segundo grupo de control (utilizando técnicas de correspondencia de la propensión) que aumentaron el ámbito de la evaluación. La valoración de la eficacia de la orientación selectiva se llevó a cabo en el plano comunitario (¿fueron las inversiones del FISE lo suficientemente progresivas en lo que respecta a las comunidades donde se realizaron los proyectos?) e individual (dentro de las comunidades con proyectos del FISE, ¿era mayor la probabilidad de que los segmentos más pobres de la población se beneficiaran de las inversiones del FISE?).

46. La evaluación mostró con valores estadísticamente significativos la eficiencia global de la selectividad y de evaluar los tipos principales de proyectos en función de criterios de selectividad. El estudio reveló que los proyectos de alcantarillado eran muy regresivos, mientras que los de saneamiento y de educación primaria eran sistemáticamente progresivos y llegaban al 17% de la población clasificada como extremadamente pobre. El resultado inmediato de la evaluación fue la suspensión de los proyectos de alcantarillado y la decisión de mejorar los contactos con las comunidades extremadamente pobres y las inversiones en ellas. El costo de esta evaluación muy compleja del proyecto del FISE representó el 1% de las inversiones realizadas por el proyecto hasta la fecha en que se realizó la evaluación.

G. Conclusiones

47. Los resultados de las encuestas del EMNV han demostrado el valor de este planteamiento. Los gobiernos han utilizado los datos para comprender el efecto de las políticas actuales, reformarlas y orientar mejor los recursos hacia determinados grupos y zonas. La insistencia en la calidad ha dado frutos en forma de reducción de errores y mayor utilidad de

los datos. No obstante, este planteamiento implica ciertas soluciones de compromiso. Los costos son relativamente altos, el menor tamaño de la muestra limita el nivel de desglose que se puede obtener y la planificación y diseño iniciales requieren mucho tiempo; no obstante, pueden obtenerse datos con rapidez una vez que ha comenzado el trabajo y las asociaciones con los responsables de las políticas aumentan el uso de los datos.

48. La incorporación de las encuestas del EMNV al sistema de encuestas de hogares de un país tiene sus ventajas. La frecuencia con que se necesitará esa encuesta depende de varios factores. En primer lugar, las necesidades analíticas del país deben impulsar la decisión de realizar una o varias encuestas a lo largo del tiempo. Si bien muchos programas gubernamentales pueden evaluarse con datos transversales (selectividad, incidencia, repercusiones utilizando técnicas de puntuación de correspondencia de la propensión), se necesitan datos transversales y datos de panel para otros tipos de análisis de los cambios a lo largo del tiempo y los efectos de las políticas y acontecimientos.

49. Una segunda consideración en lo que respecta a la frecuencia de realización de las encuestas del EMNV es la referente a la capacidad de análisis del país. Los datos deben analizarse como la mira de que sirvan a los responsables de las políticas y para conseguir que cada futura ronda de la encuesta mejore aprovechando las conclusiones de la ronda anterior. Si no es posible analizar los datos con rapidez, gran parte de la inversión en la encuesta o en su repetición puede resultar estéril. En tal caso quizá tenga sentido dejar un margen de tiempo significativo (tres años, por ejemplo) entre encuestas.

50. Finalmente, las cuestiones presupuestarias y logísticas son muchas veces tan importantes como las sustantivas a la hora de decidir con qué frecuencia o en qué momento realizar encuestas específicas. Así pues, la frecuencia con que se debe realizar una encuesta responderá a un equilibrio entre la importancia de sus resultados y las de otras encuestas. Asimismo, es importante recordar que ninguna fuente de datos es suficiente para cubrir todas las necesidades. Los registros administrativos y los datos del sistema de información sobre la gestión de los proyectos, así como un sistema de encuestas de hogares, son elementos necesarios para los gobiernos, tanto se trate de aplicarlos a la política macroeconómica como a la microeconómica. Las encuestas del EMNV, unidas a una política de encuestas frecuentes y generales, ayudarán a comprender cómo repercuten las políticas y gastos gubernamentales en la vida de su población.

ANEXO I

Encuestas del Estudio de medición de los niveles de vida efectuadas en varios países

País	Año	Recuento de hogares
Albania	1996	1 500
Albania	2002	3 600
Armenia	1996	4 920
Azerbaiyán	1995	2 016
Bolivia	1999	..
Bolivia	2000	5 032
Bolivia	2001	..
Bosnia y Herzegovina	2001	5 402
Brasil	1996-1997	4 940
Bulgaria	1995	2 500
Bulgaria	1997	2 317
Bulgaria	2001	2 633
Camboya	1997	6 010
China: Hebei y Liaoning	1995 y 1997	780
Côte d'Ivoire	1985	1 588
Côte d'Ivoire	1986	1 600
Côte d'Ivoire	1987	1 600
Côte d'Ivoire	1988	1 600
Ecuador	1994	4 500
Ecuador	1995	5 500
Ecuador	1998	5 801
Ecuador	1998-1999	5 824
Federación de Rusia [a]	1992	6 500
Gambia	1992	1 400
Ghana	1987-1988	3 200
Ghana	1988-1989	3 200
Ghana	1991-1992	4 565
Ghana	1998-1999	5 998
Guatemala	2000	7 276
Guinea	1994	4 705
Guyana	1992-1993	5 340
India: Uttar Pradesh y Bihar	1997-1998	2 250
Jamaica	1988-2000 (anual)	2 000-7 300
Kazajstán	1996	1 996

País	Año	Recuento de hogares
Kosovo	2000	2 880
Kirguistán	1993	2 000
Kirguistán	1996 (primavera)	..
Kirguistán	1996 (otoño)	1 951
Kirguistán	1997	2 962
Kirguistán	1998	2 979
Madagascar	1993	4 504
Malawi	1990	6 000
Mauritania	1987	1 600
Mauritania	1989	1 600
Mauritania	1995	3 540
Marruecos	1991	3 323
Marruecos	1998	..
Nepal	1996	3 373
Nicaragua	1993	4 200
Nicaragua	1998-1999	4 209
Nicaragua	2001	4 290
Níger	1989	1 872
Níger	1992	2 070
Níger	1995	4 383
Pakistán	1991	4 800
Panamá	1997	4 945
Papua Nueva Guinea	1996	1 396
Paraguay	1997-1988	4 353
Paraguay	1999	5 101
Paraguay	2000-2001	8 131
Perú	1985	5 120
Perú (Lima solamente)	1990	1 500
Perú	1991	2 200
Perú	1994	3 500
República Unida de Tanzanía: Kagera	1991-1994	840
República Unida de Tanzanía: nacional	1993	5 200
Sudáfrica	1993	9 000
Tayikistán	1999	2 000
Túnez	1995-1996	3 800
Uganda	1992	9 929
Viet Nam	1992-1993	4 800
Viet Nam	1997-1998	5 999

Nota: Dos puntos suspensivos (..) indican que faltan datos.

[a] La Encuesta de seguimiento longitudinal de Rusia de 1992 se realizó con financiamiento del Banco Mundial. Las encuestas posteriores no contaron con participación del Banco Mundial. Puede encontrarse más información en el sitio web de Carolina Population Center: http://www.cpc.unc.edu/projects/rlms/rlms_home.html.

ANEXO II

Presupuestación de una encuesta del EMNV

No hay dos encuestas del EMNV que sean exactamente iguales, como tampoco lo son dos ONE, ni los costos asociados con los salarios, el transporte, el equipo, etcétera, en diferentes países. Por ello es imposible decir con detalle cuánto puede costar una encuesta del EMNV en un lugar específico en un momento dado. En este capítulo se ha presentado un ejemplo de la parte correspondiente a los diferentes tipos de costos en el total de una encuesta. A continuación se presenta un pequeño ejercicio cuyo objetivo es ayudar a poner en marcha un presupuesto. Se trata simplemente de una guía rápida para estimar los costos salariales más básicos correspondientes al trabajo sobre el terreno. Utilizando esta guía con los costos reales del país pertinente, se puede llegar a una idea aproximada de lo que podría costar la encuesta.

Por término medio, según la complejidad de la encuesta y los informantes directos, un entrevistador puede realizar dos medias entrevistas por día (véase el esquema XXIII.2, página 466). En otras palabras, puede terminar una ronda de la encuesta en dos hogares. Si suponemos una semana de trabajo de seis días (independientemente de que el día libre se tome cada semana o se distribuya de alguna otra forma a lo largo del mes), un entrevistador puede ocuparse de 24 hogares por mes.

Supongamos que se necesita una muestra de 4.000 hogares. Si cada entrevistador puede completar 24 hogares por mes, se necesita un total de 167 meses-entrevistador para entrevistar a los 4.000 hogares. Si el trabajo sobre el terreno tiene lugar a lo largo de 12 meses, harán falta 14 entrevistadores. Por cada par de entrevistadores se necesita un supervisor, un encargado de la entrada de datos y un conductor y un automóvil. Así pues, el personal total sobre el terreno (sin contar la supervisión regional realizada por el personal de la ONE) comprende las siguientes personas:

14 entrevistadores
7 supervisores
7 encargados de la entrada de datos
7 conductores

Si los encargados de la planificación utilizan los parámetros indicados a continuación, los costos salariales de la parte de la encuesta correspondiente al trabajo sobre el terreno serán los siguientes:

Nota: Si bien el trabajo sobre el terreno dura sólo 12 meses, se agrega un mes adicional para sufragar los costos de capacitación (durante el cual el personal sobre el terreno recibe alguna forma de remuneración) y/o los posibles retrasos en la realización de la encuesta. Los encargados de la entrada de datos suelen mantenerse durante un mes más para finalizar y depurar el conjunto de datos, en caso necesario.

Partida	Costo por individuo por mes	Número de meses	Costo
14 entrevistadores	500	13	91 000
7 supervisores	575	13.5	54 338
7 encargados de la entrada de datos	525	14	51 450
7 conductores	300	13	27 300
Estimación aproximada de los costos salariales sobre el terreno			224 088

Según la figura XXIII.1, los costos del personal sobre el terreno, que suman las tres cuartas partes del total de los costos salariales de la encuesta, representan a su vez el 28% de los costos de la encuesta. En este caso, con un sencillo cálculo se llega a una estimación aproximada del costo de la encuesta, que sería de 1.067.086.

Indudablemente, esta cifra es sólo una estimación muy aproximativa. Se necesitan detalles sobre otros costos, como los correspondientes a la asistencia técnica, etcétera. No obstante, este simple intento inicial puede ser útil para comenzar el proceso de presupuestación de una encuesta real. El lector puede consultar el capítulo 8 de Grosh y Muñoz (1996), donde se presenta detalladamente la manera de diseñar un presupuesto realista para una encuesta del EMNV.

ANEXO III

Efecto del diseño muestral en la precisión y eficiencia de las encuestas del EMNV[1]

[1] Este anexo está basado en gran parte en el estudio anterior de Temesgen y Morganstein (2000).

A. Introducción

En otros capítulos de esta publicación se da información detallada sobre las cuestiones relativas al muestreo y, en particular, el efecto de los diseños muestrales complejos o en varias etapas en la varianza de las estimaciones obtenidas. Este efecto del diseño es frecuente en todas las encuestas que no utilizan una muestra aleatoria simple, como es el caso de las encuestas del EMNV. El efecto del diseño es una parte del error muestral global: la diferencia entre una estimación obtenida de un diseño por conglomerado en varias etapas y el que se obtendría con un diseño muestral aleatorio simple. En este anexo resumimos las cuestiones clave y describimos el efecto real del diseño muestral en varias encuestas del EMNV.

B. Cálculo de los errores muestrales, efectos del diseño y componentes afines

En una muestra aleatoria simple todas las unidades incluidas tienen una probabilidad idéntica e independiente de selección. El muestreo aleatorio simple no se utiliza casi nunca en las encuestas de hogares, debido a problemas logísticos y de costos. Como en las encuestas del EMNV, se utilizan más bien diseños muestrales más complejos en varias etapas que incorporan la estratificación y la conglomeración. Ello repercute en el cálculo de la varianza de las estimaciones y en la eficiencia de la muestra misma. Para calcular los errores de muestreo correspondientes a diseños muestrales que se aplican en más de una etapa es necesario conocer las variables que identifican los estratos, las unidades primarias de muestreo (UPM) y los procedimientos de ponderación (en su caso) utilizados en el diseño. Una vez identificadas estas variables, pueden utilizarse distintos programas de estadística para calcular las medidas necesarias[2].

[2] En este caso se ha utilizado el programa WESVAR. Otros programas que pueden utilizarse para estimar las varianzas de muestreo y algunas estadísticas afines en los diseños de encuestas complejas son CENVAR, CLUSTERS, Epi-Info, PC CARP, SUDAAN, VPLX y STATA. Algunos de estos programas pueden descargarse gratuitamente de la web.

Las medidas del error de muestreo aquí señaladas para algunas variables relativas a los hogares e individuos en las encuestas del EMNV incluyen el error estándar (*SE*), que se calcula teniendo en cuenta la complejidad del diseño muestral, el coeficiente de variación (*CV* (%)), el tamaño de la muestra (*n*), el efecto del diseño, el coeficiente de la correlación intraclase (ρ), los límites inferior y superior de los intervalos de confianza (*CI*) y el tamaño efectivo de la muestra (*EFFn*). Estos términos se definen en los capítulos II, VI, VII y en otros lugares de esta publicación.

C. Errores estándar, efectos del diseño y correlación intraclase calculados a partir de las encuestas del EMNV

Un aspecto importante al calcular los errores muestrales de las variables de una encuesta es la comparación de las eficiencias (precisión) de los diseños muestrales entre sí y con la precisión que se habría conseguido con una muestra aleatoria simple hipotética del mismo tamaño. Además de indicar la fiabilidad de los datos existentes, este ejercicio puede también ayudar a los analistas a evaluar el comportamiento de un diseño determinado y ofrecer información para el diseño de futuras encuestas. Los cuadros A.III.1, 2 y 3 comparan los efectos del diseño y medidas afines en relación con diversas variables para hacer patentes las diferencias que existen dentro de un país según las distintas variables, dentro de un país a lo largo del tiempo y entre países diferentes[3].

[3] Puede verse el informe completo en Temesgen y Morganstein (2000).

Como muestra el cuadro AIII.1, dentro de un país la misma encuesta genera efectos del diseño notablemente diferentes cuando se trata de variables distintas. El cuadro está basado en datos procedentes del EMNV de Ghana (1987) y de variables referentes a los hogares e individuos. En el caso de algunas variables, como el gasto total per cápita, donde la correlación intraclase es baja, el efecto del diseño no es elevado (1,9); en cambio, en variables como el acceso al saneamiento y el abastecimiento de agua, donde las correlaciones intraclase son altas (la infraestructura tiende a estar concentrada en áreas espaciales específicas), los efectos del diseño son considerables (7,8 y 8,0, respectivamente), y son aún mayores cuando se trata de subpoblaciones urbanas o rurales.

Cuadro AIII.1
Variación de los efectos del diseño por variable, Ghana, 1987

Variable		Estimación	*SE*	*CV* (%)	Intervalo de confianza		*n*	Efecto del diseño	*EFFn*	*ρ*
					Inferior	Superior				
Acceso a la electricidad	Total	0,267	0,019	7,265	0,229	0,305	3 138	6,034	520	0,300
	Rural	0,078	0,022	28,744	0,034	0,121	2 023	14,063	144	0,787
	Urbano	0,611	0,041	6,714	0,530	0,691	1 115	7,888	141	0,403
Tamaño del hogar	Total	4,940	0,083	1,682	4,777	5,103	3 136	2,089	1 501	0,065
	Rural	5,147	0,097	1,877	4,958	5,336	2 022	1,735	1 165	0,044
	Urbano	4,565	0,165	3,615	4,241	4,888	1 114	3,291	339	0,134
Propiedad de la tierra	Total	0,591	0,024	4,018	0,544	0,637	3 138	7,315	429	0,376
	Rural	0,747	0,033	4,393	0,683	0,811	2 023	11,520	176	0,634
	Urbano	0,308	0,035	11,413	0,239	0,376	1 115	6,453	173	0,319
Gasto total per cápita	Total	82 745,2	1 902,2	2,3	79 017,1	86 473,4	3 104	1,883	1 648	0,053
	Rural	70 908,1	2 526,4	3,6	65 956,3	75 859,8	2 001	3,100	646	0,127
	Urbano	104 219,5	3 702,1	3,6	96 963,6	111 475,4	1 103	1,759	627	0,044
Gasto per cápita en alimentos	Total	56 779,3	1 309,2	2,3	54 213,2	59 345,3	3 104	1,927	1 611	0,055
	Rural	52 382,3	1 777,9	3,4	48 897,6	55 867,0	2 001	2,577	776	0,095
	Urbano	64 756,0	2 147,9	3,3	60 546,2	68 965,8	1 103	1,580	698	0,034
Eliminación de desechos	Total	0,019	0,003	16,647	0,013	0,026	3 135	1,724	1 818	0,043
	Rural	0,010	0,003	29,044	0,004	0,016	2 020	1,704	1 185	0,042
	Urbano	0,037	0,009	23,481	0,020	0,054	1 115	2,347	475	0,079
Acceso a retrete	Total	0,590	0,025	4,159	0,542	0,638	3 135	7,808	401	0,405
	Rural	0,659	0,034	5,091	0,593	0,725	2 020	10,114	200	0,549
	Urbano	0,465	0,038	8,092	0,392	0,539	1 115	6,357	175	0,313
Acceso a agua potable	Total	0,395	0,025	6,251	0,347	0,443	3 135	7,994	392	0,416
	Rural	0,224	0,031	13,818	0,164	0,285	2 020	11,150	181	0,611
	Urbano	0,704	0,046	6,482	0,615	0,793	1 115	11,144	100	0,593

Fuente: Temesgen y Morganstein (2000).
Nota: Véanse en los cuadros AIII.4 y AIII.5 descripciones de las variables utilizadas.

En el cuadro AIII.2, basado también en datos de Ghana, se observa que los efectos del diseño pueden variar a lo largo del tiempo y en función de la variable. En este caso la diferencia entre las dos encuestas es de un año y el diseño muestral básico no cambió, pero los efectos del diseño se modificaron: la estimación correspondiente al acceso a los servicios de salud mejoró notablemente de precisión (el efecto del diseño bajó de 5,01 a 3,64); en el caso del desempleo disminuyó también, aunque no tanto. La otra variable del cuadro, la alfabetización de adultos, se midió con menos precisión en el segundo año de la encuesta.

Finalmente, como era de prever, los efectos del diseño en los distintos países pueden presentar significativas variaciones. En el cuadro AIII.3 se observa cómo las encuestas de Côte d'Ivoire y de Pakistán produjeron efectos del diseño bastante diferentes en relación con las mismas variables. Este resultado dependía tanto de los diferentes diseños muestrales utilizados en los países como de las diferentes características de dichos países.

Cuadro AIII.2
Variación de los efectos del diseño a lo largo del tiempo, Ghana, 1987 y 1988

Variable		Estimación	*SE*	*CV* (%)	Intervalo de confianza		*n*	Efecto del diseño	*EFFn*	*ρ*
					Inferior	Superior				
Ghana, 1987										
Alfabetización de adultos	Mujer	0,402	0,021	5,103	0,362	0,442	1 339	2,342	572	0,080
	Hombre	0,613	0,018	2,953	0,578	0,649	1 381	1,910	723	0,054
	Total	0,509	0,016	3,192	0,477	0,541	2 720	2,875	946	0,112
Acceso a los servicios de salud	Mujer	0,443	0,016	3,625	0,411	0,474	2 756	2,876	958	0,112
	Hombre	0,423	0,017	4,017	0,390	0,457	2 542	3,011	844	0,120
	Total	0,433	0,015	3,517	0,403	0,463	5 298	5,013	1 057	0,239
Desempleo	Mujer	0,039	0,004	10,063	0,031	0,047	4 011	1,655	2 424	0,039
	Hombre	0,047	0,004	9,136	0,038	0,055	3 543	1,454	2 437	0,027
	Total	0,043	0,003	7,666	0,036	0,049	7 554	1,983	3 810	0,059
Ghana, 1988										
Alfabetización de adultos	Mujer	0,390	0,022	5,526	0,348	0,432	1 289	2,519	512	0,090
	Hombre	0,587	0,020	3,397	0,548	0,626	1 226	2,013	609	0,060
	Total	0,486	0,018	3,654	0,451	0,521	2 515	3,179	791	0,130
Acceso a los servicios de salud	Mujer	0,375	0,013	3,558	0,348	0,401	2 921	2,215	1 319	0,072
	Hombre	0,365	0,015	4,118	0,335	0,394	2 606	2,539	1 026	0,092
	Total	0,370	0,012	3,346	0,346	0,394	5 527	3,635	1 521	0,157
Desempleo	Mujer	0,036	0,003	9,593	0,029	0,042	3 852	1,307	2 946	0,018
	Hombre	0,034	0,003	9,885	0,027	0,041	3 260	1,123	2 904	0,007
	Total	0,035	0,003	7,306	0,030	0,040	7 112	1,372	5 185	0,022

Fuente: Temesgen y Morganstein (2000).
Nota: Las descripciones de las variables utilizadas pueden verse en los cuadros AIII.4 y AIII.5, *infra*.

En resumen, el pequeño tamaño de las muestras utilizadas en las encuestas del EMNV y el que se trate de muestras en varias etapas obliga a buscar soluciones de compromiso en lo que respecta a la precisión de las estimaciones. Por ejemplo, el valor del efecto del diseño correspondiente a "Alfabetización de adultos" de todos los individuos en los datos de Côte d'Ivoire de 1988 es elevado: 6,7. Este efecto del diseño significa que la precisión de la estimación con un tamaño (n) de 1.660 es el equivalente al obtenido utilizando una muestra MAS de sólo 249. En cambio, si consideramos únicamente a los individuos "urbanos", observamos que el efecto del diseño es algo inferior (2,7), aunque todavía superior a 1, lo que significa que una muestra de 915 personas tiene el equivalente de precisión de 344 personas si se utiliza una MAS. El hecho de que el efecto del diseño pueda ser bastante considerable y la variación de dichos efectos en las distintas variables, momentos y países significa que los analistas deben reconocer y tener encuesta el diseño de la muestra al utilizar los datos y sobre todo al realizar comprobaciones estadísticas de importancia. Ello pone también de manifiesto las dificultades que plantea el diseño de muestras suficientes para encuestas de hogares sobre varios temas. Si se intenta reducir el efecto del diseño de una variable, el resultado podría ser un mayor efecto del diseño en otras. Una norma práctica es considerar ante todo la variable o variables de mayor importancia en la encuesta, en la medida de lo posible.

Cuadro AIII.3
Variación de los efectos del diseño en los diferentes países

Variable		Estimación	*SE*	*CV* (%)	Intervalo de confianza		*n*	Efecto del diseño	*EFFn*	*ρ*
					Inferior	Superior				
Côte d'Ivoire, 1988										
Alfabetización de adultos	Total	0,567	0,031	5,538	0,506	0,629	1 660	6,676	249	0,378
	Rural	0,411	0,042	10,212	0,329	0,493	745	5,415	138	0,294
	Urbano	0,738	0,024	3,217	0,691	0,784	915	2,661	344	0,111
Acceso a los servicios de salud	Total	0,417	0,029	6,883	0,361	0,473	1 849	6,260	295	0,351
	Rural	0,303	0,034	11,174	0,236	0,369	1 051	5,693	185	0,313
	Urbano	0,622	0,025	4,078	0,572	0,671	798	2,181	366	0,079
Desempleo	Total	0,038	0,007	18,837	0,024	0,052	4 979	6,991	712	0,399
	Rural	0,007	0,003	50,457	0,000	0,013	2 529	4,357	580	0,224
	Urbano	0,081	0,013	16,218	0,055	0,107	2 450	5,679	431	0,312
Pakistán, 1991										
Alfabetización de adultos	Total	0,5	0,013	2,5	0,48	0,53	6 834	4,335	1 577	0.222
	Rural	0,42	0,017	3,95	0,39	0,45	3 249	3,669	885	0.178
	Urbano	0,68	0,018	2,616	0,64	0,71	3 585	5,156	695	0.277
Acceso a los servicios de salud	Total	0,5	0,012	2,329	0,48	0,52	9 238	5,02	1 840	0.268
	Rural	0,46	0,015	3,177	0,43	0,49	4 752	4,048	1 174	0.203
	Urbano	0,61	0,017	2,74	0,57	0,64	4 486	5,185	865	0.279
Desempleo	Total	0,03	0,003	9,735	0,02	0,03	18 232	4,633	3 935	0.242
	Rural	0,02	0,003	14,955	0,02	0,03	8 934	4,706	1 898	0.247
	Urbano	0,03	0,003	8,956	0,03	0,04	9 298	2,539	3 662	0.103

Fuente: Temesgen y Morganstein (2000).
Nota: Las descripciones de las variables utilizadas pueden verse en los cuadros AIII.4 y AIII.5, *infra*.

Cuadro AIII.4
Descripción de las variables del análisis: nivel individual

Variable	Descripción	Base de la población
Desempleo	Adultos actualmente desempleados pero dispuestos a trabajar y que buscan un empleo	Personas de 15 a 64 años de edad
Acceso a los servicios de salud	Proporción de individuos que estuvieron enfermos durante el mes previo a la entrevista y que visitaron centros modernos de salud, como hospitales, dispensarios y centros de salud (pero no comadronas, curanderos u otros profesionales de medicina tradicional)	Personas que estuvieron enfermas durante el mes anterior
Alfabetización de adultos	Proporción de adultos alfabetizados (es decir, que podían leer un periódico)	Personas de 15 a 24 años de edad

Cuadro AIII.5
Descripción de las variables del análisis: hogares

Variable	Descripción
Acceso a agua potable	Proporción de hogares que tienen acceso a agua potable. En los hogares esta variable adopta un valor de 1 si el hogar obtiene su agua potable, por ejemplo, de un grifo, una tubería o un pozo con bomba. El valor es 0 si la fuente de agua potable para el hogar —un río, canal, pozo abierto, lago o pantano— se considera peligrosa para la salud.
Propiedad de tierra	Proporción de hogares que poseen tierras. En un hogar esta variable adopta un valor de 1 si el hogar posee tierras. En caso contrario, el valor es 0.
Acceso a la electricidad	Proporción de hogares que tienen acceso a la electricidad. En un hogar esta variable adopta un valor de 1 si el hogar utiliza electricidad para la iluminación y/o energía. De lo contrario, el valor es 0.

REFERENCIAS BIBLIOGRÁFICAS

Banco Mundial (2000). *Nicaragua: Ex-Post Impact Evaluation of the Emergency Social Investment Fund (FISE)*. Report No. 20400-NI. Washington, D.C.

_____ (2001). *Jamaica Survey of Living Conditions (JLSC) 1988-98: Basic Information*. Washington, D.C.: Development Economics Research Group.

_____ (2002a). *Basic Information Document: Bosnia and Herzegovina Living Standards Measurement Study Survey*. Washington, D.C.: Development Economics Research Group.

_____ (2002b). *Guatemala Poverty Assessment*. Report No. 24221-GU. Washington, D.C.

_____ (2002c). *The 1993 Nicaragua Living Standards Measurement Survey: Documentation*. Washington, D.C.: Development Economics Research Group.

Blank, Lorraine y Margaret Grosh (1999). Using household surveys to build analytic capacity. *The World Bank Research Observer*, vol. 14, No. 2 (agosto), págs. 209-227.

Demombynes, Gabriel y otros (2001). *Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries*. WIDER Discussion Paper, No. 2002/39. Helsinki: Instituto Mundial de Investigaciones de Economía del Desarrollo/ Universidad de las Naciones Unidas.

Elbers, C., J. Lanjouw y P. Lanjouw (2002). *Micro-level Estimation of Welfare*. Policy Research Working Paper, No. 2911. Washington, D.C.: Banco Mundial.

_____ (2003). Micro-level estimation of poverty and inequality. En *Econometrica,* vol. 71, No. 1, págs. 355-364.

Elbers, C. y otros (2001). *Poverty and Inequality in Brasil: New Estimates from Combined PPV-PNAD Data*. Washington, D.C.: Development Economics Research Group, Banco Mundial.

Feres, Juan Carlos (1998). Falta de respuesta a las preguntas sobre el ingreso: su magnitud y efectos en las Encuestas de hogares en América Latina. En *Conference Proceedings from the Segundo Taller Regional del Medición del Ingreso en las Encuestas de Hogares*, Buenos Aires. Noviembre. Santiago de Chile: Comisión Económica para América Latina y el Caribe (CEPAL). Document LC/R.1886.

Ghana Statistical Service (2000). *Ghana Living Standards Survey, Round Four (GLSS 4) 1998-99: Data User's Guide.* Accra.

Grosh, Margaret (1991). *The Household Survey as a Tool for Policy Change: Lessons from the Jamaican Survey of Living Conditions*. Living Standards Measurement Study Working Papers, No. 80. Washington, D.C.: Banco Mundial.

_____ (1997). The policymaking uses of multi-topic household survey data: a primer. *The World Bank Research Observer*, vol. 12, No. 2, págs. 137-160.

_____ y Paul Glewwe (1995). *A Guide to Living Standards Measurement Study Surveys and Their Data Sets.* Living Standards Measurement Study Working Paper, No. 120. Washington, D.C.: Banco Mundial.

_____, comps. (2000). *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study Surveys.* Washington, D.C.: Banco Mundial.

Grosh, Margaret y Juan Muñoz (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper, No.126. Washington, D.C.: Banco Mundial.

J. Hentschel y otros (2000). Combining household data with census data to construct a disaggregated poverty map: a case study of Ecuador. *World Bank Economic Review*, vol. 14, No. 1 (enero).

Kish, Leslie (1965). *Survey Sampling*. Nueva York: John Wiley and Sons, Inc.

Pradhan, Menno y Martin Ravallion (2000). Measuring poverty using qualitative perceptions of consumption adequacy. *Review of Economics and Statistics*, vol. 82, págs. 462-471.

Ravallion, Martin y Michael Lokshin (2001). Identifying welfare effects using subjective questions. *Economica*, vol. 68, págs. 335-357.

_____ (2002). Self-rated economic welfare in Russia. *European Economic Review*, en prensa.

Ryten, Jacob (2000). The MECOVI Program: ideas for the future: a mid-term evaluation. Documento inédito preparado para el Banco Interamericano de Desarrollo. Diciembre.

Skinner, C. J., D. Holt y T. M. F. Smith (1989). *Analysis of Complex Surveys*. Chichester, Reino Unido: John Wiley and Sons.

Temesgen, Tilahun y David Morganstein (2000). *Measurement of Sampling Errors: Application to Selected Variables in LSMS Surveys*. Washington, D.C.: Development Economics Research Group, Banco Mundial.

Capítulo XXIV
Diseño de la encuesta y diseño de la muestra en las encuestas sobre el presupuesto de los hogares

Hans Pettersson
Statistics Sweden
Estocolmo
Suecia

RESUMEN

En el presente capítulo se consideran algunas cuestiones relativas al diseño de las encuestas y al diseño muestral para las encuestas sobre el presupuesto de los hogares. Se insiste sobre todo en las encuestas realizadas en países en desarrollo. Los problemas de medición del consumo y del ingreso se examinan con cierto detalle en el apartado B. El apartado C contiene un análisis sobre algunas cuestiones fundamentales del diseño muestral, como la estratificación y la asignación muestral en el espacio (geográfica) y en el tiempo (a lo largo de toda la campaña). El apartado D contiene una descripción de la Encuesta de gasto y consumo de la República Democrática Popular Lao, de 1997/98 (LECS-2). En el apartado E se examinan algunas de las experiencias relacionadas con dicha encuesta.

Términos clave: Encuesta sobre el presupuesto de los hogares, encuesta sobre gasto y consumo, medición de los gastos, método del diario.

A. Introducción

1. El término "encuesta sobre el presupuesto de los hogares" es una expresión genérica para una amplia categoría de encuestas. Pueden recibir el nombre de "encuestas sobre el gasto familiar", "encuestas sobre gasto y consumo" o "encuestas de ingresos y gastos", pero el elemento común es el intento de reflejar partes importantes del "presupuesto" cotidiano del hogar. Algunas encuestas concebidas originalmente como encuestas sobre el presupuesto de los hogares han incorporado luego otros objetivos. Al núcleo fundamental de las preguntas sobre el consumo, los gastos y los ingresos de los hogares se han agregado módulos adicionales relativos, por ejemplo, a la salud, la nutrición y la educación. Esta práctica de integrar varios temas en una sola encuesta se está generalizando. En el presente capítulo se insiste sobre todo en las encuestas que incluyen entre sus elementos importantes la medición del presupuesto del hogar, independientemente de si se trata de una encuesta sobre varios temas o de una encuesta más especializada sobre el presupuesto.

2. Los datos sobre el consumo, el gasto y los ingresos de los hogares pueden servir para varios fines. Los datos de las encuestas pueden utilizarse en diversos estudios sobre las características socioeconómicas de la población y su distribución (por ejemplo, la difusión de la pobreza). Cuando las encuestas se realizan de forma periódica pueden emplearse para

supervisar el bienestar de diversos grupos de población. Las encuestas del Estudio de medición de los niveles de vida (EMNV) del Banco Mundial se han concebido con la finalidad específica de medir las diferencias en cuanto a la pobreza y el nivel de vida de la población. En los últimos años se ha observado gran interés por el uso de encuestas como medio para evaluar los resultados de las intervenciones gubernamentales, en particular los efectos de los proyectos de reducción de la pobreza; esos datos pueden utilizarse también para orientar la formulación de políticas relacionadas con el bienestar y con aspectos fiscales.

3. Los datos de las encuestas sobre el presupuesto de los hogares constituyen una aportación muy importante al sistema nacional de estadísticas económicas y, en particular, a las cuentas nacionales. Estas encuestas miden el consumo en el sector de los hogares y pueden reflejar también la producción de los establecimientos y operaciones agrícolas de los hogares (que representan una parte considerable de la producción nacional en los países pobres). En el sistema de estadísticas económicas se hace hincapié en los agregados nacionales. Una encuesta que atienda ante todo a las necesidades del sistema de estadísticas económicas debería concebirse de tal manera que ofreciera estimaciones de los totales nacionales. Este sistema puede resultar en algunos casos menos eficiente cuando los datos de las encuestas se utilizan para análisis orientados a las políticas y para la evaluación de las intervenciones, en cuyo caso el interés se centra en las diferencias entre los distintos grupos de población o áreas geográficas.

4. En este capítulo se insiste sobre todo en las encuestas sobre el presupuesto en los países en desarrollo, en cuanto fuente de datos para el sistema de estadísticas económicas. El capítulo está dividido en cuatro apartados principales. En el apartado B se consideran algunos problemas importantes relativos al diseño de la encuesta, sobre todo los difíciles problemas de medición, en particular la del consumo del hogar. En el apartado C se examinan las cuestiones del diseño muestral para las encuestas sobre el presupuesto de los hogares. En el apartado D se considera la Encuesta sobre gasto y consumo de la República Democrática Popular Lao, 1997/98 (LECS-2), como estudio de caso. En el apartado E se analizan las experiencias y enseñanzas aprendidas de la Encuesta sobre gasto y consumo de la RDP Lao.

B. Diseño de la encuesta

1. Métodos de recopilación de datos en las encuestas sobre el presupuesto de los hogares

5. El principal objetivo de las encuestas sobre el presupuesto de los hogares es medir el consumo total de los hogares y sus componentes. El planteamiento tradicional del problema de la medición, todavía utilizado en muchas encuestas, es recopilar información detallada. Se pide al hogar que registre por separado las compras de un gran número de artículos, tanto en cantidades físicas como en unidades monetarias. Otro planteamiento es limitar la recopilación de los datos de consumo a una lista de artículos menos detallada. Es el planteamiento normalmente adoptado por los estudios del nivel de vida del Banco Mundial (Living Standard Surveys (Deaton, 1997)).

6. Los datos sobre consumo pueden recopilarse básicamente de dos maneras:

- Mediante entrevistas de los hogares consistentes en preguntas retrospectivas referentes al consumo.
- Mediante un diario en el que el hogar registra el consumo y el gasto de cada día.

7. El método del diario requiere normalmente al menos dos visitas al hogar, una al comienzo y otra al final del período. Muchas veces se programa una visita intermedia para comprobar que las anotaciones del diario se realizan en forma adecuada. La entrevista retrospectiva podría consistir en una única visita al hogar, pero es habitual realizar dos.

2. Problemas de medición

8. ¿Cómo debería medirse el consumo de un hogar en una entrevista con preguntas retrospectivas? ¿Debería medirse en forma detallada y en relación con un gran número de artículos o sin gran detalle? El primer planteamiento produce informaciones más pormenorizadas y precisas que el segundo, pero con un costo bastante mayor. Si podemos prescindir de los detalles y tratamos fundamentalmente de estimar el consumo total, ¿permitirá el segundo planteamiento, con un pequeño número de preguntas, obtener estimaciones tan precisas como las conseguidas con el cuestionario detallado? No hay conclusiones definitivas acerca de la precisión. Deaton cita estudios de los últimos años, entre ellos una encuesta de prueba realizada en 8.000 hogares de Indonesia en que se comprobaron dos cuestionarios (Deaton, 1997). El cuestionario largo tenía 218 partidas alimentarias y 102 no alimentarias, mientras que en el cuestionario breve había 15 partidas alimentarias y 8 no alimentarias. Las estimaciones del gasto total en alimentos presentaban pocas diferencias entre ambos cuestionarios. Las estimaciones de los gastos no alimentarios eran aproximadamente un 15% mayores en el cuestionario largo (Banco Mundial, 1992, apéndice 4.2). No obstante, estos resultados no se han reproducido en otras pruebas. Ensayos semejantes realizados en El Salvador (Joliffe y Scott, 1995) y Jamaica (Statistical Institute and Planning Institute of Jamaica 1996, apéndice III) revelan mayores diferencias entre los cuestionarios. Los gastos totales eran un 40% superiores, y los gastos en alimentos un 27% mayores en el cuestionario largo en la prueba de El Salvador. En la prueba de Jamaica se obtuvieron gastos totales un 26% superiores en el cuestionario largo. Deaton formula la siguiente conclusión: "Aunque el cuestionario más breve puede dar lugar en algunos casos a una reducción espectacular de los costos y de la duración de una encuesta —en Indonesia de 80 minutos a 10—, parece que ese ahorro implica como contrapartida una pérdida de precisión" (Deaton, 1997).

9. El método del diario reduce la dependencia de los recuerdos de los informantes. No obstante, es difícil de utilizar cuando una parte considerable de la población es analfabeta. Incluso con altas tasas de alfabetización, es de prever que surjan problemas; por ejemplo, es menos probable que los hogares pobres puedan utilizar diarios, y muchos de los que podrían usarlos no los utilizan de hecho (Deaton y Grosh, 2000). La Oficina General de Estadística de Viet Nam comprobó que en las zonas urbanas muchos hogares no completaron los diarios en la Encuesta de hogares con varios objetivos de 1995 (Glewwe y Yansaneh, 2001). La duración del período de información es también un aspecto que se debe considerar si se opta por los diarios; en muchas encuestas se utilizan períodos de dos semanas, pero algunas abarcan todo un mes. Las investigaciones indican que la información sobre los gastos es menos completa al pasar de la primera a la segunda semana en los diarios de dos semanas, probablemente debido a la fatiga.

10. Muchas encuestas sobre el presupuesto de los hogares recopilan también datos sobre los ingresos. La medición de los ingresos del hogar presenta desafíos todavía mayores que la medición del consumo. El ingreso es un tema delicado para muchos informantes, sobre todo en las zonas menos necesitadas. Algunas veces existe cierta suspicacia entre los informantes, que temen que su información pueda utilizarse con fines fiscales, sobre todo en los casos en que el hogar cuenta con una empresa familiar.

11. Deben registrarse los ingresos de todos los miembros del hogar y todos los tipos de ingresos (ingresos procedentes de empresas familiares o de la agricultura, ingresos informales de actividades a tiempo parcial, rendimiento de los activos, etcétera). Los cálculos de los ingresos se complican por la existencia de donaciones en efectivo y en especie, remesas y préstamos. Los ingresos procedentes de la agricultura en los hogares de pequeños propietarios presentan problemas especiales, ya que estos hogares obtienen parte de sus alimentos de la producción de subsistencia. Asimismo, parte de los ingresos en efectivo pueden proceder de la venta de productos agrícolas, que tiene lugar de manera intermitente, lo que dificulta el registro adecuado de los ingresos durante la entrevista.

12. Es probable —en algunos casos es un hecho comprobado— que estas dificultades teóricas y prácticas de medición de los ingresos del hogar den lugar a una infravaloración de los mismos. Así lo demuestra lo ocurrido en algunas encuestas sobre ingresos y gastos. Muchas veces se observa que las estimaciones de los ingresos tal como aparecen en las encuestas son notablemente inferiores a las estimaciones del consumo; tan inferiores, que resulta difícil explicar toda esa diferencia atribuyéndola al ahorro de los hogares para el consumo. La otra posible explicación —la sobrevaloración del consumo— es menos probable. Las investigaciones indican que es mucho mayor la probabilidad de que el consumo esté infravalorado que sobrevalorado. Por ello hay razones para pensar que muchas estimaciones de los ingresos en las encuestas son demasiado bajas.

3. Períodos de referencia

13. Estrechamente relacionada con la decisión sobre el instrumento de medición (cuestionario "largo" o "corto", método del diario para el consumo de alimentos o preguntas de recuerdo, etcétera) está la relativa al período de referencia que deberán recordar los informantes. Conviene que no sea demasiado largo, para reducir en lo posible los errores de recuerdo. En un experimento del estudio sobre los niveles de vida en Ghana se investigó el efecto que podría tener la prolongación del período de referencia. El resultado fue que en el caso de 13 partidas frecuentes los gastos registrados disminuían por término medio un 2,9% por cada día agregado al período de recuerdo (Scott y Amenuvegbe, 1990). Hay algo de controversia entre los investigadores acerca de los efectos de la modificación de los períodos de recuerdo. En un estudio anterior sobre la Encuesta nacional por muestreo de la India parecía deducirse que, en relación con determinados alimentos, un período de referencia de un mes producía menos sesgo que un período de referencia de una semana (Mahalanobis y Sen, 1954). Los estudios sobre el nivel de vida realizados en los últimos años parecen confirmar los resultados de Scott pero no está claro si ello se debe a problemas de recuerdo cuando se trata de datos de un período prolongado o a efectos de delimitación (compresión) en los datos sobre períodos inmediatos (Deaton, 1997).

14. Las partidas de mayor frecuencia, como los alimentos, normalmente tienen períodos de referencia más bien breves, no superiores a un mes. La situación es diferente en los artículos menos frecuentes. El recuerdo de los gastos en este tipo de productos, como los bienes duraderos del hogar, debe abarcar un período relativamente más largo, porque un plazo demasiado breve daría lugar a grandes varianzas en las estimaciones de los totales. Así pues, la longitud del período de referencia adecuado variará según los grupos de artículos.

4. Frecuencia de las visitas

15. La mayor parte de las encuestas sobre ingresos y gastos recopilan datos mediante visitas repetidas a los hogares de la muestra. La frecuencia de las visitas a cada hogar dependerá del método de medición. El procedimiento estándar para el método retrospectivo es realizar dos visitas, separadas aproximadamente por dos semanas. En las encuestas que utilizan el método del diario las visitas de seguimiento deberían estar separadas entre sí una o dos semanas.

16. Las visitas repetidas al mismo hogar pueden provocar fatiga en el informante, lo que daría lugar al deterioro de la calidad de su información. Las ventajas del seguimiento de un hogar durante un período de tiempo más largo y el control de la calidad de los datos con visitas frecuentes al mismo hogar deben sopesarse teniendo también en cuenta la fatiga que ello puede producir.

17. Otro tipo de encuesta con visitas repetidas es aquella en que se entrevista al hogar durante dos o más períodos de referencia distribuidos a lo largo del año. Un ejemplo es la

Encuesta sobre ingresos, consumo y gastos de los hogares de Etiopía, de 1995/96, en que los hogares se visitaban dos veces en dos temporadas diferentes y se les hacían preguntas sobre el mes anterior. Esta situación se examina más adelante y más a fondo cuando se trata sobre el muestreo.

5. Falta de respuesta

18. Una característica distintiva de las encuestas sobre el presupuesto de los hogares es que pueden representar una pesada carga para los hogares que deben responder. La tasa de rechazo en general es mayor en las encuestas de presupuesto que en otras, y puede ser muy elevada en algunos segmentos de la población. A los rechazos iniciales hay que sumar los abandonos durante la encuesta. Es probable que la tasa de abandono sea mayor en esta encuesta que en otras, debido a la fatiga o a la incomodidad experimentados por el hogar cuando el entrevistador realiza visitas repetidas y sondeos detallados sobre los ingresos y los gastos.

19. No hay buenos estudios comparativos sobre los niveles de falta de respuesta en las encuestas sobre presupuesto de los países en desarrollo. Los estudios del EMNV tienen tasas de falta de respuesta de menos del 20% (Deaton y Grosh, 2000), que son considerablemente más bajas que las experimentadas en las encuestas sobre el presupuesto de los hogares en Europa occidental, donde los niveles pueden llegar a tasas de entre el 40% y el 50%. Probablemente hay una gran variación en las tasas de respuesta entre los países en desarrollo. En los países con fuerte control administrativo en el plano comunitario local, la tasa de falta de respuesta será probablemente baja.

C. Diseño de la muestra

20. Los requisitos del diseño de la muestra en el caso de una encuesta sobre presupuesto no difieren mucho de los existentes en otros tipos de encuestas de hogares. Normalmente se emplea una muestra en varias etapas, y las unidades primarias de muestreo (UPM) son zonas de empadronamiento (ZE) o unidades administrativas, como comunas, aldeas o barrios. En este apartado se considerarán algunas cuestiones específicas del diseño muestral en las encuestas sobre el presupuesto.

1. Estratificación y asignación de la muestra a los estratos

21. La estratificación de las UPM normalmente se lleva a cabo utilizando las regiones administrativas (provincias, etcétera) y, dentro de las regiones, las zonas urbanas/rurales. En lo que respecta a las encuestas sobre el presupuesto de los hogares, una mayor estratificación por nivel de ingresos hará posible una mayor eficiencia. En las ciudades y grandes urbes, normalmente es posible identificar entre dos y tres estratos de niveles de ingresos y establecer una clasificación aproximada de las UPM de acuerdo con dichos estratos (por ejemplo, zonas de ingreso alto, mediano y bajo).

22. Una encuesta sobre el presupuesto de los hogares tiene muchos usuarios con diferentes expectativas sobre los resultados, sobre todo en una encuesta de hogares con diversos objetivos, entre ellos el relacionado con el presupuesto. El encargado de la planificación muchas veces tiene que atender exigencias contradictorias de los principales usuarios. Un destino importante de los datos sobre el presupuesto de los hogares son las cuentas nacionales. Éstas requieren, ante todo y sobre todo, estimaciones nacionales fiables sobre los totales de las cuentas. Ello presupone un diseño en el que la muestra se asigne de manera uniforme entre toda la población (muestra autoponderada) o con una sobremuestra de los hogares de ingreso mediano y alto cuando el nivel de actividad económica sea más elevado.

23. Otros usuarios importantes son los encargados de la planificación gubernamental y del análisis de las políticas, que utilizan los datos para la planificación, la supervisión del bienestar y el análisis de la pobreza y que necesitan estimaciones fiables para diferentes partes del país y para diferentes grupos de población, más que estimaciones nacionales. La encuesta debería contar con un número suficiente de hogares en todas las regiones y grupos de población importantes (por ejemplo, hogares que viven en aldeas remotas o pobres). Ello supone un diseño que asigne la muestra de manera más o menos equitativa en las distintas regiones y, si es posible, que garantice una muestra adecuada de los grupos de población importantes.

24. Para atender esas exigencias contradictorias habrá que recurrir en muchos casos a soluciones de compromiso. Algunas veces lo que se hace en esos casos es utilizar la asignación cuadrática, en que la muestra se asigna a los estratos (regiones) en forma proporcional a la raíz cuadrada del tamaño del estrato (en lo que se refiere a la población o número de hogares). Esta asignación se ha utilizado en la Encuesta sobre niveles de vida de los hogares de Viet Nam y en las muestras para las encuestas de hogares en Sudáfrica.

2. Tamaño de la muestra

25. El tamaño total de la muestra en las encuestas de hogares varía de unos países a otros. En muchas encuestas el tamaño oscila entre 3.000 y 10.000 hogares, pero en los países grandes el número puede ser considerablemente mayor. Las autoridades locales a veces manifiestan una fuerte demanda de resultados con un alto grado de detalle geográfico, hasta el punto de poner en riesgo la calidad de los datos de la encuesta en algunos casos. Una muestra de gran tamaño puede "detraer" recursos de otro objetivo igualmente importante: mantener en niveles aceptables los errores no muestrales. El problema está en encontrar un equilibrio entre las demandas de los organismos administrativos subnacionales y los requisitos presupuestarios, a fin de mantener el tamaño de la muestra y los errores no muestrales en niveles manejables. Muchas veces el encargado del diseño de la encuesta se encuentra con la dificultad de explicar a los usuarios la necesidad de mantener un equilibrio entre errores muestrales y no muestrales.

3. Muestreo a lo largo del tiempo

26. Las pautas de ingresos y gastos de los grandes grupos de población pueden variar considerablemente en las diferentes estaciones. Si es posible, la encuesta debería prever una muestra suficiente para las distintas épocas. Debe prestarse especial atención a los grandes períodos de vacaciones, en que las pautas de consumo suelen desviarse considerablemente de las de otros períodos.

27. Una posible manera de resolver el problema de la estacionalidad es utilizar un período de referencia de un año. Como hemos visto, esta solución no es viable en la mayor parte de las partidas y no lo es ciertamente en el caso de los productos alimenticios. Otros planteamientos posibles son los siguientes:

- Visitas repetidas (con períodos de referencia repetidos) para los mismos hogares distribuidos a lo largo del año, con inclusión de todas las estaciones.
- Seguimiento del hogar (quizá con varias visitas) durante un período; por ejemplo, un mes. Los hogares se distribuyen a lo largo del año de acuerdo con un plan de muestreo en el que se garantiza una muestra suficiente en todas las estaciones. En este diseño se supone que al agregar datos transversales mensuales (multiplicados por 12) es posible reconstruir el año estadísticamente.

28. El segundo planteamiento ofrece probablemente la solución más común al problema de la estacionalidad. Se utiliza, entre otras, en las encuestas de gastos de la República Democrática Popular Lao, Namibia y Lesotho.

29. El primer planteamiento se ha empleado en la Encuesta sobre ingresos, consumo y gastos de los hogares de Etiopía, de 1995/96, en la que los hogares se visitaron dos veces en dos estaciones diferentes y se les hicieron preguntas sobre el mes anterior.

30. Con el segundo planteamiento se tiene en cuenta la variación estacional pero sólo en un nivel agregado. Por ese procedimiento se estimarán correctamente agregados como las medias y los totales de los ingresos o gastos totales de los hogares. No obstante, las medidas comunes de la dispersión estarán sesgadas. Los totales mensuales de los hogares individuales que se anualizan multiplicando esos totales por 12 contendrán una variación estacional (debido al hecho de que se estudia sólo un mes) y una variación no estacional aleatoria (debido al hecho de que el hogar tiene diferentes ingresos y gastos en los distintos meses que no son atribuibles a efectos estacionales). Esta variación estacional y no estacional en los totales mensuales analizados hace que la variación sea mayor de la que se habría obtenido si se hubieran observado los totales anuales. Por consiguiente, las estimaciones de la dispersión en los totales anuales estarán sesgadas si utilizamos como estimaciones las medidas de dispersión de los totales mensuales. La variación estacional puede estimarse a partir de los datos y utilizarse para reducir el sesgo. En cambio, no es posible reducir el sesgo debido a la variación dentro del hogar entre los distintos meses, ya que tenemos datos únicamente sobre un mes en cada hogar.

31. Para el analista interesado en la distribución de los gastos anuales entre los hogares, el diseño mensual presenta problemas debido al sesgo en las medidas ordinarias de dispersión (por ejemplo, la desviación estándar). Estos problemas repercuten en el análisis de la pobreza, en el que se identifican hogares individuales situados por encima o por debajo de un umbral de pobreza y se analizan las características de esos grupos. Si no se introducen correcciones, el alcance de la pobreza se sobrevalorará si menos de la mitad de la población es pobre, y se infravalorará si lo es más de la mitad. Scott demuestra mediante un modelo de cálculo que la desviación estándar de los gastos anuales está sobrevalorada un 36% en una encuesta que recopila datos de los hogares relativos a un solo mes (Scott, 1992).

D. Estudio de caso acerca de la Encuesta sobre gasto y consumo de la República Democrática Popular Lao, de 1997-1998

32. La República Democrática Popular Lao ha realizado dos encuestas sobre gasto y consumo en el último decenio. La primera (LECS-1) se llevó a cabo en 1992/93; la segunda (LECS-2), en 1997/98 (State Planning Committee, National Statistical Centre of Lao People's Democratic Republic, 1999). En el momento de la elaboración de este artículo se estaba realizando una tercera (LECS-3).

1. Condiciones generales para la realización de la encuesta

33. La República Democrática Popular Lao tenía 4,5 millones de habitantes en el último censo (1995). Su superficie es un poco mayor que la de Gran Bretaña. Las partes septentrionales y orientales son montañosas. El transporte es difícil en muchos lugares del país: el 57% de los hogares rurales vivían en aldeas que no tenían acceso a carreteras, según el censo de 1995. La República Democrática Popular Lao es todavía una sociedad predominantemente rural y agrícola. La abrumadora mayoría de la población trabaja por cuenta propia en la agricultura. La tasa de alfabetización de adultos es de aproximadamente el 60%. Aunque se hablan muchas lenguas, la lengua oficial, el lao, es comprendida por la mayor parte de la población. Las aldeas son unidades administrativas bien definidas y hay incluso una subdivisión formal dentro de las aldeas en "grupos de hogares", de 10-15 unidades. Según una evaluación aproximada (y subjetiva) de las condiciones del trabajo sobre el terreno, podría decirse que en la República Democrática Popular Lao, en comparación con un país

en desarrollo medio, es más difícil llegar a los hogares de las zonas rurales pero que, una vez que se ha llegado a ellos, es más probable que éstos cooperen.

2. Temas incluidos en la encuesta y en los cuestionarios

34. Una buena parte de los dos indicadores macroeconómicos "valor añadido" e "insumo de mano de obra en la producción" hacen referencia a la producción de los hogares en el sector de la agricultura o a sus actividades informales. Con el fin de captar los datos sobre producción de los hogares, en la segunda LECS se introdujeron tres nuevos módulos: *a*) un diario "sencillo", que se utilizó para registrar el uso del tiempo de un miembro del hogar, lo que permitiría medir la aportación de horas de mano de obra en la economía del país, y *b*) dos módulos sobre las operaciones económicas agrícolas y de los hogares. De esa manera se puede calcular el valor agregado de la producción de los hogares en la agricultura y en actividades comerciales informales.

35. En la primera entrevista se aplicaba un módulo general sobre composición del hogar, educación, empleo, fecundidad y nutrición infantil. Un módulo del diario permitía incluir todas las transacciones del hogar durante un mes. En una segunda entrevista se examinaban aspectos como la vivienda, el acceso a bienes duraderos, la tierra y el ganado. Las preguntas sobre la vivienda permitían imputar los valores del alquiler. Al final del mes se hacían preguntas sobre las compras de bienes duraderos durante los 12 meses precedentes. El cuestionario relativo a la aldea se administraba al jefe de cada una de ellas. En él se incluían elementos como las carreteras y el transporte, el abastecimiento de agua, la electricidad, los servicios de salud, los mercados locales, las escuela, etcétera.

3. Métodos de medición

36. El hecho de que en la primera LECS se utilizara el método del diario para medir las transacciones de los hogares aconsejaba la utilización de este método en la nueva LECS, siempre que hubiera funcionado de manera satisfactoria. El cambio del método de medición pondría en peligro la comparabilidad entre las encuestas. El método del diario había funcionado satisfactoriamente en la LECS-1 pero sólo con considerable apoyo a los hogares por parte de los entrevistadores. Muchos hogares no podían (o no querían) llevar debidamente el diario sin apoyo notable y frecuente del entrevistador. En tales circunstancias, este método parece ser una alternativa menos favorable. De todas formas, hay que considerar también el hecho de que muchas aldeas de la República Democrática Popular Lao son de difícil acceso. Una vez que el entrevistador se encuentra en la aldea, compensa que permanezca allí mientras duran las tres entrevistas necesarias para cada hogar, en vez de hacer que se desplace varias veces desde la aldea hasta la base de operaciones. Además, los entrevistadores podían mantener frecuentes contactos con los hogares durante su estancia en la aldea. El Centro Nacional de Estadística optó finalmente por el "método del diario con ayuda del entrevistador" para la LECS-2. Los entrevistadores permanecerían en la aldea durante todo un mes y darían a los hogares toda la asistencia necesaria para mantener el diario.

37. Se utilizó un procedimiento especial para cuantificar el consumo diario de arroz. Se midió el consumo de arroz de cada miembro del hogar durante un día para obtener una medida precisa de la ingestión en cada comida y de cada persona. Se mostraba al interesado una hoja con fotografías de seis platos con distintas cantidades de arroz (una "bola", dos "bolas", etcétera) y se le pedía que indicara qué fotografía era la que correspondía a la realidad.

38. A lo largo del mes, se seleccionó un período de 24 horas para registrar el uso del tiempo de los hogares. El diario utilizado en la LECS-2 había sido elaborado conjuntamente por Statistics Sweden y el Centro de Investigación sobre Cambios Microsociales del Consejo de Investigación Económica y Social de la Universidad de Essex. Un objetivo importante

era "aligerar" el diario, es decir, conseguir un formato que pudiera ser utilizado junto con otros instrumentos de la encuesta sin sobrecargar a los informantes. Se pedía a sólo uno de los miembros del hogar (seleccionado en forma aleatoria), de al menos 10 años de edad, que rellenara el diario sobre el uso del tiempo durante el día designado. El entrevistador elegía a los informantes al azar de manera que el número seleccionado cada día de la semana se mantuviera constante.

39. Este diario contenía 22 actividades previamente definidas, con especial insistencia en las actividades económicas. En el caso de algunas de ellas, el entrevistador buscaba información adicional en el momento de recopilación de los diarios. A quien respondía que había "trabajado como empleado" se le preguntaba si había trabajado en una explotación agrícola, en el sector gubernamental, en el sector privado o en algún otro lugar. A quien respondía "por cuenta propia" se le preguntaba qué papel había desempeñado en su negocio. Las respuestas se clasificaban de acuerdo con una lista con unas 50 categorías basadas en la Clasificación Industrial Internacional Uniforme de todas las Actividades Económicas (CIIU) y en el Sistema de Cuentas Nacionales, 1993 (Comisión de las Comunidades Europeas, Fondo Monetario Internacional, Organización de Cooperación y Desarrollo Económicos, Naciones Unidas y Banco Mundial (1993)).

4. Diseño de la muestra y trabajo sobre el terreno

40. Las zonas de empadronamiento sirven como unidades primarias de muestreo (UPM). Éstas se estratificaron en 18 provincias y, dentro de las provincias, en zonas urbanas y rurales. Las ZE rurales se estratificaron nuevamente en ZE con "acceso a carreteras" y "sin acceso a carreteras". Se asignó a cada provincia una muestra de 25 UPM. Se llevó a cabo una nueva asignación por zonas urbanas/rurales, en la que la parte urbana contaba con una fracción muestral un 50% mayor que a la de la parte rural. Las UPM se seleccionaron con un procedimiento sistemático de probabilidad proporcional al tamaño (PPT) en cada provincia, lo que dio lugar a una muestra de 450 UPM.

41. Los hogares de las UPM seleccionadas se enumeraron antes de la encuesta y se seleccionaron 20 hogares con muestreo sistemático en cada UPM, lo que arrojó una muestra total de 9.000 hogares. El muestreo a lo largo del tiempo se consiguió mediante una asignación aleatoria de la muestra provincial a lo largo del período de 12 meses, que dio lugar a dos aldeas (y en un caso tres) por mes.

42. Hacía falta un equipo de dos entrevistadores para el trabajo en las aldeas. Los entrevistadores se seleccionaron entre el personal permanente de las oficinas provinciales de estadística. Muchos habían participado en la primera LECS. La capacitación se llevó a cabo durante un período de dos semanas.

E. Experiencias y enseñanzas aprendidas

1. Métodos de medición y falta de respuesta

43. Los entrevistadores pasaron mucho tiempo en los hogares ayudando a los informantes a registrar todas las transacciones relativas al hogar y las actividades económicas y operaciones agrícolas del hogar. Hay razones para pensar que este trabajo tedioso y lento mejoró la calidad de las respuestas. Según algunas pruebas incidentales, las frecuentes visitas al hogar por parte del entrevistador establecieron en muchos casos una relación distendida y de confianza entre las partes. Además, los entrevistadores dispusieron de abundante tiempo para aclarar las relaciones, muchas veces complejas, entre consumo y producción de los hogares en la agricultura o las actividades económicas del hogar.

44. Se llevaron a cabo algunas comprobaciones de calidad. Las estimaciones del consumo de arroz resultantes de la encuesta se compararon con datos externos de la pro-

ducción agrícola y se comprobó que coincidían en términos generales. Se realizó también una comprobación sobre los niveles de consumo entre la primera y la segunda semana del diario. No había indicios de descenso de las anotaciones durante el segundo período y hubo poca diferencia en el número de entradas en el diario entre los dos períodos. Asimismo, las estimaciones del consumo total eran comparables entre los dos períodos.

45. El hecho de que hubiera diferencias muy pequeñas en el consumo en cifras agregadas entre el primero y el segundo período de dos semanas obliga a preguntarse si un diario más breve podría haber sido suficiente para registrar el consumo.

46. La tasa de falta de respuesta registrada fue baja: sólo el 3,1%. Más en concreto, fue muy baja en las zonas urbanas (sólo el 0,6%), y más elevada, pero todavía baja, en las zonas rurales (3,9%). La falta de respuesta estaba en cierto punto infravalorada. La sustitución se utilizó en algunos casos pero los procedimientos para registrar los resultados de la entrevista eran deficientes, por lo que es difícil evaluar el nivel exacto de falta de respuesta y diferenciar entre la falta de contacto y la negativa a colaborar. El número de rechazos fue muy bajo. Todas las experiencias de las encuestas de hogares de la República Democrática Popular Lao indican que los hogares se sienten obligados a participar en las encuestas gubernamentales. Además, el presidente de la aldea les pide que participen.

2. Diseño muestral y errores de muestreo

47. El análisis de la varianza y las estructuras de costos indica que el tamaño óptimo de la muestra de las UPM (zonas de empadronamiento) es del orden de 8-12 hogares. Así pues, el tamaño de muestra utilizado en la encuestas, 20 hogares, era superior al nivel óptimo (Pettersson, 2001).

48. Los cálculos indican también que la asignación equitativa de la muestra entre las provincias dio lugar a errores de muestreo en las estimaciones nacionales, que fueron aproximadamente un 20% superiores a lo que se habría conseguido con una asignación proporcional. Los coeficientes de variación fueron en general un 5% inferiores a las estimaciones de alcance nacional. La muestra de las zonas urbanas era menor que la de las zonas rurales (2.008 frente a 6.874 hogares), pero los coeficientes de variación de las estimaciones urbanas eran comparables a los de las estimaciones rurales, debido en parte a que los efectos del diseño fueron menores en las zonas urbanas.

49. Los efectos del diseño fueron relativamente elevados en las zonas rurales, considerablemente mayores que en las zonas urbanas (véase el cuadro XXIV.1). Ello se debió al hecho de que las aldeas rurales son socioeconómicamente homogéneas. Como la mayor parte de las UPM rurales constan de una aldea, las UPM serían también homogéneas. En las ciudades y grandes urbes hay relativamente poca segregación por nivel de ingresos en zonas ricas y pobres: los hogares ricos conviven con hogares pobres por toda la ciudad. Esa es la razón por la que muchas UPM urbanas contienen hogares ricos y hogares pobres, lo que hace que las UPM urbanas sean relativamente heterogéneas.

Cuadro XXIV.1
Efectos del diseño en el consumo de los hogares y posesión de bienes duraderos

	Nacional	Urbano	Rural
Consumo mensual total por hogar en kip	5,4	3,8	7,7
Consumo mensual de alimentos por hogar en kip	5,8	4,4	6,8
Proporción de hogares que poseen un vehículo automóvil	2,1	1,3	3,3
Proporción de hogares que poseen un televisor	5,4	3,1	6,8
Proporción de hogares que poseen una radio	4,5	2,7	4,8
Proporción de hogares que poseen un vídeo	5,5	3,9	6,1

50. Cada hogar de la muestra fue examinado durante un mes, y la muestra se distribuyó uniformemente a lo largo de un período de doce meses. Ello provocó problemas cuando se hicieron estimaciones de las tasas de pobreza a partir de la encuesta (véase el subapartado C.3). La variación estacional se estimó a partir d0e los datos y se utilizó para eliminar la variación estacional de las estimaciones. No pudo estimarse la variación no estacional aleatoria dentro del hogar entre unos meses y otros. El resultado fue que la dispersión del consumo de los hogares se sobrevaloró algo, y lo mismo ocurrió con las tasas de pobreza.

3. Experiencias registradas en relación con el diario sobre el uso del tiempo

51. No se conoce el número de diarios terminados por cuenta propia o con ayuda del entrevistador. No obstante, hay indicios de que los entrevistadores, en general, ofrecieron apoyo significativo a la mayor parte de los informantes, aunque pudo haber habido diferencias regionales.

52. El muestreo aleatorio de una sola persona del hogar no funcionó bien. El cálculo de la distribución por edad/sexo entre las personas que rellenaron el formulario sobre el uso del tiempo indica que los entrevistadores y supervisores no consiguieron aplicar las normas sobre la selección aleatoria. Parece que en muchos casos el entrevistador no insistió en utilizar a la persona seleccionada aleatoriamente sino que permitió sustituciones, probablemente por razones prácticas. Los cálculos indican que los hombres en edad activa (de 15 a 64 años) estaban sobrerrepresentados, mientras que ocurría lo contrario con los jóvenes (10 a 14 años) de ambos sexos y los ancianos (65 o más años), sobre todo en el caso de las mujeres (Johansson, 2000) (véase el cuadro XXIV.2). Se necesita un cambio en el procedimiento para garantizar una mejor representatividad de los datos sobre el uso del tiempo. Si lo que se pretende con este módulo de la encuesta es reflejar sobre todo las actividades económicas, quizá no sea necesario incluir a los más jóvenes y a los más ancianos. No obstante, la inclusión de estas categorías es interesante para un programa social, especialmente orientado al trabajo infantil y a la situación de los ancianos.

4. Uso de la LECS-2 para las estimaciones del PIB

53. Los resultados de la inclusión de módulos que miden los datos de valor sobre la producción de los hogares y los costos de los insumos y el uso del tiempo han sido alentadores. Ha reforzado considerablemente la base estadística de las estimaciones del producto interno bruto (PIB). La encuesta arroja ahora datos importantes para las cuentas nacionales en relación con el valor añadido en la producción de los hogares, los insumos de mano de obra en el total de la economía y el nivel y estructura del consumo privado.

54. En la nueva estimación básica del PIB para 1997, la producción de los hogares en la agricultura y en las actividades de la economía informal representaba el 64% del PIB y un porcentaje todavía superior del PIB considerado desde el punto de vista del uso.

Cuadro XXIV.2
Coeficiente entre el número efectivo y previsto de personas en la muestra del diario sobre el uso del tiempo

	Coeficiente entre resultados efectivos/previstos		
Edad	**Hombres**	**Mujeres**	**Todos**
10-14	0,41	0,49	0,45
15-64	1,33	1,04	1,18
65 +	0,59	0,29	0,43
Todos	1,11	0,90	1,00

Aproximadamente el 80% del insumo de mano de obra en la economía total procedía de la producción de los hogares en la agricultura y en actividades económicas del sector informal (Johansson, 2000).

F. Observaciones finales

55. En este capítulo se han considerado algunas cuestiones relativas al diseño de encuestas cuyo objetivo es medir el presupuesto de los hogares. Se ha insistido en las encuestas en las que se estima el total del consumo de los hogares y la producción y en las que esas estimaciones sirven como aportación para las cuentas nacionales y las estadísticas económicas nacionales en general. Puede verse un examen más detenido de las cuestiones del diseño en Deaton y Grosh (2000) y Naciones Unidas (1989), por ejemplo.

56. El estudio de caso utilizado en este capítulo es algo atípico en lo que respecta al tiempo de entrevistador utilizado por hogar. Las consideraciones de la exactitud de las mediciones y las condiciones del trabajo sobre el terreno aconsejaron para la encuesta de la República Democrática Popular Lao este tipo de diseño, que requiere una gran cantidad de recursos. El uso del método del diario en una población con baja tasa de alfabetización obligó en muchos casos a prestar ayuda a los hogares prácticamente todos los días. El método del diario con ayuda del entrevistador se consideró necesario para reflejar con exactitud el consumo de los hogares en este país. En otros podrían conseguirse estimaciones de calidad aceptable con métodos menos costosos.

REFERENCIAS BIBLIOGRÁFICAS

Banco Mundial (1992). Indonesia: public expenditures, prices and the poor. Indonesia Resident Mission 11293-IND, Yakarta. Citado en Deaton (1997).

Comisión de las Comunidades Europeas, Fondo Monetario Internacional, Organización de Cooperación y Desarrollo Económicos. Naciones Unidas y Banco Mundial (1993). *Sistema de Cuentas Nacionales, 1993*. No. de venta S.94.XVII.4.

Deaton, A. (1997). *The Analysis of Household Surveys. A Micro Econometric Approach to Development Policy*. Baltimore, Maryland y Londres: Johns Hopkins University Press.

_____ y M. Grosh (2000). Consumption. En *Designing Household Survey Questionnaires in Developing Countries: Lessons from 15 Years of Living Standards Measurement Study,* M. Grosh y P. Glewwe, comps. Washington, D.C.: Banco Mundial.

Glewwe, P. y Yansaneh, I. (2001). *Recommendations for Multi-Purpose Household Surveys from 2002 to 2010*. Report of Mission to the General Statistics Office, Viet Nam.

Johansson, S. (2000). *A Household Survey Program for Lao PDR. Report on a Short-Term Mission to Vientiane, August 7-21, 2000*. Estocolmo: International Consulting Office, Statistics Sweden.

Joliffe, D. y K. Scott (1995). The sensitivity of measures of household consumption to survey design: results from an experiment in El Salvador. Washington, D.C.: Departamento de Investigaciones sobre Políticas de Desarrollo, Banco Mundial.

Mahalanobis, P. C. y S. B. Sen (1954). On some aspects of the Indian National Sample Survey. *Bulletin of the International Statistical Institute*, vol. 34.

Naciones Unidas (1989). *National Household Survey Capability Programme: Household Income and Expenditure Surveys: A Technical Study*. DP/UN/INT-88-X01/6E. Departamento de Cooperación Técnica para el Desarrollo y Oficina de Estadística.

Pettersson, H. (2001). *Sample Design for the Household Surveys: Report from a Mission to the National Statistics Centre, Lao P.D.R. February 19-March 2, 2001*. Estocolmo: International Consulting Office, Statistics Sweden.

Rydenstam, K. (2000). The "light" time diary approach: report on some Lao PDR and Swedish actions and experiences. Documento para la reunión del Grupo de Expertos en métodos para la realización de encuestas sobre el empleo del tiempo, 23-27 de octubre de 2001.

Scott, C. (1992). Estimation of annual expenditure from one-month cross-sectional data in a household survey. *Inter-Stat Bulletin*, vol. 8, págs. 57-65.

_____ y B. Amenuvegbe (1990). *Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana*. Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper, No. 6. Washington, D.C.: Banco Mundial.

State Planning Committee, National Statistical Centre of Lao People's Democratic Republic (1999). *The household of Lao PDR: Social and economic indicators: Lao Expenditure and Consumption Survey 1997/98*. Vientiane.

Statistical Institute and Planning Institute of Jamaica (1996). *Jamaica Survey of Living Conditions 1994*. Kingston.

Capítulo XXV
Encuestas de hogares en países en transición

Jan Kordos
Escuela de Economía de Varsovia
Oficina Central de Estadística
Varsovia
Polonia

RESUMEN

En este capítulo se presenta un panorama de los principales aspectos del diseño y realización de encuestas de hogares por muestreo (encuestas de hogares por muestreo) en los países en transición durante el decenio 1991-2000. Además, se presenta información de 14 países en transición sobre los aspectos prácticos de estas encuestas. Las oficinas de estadística de esos países facilitaron esta información en 2001 respondiendo a cuestionarios especiales y, en algunos casos, actualizándolos posteriormente.

El capítulo consta de dos apartados. En el apartado A se presenta una valoración general de las encuestas de hogares en los países en transición. El apartado B contiene estudios de casos de encuestas de hogares por muestreo en algunos países en transición.

En el apartado A se presenta una síntesis de las principales características de las encuestas de hogares en los países en transición. En particular, se consideran dos tipos principales de encuestas: las encuestas sobre el presupuesto de los hogares (HBS) y las encuestas de población activa (LFS). Se consideran las siguientes características de las encuestas: marco muestral; diseño de la muestra, tamaño de las muestras, método de estimación, estimación de los errores de muestreo, tasas de falta de respuesta, costos de las encuestas y efectos del diseño. Los países en transición tenían ya cierta tradición de encuestas sobre presupuesto de los hogares, aunque hubo que efectuar readaptaciones en cada país. La encuesta de población activa es un tipo de encuesta totalmente nuevo y se ha introducido en diferentes países en transición únicamente en el último decenio, en algunos casos con asistencia técnica del exterior. El apartado A concluye con algunas recomendaciones para mejorar las encuestas de hogares por muestreo en los países en transición, teniendo en cuenta los censos de población y vivienda de 2000.

En el apartado B se presentan estudios de casos de Estonia, Hungría, Letonia, Lituania, Polonia y Eslovenia. En las descripciones se destacan las principales características de las encuestas sobre presupuesto de los hogares y de población activa y otras encuestas de hogares en cada país.

Términos clave: Encuesta sobre el presupuesto de los hogares, encuesta de población activa, costo de la encuesta, efecto del diseño, error de muestreo, tasa de falta de respuesta.

A. Evaluación general de las encuestas de hogares en los países en transición

1. Introducción

1. El objetivo del presente apartado es describir algunos aspectos del diseño y realización de las encuestas de hogares en algunos países en transición, en particular algunos de los países de Europa central y oriental y la Federación de Rusia, durante 1991-2000. El hecho de que haya grandes diferencias entre los distintos tipos de encuestas de hogares por muestreo en lo que respecta a la materia, unidades de respuesta, periodicidad, diseño de la muestra y metodologías de recopilación, da lugar a diferentes niveles de costos y de tasas de falta de respuesta. En este capítulo se consideran el diseño y realización de dos tipos de encuestas de hogares por muestreo: la encuesta sobre el presupuesto de los hogares (HBS) y la encuesta de población activa (LFS). No obstante, se mencionan también otras encuestas de hogares realizadas por los países en transición durante el decenio estudiado.

2. Antes de considerar las encuestas de hogares por muestreo en los países en transición durante el decenio se presentará una descripción general de las encuestas de hogares en esos países antes del período de transición como base para comprender el ulterior desarrollo de las encuestas de hogares en estos países.

3. Al preparar este capítulo se elaboró un cuestionario especial, que se envió a los 14 países en transición siguientes:

- Belarús,
- Bulgaria,
- Croacia,
- Eslovaquia,
- Eslovenia,
- Estonia,
- Federación de Rusia,
- Hungría,
- Letonia,
- Lituania,
- Polonia,
- Rumania,
- República Checa,
- Ucrania.

4. Ocho países prepararon documentos de alcance general que se publicaron en *Statistics in Transition,* vol. 5, No. 4 (junio de 2002).

5. Se presta especial atención al diseño y realización de encuestas de hogares por muestreo en estos países, con especial insistencia en cuestiones como los marcos muestrales, el diseño muestral, el tamaño de la muestra, los métodos de estimación de parámetros y errores de muestreo, las tasas de falta de respuesta, los costos de las encuestas y los componentes de costos, los efectos del diseño y su uso en el análisis estadístico. Se describen también los futuros planes para mejorar las encuestas después de la ronda de censos de población y censos de vivienda de 2000.

2. Encuestas de hogares por muestreo en los países de Europa central y oriental y la URSS antes del período de transición (1991-2000)

6. No es fácil evaluar objetivamente las encuestas de hogares por muestreo en los países de Europa central y oriental y la URSS antes del período de transición. Como es conocido, estos países tenían un sistema centralizado de estadística, y los censos o registros

completos eran la forma principal de recopilación de datos. No obstante, hay publicaciones en las que se describen las encuestas de hogares de esos países durante el período en cuestión y es un hecho que se celebraron conferencias, seminarios y reuniones de grupos de trabajo para examinar los métodos de las encuestas.

7. Los antiguos países comunistas —los países de Europa oriental y central y la Unión Soviética— tenían un sistema de encuestas de hogares por muestreo, las más importantes de las cuales eran las encuestas sobre el presupuesto familiar (FBS). Se realizaban también periódicamente encuestas en gran escala sobre las condiciones de vida y encuestas sobre ingresos, microcensos, encuestas de salud, encuestas sobre el empleo del tiempo y diferentes tipos de encuestas sociales y demográficas por muestreo.

8. A partir del decenio de 1950, las encuestas sobre el presupuesto familiar se establecieron según la metodología soviética basada en el llamado método sectorial (Postnikov, 1953). Ello supone la elección de hogares de entre los empleados de determinadas empresas en cada sector de actividad. Los hogares seleccionados que participaban en la encuesta durante varios años llevaban diarios de ingresos y gastos. La muestra no rotaba y estaba formada únicamente por hogares con personas empleadas en empresas socialistas, con exclusión de quienes vivían demasiado lejos de la empresas seleccionadas. Los hogares se seleccionaban de acuerdo con un diseño en dos etapas. En la primera fase de selección se designaba un número determinado de empresas (u otras unidades de lugares de trabajo) con una probabilidad proporcional al número total de empleados en las mismas. En la segunda fase, en cada empresa seleccionada se elegía sistemáticamente el mismo número de hogares entre una lista de empleados estratificada por tipo de grupo económico. Cada grupo se ordenaba primero en función de la magnitud de los salarios. En cada fase de selección se elegían unidades muestrales sistemáticamente comenzando por el punto medio del "intervalo de muestreo". Se suponía que este método de selección muestral era autoponderado en cada tipo de actividad. Después de la selección se aplicaba un procedimiento especial para comprobar la representatividad de la muestra, utilizando datos sobre los sueldos y salarios medios. Al comienzo, las tasas de falta de respuesta eran bajas y había notables diferencias entre los países.

9. En los años 1959-1962 se hizo especial hincapié en mejorar y unificar las FBS. Para ello, un Comité Permanente de Estadística del Consejo de Ayuda Mutua Económica (CAME) estableció un grupo de trabajo especial formado por países de Europa oriental y central y por la Unión Soviética. Se consiguieron algunos progresos en esferas metodológicas, como definiciones de conceptos, clasificaciones y diseño del cuestionario. Algunos países pusieron en tela de juicio el planteamiento sectorial, señalando las desventajas de tener el mismo hogar en la encuesta durante varios años. En algunos países, como las tasas de falta de respuesta habían aumentado de forma ininterrumpida, se propuso que se aplicara un método de rotación en la selección de la muestra y que se redujera el plazo de participación del mismo hogar en la encuesta. En el decenio de 1960 algunos países experimentaron un "planteamiento territorial", que consistía esencialmente en un diseño de probabilidad geográfica en el que los hogares se seleccionaban a partir de las zonas de empadronamiento y las viviendas estratificadas por región. La rotación de los hogares de la muestra redujo los períodos de participación de éstos en la encuesta. (Glowny Urzad Statystyczny (GUS), 1971a; Kordos 1985, 1996)[1]. En algunos países de Europa central y oriental (Bulgaria, Checoslovaquia, Hungría, Polonia, Rumania) comenzó a cambiar la metodología de las HBS.

[1] GUS = Oficina Central de Estadística de Polonia.

10. Después de algunos experimentos, Polonia aceptó en 1971 el planteamiento territorial para las HBS, y en 1982 se aplicó el método de rotación (Glowny Urzad Statystyczny (GUS), 1971a; Kordos, 1982, 1985; Lednicki, 1982).

11. En Hungría, después del establecimiento del Sistema unificado de encuestas de hogares (USHS) a mediados del decenio de 1970, la encuesta sobre el presupuesto de los hogares se convirtió en una encuesta continua durante el período 1976-1982, en los años

1983-1991 se realizó con periodicidad bienal y desde 1993 es de nuevo una encuesta continua. Las encuestas de ingresos, introducidas en 1963, se realizaron dos veces por decenio. Hubo varias encuestas de hogares en el marco del USHS, sobre todo en el decenio de 1980. Por ejemplo, una encuesta sobre el empleo del tiempo, una encuesta sobre el prestigio (prestigio asociado a las diversas ocupaciones), una encuesta sobre las condiciones de vida y la estratificación social, etcétera (Mihalyffy, 1994; Éltetö y Mihalyffy, 2002).

12. Se realizaron otras encuestas de hogares en estos países durante el período previo a la transición. El Comité Permanente de Estadística del CAME incluyó en su plan de trabajo de 1968-1970 un tema sobre las "posibilidades de mayor aplicación de los métodos de muestreo en las investigaciones estadísticas de los países miembros del Consejo de Ayuda Mutua Económica". En abril de 1970, Polonia se encargó de organizar un seminario sobre el tema y de preparar el documento principal (Kordos, 1970). En el seminario participaron nueve países (Bulgaria, Checoslovaquia, la República Democrática Alemana, Hungría, Mongolia, Polonia, Rumania, la Unión Soviética y Yugoslavia). Cada país presentó un documento en ruso, y todos ellos se publicaron posteriormente en polaco en un volumen especial [Glowny Urzad Statystyczny (GUS), 1971a]. Se presentaron también documentos metodológicos, que se publicaron en un segundo volumen [Glowny Urzad Statystyczny (GUS), 1971b]. Gracias a esos documentos es posible evaluar en términos generales qué tipo de encuestas de hogares se realizaron en esos países hasta 1970.

13. Hubo varias conferencias internacionales sobre las encuestas de hogares y, en particular, acerca de las encuestas sobre el presupuesto de los hogares. Los estadísticos polacos participantes en esas reuniones internacionales prepararon informes amplios, que se publicaron en revistas polacas de estadística. El autor participó en el Seminario de estadística europeo sobre las encuestas de hogares que tuvo lugar en Viena en 1961 (Kordos, 1963), y en la Segunda conferencia internacional sobre metodología de las encuestas de hogares, celebrada en Ginebra en 1981 (Kordos, 1981).

14. Las encuestas de hogares en esos países se examinaron también en la Conferencia internacional sobre estadísticas económicas para las economías en transición: Europa oriental en el decenio de 1990, que tuvo lugar en la ciudad de Washington, del 14 al 16 de febrero de 1991 (Garner y otros, 1993).

15. El análisis de esas publicaciones permite comprobar que se utilizaron métodos de muestreo en las esferas de agilización del procesamiento de datos de los censos de población y vivienda (Bulgaria, Checoslovaquia, la República Democrática Alemana, Polonia, Yugoslavia); microcensos (Checoslovaquia, Hungría, Polonia, la URSS, Yugoslavia); condiciones de vida (Bulgaria, Hungría, Polonia, Rumania, la URSS); encuestas de verificación después de los censos de población y vivienda (Bulgaria, Checoslovaquia, Hungría, Mongolia, Polonia, Rumania, la URSS, Yugoslavia), y encuestas sobre el empleo del tiempo (Bulgaria, Hungría, Polonia, Rumania). Se observaron grandes diferencias en el desarrollo estadístico de esos países, lo que influyó en cierta manera en el progreso de las encuestas de hogares durante el período de transición.

3. Encuestas de hogares en el período de transición

16. En este apartado se examinan la metodología y realización de las encuestas de hogares por muestreo que tuvieron lugar durante el período de transición (1991-2000). Las encuestas se ampliaron y modificaron considerablemente en este período. Las encuestas sobre el presupuesto de los hogares se mejoraron, y continúan mejorándose, y por primera vez en cada país se ha introducido, o se va a introducir en breve, una nueva encuesta: la encuesta de población activa. Asimismo, se están poniendo en marcha otras nuevas encuestas de hogares por muestreo: encuestas sobre el bienestar y la salud de la población, encuestas sobre las condiciones de vida de la población y otras encuestas demográficas y sociales.

17. Comenzaremos examinando las HBS y las LFS. Se describen también en términos generales otras encuestas periódicas o realizadas con carácter excepcional. Luego se prestará especial atención a algunos aspectos metodológicos comunes a todas las encuestas de hogares, como la construcción del marco muestral, el diseño de la muestra, el método de estimación, el error de muestreo, el efecto del diseño, los costos de la encuesta, la falta de respuesta y los planes para introducir futuras mejoras en las encuestas de hogares.

18. En el decenio en estudio, en casi todos los países en transición se revisó la HBS y se introdujo una nueva encuesta. Como no había LFS antes del período de transición, se formularon y realizaron encuestas de nueva planta. En el cuadro XXV.1 *infra* puede observarse el año de comienzo de la nueva HBS y de la nueva LFS, su periodicidad y el último año de revisión.

19. Como puede verse en el cuadro, las nuevas HBS, después de revisarse y ajustarse en conformidad con los requisitos de la Oficina Estadística de las Comunidades Europeas (Eurostat) (Eurostat, 1997), eran en general encuestas permanentes. Las LFS se introdujeron en los países en transición durante el período 1992-1999.

4. Encuestas sobre el presupuesto de los hogares

20. La realización de las encuestas sobre el presupuesto de los hogares tiene una larga tradición en los países en transición. Ha habido gran interés por estas encuestas debido a su función especial en el análisis de las condiciones de vida de la población y en el cálculo de los índices de los precios de consumo. Se experimentaron varios métodos y se realizaron varios intentos de mejorar la metodología y la organización. En algunos países, como Bulgaria, Hungría, Polonia y Rumania, las mejoras en la metodología de las encuestas habían comenzado en los decenios de 1970 y 1980. Al comienzo de los años noventa, otros países comenzaron a cambiar la metodología de las HBS. Las encuestas se revisaron y se ajustaron a los requisitos de Eurostat (Eurostat, 1997). Eurostat ha mostrado gran empeño en ayudar a los Estados Miembros y a otros países interesados a mejorar los métodos y procedimientos de sus encuestas dándoles orientación y apoyo técnico directo (Eurostat, 1995, 1996, 1998a, 1998b). Así pues, se han adoptado nuevos conceptos, definiciones y clasificaciones y se han

Cuadro XXV.1
Nuevas encuestas sobre el presupuesto de los hogares y de población activa en algunos países en transición, 1992-2000

	Año inicial		Periodicidad		Año de la última revisión	
País	**HBS**	**LFS**	**HBS**	**LFS**	**HBS**	**LFS**
Belarús	1995	–	Trim.	–	1995	–
Bulgaria	1992	1993	Perm.	Trim.	2000	2001
Croacia	1998	1996	Perm.	Sem.	2000	2000
Eslovaquia	2003	1993	Perm.[a]	Perm.	2002	1999
Eslovenia	1999	1993	Perm.	Perm.	1997	Anual[b]
Estonia	1995	1995	Perm.	Perm.	1999	2000
Federación de Rusia	1997	1992	Perm.[c]	Trim.	1996	1998
Hungría	1976	1996	Perm.	Perm.	1997	1997
Letonia	1995	1995	Perm.	Sem.	1998	1999
Lituania	1992	1994	Perm.	Sem.	1996	1997
Polonia	1982	1992	Perm.	Perm.[d]	2000[e]	1999
República Checa	1991	1993	Perm.	Perm.[f]	1999	2000
Rumania	2001	1994	Perm.	Perm.[g]	2000	2001
Ucrania	1999	1999	Perm.	Trim.	2000	1999

Fuente: Datos procedentes de cuestionarios presentados por algunos países.

Nota: Trim. = trimestral; Perm. = permanente; Sem. = semestral; (–) = datos no aplicables.

[a] Permanente desde 1957, pero se revisó en 2002. Nueva encuesta a partir de 2003.

[b] Cada año se agregan preguntas especiales.

[c] Permanente desde 1952 pero revisada en 1996. En 1997 comenzó una nueva encuesta.

[d] Desde el cuarto trimestre de 1999.

[e] Desde 1982, revisada tres veces.

[f] Desde 1996.

[g] Desde 2000.

elaborado nuevos diarios y cuestionarios. Por primera vez, las encuestas se están utilizando también como elemento integrante de las cuentas nacionales, para medir el consumo final de los hogares en cifras agregadas.

21. Todas las HBS se limitan a la población residente en hogares privados. Quedan excluidos los hogares colectivos o institucionales (hospitales, residencias de ancianos, internados, centros penitenciarios, cuarteles militares, etcétera). Los 14 países en transición, con excepción de la República Checa y de Eslovaquia, han remodelado su HBS.

5. Encuestas de población activa

22. En los países en transición las encuestas de población activa representan un nuevo concepto, elaborado únicamente después de 1992. Eurostat y representantes de las oficinas nacionales de estadística y de los ministerios de trabajo, con el fin de examinar los aspectos técnicos de estas encuestas, se reunieron periódicamente varias veces al año en las reuniones que el Grupo de Trabajo sobre estadísticas de empleo celebró en Luxemburgo (Eurostat, 1998a, 1998b). En consecuencia, las LFS se realizaron de acuerdo con las recomendaciones de la Organización Internacional del Trabajo (OIT) y los métodos y definiciones de Eurostat (Eurostat, 1998a).

23. Desde 1989, la Oficina de Estadística de la OIT había ayudado activamente a los países de Europa central y oriental y a la antigua URSS a revisar y reestructurar radicalmente sus sistemas de estadísticas de población activa a fin de tener en cuenta los nuevos requisitos derivados de su transición hacia una economía de mercado. Esta asistencia técnica se ofreció en forma de varias sesiones de capacitación, seminarios, conferencias y visitas de expertos.

24. Con respecto a la LFS, los expertos de la OIT realizaron misiones a la Federación de Rusia (en 1992, para la preparación de una LFS piloto; en mayo de 1993 para una LFS en toda regla); Ucrania (noviembre de 1991 y noviembre de 1992, para la preparación de una LFS; en noviembre de 1993, para una encuesta de prueba); Bulgaria (diciembre de 1991, julio y octubre de 1992, abril de 1993 y febrero de 1994); Eslovenia (octubre de 1993); Belarús (noviembre de 1993 y septiembre de 1994, para la preparación de una encuesta piloto y actividades de seguimiento); Kazajstán (marzo y junio de 1993, para examinar la viabilidad de poner en marcha una LFS piloto). Además, se organizaron tres sesiones de capacitación en el empleo acerca de la preparación y realización de una LFS para especialistas de Rusia y Ucrania en Noruega (1991) y Alemania (1991 y 1992).

25. En 1994 (31 de agosto al 2 de septiembre), la OIT organizó la Conferencia internacional sobre la reestructuración de las estadísticas del trabajo en los países en transición, que tuvo lugar en Minsk. El objetivo inmediato de la conferencia era tomar nota de los logros alcanzados y de la tarea todavía pendiente con el fin de disponer de estadísticas fiables y coherentes sobre el mercado de trabajo para atender las necesidades de formulación de políticas y de información en los países en transición. Todos los documentos preparados para esta conferencia internacional se publicaron en un número especial de *Statistics in Transition,* vol. 2, No. 1, marzo de 1995.

26. Algunos aspectos del diseño y realización de las LFS en 14 países en transición deben ser objeto de especial atención. Como se ha visto en el cuadro XXV.1, 13 de los países en transición han comenzado ya las LFS y Belarús tiene previsto comenzar una en breve. Siete de los 14 países (Eslovaquia, Eslovenia, Estonia, Hungría, Polonia, la República Checa y Rumania) realizaron encuestas permanentes, lo que significa que las semanas de referencia se distribuyeron uniformemente a lo largo de todo el año. En tres países (Bulgaria, la Federación de Rusia y Ucrania), la encuesta se llevó a cabo con periodicidad trimestral; en otros tres (Croacia, Lituania y Letonia), dos veces al año. En Estonia, hasta 1999, la encuesta se realizó anualmente (en primavera), pero desde 2000 se ha convertido en una encuesta trimestral

permanente. Todos los países tienen previsto revisar su LFS en el futuro próximo, utilizando los resultados de los censos de población y vivienda como base para mejorar el marco muestral, el diseño de la muestra y el método de estimación.

6. Características comunes de los diseños muestrales y realización de las HBS y las LFS

27. Las HBS y las LFS son dos encuestas de hogares significativamente distintas. No obstante, en la medida en que algunas características metodológicas y de ejecución (como el marco muestral, diseño de la muestra, método de estimación, estimación del error muestral, efecto del diseño, costo y tasas de falta de respuesta, así como los planes futuros para mejorar las encuestas) son comunes a ambas, conviene considerarlas simultáneamente.

28. Cada país ha seguido procedimientos bastante parecidos para reclutar y capacitar a los entrevistadores. En general, éstos no se contratan y capacitan en exclusiva para las HBS o LFS, pues participan también en otras encuestas de hogares en el país. En todas las HBS la recopilación de datos supone una combinación de diarios mantenidos por los hogares o individuos, en general con anotaciones realizadas cada día, y una o más entrevistas.

29. En lo que respecta a la LFS, la entrevista personal cara a cara es el principal modo de recopilación de datos. La "persona de referencia" facilita información sobre el hogar, y cada individuo rellena un cuestionario personal. La entrevista mediante sustitutos es rara, pero la mayor parte de los países consideran que es una fuente válida de información. En las situaciones en que no es posible establecer contacto personal con un individuo, la mayoría de los países admiten la posibilidad de que el entrevistador deje el cuestionario para que lo rellene el informante ("autoadministración"). Esta opción goza de mayor aceptación que la entrevista mediante sustitutos. Dado el contenido del cuestionario, no se han utilizado demasiado las entrevistas por teléfono, pero están comenzando a emplearse este sistema con ayuda de computadora (CATI) (Estonia). La mayoría de los países utilizan el sistema de entrevista tradicional con lápiz y papel.

Marco muestral para las HBS y las LFS

30. Los censos de población y vivienda son la base para la construcción del marco muestral de las encuestas de hogares en varios países (Bulgaria, Hungría, Polonia y Rumania). Se utilizan los datos del censo para crear unidades primarias de muestreo (UPM) basadas en las zonas de empadronamiento censal (ZE), normalmente ajustadas a las demandas específicas de la encuesta. En la mayor parte de los casos, las viviendas sirven como unidades secundarias de muestreo (USM). Normalmente, las viviendas de algunas UPM se actualizan cada año. La actualización supone una estimación del descenso del total de viviendas debido a la demolición de edificios y cambios en las fronteras de los distritos como consecuencia de la modificación de las divisiones administrativas del país (Glowny Urzad Statystyczny (GUS), 1999; Kordos, 1982, 1996; Lednicki, 1982; Martini, Ivanova y Novosyolva, 1996; Mihalyffy, 1994).

31. Algunos países de la antigua Unión Soviética, por ejemplo, Belarús, Estonia, Letonia y Lituania, utilizan registros de población (RP) y direcciones de dicho registro y otros documentos administrativos disponibles como marcos de muestreo. (Lapins y Vaskis, 1996; Martini, Ivanova y Novosyolova, 1996; Šniukstiene, Vanagaite y Binkauskiene, 1996; Traat, Kukk y Sostra, 2000).

32. En la Federación de Rusia, se utilizó eficazmente el microcenso de 1994 como marco de muestreo para la HBS y la LFS (Goskomstat, 2000).

33. En general, la población objetivo incluye todos los hogares privados en todo el territorio nacional de cada país, con pequeñas excepciones. En algunos casos no se incluyen

algunos pequeños grupos de población, sobre todo debido a limitaciones en la cobertura del marco muestral disponible.

34. Hay planes para utilizar los resultados de la ronda de censos de población de 2000 como marcos muestrales para la HBS y la LFS y otras encuestas de hogares en el futuro (Éltetö y Mihalyffy, 2002; Kordos, Lednicki y Zyra, 2002; Lapins y otros, 2002; Kurvits, Sõstra y Traat, 2002).

Tamaño y asignación de la muestra

35. En el año 2000 el intervalo de los tamaños de la muestra para la HBS oscilaba desde 1.028 hogares en Eslovenia y 1.300 en Eslovaquia hasta 36.163 en Polonia y 48.675 en la Federación de Rusia. En el cuadro XXV.2 pueden verse los tamaños de la muestra de la HBS y de la LFS en los países en transición en 2000.

36. En general, los grandes países, debido a su mayor necesidad de resultados desglosados y a su mayor capacidad, necesitaban tamaños muestrales mayores pero, naturalmente, no en proporción a su tamaño. En algunos países la muestra se distribuyó proporcionalmente en las diferentes regiones geográficas con el fin de lograr la máxima precisión de las estimaciones en el plano nacional. No obstante, tres países, Hungría, Polonia y la Federación de Rusia, eligieron asignaciones desproporcionadas, otorgando tasas más elevadas a las regiones menores y garantizando así un tamaño mínimo de la muestra para cada región del país.

Cuadro XXV.2
Tamaño de la muestra, diseño de la muestra y métodos de estimación en la HBS y la LFS (2000), en algunos países en transición

	Tamaño de la muestra		Diseño de la muestra		Método de estimación	
País	HBS	LFS	HBS	LFS	HBS	LFS
Belarús	6 000	_	2 etapas	_	Ponder.	_
Bulgaria	6 000	24 000	2 etapas PPT	2 etapas PPT	Directo	Ponder.
Croacia	2 865	12 843	2 etapas PPT	2 etapas PPT	Ponder.	Directo
Eslovaquia	1 300	10 250	Cuota	2 etapas PPT	_	Ponder.
Eslovenia	1 028	7 000 [a]	2 etapas PPT	1 etapa persona	Ponder.	Ponder.
Estonia	9 840	9 127	RP PPT	RP Pr. ig.	Ponder.	Ponder.
Federación de Rusia	48 675	123 041	2 etapas PPT	2 etapas PPT	Microc. ponder.	Microc. ponder.
Hungría	11 862 [b]	36 500 [c]	3 etapas PPT	3 etapas [d] PPT	Calibr.	Calibr.
Letonia	3 847	7 940	2 etapas PPT	2 etapas PPT	Ponder.	Ponder.
Lituania	10 680	6 000	PPT persona	PPT persona	Ponder.	Ponder.
Polonia	36 163 [e]	24 400 [f]	2 etapas PPT	2 etapas PPT	LFS ponder.	Demogr. ponder.
República Checa	3 250	31 800	Cuota	2 etapas PPT	Último microcenso	Ponder.
Rumania	17 827	17 600	2 etapas PPT	2 etapas PPT	Ponder.	Ponder.
Ucrania	12 534	38 695	2 etapas PPT	2 etapas PPT	Ponder.	Ponder.

Fuente: Datos procedentes de cuestionarios presentados por algunos países.

Nota: "Último microcenso" fue el de 1995; "Microc. ponder." = microcenso ponderado; "RP" = registro de población; "LFS ponder." = ponderaciones utilizadas en la LFS; "Demog. ponder." = ponderaciones utilizadas a partir de los datos de control de posestratificación de las proyecciones demográficas; "Calibr." = método de calibración; "Pr. ig." = probabilidad igual. (–) = datos no aplicables.

[a] Cifra trimestral.

[b] El número de hogares que cooperó con la encuesta fue de 10.191. Para conseguir este resultado, el entrevistador tuvo que visitar nada menos que 17.243 direcciones.

[c] Selección trimestral.

[d] Excepto en las ciudades autorrepresentadas, en cuyo caso la selección fue en dos etapas.

[e] Este tamaño de la muestra se consiguió únicamente en 2000. En el año anterior, el total era de unos 32.000 hogares.

[f] Número trimestral de las viviendas seleccionadas. Cada trimestre, se selecciona el mismo número de viviendas.

37. En el año 2000

a) HBS: La Federación de Rusia tuvo el tamaño más grande (48.675 hogares), seguida de Polonia (36.163), Ucrania (12.534) y Hungría (11.862). Los países con la muestra más pequeña fueron Eslovaquia (1.300) y Eslovenia (1.028);

b) LFS: La Federación de Rusia tuvo la muestra más grande (123.041), seguida de Ucrania (38.695), Hungría (36.500, trimestral), Polonia (24.400, trimestral), República Checa (31.800), Bulgaria (24.000), Rumania (17.600) y Eslovaquia (10.250). Todos los demás países utilizaron muestras de tamaño inferior a 10.000.

Diseño de la muestra y selección

38. En los diez últimos años los países en transición aplicaron diferentes diseños de la muestra para las HBS y las LFS. Se utilizaron criterios diversos para la estratificación de las UPM antes de la selección. El criterio más frecuente fue la región geográfica y/o el entorno urbano/rural. La estratificación por tamaño de la población de la localidad se utilizó también en algunos países (por ejemplo, Hungría, Polonia, la Federación de Rusia y Ucrania).

39. La mayor parte de las encuestas se basaron en el muestreo en dos etapas: selección de las unidades primarias de muestreo (UPM), en la primera etapa, seguida de la selección de un pequeño número de viviendas u hogares dentro de cada UPM seleccionada, en la segunda etapa. Normalmente, las probabilidades de selección en las dos etapas se equilibraban con el fin de obtener una muestra "autoponderada" de los hogares dentro de los diferentes campos: las UPM se seleccionan con probabilidad proporcional al tamaño (PPT), normalmente en proporción con el número de viviendas, y en algunas UPM se eligió el mismo número de unidades secundarias de muestreo (USM). Se utilizaron muestras directas (en una sola etapa) de las viviendas, hogares o personas en grandes ciudades de Letonia y Lituania. Por el contrario, en Hungría, en las pequeñas localidades, la muestra se seleccionó en tres etapas: grandes áreas en la primera, agrupaciones menores en la segunda y direcciones u hogares en la última.

Rotación de la muestra

40. La carga de los hogares puede reducirse mediante rotación periódica de la muestra, para que no tengan que responder siempre los mismos. No obstante, la rotación de las unidades aumenta el costo de la encuesta por las siguientes razones: necesidades adicionales de mantenimiento de la muestra, posibilidad de nuevas actividades de capacitación de entrevistadores, costos de recopilación inicial de la información de referencia y dificultades para preparar nuevas unidades que puedan facilitar datos. La rotación parcial de las unidades de la muestra con una tasa fija es un compromiso entre la rotación total —la sustitución del 100% de las unidades, que resulta muy costosa y ofrece estimaciones del cambio poco precisas— y la ausencia total de rotación —una encuesta por panel—, que da lugar a una distribución inaceptable de la carga de la respuesta. Los planes de rotación mantienen una unidad dentro de la muestra durante un determinado período durante el cual ésta ya no puede ser seleccionada de nuevo para la misma encuesta por un determinado período mínimo.

41. En la mayor parte de los países en transición se utiliza una u otra forma de rotación de la muestra tanto en la HBS como en la LFS. Por ejemplo, en Estonia, Polonia y Rumania se ha aplicado la pauta 2–(2)–2, es decir, dos trimestres dentro de la muestra, dos trimestres fuera, otros dos trimestres más dentro, y salida definitiva.

Ponderación de los resultados

42. Las tasas de falta de respuesta en la HBS suelen ser elevadas, y cambian considerablemente la estructura socioeconómica de los hogares de la muestra. Para reducir este

efecto, los resultados de la muestra se someten a ponderación. Tanto el error de muestreo como el error de falta de respuesta pueden reducirse sustancialmente cuando se dispone de abundante información auxiliar y se usa para la reponderación mediante el método de calibración. Hungría fue el único país donde se utilizó la calibración en ambos tipos de encuesta (Éltetö y Mihalyffy, 2002; véase también Deville y Särndal, 1992).

43. La información sobre las características básicas de las unidades del marco puede ser útil para el diseño y selección de la muestra. Dicha información puede utilizarse sobre todo para calcular ponderaciones, que se aplican para reducir el efecto de la falta de respuesta. Con ese fin debe disponerse de información sobre las características de las unidades, tanto de las que responden como de las que no lo hacen.

44. En primer lugar, cada hogar de la muestra se pondera mediante el inverso de la probabilidad con que se seleccionó. La ponderación de la falta de respuesta supone la división de la muestra en clases adecuadas de ponderación, y dentro de cada una de esas clases los informantes se ponderan para tener en cuenta los casos de falta de respuesta en dicha clase. En algunas circunstancias se utilizan ponderaciones adecuadas de fuentes externas. Asimismo, en lo que respecta a la HBS, se aplican ponderaciones adecuadas de la LFS (por lo que se refiere al tamaño de los hogares y la relación entre zonas rurales y urbanas) (Polonia).

45. En los países bálticos se utilizan procedimientos especiales para obtener una muestra autoponderada de los hogares a partir de un registro de población (Lapins y Vaskis, 1996; Šniukstiene, Vanagaite y Binkauskiene, 1996; Traat, Kukk y Sõstra, 2000).

46. Los datos de las LFS se utilizan simultáneamente para el análisis de los hogares y de las personas. Por ello es necesario utilizar un procedimiento de ponderación que garantice la total coherencia del análisis en relación con ambos tipos de unidades. Toda la ponderación de la muestra original se aplica a los hogares, es decir, que el procedimiento garantiza que las personas dentro de un hogar reciban todas la misma ponderación.

47. Las ponderaciones se deducen en forma secuencial. En cualquiera de los pasos posteriores al primero, se calculan a partir de los valores de la muestra ya ponderados de conformidad con los resultados de los pasos precedentes. La ponderación final de una unidad es el producto de los factores de ponderación determinados en cada paso. Las ponderaciones calculadas en cada paso se normalizan; en otras palabras, se reducen a escala para que el valor medio por unidad muestral equivalga a 1,0 y la suma de las ponderaciones sea igual al tamaño de la muestra original.

48. Si bien se utiliza un conjunto común de procedimientos en todas las encuestas, las variables específicas implicadas en cada caso y las fuentes de los datos utilizados varían de una encuesta a otra. De todas formas, algunas variables suelen ser importantes prácticamente en todas las circunstancias, como la ubicación geográfica del lugar, el tamaño y la composición del hogar y la distribución de la población por edad, sexo y otras características básicas (Verma, 1995).

Estimación de los errores estándar

49. La mayoría de los países aplican diseños muestrales complejos para la HBS y la LFS y, por lo tanto, se ven obligados a incorporar estas características complejas en el cálculo de la variación de las varianzas muestrales (Wolter, 1985). No se dispone de expresiones analíticas de la varianza para estimar el error de muestreo de las estimaciones complejas; por ello, se utilizan métodos de aproximación. Los países han utilizado los siguientes métodos: grupo aleatorio (por ejemplo, Polonia en la HBS hasta 2000, y en la LFS hasta 1999), *jackknife* (Hungría), serie de Taylor (Polonia para la LFS desde el cuarto trimestre de 1999), media muestra equilibrada (Polonia para la HBS desde 2001) y un método analítico adaptado (la

Federación de Rusia). Algunos países (Estonia, Letonia y Eslovenia) utilizan el programa para el análisis estadístico de datos correlacionados Software for the Statistical Analysis of Correlated Data (SUDAAN), programa bien conocido que se utiliza para calcular los errores estándar en los diseños complejos.

Tasas de falta de respuesta en la HBS y en la LFS

50. Si consideramos las tasas medias de falta de respuesta en la HBS de algunos países en transición durante los últimos cuatro años, pueden identificarse los tres grupos siguientes a partir de los datos contenidos en el cuadro XXV.3:

a) Grupo con tasa elevada de falta de respuesta (más del 40%): Estonia (43,6%), Polonia (43,4%), Bulgaria (41,7%) y Hungría (40,0%);

b) Grupo con tasa intermedia de falta de respuesta (más del 20% y menos del 30%): Federación de Rusia (25,6%), Ucrania (25,0%), Letonia (24,5% y Lituania (22,2%);

c) Grupo con tasa baja de falta de respuesta (inferior al 20%): Croacia (19,0%), Eslovenia (18,5%) y Rumania (11,0%).

51. Como puede observarse si se comparan los cuadros XXV.3 y XXV.4, las tasas de falta de respuesta en la HBS fueron muy superiores a las registradas en la LFS en todos los países. Además, en algunos de ellos se observan signos claros de un aumento de las tasas de falta de respuesta a lo largo del tiempo en ambos tipos de encuesta. En lo que respecta a la LFS, puede observarse cierto aumento de las tasas de no respuesta en los siguientes países:

a) Polonia (4,5% en 1992, frente al 22,1% de 2000);

b) Bulgaria (10,1% en 1993, frente al 17,2% de 2000);

c) República Checa (16% en 1993, frente al 24% de 2000);

d) Croacia (6,3% en 1996, frente al 15,7% de 2000);

e) Rumania (2,6% en 1994, frente al 8,9% de 2000);

f) Eslovenia (9,0% en 1992, frente al 12,0% de 2000).

Cuadro XXV.3
Tasas de falta de respuesta en la HBS en algunos países en transición, 1992-2000 (*porcentaje*)

	Tasa de falta de respuesta en el año								
País	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Bulgaria	..	33,0	34,2	35,6	37,9	49,0	41,1	39,7	37,0
Croacia	..	..	..	..	..	..	19,0	21,0	17,0
Eslovaquia	No registrada								
Eslovenia	..	24,6	22,1	28,0	34,6	19,5	18,4	17,6	18,6
Estonia	..	..	..	44,4	50,2	44,9	46,6	47,5	35,2
Federación de Rusia	10,4	10,5	5,9	11,5	31,4	47,5	25,0	13,9	16,0
Hungría	..	36,7	40,4	32,6	43,3	40,6	40,9	39,6	39,0
Letonia	..	..	..	..	26,1	24,1	21,9	23,1	28,7
Lituania	..	..	..	..	24,0	20,3	22,7	22,8	22,8
Polonia	23,2	27,6	25,3	25,1	31,4	34,3	40,7	49,4	49,2
República Checa	No registrada								
Rumania	..	..	..	8,0	10,2	9,6	10,4	11,6	13,4
Ucrania	..	..	..	..	..	..	..	24,2	25,7

Fuente: Cuestionarios especiales de los países.
Nota: Dos puntos suspensivos (..) indican que faltan datos.

Cuadro XXV.4
Tasas de falta de respuesta en la LFS en algunos países en transición, 1992-2000 (*porcentaje*)

País	Tasa de falta de respuesta en el año								
	1992	1993	1994	1995	1996	1997	1998	1999	2000
Bulgaria	..	10,1	8,8	8,5	11,0	14,4	16,7	16,2	17,2
Croacia	..	..	..	..	6,3	14,0	18,1	15,0	15,7
Eslovaquia	..	..	6,2	5,9	5,1	5,0	5,6	5,9	5,7
Eslovenia	9,0	7,9	9,8	9,7	10,0	12,5	12,4	11,7	12,0
Estonia	..	..	..	7,4	..	13,5	13,4	13,2	9,9
Federación de Rusia	4,6	6,8	5,9	4,5	5,5	5,8	5,8	5,3	4,5
Hungría	..	10,3	8,1	11,4	13,6	14,3	12,2	8,9	9,2
Letonia	..	..	..	13,7	13,3	12,4	9,8	9,4	10,1
Lituania	..	..	..	..	..	9,6	9,0	8,7	8,9
Polonia	4,5	5,3	8,9	9,9	10,0	9,6	11,6	18,2	22,1
República Checa	..	16	16	18	20	19	21	22	24
Rumania	..	..	2,6	2,3	6,4	6,7	7,4	7,9	8,9
Ucrania	..	..	..	..	..	..	..	29,2	28,3

Fuente: Cuestionarios especiales de los países.
Nota: Dos puntos suspensivos (..) indican que faltan datos.

52. Los datos del cuadro XXV.4 indican que las tasas de falta de respuesta presentaron considerables diferencias en los distintos países, que pueden dividirse en tres grupos en función del nivel de falta de respuesta:

a) Grupo con tasa elevada de falta de respuesta (superior al 15%): Ucrania (28,8%), la República Checa (21,5%), Bulgaria (16,1%), Croacia (15,7%) y Polonia (15,4%);

b) Grupo con tasas moderadas de falta de respuesta (del 10% al 15%): Estonia (12,5%), Eslovenia (12,2%), Hungría (11,2%) y Letonia (10,4%);

c) Grupo con tasas bajas de falta de respuesta (por debajo del 10%): Lituania (9,1%), Rumania (7,7%), Eslovaquia (5,6%) y Federación de Rusia (5,4%).

Costos de las encuestas de hogares

53. En toda encuesta por muestreo hay que dar respuesta a dos preguntas importantes:

a) ¿Cuál es el costo total de la encuesta?

b) ¿Cuál es el grado de precisión de las estimaciones principales?

54. No es fácil evaluar los costos de las encuestas de hogares en los países en transición. Algunos de ellos sólo indican el costo total directo de la recopilación de datos, que incluye las entrevistas, los viajes, los costos de material y los servicios vinculados con la recopilación de datos; se excluyen otros componentes de los costos, como la preparación de la encuesta, los medios de imputación metodológica, el procesamiento de los datos y la redacción y publicación del informe.

55. A pesar de la crucial importancia del presupuesto, la estimación de costos es uno de los aspectos menos desarrollados de la planificación de encuestas. Uno de los problemas es que el mantenimiento de registros detallados de los costos representa una carga muy pesada. Otro es la dificultad de distinguir entre los costos de las actividades conjuntas, especialmente los gastos administrativos e indirectos. No obstante, la elaboración y mantenimiento de un sistema general de registro de costos puede ser muy rentable con respecto a la planificación futura y la capacidad de atraer el apoyo necesario de los programas de datos (Naciones Unidas, 1984).

56. En Polonia (Kordos, Lednicki y Zyra, 2002), el costo directo de la HBS en 2000 fue de 4.567.000 euros, de los cuales 3.571.394 (78,2%) correspondieron a los costos de las entrevistas, 146.144 (3,2%) a gastos de viaje y 429.298 (9,4%) a incentivos. Dado que en 2000 el tamaño de la muestra de los hogares encuestados era de 36.163, el costo medio por hogar fue de 126,3 euros.

57. Se realizaron cálculos semejantes con respecto a la LFS de 2000. Los costos directos totales de la encuesta fueron 1.094.200 euros, de los que 878.642,6 (80,3%) correspondieron a las entrevistas y 45.956,4 (4,2%) a los viajes. No hubo costos en concepto de incentivos en la LFS. Teniendo en cuenta que en 2000 se entrevistaron casi 80.000 hogares, el costo de una entrevista fue por término medio de 13,7 euros. El costo de la HBS fue unas 10 veces superior al de la LFS, debido sobre todo al hecho de que la HBS es un proceso muy lento, que supone varias entrevistas de los mismos informantes y el uso de diarios y documentos justificantes. Por el contrario, para la LHS basta con una sola entrevista.

58. Hungría aportó datos interesantes acerca de los costos de la HBS y de la LFS en el año 2000 (Éltetö y Mihalyffy, 2002). En los cuadros XXV.5 y XXV.6. se presentan evaluaciones detalladas de la estructura de costos de la HBS y de la LFS. Los gastos correspondientes a la LFS (432.000 euros) superaron los de la HBS (326.000 euros). No obstante, considerando que en la LFS un hogar se visitaba cuatro veces al año y no se daba ningún incentivo a los hogares cooperadores, los gastos por hogar fueron considerablemente más bajos que los de la HBS (27,5 euros por hogar en la HBS frente a 8,4 euros en la LFS). En los cuadros XXV.5 y XXV.6 puede verse la estructura de costos de la HBS y la LFS, tanto en cifras absolutas (en euros) como en porcentajes.

Cuadro XXV.5
Estructura de costos de la HBS en Hungría en el año 2000

	Costo en	
Componente del costo	**Euros**	**Porcentaje**
Diarios mensuales	148 650	45,6
Cuestionarios de final del año	35 865	11,0
Visita a los hogares que no han respondido	4 345	1,3
Incentivos a los hogares que cooperan	75 855	23,3
Prima a los entrevistadores	18 585	5,7
Costos del material	42 700	13,1
Total	**326 000**	**100,0**

Fuente: Éltetö y Mihalyffy (2002).

Cuadro XXV. 6
Estructura de costos de la LFS en Hungría en el año 2000

	Costo en	
Componentes del costo	**Euros**	**Porcentaje**
Visitas a los hogares	22 032	5.1
Cuestionarios de los hogares	65 232	15.1
Cuestionarios de actividades	212 110	49.1
Cuestionarios suplementarios	42 336	9.8
Prima a los entrevistadores	33 695	7.8
Costos del material	56 595	13.1
Total	**432 000**	**100.0**

Fuente: Éltetö y Mihalyffy (2002).

Efectos del diseño

59. Como puede desprenderse de la descripción de las encuestas de hogares en algunos países en transición, casi todas las encuestas de hogares por muestreo están basadas en un proceso en varias etapas. Ello significa que el cálculo de los efectos del diseño es un requisito necesario para el análisis estadístico de los datos procedentes de esas encuestas (Kish y Frankel, 1974).

60. Veamos un ejemplo de la Federación de Rusia (Goskomstat, 2000, págs. 219-220) en el que el tamaño de la muestra de la LFS trimestral se determinó por separado con respecto a cada una de las regiones de la Federación de Rusia. El tamaño de la muestra se determinó para diversos niveles de la tasa de empleo verdadera. El nivel deseado de precisión de la estimación se fijó en el 1,5%, el 5% y el 8% para la Federación de Rusia, las regiones de tamaño grande e intermedio y las regiones pequeñas, respectivamente. El efecto del diseño se calculó de conformidad con la fórmula de la ecuación (7) del capítulo VI, y sobre la base de los datos de la encuesta por muestreo sobre empleo y desempleo de 1998. Los efectos del diseño calculados se situaban entre 1,52 y 2,14. Se calcularon los efectos del diseño en relación con varias características de la HBS y de la LFS.

7. Observaciones finales

61. En el apartado hemos presentado diferentes aspectos del diseño muestral y de la realización de encuestas de hogares por muestreo, con especial atención a las dos más importantes: la HBS y la LFS. A partir de este examen general de las encuestas de hogares pueden extraerse algunas conclusiones. Las encuestas de hogares por muestreo en países en transición se revisaron y armonizaron de acuerdo con los nuevos requisitos de la economía de mercado y las recomendaciones de Eurostat (1995; 1996; 1997; 1998a), con algunas diferencias entre los países vinculadas a sus experiencias anteriores y a las posibilidades existentes. Aunque el progreso ha sido evidente en la elaboración de encuestas de hogares, se necesita mayor atención de las oficinas de estadística al cálculo y presentación de los errores estándar, a la evaluación de los componentes del costo y al cálculo y publicación de los efectos del diseño y su utilización en el análisis estadístico. Además, hay problemas específicos que afectan a países concretos, como las bajas tasas de respuesta y el tamaño insuficiente de la muestra en algunos campos. Hay problemas muy importantes y graves relacionados con la comparabilidad de los resultados entre países. Eurostat debe ocuparse de resolver estos problemas, ya que afectan a la integración y armonización de las encuestas de hogares por muestreo. Los países en transición tienen sus propios planes para el futuro desarrollo de las encuestas de hogares. Uno de ellos supone la utilización de los resultados de la ronda de censos de población y vivienda de 2000. Sus datos ofrecen oportunidades para mejorar los marcos y diseños muestrales y los métodos de estimación, sobre todo en los campos de menor magnitud.

62. Los distintos estudios de casos sobre países concretos en transición que se presentan en el siguiente apartado ofrecen una imagen más detallada de los problemas del diseño, realización y análisis relacionados con las encuestas de hogares en esos países. Después de los estudios se adjunta una lista completa de referencias bibliográficas, que pueden utilizarse para estudiar diferentes aspectos de las encuestas de hogares en los países en transición.

B. Encuestas de hogares por muestreo en países en transición: estudios de casos

63. Los estudios de casos que se presentan en este apartado fueron preparados por autores de los siguientes países en transición: Eslovenia, Estonia, Hungría, Letonia, Lituania y Polonia. Se publicaron artículos más completos de ocho países en transición en *Statistics*

in Transition, volumen 5, No. 4 (junio de 2002). La información sobre las características principales de la HBS, la LFS y otras encuestas de hogares en cada país que se presenta a continuación sirve como complemento de la información del apartado A.

1. Encuesta de población activa y Encuesta sobre el presupuesto de los hogares en Eslovenia*

* Texto preparado por Vasja Vehovar, Facultad de Ciencias Sociales, Universidad de Ljubljana, y Metka Zaletel, Tatjana Novak, Marta Arnež y Katja Rutar, Oficina de Estadística de la República de Eslovenia.

Introducción

64. La República de Eslovenia consiguió la independencia en los primeros años noventa. Antes, las actividades estadísticas de la antigua Yugoslavia estaban centralizadas en la Oficina Federal de Estadística. En aquellas fechas las encuestas de hogares no ocupaban un lugar destacado en el programa nacional de estadística. Tras la independencia, la Oficina de Estadística de Eslovenia se transformó rápidamente de oficina regional en nacional. El proceso de transición fue relativamente suave debido, en parte, al hecho de que los altos cargos se mantuvieron prácticamente sin cambios y retuvieron el poder durante todo el período de transición.

Encuesta de población activa (LFS)

Antecedentes

65. La primera LFS fue realizada en 1989 por la Facultad de Ciencias Sociales de la Universidad de Ljubljana (Vehovar, 1997). La Oficina de Estadística de la República de Eslovenia asumió la plena responsabilidad de la LFS en 1995.

66. Las muestras de la LFS para 1989-1995 se formularon y realizaron en gran parte con carácter esporádico debido a la incertidumbre de los presupuestos anuales. Desde 1992, el diseño fue una muestra en conglomerados en tres etapas con 3.000 nuevos hogares cada año. Las unidades se mantenían en la muestra durante tres años consecutivos y el tamaño total era de unas 8.500 unidades.

Revisión

67. En 1997 tuvo lugar una importante revisión en respuesta a las peticiones de resultados más frecuentes (es decir, trimestrales) y más detallados (es decir, regionales). Las directrices de Eurostat fueron también estímulos importantes para la revisión (Eurostat, 1998a).

68. La LFS se transformó en una encuesta permanente por panel con selección de la muestra y publicación de los resultados con periodicidad trimestral. Cada muestra trimestral se divide en seis intervalos de dos semanas. El período de referencia para las entrevistas es la semana (lunes a domingo) anterior a éstas. Se aplica el modelo de rotación 3-1-2, de manera que los hogares son entrevistados durante tres trimestres consecutivos, luego se omiten durante un trimestre y vuelven a incluirse otros dos trimestres. Este modelo produce una superposición del 60% entre dos trimestres consecutivos y del 40% entre dos años consecutivos.

69. El marco muestral de la LFS es el registro central de la población y la información resultante de la estratificación. Las definiciones de los estratos están basadas en seis tipos de asentamientos (según el tamaño de éstos y la proporción de la población que son agricultores) y 12 regiones geográficas. Después de su integración, queda un total de 47 estratos.

70. En cada estrato, la muestra se selecciona utilizando el muestreo sistemático con comienzo aleatorio. Se aplica una estratificación implícita mediante una clasificación de datos por asentamiento, calle y número de edificio. La tasa de muestreo en cada estrato se ajusta de acuerdo con la tasa prevista de falta de respuesta. No se aplican sustituciones sobre el

terreno, ya que se ha comprobado que estas sustituciones ofrecen pocas ventajas y plantean considerables problemas (Vehovar, 1999).

71. En cada trimestre se seleccionan 2.000 nuevas unidades. Además, se incluyen unos 5.000 hogares (informantes) de los cuatro trimestres anteriores. Así pues, se seleccionan unos 7.000 hogares por trimestre (2.000 de la muestra entrante y 5.000 de la muestra continua). Se prevé que respondan 6.000 de esos hogares. El número total de entrevistas individuales terminadas es de aproximadamente 20.000.

Realización

72. Todos los hogares de la muestra trimestral entrante son entrevistados personalmente (entrevista directa) con ayuda de computadoras (CAPI). Hay unos 30 entrevistadores experimentados en la LFS, todos ellos equipados con computadoras portátiles. Las entrevistas repetidas se realizan desde el centro telefónico de la oficina de estadística mediante el procedimiento CATI, con excepción de los hogares que no tienen teléfono y de los que no pueden participar en una entrevista telefónica. La tasa de cobertura telefónica nacional es de aproximadamente el 95%. Antes de las entrevistas, cada hogar recibe previamente una carta con una descripción de la encuesta y un folleto con los resultados de la LFS de encuestas anteriores. No se ofrecen incentivos.

73. En las entrevistas personales de la parte entrante de la muestra, las tasas de falta de respuesta han sido del 17%-18%, y las tasas de respuesta negativa, del 12%-13%. En las entrevistas telefónicas repetidas de los hogares ya incluidos en el panel, las tasas de falta de respuesta han sido ligeramente inferiores (10%-11%), lo mismo que las tasas de respuesta negativa (6%-7%). La tasa de falta de respuesta de la LFS creció considerablemente a partir de 1991, pero se ha estabilizado en los últimos cuatro años.

Errores muestrales y su publicación

74. Los datos se ponderan en función de la probabilidad desigual de selección y de la falta de respuesta unitaria. Se aplica la posestratificación de acuerdo con la distribución demográfica conocida por edades (8 grupos), sexo y región (12 regiones). El hecho de que la posestratificación se realice en el plano individual significa que los miembros del mismo hogar pueden recibir ponderaciones diferentes.

75. Los errores de muestreo y los efectos del diseño se estiman en general únicamente con respecto a las variables fundamentales: tasa de desempleo y coeficiente empleo/población. Los efectos del diseño son relativamente bajos, por ejemplo, de 1,3 en el caso de la tasa de desempleo.

76. Se calculan habitualmente los coeficientes de variación (CV) de las estimaciones. Las estimaciones con CV inferiores al 10% se publican sin ninguna restricción; las estimaciones con CV del 10% al 20% se publican en un solo tramo. Las CV del 20% al 30% se publican con un doble tramo. Cuando el CV es superior al 30%, los resultados se sustituyen por un punto (.), que significa "distinto de cero pero poco fiable".

77. Los resultados de la encuesta se publican trimestralmente en los *Informes estadísticos resumidos*, el *Anuario de Estadística* y en varias otras publicaciones de Eslovenia. La serie especial "Resultados de la encuesta" contiene resultados detallados sobre la encuesta y la metodología. Los datos aparecen también en publicaciones de otras organizaciones, como el Banco Mundial, el Fondo de las Naciones Unidas para la Infancia (UNICEF) y Eurostat. Los microdatos son también analizados por investigadores no pertenecientes a la Oficina de Estadística.

Encuesta sobre el presupuesto de los hogares (HBS)

Antecedentes

78. La primera encuesta sobre el consumo de los hogares se había realizado en el decenio de 1960. Hasta 1997 se llevó a cabo de acuerdo con la metodología relativamente avanzada e innovadora diseñada por la Oficina Federal de Estadística de Yugoslavia. El diseño comprendía una muestra por conglomerados en dos etapas con estratificación de las UPM en la primera etapa. Las unidades primarias de muestreo eran las zonas de empadronamiento (ZE), muestreadas con probabilidad proporcional al tamaño (PPT). En la segunda etapa, los individuos se seleccionaban a partir del registro central de población; también determinaban el hogar. En cada UPM se entrevistaban cinco hogares. Hasta 1993 se utilizó el procedimiento de sustitución para poder disponer de cinco unidades con las respuestas pertinentes; no obstante, a partir de 1994 las "tomas" por conglomerado pasaron de 6 a 8 personas dentro de cada UPM, característica del diseño que obligaba a introducir correcciones adicionales con ponderaciones. Se realizaban periódicamente dos encuestas diferentes de la HBS: una con periodicidad trimestral y otra encuesta con intervalos de cinco años. En la última HBS anual, los hogares eran 3.270; en la encuesta trimestral, eran 1.000. En la encuesta anual, la entrevista se realizaba al final del año y abarcaba todo el período de doce meses, mientras que en la entrevista trimestral los hogares de la muestra se entrevistaban cuatro veces al año.

Remodelación de la muestra de la HBS

79. La principal razón de la remodelación de la encuesta fueron las nuevas orientaciones de Eurostat (Eurostat, 1997).

80. El registro de población se utiliza para seleccionar a los informantes individuales. Estas personas determinan también los hogares. Se utilizan ponderaciones para introducir los ajustes correspondientes a las probabilidades desiguales de selección de personas y hogares. Se excluyen los hogares institucionales. La muestra anual comprende 1.200 hogares informantes. Como se trata de una muestra demasiado pequeña para poder aplicar el modelo "nórdico", se fusionan los datos de las muestras de tres años consecutivos y los resultados vuelven a calcularse con respecto al año intermedio. De esta manera se puede obtener una muestra de 3.600 hogares.

81. Se utiliza la asignación proporcional en 47 estratos. Debido al tamaño relativamente pequeño de la muestra y al gran número de estratos, la estratificación se realiza sólo de manera implícita. En los pequeños asentamientos (menos de 1.000 habitantes), las zonas de empadronamiento sirven como UPM y se seleccionan con probabilidad proporcional al tamaño (PPT). En cada UPM se seleccionan cuatro hogares informantes. En las ciudades de tamaño intermedio y grande se aplica el método de muestreo aleatorio simple (MAS). En consecuencia, los efectos del diseño son relativamente bajos, aproximadamente 1,2 en las variables fundamentales. Las unidades se seleccionan por separado en cada trimestre y se asignan a las 12 semanas del trimestre correspondiente. La decimotercera semana se utiliza para el trabajo pendiente con quienes no han respondido.

Realización

82. Una semana antes de la primera visita se envía una carta previa, junto con el incentivo: una calculadora de bolsillo. Como se trata de una encuesta permanente, puede realizarse con un número menor de entrevistadores (por ejemplo, 20).

83. Los entrevistadores registran todos los contactos efectivos o intentos de contacto con un hogar en un formulario especial. La situación con respecto a la vivienda, el hogar y la persona de referencia de cada unidad es, por lo tanto, muy clara, y lo mismo ocurre

con el número de intentos de contacto, número de diarios rellenados y posibles razones de la falta de respuesta.

84. Los datos se recopilan utilizando el cuestionario rellenado por el entrevistador y los diarios mantenidos por los miembros del hogar. Casi todas las entrevistas se realizan con ayuda de computadoras (CAPI).

85. Los hogares llevan el diario durante 14 días. Durante ese tiempo anotan periódicamente la información correspondiente a sus gastos diarios. Se considera que los hogares responden si realizan al menos el cuestionario básico de la entrevista, ya que éste contiene dos tercios de los datos. La tasa de respuesta en los cuestionarios de las entrevistas es relativamente elevada y estable (aproximadamente el 81%). En cambio, la tasa de respuestas completas, incluidos los diarios, es más baja: en torno al 70%.

Errores muestrales y publicación

86. Si faltan todos los diarios de un hogar dado, los datos se imputan utilizando el método *hot-deck* a partir de un donante similar. Se utiliza el mismo procedimiento para el caso de falta de respuestas sobre partidas concretas. Cada valor omitido se sustituye con los datos correspondientes del informante anterior dentro de la misma clase de imputación definida por el tamaño del hogar y las características sociodemográficas. En particular, el ingreso individual omitido se sustituye por el ingreso de un donante que se encuentre en una situación semejante de empleo y nivel de instrucción.

87. El método de cálculo de las ponderaciones del diseño y el de la posestratificación son semejantes al de la muestra de la LFS. Además, se establecen factores específicos de expansión para compensar los diferentes períodos de referencia. El coeficiente del nuevo cálculo es básicamente el coeficiente del período de referencia de la encuesta (un año) con el período de referencia de la variable individual. Se necesitan también ponderaciones especiales al combinar los datos de tres años consecutivos. En el cálculo de una fecha específica se utilizan, por tanto, los datos de los tres años, de manera que la mitad de los datos hace referencia al período anterior a esa fecha y la otra mitad al período posterior.

88. La metodología y los resultados de la HBS se describen en las publicaciones antes citadas sobre la LFS.

Conclusiones

89. Antes de que se consiguiera la independencia en 1991, las encuestas de hogares por muestreo no eran un instrumento frecuente de recopilación de datos en Eslovenia. No obstante, a diferencia de otros países en transición, Eslovenia había realizado periódicamente encuestas HBS a partir de mediados del decenio de 1960, y la serie de LFS anuales desde los últimos años ochenta.

90. Después de la independencia, la Oficina de Estadística de Eslovenia experimentó una transición eficaz y sin sobresaltos. Ahora realiza periódicamente las series estándar de encuestas de hogares. Las encuestas socioeconómicas básicas (LFS, HBS) están casi totalmente armonizadas con las prescripciones de Eurostat (Oficina de Estadística de la República de Eslovenia, 2001). Se han realizado también otras encuestas de hogares, en particular las siguientes: Encuesta sobre el consumo de energía de los hogares (HECS), 1997; Encuesta sobre el empleo del tiempo (TUS), 2000/2001; Encuesta mensual sobre la actitud de los consumidores (CAS); Encuesta trimestral sobre los viajes de la población nacional (QSTDP), y Encuesta anual sobre delincuencia y la victimización (2000, 2001).

91. Hay todavía margen para mejorar el sistema de encuestas de hogares por muestreo. Eslovenia tiene un sistema estadístico basado en registros abundantes y precisos (datos fiscales, base de datos de empleados, base de datos de los seguros, etcétera) que puede aso-

ciarse de manera eficiente y eficaz con los sistemas geográficos y los datos del censo. Gracias a ello pueden obtenerse ventajas adicionales que podrían aplicarse al diseño de muestras óptimas y a la estimación.

2. Encuesta de hogares por muestreo de Estonia*

* Texto preparado por Imbi Traat, Instituto de Estadísticas Matemáticas, Universidad de Tartu (correo electrónico: imbi@ut.ee).

Introducción

92. La Oficina de Estadística de Estonia realizó dos grandes encuestas de hogares en 1995: la Encuesta sobre la población activa (LFS) y la Encuesta sobre el presupuesto de los hogares (HBS).

93. La HBS es una encuesta permanente; sus resultados se publican con periodicidad trimestral y anual. En 1999, la encuesta se revisó en el marco de un proyecto del Banco Mundial, sus diarios se modificaron y los procedimientos de muestreo y ponderación se armonizaron con los datos disponibles más recientes. La descripción de la encuesta puede verse en Traat, Kukk y Sõstra (2000) y, con mayor detalle, en Traat (1999).

94. La LFS se había aplicado en 1995 como encuesta independiente. La siguiente versión, en 1997, se llevó a cabo con una metodología modificada. Posteriormente se realizó con periodicidad trimestral hasta 2000, en que se convirtió en encuesta permanente. La encuesta se describe en Kurvits, Sõstra y Traat (2002) y en Statistical Office of Estonia (1999).

95. Además, la Oficina de Estadística de Estonia ha realizado muchas otras encuestas de hogares o basadas en la población, que corresponden a una serie de estudios semejantes en otros países europeos, y la información resultante se ha utilizado para realizar comparaciones nacionales e internacionales. Entre ellas se incluyen la Encuesta sobre educación de los adultos, de 1997; la Encuesta sobre el empleo del tiempo, de 1999-2000; la Encuesta sobre las condiciones de vida, de 1994 y 1999; y Comportamiento sanitario de la población adulta de Estonia, de 2000 (Kurvits, Sõstra y Traat, 2002).

96. A continuación se describen brevemente la Encuesta sobre el presupuesto de los hogares y la Encuesta sobre la población activa de Estonia.

Contenido de los datos

97. La HBS es una encuesta basada en diarios. Cada hogar incluido en la muestra facilita información sobre su consumo y gasto en alimentos durante una semana, así como sobre otros ingresos y gastos de un mes. Existe también una entrevista previa referente a la composición de los hogares y otras informaciones básicas y una breve entrevista posterior acerca de los cambios en la composición del hogar.

98. El programa de recopilación de datos de la LFS ha sido más amplio que el de una encuesta normal de población activa, sobre todo en 1995-1999, en que se recopiló información retrospectiva. En ella se recogía la situación laboral del informante (empleado, desempleado o inactivo) durante el intervalo transcurrido desde la encuesta anterior. En relación con dicho estado se anotaban la fecha inicial y final y otros datos pertinentes. El módulo estándar de la encuesta de población activa se centra en la semana de referencia y en él se hacen preguntas a las personas empleadas acerca de la ocupación, tiempo habitual y efectivo de trabajo, actividad económica de la empresa/organización, etcétera. A las personas desempleadas se les formulaban preguntas acerca de las medidas adoptadas para encontrar empleo, la continuidad de la búsqueda de él y las características del que estaban buscando, etcétera.

Recopilación de datos

99. El Departamento de Entrevistadores (establecido en 1994) de la Oficina de Estadística de Estonia se encarga de la recopilación de datos en varias encuestas. Los 15 coordina-

dores de condado organizan la labor de 130 entrevistadores, distribuidos en toda la nación. En las zonas rurales un entrevistador realiza por término medio entre 10 y15 entrevistas por encuesta en un mes, y en las zonas urbanas, entre 15 y 20, aunque su volumen de trabajo varía según el tamaño de las muestras regionales. El trabajo de entrevistador es un segundo empleo para aproximadamente la mitad de los entrevistadores. Se les paga por entrevista terminada y por los intentos de establecer contacto con personas que no responden.

100. La entrada y la codificación de los datos se realizan en la Oficina de Estadística utilizando el sistema de procesamiento de la encuesta denominado Blaise. La primera comprobación lógica se incluye en el programa de entrada de datos. Para el procesamiento de datos y otras comprobaciones más complejas se utilizan el programa Statistical Analysis System (SAS) en el caso de la LFS y FoxPro en el de la HBS.

101. El costo total en euros de la HBS y la LFS en 2000 fue de 153.000 y 128.000, respectivamente. Los salarios de los entrevistadores, el transporte y la comunicación representaron aproximadamente el 70% del costo total, y la entrada de datos, en torno al 15%.

Diseño de la muestra

102. La población objetivo de la HBS comprende todos los hogares de Estonia, excluidos los institutos. La población objetivo de la LFS comprende los residentes de 15 a 74 años de edad.

103. El marco muestral de ambas encuestas es el Registro de población. Las unidades de muestreo son personas y se muestrean sistemáticamente a partir de la lista del Registro de población. Los estratos (tres en el caso de la HBS y cuatro en el de la LFS) con tasas de muestreo diferentes se utilizan para obtener una mejor cobertura regional.

104. En ambas encuestas se utiliza información auxiliar del marco. La frecuencia de la dirección en el marco determina la probabilidad de inclusión de dicha dirección. La muestra se divide en dos partes, sometidas a reglas diferentes: la muestra de direcciones (los registros con direcciones completas) y la muestra de personas (los registros con direcciones desconocidas o incompletas). Existen direcciones desconocidas o incompletas en las regiones rurales, donde la dirección es sólo el nombre de la aldea sin ninguna otra información.

105. En la muestra de direcciones se incluyen todos los hogares que viven en las direcciones muestreadas. En la muestra de personas se incluyen únicamente los hogares con las personas seleccionadas. Se intenta localizar el hogar pertinente dentro de un condado. Aproximadamente el 15% de los hogares se muestrean mediante personas seleccionadas.

106. El diseño de la muestra utiliza el muestreo de probabilidad proporcional al tamaño (PPT), en que el tamaño es la frecuencia de la dirección en el marco o el tamaño del hogar (según la información recibida de éste). En el caso de la HBS, ésta es la muestra final. Sus probabilidades de inclusión PPT se utilizan para deducir estimadores. En cuanto a la LFS, ésta es la muestra de la primera fase de hogares/direcciones cuyo número de miembros en edad de trabajar, si no se puede conocer a través del Registro, se determina con ayuda de las autoridades locales. El objetivo del muestreo de la segunda fase es obtener una muestra de probabilidad igual de hogares (y sus miembros). En la muestra final se incluyen todos los hogares (direcciones) con un miembro en edad de trabajar y, mediante muestreo sistemático, la mitad de los hogares con dos miembros en edad de trabajar, un tercio de los hogares con tres miembros en edad de trabajar, etcétera.

107. En la muestra de la HBS actual se incluyen 820 hogares cada mes.

108. Desde 2000, los hogares de la LFS se han renovado de acuerdo con un plan de rotación de 2-2-2. Los hogares son entrevistados cuatro veces: durante dos trimestres consecutivos y, después de una pausa de dos trimestres, en los dos trimestres correspondientes del año siguiente. De acuerdo con este plan de rotación, en un trimestre cualquiera, el 25%

de los hogares participa por primera vez y el 50% son hogares que fueron entrevistados en el trimestre anterior. De esta manera hay una superposición del 50% entre los trimestres adyacentes y también entre los mismos trimestres de los años adyacentes.

Falta de respuesta

109. En las encuestas basadas en diarios, el aumento de la carga de la respuesta tiende a generar tasas de falta de respuesta más elevadas (véase el apartado A de este capítulo).

110. En general, las "negativas" representan cerca del 50% del total de la falta de respuesta, y los casos de "ausencia del hogar" son aproximadamente el 25%.

111. La tasa de falta de respuesta en la LFS ha sido siempre mucho menor que en la HBS. Además, la tasa de negativas ha aumentado debido a los límites temporales para el trabajo sobre el terreno como consecuencia de la transición al sistema de encuesta permanente. Además, la inclusión de los hogares en cuatro ocasiones durante la LFS ha dado lugar a un aumento de las tasas de falta de respuesta.

Ponderación

112. La ponderación se utiliza tanto en la HBS como en la LFS. En la HBS se calculan las ponderaciones de los hogares; en la LFS, las de las personas.

113. En la HBS, las tasas de respuesta y los niveles de ingresos/gastos determinan los seis grupos de ponderación. Las ponderaciones muestrales iniciales se multiplican por los inversos de las tasas de respuesta del grupo. Las ponderaciones de respuesta se calibran luego mediante distribuciones por sexo/edad (cinco clases) sobre la base de estadísticas demográficas conocidas.

114. En la LFS, las ponderaciones se formulan en una secuencia de varios pasos (Verma, 1995). La ponderación inicial de un informante es el tamaño de la población objetivo (personas entre 15 y 74 años) dividida por el número de informantes calculados dentro de cada uno de los cuatro estratos. Luego se forman seis grupos regionales de ponderación de tamaño razonablemente uniforme con tasas de respuesta diferentes R_j. Dentro de cada grupo, el factor de corrección de ponderación de un informante individual es $w_j^{(o)} = R/R_j$, donde R es la tasa de respuesta global (media). Luego se utiliza el método del coeficiente de inclinación con cinco iteraciones para calibrar las distribuciones de la muestra con respecto a los puntos de referencia de la población utilizando el sexo, la edad (grupos de cinco años) y el lugar de residencia (15 condados y la capital).

Parámetros y estimadores

115. La mayor parte de los parámetros estimados en la HBS y en la LFS son totales y coeficientes. Se utilizan los estimadores ponderados de Horvitz-Thompson o sus coeficientes.

116. Las estimaciones de la varianza se calculan utilizando SUDAAN. Como este programa no acepta el diseño exacto de la LFS y la HBS, se utiliza el diseño disponible más próximo de SUDAAN: muestreo conglomerado estratificado con sustitución de probabilidad desigual, con los hogares como conglomerados. Debido al supuesto de muestreo con sustitución, las estimaciones sobrevaloran ligeramente las varianzas verdaderas.

Novedades previstas para el futuro

117. El censo de población de 2000 contiene información sobre los hogares e individuos de Estonia. El sistema de ponderación de la HBS y de la LFS se revisará teniendo en cuenta los datos disponibles del censo, que reflejan la situación demográfica de Estonia con mayor precisión que los datos utilizados anteriormente.

118. Se procurará mejorar otras fases de la encuesta. Por ejemplo, en 2002, se comprobó para la LFS un nuevo método de recopilación de datos: la entrevista por teléfono con ayuda de computadora (CATI), con 10 computadoras portátiles. Se realizará un ensayo final de entrevista cara a cara en el primer contacto y de entrevista telefónica en las tres entrevistas posteriores.

3. Diseño y realización de la Encuesta sobre el presupuesto de los hogares y la Encuesta de población activa en Hungría*

* Texto preparado por Ödön Éltetö y László Mihalyffy, Oficina Central de estadística, P.O. Box 51, H-1525 Budapest, Hungría.

Encuesta sobre el presupuesto de los hogares

119. La HBS tiene una larga tradición en Hungría. Comenzó en el decenio de 1950 y estuvo basada inicialmente en muestras por cuotas. Luego utilizó el diseño basado en la probabilidad en el contexto del Sistema unificado de encuestas de hogares (USHS) a mediados del decenio de 1970. El marco muestral ha consistido siempre en distritos de empadronamiento (DE), actualizados después de cada censo decenal, el más reciente de los cuales tuvo lugar en 2002. Entre 1976 y 1982, la HBS había sido una encuesta permanente; entre 1983 y 1991, se realizó cada dos años, y desde 1993 ha vuelto a ser una encuesta permanente.

120. La muestra de la HBS se selecciona en tres etapas, con excepción de las ciudades autorrepresentadas, es decir, las que tienen 7.000 o más viviendas, en cuyo caso el proceso de selección tiene únicamente dos etapas. Cuando se trata de localidades no autorrepresentadas, las unidades primarias de muestreo (UPM) son las localidades, las unidades secundarias (USM) son los DE y las unidades últimas de muestreo son las viviendas. En las ciudades autorrepresentadas, las UPM son los DE.

121. Las localidades se estratifican por tamaño, lo que da lugar a ocho estratos, y por condado. En general, la muestra no se asigna proporcionalmente a los estratos. La tasa de muestreo es más baja en las localidades pequeñas que en las grandes ciudades, sobre todo en Budapest. El tamaño de la muestra anual se distribuye uniformemente a lo largo de los meses.

122. A los hogares que consienten en participar en la encuesta se les pide que informen sobre sus ingresos y gastos diarios durante un período de un mes. Durante ese período, los entrevistadores recopilan datos adicionales sobre el hogar, como la edad y la estructura ocupacional del mismo, el tipo, tamaño y equipamiento de la vivienda, los bienes duraderos de consumo, etcétera. Además, al comienzo del año siguiente el entrevistador visita de nuevo el hogar para solicitar a los miembros información sobre gastos menos frecuentes y más elevados efectuados durante el año y sobre ciertos tipos de ingresos anuales.

123. El hecho de que cada dos años los entrevistadores visiten todos los hogares de su DE para recopilar datos demográficos y económicos, como el tamaño del hogar, la edad, el nivel de instrucción y la actividad económica del jefe de hogar, constituye un aspecto importante de la HBS. Estos datos se utilizan fundamentalmente para la sustitución: debido a la tasa más bien alta de falta de respuesta, está autorizado el uso de hogares sustitutos (dos en las grandes ciudades y uno en los demás lugares). El hogar sustituto se selecciona dentro del mismo estrato que el hogar seleccionado originalmente y dentro del mismo DE asignado al entrevistador original.

124. Todos los años, la rotación de un tercio de los hogares es tal que el tamaño de la muestra dentro de cada DE se mantiene constante (seis viviendas). Como los hogares que no responden pueden sustituirse, el número real de hogares que cooperan en la encuesta puede ser mayor o menor de los seis iniciales. Por ello, la tasa de rotación de un DE dado puede ser superior o inferior a un tercio. Un hogar que ha participado en la encuesta durante tres años consecutivos queda excluido permanentemente.

125. En 2000, la muestra de la HBS incluyó casi 1.980 distritos de empadronamiento en 262 localidades, y el número de hogares inicialmente seleccionados fue de 11.862.

126. Como los entrevistadores se encuentran también con frecuencia con casos de rechazo u otros tipos de falta de respuesta en las direcciones sustitutas, el número de entrevistas finales es menor del tamaño de la muestra planificado. Por ejemplo, en 2000, en vez de los 11.862 (1.977 x 6) hogares, sólo realizaron la encuesta 10.191, y para conseguir este resultado los entrevistadores tuvieron que visitar nada menos que 17.243 direcciones. Las tasas de falta de respuesta habían aumentado después de 1993, alcanzando el 43,3% en 1996, y luego disminuyeron ligeramente. En 2000, la tasa total de falta de respuesta fue del 39%, y los rechazos representaron casi el 27%; el resto correspondió a viviendas vacías, ausencias del hogar, direcciones inválidas y otros factores. Dado el problema de la falta de respuesta, el logro del tamaño previsto de la muestra es particularmente difícil en la capital y en algunas grandes ciudades. Aunque hasta final de 2002 los hogares que cooperaban habían recibido un incentivo monetario para que facilitaran datos, la cantidad no era lo bastante considerable como para motivar a muchos hogares a cooperar con la encuesta, por lo que el programa de incentivos se canceló. No obstante, se produjo un cambio a favor en el sistema de remuneración de los entrevistadores, que los motivó a intensificar sus esfuerzos por persuadir a los hogares a que cooperaran en la encuesta. En términos globales, la tasa de negativas a participar disminuyó del 34,4% en 1996 al 26,9% en 2000.

127. El diseño de la muestra de la HBS garantiza las condiciones necesarias para utilizar el conocido estimador de Horvitz-Thompson. Los totales son sumas ponderadas de las observaciones, y las ponderaciones del diseño son los recíprocos de las probabilidades de inclusión. En cada uno de los 98 estratos del diseño de la muestra de la HBS, la ponderación del diseño es única, y se define como el coeficiente entre el número de viviendas no vacantes del estrato de la población y el número de entrevistas terminadas. Debido a la falta de respuesta de unidades completas, y también a las posibles deficiencias de cobertura, las ponderaciones del diseño no son adecuadas para calcular los datos de la HBS, por lo que deberían utilizarse ponderaciones calibradas. Durante el proceso de calibración, las ponderaciones del diseño se ajustan utilizando las siguientes variables auxiliares:

- Grupo de edad-sexo (2 × 4 categorías);
- Actividad económica (9 categorías);
- Nivel de instrucción (3 categorías);
- Tipo de hogar (3 categorías).

128. En el caso de los datos trimestrales se realiza la calibración en tres grandes zonas: la capital, las ciudades con derechos de condado y el resto del país. En cuanto a los datos anuales, el desglose de las zonas para la calibración es más detallado. Se consideran también las siete regiones del condado (regiones de nivel, II de la Nomenclatura de Unidades Territoriales Estadísticas (NUTS), en términos de Eurostat).

129. Las ponderaciones calibradas de la HBS se calculan utilizando el procedimiento de ponderación del coeficiente de inclinación generalizado.

130. Las estimaciones de los errores de muestreo para las partidas detalladas de ingresos y gastos obtenidas a partir de los datos de la HBS se calculan y publican periódicamente. Los cálculos se realizan utilizando la opción *jackknife* estratificada del programa VPLX elaborado por el R. E. Fay. En aplicaciones futuras se prevé la utilización del método *bootstrap*, en particular en el caso de los cuantiles estimados.

131. La HBS es una de las encuestas más costosas de la Oficina Central de Estadística. En 2000, los gastos directos en la encuesta —con exclusión de los salarios del personal de la oficina central y de las oficinas de los condados— a saber, las remuneraciones de los entrevistadores, los incentivos a los hogares cooperadores y los costos materiales, ascendieron a 84.769.000 forint húngaros (Ft), que correspondían a aproximadamente a 326.000 euros.

132. El efecto del diseño en 2000 fue aproximadamente de 2 en el caso de los ingresos disponibles netos, 2,5 en el de los gastos para alimentos y 2 en el total de gastos personales.

133. Los resultados de la encuesta se publican anualmente en húngaro y en inglés, con un breve análisis de los datos. La última publicación con los datos de la HBS de 2001 apareció en 2002 con el título *Household Budget Survey: 2001 Annual Report* (Central Statistical Office, Budapest, 2002). La publicación está también disponible en cederrón.

Encuestas de población activa

134. La LFS es una nueva encuesta de hogares introducida por la Oficina Central de Estadística en 1992. Su muestra se seleccionó en 1991 utilizando como marco el censo de 1990. Se consideraron como ciudades autorrepresentadas las de 15.000 habitantes o más. La muestra inicial de un trimestre constaba de 9.960 DE en 670 localidades con 3 direcciones de cada DE, lo que dio lugar a una muestra trimestral de 9.960 x 3 = 29.880 direcciones.

135. En la segunda mitad del decenio de 1990 aumentó la demanda de datos más pormenorizados y fiables de la LFS regional, y el tamaño de la muestra se incrementó un 40%. Creció también el número de localidades incluidas en la muestra, especialmente el número de DE. En 2000, la muestra contenía 12.829 DE de 754 localidades y, por lo tanto, se visitaron trimestralmente casi 36.500 hogares. Puede encontrarse información más detallada sobre la muestra ampliada en Éltetö (2000).

136. En la actualidad, la recopilación de datos tiene lugar cada mes, y se ha tomado como período de referencia la semana que contiene el duodécimo día del mes, mientras que la semana siguiente sería el período de recopilación de los datos. Los datos de la LFS se recogen sobre todo mediante entrevistas directas utilizando cuestionarios tradicionales de papel, aunque hay planes de ir intensificando el uso de las entrevistas telefónicas, sobre todo en el caso de las entrevistas repetidas. En las direcciones de la muestra de la LFS pueden incluirse y son objeto de entrevistas todas las personas con edades comprendidas entre 15 y 74 años.

137. En virtud del sistema de rotación aplicado en la LFS, los hogares seleccionados se mantienen en la muestra durante seis trimestres consecutivos; después se excluyen. En otras palabras, por efecto de la rotación, cada trimestre se elimina un sexto de la muestra.

138. Las ponderaciones del diseño de la muestra de la LFS se calculan de la misma manera que en la HBS. Las ponderaciones finales de la muestra de la LFS se determinan también utilizando el enfoque del coeficiente de inclinación. En la calibración de las ponderaciones de la muestra de la LFS se utilizan las siguientes variables auxiliares en los 19 condados y en la capital:

- Edad-sexo (2×10 categorías);
- Residencia en ciudades con derechos de condado u otros lugares (2 categorías).

139. Los errores de muestreo de los datos trimestrales de la LFS contenidos en el "Cuadro inicial" se procesan con el modelo *jackknife* estratificado utilizando el programa VPLX. Los errores muestrales de los datos mensuales se calculan también, pero no se publican. En lo que respecta al error de muestreo, la LFS cumple los requisitos de precisión de Eurostat contenidos en el Reglamento del Consejo (CE) No. 577/98 de 9 de marzo de 1998.

140. En la LFS las tasas de falta de respuesta —en particular las de negativas a participar— son muy inferiores a las de la HBS. Desde el comienzo, hasta 1997, un ligero aumento de las tasas de falta de respuesta habría permitido llegar a un máximo del 14,3%. Después de 1997 la tasa total descendió, situándose en el 9,2% en 2000. La tasa de negativas a participar aumentó también al principio, alcanzando el 7% en 1996 y 1997, y luego disminuyó, hasta situarse en el 3,2% en 2000.

141. La LFS es una operación costosa. Los gastos directos de la encuesta en el año 2000 sumaron 109.802.000 forint, que corresponderían a 422.000 euros. No obstante, teniendo en cuenta que cada hogar se visita cuatro veces al año y que no recibe ningún incentivo por colaborar, los gastos por hogar son bastante más bajos que los de la HBS.

142. Debido sobre todo a que la muestra de la LFS contiene muchas más UPM que la HBS, el efecto del diseño es considerablemente más bajo. En 2001 fue de 1,4 en lo que respecta a la tasa total de desempleo, y de 0,8 en la tasa de participación femenina.

143. La LFS se complementa con un módulo en el que se abordan temas como la situación de las mujeres trabajadoras o de las que están disfrutando de licencia por maternidad, etcétera. Estos módulos se incluyen, por término medio, en tres de los cuatro trimestres. Uno de los tres módulos, en general el del segundo trimestre del año, incluye el tema recomendado por Eurostat para ese año. Tanto los cuestionarios básicos de la LFS como los correspondientes a los módulos de Eurostat contienen toda la información requerida por ésta.

144. Los datos trimestrales y anuales de la LFS se publican en boletines bilingües.

145. Cabe concluir que tanto la HBS como la LFS son encuestas de hogares muy importantes de la Oficina Central de Estadística de Hungría. Los datos de la HBS se utilizan no sólo para calcular las ponderaciones para el índice de precios de consumo, sino también para estimar el consumo de los hogares dentro de los cálculos de las cuentas nacionales para presentar los valores trimestrales y anuales del producto interno bruto (PIB). Además, sus datos son de importancia fundamental para algunas esferas de investigación, como las condiciones de vida de varios estratos sociales, las pautas de gasto de diferentes tipos de hogares y los cambios registrados en ese terreno, la demanda de consumo de los diferentes tipos de productos, etcétera. Conviene mencionar asimismo que para conseguir un mayor cumplimiento de los requisitos de Eurostat, los gastos de 2001 se han agrupado de acuerdo con el sistema de la Clasificación del consumo individual por finalidades (CCIF) (Naciones Unidas, 2000, Parte III).

146. Aunque la información sobre el número de personas sin empleo registradas puede obtenerse a través de otras fuentes, los datos de la LFS presentan características especiales por lo que respecta tanto al concepto como al nivel de detalle. La información sobre la situación efectiva y los cambios registrados en el mercado de trabajo contenida en la LFS es indispensable para el gobierno central, para los gobiernos locales y para los investigadores. La tasa de desempleo oficial basada en los datos de la LFS es uno de los indicadores económicos más importantes.

4. Diseño y realización de las encuestas de hogares en Letonia*

* Texto preparado por Janis Lapins, del Departamento de Estadística del Banco de Letonia; Edmunds Vaskis y Zaiga Priede, de la Oficina Central de Estadística de Letonia, y Signe Balina, de la Universidad de Letonia, Riga.

Encuesta sobre el presupuesto de los hogares de Letonia

147. La Encuesta sobre el presupuesto de los hogares (HBS) es una encuesta permanente que se ha llevado a cabo desde 1995. Se remodeló en mayo de 2001.

148. La HBS se introdujo con asistencia técnica del Banco Mundial en septiembre de 1995. Estaba ya en la fase preparatoria cuando se estableció el requisito de que los resultados deberían estar de acuerdo con las prescripciones de Eurostat.

Ámbito de la encuesta

149. La población objetivo de la HBS son todos los hogares privados de Letonia. Las personas que viven en hogares institucionales (residencias para ancianos, centros para niños discapacitados, residencias estudiantiles, hoteles, cuarteles, hospitales, sanatorios, instituciones penitenciarias, etcétera) y las personas sin hogar están excluidas de la encuesta actual.

Muestreo

150. La muestra representa toda la población así como sus grupos más característicos. Cada mes se entrevistan 342 hogares. Cada hogar incluido en la muestra es entrevistado una sola vez.

151. Se aplica el muestreo probabilístico estratificado en dos etapas. Los hogares se estratifican por grado de urbanización y por asignación geográfica. La asignación de la muestra entre los estratos se hace en forma proporcional al tamaño de la población dentro de los estratos. En las zonas urbanas se ha elegido como marco muestral el registro de población, mientras que en las áreas rurales se han utilizado las listas de hogares.

152. Seis distritos administrativos de Riga, la capital, junto con las seis grandes ciudades, forman 12 estratos autorrepresentados. Todas las demás ciudades se utilizan como UPM en las áreas urbanas restantes, que se distribuyen entre 10 estratos que se establecen combinando 5 regiones y 2 grupos de tamaño. En la primera etapa de muestreo, se seleccionan las UPM dentro de cada estrato con probabilidad proporcional al número total de habitantes. En la segunda etapa se seleccionan personas de 15 o más años de edad mediante muestreo aleatorio simple.

153. En la zonas rurales, los hogares se distribuyen entre cinco estratos o regiones geográficas. Por norma general, se utilizan como UPM las *pagasts* (parroquias civiles, que son los territorios rurales administrativos de menor tamaño); algunas de las *pagasts* pequeñas se agregan a un territorio adyacente. Dentro de cada estrato, las UPM se seleccionan con probabilidad proporcional al número de hogares. En la segunda etapa, los hogares se seleccionan utilizando el muestreo aleatorio simple.

Costo de la encuesta

154. La HBS es una de las actividades estadísticas más costosas. En la HBS de 2001, el costo de la encuesta por hogar fue de 24 lats letones (LVL) (aproximadamente, 40 dólares de los EE. UU.). Las partidas principales de gastos están relacionadas con las actividades sobre el terreno. La remuneración de los entrevistadores representó el 44% del costo total de la encuesta, seguida de los incentivos a los informantes (16%), salarios de los supervisores (14%) y costos de transporte (8%).

Error de muestreo

155. En la HBS, las varianzas de las estimaciones seleccionadas correspondientes a los ámbitos de mayor interés (capital y seis grandes ciudades, otras ciudades y zonas rurales) se calculan utilizando el programa SUDAAN. Tomando como base estos valores, se estiman las varianzas y los efectos del diseño de alcance nacional.

Falta de respuesta

156. El nivel total de la falta de respuesta fue del 26,1% en 2000. La principal razón de la falta de respuesta fue la negativa a participar, incluida la de los hogares que interrumpieron la participación durante el mes de la encuesta (46,0% de todos los casos de falta de respuesta), seguida de las personas ausentes del hogar (31,8%) o que no podían participar debido a enfermedad o a su avanzada edad (11,6%). El nivel de falta de respuesta fue mucho mayor en las zonas urbanas (31,9%) que en las rurales (12,2%).

157. Los hogares que se niegan a participar en la encuesta o que no responden a las preguntas de la misma, así como los que no se encuentran en la dirección indicada, pueden repercutir en la precisión de los resultados, y que no debe pasarse por alto. Con el fin de mantener el tamaño de la muestra en el nivel elegido, se aplicó el método de muestreo secuencial.

Cada hogar que se negaba a participar o que no respondía a las preguntas se sustituía por otro de la lista de reserva, al que se le realizaba posteriormente la entrevista.

Remodelación de la HBS en 2001-2002

158. La remodelación más reciente de la muestra de la HBS se realizó tomando como base el censo de población que se llevó a cabo en la primavera de 2000. Los instrumentos de la encuesta se modificaron significativamente y se introdujo el período unificado de referencia retrospectiva de 12 meses en el caso de los bienes duraderos, las adquisiciones y pagos realizados en forma ocasional, los ingresos estacionales procedentes del trabajo remunerado y los ingresos y gastos en efectivo correspondientes a la producción agrícola del hogar. La HBS anterior se terminó al final de 2000.

159. A partir de enero de 2002 se coordinaron las muestras de dos encuestas, la HBS y la LFS. En ambas encuestas la muestra anual de los hogares se distribuye uniformemente a lo largo del tiempo (participa en la encuesta el mismo número de hogares en cada una de las 52 semanas del año). La muestra de las UPM está también uniformemente distribuida en los territorios dentro de cada trimestre.

160. En la nueva HBS y en la LFS permanente se utiliza la misma red de entrevistadores. En las anteriores HBS y LFS se utilizaron redes de entrevistadores independientes. Además, los entrevistadores de las zonas rurales se contratan entre la población local. Con el nuevo diseño, los entrevistadores son móviles y pueden trabajar en diferentes territorios administrativos. Ello permite la distribución de la muestra en forma más amplia en los territorios rurales (en la nueva HBS, la muestra anual se distribuye en 208 UPM rurales diferentes). Al mismo tiempo, la carga de trabajo de los entrevistadores se distribuye ahora de manera más uniforme, y el transporte resulta más económico. La estructura revisada de entrevistadores de la Oficina Central de Estadística se instituyó en enero de 2002.

Encuesta de población activa de Letonia

161. Durante el período 1995-2001, la Encuesta de población activa de Letonia se llevó a cabo con periodicidad semestral, en mayo y noviembre. La LFS permanente revisada se realizó en enero de 2002.

162. La LFS de Letonia se preparó de conformidad con la metodología internacionalmente aprobada sobre encuestas de población activa de la Organización Internacional del Trabajo (OIT), que garantiza la comparabilidad de la información con otros países (Eurostat, 1998a; 1998b).

Ámbito de la encuesta

163. La población incluida en la LFS comprende todos los residentes de Letonia de 15 años o más que viven en hogares privados. Las personas que se encuentran en instituciones como residencias para ancianos, centros para niños discapacitados, hoteles, cuarteles, hospitales, sanatorios, instituciones penitenciarias, etcétera, así como las personas sin hogar, están excluidas de la encuesta.

164. Con el fin de aplicar las recomendaciones de Eurostat y de reducir los costos de la encuesta, se entrevista también a todos los individuos de este grupo de edad que viven en el mismo hogar con las personas incluidas en la muestra. El tamaño de la muestra nacional de una oleada de la encuesta es de 7.940 hogares.

165. Todas las preguntas de la encuesta hacen referencia a la semana civil (lunes a domingo) anterior al día de la entrevista. Los datos se recopilan normalmente mediante entrevistas directas, con lápiz y papel. Si un informante no quiere recibir al entrevistador en su casa, se le pide que responda a la entrevista por teléfono.

Muestreo

166. La muestra de las áreas urbanas se toma del registro de población. La de las zonas rurales está basada en listas completas de hogares. Desde 1998, la muestra rural se ha basado en el registro de hogares elaborado en la Oficina Central de Estadística de Letonia.

167. La LFS comprende 7 grandes ciudades, 32 ciudades de tamaño intermedio y todas las *pagasts*. En cada oleada de la encuesta se entrevistan casi 16.000 personas. Para la elaboración de la muestra, el procedimiento de muestreo en una sola etapa (en las grandes ciudades y zonas rurales) o en dos etapas (ciudades intermedias) se aplica con estratificación basada en la división territorial administrativa del país. En las áreas urbanas se selecciona una muestra aleatoria simple de personas de 15 o más años dentro de cada UPM seleccionada. En las zonas rurales se selecciona dentro de cada *pagast* una muestra aleatoria simple de hogares.

168. De acuerdo con el plan de rotación para la muestra de la LFS, las personas de cada hogar se incluyen en la encuesta tres veces. Dentro de cada oleada de la encuesta, la tasa de sustitución de la muestra es de un tercio de los hogares en cada ciudad grande, ciudad intermedia o *pagast*.

Falta de respuesta

169. La tasa total de falta de respuesta alcanzó el 10,1% en 2000. En las zonas rurales fue más baja (sólo el 8,5%) que en las zonas urbanas (11,4%). El porcentaje de negativas a participar en las zonas rurales fue particularmente reducido (sólo del 0,5%, aproximadamente). Para aumentar la tasa de respuesta se ha recurrido al método de las entrevistas mediante sustitutos. Este procedimiento se aplicó aproximadamente en la tercera parte de las entrevistas realizadas.

Imperfecciones del marco

170. No todas las personas de la muestra vivían en la dirección que, según el registro, sería su unidad de vivienda. Es costoso, lento y algunas veces prácticamente imposible localizar y entrevistar a esas personas en el lugar donde viven realmente, por lo que los entrevistadores tienen que realizar la entrevista con los hogares que viven de hecho en las direcciones de la muestra. El análisis de los casos de falta de participación revelaba que sólo el 2,0% de todos esos casos (2,3% en las zonas rurales) parecían estar relacionados con alguna imperfección del marco (viviendas vacías, casas demolidas, direcciones inexistentes, etcétera).

Remodelación de la LFS en 2001-2002

171. El cuestionario de la LFS se revisó en 2001 para hacerlo totalmente compatible con las prescripciones de la Unión Europea. La LFS se lleva a cabo ahora como encuesta permanente.

172. Desde enero de 2002 se han introducido también cambios significativos en el diseño de la muestra para la LFS. Se utiliza la misma red de entrevistadores para la LFS y para la HBS. En consecuencia, a partir de enero de 2002 se coordinaron las muestras de ambas encuestas. Se prevé que la coordinación de las muestras de estas dos principales encuestas de hogares promoverá una utilización más eficaz de los recursos de la encuesta.

173. La capacitación de los entrevistadores tuvo lugar en diciembre de 2001. La LFS permanente comenzó en enero de 2002.

Otras encuestas de hogares

174. La Encuesta sobre las condiciones de vida (LCS) se puso en marcha en 1994 y en 1999, en el marco del Proyecto Norbalt, financiado por el Gobierno de Noruega, y en estrecha cooperación con el Fafo Institute (Instituto de Ciencias Sociales Aplicadas de Oslo).

175. En la segunda mitad del decenio de 1990 se iniciaron otras encuestas de hogares, entre ellas la Encuesta sobre familia y fecundidad (1995), la Encuesta sobre el empleo del tiempo (1996), la Encuesta sobre la confianza de los consumidores (1993-1999), la Encuesta sobre el consumo de energía por los hogares (1996), la Encuesta sobre el turismo interno (1998), la Encuesta sobre el uso de computadoras personales en los hogares (1998), la Encuesta especial sobre el módulo de la pobreza (1998) y la Encuesta sobre la actitud ante los problemas del suicidio (1999), entre otras.

176. Desde 1996 se ha realizado en tres o cuatro ocasiones al año la encuesta sobre viajes transfronterizos. En ella se consideran los desplazamientos de viajeros, tanto de residentes letones que regresan del extranjero como de extranjeros que abandonan Letonia.

177. En general, los resultados de las encuestas se publican en letón e inglés y en forma impresa y electrónica. Para las actividades de investigación, la Oficina Central de Estadística ha garantizado a los usuarios de Letonia y del exterior el acceso a archivos anónimos de microdatos.

Observaciones finales

178. Esperamos que el desarrollo del nuevo servicio de entrevistadores móviles y altamente cualificados permita una mayor flexibilidad en la planificación y ejecución de las nuevas encuestas por muestreo y otras encuestas especiales.

179. La Oficina Central de Estadística tiene también previsto introducir metodologías modernas de recopilación de datos. Una de las primeras medidas será iniciar la aplicación de la tecnología de la entrevista personal con ayuda de computadora (CAPI) en los cinco próximos años.

5. Encuestas de hogares por muestreo en Lituania*

* Texto preparado por Danute Krapavickaite, del Instituto de Matemáticas e Informática (Calle Akademijos, 4, LT 2600 Vilna) y el Departamento Lituano de Estadística (Avenida Gedimino, 29, 2746 Vilna).

Introducción

180. La Encuesta sobre el presupuesto de los hogares (HBS) fue la primera encuesta por muestreo realizada en Lituania. Se llevó a cabo por primera vez durante un período de 12 meses en 1936-1937. La HBS fue la única encuesta periódica por muestreo utilizada para obtener estadísticas relativas a la economía planificada de Lituania. Una vez que Lituania consiguió la independencia en 1990, comenzó a instaurarse la economía de mercado. Hubo que introducir un nuevo cuestionario para recopilar más datos y elaborar un nuevo diseño muestral para incluir al sector privado y los resultados publicados tuvieron que remodelarse para ofrecer a los usuarios datos comparables con los resultados de otros países. La principal revisión de la HBS se llevó a cabo con ayuda de expertos del Banco Mundial en 1996, como se describe en Šniukštiene, Vanagaite y Binkauskiene (1996). El diseño de la muestra y el método de estimación no han cambiado desde entonces.

181. La otra encuesta periódica de hogares es la Encuesta de población activa (LFS), que comenzó en 1994. El registro de población se modernizó en 1996, y desde entonces se ha utilizado para la selección de la muestra como marco muestral en la mayoría de las encuestas de hogares, incluida la LFS.

182. Otras encuestas de hogares, muchas de las cuales se realizaron con carácter excepcional, abarcaban temas como las condiciones de vida (1997), el empleo del tiempo (1998), los ancianos (1999), el consumo de energía de los hogares (1997), la accesibilidad de los servicios de atención de la salud (1998) y la disponibilidad de computadoras en los hogares (2000).

Estimaciones y errores de la Encuesta de población activa

Diseño muestral

183. La población de la LFS está integrada por residentes de Lituania con edad superior a 15 años. La muestra se construye de la manera siguiente: una vez seleccionada una muestra aleatoria simple de aproximadamente 3.000 personas tomadas del registro de población, se agregan los miembros de sus hogares, aun cuando no estén en el registro. La proporción de mujeres entre los informantes ha sido del 52,5%.

Rotación de la muestra

184. Con el fin de evitar grandes cambios en los resultados de las diferentes encuestas, en cada una de ellas rota únicamente un tercio del total. Cada hogar seleccionado participa en dos encuestas, queda excluido en una, realiza otra más y luego sale del sistema.

Precisión de las estimaciones

185. La distribución de los informantes de la encuesta por áreas urbanas/rurales, edad y sexo difiere ligeramente de las distribuciones correspondientes basadas en los datos del censo. La posestratificación de la muestra se procesó en 12 grupos de edad, 2 grupos de sexo y 10 condados, con un total de 240 grupos de ponderación.

186. Se utilizan diferentes sistemas de ponderación para estimar el número de personas empleadas y desempleadas. Con el fin de mejorar la precisión de las estimaciones de los desempleados se utilizan también para la posestratificación los índices de las oficinas de empleo. En Krapavickaitë, Klimavicius y Plikusas (1997) se describe la estimación de la varianza en el diseño de muestras de tamaño fijo.

Costo de la encuesta

187. El costo de una encuesta es de aproximadamente 70.000 litai de Lituania (tipo de cambio en dólares de 2000: 1 dólar = 4 litai). La impresión de los cuestionarios y su distribución entre los informantes representan el 14% del costo total, y el 86% corresponde a la remuneración de los entrevistadores, los gastos de transporte de éstos y el costo del envío por correo de los cuestionarios terminados a Vilna. No se incluyen los gastos relacionados con el trabajo metodológico de diseño de la muestra y de preparación del cuestionario, selección de la muestra y entrada, edición y procesamiento de los datos.

Encuesta sobre el presupuesto de los hogares

Diseño de la muestra

188. La HBS es una encuesta permanente. La muestra se establece una vez al año, se divide en 12 partes y se distribuye por meses. Todos los hogares participan en la encuesta durante un mes. La población de los hogares privados de Lituania se divide en tres estratos, de acuerdo con el tipo de residencia. Se selecciona una muestra aleatoria simple de 4.476 personas de 16 años o más a partir del registro de población de las grandes ciudades: Vilna, Kaunas, Klaipėda, Šiauliai y Panevėžys. Se obtiene una muestra aleatoria de 20 conglomerados con probabilidad proporcional al tamaño a partir de los 140 conglomerados de pequeñas ciudades, y una muestra aleatoria de 33 conglomerados con probabilidad proporcional a su tamaño a partir de la población de 463 conglomerados en la primera etapa. De cada uno de los conglomerados seleccionados se extrae una muestra aleatoria simple de personas. Se entrevista a todas las personas que residen en los hogares seleccionados. Si hay varios hogares en la misma dirección, se incluye en la muestra el hogar de la persona con la fecha de nacimiento más próxima.

Precisión de las estimaciones

189. Para la estimación de la HBS se utilizan ponderaciones del diseño. Los efectos del diseño de las estimaciones son superiores a la unidad. Ello significa que en el futuro habrá que utilizar información auxiliar para obtener estimaciones más precisas.

Costo de la encuesta

190. El costo anual total de la encuesta es de aproximadamente 900.000 litai, desglosados como sigue: 61%, remuneración de los entrevistadores; 18%, impuestos; 14%, pago a los hogares; 5%, transporte, y 2%, otros gastos.

Difusión de los resultados de las encuestas

191. Los resultados de las encuestas los publica Estadísticas de Lituania, y los principales de ellos aparecen en la revista mensual *Economic and Social Development in Lithuania*. Todos los resultados se difunden en números especiales dedicados a temas como la población activa, el empleo y el desempleo (datos de las encuestas) y los ingresos y gastos de los hogares.

Observaciones finales

192. Según los resultados provisionales del censo de población y vivienda de 2001, la población total de Lituania se estima en 3.491.000 residentes habituales. Esta cifra es inferior en 202.000 personas a la obtenida de los datos demográficos publicados el 1 de enero de 2001. Después de concluir los resultados del censo, Estadísticas de Lituania dispondrá de datos demográficos más fiables, que serán la base para mejorar las futuras encuestas de hogares, y se prevé que podrá reducirse el error sistemático de esas encuestas.

6. Encuestas de hogares en Polonia durante el período de transición*

* Texto preparado por Jan Kordos, de la Escuela de Economía de Varsovia, y Bronislaw Lednicki y Malgorzata Zyra, de la Oficina Central de Estadística (Avenida Niepodleglosci, 208, 00-925 Varsovia.

Introducción

193. Las encuestas de hogares de Polonia tienen una tradición relativamente prolongada (Glowny Urzad Statystyczny (Oficina Central de Estadística de Polonia) (GUS), 1987, 1998a, 1999; Kordos, 1985, 1996; Lednicki, 1982). En el decenio de 1980 comenzó la introducción gradual del Sistema Integrado de Encuestas de hogares (ISHS). Se puso en marcha en 1982 y se terminó en 1992 (GUS, 1987; Kordos, 1985).

194. El componente más importante de dicho sistema era la Encuesta sobre el presupuesto de los hogares (HBS), que estaba basada en un muestreo en dos fases, la rotación trimestral de los hogares a lo largo de un año y la rotación de un tercio de los hogares en los tres años siguientes. Ello significa que dos tercios de los hogares se incluían en el panel durante cuatro años consecutivos. Había también un ciclo de cuatro años en la encuesta de submuestras. Este programa se canceló en 1992. Al mismo tiempo, es decir, durante el período 1983-1992, las submuestras seleccionadas para la HBS se utilizaron para más de 30 encuestas sociales en las que se empleaban módulos temáticos.

195. Los intentos de integrar las encuestas de hogares realizadas en el decenio de 1980 facilitaron considerablemente su ajuste a las normas europeas (GUS, 1997). Se necesita todavía una integración mayor y nuevas mejoras de la metodología de la encuestas de hogares (Kordos, 1998).

Las encuestas de hogares en el período de transición

196. Las encuestas se ampliaron y modificaron considerablemente después de 1990. La HBS todavía sigue mejorando. En 1992 se introdujo por primera vez una nueva LFS, y se

han introducido nuevas encuestas de hogares, en particular la encuesta sobre las condiciones de vida, la encuesta sobre el estado de salud de los hogares, la encuesta sobre el empleo del tiempo, el microcenso de población y distintas encuestas de posenumeración.

Encuesta sobre el presupuesto de los hogares

197. Las encuestas sobre el presupuesto de los hogares se remontan a hace casi 45 años (GUS, 1999; Kordos, 1996; Lednicki, 1982). Se experimentaron varios métodos y se realizaron varios intentos de mejorar la ejecución. Al comienzo del decenio de 1990 se modificó la metodología de la encuesta. En el nuevo método para la realización de la HBS introducido en 1992 se modificó la clasificación de ingresos y gastos y la de los tipos socioeconómicos de la encuesta. Por primera vez se incluyeron en la encuesta todos los tipos de hogares individuales de Polonia, con un total aproximado de 32.000 hogares. En 1997 se trató de mejorar la integración de las encuestas de hogares[2]. En 2000 se llevó a cabo una remodelación de la HBS y se modificaron algunos componentes metodológicos (Kordos, Lednicki y Zyra, 2002). Está previsto mejorar la HBS y su integración con otras encuestas de hogares, en gran parte motivado por recomendaciones de Eurostat (Eurostat, 1997).

[2] Véase el reglamento interno No. 20 del Presidente de la Oficina Central de Estadística, del 30 de octubre de 1997, sobre el establecimiento del Grupo de Trabajo para la mejora de la metodología y la integración de las encuestas de hogares.

Encuesta de población activa

198. La encuesta sobre la actividad económica de la población se realizó en Polonia por primera vez en mayo de 1992, y se repitió trimestralmente hasta el tercer trimestre de 1999 (Szarkowski y Witkowski, 1994). Se preparó siguiendo las recomendaciones de la OIT. Cada trimestre se entrevistaban aproximadamente 24.000 hogares y personas de 15 años o más que eran miembros de esos hogares. En algunas ocasiones se incluyeron en la encuesta módulos sobre temas sociales seleccionados, lo que amplió considerablemente la oportunidad de hacer análisis sociales y económicos y la variedad de los resultados publicados.

199. Los resultados de la encuesta se publican cada trimestre. La remodelación de la LFS tuvo lugar en 1999 para ajustarla a la nueva división administrativa del país y mejorar su eficiencia de acuerdo con las prescripciones de Eurostat (Eurostat, 1998b; Verma, 1995).

Microcenso de población y vivienda de 1995

200. En el último decenio del siglo se llevaron a cabo varias encuestas de hogares especiales, la mayor de las cuales fue el Microcenso de 1995. En mayo de ese año se realizó una encuesta por muestreo en gran escala (microcenso) de la población y la vivienda (Bracha, 1996; GUS, 1998a). Éste era el tercer microcenso, precedido por los de 1974 y 1984. Conviene añadir que los censos ofrecen una oportunidad de recopilar datos sobre las personas discapacitadas, la migración y otros temas relacionados con las ciencias sociales.

201. El microcenso de 1995 abarcó el 5% de la población, es decir, casi 600.000 hogares. El censo de población y vivienda completo se llevó a cabo en mayo y junio de 2002; el anterior había tenido lugar en 1988.

Encuestas sobre las condiciones de vida

202. Además de la HBS, en 1997 se tomó la decisión de realizar una encuesta sobre diversos aspectos de las condiciones de vida de la población (Kordos, Lednicki y Zyra; 2002). La encuesta se preparó con gran esmero en cooperación con expertos del Instituto Nacional Francés para los estudios estadísticos y económicos (Institut national de la statistique et des études économiques (INSEE)) y se realizó con una gran muestra por primera vez a mediado de 1997. Se repitió en menor escala cada año utilizando submuestras de panel y en mayor escala cada varios años.

203. En total, en la encuesta intervinieron 12.524 hogares, y la tasa de respuesta fue del 87% (hogares) y del 86% (personas adultas). A mediados de 1998 se repitió la encuesta en menor escala.

204. La muestra de 1999 constaba de dos submuestras: la submuestra seleccionada en 1998 (panel), y una nueva submuestra, cuyo tamaño era igual a la submuestra de panel de 1998. De esta manera, cada año había una submuestra de panel y una nueva submuestra seleccionada a partir del marco muestral actualizado.

205. En 2001 se llevó a cabo una nueva encuesta en gran escala sobre las condiciones de vida, con una muestra de aproximadamente 24.000 hogares, con 18.052 informantes y una tasa de falta de respuesta del 25%. La encuesta continuará hasta la introducción, en 2005, de una Encuesta sobre ingresos y condiciones de vida (EU-SILC), preparada de acuerdo con el programa de Eurostat (Eurostat, 2001).

Encuesta sobre el estado de salud de la población

206. Esta encuesta se realizó en abril de 1996 y se incluyeron en ella 192.000 hogares. La tasa de respuesta fue del 88,6%. Ésta fue la primera encuesta sobre el estado de salud de la población en Polonia realizada en tan gran escala.

207. La encuesta estuvo basada en las recomendaciones de la Organización Mundial de la Salud (OMS), que hacen posible la comparación de los resultados con otros países europeos, sobre todo los Estados Miembros de la Unión Europea y los países de la región de la Comisión Económica para Europa (CEPE).

Encuesta sobre el empleo del tiempo

208. La Oficina Central de Estadística de Polonia (GUS) realizó encuestas sobre el empleo del tiempo en 1969, 1976 y 1984 (Kordos, 1988b). En 1996, GUS llevó a cabo una encuesta en pequeña escala sobre el empleo del tiempo con una muestra de 1.000 hogares y con inclusión de personas de 10 años o más. Uno de los objetivos de la encuesta era verificar la aplicabilidad de la metodología propuesta por Eurostat (GUS, 1998b). En 2004 se llevará a cabo una encuesta en gran escala sobre el empleo del tiempo.

Aspectos metodológicos comunes de las encuestas de hogares

Marcos muestrales

209. Los censos de población son la base para los marcos muestrales utilizados por las encuestas de hogares en Polonia. Las UPM se elaboran utilizando distritos estadísticos de empadronamiento (DEE) o zonas de empadronamiento censal (ZE) normalmente ajustados a las demandas específicas de una encuesta. Las viviendas suelen servir como unidades secundarias de muestreo (USM). Las viviendas de los DEE o de las ZE se actualizan anualmente, y ello supone un aumento del número de viviendas debido a las nuevas construcciones, una disminución de su número por demoliciones y cambios en los límites de los distritos a causa de los cambios ocurridos en la división administrativa del país. En cada distrito, el marco muestral contiene información sobre las direcciones y estimaciones del número de miembros de la población y el número de viviendas (GUS, 1998a).

210. Para la selección de la muestra de la HBS y de la LFS hubo que fusionar los DEE o las ZE adyacentes para conseguir el tamaño mínimo necesario para cada UPM. Por ejemplo, se constituyeron 29.172 UPM para la HBS a partir de los 33.023 DEE (en las zonas urbanas, las unidades primarias de muestreo tenían al menos 250 viviendas; en las zonas rurales, 150).

Diseños muestrales de las encuestas de hogares

211. Normalmente, en cada encuesta de hogares se utiliza la selección muestral en dos etapas y las UPM se seleccionan con probabilidad proporcional al tamaño (PPT). La estratificación está basada en la región (*voivodship*), la pertenencia a zonas urbanas/rurales y, en algunos casos, el tamaño de la localidad. En las encuestas permanentes, como la HBS y la LFS, se ha utilizado una pauta de rotación diferente, y los resultados finales se ponderan para reducir al mínimo los efectos de la falta de respuesta.

Diseños muestrales para la HBS

212. En los últimos 45 años se aplicaron diseños muestrales diferentes para la HBS (GUS, 1999; Kordos, 1996; Lednicki, 1982). Aquí examinaremos el diseño muestral de la HBS más reciente, que ha tenido lugar desde 2000. En lo que respecta al período 1992-2000, los diseños muestrales se describen con detalle en Kordos, Lednicki y Zyra (2002).

213. Desde 2001, se han seleccionado dos submuestras de 675 UPM a partir de un total de 29.172 UPM. Las UPM se estratifican de acuerdo con las 16 *voivodships*, según el tamaño de la localidad. Las grandes ciudades constituyen estratos independientes. El número de estratos de cada *voivodship* oscila entre 3 y 12. En total, hay 96 estratos. La asignación de la muestra a los estratos es proporcional a la población total de las viviendas de cada estrato. Las UPM se seleccionan con probabilidad proporcional al número de viviendas, de acuerdo con la fórmula Hartley-Rao. En cada UPM se seleccionan 24 viviendas durante dos años (2 viviendas por cada mes, y las mismas viviendas se mantienen en ambos años). Además, en cada UPM se seleccionan 150 viviendas de manera independiente como submuestra de reserva, que se utilizarán en caso de falta de respuesta. Cada año se selecciona una nueva submuestra de 675 UPM para un período de dos años.

Ponderación en la HBS

214. Las tasas de falta de respuesta en la HBS suelen ser elevadas, y repercuten considerablemente en la estructura socioeconómica de los hogares de la muestra. Para reducir esos efectos, los resultados de la muestra se someten a ponderación.

215. En primer lugar, cada hogar de la muestra se pondera en proporción inversa a la probabilidad con que se seleccionó. Se utilizan ponderaciones de las fuentes externas. En el caso de la HBS, se aplican ponderaciones adecuadas adicionales de la LFS (para el tamaño de los hogares y la proporción urbana y rural de la población).

Método de estimación del error estándar

216. Hasta 2000 se utilizó el método del grupo aleatorio de estimación del error muestral. Desde 2001 se ha utilizado un método de semimuestras equilibradas.

Diseño muestral para la LFS y su remodelación en 1999

217. Para la LFS se seleccionó una muestra en dos etapas con estratificación. Las UPM eran las ZE en las ciudades; en las zonas rurales, los DEE (en algunos casos, las unidades de muestreo se crearon agrupando dos o más ZE o DEE adyacentes, para atenerse a las prescripciones sobre el tamaño mínimo). Las viviendas sirvieron como unidades muestrales de la segunda fase (Szarkowski y Witkowski, 1994).

Remodelación de la LFS en 1999

218. Desde el cuarto trimestre de 1999 la LFS se ha realizado como encuesta permanente. Las UPM y las USM se seleccionaron del mismo modo que en la encuesta anterior, pero se modificó la asignación de la muestra de acuerdo con las 16 *voivodships*. Para alcanzar una mayor precisión en las estimaciones por *voivodship*, el tamaño de la muestra de cada una

de éstas se asignó en forma casi proporcional a la raíz cuadrada del número de viviendas que la integraban. Los tamaños de los estratos creados dentro de las *voivodships* eran proporcionales al tamaño de las localidades.

219. Las UPM dentro de los estratos se seleccionaron con probabilidad proporcional al número de viviendas de una UPM. Luego se seleccionó un número determinado de viviendas (de cuatro a nueve) de cada UPM. Cada 13 semanas de un trimestre[3], los entrevistadores visitan un número determinado de viviendas de muestreo aleatorio (1.880-1.900) y recopilan datos referentes a la actividad económica durante la semana precedente. La encuesta abarca todas las personas de 15 años o más que viven en las viviendas seleccionadas. Cada semana se cambia una muestra de viviendas. Las muestras semanales son resultado de una división aleatoria de una muestra trimestral en 13 partes. La muestra trimestral oscila entre 24.440 y 24.700 viviendas (GUS, 2000).

[3] Según los reglamentos de Eurostat, el término "trimestre" tal como se aplica actualmente a la LFS es ligeramente distinto del trimestre civil: cada trimestre de la LFS consta de 13 semanas y comienza siempre en lunes. Así pues, el primer trimestre de 2000 duró del 3 de enero al 3 de abril.

220. Se aplica la siguiente pauta de rotación de los hogares: dos trimestres en la encuesta, dos trimestres fuera, otros dos trimestres dentro y salida definitiva del sistema (pauta de rotación 2–(2)–2).

Ponderación de los resultados de LFS

221. La ponderación se realiza en tres etapas (véase información más detallada en Kordos, Lednicki y Zyra, 2002).

Estimación de los errores estándar

222. Hasta 1999, los errores estándar de las estimaciones se calcularon de acuerdo con el método del grupo aleatorio. Desde la remodelación de la LFS en el cuarto trimestre de 1999, se ha utilizado la técnica de linealización de Taylor.

Costos de las encuestas de hogares

223. La Oficina Central de Estadística tiene un sistema para evaluar el costo de las encuestas de hogares. En cada encuesta por muestreo se evalúa el costo directo de la encuesta utilizando las experiencias anteriores sobre el terreno y algunas recomendaciones administrativas. En esta evaluación se incluyen los costos de las entrevistas sobre el terreno, los gastos de viaje, los costos materiales, los servicios vinculados con la encuesta, los incentivos para aumentar la participación, los impuestos, etcétera (GUS, 2001), pero no la codificación y edición, la ejecución de programas de computadora, las contribuciones metodológicas de los costos indirectos y generales y el costo de personal involucrado en varios proyectos.

224. Como ejemplos, se presentan en el apartado A del presente capítulo los elementos de costos de la HBS y la LFS en Polonia en el año 2000.

Efectos del diseño

225. En las encuestas de Polonia, es decir, la HBS y la LFS, se calcularon los efectos del diseño en relación con varias características (Kordos, Lednicki y Zyra, 2002). Como práctica, y para su comparación con otros países, se calcularon los efectos del diseño con varios parámetros para los años 2000 y 2001.

226. En relación con algunas características de la HBS, los efectos del diseño y los errores estándar relativos (porcentaje presentado entre paréntesis) fueron los siguientes: ingreso total, 4,24 (1,1); gasto total, 4,16 (1,0); gastos en alimentos, 3,53 (0,4); vestido y zapatos: 2,72 (1,5); mantenimiento de la vivienda, 4,04 (1,3); atención de salud personal, 3,28 (1,7); transporte y comunicación, 2,16 (4,5); educación, 2,50 (3,9).

227. En la LFS, los efectos del diseño de 2000 y 2001 se calcularon con respecto al número total de desempleados en los diferentes grupos de la clasificación cruzada sobre la

base de las zonas urbanas/rurales, el tamaño de las localidades (clases de ciudad) y el nivel de instrucción. La mayor dispersión correspondió a las clases de ciudades, con efectos del diseño que oscilaban entre 1,7 y 3,55.

228. Como puede desprenderse de las estimaciones mencionadas, los efectos del diseño en lo que respecta a los datos de la HBS y la LFS fueron normalmente mayores de 1, y en algunas características fueron incluso superiores a 4. Por ello los errores estándar basados en los supuestos de muestras aleatorias simples solían infravalorar los errores estándar resultantes del diseño muestral complejo aplicado.

Falta de respuesta en las encuestas de hogares

229. Como se examina en el apartado A, las tasas de falta de respuesta aumentaron tanto en la HBS como en la LFS durante el último decenio. Las principales razones de este aumento fueron las negativas a participar y las "ausencias" del hogar. En la HBS, las tasas de negativa a participar pasaron del 10,2% en 1992 al 25,0% en 2000, y las tasas de ausencia subieron del 4,5% en 1992 al 14,5% in 2000.

230. En cifras medias anuales, las tasas de falta de respuesta en la LFS aumentaron de forma constante durante la mayor parte del período comprendido entre 1992-2000, pasando del 4,5% en 1992 al 22,1% en 2000, y lo mismo ocurrió con las tasas de negativa a participar, que pasaron del 2,0% en 1992 al 10,9% en 2000. Las tasas de falta de respuesta crecieron de forma significativa entre 1992-2000, y las dos razones principales fueron las negativas y las ausencias.

231. Las tasas de falta de respuesta en la LFS difieren dependiendo del tamaño de las localidades: la mayor tasa corresponde a Varsovia, y la menor a las zonas rurales. En el año 2000, las tasas anuales ponderadas de falta de respuesta en función del tamaño de las localidades fueron las siguientes: Varsovia: 54,5%; grandes ciudades (de 500.000 a 1 millón de habitantes): 32,6%; grandes ciudades (de 100.000 a 500.000 habitantes): 33,3%; ciudades (de 20.000 a 100.000 habitantes): 23,1%; ciudades (por debajo de 20.000 habitantes): 19,0%; y zonas rurales: 11,1%.

Observaciones finales

232. En este subapartado se han presentado descripciones generales de las encuestas de hogares en el período de transición de la Oficina Central de Estadística de Polonia (GUS), con especial atención a dos encuestas permanentes: la HBS y la LFS. GUS tiene una larga tradición de realizar encuestas de hogares y amplia experiencia en este terreno. Ello representó un beneficio al comienzo del período de transición, ya que facilitó la remodelación de las encuestas y el diseño de otras nuevas.

233. La asimilación de los resultados del censo de población y vivienda de 2002 constituirá una de las tareas más importantes de las encuestas de hogares en los próximos años. El censo permitirá obtener tanto marcos muestrales actualizados para las encuestas de hogares como información auxiliar para lograr una mayor precisión de las estimaciones y para los métodos de estimación de zonas pequeñas que son ahora objeto de estudio.

234. Se ha comenzado a preparar una nueva encuesta de hogares —la EU-SILC, que se introducirá en 2005 (Eurostat, 2001)— y se están mejorando las encuestas actuales para adaptarlas a las normas de la Unión Europea.

Agradecimientos

Los Presidentes de las oficinas centrales de estadística de Belarús, Bulgaria, Croacia, Eslovaquia, Eslovenia, Estonia, la Federación de Rusia, Hungría, Letonia, Lituania, Polonia, la República Checa, Rumania y Ucrania aceptaron presentar información sobre sus encuestas de hogares por

muestreo. El autor desearía manifestar su agradecimiento por los datos presentados y por las valiosas observaciones ofrecidas por numerosos colegas de esos países, así como su reconocimiento a los autores de las contribuciones al apartado B, que pretende complementar el material presentado en el apartado A.

El autor agradece asimismo los valiosos comentarios y sugerencias de algunos revisores académicos y de participantes en la reunión del Grupo de Expertos sobre análisis de las características prácticas de las encuestas en los países en desarrollo y en transición, celebrada en Nueva York del 8 al 11 de octubre de 2002.

REFERENCIAS BIBLIOGRÁFICAS

Bracha, Cz. (1996). Schemat losowania próby do Mikrospisu 1995 (Sample design for the 1995 microcensus). *Wiadomosci Statystyczne*, No. 3, págs. 12-18.

Deville, J.-C. y C. E. Särndal (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, vol. 87, págs. 376-382.

Éltetö, Ö. (2000). Enlargement of the sample of the Hungarian LFS to get reliable small area estimates for labour market indicators. *Statistics in Transition*, vol. 4, No. 4, págs. 549-563.

_____ y L. Mihalyffy (2002). Household surveys in Hungary. *Statistics in Transition*, vol. 5, No. 4, págs. 521-540.

Eurostat (1995). *The Future of European Social Statistics: Guidelines and Strategies.* Luxemburgo.

_____(1996). *The Future of European Social Statistics: Use of Administrative Registers and Dissemination Strategies.* Luxemburgo.

_____(1997). *Family Budget Surveys in the EC: Methodology and Recommendations for Harmonisation.* Population and Social Conditions 3, Methods E. Luxemburgo.

_____(1998a). *Labour Force Survey: Methods and Definitions*. Luxemburgo.

_____(1998b). *Labour Force Survey in Central and Eastern European Countries.* Methods and Definitions (Provisional). Luxemburgo.

_____(2001). Meeting of the Working Party Statistics on Income and Living Conditions (EU-SILC), 10 y 11 de diciembre de 2001. Luxemburgo. Documento de trabajo.

Garner, T. I. y otros (1993). Household surveys of economic status in Eastern Europe: an evaluation. En *Economic Statistics for Economies in Transition: Eastern Europe in the 1990s.* Washington, D.C.: United States. Bureau of Labour Statistics, págs. 309-353.

Glowny Urzad Statystyczny (GUS) (Oficina Central de Estadística de Polonia) (1971a). Badania statystyczne metodą reprezentacyjną w krajach socjalistycznych (Encuestas por muestreo en los países socialistas). *Biblioteka Wiadomości Statystycznch,* tomo 14 (Varsovia), pág. 220.

_____(1971b). Wybrane problemy metodologiczne badań reprezentacyjnych (Algunos problemas metodológicos de las encuestas por muestreo). *Biblioteka Wiadomości Statystycznch,* tomo 15 (Varsovia), pág.151.

_____(1987). *Problemy integracji statystycznych badan gospodarstw domowych* (Problemas de integración de las encuestas estadísticas de hogares). *Biblioteka Wiadomości Statystycznch,* tomo 34 (Varsovia).

_____(1997). *Stan dostosowania polskiej statystyki publicznej do standardów Unii Europejskiej. Harmonogram prac dostosowawczych* (Armonización de las estadísticas de Polonia con las normas de la Unión Europea). Varsovia. Mimeografiado.

_____(1998a). *Metodologia i organizacja mikrospisów* (Metodología y organización de microcensos). *Statystyka w Praktyce*. Varsovia.

_____(1998b). Budzet czasu ludnosci 1996 (Encuesta sobre el empleo del tiempo de 1996). En *Studia i Analizy Statystyczne.* Varvosia.

_____(1999). Metodyka badania budzetów gospodarstw domowych (Metodología de la encuesta sobre el presupuesto de los hogares). En *Zeszyty Metodyczne i Klasyfikacje.* Varsovia.

_____(2000a). Aktywnosc ekonomiczna ludnosci Polski: I kwartał 2000 (Encuesta sobre la población activa en Polonia: primer trimestre de 2000). *Informacje i Opracowania Statystyczne.*Varsovia.

_____(2001). *Zasady wyceny kosztów prac statystycznych realizowanych przez służby statystyki publicznej w roku 2003* (Principio de la evalución de costos de la labor estadística en las estadísticas oficiales en 2003). Varsovia. Mimeografiado.

Goskomstat (2000) *Methodologitcheskie polojenia po statistike* (Principio metodológico de las estadísticas), Vypusk treti (3a. ed.). Moscú, 294 págs.

Groves, R. M. (1989). *Survey Errors and Survey Costs.* Nueva York: John Wiley and Sons.

_____ y M. P. Couper (1998). *Non-response in Household Interview Surveys*. Nueva York: John Wiley and Sons.

Kish, L. y M. R. Frankel (1974). Inference from complex samples (with discussion). *Journal of the Royal Statistical Society, s*erie B, vol. 36, págs. 1-37.

Kordos, J. (1963). Seminarium statystyczne w Wiedniu (Seminario de estadística en Viena). *Przegląd Statystyczny*, No. 2, págs. 307-310.

_____(1970). Mozliwosci szerszego stosowania metody reprezentacyjnej w badaniach statystycznych krajów-czlonków RWPG (Posibilidades de una mayor aplicación de los métodos de muestreo en las investigaciones estadísticas de los países miembros del Consejo de Ayuda Mutua Económica). *Przegląd Statystyczny,* No. 1, págs. 33-50.

_____(1981). Problemy badań gospodarstw domowych: drugie spotkanie statystyków europejskich w Genewie (Problemas de las encuestas de hogares: segunda reunión de estadísticos europeos). *Wiadomosci Statystyczn,* No. 11, págs. 36-38.

_____(1982). Metoda rotacyjna w badaniach budzetow rodzinnych w Polsce (Método de rotación en las encuestas sobre el presupuesto de los hogares en Polonia). *Wiadomosci Statystyczne*, No. 9.

_____(1985). Towards an integrated system of household surveys in Poland. *Bulletin of the International Statistical Institute* (Amsterdam), vol. 51, libro 2, págs. 13-18.

_____(1988a). Jakość danych statystycznych (Calidad de los datos estadísticos). *Państwowe Wydawnictwo Ekonomiczne* (Varsovia), págs. 204.

_____(1988b). Time-use surveys in Poland. *Statistical Journal of the United Nations ECE*, vol. 5, págs. 159-168.

_____(1996). Forty years of the household budget surveys in Poland. *Statistics in Transition*, vol. 2, No. 7, págs. 1119-1138.

_____(1998). Social Statistics in Poland and its Harmonisation with the European Union Standards, *Statistics in Transition*, vol. 3, No.4, págs. 617-639.

_____(2001). Some data quality issues in statistical publications in Poland. *Statistics in Transition*, vol. 5, No. 3, págs. 475-489.

_____, B. Lednicki and M. Zyra (2002). The household sample surveys in Poland. *Statistics in Transition*, vol. 5, No. 4, págs. 555-589.

Krapavickaitë, D. (2002). The household sample surveys in Lithuania. *Statistics in Transition*, vol. 5, No. 4, págs. 591-603.

_____, G. Klimavicius y A. Plikusas (1997). On some estimators in cluster sampling. *Proceedings of the XXXVIII Conference of the Lithuanian Mathematical Society*, págs. 298-303.

Kurvits, M. A., K. Sõstra e I. Traat (2002). Estonian household surveys: focus on the labour force survey. *Statistics in Transition*, vol. 5, No. 4.

Lapins, J. y E. Vaskis (1996). The new household budget survey in Latvia. *Statistics in Transition*, vol. 2, No. 7, págs. 1085-1102.

Lapins, J y otros (2002). Household sample surveys in Latvia. *Statistics in Transition*, vol. 5, No. 4, págs. 617-641.

Lednicki, B. (1982). Schemat losowania i metoda estymacji w rotacyjnym badaniu budzetów gospodarstw domowych (Diseño muestral y método de estimación en las encuestas sobre el presupuesto de los hogares por rotación). *Wiadomosci Statystyczne*, No. 9.

Martini, A., A. Ivanova y S. Novosyolova (1996). The income and expenditure survey of Belarus: design and implementation. *Statistics in Transition*, vol. 2, No. 7, págs. 1063-1084.

Mihalyffy, L. (1994). The unified system of household surveys in the decade 1992-2001. *Statistics in Transition*, vol. 1, No. 4, págs. 443-462.

Naciones Unidas (1964). *Recomendaciones para la preparación de informes sobre encuestas a base de muestras (edición provisional*). Informes estadísticos, Serie C, No.1, Rev.2. ST/STAT/SER. C/1/Rev.2. Nueva York.

_____(1984). *Manual de encuestas de hogares (edición revisada*), Estudios de métodos, No. 31. Nueva York. No. de venta: S.83.XVII.13.

_____(2000). *Clasificaciones de gastos por finalidades: Clasificación de las funciones del gobierno (CFG*). *Clasificación del consumo individual por finalidades (CCIF*), *Clasificación de las finalidades de las instituciones sin fines de lucro que sirven a los hogares (CFISFL*), *Clasificación de los gastos de los productores por finalidades (CGPF*). Informes estadísticos, No. 84. No. de venta: S.00.XVII.6.

Oficina de Estadística de Estonia (1999). *Estonian Labour Force Survey 1998: Methodological Report*, Tallin.

Oficina de Estadística de la República de Eslovenia (2001). *Slovenian Statistical System: A Global Assessment,* 2001. A Phare project. Ljubljana.

Postnikov, S. (1953). O metodach otbora semei rabocich, slujašcich i kolchoznikov dla obsledovanjia ich budjeta (Sobre los métodos de selección de familias de trabajadores, empleados no trabajadores y empleados de cooperativas para la investigación de sus presupuestos). *Vestink Statistiki*, No. 3, págs. 14-25

Särndal, C-E., B. I. Swensson y J. Wretman (1992). *Model Assisted Survey Sampling*. Nueva York: Springer-Verlag.

Šniukstiene, Z., G. Vanagaite y G. Binkauskiene (1996). Household Budget Survey in Lithuania. *Statistics in Transition*, vol. 2, No. 7, págs. 1103-1117.

Szablowski, J., J. Wesolowski y R. Wieczorkowski (1996). Indeks zgodnosci jako miara jakosci danych: na podstawie wyników spisu kontrolnego do Mikrospisu 1995 (El índice de ajuste como medida de la calidad de los datos: sobre la base de la encuesta posterior al empadronamiento del microcenso de 1995). *Wiadomosci Statystyczne*, No. 4. págs. 43-49.

Szarkowski, A. y J. Witkowski (1994). The Polish Labour Force Survey. *Statistics in Transition*, vol. 1, No. 4, págs. 467-483.

Traat, I. (1999). *Redesign of the Household Budget Survey*: *Final Report of the Sampling Group*. Tallin: Oficina de Estadística de Estonia.

_____, A. Kukk y K. Sõstra (2000). Sampling and estimation methods in the Estonian household budget survey. *Statistics in Transition*, vol. 4, No. 6, págs. 1029-1046.

Unión Europea (1998). Reglamento No. 577/98 del consejo de 9 de marzo de 1998 relativo a la organización de una encuesta muestral sobre la población activa en la Comunidad. *Diario Oficial de las Comunidades Europeas*, 14/3/98, págs. I.77/3-I.77/7.

Vehovar, V. (1997). The Labour Force Survey in Slovenia. *Statistics in Transition*, vol. 3, No. 1 (junio), págs. 191-199.

_____(1999). Field substitution and unit non-response. *Journal of Official Statistics,* vol. 15, No. 2, págs. 335-350.

_____ y M. Zaletel (1995). Non-response trends in Slovenia. *Statistics in Transition*, vol. 2, No. 5, págs. 775-788.

Verma, V. (1995). *Technical Report on the Turkey Labour Force Survey*. Project TUR/86/081. Ginebra: Organización Internacional del Trabajo.

Wolter, K. M. (1985). *Introduction to Variance Estimation*. Nueva York: Springer-Verlag.