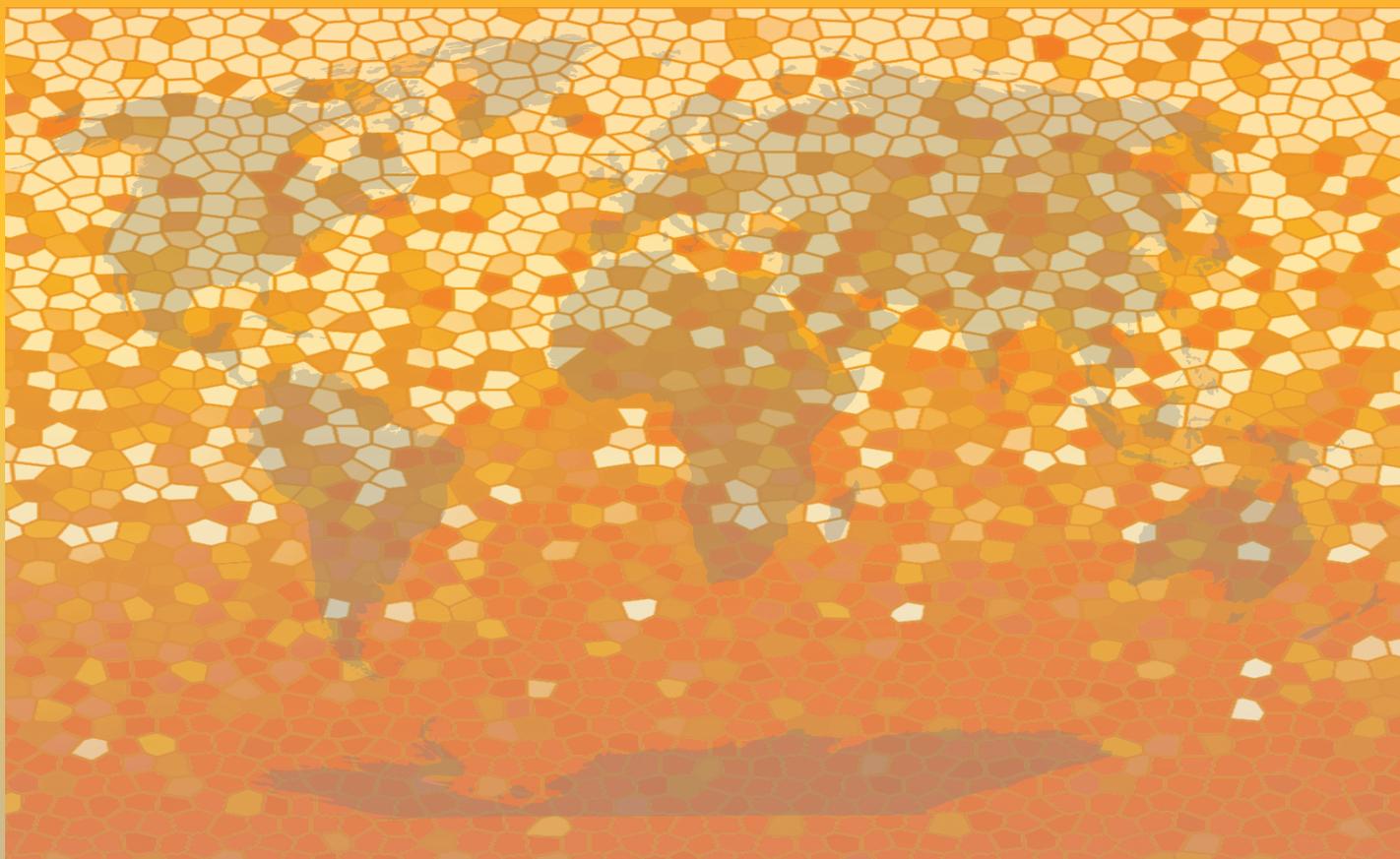


A s u n t o s e c o n ó m i c o s

y s o c i a l e s

Manual de revisión de datos de los censos de población y vivienda

Revisión 1



Naciones Unidas

Departamento de Asuntos Económicos y Sociales
División de Estadística

Estudios de métodos

Serie F No. 82/Rev.1

Manual de revisión de datos de los censos de población y vivienda

Revisión 1



Naciones Unidas
Nueva York, 2011

Departamento de Asuntos Económicos y Sociales

El Departamento de Asuntos Económicos y Sociales de las Naciones Unidas es un punto de contacto fundamental entre las políticas mundiales en las esferas económica, social y ambiental y la acción nacional. El Departamento trabaja en tres esferas relacionadas entre sí: i) compila, produce y analiza una amplia gama de datos e información de tipo económico, social y ambiental que aprovechan los Estados Miembros de las Naciones Unidas para examinar problemas comunes y hacer un balance de las opciones en materia de políticas; ii) facilita las negociaciones de los Estados Miembros en muchos órganos intergubernamentales sobre el curso a seguir en forma conjunta para abordar los desafíos mundiales actuales o en ciernes, y iii) asesora a los gobiernos interesados sobre las formas y los medios de traducir los marcos normativos desarrollados en las conferencias y cumbres de las Naciones Unidas en programas a nivel de países y, mediante la asistencia técnica, ayuda a aumentar la capacidad nacional.

Nota

Las denominaciones empleadas en esta publicación y la forma en que aparecen presentados los datos que contiene no entrañan, por parte de la Secretaría de las Naciones Unidas, juicio alguno sobre la condición jurídica de ninguno de los países, territorios, ciudades o regiones o de sus autoridades ni respecto de la delimitación de sus fronteras o límites.

El término “país” utilizado en esta publicación se aplica, según los casos, a determinados territorios o zonas.

Las denominaciones “regiones desarrolladas” y “regiones en desarrollo” se utilizan para facilitar la tabulación, sin que representen forzosamente un juicio sobre la etapa alcanzada por un determinado país o zona en el proceso de desarrollo.

Las firmas de los documentos de las Naciones Unidas se componen de letras mayúsculas combinadas con cifras.

ST/ESA/STAT/SER.F/82/Rev.1

PUBLICACIÓN DE LAS NACIONES UNIDAS

No. de venta: S.09.XVII.11

ISBN 978-92-1-361245-3

Copyright © Naciones Unidas, 2009

Reservados todos los derechos

Prefacio

Las Naciones Unidas han publicado a lo largo de los años una serie de manuales e informes técnicos para ayudar a los países a planificar y realizar censos de población y de vivienda más perfeccionados y eficaces en función de los costos. Los manuales e informes se revisan y se actualizan periódicamente a fin de reflejar los nuevos adelantos y problemas relativos al levantamiento de censos. La presente publicación es parte de una serie de manuales que se han elaborado para ayudar a los países a prepararse para las rondas de censos del año 2000 y más adelante. Los otros manuales de la serie son los siguientes:

- a) *Manual sobre gestión de censos de población y habitación* (publicación de las Naciones Unidas, No. de venta: S.00.XVII.15);
- b) *Manual de sistemas de información geográfica y cartografía digital*, Serie F, No. 79 (Publicación de las Naciones Unidas, No. de venta: S.00.XVII.12);
- c) *Guide for the Collection of Economic Characteristics in Population Censuses* (de próxima aparición).

En *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008) se examinan las ventajas de aplicar un sistema de control y mejoramiento de la calidad en las etapas iniciales del censo, que será decisivo para el éxito de todas las operaciones censales, así como la importancia del plan de revisión que debería formularse como parte del programa general del censo e integrarse a los demás planes y procedimientos censales. Será especialmente útil para los usuarios de este *Manual de revisión de datos de los censos de población y vivienda* referirse a los *Principios y recomendaciones*, donde se proporciona considerable información de los antecedentes sobre los procedimientos de revisión de datos descritos en los capítulos III, IV y V.

La finalidad de la presente publicación es ofrecer a los países un amplio panorama de la metodología de revisión de los datos de censos y encuestas y suministrar a los funcionarios interesados información sobre la aplicación de diversos enfoques de la revisión de los datos censales. También se propone alentar a los países a conservar un historial de sus experiencias de revisión, promover la comunicación entre los especialistas temáticos y los especialistas en procesamiento de datos y documentar las actividades llevadas a cabo en el censo o en la encuesta en curso, a fin de evitar duplicar esas actividades en el próximo censo o encuesta.

En el *Manual* se pasa revista a las ventajas e inconvenientes respectivos de la revisión manual y por ordenador. En los grandes censos la corrección manual no suele ser viable económicamente. Las condiciones de este tipo de corrección suelen especificarse en programas informáticos diseñados especialmente, que llevan a cabo un escrutinio automático de los errores y una atribución de valores basada en otra información sobre esa persona u hogar o sobre otras personas u hogares. Una parte importante del *Manual* trata de la corrección automática de los datos.

La revisión de los datos por ordenador desempeña un papel importante en la localización y corrección de errores. En la etapa de revisión por ordenador es posible efectuar comprobaciones de coherencia detalladas en consulta con los especialistas temáticos. Los errores observados pueden corregirse consultando los impresos originales o automáticamente. La revisión automática permite acelerar el procesamiento de los datos, pero también exige un minucioso control de la calidad de los datos que se introducen.

Esta publicación consta de una introducción y de cinco capítulos. En la introducción se describe el proceso censal y los distintos tipos de errores que se producen en un censo. El capítulo I abarca los aspectos básicos de la revisión de datos censales. En los capítulos II a V se describen procedimientos y técnicas para revisar los datos censales en distintas etapas del procesamiento. Las consideraciones técnicas, especialmente las pertinentes para la programación, se tratan en los anexos.

Aunque el presente *Manual* se concentra en la revisión de datos de los censos de población y de vivienda, muchos de los conceptos y técnicas que describe se aplican también a las encuestas.

Se agradece cordialmente la contribución de Michael J. Levin, del Harvard Center for Population and Development Studies. Merecen también reconocimiento Michael Bankier, Wesley Benjamin, Marcel Bureau, Sean Crowe, Sylvain Delisle y Darryl Janes, de Statistics Canada, que revisaron el texto provisional y ofrecieron valiosos comentarios para su versión final.

Índice

	<i>Página</i>
Introducción	1
A. Propósito del <i>Manual</i>	1
B. El proceso censal	2
C. Errores en el proceso censal	3
1. Errores de cobertura	3
2. Errores de contenido	3
a) Errores en el diseño del cuestionario	3
b) Errores de los empadronadores	4
c) Errores de los declarantes	4
d) Errores de codificación	4
e) Errores en la introducción de datos	5
f) Errores en la revisión por ordenador	5
g) Errores de tabulación	5
D. Estructura del <i>Manual</i>	6
I. La revisión de datos en los censos y en las encuestas	7
A. Reseña histórica de la revisión de datos	7
B. El equipo de revisión de datos	8
C. Prácticas de revisión de datos: datos revisados y datos no revisados .	9
D. Elementos básicos de la revisión de datos	12
1. Perjuicios de una revisión excesiva	15
a) Puntualidad	15
b) Costo	15
c) Distorsión de los valores reales	15
d) Una falsa sensación de seguridad	15
2. Tratamiento de los valores desconocidos	16
3. Modificaciones erróneas	16
4. Determinación de niveles de tolerancia	16
5. Enseñanzas extraídas del proceso de revisión de datos	17
6. Garantía de calidad	17
7. Costos de la revisión	18
8. Atribución de valores	18
9. Archivo	19
II. Práctica de la revisión de datos	21
A. Consideraciones relativas a la codificación	23
B. Corrección manual frente a corrección automatizada	27
C. Directrices para la corrección de datos	29
D. Comprobaciones de validez y coherencia	32
1. Método secuencial de revisión de datos	33
2. Método de revisión de datos con variables múltiples	34
E. Métodos de corrección y atribución de datos	37

	Página
1. La técnica de atribución estática.	37
2. La técnica de atribución dinámica.	38
3. Consideraciones relativas a la atribución dinámica	42
a) Consideraciones geográficas.	42
b) Utilización de elementos relacionados	42
c) Efectos en las matrices del orden de las variables	42
d) Complejidad de las matrices de atribución.	42
e) Elaboración de matrices de atribución	43
f) Matrices de atribución estándar.	44
g) Casos en que no se utiliza la atribución dinámica.	47
h) Tamaño apropiado de las matrices de atribución	49
i) Problemas que se plantean cuando la matriz de atribución es demasiado grande	49
ii) Comprensión del funcionamiento de la matriz de atribución	49
iii) Problemas que se plantean cuando la matriz de atribución es demasiado pequeña	49
iv) Elementos que presentan dificultades en las matrices de atribución	49
4. Comprobación de las matrices de atribución.	50
a) Establecimiento de la matriz estática inicial.	50
b) Mensajes de error	50
c) Enumeración de errores adaptada a necesidades concretas.	51
d) ¿Cuántas veces debe aplicarse el programa de revisión?	53
5. Señales de alerta de atribución de valores.	54
F. Otros sistemas de revisión de datos	56
III. Revisiones estructurales	59
A. Revisión de datos geográficos	60
1. Ubicación del local de habitación (localidad)	60
2. Población urbana y rural	60
B. Comprobaciones de la cobertura del censo.	61
1. Empadronamiento <i>de facto</i> y <i>de jure</i>	61
2. Jerarquía de los hogares y de las unidades de habitación.	62
3. Fragmentos de cuestionarios	62
C. Estructura de los registros de vivienda	62
D. Correspondencia entre los registros de vivienda y de población.	63
1. Viviendas desocupadas y ocupadas.	63
a) Decidir que la unidad de habitación constará como desocupada	63
b) Volver a visitar la unidad de habitación en varias ocasiones para completar los cuestionarios	63
c) Utilizar otra unidad de habitación para sustituir a las personas que faltan	64
2. Duplicación de hogares y de unidades de habitación	64
3. Omisión de hogares y de unidades de habitación	64
4. Correspondencia entre el número y la suma de los ocupantes	65
a) Cuando el número de ocupantes es mayor que la suma de los ocupantes	65
b) Comprobación del número de personas por sexo	65
c) Numeración secuencial	66

	<i>Página</i>
5. Correspondencia entre los ocupantes y el tipo de edificio/hogar	66
E. Duplicación de registros	66
F. Poblaciones especiales.	67
1. Personas que viven en locales colectivos.	67
a) Cuando los locales colectivos constituyen un tipo de registro diferente	67
b) Cuando una variable permite distinguir los registros colectivos de otros registros	67
c) Cuando se ha omitido el código correspondiente a “tipo de colectivo”	68
d) Cuando aparece el código correspondiente a un colectivo pero todas las personas están emparentadas	68
e) Distinción entre varios tipos de colectivos	68
2. Grupos de difícil empadronamiento	68
a) Migrantes estacionales.	68
b) Personas sin hogar	69
c) Nómadas y personas que viven en zonas de difícil acceso	69
d) Residentes civiles temporalmente ausentes del país	69
e) Extranjeros civiles que no cruzan la frontera a diario y se hallan en el país temporalmente, con inclusión de las personas no documentadas o los transeúntes que se encuentran en barcos fondeados en puertos en el momento del censo	69
f) Refugiados	70
g) Personal militar, naval y diplomático y sus familiares que están fuera del país, y personal militar, naval y diplomático extranjero y sus familiares que se encuentran en el país	70
h) Civiles extranjeros que cruzan la frontera a diario para trabajar en el país	70
i) Civiles residentes que cruzan la frontera a diario para trabajar en otro país.	70
j) Marineros mercantes y pescadores residentes en el país que se encuentran en el mar en el momento del censo (incluso aquellos que no tienen otro lugar de residencia que sus camarotes)	70
G. Determinación del jefe del hogar y del cónyuge	71
1. Revisión de la variable de jefe del hogar	71
a) El orden de las relaciones.	71
b) Cuando el jefe del hogar no es la primera persona	71
i) Asignación de un indicador al registro del jefe del hogar.	72
ii) Determinación de la primera persona como jefe del hogar.	73
iii) Reasignación de los códigos de relación de modo que la primera persona sea el jefe del hogar.	73
c) Hay más de un jefe del hogar	73
d) No se ha consignado un jefe del hogar	74

	<i>Página</i>
2. Revisión de los datos del cónyuge	74
a) Cuando se encuentra exactamente un cónyuge en sociedades monógamas.....	74
b) Cuando se encuentra más de un cónyuge en sociedades monógamas.....	74
c) Los cónyuges en sociedades polígamas.....	74
d) Otras características de los jefes de hogar y los cónyuges ..	75
H. Edad y fecha de nacimiento	75
1. Cuando consta la fecha de nacimiento pero no la edad	75
2. Cuando hay discrepancia entre la edad y la fecha de nacimiento	75
I. Recuento de asientos inválidos	75
IV. Revisión de los elementos de población	77
A. Características demográficas.....	78
1. Relación	79
a) Revisión de la variable de relación	79
b) Cuando el jefe del hogar debe constar en primer lugar	79
c) Cuando las relaciones se codifican en sentido inverso	79
d) Cuando aparecen cónyuges polígamos	80
e) Cuando aparecen varios progenitores	80
f) Cuando en los censos se compilan las relaciones con especificación de sexo.....	80
g) Cuando hay una discrepancia entre la relación y el estado civil.....	80
2. Sexo	81
a) Cuando el código de sexo es válido pero se ha consignado el mismo valor en la variable de sexo para el jefe del hogar y el cónyuge	81
b) Cuando hay información sobre fecundidad para una persona de sexo masculino o cuando no la hay para una mujer adulta	81
c) Cuando el código de sexo es inválido y hay un cónyuge ...	82
d) Cuando el código de sexo del cónyuge es inválido	82
e) Cuando el código de sexo es inválido y existe información que denota sexo femenino	82
f) Cuando el código de sexo es inválido y la persona es marido del cónyuge	82
g) Cuando el código de sexo es inválido y no hay información suficiente para determinar el sexo	82
h) Nota sobre las proporciones de los valores atribuidos a la variable de sexo.....	83
3. Fecha de nacimiento y edad	83
a) Edad y fecha de nacimiento.....	83
b) Relación entre la fecha de nacimiento y la edad.....	84
c) Cuando la edad calculada queda por encima del límite superior.....	84
d) Revisión del valor de la edad.....	85
e) Revisión del valor de la edad cuando constan el jefe del hogar y el cónyuge	85

	<i>Página</i>
f) Revisión del valor de la edad del jefe del hogar cuando no consta el cónyuge pero consta un hijo	85
g) Revisión del valor de la edad del jefe del hogar cuando consta un progenitor	86
h) Revisión del valor de la edad del jefe del hogar cuando consta un nieto	86
i) Revisión del valor de la edad del jefe del hogar cuando no se dispone de ninguna otra edad	86
j) Revisión del valor de la edad del cónyuge cuando ya se ha determinado la edad del jefe del hogar	87
k) Revisión del valor de la edad de otras parejas casadas del hogar cuando se conoce la edad de una de las personas	87
j) Revisión del valor de la edad de un hijo cuando ya se ha determinado la edad del jefe del hogar	87
m) Revisión del valor de la edad de un progenitor cuando ya se ha determinado la edad del jefe del hogar	88
n) Revisión del valor de la edad de un nieto cuando ya se ha determinado la edad del jefe del hogar	88
o) Revisión del valor de la edad de todas las demás personas	88
4. Estado civil	88
a) Revisión del valor del estado civil	89
b) Asignación del valor del estado civil cuando no se utiliza la atribución dinámica	89
c) Asignación del valor del estado civil cuando se utiliza la atribución dinámica	89
d) El cónyuge debe estar casado	89
e) El cónyuge en una pareja casada	89
f) Si hay un cónyuge, el jefe del hogar debe estar casado	90
g) Jefe del hogar sin cónyuge y sin hijos	90
h) La atribución se utilizará como último recurso	90
i) Relación entre la edad y el estado civil en el caso de personas jóvenes	90
5. Edad del primer matrimonio	90
a) La edad al contraer matrimonio debe quedar en blanco para las personas que nunca se han casado	91
b) Debe constar un valor para las personas que alguna vez han estado casadas	91
6. Fecundidad: hijos nacidos vivos e hijos vivos	91
a) Elementos de fecundidad que se compilan	92
b) Normas generales para la revisión del valor de la fecundidad	92
c) Relación entre los hijos nacidos vivos y los hijos vivos	93
d) Procedimiento de revisión cuando sólo se han declarado valores para los hijos nacidos vivos	93
e) Procedimiento de revisión cuando se han comunicado valores para los hijos nacidos vivos y los hijos vivos	93
f) Procedimiento de revisión cuando se han declarado valores para los hijos nacidos vivos, los hijos vivos y los hijos fallecidos	95

	<i>Página</i>
i) Cuando se han declarado los tres elementos	96
ii) Cuando se han declarado dos elementos	96
iii) Cuando se ha declarado un elemento	96
iv) Cuando no se ha declarado ningún elemento	96
g) Procedimiento de revisión cuando se han declarado valores para los hijos nacidos vivos, los hijos que viven en el hogar, los hijos que viven fuera del hogar y los hijos fallecidos.	97
i) Cuando se han declarado los cuatro elementos	97
ii) Cuando se han declarado tres de los cuatro elementos	97
iii) Cuando se han declarado dos de los cuatro elementos	98
iv) Cuando sólo se ha declarado un elemento	99
v) Cuando no se ha declarado ningún elemento	100
h) Caso especial en que hay cinco o más elementos.	100
i) Importancia de que un único donante sea la fuente de todos los elementos de fecundidad	100
j) Relación de los hijos propios con los hijos que se encuentran en el hogar y los hijos supervivientes	101
7. Fecundidad: fecha de nacimiento del último hijo nacido vivo y nacimientos en los 12 meses anteriores al censo.	101
8. Fecundidad: edad de la madre en el momento del nacimiento del primer hijo nacido vivo	102
9. Mortalidad	103
a) Edad y sexo de la persona fallecida	104
b) Causa del fallecimiento	104
c) Mortalidad materna	104
d) Mortalidad infantil.	104
10. Orfandad de madre o de padre y número de línea de la madre	105
B. Características de migración	106
1. Lugar de nacimiento.	106
a) Relación entre los asientos correspondientes al país de nacimiento y a los años que se ha vivido en el distrito	106
b) Asignación de un valor “desconocido” a los asientos inválidos del lugar de nacimiento.	107
c) Utilización de la atribución estática para el valor del lugar de nacimiento	107
d) Utilización de la atribución dinámica para el valor del lugar de nacimiento	107
e) Asignación de un valor del lugar de nacimiento cuando consta la madre de la persona	107
f) Asignación de un valor del lugar de nacimiento al hijo del jefe del hogar	108
g) Asignación de un valor del lugar de nacimiento al hijo, pero no al jefe del hogar.	108
h) Asignación de un valor del lugar de nacimiento a mujeres adultas que tienen marido	108
i) Asignación de un valor del lugar de nacimiento a mujeres adultas que no tienen marido	108
j) Asignación de un valor del lugar de nacimiento a personas de sexo masculino	109

	<i>Página</i>
2. Ciudadanía	109
a) Revisión del valor de la ciudadanía	109
b) Relación entre el grupo étnico/raza y la ciudadanía	110
c) Relación entre la naturalización y la ciudadanía	110
d) Relación entre la duración de la residencia y la ciudadanía	110
3. Duración de la residencia	110
a) Revisión del valor de la duración de la residencia	110
b) Residencia <i>de facto/de jure</i> y duración	110
c) Relación entre la edad y la duración de la residencia	111
d) Relación entre el lugar de nacimiento y la duración de la residencia	111
e) Personas que siempre han vivido en el lugar	111
f) Duración de la residencia de la persona a partir de la duración de la residencia de la madre	111
g) Duración de la residencia de la persona a partir de la duración de la residencia de un hijo	112
h) Duración de la residencia de una persona cuando no se disponga de otra información	112
4. Lugar de residencia anterior	113
a) Revisión del valor del lugar de residencia anterior	113
b) Lugar de residencia anterior cuando las fronteras han cambiado	113
c) Cuando la persona no se ha trasladado desde que nació	113
d) Utilización de otras personas de la unidad	113
e) No hay otra persona apropiada para determinar la residencia anterior	113
5. Lugar de residencia en una fecha especificada del pasado	114
6. Año de llegada	114
a) Relación de la edad con el año de llegada	114
b) Relación del lugar de nacimiento con el año de llegada	115
c) Personas que siempre han vivido en el lugar	115
d) Año de llegada de la persona desde el año de llegada de la madre	115
e) Año de llegada del hijo con respecto al año de llegada del jefe de hogar	116
f) Año de llegada de la persona cuando no se dispone de otra información	116
7. Relación entre duración de la residencia y año de llegada	116
8. Residencia habitual	117
C. Características sociales	118
1. Grado de alfabetización	118
2. Asistencia escolar	119
a) Revisión del valor de la asistencia escolar	119
b) Matrícula a tiempo completo o a tiempo parcial	119
c) Congruencia entre la asistencia escolar y la actividad económica	119
d) Asignación de valores a asientos inválidos o incongruentes de la variable de asistencia escolar	120

	<i>Página</i>
3. Nivel de instrucción (grado o nivel máximo terminado)	120
a) Revisión del valor del nivel de instrucción	120
b) Edad mínima para el nivel de instrucción	120
c) Relación entre la edad y el nivel de instrucción	120
4. Especialidad y titulación académica	121
5. Religión.	121
Revisión del valor de la religión	121
a) No consta la religión del jefe del hogar pero consta la religión de otra persona de la unidad de habitación	121
b) No consta la religión del jefe del hogar ni de ninguna otra persona de la unidad de habitación	122
c) No consta la religión de alguna persona distinta del jefe del hogar	122
6. Idioma.	122
a) Revisión del valor del idioma	122
b) Revisión del valor del idioma: el jefe del hogar	122
c) Revisión del valor del idioma: personas distintas del jefe del hogar	122
d) Revisión del valor del idioma: utilización del origen étnico o del lugar de nacimiento.	123
e) Revisión del valor del idioma: lengua materna	123
f) Revisión del valor del idioma: capacidad de hablar un idioma determinado	123
7. Grupo étnico y pueblos indígenas	123
a) Revisión del valor del grupo étnico	124
b) Revisión del valor del grupo étnico: el jefe del hogar	124
c) Revisión del valor del grupo étnico: personas distintas del jefe del hogar.	124
d) Revisión del valor del grupo étnico: utilización del idioma y del lugar de nacimiento.	124
8. Discapacidad	125
a) Preguntas del censo relativas a la discapacidad	125
b) Revisión del valor de la discapacidad.	125
c) Discapacidades múltiples.	125
d) Revisión del valor de las causas de la discapacidad.	126
D. Características económicas	126
1. Población activa	126
a) Categorías relativas a la actividad económica	127
i) Población desempleada	127
ii) En busca de empleo	128
iii) No actualmente activa	128
iv) Por qué no busca trabajo	128
b) Revisión del valor de la actividad económica.	129
i) Personas empleadas	129
ii) Actividad económica de las personas desempleadas.	129
iii) Actividad económica de los estudiantes y de los jubilados.	129

	<i>Página</i>
iv) Cuando la actividad económica no tiene un asiento válido y se han declarado variables de empleo.	129
v) Cuando la actividad económica no tiene un asiento válido y se han declarado variables de desempleo. . . .	130
vi) Cuando la actividad económica no tiene un asiento válido y no se ha declarado ninguna de las variables económicas.	130
2. Tiempo trabajado	130
3. Ocupación	130
4. Rama de actividad económica	131
5. Categoría en el empleo.	131
6. Ingresos	132
7. Sector institucional del empleo.	133
8. Empleo en el sector informal	133
9. Lugar de trabajo	133
V. Revisión de los elementos de vivienda.	135
A. Temas básicos y adicionales.	137
1. Edificio: descripción del edificio (tema básico)	137
2. Ubicación del local de habitación (tema básico)	138
3. Ocupación (o desocupación) (tema básico)	138
4. Tipo de propiedad (tema básico)	139
5. Número de cuartos (tema básico).	140
6. Número de dormitorios (tema adicional).	140
7. Superficie útil (tema adicional).	141
8. Sistema de suministro de agua (tema básico).	141
9. Agua potable; fuente principal de (tema básico)	142
10. Tipo de retrete (tema básico).	143
11. Instalaciones de alcantarillado (tema básico).	143
12. Instalaciones de baño (tema básico).	144
13. Cocina, disponibilidad de (tema básico).	144
14. Combustible utilizado para cocinar (tema básico)	145
15. Tipo de alumbrado y/o electricidad (tema básico).	146
16. Eliminación de residuos sólidos, tipo principal de (tema básico).	146
17. Tipo de calefacción, y energía utilizada (tema adicional).	147
18. Agua caliente, disponibilidad de (tema adicional)	147
19. Gas por tubería, disponibilidad de (tema adicional).	147
20. Uso de la unidad de habitación (tema adicional).	148
21. Ocupación por uno o más hogares (tema básico)	148
22. Número de ocupantes (tema básico)	148
23. Tipo de edificio (tema básico).	149
24. Año o período de construcción (tema adicional)	149
25. Viviendas en el edificio, número de (tema adicional).	149
26. Materiales de construcción de las paredes exteriores (tema básico)	150
27. Material de construcción del piso y el tejado (tema adicional)	150
28. Ascensor, disponibilidad de (tema adicional)	151
29. Edificio de tipo agrícola (tema adicional).	151

	<i>Página</i>
30. Estado de conservación (tema adicional)	151
31. Características del jefe del hogar u otro miembro de referencia del hogar (tema básico).	152
32. Tenencia o título de ocupación (tema básico).	152
33. Costos de alquiler o de la vivienda ocupada por su propietario (tema adicional).	152
34. Amueblado o no amueblado (tema adicional)	153
35. Sistemas de tecnología de la información y las comunicaciones, disponibilidad de (tema básico)	153
36. Automóviles, número de (tema adicional)	154
37. Aparatos electrodomésticos, disponibilidad de (tema adicional)	154
38. Espacio exterior disponible para uso del hogar (tema adicional).	155
B. Unidades de habitación ocupadas y desocupadas	155
 Anexos	
I. Variables derivadas.	157
A. Variables derivadas para los registros de vivienda.	157
1. Ingresos del hogar.	157
2. Ingresos familiares	158
3. Núcleo familiar	158
4. Tipo de hogar.	159
5. Composición de los hogares	159
6. Composición de la familia.	160
7. Posición en el hogar y en la familia	160
8. Efectos del VIH/SIDA en la estructura de los hogares	161
9. Familiares.	161
10. Trabajadores de la familia	162
11. Instalación completa de agua corriente.	162
12. Cocina completa	162
13. Alquiler bruto	163
14. Índice de riqueza.	163
B. Variables derivadas para los registros de población.	163
1. Situación laboral	163
2. Hijos propios	164
3. Progenitores en el hogar	164
4. Año en curso en la escuela.	165
5. Meses transcurridos desde el último nacimiento	165
II. Relación entre el formato del cuestionario y la introducción de datos	167
III. Escaneado frente a introducción manual de los datos	171
A. Introducción de los datos	171
1. El escáner	171
2. Introducción directa de datos a través del teclado	172
3. Introducción interactiva de datos	173
B. Verificación.	174
1. Verificación dependiente.	175
2. Verificación independiente	176
C. Revisión cuando se trata de datos escaneados	176
D. Conclusión	177

	<i>Página</i>
IV. Modelos de diagramas de secuencia	179
V. Métodos de atribución de valores	185
VI. Programas informáticos de revisión	189
Glosario	193
Referencias bibliográficas	199

Cuadros

1. Muestra de población, por grupo de edad de 15 años y sexo, utilizando datos sin revisar y revisados.	10
2. Población y cambios en la población por grupos de edad con intervalos de 15 años, con inclusión de los valores desconocidos: 2000 y 2010	11
3. Población y cambio de población por grupo de edad con intervalos de 15 años, con exclusión de los valores desconocidos: 2000 y 2010	12

Gráficos

1. Ejemplos de códigos comunes para determinados elementos	26
2. Un hogar hipotético típico, con información sobre las variables de relación, sexo y fecundidad de los miembros	30
3. Ejemplo de hogar en el que el jefe del hogar y el cónyuge son del mismo sexo	30
4. Ejemplo de hogar con las edades de algunos de sus miembros	31
5. Ejemplo de hogar con posibles incongruencias en las edades declaradas. .	32
6. Ejemplo de normas para una revisión con variables múltiples de determinadas características de población.	35
7. Ejemplo de un jefe de hogar y un cónyuge del mismo sexo en un conjunto de datos sin depurar y su solución	35
8. Ejemplo de especificaciones de revisión para corregir la variable relativa al sexo, redactadas en pseudocódigo.	36
9. Ejemplo de análisis de una revisión con variables múltiples en el caso de una viuda muy joven con tres hijos	36
10. Modelo de hogar como ejemplo de valores utilizados para la atribución dinámica	39
11. Matriz estática inicial para la variable de edad, basada en el sexo y en las relaciones.	40
12. Ejemplo de matriz de atribución dinámica después de un cambio	41
13. Ejemplo de una matriz de atribución dinámica después de múltiples cambios	41
14. Ejemplo de un jefe del hogar y de su progenitor a los que no se ha asignado un valor de idioma	44
15. Valores iniciales de una matriz de atribución dinámica para la variable de idioma	45
16. Ejemplo de miembros de un hogar a los que no se han asignado valores en la variable de idioma.	45
17. Ejemplo de jefe del hogar y un hijo, en el que no consta la edad del hijo ..	47
18. Ejemplo de un jefe de hogar y un hijo en el que se desconocen la edad y el nivel de escolarización del hijo	48

	<i>Página</i>
19. Ejemplo de un conjunto de valores para una matriz de atribución estática y modelo de código de atribución	50
20. Ejemplo de informe resumido sobre el número de atribuciones por error.	51
21. Modelo de informe sobre los errores de un cuestionario	51
22. Ejemplo de enumeración suplementaria de error por cuestionario, con múltiples variables	52
23. Ejemplo de registros de población con señales de los valores atribuidos . .	55
24. Ejemplo de alerta en el caso de una mujer joven con el valor de fecundidad en blanco y una señal de alerta añadida.	55
25. Ejemplo de hogar cuyo jefe ha sido empadronado como primera persona	71
26. Ejemplo de hogar cuyo jefe ha sido empadronado como quinta persona . .	72
27. Ejemplo de un hogar con información sobre la fecundidad	94
28. Valores iniciales para determinar los hijos vivos cuando los asientos de edad de la madre e hijos nacidos vivos son válidos	95
29. Ejemplos de matrices de atribución que pueden elaborarse para pares de elementos de información conocidos.	99
 <i>Recuadros</i>	
1. Resultados que deberían obtenerse con la revisión de los datos censales . .	9
2. Directrices principales para la corrección de datos.	30
3. Directrices para las revisiones estructurales	59
4. Revisión y atribución de un valor para la edad	84
 <i>Anexos-Gráficos</i>	
A.II.1. Modelo de impreso de cuestionario con páginas personales	167
A.II.2. Ejemplo de secuencia en un cuestionario con páginas personales.	167
A.II.3. Modelo de página sobre un hogar en un cuestionario con todas las personas en la misma página.	168
A.II.4. Ejemplo de secuencia en un cuestionario con páginas sobre un hogar y varias personas por página.	169
A.II.5. Ejemplo de una página sobre un hogar con varias personas sin problemas de introducción de datos	170
A.II.6. Ejemplo de una página sobre un hogar con varias personas con posibles problemas de introducción de datos	170
A.IV.1. Modelo de diagrama de secuencia para determinar el jefe del hogar	180
A.IV.2. Modelo de diagrama de secuencia para determinar la existencia de un cónyuge en el hogar	181
A.IV.3. Modelo de diagrama de secuencia para revisar la variable de sexo del jefe del hogar y del cónyuge.	182

Siglas

HF	Hijos fallecidos
HNV	Hijos nacidos vivos
HV	Hijos vivos
HVFH	Hijos que viven fuera del hogar
HVH	Hijos que viven en el hogar
ICIDH	Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías
MAVP	Metodología de atribución basada en el vecino más próximo
OIT	Organización Internacional del Trabajo
SAS	Statistical Analysis System
SCN	Sistema de Cuentas Nacionales
SIG	Sistemas de información geográfica
SPSS	Statistical Package for the Social Sciences
UNESCO	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura

Introducción

A. Propósito del Manual

1. Un censo o una encuesta bien diseñados¹ cuyo producto final presente un mínimo de errores son un recurso de valor inestimable para una nación. Para que los resultados de un censo o de una encuesta sean correctos, los datos deben contener el menor número posible de errores y discrepancias, especialmente después de la etapa de procesamiento de los datos.

*El procedimiento que se utiliza para descubrir errores en los registros de datos y entre ellos, tanto en la etapa de recopilación y captación de datos como más tarde, así como para corregir elementos concretos, es la **revisión** de los datos de los censos de población y vivienda.*

2. Los datos de los censos o de las encuestas nunca son perfectos. Ya hace tiempo que en los países se reconoce que los datos obtenidos en los censos y en las encuestas presentan problemas, y por ello se han adoptado diversos métodos para responder a las insuficiencias de datos y a las respuestas discordantes. Sin embargo, debido a los largos intervalos que median entre los censos, a menudo no están bien documentados los procedimientos que se utilizaron en la revisión de los datos. Como consecuencia de ello, en cada nuevo censo o encuesta los países tienen que reinventar el proceso utilizado en anteriores actividades de recopilación de datos.

3. El presente *Manual de revisión de datos de los censos de población y vivienda* tiene por objeto subsanar el déficit de conocimientos sobre la metodología de revisión de datos de los censos y las encuestas y proporcionar a los funcionarios interesados información sobre la aplicación de distintos métodos para la revisión de los datos censales. Tiene también el propósito de alentar a los países a conservar un historial de sus experiencias de revisión de datos, mejorar la comunicación entre los especialistas temáticos y los especialistas en procesamiento de datos y documentar las actividades llevadas a cabo en el censo o la encuesta en curso a fin de evitar duplicar esas actividades en el futuro.

4. El *Manual* servirá de referencia tanto para los especialistas² temáticos como para los especialistas en procesamiento de datos cuando ambos grupos colaboren en la formulación de especificaciones y en la elaboración de programas para la revisión de datos de los censos y encuestas. Los métodos se presentan en forma de “recetas”, lo cual permitirá a los países adoptar los tipos de revisión de datos más apropiados a su situación estadística en cada momento. La presente publicación tiene también por objetivo promover una mejor comunicación entre esos especialistas en su labor de formulación y aplicación de un programa de revisión de datos.

5. En los párrafos que siguen se describen el proceso censal, los tipos de errores que se producen en un censo y los elementos básicos de la revisión de los datos cen-

¹ Un censo es un recuento completo. Una encuesta normalmente abarca una proporción menor de la población total. Las revisiones aquí descritas deberían ser válidas para ambos proyectos.

² En el presente *Manual* se entiende por especialistas temáticos a las personas que desempeñan su actividad en los ámbitos de la población, la vivienda y esferas conexas.

sales. En los capítulos siguientes se exponen los procedimientos y técnicas de revisión de datos en las distintas etapas de procesamiento. Aunque este manual se centra en la revisión de datos de los censos de población y vivienda, muchos de los conceptos y técnicas son también aplicables a las encuestas.

B. El proceso censal

6. Un censo de población y/o vivienda es el conjunto de las operaciones consistentes en recoger, recopilar, evaluar, analizar y divulgar datos demográficos y/o de vivienda, económicos y sociales relativos a todos los habitantes y sus locales de habitación (Naciones Unidas, 2008). Tradicionalmente, los censos se llevan a cabo en momentos determinados en todo un país o en una parte bien delimitada de él. Recientemente, algunos países han comenzado a realizar encuestas permanentes relativas a la totalidad del país, utilizando un formulario “largo”, para disponer de una cobertura cronológica completa. En uno y otro caso, el censo brinda una imagen de la población y de la vivienda en un momento dado.

7. El objetivo fundamental de un censo es proporcionar información sobre el tamaño, la distribución y las características de la población de un país. Los datos censales se utilizan para la formulación de políticas, la planificación y la administración, así como para la gestión y la evaluación de programas en los campos de la educación, la fuerza de trabajo, la planificación de la familia, la vivienda, la salud, el transporte y el desarrollo rural. Una aplicación administrativa básica es la demarcación de entidades políticas territoriales y la asignación de representación ante los órganos de gobierno. El censo constituye también un recurso inestimable para los investigadores, ya que proporciona datos para el análisis científico de la composición y la distribución de la población y para formular modelos estadísticos de pronóstico de su crecimiento futuro. El censo ofrece al comercio y a la industria los datos básicos que necesitan para evaluar la demanda de vivienda, escuelas, mobiliario, alimentos, vestido, instalaciones recreativas, suministros médicos y otros bienes y servicios.

8. Todos los censos y encuestas tienen ciertas características principales comunes, a saber: *a)* la labor preparatoria; *b)* el empadronamiento o recopilación; *c)* el procesamiento de datos, que incluye la introducción (teclado o escaneado), la revisión y la tabulación de los datos; *d)* la construcción de bases de datos y la difusión de los resultados; *e)* la evaluación de los resultados; y *f)* el análisis de los resultados.

9. La labor preparatoria comprende muchos elementos, como la determinación de la base jurídica del censo, la elaboración de un presupuesto, el establecimiento del calendario, la organización administrativa, la cartografía, la creación de un listado de unidades de vivienda, la elaboración del programa de tabulación, la preparación del cuestionario y la formulación de planes y la capacitación de personal para tareas de empadronamiento, de ensayos previos, de procesamiento de datos y de difusión.

10. El proceso de empadronamiento dependerá del método que se haya elegido, del momento de inicio y de la duración del período de empadronamiento, del nivel de supervisión y de la posibilidad de utilizar una muestra, así como de la forma de aplicarla. Una vez que se han reunido los datos, estos deben ser codificados, captados, revisados y tabulados. El procesamiento de datos producirá bases de datos de micronivel y de macronivel. Las oficinas nacionales de estadística o del censo utilizan esas bases de datos en tabulaciones, análisis de series cronológicas, operaciones de generación de gráficos y mapas y sistemas de información geográfica (SIG) para fines de cartografía temática y otras técnicas de difusión. Para evaluar el contenido y la cobertura de los resultados se aplican diversos métodos, como el análisis demográfico y las encuestas

de verificación. Por último, los resultados se analizan de distintas maneras, entre las que se incluyen resúmenes descriptivos de los resultados, análisis de los resultados del censo orientados a la formulación de políticas, y estudios analíticos detallados de uno o más aspectos de la situación demográfica y social del país.

C. Errores en el proceso censal

11. Los datos obtenidos en un censo pueden presentar errores de distinta naturaleza, que de manera general pueden clasificarse como errores de cobertura y errores de contenido.

1. Errores de cobertura

12. Los errores de cobertura se deben a omisiones o duplicaciones de personas o unidades de habitación durante el empadronamiento. Estos errores pueden deberse, entre otras razones, a que los mapas o las listas de las zonas de empadronamiento son incompletos o inexactos, a que los empadronadores no hayan visitado todas las unidades de las zonas que les correspondían, a que algunas personas hayan sido contadas dos veces, a que se hayan omitido personas que no querían ser empadronadas, a que se hayan tratado erróneamente determinadas categorías de personas, como los visitantes o los extranjeros no residentes, o a que se hayan perdido o destruido los registros del censo después del empadronamiento. En la medida de lo posible, los errores de cobertura deberían corregirse sobre el terreno. En el proceso de revisión de datos llevado a cabo en la oficina se eliminarán los registros duplicados. Sin embargo, hay que procurar determinar si se trata de duplicaciones de personas o de hogares. La información sobre los mellizos, por ejemplo, puede ser idéntica, si se exceptúa el número de secuencia. Por tanto, las normas de revisión de datos que se aplican durante este proceso determinarán cuándo se aceptará y cuándo se rechazará la información que parezca estar duplicada y cuándo se efectuarán cambios mediante la atribución de nuevos valores.

13. La revisión estructural, que se examina en el capítulo III, permite comprobar en cada hogar si el número de registros personales es correcto, si la secuencia es adecuada y si hay duplicaciones de personas.

2. Errores de contenido

14. Los errores de contenido se deben a que se han comunicado o registrado incorrectamente las características de las personas, los hogares y las unidades de habitación. Los errores de contenido pueden obedecer a que las preguntas están mal diseñadas o a que su secuencia es incorrecta, a una mala comunicación entre el declarante y el empadronador, así como a equivocaciones en la codificación y la introducción de datos, a errores en la revisión manual o informatizada de los datos y a la tabulación errónea de los resultados. Habrá que establecer un método apropiado de seguimiento de la revisión de los datos, denominado también rastro de verificación de cambios, que deberá irse conservando en cada etapa del proceso a fin de asegurar que no se pierdan datos. En los apartados que figuran a continuación se explican los distintos errores mencionados.

a) Errores en el diseño del cuestionario

15. Las preguntas o instrucciones mal redactadas son una fuente de errores de contenido. Hay que prestar la máxima atención al tipo de cuestionario, a su formato y a la redacción exacta y la secuencia de las preguntas, puesto que los defectos de un

cuestionario mal diseñado no podrán corregirse durante el empadronamiento ni más tarde. Hay que efectuar ensayos previos para reducir al mínimo los errores que puedan surgir a causa de un mal diseño de los cuestionarios. Si, por ejemplo, las pautas de omisión no están claramente indicadas o no se han introducido en el lugar apropiado, el empadronador puede saltarse erróneamente ciertas secciones del cuestionario y no reunir toda la información pertinente.

b) Errores de los empadronadores

16. A menos que el censo se administre mediante un cuestionario que llena el propio declarante, se producirá una interacción entre los empadronadores y los declarantes. El empadronador puede equivocarse al hacer las preguntas, porque abrevie o modifique la redacción de estas o porque no explique perfectamente su significado al declarante. El empadronador también puede agregar errores al registrar las respuestas. Contar con empadronadores experimentados o bien capacitados son factores decisivos para la calidad de los datos recopilados. Los empadronadores deben recibir un entrenamiento apropiado en todos los aspectos de los procedimientos censales. Hay que lograr que comprendan por qué es importante su función en el proceso censal y qué lugar ocupa el empadronamiento entre las demás etapas del censo. Además, dado que los empadronadores proceden de medios muy diversos y tienen niveles de educación muy diferentes, la capacitación debe estar diseñada de manera que permita a todos los empadronadores aprender cómo deben hacer las preguntas para obtener una respuesta apropiada.

c) Errores de los declarantes

17. Pueden introducirse errores en los datos cuando los declarantes no entienden bien algunos elementos. También pueden introducirse errores como consecuencia de respuestas deliberadamente equivocadas o de respuestas de terceros (cuando contesta al cuestionario alguien distinto de la persona a la que se refiere la información). La calidad de las respuestas puede mejorarse dando publicidad al censo e impartiendo capacitación a los empadronadores para que expliquen de la mejor manera el propósito del censo y los motivos de las diversas preguntas. Algunos países utilizan cuestionarios que llena el propio declarante, de forma que no hay interacción entre este y el empadronador. En el caso de los impresos que llena el propio declarante, pueden producirse errores cuando no entiende bien las preguntas o las instrucciones.

18. Es mejor corregir los errores de los declarantes y de los empadronadores en la etapa del empadronamiento, cuando los impresos, los declarantes y los empadronadores todavía están disponibles. Los supervisores deben tener la capacidad de instruir a los empadronadores. Los supervisores también deben tener la posibilidad de comprobar periódicamente durante el empadronamiento los datos reunidos por los empadronadores, para evitar que introduzcan en los datos errores sistemáticos. Lo ideal sería que los supervisores resolvieran los errores de los empadronadores y de los declarantes sobre el terreno, antes de que los cuestionarios se envíen a las oficinas regionales o centrales.

d) Errores de codificación

19. Pueden producirse errores en la codificación cuando el operador asigna un código equivocado a la información. Durante la introducción de datos pueden generarse errores por fallos en el teclado. En general, la falta de supervisión y comprobación en esta etapa demorará la publicación de los datos, ya que la detección y corrección de los errores resulta más difícil después. A menudo se hacen revisiones manuales antes de la operación de codificación o mientras se lleva a cabo.

e) Errores en la introducción de datos

20. En el programa informático de introducción de datos pueden incluirse ciertas comprobaciones basadas en el margen de variación o en la coherencia, para impedir que se introduzcan datos inválidos. Un sistema inteligente de introducción de datos comprueba que el valor de cada campo o elemento esté dentro de un margen de variación permisible para los valores correspondientes a ese elemento. Este sistema aumenta la probabilidad de que el operador introduzca datos razonables y alivia parte de la carga de revisión de datos en etapas posteriores del proceso de preparación de los datos. Sin embargo, esas comprobaciones pueden frenar la introducción de datos. Por lo tanto, al determinar el número de comprobaciones de coherencia que se realizarán durante la introducción de datos habrá que tener muy en cuenta la necesidad de mantener una velocidad de introducción razonable. Hay que establecer un equilibrio de antemano, de forma que los encargados de la introducción de datos no dediquen demasiado tiempo a actividades de comprobación. La comprobación de la introducción de datos mejora inevitablemente la calidad de los mismos. Los impresos cuyos datos ya se han introducido pueden comprobarse volviendo a introducir la misma información, a menudo aplicando técnicas de muestreo. La relación entre el formato del cuestionario y la introducción de los datos se examina en el anexo II. En el anexo III se compara el escaneado con la introducción manual de los datos.

f) Errores en la revisión por ordenador

21. Una de las etapas cruciales del procesamiento de los datos censales es la de depuración o revisión de datos. En este proceso se modifican o se corrigen los datos inválidos o discordantes atribuyendo a los elementos datos plausibles cuando no ha habido respuestas o cuando la información no es coherente. Paradójicamente, cualquiera de esas operaciones de revisión puede dar lugar a nuevos errores.

g) Errores de tabulación

22. En la fase de tabulación pueden producirse errores debidos a equivocaciones en el procesamiento de datos o a la utilización de información “desconocida” (es decir, que no se ha suministrado). Los errores que se producen en esta etapa son difíciles de corregir sin introducir nuevos errores. La falta de comprobaciones entre tabulaciones y los errores de imprenta provocan errores en la etapa de publicación. En vez de intentar efectuar correcciones en los propios cuadros, es importante que el sistema de procesamiento prevea efectuar una revisión adicional cuando aparezcan incongruencias en los cuadros. Si los errores se van transmitiendo de una etapa a otra del proceso, se pondrán de manifiesto cuando se publiquen los resultados, y estos tendrán un valor cuestionable. Si las “correcciones” se introducen en la fase de la tabulación —por ejemplo, si se encuentran valores desconocidos y se colocan en los “totales” pero no en la distribución—, los cuadros no podrán ser reproducidos por otros analistas y tendrán menos valor global. Es más prudente entender el procesamiento del censo como un sistema de retroalimentación, de manera que los cambios en el conjunto de los datos se introduzcan durante la revisión y no durante la tabulación. Antes de publicar los cuadros es esencial llevar a cabo una comprobación exhaustiva para asegurarse de que se han preparado todas las tabulaciones previstas para todas las unidades geográficas abarcadas. Si bien las comprobaciones basadas en el margen de variación o en la coherencia realizadas en la etapa de revisión pueden reducir la mayor parte de los errores, es esencial hacer una comprobación general después de la tabulación, que a veces recibe el nombre de “revisión de macronivel”. Los cuadros deben ser examinados por personas capacitadas y experimentadas en comprobar que las cifras que constan en las

distintas casillas concuerdan con la situación que se conoce a nivel local. En algunos casos una rápida consulta de los impresos del censo puede poner de relieve errores de codificación. También puede ser útil calcular determinadas proporciones y tasas de crecimiento y compararlas con las cifras de censos anteriores u otras cifras publicadas por estudios basados en un muestreo. Sin embargo, la comparación con otras cifras basadas en encuestas o en datos administrativos sólo debe intentarse si los conceptos utilizados son comparables. Si se encuentran errores en las tabulaciones finales, las correcciones deberán hacerse primero en el conjunto de datos de micronivel, en parte para permitir que otros encargados del procesamiento de datos en la oficina nacional de estadística o del censo construyan cuadros comparables. Además, puesto que las oficinas nacionales de estadística o del censo a veces ponen a disposición de los investigadores y de otros usuarios de los sectores público y privado partes de los archivos de datos de micronivel, es importante que los cuadros puedan ser reproducidos por esos otros usuarios.

23. Como se ha indicado, el proceso censal entraña varias operaciones secuenciales e interrelacionadas, y en cada una de ellas pueden producirse errores. Es importante recordar que las revisiones efectuadas por ordenador son parte de un sistema de retroalimentación y que la revisión por ordenador no sólo produce material para las tabulaciones sino que también crea nuevas relaciones con las etapas anteriores de reunión y procesamiento de datos sobre el terreno. Para una oficina nacional de estadística o del censo, la mejor forma de evitar los problemas causados por la revisión por ordenador es aumentar al máximo el grado de revisión sobre el terreno. La oficina nacional de estadística o del censo también deberá asegurarse de que la codificación y la introducción de datos sean correctas, y deberá haber una corriente constante de información entre todas las operaciones de introducción de datos, de revisión y de tabulación.

D. Estructura del *Manual*

24. En el capítulo I se examina la función de la revisión de datos en los censos y encuestas. Los otros capítulos tratan de temas concretos. En el capítulo II se trata de la práctica de la revisión, incluidos los métodos de atribución de valores. En el capítulo III se presentan las revisiones estructurales, las revisiones que examinan al mismo tiempo elementos de vivienda y de población, y determinados procedimientos útiles para el resto de las revisiones, como el de determinar si hay un jefe del hogar y no más de uno. En el capítulo IV se examina la revisión de los elementos de población, y en el capítulo V se trata de la revisión de los elementos de vivienda. Por último, en una serie de anexos se examinan cuestiones concretas relativas a la revisión y la atribución de valores en los censos de población y vivienda.

Capítulo I

La revisión de datos en los censos y en las encuestas

A. Reseña histórica de la revisión de datos

25. Antes de la aparición de los ordenadores, en la mayoría de las operaciones censales se contrataba a un gran número de empleados semicualificados para revisar los impresos. No obstante, aun cuando el número de elementos fuera limitado, la complejidad de las relaciones entre ellos hacía que con estas simples comprobaciones ni siquiera se pudiera comenzar a resolver todas las incongruencias que podían presentar los datos. Cada empleado interpretaba las normas casi a su manera, e incluso un mismo empleado podía actuar de manera poco sistemática.

26. La labor de revisión de los datos censales cambió notablemente con la introducción de los ordenadores. Estos permitían sacar a la luz muchas más discrepancias que la revisión manual. Las especificaciones para la revisión se fueron perfeccionando y haciendo cada vez más complejas. Fue posible automatizar la atribución de valores y se crearon normas para el proceso correspondiente (Nordbotten, 1963; Naus, 1975). Al mismo tiempo, el proceso permitía aumentar el contacto con los declarantes, o por lo menos con los cuestionarios ya llenados por ellos. Muchos equipos de revisión empezaron a pensar que cuanto más se revisara, mejores serían los resultados, y que estos serían más exactos cuanto más compleja fuera la revisión. Los programas producían miles de mensajes de error, lo cual exigía examinar manualmente los impresos originales o, en el caso de algunos estudios, volver a entrevistar a los declarantes³.

27. Con los ordenadores ha resultado cada vez más fácil hacer cambios en el conjunto de datos. A veces esos cambios han servido para corregir registros o elementos. Muchos registros pasaban por el ordenador varias veces y los errores e incongruencias eran revisados por personas distintas cada vez (Boucher, 1991; Granquist, 1997).

28. De todo este proceso salieron varios programas informáticos generales para la revisión de datos censales, algunos de los cuales todavía se utilizan. Inicialmente, los programas se elaboraron para ordenadores centrales; más adelante algunos se modificaron para que pudieran utilizarse en ordenadores personales. En este período, Fellegi y Holt (1976) elaboraron un nuevo método general para la revisión y atribución de datos censales, que no se aplicó inmediatamente pero que se está adoptando cada vez más a medida que se va haciendo más compleja la labor de revisión en las oficinas nacionales de estadística y del censo.

29. Un importante progreso en la actividad de revisión de los datos censales se produjo en el decenio de 1980, cuando las oficinas nacionales de estadística o del censo comenzaron a utilizar ordenadores personales para introducir, revisar y tabular

³ El anexo V contiene una exposición de los distintos tipos de métodos de atribución; en los anexos V y VI se trata de algunos programas de revisión por computadora.

sus datos. Por primera vez los encargados del procesamiento de datos podían efectuar revisiones en línea en la fase de introducción de los datos o muy poco después. Para las encuestas y para los censos de países pequeños, los funcionarios competentes podían elaborar programas que permitieran descubrir errores durante la fase de reunión de datos o cuando estos se estaban introduciendo directamente en la máquina. La revisión por computadora permitía mantener un contacto más directo y constante con los declarantes para resolver los problemas que se habían observado en el proceso de revisión (Pierzchala, 1995).

30. En los primeros años se registraron éxitos considerables en el proceso de efectuar comprobaciones, cada vez más complejas y exhaustivas, de los datos de los censos o de las encuestas. Los equipos de revisión creaban especificaciones de revisión cada vez más complicadas y los especialistas en procesamiento de datos dedicaban meses a elaborar diagramas de secuencia y gráficos de decisión. Los analistas raramente evaluaban los programas. Daba la impresión de que la revisión podría corregir todos los problemas que surgieran en las fases anteriores de recopilación, codificación e introducción de datos. No obstante, numerosos analistas observaron que en muchos casos toda esta labor adicional de revisión era perjudicial para los datos, o que por lo menos retrasaba los resultados y creaba sesgos en los resultados. En algunas ocasiones el programa examinaba tantas veces los datos, corrigiendo primero un elemento y después otro, que los resultados se alejaban mucho de los datos iniciales no revisados.

31. En muchos censos y encuestas de gran tamaño, una revisión tan amplia provocaba demoras considerables en el proceso del censo o de la encuesta. Los empleados dedicaban mucho tiempo a buscar datos en los impresos; los especialistas en procesamiento de datos seguían elaborando aplicaciones para examinar una cantidad de casos reducida. Granquist (1997) señala que muchos estudios han demostrado que, por lo que respecta a buena parte de esta labor adicional, “la mejora de la calidad es marginal, nula o incluso negativa; hay muchos tipos de errores sistemáticos graves que no pueden ser identificados en la revisión”.

32. A medida que las organizaciones nacionales de estadística vayan realizando nuevos censos y encuestas es posible, e incluso probable, que se lleve a cabo una amplia labor de revisión de los datos por ordenador. Por consiguiente, cada oficina nacional de estadística o del censo deberá tratar de determinar el grado de revisión por ordenador que resulta más apropiado para sus finalidades.

B. El equipo de revisión de datos

33. Cuando las oficinas nacionales de estadística se preparan para un censo deben considerar diversas formas posibles de mejorar la calidad de su labor. Una de ellas es la creación de un equipo de revisión de datos. El proceso de revisión debe estar a cargo de un equipo integrado por los directores del censo, por especialistas temáticos y por especialistas en procesamiento de datos. Este equipo debe establecerse tan pronto como comiencen los preparativos del censo, preferiblemente durante la fase de redacción del cuestionario. El equipo de revisión es importante desde un principio, y lo seguirá siendo durante todo el proceso de revisión de los datos. Si se dedica la debida atención a constituir el equipo y a formular y aplicar las normas de revisión y atribución de valores, la realización del censo será más rápida y eficiente.

34. La celebración de reuniones entre los funcionarios del censo y la comunidad de usuarios con respecto a las tabulaciones y otros productos basados en los datos puede ofrecer ideas sobre los tipos de revisión que será necesario llevar a cabo. Es frecuente que los usuarios soliciten un determinado cuadro o tipo de cuadros, que

puede que requieran una revisión adicional para eliminar posibles incongruencias. El equipo de revisión debería contemplar la realización de esos cuadros durante el período de revisión inicial, en vez de construirlos como cuadros especiales después del procesamiento de los datos censales. Elaborar las normas de revisión de datos y los programas informáticos correspondientes durante un ensayo previo o un ensayo general permite poner a prueba los propios programas y crea ciclos más rápidos para las distintas etapas del proceso de revisión y atribución de valores. A continuación, el equipo de revisión averigua los efectos de esos procesos y, de ser necesario, adopta medidas de corrección.

35. Los especialistas temáticos y los especialistas en procesamiento de datos deben colaborar en la elaboración de las normas de revisión y de atribución de valores. El equipo de revisión formula un plan de escrutinio de errores y de revisión en una etapa inicial de los preparativos del censo. El equipo de revisión del censo o de la encuesta redactará y pondrá por escrito normas de coherencia y de corrección.

36. Además de formular las normas de revisión y de atribución de valores, los especialistas temáticos y los especialistas en procesamiento de datos deben colaborar en todas las etapas del censo o de la encuesta, incluida la etapa de análisis. El riesgo de revisar demasiado es tan grande como el riesgo de no revisar bastante y de que el conjunto de datos contenga información no depurada o falaz. Por tanto, ambos grupos deben hacerse responsables del mantenimiento adecuado de sus bases de metadatos. El equipo de revisión también debe utilizar de manera eficiente las fuentes administrativas y los registros de encuestas de que se disponga para mejorar censos o encuestas ulteriores.

37. Cuando las oficinas nacionales de estadística o del censo utilizaban ordenadores centrales, la comunicación entre los especialistas temáticos y los especialistas en procesamiento de datos era limitada. Esta división persistió durante algún tiempo después de la aparición de los microordenadores, pero los programas informáticos son ahora más fáciles de utilizar y muchos especialistas temáticos pueden elaborar y ensayar sus propios planes de tabulación y sus normas de revisión. Si bien los especialistas temáticos no suelen encargarse de procesar los datos, sí suelen tener un conocimiento de las etapas que siguen para ello los especialistas en procesamiento de datos.

Recuadro 1

Resultados que deberían obtenerse con la revisión de los datos censales

Con la revisión de los datos censales deberían obtenerse los siguientes resultados:

- ☞ Suministrar a los usuarios datos censales de alta calidad.
- ☞ Determinar los tipos y las fuentes de los errores.
- ☞ Proporcionar resultados del censo corregidos.

C. Prácticas de revisión de datos: datos revisados y datos no revisados

38. Los países realizan actividades de revisión del censo para mejorar los datos y su presentación. En esta sección se pone de relieve un problema al que se enfrentan las oficinas nacionales de estadística o del censo cuando se publican datos censales no revisados. Se presenta un ejemplo de esos problemas utilizando un conjunto hipotético de datos.

39. La oficina nacional de estadística o del censo de un país ficticio se encuentra con el dilema de intentar ofrecer servicios a diversos usuarios. Es posible que algunos de ellos deseen que, para fines de análisis o investigación, se incluyan los valores desconocidos, mientras que otros quieran datos con un mínimo de “ruido” (posible error) para sus fines de planificación o formulación de políticas. Si la oficina nacional de estadística o del censo difunde un cuadro con datos no revisados, como el que figura en las casillas correspondientes a este caso en el cuadro 1, los analistas y los encargados de la formulación de políticas tendrán que recurrir a hipótesis para poder hacer uso de los datos. Con el cuadro 1 se quiere ilustrar esta cuestión utilizando un pequeño número de personas. Se ha supuesto que en el caso de 23 personas de ese país no consta el sexo⁴ y que en el de otras 15 no consta la edad. Esas omisiones pueden ser consecuencia de faltas de respuesta o de errores en la introducción de los datos. En dos de los casos no consta ni el sexo ni la edad.

⁴ En esta publicación se utilizan indistintamente los términos “sexo” y “género”.

Cuadro 1
Muestra de población, por grupo de edad de 15 años y sexo, utilizando datos sin revisar y revisados

Grupo de edad	Sexo						
	Datos no revisados			No declarado	Datos revisados		
	Total	Hombres	Mujeres		Total	Hombres	Mujeres
Total	4 147	2 033	2 091	23	4 147	2 045	2 102
Menos de 15 años	1 639	799	825	15	1 743	855	888
15 a 29 años	1 256	612	643	1	1 217	603	614
30 a 44 años	727	356	369	2	695	338	357
45 a 59 años	360	194	166	0	341	182	159
60 a 74 años	116	54	59	3	114	53	61
75 años o más	34	12	22	0	37	14	23
No declarado	15	6	7	2			

40. La mayoría de los usuarios desearán adoptar sus propias decisiones acerca de qué hacer con los valores desconocidos. Un enfoque lógico, aunque tal vez ingenuo, consistiría en distribuir los valores desconocidos en la misma proporción que los valores conocidos. Si la oficina nacional de estadística o del censo opta por atribuir un valor a las variables desconocidas, el equipo de revisión de datos puede decidir consignar 12 hombres y 11 mujeres, una distribución que se aproxima al 50%, aunque con un pequeño sesgo, porque en el censo se han empadronado más personas de sexo femenino. En este caso los resultados serán coherentes con los datos que se presentan en las casillas correspondientes a “Datos revisados” del cuadro 1.

41. Existen otras opciones para tratar los valores desconocidos. El equipo de revisión de datos puede decidir atribuir esos valores basándose solo en la distribución por sexos y haciendo caso omiso de otra información disponible, como la relación entre cónyuges, si una persona de sexo desconocido consta como madre de otra persona, o si una persona de sexo desconocido tiene un asiento positivo en la variable del número de hijos nacidos vivos. Otra estrategia de atribución sería tomar en cuenta una o más de esas otras variables.

42. Otra posibilidad que tiene la oficina nacional de estadística o del censo es atribuir valores basándose en la distribución por edades. En la población de muestra del cuadro 1 había 15 casos de los cuales no constaba la edad. Esos datos también podrían distribuirse en la misma proporción que los valores conocidos; esta es

también una estrategia de atribución lógica. Sin embargo, el equipo de revisión de datos probablemente obtendría mejores resultados si tuviera en cuenta otras variables y combinaciones, como las edades relativas del marido y de la esposa, del padre y el hijo o del abuelo y el nieto, o bien la presencia de niños en edad escolar, jubilados y personas que forman parte de la fuerza de trabajo.

43. En el cuadro 1 los datos revisados están más depurados porque se han suprimido los valores desconocidos (véanse las columnas que figuran bajo el epígrafe “Datos revisados”). En esta parte del cuadro no hay valores desconocidos, ya que el programa les asigna otras respuestas. Sin embargo, muchos demógrafos y otros especialistas temáticos han deseado tradicionalmente que los valores desconocidos figurasen en los cuadros, como sucede con los datos no revisados del cuadro 1, porque consideran que este procedimiento les permite utilizar las cifras en diversos tipos de evaluaciones para medir la eficacia de los procedimientos del censo o para ayudar a planificar futuros censos y encuestas. Pueden lograrse ambos objetivos —un cuadro con datos revisados para los usuarios sustantivos y un cuadro con datos no revisados para fines de evaluación— si se hacen dos tabulaciones, que incluyan y excluyan respectivamente los valores desconocidos.

44. Las oficinas de estadística deberían hacer todo lo posible por mantener los datos recopilados originales. Debería archivar un conjunto completo de los datos originales introducidos, tanto como parte del registro histórico, como para referencia, por si acaso el personal decide volver a revisar una parte del conjunto de datos desde el comienzo. No obstante, los valores originales de elementos cruciales, como la edad, el sexo y la fecundidad, deben mantenerse en algún lugar de cada registro para que los demógrafos y otros especialistas puedan analizar los resultados de las revisiones.

45. Otro problema que presenta la inclusión de los valores desconocidos en los cuadros publicados es que esos valores pueden afectar al análisis de las tendencias. Las nuevas tecnologías hacen que este análisis sea mucho más fácil que antes. Por ejemplo, en el cuadro 2 se presenta una distribución por edades de dos censos consecutivos. En este pequeño país el número de valores desconocidos se redujo de 217 (es decir, alrededor del 6,5% de las respuestas) en el año 2000 a apenas 15 (es decir, menos del 1% de las respuestas) en 2010.

46. En este caso la oficina nacional de estadística o del censo debe tratar de aclarar el efecto que tendrán esas proporciones distintas de valores desconocidos en ese censo concreto y determinar los cambios que se han producido entre los dos censos. Por ejemplo, el 6,5% de valores desconocidos en el censo de 2000 hace difícil com-

Cuadro 2

Población y cambios en la población por grupos de edad con intervalos de 15 años, con inclusión de los valores desconocidos: 2000 y 2010

Grupo de edad	Total		Cambio en el total	Cambio porcentual	Porcentaje	
	2010	2000			2010	2000
Total	4 147	3 319	828	24,9	100,0	100,0
Menos de 15 años	1 639	1 348	291	21,6	39,5	40,6
15 a 29 años	1 256	902	354	39,2	30,3	27,2
30 a 44 años	727	538	189	35,1	17,5	16,2
45 a 59 años	360	200	160	80,0	8,7	6,0
60 a 74 años	116	89	27	30,3	2,8	2,7
75 años o más	34	25	9	36,0	0,8	0,8
No declarado	15	217	-202	-93,1	0,4	6,5

para cómo han cambiado entre los dos censos las distribuciones porcentuales de los grupos de edad de 15 años. El porcentaje de personas de entre 15 y 29 años de edad parece haber subido sólo del 27% al 30% durante el decenio, pero la distribución de los valores desconocidos podría modificar el análisis.

47. En el cuadro 3, que es un cuadro revisado, los valores desconocidos se han distribuido proporcionalmente o aplicando algún otro método de atribución. En este caso es mucho más fácil observar en los dos censos los cambios, tanto absolutos como porcentuales, y la distribución de los grupos de edad. Por supuesto, para obtener resultados precisos y fidedignos, los equipos de revisión de datos deberán asegurarse de que haya coherencia entre los dos censos y/o encuestas y coherencia interna. La entrada correspondiente a “no declarado” se ha suprimido.

Cuadro 3

Población y cambio de población por grupo de edad con intervalos de 15 años, con exclusión de los valores desconocidos: 2000 y 2010

Grupo de edad	Total		Cambio en el total	Cambio porcentual	Porcentaje	
	2010	2000			2010	2000
Total	4 147	3 319	828	24,9	100,0	100,0
Menos de 15 años	1 743	1 408	335	23,8	42,0	42,4
15 a 29 años	1 217	952	265	27,8	29,3	28,7
30 a 44 años	695	578	117	20,2	16,8	17,4
45 a 59 años	341	230	111	48,3	8,2	6,9
60 a 74 años	114	109	5	4,6	2,7	3,3
75 años o más	37	42	-5	-11,9	0,9	1,3

D. Elementos básicos de la revisión de datos

48. La revisión de datos comprende la inspección sistemática de las respuestas inválidas e incongruentes y la posterior corrección manual o automática (utilizando los “valores desconocidos” o la atribución dinámica), con arreglo a ciertas normas predeterminadas. Algunas operaciones de revisión requieren correcciones manuales, que se realizan en la oficina. Otras operaciones de revisión consisten en correcciones hechas electrónicamente por ordenador. Las publicaciones censales probablemente contendrán cierta proporción de datos sin sentido si las oficinas nacionales de estadística o del censo no revisan los resultados del censo o de la encuesta. Este tipo de revisión reduce la distorsión de las estimaciones, facilita el procesamiento y aumenta la confianza de los usuarios. Además, según Pullum, Harpham y Ozsever (1986), “... lo principal que se consigue con la revisión o la depuración de los datos es, en primer lugar, ver si las distintas respuestas son coherentes entre sí y con el formato básico del instrumento de encuesta”.

49. Los archivos de datos censales no depurados contienen errores de muchos tipos. En el procesamiento de datos los errores se dividen en dos categorías: los que pueden bloquear la continuación del procesamiento y los que producen resultados inválidos o incongruentes sin interrumpir el normal desarrollo de las posteriores operaciones de procesamiento. Como se indica en *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008, párr. 1.311), han de corregirse todos los errores del primer tipo y cuantos sea posible del segundo. Por consiguiente, la finalidad básica de la revisión de los datos censales en la etapa de procesamiento será identificar el mayor número posible de errores y modificar el conjunto

de datos de manera que los asientos sean válidos y coherentes. Sin embargo, en el procesamiento no pueden corregirse todos los errores del censo, incluidas las respuestas al cuestionario que tal vez sean coherentes a nivel interno pero que en realidad representan casos de información incorrecta suministrada por los declarantes, ni los errores de anotación cometidos por los empadronadores.

50. Las normas de revisión de datos suelen ser de dos tipos: 1) **definitivas**, que identifican los errores con certeza, y 2) **indicativas**, que señalan datos problemáticos (Granquist y Kovar, 1997, pág. 420). En la revisión definitiva se indican los valores que indudablemente son un error, mientras que en la revisión indicativa se señalan datos que probablemente son inválidos o incongruentes. Los errores definitivos, que son los errores apreciados por las revisiones definitivas, comprenden los valores inválidos u omitidos y los que presentan incoherencias. Por el contrario, las revisiones indicativas señalan los datos que no cumplen unos criterios de revisión predominantemente subjetivos, elementos con un valor relativamente alto o bajo en comparación con otros datos del mismo cuestionario y otros valores problemáticos. Con miras a mantener la confianza en el censo, particularmente cuando la oficina nacional de estadística o la del censo decide difundir datos de micronivel, el proceso de revisión deberá ser capaz no solo de identificar sino de resolver los errores definitivos. Los errores que saca a la luz la revisión indicativa son más difíciles de corregir, presentan menos ventajas que en el caso de la detección y solución de los errores definitivos e incrementar el costo del proceso total.

51. Puesto que todos los elementos de un censo se han incluido específicamente porque los encargados de la planificación y de la formulación de políticas los necesitan, el número de casos detectados en la revisión indicativa que habrá que resolver durante la fase de revisión y atribución será relativamente mayor que en el caso de las encuestas. No obstante, al determinar los procedimientos de revisión definitiva de un censo el personal temático debería investigar los procedimientos de revisión formulados para los censos piloto y los formulados durante el procesamiento, a fin de asegurarse de que cada uno de ellos produzca los beneficios esperados en relación con su costo. Esas investigaciones deberán formar parte de la evaluación del censo. Como señalan Granquist y Kovar (1997, pág. 422), “los datos sobre tasas de acierto, es decir, la proporción de señales de alerta que tienen como resultado modificaciones de los datos originales, raramente se comunican en las evaluaciones o estudios de los procesos de revisión de datos”.

52. Otra serie de técnicas y terminología es la referente a la revisión de micronivel y de macronivel. Como se ha indicado, los procedimientos de revisión de los censos y encuestas detectan errores en los registros de datos y entre ellos. En el presente *Manual* se describe la revisión de micronivel, que trata de las formas de garantizar la validez y la coherencia de los registros de datos y de las relaciones entre los registros de un mismo hogar. Con ese método se comprueban los datos editados con las frecuencias y tolerancias previstas, con la finalidad de determinar los distintos problemas existentes en los datos; si se encuentran “errores”, la revisión de macronivel puede introducir un cambio general en los datos agregados, devolver un registro para que sea procesado de nuevo o agregar nuevas revisiones de micronivel con el fin de corregir el problema. Por ejemplo, en un país puede haber un porcentaje muy elevado de personas que no comunican su edad. Después de atribuir valores de edad para obtener un conjunto de datos completo, las comprobaciones de macronivel, o de los datos agregados, permitirán asegurar que no haya un sesgo en los valores atribuidos debido a que las personas de más edad dejan de comunicar selectivamente esa información. El equipo de revisión de datos posiblemente decida adoptar medidas con miras a eliminar el riesgo de que se cree un sesgo potencial, según los resultados que arroje el análisis.

Tanto la revisión de macronivel como la de micronivel requieren detenidos ensayos antes de ser utilizadas.

53. Como se ha señalado, el proceso de revisión debería permitir conservar los datos originales en la mayor medida posible. El equipo de revisión necesita contar con datos depurados de alta calidad, pero también necesita conservar lo que la organización ha recopilado sobre el terreno. Los datos originales deben conservarse a lo largo de todas las etapas del procesamiento por ordenador, por si el equipo de revisión necesita volver a examinar todo el proceso. A veces vuelven a examinarse los datos originales cuando el equipo descubre que se ha producido un error sistemático en el proceso de revisión. En otros casos el examen se realiza porque se descubre que falta una parte del conjunto de datos, o que hay una duplicación, y es necesario reconstituir y volver a revisar el conjunto de datos.

54. En ocasiones la fuente de error se encuentra fuera de la oficina de procesamiento. Banister (1980, pág. 2) señala que “si sabemos que una elevada proporción de algún subgrupo no ha respondido a una determinada pregunta del censo, esto significa que los entrevistados no comprendieron la pregunta, que se resistían a responderla o que se mostraban apáticos en cuanto a cooperar con el censo”. Por tanto, la autora argumenta que las tasas de falta de respuesta de determinados subgrupos deberían incluirse en los medios de almacenamiento de datos del censo y en los cuadros publicados. Actualmente es más probable que las oficinas nacionales de estadística o del censo conserven esos datos en un disco compacto o en otros medios, para que puedan utilizarlos los investigadores.

55. Cada vez se pone más de manifiesto que por muchos esfuerzos que se hagan, la revisión de los datos por computadora nunca podrá reemplazar a una reunión de datos censales de gran calidad. Las oficinas nacionales de estadística o del censo saben que a partir de cierto punto la revisión por ordenador no sólo tiene efectos limitados sino que llega a ser contraproducente porque añade al conjunto de datos más errores que los que corrige. Modificar un elemento del censo no es lo mismo que corregirlo. Por lo tanto, los miembros del equipo de revisión de datos deben colaborar para determinar el inicio, el punto medio y la conclusión del proceso de revisión.

56. La revisión y la atribución de valores pueden o no mejorar la calidad de los datos, pero un conjunto de datos depurados facilita mucho el análisis y la utilización. El proceso comienza con el diseño del cuestionario del censo. Los demógrafos y otros especialistas temáticos suelen determinar su contenido, a menudo en consulta con los grupos de usuarios. De todas formas, en definitiva, los datos del censo no se producen primordialmente “para los puristas de la demografía, sino para un público mucho más amplio de académicos, de encargados de la formulación de políticas y de personas no especializadas” (Banister, 1980, pág. 17). Sin embargo, producir un censo limpio de asientos inválidos o incongruentes es esencial cuando está en juego la credibilidad del propio censo y de la oficina nacional de estadística o del censo. Como señala Banister, “las organizaciones encargadas del levantamiento de censos pueden citar casos de periodistas que han escrito artículos humorísticos o de ciudadanos indignados que han dirigido cartas a los funcionarios del censo, movidos por cuadros publicados en los que figuran abuelos de tres años de edad y personas que se trasladan a su trabajo en ferrocarriles inexistentes”.

57. El problema consiste en determinar hasta dónde hay que llegar para obtener un conjunto de datos de buena calidad. Como se ha señalado antes, la aparición de los ordenadores, que al principio eran centrales y más tarde microordenadores, ha permitido una automatización virtualmente completa del proceso de revisión de datos. En muchas oficinas nacionales de estadística o del censo los especialistas temáticos se han convertido efectivamente en entusiastas de la revisión de datos. Así, ahora las

oficinas llevan a cabo muchas pruebas de coherencia que antes resultaban difíciles, particularmente las que entrañan comprobaciones entre registros y comprobaciones entre hogares. Lamentablemente, esta característica de los microordenadores también ha provocado muchos problemas, el mayor de los cuales es la revisión excesiva.

1. Perjuicios de una revisión excesiva

58. Un exceso de revisión tendrá un efecto perjudicial en el proceso por diversos motivos, entre los que se incluyen la puntualidad, el costo y la distorsión de los valores reales. También crea una falsa sensación de seguridad con respecto a la calidad de los datos. Estos problemas se describen a continuación.

a) Oportunidad

59. Cuantas más revisiones realice una oficina nacional de estadística o del censo, más tiempo se necesitará para todo el proceso. El problema principal consiste en determinar hasta qué punto el tiempo adicional aumenta la calidad del producto censal. El equipo de revisión de datos deberá evaluar, mientras desempeña su labor y una vez concluida esta, los beneficios netos que el tiempo y los recursos adicionales han supuesto para el producto general del censo. A menudo, los beneficios son tan pequeños en relación con el tiempo invertido, que es preferible que los datos presenten pequeñas deficiencias que privar a los usuarios principales de la posibilidad de recibir puntualmente la información.

b) Costo

60. Análogamente, el costo del proceso censal aumenta a medida que pasa el tiempo. Cada oficina nacional de estadística o del censo debe decidir, a medida que aumenta el volumen y la complejidad de su labor de revisión de datos, si los incrementos de costos están justificados y si puede permitirse esos costos adicionales.

c) Distorsión de los valores reales

61. Aunque el propósito del proceso de revisión es mejorar la calidad de los datos, el aumento del volumen y de la complejidad de la revisión también puede tener efectos negativos. A veces los equipos de revisión modifican equivocadamente los elementos por motivos como la mala comunicación entre los especialistas temáticos y los especialistas en procesamiento de datos, los errores que contiene un programa muy complejo y avanzado o la repetida modificación de un elemento del censo durante el proceso de revisión. Las oficinas nacionales de estadística o del censo procurarán evitar este tipo de problemas siempre que sea posible. Granquist y Kovar (1997) señalan, por ejemplo, que atribuir un valor a la edad de un marido y de una esposa utilizando una diferencia de edades establecida entre ellos puede ser útil pero también crear un sesgo artificial en los datos cuando hay muchos casos de este tipo.

d) Una falsa sensación de seguridad

62. La revisión excesiva otorga a los funcionarios de la oficina nacional de estadística o del censo y a otros usuarios una falsa sensación de seguridad, especialmente cuando las oficinas no aplican ni documentan medidas de garantía de la calidad. Además, en las tabulaciones de los censos aparecerán resultados raros por mucho que se esfuerce el equipo de revisión de datos, por lo que es importante advertir a los usuarios de que pueden producirse pequeños errores. Esto es especialmente cierto en la actualidad, cuando muchos países publican datos de micronivel tomados de muestras. Las oficinas

nacionales de estadística o del censo puede que no deseen publicar datos perjudiciales para el proceso de planificación, de manera que habrá que prestar mucha atención para asegurar que todas las variables cruciales se han revisado de manera apropiada y puedan utilizarse para la planificación. Por ejemplo, ninguna oficina nacional de estadística o del censo querrá publicar datos de micronivel o tabulaciones con valores desconocidos para las variables de sexo o edad. Por otra parte, variables como la discapacidad o el grado de alfabetización funcionan igualmente bien con un nivel de revisión menor. Si bien pueden producirse algunas incongruencias en las tabulaciones cruzadas porque las oficinas nacionales de estadística o del censo no pueden revisar todos los pares de variables, los equipos de revisión deberían comprobar las combinaciones más importantes y disponer de procedimientos de corrección cuando se encuentren incongruencias.

2. Tratamiento de los valores desconocidos

63. El equipo de revisión de datos deberá decidir en una fase temprana de la planificación del censo la forma de tratar los casos no declarados o desconocidos. Como ya se ha señalado, los cuadros con columnas o entradas de valores desconocidos no proporcionan información y no siempre son útiles, de manera que los planificadores de la mayoría de países prefieren que se atribuyan valores a esas variables. Si no se aplica algún método para tratar los valores desconocidos, muchos usuarios distribuirán tales valores en la misma proporción que los datos conocidos, lo cual equivale a atribuir los valores desconocidos después de la tabulación. Además, algunos valores informativos inválidos no pueden tratarse como valores en blanco, y deben aclararse. Por ejemplo, “América del Sur” como lugar de nacimiento debe atribuirse a un país (por ejemplo, el Perú). Es necesario que el equipo de revisión de datos decida la manera de tratar sistemáticamente los valores desconocidos.

3. Modificaciones erróneas

64. Aunque podrían hacerlo, las oficinas nacionales de estadística o las del censo no suelen trabajar con modelos cuando formulan sus normas de revisión de datos. Los equipos de revisión deberán formular normas que se adapten a las características reales de la situación de la población o a la de la vivienda. Todos los datos deben ajustarse a las normas de revisión. Por ejemplo, puede haber una serie de normas que requiera que el hijo de un jefe del hogar tenga como mínimo 15 años menos que el jefe del hogar. Sin embargo, el hijo del jefe del hogar puede ser un hijo social y no un hijo biológico, ya que puede tratarse del hijo biológico del cónyuge pero no del jefe del hogar. Por lo tanto, la diferencia de edad podría ser inferior a 15 años. Puesto que en la mayoría de países no se planifica separadamente para los hijos y los hijastros, si en las circunstancias indicadas las normas de revisión de datos hacen que se modifique la edad del hijo, puede haber una incongruencia entre los datos resultantes y los valores del nivel de instrucción, la participación en la fuerza de trabajo y otras variables. Así pues, esta norma debería someterse a prueba, para comprobar sus resultados, antes de aplicarla efectivamente.

4. Determinación de niveles de tolerancia

65. El equipo de revisión de datos deberá establecer “niveles de tolerancia” para cada elemento, y a veces para combinaciones de elementos. Los niveles de tolerancia indican el número de respuestas inválidas o incongruentes que se aceptarán antes de que los equipos de revisión adopten medidas correctivas. En la mayoría de elementos de un censo, por ejemplo, un pequeño porcentaje de los declarantes no dará respues-

tas “aceptables”, por distintos motivos. Respecto de algunos elementos como la edad y el sexo, que se utilizan en combinación con tantos otros elementos para fines de planificación, el nivel de tolerancia podría ser bastante bajo. Cuando el porcentaje de respuestas en blanco o incongruentes sea bajo (menos del 1% o el 2%), unas normas de revisión razonables probablemente no tendrán efectos en la utilización de los datos. Cuando el porcentaje sea elevado (del 5% al 10% o más, según la situación), una atribución de valores simple o incluso una atribución compleja pueden distorsionar los resultados del censo.

66. A fin de reducir al mínimo las respuestas en blanco, las oficinas nacionales de estadística o del censo deben velar por que los empleados del censo procuren por todos los medios obtener la información sobre el terreno. Si en un país determinado se decide que no se necesita tanta precisión respecto de algunos elementos, como el grado de alfabetización o la discapacidad, el nivel de tolerancia para esos elementos podría ser muy superior. A veces los equipos de revisión de datos pueden corregir elementos que presentan demasiados errores haciendo que los empadronadores regresen sobre el terreno, realizando nuevas entrevistas por teléfono o aplicando su conocimiento de una zona determinada. Sin embargo, suele ser demasiado costoso regresar sobre el terreno o realizar otras operaciones complementarias, y es posible que la oficina nacional de estadística o del censo decida no utilizar el elemento o, si lo utiliza, agregarle una nota cautelar.

67. Una cuestión que se plantea es la de quién debe encargarse de determinar el nivel de tolerancia correspondiente a un elemento. Podría ser que el equipo de revisión de datos, integrado por especialistas temáticos y especialistas en procesamiento de datos, fuera el que decidiera los niveles de tolerancia. El personal temático deberá seguir utilizando los elementos durante algún tiempo, y por tanto tendrá interés profesional en velar por que se obtengan datos de la máxima calidad. Sin embargo, es posible que los especialistas en procesamiento de datos consideren que en la práctica no pueden elaborar programas de revisión apropiados para reducir la tolerancia a niveles aceptables, o que los propios datos no admitan ningún programa que logre mantenerlos dentro de los límites de tolerancia.

5. Enseñanzas extraídas del proceso de revisión de datos

68. A medida que se van revisando los datos hay que efectuar análisis detallados de las experiencias positivas y negativas adquiridas, a fin de mejorar la calidad del censo o de la encuesta en curso y de los censos y encuestas futuros. El equipo de revisión de datos debe procurar siempre determinar qué es lo que funciona adecuadamente y qué es lo que no. También deberá determinar si los aspectos del proceso que funcionan adecuadamente pueden mejorarse y agilizarse de modo que los datos lleguen a los usuarios con mayor rapidez. Durante el proceso censal, cuanto antes detecte los errores la oficina nacional de estadística o la oficina del censo, más probable será que se corrijan.

6. Garantía de calidad

69. La garantía de calidad es importante en todas las operaciones censales. Por consiguiente, deberán haberse establecido mecanismos de garantía de la calidad para supervisar los progresos de la fase de revisión y atribución de valores por computadora. Los rastros de verificación de cambios, las evaluaciones de resultados y las estadísticas de diagnóstico son fundamentales para analizar la calidad de la labor de revisión y la rapidez del procesamiento (Granquist y Kovar, 1997; Statistics Canada, 1998).

7. Costos de la revisión

70. El presente *Manual* puede ayudar a los países a reducir los elevados costos que, tanto en tiempo como en recursos, supone realizar la labor de revisión y atribución de los datos de los censos o encuestas. Como señalan Granquist y Kovar (1997, pág. 418), incluso “en el decenio de 1990 la labor de revisión de datos es esencialmente tan costosa como lo era en el decenio de 1970, aunque el proceso se ha racionalizado en gran medida gracias a la constante aplicación de los adelantos tecnológicos”. Para la mayoría de los países, los trabajos de revisión representan un volumen tremendo de tiempo y de recursos, de manera que cada país debe decidir sobre el rendimiento de su inversión. Según los mismos autores, a comienzos del decenio de 1990 el costo de la revisión de los datos de las encuestas por hogares representaba aproximadamente el 20% del presupuesto total de los censos de todo el mundo.

71. La revisión excesiva puede demorar los resultados del censo. Aunque los funcionarios de las oficinas nacionales encargadas de los censos o encuestas sólo puedan aportar pruebas incidentales de este tipo de experiencias con los censos, un estudio llevado a cabo por Pullum, Harpham y Ozsever (1986) concluyó que la revisión automatizada del Estudio Mundial sobre la Fecundidad supuso un retraso de alrededor de un año en la publicación de los resultados. Sería mejor que las oficinas nacionales de estadística o del censo dedicaran sus recursos a aumentar la calidad del empadronamiento correspondiente al censo o a la encuesta.

8. Atribución de valores

72. La atribución de valores es el proceso que se sigue para resolver los problemas que presentan las respuestas en blanco, inválidas o incongruentes aparecidas durante la revisión de los datos. La atribución consiste en modificar una o más de las respuestas o de los valores omitidos en un registro o en varios registros que se están revisando, para asegurar la plausibilidad y la coherencia interna de los registros. El contacto con el declarante o un estudio manual del cuestionario pueden eliminar algunos problemas en una etapa temprana del proceso. Sin embargo, debido a consideraciones relativas al volumen de respuestas, al costo y a la puntualidad suele ser imposible resolver todos los problemas en esas etapas tempranas. En esos casos se recurre a la atribución de valores para resolver los restantes elementos inaceptables, ya que es conveniente producir un archivo completo y coherente con datos atribuidos. La mejor labor de atribución es la que realizan los miembros del equipo que tienen pleno acceso a los datos de micronivel y buena información complementaria.

- a) El registro que se obtiene después de la atribución de valores debe parecerse mucho al registro que presentaba deficiencias. Suele ser mejor atribuir valores a un número mínimo de variables, de forma que se conserve el mayor número posible de datos del declarante. La hipótesis subyacente (que no siempre es cierta en la práctica) es que es más probable que un declarante no cometa varios errores sino uno o dos;
- b) El registro obtenido mediante atribución debe cumplir satisfactoriamente todas las normas de revisión;
- c) Los equipos de revisión de datos deberían indicar los valores obtenidos mediante atribución e identificar claramente los métodos y las fuentes de atribución;
- d) El equipo de revisión de datos debería conservar los valores atribuidos y no atribuidos de un determinado registro, a fin de evaluar el grado de atribución y sus efectos.

9. Archivo

73. Una parte de la labor de garantía de calidad del censo o de la encuesta consiste en documentar todos los procesos y a continuación archivar esa documentación. Las oficinas nacionales de estadística o del censo deben conservar los registros de datos depurados y no depurados para su análisis posterior. Algunos procedimientos, como muchas formas de lectura por escáner, permiten conservar automáticamente la imagen original. Por otra parte, inmediatamente después de introducir los lotes de datos, estos deben ser concatenados y conservados para su posible análisis. No obstante, con uno u otro procedimiento, es importante archivar las copias originales de los archivos no revisados. Es más, deberían conservarse copias de los datos no revisados atribuidos en varios lugares de la oficina de estadística, así como en otros lugares del país e incluso fuera de él.

74. La documentación debe ser completa, de manera que los planificadores de los censos o las encuestas puedan reconstruir más adelante los mismos procesos a fin de asegurar la compatibilidad con el censo o con la encuesta que se está examinando. Los procesos y los resultados deben poder reproducirse. Por último, tanto los datos no depurados como los datos depurados deberán almacenarse en varios lugares y deben adoptarse medidas apropiadas para asegurar que sigan disponibles en el futuro.

75. Como se observa en otro lugar, parte de la documentación implica dos tipos de informes de revisión. El primero presenta una estadística sumaria con el número y el porcentaje de errores (basada en denominadores adecuados, como el total de unidades de habitación, la población total, la población en edad de trabajar, las mujeres adultas, etcétera). El segundo contiene al menos una muestra de la estructura del “caso”, con el registro no revisado de hogares o viviendas, la lista de errores y sus soluciones para la unidad de vivienda o los componentes de esta y los datos revisados sobre la unidad de vivienda o el hogar.

76. Los dos conjuntos de errores deben presentarse en los niveles geográficos lógicos, y sin duda ninguna en relación con las grandes divisiones civiles; por otro lado, la presentación de listas de errores de nivel geográfico inferior podría ayudar a seleccionar los problemas en la capacitación de patrocinadores, el control de calidad y otras cuestiones relacionadas con el empadronamiento.

Capítulo II

Práctica de la revisión de datos

77. El presente capítulo ofrece un panorama general de las aplicaciones del proceso de revisión y atribución. Contiene un marco para el flujo general de la revisión del censo o la encuesta que parte de los datos brutos escaneados o introducidos manualmente y, pasando por la revisión de la estructura y el contenido, llega hasta la presentación de un conjunto de datos editados⁵. Los ejemplos se han seleccionado con el fin de ilustrar qué tipos de problemas pueden presentar a los usuarios los datos no revisados y por qué los datos revisados son más útiles. Se consideran cuestiones asociadas con la introducción manual y la codificación, en cuanto relacionadas con el proceso de revisión preliminar. Se tienen también en cuenta cuestiones generales de la revisión con computadora junto con las orientaciones sobre temas como la comprobación de la validez y la coherencia. Se examinan con detalle los dos tipos genéricos de revisión con computadora y las técnicas de atribución estática (*cold deck*) y de atribución dinámica (*hot deck*).

78. Independientemente de que los datos del censo se escaneen o se introduzcan manualmente, cabe hablar de un determinado flujo general. El equipo de revisión del censo comienza con los datos no revisados. En la mayor parte de los casos todos los datos han sido codificados previamente por el empadronador o por el personal de la oficina, por lo que el conjunto de datos está ya listo para la revisión de la estructura. En algunos casos se necesita una operación para convertir los datos escaneados de manera que pueden ser utilizados con otra forma de lectura mecánica con vistas al proceso de revisión, de acuerdo con el programa de revisión que se vaya a utilizar. En ciertos casos, los datos escaneados requieren también de una segunda operación de codificación automatizada para que puedan rellenarse informaciones como las referentes al lugar de nacimiento, la rama de actividad y la ocupación.

79. En uno y otro caso, los datos no revisados deben aparecer de tal forma que los programas informáticos puedan realizar las revisiones estructurales (tal como se describen detalladamente en el capítulo III). La revisión estructural implica una comprobación para garantizar que todas las grandes divisiones civiles estén presentadas en orden geográfico o numérico y que dentro de cada gran división civil aparezca cada división civil de importancia secundaria, también en orden geográfico o numérico. Luego, dentro de cada división civil secundaria debe aparecer cada localidad, de acuerdo con un orden geográfico o numérico. Este procedimiento continúa hasta el nivel geográfico más bajo. Como se describe en el siguiente capítulo, deben establecerse procedimientos adecuados para garantizar que cada unidad de vivienda aparezca una sola vez en el conjunto de datos.

80. La revisión estructural debe comprobar también que todos los tipos de registro estén presentes cuando convenga y que ninguno se repita si no es necesario. Así, en el caso de un censo de población y de vivienda deben ir en primer lugar los registros de población o los de vivienda, y esa convención debe respetarse en todo el conjunto de datos. En la mayoría de los casos sólo estarán presentes los registros

⁵ Cuando inicialmente se escribió este Manual para los censos de 2000, casi todos los países introducían sus datos a mano. Ahora, casi todos los países utilizan el escáner, algunas veces complementado con la introducción manual. El presente Manual trata de tener en cuenta el escaneado para las revisiones de estructura y contenido. Mientras se preparaba la edición fueron apareciendo nuevas tecnologías, que implicaban el uso de los asistentes personales digitales (PDA), e Internet para la recopilación de datos y la revisión interactiva (véase, por ejemplo, Ireback (2000) sobre el uso de la Internet en Suecia). Lo mismo que los países tecnológicamente en desarrollo tenían problemas con el escaneado en los primeros años 2000, ahora muchos países consideran que todavía debe perfeccionarse el uso de los PDA.

de vivienda, por lo que es preciso resolver el problema de los registros excedentes; los programadores deben suministrar también registros de vivienda a los hogares sin registros de vivienda. De la misma manera, en las unidades de vivienda ocupadas (definidas normalmente como tales en el registro de vivienda) debe haber personas presentes, pero no en las unidades desocupadas.

81. Con respecto a la estructura, es importante señalar que, inevitablemente, la revisión de la misma será examinada de nuevo durante la revisión del contenido, e incluso posteriormente, a medida que vayan apareciendo pequeñas deficiencias durante los distintos procesos. Este es el procedimiento normal del censo, y debería contarse con él; posteriormente, se tendrán en cuenta las necesidades de tiempo, personal y equipo.

82. A partir de ahí puede comenzar la revisión de contenido. Cada información relativa a la población y a la vivienda debe considerarse por su cuenta, y normalmente también en forma conjunta, para determinar la validez considerada de cada información y el mejor ajuste entre ellas. En los capítulos IV y V se examinan los diversos elementos relativos a la población y la vivienda que aparecen en *Principios y recomendaciones para los censos de población y habitación, Revisión 2*.

83. Una vez terminada la revisión del contenido debe establecerse un conjunto de datos completamente revisados. Los datos no revisados deberán almacenarse en varios lugares seguros, y las partidas no revisadas importantes (o todas ellas) deberán aparecer también al final de los distintos tipos de registros. Una vez más, es importante señalar que a medida que se elaboran los cuadros quizá deban revisarse las revisiones de contenido, a fin de resolver los problemas específicos que puedan resultar de las tabulaciones cruzadas.

84. El propósito de la revisión de los datos de los censos y encuestas es descubrir omisiones e incongruencias en los registros de datos. La atribución se utiliza para corregirlos. En la revisión de datos se establecen procedimientos específicos para resolver las omisiones y los diversos tipos de asientos inaceptables. La atribución modifica los valores inválidos y resuelve las incongruencias que se encuentran en el conjunto de datos. El producto que se obtiene es un archivo depurado de datos de micronivel preparados para su tabulación, que contiene valores aceptables y coherentes para todos los elementos pertinentes de cada unidad de habitación y de cada persona empadronada.

85. Es importante subrayar que, por mucho esfuerzo que se le dedique, la revisión no podrá reemplazar a una labor de empadronamiento de buena calidad. El proceso de revisión de datos funciona bien cuando las atribuciones se utilizan para resolver omisiones e incongruencias aleatorias. Sin embargo, si se producen errores sistemáticos durante la reunión de datos, la revisión no podrá mejorar la calidad de los datos por muy perfeccionados que sean los procedimientos que se apliquen. La elección de los temas que habrá que investigar es de importancia decisiva para la calidad de los datos obtenidos. Al ser entrevistados, los declarantes deben tener la voluntad y la capacidad de suministrar información apropiada y suficiente. Por tanto, tal vez sea necesario evitar temas que puedan dar lugar a temores, prejuicios locales o supersticiones, o preguntas demasiado complicadas y difíciles que el declarante medio no puede responder fácilmente en el contexto de un censo de población. La redacción exacta de cada pregunta para obtener la respuesta más fidedigna dependerá necesariamente de las circunstancias nacionales y deberá ser comprobada debidamente antes del censo. Por consiguiente, es de la máxima importancia que las oficinas nacionales de estadística o del censo asignen recursos suficientes para obtener datos censales de la máxima calidad.

86. Para ejecutar la fase de revisión del proceso por medio de ordenador, el equipo de revisión de datos preparará instrucciones o especificaciones por escrito,

algoritmos, diagramas de secuencia y pseudocódigos. Un pseudocódigo es un conjunto de especificaciones o instrucciones de revisión, como las que pueden verse en el gráfico 8.

87. Los diagramas de secuencia ayudan a los especialistas temáticos a comprender las distintas relaciones existentes entre las variables y facilitan la redacción de las instrucciones para la revisión de datos. En el anexo IV se presentan ejemplos de diagramas de secuencia. Los especialistas temáticos colaboran con los especialistas en informática para redactar las instrucciones de revisión de datos, en las cuales se describen las operaciones correspondientes a cada elemento. Las instrucciones de revisión de datos deben ser claras y concisas y no presentar ambigüedades, ya que servirán de base para el programa informático de revisión.

88. Todo el equipo de revisión de datos del censo —tanto los especialistas temáticos como los especialistas en procesamiento de datos— debe estar bien familiarizado con el procesamiento y el análisis de datos demográficos. El personal poco cualificado podría introducir fortuitamente errores y sesgos adicionales en el censo.

A. Consideraciones relativas a la codificación

89. Durante gran parte de la segunda mitad del siglo XX, como se ha señalado antes, los países introducían los datos manualmente. Ahora, la mayoría de los países escanean los censos, aunque en muchos casos continúan utilizando el teclado. Aun cuando los formularios se escaneen, algunas variables deben traducirse de palabras a números. El proceso de generación de números y caracteres alfanuméricos se conoce con el nombre de **codificación**.

90. Algunos programas de revisión pueden reconocer y utilizar fácilmente datos alfanuméricos, pero la mayoría de ellos tienen dificultades para clasificar, resumir y deducir porcentajes, medianas, etcétera, cuando se incluyen datos no numéricos.

91. Los códigos compuestos en su totalidad por caracteres alfabéticos o caracteres alfabéticos combinados con números (llamados alfanuméricos) deben evitarse en lo posible. Cuando se escanean los formularios, los caracteres alfanuméricos no plantean un problema significativo; en cambio, para muchos programas informáticos su uso requiere considerable manipulación o al menos una gran atención. Muchos programas de revisión pueden procesar caracteres alfanuméricos sólo si aparecen entre comillas o se distinguen de alguna otra manera.

92. Al elaborar un plan de codificación, el personal encargado del censo y de las encuestas debe considerar el rendimiento de cada inversión de tiempo, energía y fondos. Las consideraciones relativas a la codificación no tienen demasiada importancia cuando se trata de países o encuestas de pequeña magnitud, ya que el volumen de procesamiento requerido es mucho menor que en un censo. Asimismo, los datos escaneados no se ven tan afectados por los problemas asociados con las columnas adicionales de información.

93. Por otro lado, por ejemplo, cuando un censo o encuesta utiliza para la *relación* dos columnas en lugar de una, el escaneado introducirá errores que no se darían si hubiera una sola columna de información; es decir, con los códigos 1 a 9, el escáner puede captar un carácter alfabético, o un espacio en blanco, o una marca parásita convertida a un carácter legible. No obstante, estas cuestiones se resuelven fácilmente en la revisión, como se describe con mayor detalle en otra parte de la publicación.

94. Sin embargo, cuando hay dos columnas, por ejemplo, los códigos 1 a 10, existe el riesgo de introducir todo un nuevo espectro de errores. En vez de los valores

aceptables, de 1 a 9, ahora habría valores de entrada que se situarían entre 0 y 99, además de los mencionados caracteres alfabéticos, respuestas en blanco y marcas parásitas. Cuando los revisores reciben un valor de 13 deben comenzar a tomar decisiones estratégicas sobre qué hacer con ese valor. ¿Debería ser 3, en cuyo caso el 1 inicial sería un error? ¿Debería ser 10, y el error estaría en el 3? En la mayoría de los casos, los especialistas ofrecen las especificaciones de revisión pertinentes, pero la presencia de estos valores aumenta automáticamente el tiempo y la complejidad de la revisión y podría reducir la calidad del conjunto de datos final.

95. Uno de los problemas más comunes, que se examina con detalle más adelante, es el relacionado con la serie relativa a la fecundidad. Muchos países recopilan ahora información sobre los hijos que viven en el hogar, los hijos que viven en otro lugar y los hijos fallecidos, y algunas veces sobre la suma de todos ellos, desglosada por sexo. Así pues, los países pueden tener hasta 12 informaciones distintas. La cuestión que se plantea aquí es cuántos dígitos deberían utilizarse en cada uno de esos elementos. Cuando se utilizan dos columnas, el valor correspondiente a los hijos varones en la casa podría oscilar entre 0 y 99; cuando se utiliza sólo una columna, el margen va sólo de 0 a 9. No obstante, ya que es sumamente improbable que una mujer tenga más de nueve hijos varones en el hogar, la utilización de dos dígitos introduce una elevada probabilidad de detectar las marcas parásitas o los errores de lectura del escáner, que leería 9 en vez de 0, por ejemplo, lo que daría lugar a un total de 91 hijos en vez de 01.

96. Así pues, por lo que se refiere a los niños y niñas que se encuentran en la casa, que están actualmente en otro lugar o que han fallecido, sería más conveniente utilizar una sola columna. En cambio, para el total de hijos en el hogar, hijos en otros lugares, total de hijos fallecidos y total de hijos, quizá fuera más adecuado utilizar dos columnas. Ello depende en gran parte de los niveles de fecundidad del país. En algunos casos un hogar puede tener excepcionalmente más de nueve personas en una categoría determinada; pero, como ocurre siempre en los censos, la oficina de estadística deberá tratar de buscar una solución de compromiso entre el número de errores y la calidad de los datos.

97. Por lo que se refiere a las variables ordinales, cabe señalar la siguiente serie de códigos de relación:

- 1 Jefe de hogar (o cabeza del hogar)
- 2 Cónyuge
- 3 Hijo
- 4 Hermano
- 5 Progenitor
- 6 Nieto
- 7 Otro familiar
- 8 Persona no emparentada

En la mayoría de los países este conjunto de códigos estándar abarca la mayor parte de las relaciones. Algunos países añaden un código "0" para el jefe del hogar, y luego pueden agregar una décima categoría.

98. Estos códigos pueden utilizarse para obtener la composición del hogar, como se observa en el anexo I relativo a las variables derivadas. No obstante, muchos países, en particular los que sufren la epidemia del VIH/SIDA, necesitan información mucho más detallada de la que pueden ofrecer esos códigos. Dichos países quizá necesiten información específica sobre los yernos y nueras, suegros, abuelos, sobrinos y sobrinas, etcétera. En tal caso, la oficina de estadística necesitará códigos adicionales de dos dígitos para realizar su misión.

99. Cuando un país decide utilizar varias columnas debe decidir también cómo. En el ejemplo anterior se supone que los códigos de relación serán secuenciales. No obstante, una vez que se ha tomado la decisión de utilizar dos columnas, los especialistas temáticos pueden decidir la importancia atribuida a las columnas. Por ejemplo:

- 10 Jefe de hogar
- 11 Cónyuge
- 12 Hermano
- 13 Cónyuge del hermano
- 21 Hijo
- 22 Hijo adoptado
- 23 Hijastro
- 24 Sobrino
- 31 Progenitor
- 32 Suegro
- 33 Tío
- 41 Nieto
- 77 Otro familiar
- 88 Persona no emparentada
- 90 Población en instituciones

100. Este esquema codifica la generación en la primera columna: 1 para la generación del jefe de hogar, 2 para una generación inferior, 3 para una generación superior, 4 para dos generaciones inferiores, etcétera, y luego enumera los tipos de familiares dentro de cada una de las categorías. Si bien estos valores pueden ser útiles en la reconstrucción de la familia, el personal de la oficina y algunos usuarios generales quizá los encuentren demasiado complicados.

101. No obstante, este tipo de codificación debería considerarse en el caso de determinadas variables sociales y económicas. Por lo que se refiere al grupo étnico, por ejemplo, el primer dígito denotaría la agrupación tribal o étnica principal, y el segundo la agrupación tribal o étnica secundaria. Cuando existen más de 10 agrupaciones secundarias, obviamente deberán utilizarse dos dígitos en la primera columna.

102. De la misma manera, en los elementos que implican tres o cuatro dígitos, como la ocupación o la rama de actividad, el primer dígito denotaría la ocupación o rama de actividad principal; el segundo, la ocupación o rama de actividad secundaria; y el tercero, la ocupación o rama de actividad específica. Dado que la mayoría de los planes internacionales de codificación tienen ya inscritos los niveles en los códigos, la oficina de estadística no tiene que realizar ninguna labor adicional.

103. Cuando elaboren listas de códigos para el programa de revisión y para las tabulaciones ulteriores, las oficinas nacionales de estadística o del censo tal vez deseen establecer códigos comunes para algunos elementos. En muchos países, por ejemplo, los códigos relativos a lugares (lugar de nacimiento, lugar de nacimiento de los padres, lugar de residencia anterior, lugar de trabajo), idioma, grupo étnico/raza y ciudadanía son muy similares. Un plan de codificación común para la variable "lugar" podría aplicarse mediante un código de tres cifras, la primera de las cuales representaría el continente, la segunda la región y la tercera el país concreto. Las oficinas nacionales de estadística o del censo también pueden utilizar los códigos numéricos de los países establecidos por organismos internacionales como la División de Estadística de las Naciones Unidas (Naciones Unidas, 1999). Un conjunto de códigos comunes para variables estrechamente relacionadas puede reducir los errores de codificación y ayudar a los encargados del procesamiento de datos durante la labor de revisión.

Los códigos comunes también permiten que, de ser necesario, los encargados del procesamiento de datos trasladen una anotación de un elemento a otro.

104. La estructura de la codificación puede facilitar el proceso de codificación y también el ulterior procesamiento en las fases de revisión, tabulación y análisis. Para los países grandes con muchos inmigrantes o grupos étnicos, los códigos basados en el continente, la región y el país, con distintos códigos o cifras asignados a cada una de esas unidades, son preferibles a una mera enumeración.

105. En el gráfico 1 se presentan ejemplos de códigos comunes para elementos como el lugar de nacimiento, la ciudadanía, el idioma y el grupo étnico. En el caso de Filipinas, los códigos correspondientes a quienes hablan ilokano y tagalo son distintos del código general de los idiomas de Filipinas. Dependiendo de cuál sea la situación concreta del país, esos códigos también podrían ser distintos entre sí. A pesar de que lo habla más de un grupo, el idioma inglés tiene un único código. Así, los códigos relativos al lugar de nacimiento, la ciudadanía y el grupo étnico del Canadá y los Estados Unidos son ligeramente diferentes. Para las personas nacidas en Francia, de ciudadanía francesa, que hablan francés y cuyo grupo étnico es el francés, se utiliza el mismo código. Por lo tanto, si falta uno de esos elementos, y si el equipo de revisión decide que esta solución es apropiada, el encargado del procesamiento de datos puede aplicar el código de una de las otras variables.

106. Si los elementos de un cuestionario que forman un grupo no son independientes entre sí, es probable que los funcionarios de la oficina nacional encargada de los censos y de las encuestas no hagan todas las preguntas correspondientes. El equipo de revisión deberá decidir caso por caso cuándo habrá que utilizar directamente otros elementos para la asignación de valores y cuándo habrá que recurrir a otras variables disponibles.

Gráfico 1

Ejemplos de códigos comunes para determinados elementos

Grupo	Lugar de nacimiento	Ciudadanía	Idioma	Etnia
Francia/francés	10	10	10	10
España/español	20	20	20	20
América Latina	25	25	20	25
Filipinas/filipino	30	30	30	
ilokano			32	
tagalo			32	
Inglaterra/inglés	40	40	40	40
Canadá	50	50	40	50
Estados Unidos	52	52	40	52

107. Otro problema es el que se produce cuando entre los distintos censos (o entre un censo y una encuesta) difieren las definiciones relativas a variables como el trabajo o el grupo étnico. La oficina nacional de estadística o del censo deberá decidir cómo tener en cuenta esos cambios, tanto para los datos que se están revisando en ese momento como para los conjuntos de datos del censo anterior, a fin de poner de manifiesto las tendencias. Si se dispone de los datos originales no depurados, los encargados del procesamiento de datos pueden modificar las variables apropiadas y volverlas a procesar.

108. Es posible incluso que un país europeo utilice un único código de país de origen para todos los países del Asia meridional cuando sólo es necesario identificar algunos casos sueltos. Sin embargo, dada la evolución de las pautas de migración, a

lo largo de toda la labor de procesamiento de la encuesta o censo siguientes tal vez se requieran códigos separados para la India, Bangladesh, el Pakistán, Sri Lanka y otros países del Asia meridional.

B. Corrección manual frente a corrección automatizada

109. La revisión manual de los datos de un censo puede llevar meses o años y presenta muchas posibilidades de error humano. La revisión manual es un método inferior a la revisión por ordenador, en parte porque es imposible crear o reconstruir una secuencia de revisiones en el proceso de corrección manual. La revisión por ordenador, o automatizada, reduce el tiempo necesario y la posibilidad de introducción de errores humanos. En la revisión, tanto computarizada como manual, se comprueba la validez de un asiento buscando un valor aceptable, pero los programas informáticos también comprueban el valor del asiento comparándolo con variables conexas para determinar su coherencia. Por último, un aspecto muy importante es que la revisión automatizada permite crear un rastro de las revisiones y, por tanto, el proceso puede reproducirse, mientras que la revisión manual no ofrece esta posibilidad.

110. En los primeros años de introducción de datos con ordenador no era posible realizar una revisión en el momento de la entrada de datos; es decir, todas las correcciones debían hacerse manualmente, en el marco de las operaciones de codificación y comprobación en la oficina, o formaban parte de las operaciones informáticas que seguían a la introducción manual de los datos. Los programas más recientes incluyen funciones de revisión, que evitan entradas inválidas a no ser que sean impuestas por los encargados de introducir los datos, y que se puedan señalar contradicciones, para su corrección por los encargados de introducir los datos en forma manual o por un programador informático. Tras la introducción del escáner ha ocurrido algo semejante: en los primeros años no era posible revisión alguna durante la entrada de los datos; en cambio, recientemente, en los sistemas de escaneado se han introducido revisiones de la validez y conversiones de datos y recodificaciones.

111. Cuando en los censos y encuestas se recoge un gran volumen de datos, el personal no siempre puede consultar los documentos originales para corregir los errores. Aun cuando se disponga de los cuestionarios originales, los datos que constan en ellos puede que a veces estén equivocados o sean incongruentes. Un sistema computarizado de revisión y atribución de datos corrige y cambia inmediatamente los datos erróneos y genera informes de todos los errores descubiertos y de las modificaciones efectuadas. Las revisiones por ordenador deben planearse minuciosamente, de manera que el personal pueda dedicar su tiempo a otras tareas de procesamiento de datos. El examen de grandes cantidades de datos mediante un sistema informático puede requerir mucho tiempo, pero nunca tanto como en la corrección manual.

112. La corrección manual adopta varias formas. Un ejemplo sencillo sería el de un error en la respuesta relativa a la variable del sexo: un supervisor comprueba el trabajo de un empadronador y encuentra un error evidente, como asignar “sexo masculino” a una persona que se llama “María”. Si cambia la respuesta por “sexo femenino”, el supervisor realiza una revisión manual. Si el supervisor no corrige el cuestionario sino que lo envía a una oficina establecida sobre el terreno, los empleados de esa oficina pueden observar el problema y corregirlo manualmente. En la oficina central, durante la codificación, los codificadores pueden observar la discrepancia entre el nombre y el sexo y efectuar entonces la corrección manual. En caso de que los codificadores no adviertan el problema, cuando los empleados correspondientes introduzcan los datos del cuestionario tal vez descubran la discrepancia entre el nombre y el sexo y efectúen la corrección manual antes de introducir los datos.

113. Sin embargo, si el error pasa inadvertido y el empleado introduce el código correspondiente a “sexo masculino”, pueden seguirse varios procedimientos distintos a partir de ese momento. Respecto de los elementos relacionados con el género, como el bloque relativo a la fecundidad, el programa de revisión puede alertar del hecho de que se trata de una persona de sexo masculino con información sobre fecundidad y producir un mensaje que lo indique, mientras se están introduciendo los datos. El encargado de introducir los datos puede examinar el cuestionario en ese momento, descubrir que realmente se trata de una mujer y hacer la corrección manualmente. Otra posibilidad, si la oficina nacional de estadística o del censo utiliza un programa de revisión independiente de la introducción de datos, es que el programa informático alerte de que esa persona de sexo masculino tiene información sobre fecundidad. Entonces, utilizando la información geográfica, los funcionarios pueden encontrar el cuestionario original en los archivos, extraerlo y determinar que el valor del sexo del declarante, cuyo nombre es “María”, se ha consignado erróneamente como “masculino” en vez de “femenino”. En ese momento los funcionarios pueden transmitir la información al encargado de introducción de los datos, que recuperará el registro y efectuará la corrección manual.

114. El ejemplo anterior demuestra tanto las ventajas como los inconvenientes de la revisión manual. En cualquiera de las etapas indicadas un empleado del censo podría detectar el error —la discrepancia entre el nombre y el sexo— y efectuar la corrección. Sin embargo, las oficinas nacionales de estadística o del censo que utilizan la revisión manual probablemente cuenten con funcionarios que comprueban esta relación en cada una de las etapas. A esta actividad se dedica una gran cantidad de energía, y los resultados probablemente difieren muy poco, particularmente con respecto a los valores agregados, de los que se habrían obtenido si el personal no hubiera recibido instrucciones de realizar revisiones manuales.

115. Inicialmente, la única forma de efectuar correcciones en un conjunto de datos era introducir los cambios manualmente. Algunos países todavía no se sienten seguros con la corrección automática, de forma que utilizan la corrección manual en alguna de las etapas descritas. Si el conjunto de datos es pequeño, si la puntualidad no es una consideración decisiva o si se dispone de mucha mano de obra, la corrección manual será útil en muchos casos. Su ventaja es que si la información del cuestionario es completa y correcta, y si la discrepancia puede resolverse efectivamente examinando el impreso, probablemente habrá cierto mejoramiento de la calidad del censo o de la encuesta (el equipo de revisión tendrá que suponer, por ejemplo, que “María” no es “Mario”; que si hay datos sobre fecundidad era correcto obtenerlos para esa persona; y que los datos no se obtuvieron equivocadamente). En realidad, los procedimientos de revisión y atribución de datos raramente mejoran la calidad de la reunión de datos. Solamente modifican ciertos elementos.

116. A veces, examinar un cuestionario para efectuar correcciones manuales resulta infructuoso. Por el motivo que sea, la información no está consignada. Es posible que una persona no quiera divulgar su edad, de manera que el elemento correspondiente del cuestionario quedará en blanco. En tal caso, examinar el cuestionario no solucionará el problema. Entonces el equipo de revisión deberá adoptar una decisión sobre la forma de resolver la situación. En el caso de la corrección manual, la oficina nacional de estadística o del censo deberá asignar un valor “desconocido” o utilizar alguna serie de valores para atribuir un valor al elemento de edad.

117. A menos que se establezca contacto con el declarante, la corrección manual reduce inevitablemente la calidad y la coherencia porque exige más tiempo y es más costosa. Los ordenadores no se cansan, y son más rápidos; no tienen problemas personales que puedan afectar al mantenimiento de cierto nivel de calidad o de coherencia

y en la mayoría de los casos hacen que el procesamiento sea más económico. Actualmente la mayoría de los países utilizan algún tipo de corrección automática.

118. Las omisiones y las respuestas incongruentes reducen la calidad de los datos y hacen difícil presentar unos cuadros censales de fácil comprensión. Algunos usuarios prefieren tabular las omisiones y las respuestas incongruentes en la categoría de “no declarado”, mientras que otros prefieren distribuir esos casos proporcionalmente entre los elementos declarados que no presenten discrepancias. Hay otros que recomiendan normas para atribuir valores “probables” a las variables omitidas o a las respuestas incongruentes. La utilización de ordenadores permite atribuir de manera eficiente las respuestas sobre la base de otra información del cuestionario o de otra información comunicada relativa a una persona o a una unidad de habitación con características similares.

119. El proceso de revisión de datos puede aprovecharse de que el ordenador tiene capacidad para examinar muchas características. Así, los procedimientos de revisión de datos en los que intervienen muchas características interrelacionadas pueden tener como resultado la atribución de respuestas más razonables que las que podría producir una revisión simple. Por otra parte, una revisión mal diseñada puede producir datos censales deficientes. El equipo de revisión de datos debe estar integrado por especialistas temáticos experimentados en distintas disciplinas pertinentes, además de los especialistas en procesamiento de datos. Los miembros del equipo de revisión de datos deben seleccionar minuciosamente las variables cuya coherencia habrá que someter a prueba a fin de determinar las especificaciones de revisión y atribución de valores. Los productos de los programas deberían incluir los porcentajes de respuestas que fueron modificadas o a las que se atribuyeron valores. Así los analistas estarán en mejores condiciones para juzgar la calidad de los datos; por ejemplo, un elevado porcentaje de atribuciones alertaría de que hay que utilizar los datos con prudencia.

120. El rastro de verificación —es decir, la secuencia de revisiones— indica los cambios hechos en cada variable. Esta secuencia se utiliza para seguir el historial de las respuestas desde que se recibieron los datos y a todo lo largo del proceso de revisión y atribución.

C. Directrices para la corrección de datos

121. Tanto si se hace de manera manual como automáticamente, la revisión debe lograr que los datos sean tan representativos de la situación real como sea posible, mediante la eliminación de las omisiones y los asientos inválidos y la modificación de los valores incongruentes.

122. Puede tomarse como ejemplo el diagrama del gráfico 2, relativo a un hogar determinado. El diagrama presenta un hogar con información coherente para las variables de relación y de sexo. El jefe del hogar es de sexo masculino y en su caso no consta información sobre la fecundidad; el cónyuge es de sexo femenino y en su caso sí que se incluye la correspondiente información sobre la fecundidad.

123. Sin embargo, en muchos casos la información será incongruente. Se plantean, pues, las cuestiones siguientes: ¿Qué proceso de revisión debería aplicarse a un hogar con asientos incongruentes? ¿Cómo debería proceder el equipo de revisión si se indica que el jefe del hogar y el cónyuge son de sexo masculino, como sucede en el gráfico 3? Anteriormente, la norma de revisión típica habría consistido en suponer que la primera persona de una pareja es de sexo masculino, particularmente si esa persona es el jefe del hogar, y que la segunda persona, o el cónyuge, es de sexo femenino.

Recuadro 2

Directrices principales para la corrección de datos

Al establecer los procedimientos de revisión puede ser útil tener presentes las siguientes sugerencias para la corrección de datos:

- ☞ Hacer el menor número posible de cambios en los datos consignados originalmente;
- ☞ Eliminar las incongruencias obvias entre los asientos;
- ☞ Proporcionar valores para los asientos erróneos o los elementos omitidos utilizando como guía otras variables correspondientes a la unidad de habitación, a la persona o a otras personas del hogar o de un grupo comparable, siempre de conformidad con los procedimientos especificados. En algunas ocasiones la categoría “no declarado” será apropiada para determinados elementos.

Gráfico 2

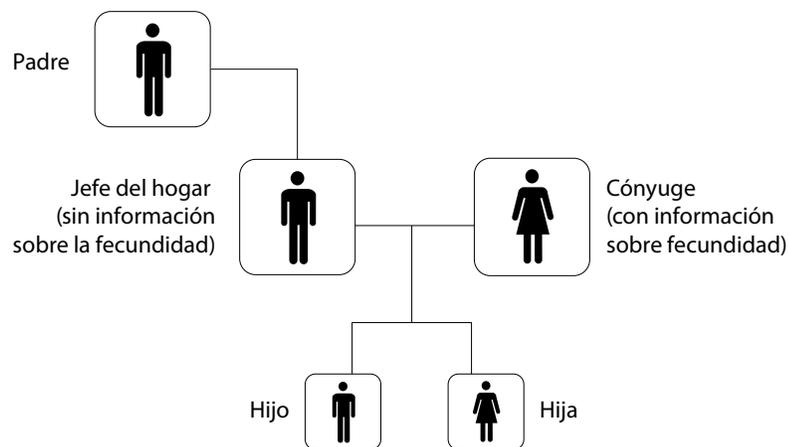
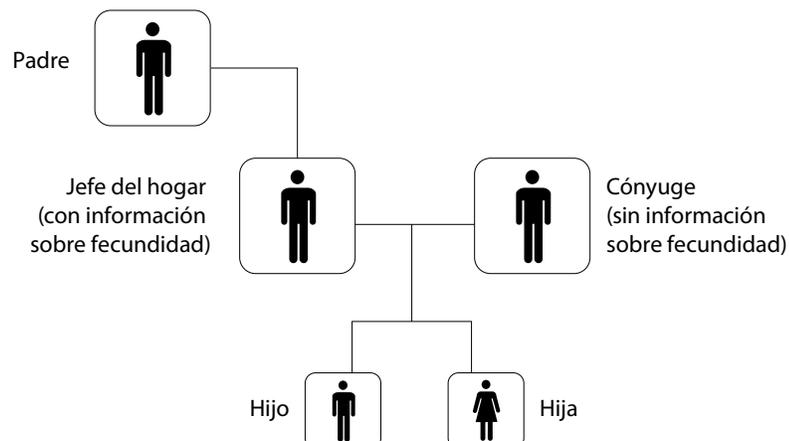
Un hogar hipotético típico, con información sobre las variables de relación, sexo y fecundidad de los miembros


Gráfico 3

Ejemplo de hogar en el que el jefe del hogar y el cónyuge son del mismo sexo


124. Si en este caso resultara que el jefe del hogar era la esposa y no el marido, la norma de revisión adoptada sería equivocada y la oficina nacional de estadística o del censo se encontraría con cuatro errores:

- a) El sexo del jefe del hogar estaría equivocado;
- b) El sexo del cónyuge estaría equivocado;
- c) El jefe del hogar, que es una mujer, perdería su información sobre la fecundidad;
- d) Se asignarían erróneamente datos de fecundidad al cónyuge, que es de sexo masculino.

Obviamente, este no es un buen procedimiento de revisión.

125. En contraste, cuando al aplicar un buen procedimiento de revisión se descubre que el jefe del hogar y el cónyuge son del mismo sexo, se comprobará la información sobre fecundidad de ambas personas. Puesto que solamente el jefe de familia tiene datos sobre fecundidad, se cambiará el valor de la variable de sexo. De este modo se cumplirán las normas de revisión relativas a esos elementos.

126. El gráfico 4 es otro ejemplo del mismo problema. En la mayoría de países se considera que la edad de procrear está entre los 15 y los 49 años. Puede suceder que una mujer declare haber tenido un hijo a los 52 años de edad, según se pone de manifiesto directamente en el número de línea correspondiente a la madre del niño o en la diferencia de edad calculada (la diferencia de edades entre madre e hijo biológico no puede ser de más de 50 años, aunque en el caso de los niños adoptados las diferencias de edad podrían ser mayores). El equipo de revisión deberá decidir si la diferencia de edad es aceptable o si hay que cambiarla, sustituyendo en la revisión una u otra de las edades. Si la modificación amplía el intervalo de años aceptable como edad de procrear y otras mujeres comunican haber tenido hijos a edades más avanzadas, pueden introducirse más anomalías en el conjunto de datos si la equivocación está en la edad que se ha comunicado. También en este caso el equipo de revisión decidirá si las edades declaradas son apropiadas para ciertas variables.

127. El gráfico 5 presenta otra situación posible. Puede suponerse que en la revisión se encuentra una mujer de 70 años de edad con hijos de 10 y 8 años, lo cual es posible porque el marido podría haber tenido los hijos con una esposa anterior. En tales circunstancias, los hijos están relacionados con el jefe del hogar, y no con

Gráfico 4

Ejemplo de hogar con las edades de algunos de sus miembros

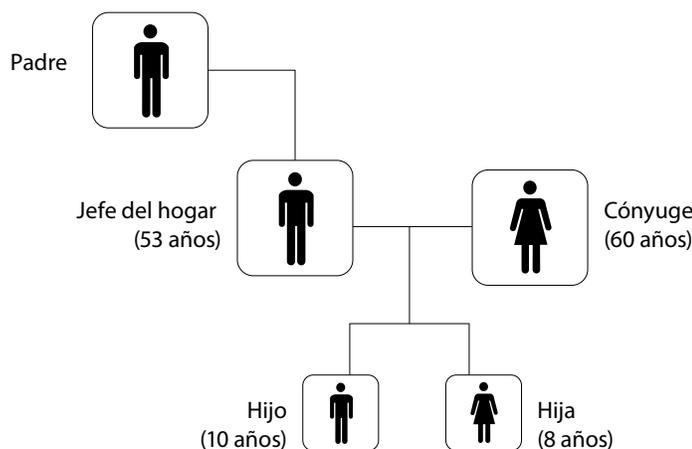
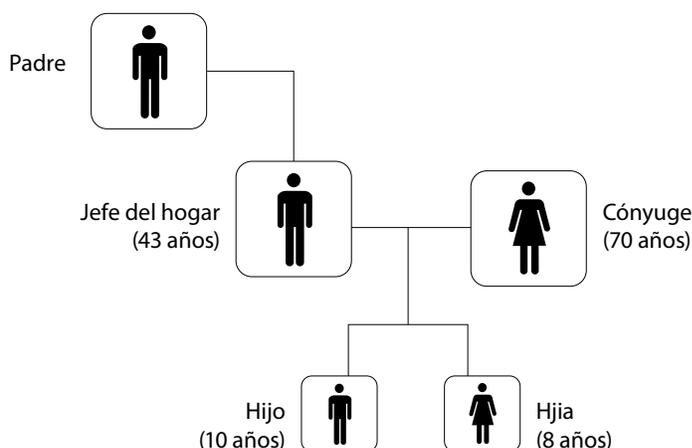


Gráfico 5
Ejemplo de hogar con posibles incongruencias en las edades declaradas



el cónyuge *per se*, aunque tal vez sea más probable que los encargados de introducir los datos hayan cometido un error consistente en escribir “7” cuando debían haber escrito “4” para indicar una edad de 40 años. Puede suponerse que, por el motivo que sea, los especialistas temáticos requieren que el encargado del procesamiento de datos modifique la edad de la madre o la del hijo cuando haya un intervalo de más de 50 años entre ellas. Este requisito exigirá otra revisión más complicada. Puesto que la mujer tiene 70 años y el primer hijo 10, el equipo de revisión debe decidir qué edad se debe modificar. Tal vez decida modificar la edad del primer hijo, asignándole 20 años, lo cual resolvería el problema para ese primer hijo, o puede cambiar la edad del cónyuge. En el primer caso, seguirá habiendo un problema con respecto a la edad del segundo hijo, que también requiere modificación.

128. Si se consideran únicamente las edades de la madre y de un hijo, la atribución asignaría al azar la edad, que sería correcta alrededor del 50% de las veces. Sin embargo, cuando en la revisión se examina también la edad del marido, lo más probable es que el equipo de revisión modifique la edad del cónyuge, sobre la base de esa información adicional. Ese único cambio haría más compatibles las edades de todos los miembros de la familia.

D. Comprobaciones de validez y coherencia

129. Uno de los requisitos principales de la revisión es que ningún elemento debe contener valores inválidos. Las respuestas a todos los elementos interrelacionados, por otra parte, deben presentar una coherencia dentro de cada registro y entre ellos. Entradas inválidas son las que resultan inaceptables por razones técnicas o estéticas. Por ejemplo, en el caso del género sólo están permitidos los códigos de hombre y mujer. Cualquier otro valor sería inaceptable y debería sustituirse por “desconocido” o por uno de los dos sexos aceptables; dado que la mayoría de los países realiza la planificación y la formulación de muchas variables de las políticas sobre la base del sexo, la existencia de valores desconocidos en el conjunto de datos complicaría los esfuerzos por obtener los valores necesarios para esa tarea. De la misma manera, las tabulaciones con incongruencias como “paredes de paja y techo de hormigón”, “mujeres de 13 años con 20 hijos”, “3 años de edad con título de doctorado” darían una pobre imagen de la oficina de estadística aun cuando los pocos casos de incongruencia no perjudicaran la planificación efectiva de un país.

130. Cuando se atribuyan valores habrá que tener en cuenta —en la mayor medida posible y no necesariamente de manera secuencial con respecto a variables conexas— toda la información sobre variables interrelacionadas al mismo tiempo. Sin embargo, en algunos casos en el proceso de revisión puede hacerse una comprobación de coherencia antes de determinar la validez de un asiento. Si se atribuye un valor como consecuencia de la comprobación de coherencia, hay que comparar el valor del asiento original para determinar si se trata efectivamente de un cambio. Si no lo es, el asiento original permanecerá tal como está.

1. Método secuencial de revisión de datos

131. En este procedimiento se comienza con el primer elemento que hay que revisar, que suele ser la primera variable del cuestionario, y se van examinando sucesivamente los elementos hasta concluir la revisión de todos ellos. El planteamiento habitual es tener primero en cuenta las tasas de respuesta y la importancia relativa de los diversos elementos. Las revisiones, normalmente, empiezan con el sexo y la edad, debido a su importancia, sobre todo en la atribución dinámica. Si bien el procedimiento secuencial no respeta las relaciones entre los elementos, aporta un marco adecuado a partir del cual se puede llevar a cabo la revisión.

132. El valor de un elemento puede modificarse más de una vez durante el proceso de revisión. Este procedimiento puede introducir uno o más errores en el conjunto de datos. Un valor atribuido puede ser incongruente con otros datos. Aun cuando los elementos se revisen secuencialmente, una determinada variable debería ser revisada teniendo en cuenta todas las demás, de ser posible. Por ejemplo, la edad de un hijo cuyo valor ha sido atribuido a partir de la edad de la madre puede ser incongruente con los años de escolarización del niño o con los años que este ha vivido en el distrito. En este caso habrá que atribuir un nuevo valor a la edad hasta que el asiento sea congruente. El valor atribuido a la edad será una variable intermedia hasta que se haya hecho una asignación definitiva. En los procedimientos de revisión las variables intermedias que se hayan atribuido no deben consignarse como cambios hasta que se haya hecho una asignación definitiva.

133. Si bien el programa de revisión podría aceptar un valor en blanco o “no declarado” respecto de algunos elementos y condiciones, los datos relacionados pueden suministrar información para la mayoría de los elementos que se han dejado en blanco o que contienen valores erróneos. Los valores suministrados de esta forma pueden ser correctos o no. No obstante, la amplia capacidad y la velocidad del ordenador para comparar los distintos valores almacenados permitirá determinar valores sustitutos que describan la situación de manera razonable. En la mayoría de los casos las tabulaciones resultantes serán a veces más congruentes que las hechas a partir de registros no depurados o de los registros en los que la atribución convierte todos los asientos inaceptables en valores “no declarados”.

134. El programa de revisión también deberá hacer comprobaciones estructurales (véase el capítulo III) de los elementos de población (véase el capítulo IV) y de los elementos de vivienda (véase el capítulo V). Además, los procedimientos de revisión de datos deberían incluir probablemente la creación en el registro de una o varias variables recodificadas necesarias para la tabulación, como se indica en el anexo 1.

135. Es sumamente importante evitar la revisión circular, que se produce cuando se modifican uno o varios elementos y más adelante se vuelven a modificar de forma que quedan como estaban antes. En otra parte del *Manual* se observa que el personal debe proceder a varias aplicaciones del programa para asegurarse de que se han revisado completamente todos los elementos. Es posible que se hayan fijado crite-

rios de revisión que modifiquen los datos durante una primera aplicación pero que al ser aplicados una segunda vez a los datos modificados vuelvan a cambiarlos de forma que los datos queden con su configuración original. Este procedimiento puede irse repitiendo en aplicaciones sucesivas. El equipo de revisión deberá evitar la adopción de esos criterios en el proceso de revisión.

2. Método de revisión de datos con variables múltiples

136. Es posible que el procedimiento secuencial de revisión de los datos de censos y encuestas presentado en la sección 1 *supra* no siempre produzca los mejores resultados; es decir, los que más se aproximan a la distribución real de las variables. Como se ha señalado, si el procedimiento secuencial se aplica sin tomar las precauciones apropiadas, suele provocar problemas en la revisión.

137. Otro procedimiento es el de la revisión con variables múltiples, que se basa en el sistema de Fellegi-Holt. Este procedimiento requiere más conocimientos informáticos y ordenadores con mayor capacidad, pero permite obtener resultados que probablemente se aproximarán más a la “realidad”. En el anexo V (“Métodos de atribución de valores”) figuran diferentes tipos de revisión con variables múltiples. Para aplicar este método es necesario establecer un conjunto de instrucciones o afirmaciones positivas para comprobar la relación entre las variables. A continuación se comprueba cada una de las afirmaciones comparándola con los datos del hogar para determinar si todas las afirmaciones son ciertas. Respecto de cada afirmación falsa el programa de revisión efectuará un seguimiento, caso por caso, de los asientos inválidos o de las incongruencias. Una vez hechas todas las pruebas, el sistema de revisión y atribución de valores evaluará la mejor forma de modificar el registro para que supere todos los criterios de revisión. Los equipos de revisión suelen aplicar un enfoque de cambios mínimos y modifican el menor número posible de variables para obtener un registro aceptable.

138. Las 11 instrucciones que figuran en el gráfico 6 brindan un ejemplo de las normas que podrían aplicarse en un procedimiento de revisión con variables múltiples de determinadas características de la población. En el ejemplo, el jefe del hogar debe tener 15 años de edad o más. En los programas genéricos de revisión sería mejor utilizar “X” años, siendo X el mínimo fijado para el país. Las instrucciones del ejemplo se concentran en otras variables primarias importantes como la relación, el sexo, la edad, el estado civil y la fecundidad. Las variables están estrechamente relacionadas, por lo cual los equipos de revisión deberían examinarlas conjuntamente para revisar los datos con la máxima eficiencia. Cabe señalar que aunque todas las variables sean importantes, algunas son más decisivas que otras para la presentación de los datos.

139. En el gráfico 6 puede verse un caso sencillo en el que por alguna razón tanto el jefe del hogar como el cónyuge tienen el mismo sexo: ambos aparecen como varones y uno es un varón con información sobre la fecundidad. Es evidente que el sexo está equivocado (como se observa en el resumen final) y que el varón con información sobre la fecundidad debe aparecer como mujer.

140. En el ejemplo del gráfico 7 ambos cónyuges pertenecen a la misma población que los del gráfico 6. Se ha declarado que ambos son de sexo masculino. En este caso el procedimiento de revisión es simple y directo. La variable en la que se haya observado el mayor número de errores es la que se revisará en primer lugar. En el gráfico 7 el programa de revisión ejecuta el procedimiento de atribución de un valor a la variable de “sexo” porque, sobre la base de los datos del gráfico 6, esa variable es la que más errores presenta con respecto a la relación y el sexo y respecto a la fecundidad

Gráfico 6
Ejemplo de normas para una revisión con variables múltiples de determinadas características de población

No.	Norma	Relación	Sexo	Edad	Estado civil	Fecundidad
1	El jefe del hogar debe tener 15 años de edad o más					
2	El cónyuge debe tener 15 años de edad o más					
3	Un cónyuge debe estar casado					
4	Si hay un cónyuge, el jefe de hogar debe estar casado					
5	Si hay un cónyuge, el sexo del jefe del hogar y el del cónyuge deben ser distintos	1	1			
6	Una persona menor de 15 años de edad no debe estar casada					
7	No debe haber datos sobre fecundidad para una persona de sexo masculino		1			1
8	No debe haber datos sobre fecundidad para una persona de sexo femenino menor de 15 años de edad					
9	La variable de fecundidad no debe estar en blanco para una persona de sexo femenino de 15 años de edad o más					
10	Un hijo debe ser más joven que el jefe del hogar					
11	Un progenitor debe tener más años que el jefe del hogar					
	Totales	1	2			1

Nota: "1" indica que existe incongruencia entre dos o más elementos. Por ejemplo, en la norma 5, dado que el jefe del hogar y el cónyuge son del mismo sexo, se encuentran problemas en la revisión de la relación y el sexo; por esa razón aparece un 1 en esas casillas.

y el sexo. Cuando el programa de revisión comprueba la fecundidad y descubre que hay información sobre fecundidad para el jefe del hogar pero no para el cónyuge, se atribuye el valor "sexo femenino" al jefe del hogar. Por último, cuando el equipo de revisión vuelva a comprobar la serie de afirmaciones y todas ellas sean ciertas, cualquier otra revisión sería inútil.

141. Las especificaciones para ese tipo de revisión pueden redactarse como se indica en el gráfico 8. Si la información sobre la fecundidad es completa en ambos casos, la revisión funcionará. No obstante, es obvio que la revisión no está completa ya que sólo tiene en cuenta el caso en que la información sobre la fecundidad está completa y es precisa tanto en lo que respecta al jefe del hogar como al cónyuge.

Gráfico 7
Ejemplo de un jefe de hogar y un cónyuge del mismo sexo en un conjunto de datos sin depurar y su solución

Persona	Relación	Sexo	Hijos nacidos
Datos sin depurar			
1	Jefe del hogar	Masculino	03
2	Cónyuge	Masculino	EN BLANCO
Datos después de revisar el sexo			
1	Jefe del hogar	Femenino	03
2	Cónyuge	Masculino	EN BLANCO

Gráfico 8

Ejemplo de especificaciones de revisión para corregir la variable relativa al sexo, redactadas en pseudocódigo

Si el SEXO del JEFE DEL HOGAR = SEXO del CÓNYUGE
 Si la FECUNDIDAD del JEFE DEL HOGAR no está en blanco
 Si la FECUNDIDAD del CÓNYUGE está en blanco
 (si el SEXO del jefe del hogar no consta ya como femenino) Cambiar SEXO = femenino
 (si el SEXO del cónyuge no consta ya como masculino) Cambiar SEXO = masculino
 De no ser así *¡¡Tomar medidas porque ambos tienen el mismo sexo y ambos tienen información sobre fecundidad!!*
 [Las "medidas" podrían consistir en utilizar el sexo del jefe anterior, o alternar el sexo del jefe, o utilizar coeficientes de sexos de todos los jefes para una respuesta adecuada, etcétera]
 Terminar - si
 De no ser así *¡¡Éste es el caso cuando el valor de fecundidad del jefe del hogar está en blanco!!*
 Si la FECUNDIDAD del CÓNYUGE no está en blanco
 (si el SEXO del jefe del hogar no consta ya como masculino) Cambiar SEXO = masculino
 (si el SEXO del cónyuge no consta ya como femenino) Cambiar SEXO = femenino
 De no ser así *¡¡Tomar medidas porque NINGUNO de ellos tiene información sobre fecundidad!!*
 [Las "medidas" podrían consistir en utilizar el sexo del jefe anterior, o alternar el sexo del jefe, o utilizar coeficientes de sexos de todos los jefes para una respuesta adecuada, etcétera]
 Terminar - si
 Terminar - si
 Terminar - si

142. En el gráfico 9 se presenta un ejemplo en el cual un procedimiento de revisión examina a un jefe del hogar de sexo femenino de 13 años de edad que consta como viuda pero que tiene tres hijos, según la información introducida. Cuando el programa aplica las normas de revisión, el resultado es el siguiente:

Gráfico 9

Ejemplo de análisis de una revisión con variables múltiples en el caso de una viuda muy joven con tres hijos

No.	Norma	Relación	Sexo	Edad	Estado civil	Fecundidad
1	El jefe del hogar debe tener 15 años de edad o más	1		1		
2	El cónyuge debe tener 15 años de edad o más					
3	Un "cónyuge" debe estar casado					
4	Si hay un cónyuge, el jefe de hogar debe estar casado					
5	Si hay un cónyuge, el sexo del jefe del hogar y el del cónyuge deben ser distintos					
6	Una persona menor de 15 años de edad no debe estar casada			1	1	
7	No debe haber datos sobre fecundidad para una persona de sexo masculino					
8	No debe haber datos sobre fecundidad para una persona de sexo femenino menor de 15 años de edad		1	1		1
9	La variable de fecundidad no debe estar en blanco para una persona de sexo femenino de 15 años de edad o más					
10	Un "hijo" debe ser más joven que el jefe del hogar					
11	Un "progenitor" debe tener más años que el jefe del hogar					
Totales		1	1	3	1	1

143. Examinamos también aquí el caso de una viuda de 13 años de edad que es jefe del hogar y tiene tres hijos. La primera revisión, descrita en la norma 1 —un jefe de hogar de 15 años de edad o más— suscita problemas debido a que el jefe del hogar tiene menos de 15 años. Como tiene 13 años de edad, los recuadros correspondientes a “relación” y “edad” están marcados, ya que hay una incongruencia entre esas dos variables. Ella no es cónyuge, por lo que no se aplican las normas 2 y 3. Las normas 4 y 5 tampoco se aplican por la misma razón: se refieren únicamente al cónyuge. No obstante, dado que, en virtud de la norma 6, una persona de menos de 15 años de edad (en este caso, 13 años) no debería estar casada y nuestra muchacha de 13 años está viuda, se incumple la norma. En la medida en que la norma 7 se refiere a los varones, no se aplica. En virtud de la norma 8, las mujeres de menos de 15 años de edad no deben tener información relativa a la fecundidad; en la medida en que esta persona tiene información relativa a la fecundidad, se incumple la norma. Las normas 9, 10 y 11 no se aplican a esta persona.

144. Con arreglo a la serie de afirmaciones, la variable correspondiente a la edad es la que presenta más errores y la que habrá que modificar primero. Una vez cambiado el valor correspondiente a la edad, se realizan de nuevo las pruebas y habrá concluido la revisión si con ese cambio se resuelven todas las incongruencias. En otro caso, el programa revisará la variable siguiente que presente el siguiente mayor número de incongruencias.

E. Métodos de corrección y atribución de datos

145. Como se ha mencionado, en los registros de datos de todos los censos y encuestas hay lagunas debidas a información “no declarada”, “desconocida” u omitida por cualquier otro motivo. También hay asientos inválidos a causa de equivocaciones de los declarantes, los empadronadores o los encargados de introducir los datos. Los métodos de corrección varían según el elemento de que se trate. En la mayoría de casos es posible asignar a los elementos códigos válidos, con una garantía razonable de que serán correctos, utilizando respuestas correspondientes a otros elementos de la misma persona o del mismo hogar o de los registros de otras personas u hogares.

146. En el presente *Manual* se describen dos técnicas por computadora para corregir los datos inaceptables. Una de ellas es el método de atribución estática, que se utiliza principalmente para las omisiones o los valores desconocidos. La otra es el método de atribución dinámica, que puede utilizarse tanto para las omisiones como para los datos incongruentes o inválidos. Los diferentes programas informáticos, que utilizan diversas metodologías, emplean la atribución estática y la atribución dinámica de distintas maneras, como se ilustra en los anexos.

1. La técnica de atribución estática

147. En el procedimiento de atribución estática el programa de revisión asigna a un elemento omitido una respuesta particular tomada de un conjunto predeterminado, o bien atribuye un valor proporcional basado en la distribución de respuestas válidas. En el método de atribución estática el programa no actualiza el conjunto de variables original. Los valores de la matriz estática inicial no cambian una vez procesados los registros de la primera, la segunda, la décima o cualquier otra persona, sino que los valores originales proporcionan el valor que se atribuirá a cualquier dato omitido.

148. La atribución estática es un método estocástico, al igual que la atribución dinámica, pero los valores no cambian en el tiempo. Este procedimiento se describe en el anexo V.

149. En la atribución estática se aplica a veces un método proporcional, con arreglo al cual las respuestas se asignan sobre la base de porcentajes predeterminados. Un ejemplo de la distribución proporcional de respuestas sería una tabulación de datos válidos —es decir, datos procedentes de elementos que constan debidamente, y no de omisiones— sobre el tiempo semanal de trabajo de hombres de 33 años de edad empleados en la agricultura, la cual indicase que el 25% trabajó 50 horas por semana, el 40% trabajó 60 horas por semana y el 35% trabajó 70 horas por semana. Las respuestas en blanco o inválidas relativas al elemento del tiempo semanal de trabajo de los hombres de 33 años de edad empleados en la agricultura serían sustituidas en el 25% de los casos por 50 horas; en el 40% de los casos, por 60 horas, y en el 35% de los casos, por 70 horas. Sin embargo, a menos que se disponga de datos fidedignos de censos, encuestas u otras fuentes anteriores, esta técnica requiere una tabulación previa de las respuestas válidas del censo en curso, lo cual quizá no sea viable por motivos económicos u operacionales.

2. La técnica de atribución dinámica

150. Otro método para eliminar los valores desconocidos de los datos es la técnica de atribución dinámica, que se utiliza para asignar valores a asientos en blanco, desconocidos, incorrectos o incongruentes. Este método fue elaborado originalmente por la Oficina del Censo de los Estados Unidos, pero después lo han ido perfeccionando otras instituciones. La atribución dinámica utiliza una o más variables para estimar una respuesta probable cuando aparece un elemento desconocido (o, en ciertas circunstancias, varios elementos desconocidos) en el conjunto de datos. La atribución dinámica es cada vez más habitual en las revisiones de los censos porque es sencilla y produce resultados claros y fáciles de reproducir. Además, al eliminarse los valores desconocidos es más fácil obtener las tendencias entre censos y encuestas, ya que el analista no tiene que ocuparse de los elementos desconocidos caso por caso.

151. En la técnica de atribución dinámica los datos conocidos acerca de personas con características similares determinan la información más apropiada que se utilizará cuando se desconozca alguna información sobre otra persona. Esas características comprenden el sexo, la edad, la relación con el jefe del hogar, la situación económica y la educación. La propia matriz de atribución es un conjunto de valores parecido al de los naipes de una baraja. Esas matrices almacenan la información, que se utiliza cuando se encuentran valores desconocidos. La baraja cambia constantemente al actualizarse y/o “barajarse” lógicamente los naipes, de manera que los valores atribuidos van cambiando a medida que se van procesando los datos; de ahí el calificativo de atribución “dinámica”.

152. Los valores almacenados representan información sobre los “vecinos más próximos” con información semejante. Téngase en cuenta que el vecino más próximo suele ser el vecino *previo* más próximo ya que, sobre todo en el enfoque secuencial antes descrito, las unidades de habitación y las personas de esas unidades se consideran sólo una vez, y luego el programa continúa. Por ello, dentro de una aldea, cuando se desconoce si una persona es huérfana de madre, por ejemplo, el conjunto de valores contiene información sobre la persona más reciente encontrada con el mismo sexo y edad y una información válida sobre la orfandad de madre. Este enfoque es particularmente importante en los países que tienen movimientos migratorios relativamente considerables o una gran difusión del VIH/SIDA u otros fenómenos que dan lugar a una actividad estadística desacostumbrada. De la misma manera, es más probable que las características de vivienda sean semejantes dentro de una aldea o conjunto de aldeas que entre una aldea o conjunto de aldeas y otras partes del país.

153. Un ejemplo sencillo consistiría en almacenar un único valor. Por ejemplo, si el valor atribuido a la variable de sexo de una persona es inválido por algún motivo, se le asignará arbitrariamente un valor inicial (masculino o femenino), que se acepta como tal. Este valor se atribuye a la primera persona cuya variable de sexo tenga un valor desconocido. No obstante, si en el caso de la primera persona este asiento es válido, el sexo de la primera persona sustituirá el valor inicial. Si el sexo de la segunda persona es desconocido, la matriz de atribución asignará el valor almacenado para la variable de sexo. En este caso el valor que se atribuye es el sexo de la primera persona. En definitiva, cuando en el proceso de revisión se encuentra un valor aceptable para un elemento, este se incluye en la matriz de atribución. Cuando se encuentra un valor inaceptable, la atribución lo sustituye por el valor aceptable de la matriz de atribución.

154. Uno de los problemas que presenta el método de atribución dinámica descrito es que si dos elementos distintos tienen valores desconocidos, tal vez no pueda usarse el mismo “donante” para asignar respuestas válidas. Es posible que cada valor provenga de una persona “real”, pero puede tratarse de personas diferentes. Sería mejor asignar ambas variables al mismo tiempo, tomadas de la misma persona. Sin embargo, la programación de esas matrices más complejas puede presentar dificultades.

155. Los datos que figuran en el gráfico 10 se refieren a un hogar integrado por un conjunto de 10 personas. Las informaciones en blanco —señaladas con “X” y “XX”— indican la ausencia de datos. Muchas veces se utilizan los números 9 y 99 para indicar la ausencia de información; en este caso sobre el sexo (un 9 para un solo dígito) y sobre la edad (99 para dos dígitos). No obstante, en algunos casos se necesita el valor 9 para otro valor real, por ejemplo, en un número limitado de códigos de relación; por ello estos valores deben utilizarse con gran moderación; de hecho, si se puede utilizar otro valor, como “X”, “.” o “..”, esa sería probablemente la mejor opción. Téngase en cuenta que hay otras variables que pueden utilizarse en la atribución —como la educación y la ocupación— que no se han incluido en este breve ejemplo.

156. Si el valor inicial de la matriz de atribución que se denomina MATRIZ DE SEXO es masculino (código = 1), la matriz de atribución tendrá una indicación semejante a la siguiente: SEXO = 1.

157. Una vez concluido el procesamiento de los datos de la persona 1, el valor seguirá siendo 1. No obstante, el valor se transformará en 2 después de la revisión de la segunda persona, ya que esa persona es de sexo femenino. Entonces la variable tendrá una indicación similar a la siguiente: MATRIZ DE SEXO = 2.

Gráfico 10

Modelo de hogar como ejemplo de valores utilizados para la atribución dinámica

Número de identificación	Relación	Sexo	Edad
1	1	1	39
2	2	2	35
3	3	1	13
4	3	X	10
5	4	2	40
6	4	1	XX
7	4	2	13
8	5	X	XX
9	5	1	44
10	5	2	36

Nota: X, XX = información omitida.

158. En cada asiento válido para la variable de sexo de una persona revisada el código correspondiente al sexo de esa persona sustituye el valor de la matriz de atribución. Cuando se revisa la tercera persona, el valor de atribución vuelve a ser 1, es decir, masculino.

159. En el caso de la cuarta persona se desconoce el valor de la variable de sexo, de forma que en la revisión se acude al valor de la matriz de atribución, que en este caso es masculino, y se sustituye el valor desconocido por el valor de la matriz de atribución. La persona 5 es de sexo femenino, de forma que ese valor sustituye al valor anterior de la matriz de atribución de la persona 3 (sexo masculino). Este proceso continúa hasta la persona 8.

160. El proceso de revisión recurre de nuevo a la atribución, y a la persona 8 se le asigna el valor de sexo femenino, ya que el valor de la matriz de atribución obtenido de la persona 7 es femenino. En el procedimiento de revisión se ha utilizado dos veces la matriz de atribución para obtener valores: una vez para obtener un valor de sexo masculino y otra para obtener uno de sexo femenino. Dado que los sexos aparecen aproximadamente con la misma frecuencia, a largo plazo se atribuirá cada valor de sexo aproximadamente el 50% de las veces. Una vez revisada la totalidad de las 10 personas, la variable será: MATRIZ DE SEXO = 2.

161. Aunque una matriz de atribución puede asignar un valor a la variable de sexo de esta manera, hay otras formas más complejas de aplicar el procedimiento. Por ejemplo, el programa de revisión puede utilizar las variables de relación con el jefe del hogar y de sexo para ayudar a determinar la edad de una persona. Obsérvese la siguiente lista parcial de códigos de relación:

- 1 = Jefe del hogar
- 2 = Cónyuge
- 3 = Hijo
- 4 = Otro familiar
- 5 = Persona no emparentada

162. El encargado del procesamiento de datos puede establecer para la variable de edad valores iniciales que se aproximen a la situación real basándose en la relación y el sexo. Esos valores no son muy importantes, ya que es casi seguro que en la revisión serán sustituidos antes de ser utilizados. Además, el proceso de revisión obliga a atribuir muchos valores, de forma que pocos valores iniciales afectarán a las tabulaciones finales. Esos valores podrían ser los que se presentan en el gráfico 11.

Gráfico 11

Matriz estática inicial para la variable de edad, basada en el sexo y en las relaciones

	Relaciones				
	Jefe de hogar	Cónyuge	Hijo/hija	Otro familiar	Persona no emparentada
	(1)	(2)	(3)	(4)	(5)
Masculino (1)	35	35	12	40	40
Femenino (2)	32	32	12	37	37

163. A continuación pueden volver a examinarse las 10 personas presentadas en el gráfico 10. Puesto que la primera persona de nuestra muestra consta como jefe del hogar (código = 1) y es de sexo masculino (código = 1), su edad (39) sustituye el primer elemento (coordenadas 1,1) durante la atribución. Entonces la matriz contendrá los valores que se presentan en el gráfico 12.

Gráfico 12

Ejemplo de matriz de atribución dinámica después de un cambio

	Relaciones				Persona no emparentada
	Jefe de hogar	Cónyuge	Hijo/hija	Otro familiar	
	(1)	(2)	(3)	(4)	(5)
Masculino (1)	39	35	12	40	40
Femenino (2)	32	32	12	37	37

164. La segunda persona es un cónyuge (código = 2) de sexo femenino (código = 2), de manera que su edad (35) sustituye el valor del segundo renglón de la segunda columna, por lo que se modifican esos valores en la matriz. Las edades de otras personas del hogar sustituyen de manera similar los valores de la matriz de atribución, hasta la quinta persona.

165. Cabe observar que en el anterior procedimiento de atribución de la variable de sexo se asignó el valor de sexo 1 a la persona 4. Dado que las normas de revisión exigen atribuir un valor a la variable de sexo, la matriz no se actualiza con la edad de esa persona. La actualización se efectuará sólo con valores tomados de registros en los que tanto el sexo como la relación tengan inicialmente asientos correctos. Sin embargo, cuando en la revisión se llega a la persona 6, se observa que la edad es desconocida. La persona es de sexo masculino y es “otro familiar” del jefe del hogar. Por lo tanto, se utiliza el elemento de la matriz de atribución correspondiente a las personas de sexo masculino cuyo grupo de relación sea “otro familiar” (la cuarta columna del primer renglón) y se asigna el valor de edad correspondiente a esa categoría (“sexo masculino, otro familiar”, que en este caso es 40).

166. No constan ni el sexo ni la edad de la octava persona. En la revisión se atribuye el valor de sexo femenino y a continuación se asigna la edad sobre la base de este valor de sexo y del código de relación (5). En este caso, la edad es 37.

167. Aunque en la revisión se atribuyó un valor a la edad a partir de la relación conocida, se utilizó un valor de sexo previamente asignado para las demás variables. En este caso la utilización de valores asignados para nuevas atribuciones es un ejemplo de un procedimiento de revisión deficiente (véase el apartado *d*) de la sección 3 *infra*). Sería mejor buscar otros elementos con valores conocidos, como el estado civil, para aplicarlos en la atribución.

168. Después de la décima persona, los valores de la matriz de atribución son los que se indican en el gráfico 13. En este ejemplo, para ambas atribuciones se utilizó la matriz estática inicial. Normalmente en la atribución sólo se utilizarán unos pocos valores iniciales, o ninguno. En la mayoría de los casos se asignarán valores tomados de la población empadronada.

Gráfico 13

Ejemplo de una matriz de atribución dinámica después de múltiples cambios

	Relaciones				Persona no emparentada
	Jefe de hogar	Cónyuge	Hijo/hija	Otro familiar	
	(1)	(2)	(3)	(4)	(5)
Masculino (1)	39	35	13	40	44
Femenino (2)	32	35	12	13	36

3. Consideraciones relativas a la atribución dinámica

a) Consideraciones geográficas

169. Si el programa de revisión aplica técnicas de atribución dinámica para asignar los valores no consignados, debería procurar utilizar datos clasificados con arreglo a la zona geográfica más pequeña que se haya definido. Este procedimiento aumentaría la probabilidad de obtener una respuesta correcta, ya que las personas que viven en la misma zona geográfica pequeña suelen ser bastante homogéneas con respecto a sus características demográficas y de otro tipo. Si la población no es homogénea, no habrá correlación, de manera que el equipo de revisión deberá examinar las variables caso por caso. Asimismo, como se examinará más adelante, hay variables que no son aplicables en determinados ámbitos —como la calefacción central en los lugares muy calurosos—, lo que debe tenerse en cuenta en la revisión.

b) Utilización de elementos relacionados

170. Antes de recurrir a la atribución dinámica para obtener valores de variables desconocidas habría que procurar utilizar elementos relacionados para asignar un valor que tenga probabilidades de ser correcto. Por ejemplo, si falta el estado civil de una persona, el programa de revisión determinará si la persona tiene un cónyuge en el hogar. De ser así, el programa asignará el código correspondiente al valor de persona casada, sin utilizar la matriz de atribución. Sin embargo, cuando no exista ninguna prueba de este tipo el programa tendrá que recurrir a un valor de la matriz de atribución.

c) Efectos en las matrices del orden de las variables

171. Las oficinas nacionales de estadística o del censo que utilizan matrices de atribución deberían considerar qué variables necesitan al fijar el orden de sus revisiones. En el caso de los elementos de población, las oficinas procurarán revisar las variables de sexo y edad al comienzo, de forma que puedan utilizarlas en las otras matrices de atribución. De manera general, no deben incluirse variables no depuradas en las matrices de atribución, aunque la mayoría de los programas informáticos aceptarán valores desconocidos en determinados renglones o columnas. Las tasas de respuesta y la distribución de los atributos dentro de las variables ayudarán a determinar las variables óptimas, y los atributos más útiles dentro de esas variables, como ayuda para la atribución dinámica. Las matrices de atribución subsiguientes podrán utilizar los elementos una vez revisados. Sin embargo, siempre que sea posible, las oficinas de estadística deberían considerar la posibilidad de excluir de la matriz de atribución los datos revisados.

172. Si, por ejemplo, en el proceso de revisión se atribuye el valor de edad sobre la base del sexo y la relación, las casillas de esta matriz de atribución (sexo por relación) no deberán actualizarse si el sexo o la relación son valores atribuidos. Por regla general, sólo cuando la edad, el sexo y la relación tengan valores válidos y coherentes se podrá introducir la edad en la casilla correspondiente a los valores apropiados de sexo y relación. Sin embargo, a veces es inevitable utilizar datos depurados debido a otros factores. Es importante señalar que la mayoría de los países no tienen en cuenta esa sugerencia y realizan la atribución partiendo de valores previamente atribuidos. Una posible solución sería utilizar señales de alerta de atribución de valores para desalentar el uso de datos atribuidos para emparejar un donante con una unidad donde exista un error.

d) Complejidad de las matrices de atribución

173. La oficina nacional de estadística o del censo podrá incrementar la probabilidad de obtener sistemáticamente un valor “correcto” de la matriz de atribución

si hace que esta sea más detallada. Por ejemplo, el programa puede atribuir un valor de estado civil utilizando únicamente la relación. Sin embargo, la probabilidad de los casos de viudez o de divorcio aumenta con la edad. Por lo tanto, es lógico atribuir el valor de estado civil sobre la base de la edad y la relación. Utilizando la edad y la relación de la persona de que se trate, el programa de revisión tomará el valor de estado civil de una persona con las mismas características que conste en el registro válido inmediatamente precedente de los que están almacenados en la matriz de atribución.

174. No obstante, el procedimiento descrito puede crear nuevos problemas. La oficina nacional de estadística o del censo suele revisar los elementos de un cuestionario siguiendo un orden determinado en el que, aplicando un criterio secuencial, la edad se examina después del estado civil. De ser así, cuando en un registro falten tanto el estado civil como la edad será imposible tomar el valor de estado civil del registro inmediatamente precedente que tenga los mismos valores de edad y relación⁶. Por tanto, es posible que el programa no pueda determinar la categoría de edad correspondiente a ese registro. Otra solución sería que en la matriz de atribución hubiera un renglón o una columna para los elementos “no declarados”. Este procedimiento permitiría al programa asignar un valor al estado civil utilizando la categoría de estado civil del registro inmediatamente precedente que tenga la misma relación y un valor de edad “no declarado”. Sin embargo, hay dos factores que desaconsejan este procedimiento. Uno es que los casos “no declarados” en una misma combinación son tan pocos que resultaría difícil actualizar la matriz de atribución respecto del elemento que falta. En segundo lugar, es esencialmente imposible obtener valores apropiados de atribución estática para esas combinaciones de valores “desconocidos”, ya que no existen en el mundo “real”.

175. Dar solución al problema descrito aumenta el trabajo del encargado del procesamiento de datos, pero el resultado es un producto más depurado. En primer lugar, el programa de revisión efectúa una prueba para determinar si los elementos tienen códigos válidos. Si en el registro de la persona que se examina no hay un código válido para ese elemento, la matriz de atribución no utilizará el elemento para ese registro. Los encargados del procesamiento de datos pueden facilitar el proceso creando una matriz de atribución más simple. Para proseguir con el ejemplo anterior, si el programa debe atribuir un valor de estado civil que falta, la matriz de atribución tendrá normalmente dos dimensiones: la edad y la relación. Si una vez realizadas las pruebas el programa no encuentra ningún código válido para la edad, atribuirá un valor al estado civil basándose únicamente en la relación. Dado que la revisión de la relación se efectúa antes que la del estado civil, el código de relación será válido. El programa utiliza esos mismos principios para todos los procedimientos de atribución dinámica.

e) Elaboración de matrices de atribución

176. Los especialistas temáticos, en colaboración con los especialistas en procesamiento de datos, deberían preparar matrices de atribución apropiadas. (Algunos equipos de revisión de datos utilizan múltiples matrices de atribución.) Sólo se utilizan respuestas válidas para actualizar las matrices de atribución; los equipos de revisión no utilizan valores asignados ni atribuidos. Tanto los especialistas temáticos como los especialistas en procesamiento de datos deben comprobar las especificaciones de revisión y las series de valores de las matrices dinámicas, para verificar que son coherentes y completas.

177. La formulación de una matriz de atribución exige mucho tiempo y reflexión y obliga a investigar la posibilidad de utilizar registros administrativos y los resultados de censos o encuestas anteriores, en particular en relación con los valores

⁶ La práctica de revisión más recomendable no implica el uso de valores revisados en la atribución dinámica. Algunas veces esta práctica es difícil de aplicar, debido a problemas cronológicos relacionados con los resultados o a dificultades en la programación informática. En tales casos se atribuiría una de las diversas variables, y su valor se colocaría en las atribuciones dinámicas adecuadas y luego se utilizaría para atribuir las variables posteriores.

de la atribución dinámica. Incluso después de la labor de investigación y desarrollo, los encargados de la revisión no deben aplicar al azar las matrices de atribución. Cuando estas carecen de coherencia interna se necesita un esfuerzo considerable para reconciliarlas. Cuando las matrices de atribución no aplican las convenciones ordinarias, el personal deberá examinar cada una por separado.

178. Aunque en los ejemplos del presente *Manual* cada una de las casillas de las matrices de atribución tiene un solo valor, algunos equipos de revisión mantienen más de una posibilidad para cada casilla. Sería como una especie de matriz en dos dimensiones, con una tercera. Esas casillas ofrecen una dimensión adicional. Por ejemplo, si se desconocen las edades de todos los hijos de una familia en la que hay cuatro hijos varones, el ordenador no asignará el mismo valor cuatro veces —con lo que los hijos se convertirían en cuadrillizos—, sino cuatro edades diferentes. Sin embargo, incluso así podría asignarse el mismo valor más de una vez, según los valores que haya almacenados en las matrices.

f) Matrices de atribución estándar

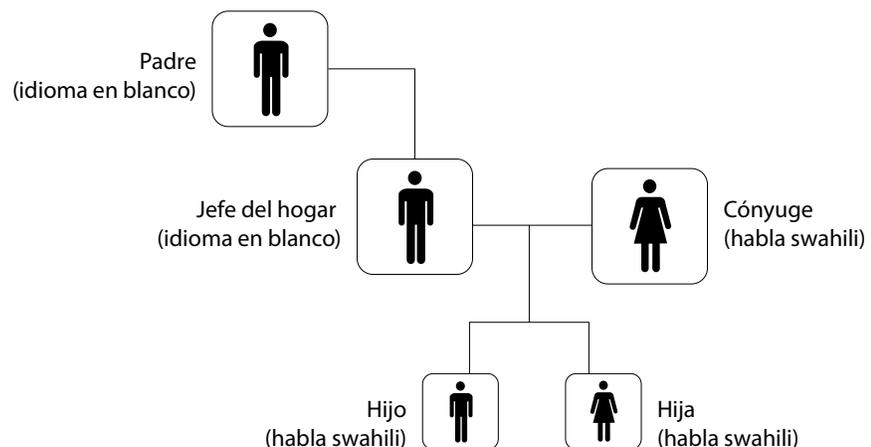
179. Las matrices de atribución estándar pueden agilizar el proceso de revisión. Es posible ensayar y aplicar rápidamente matrices de atribución con dimensiones estándar para diversas variables sociales y económicas, como los grupos de edad y el sexo.

180. Es posible que la oficina nacional de estadística o del censo desee elaborar una matriz de atribución para determinar un código de idioma cuando no conste ese valor. Lo primero que con toda seguridad buscará el programa de revisión será otra persona del propio hogar que comunique que habla un determinado idioma. Si no se da el caso, el programa puede seleccionar el idioma de una persona anterior del mismo sexo y grupo de edad (habiendo actualizado la matriz de atribución cuando los tres elementos fueran válidos). El procedimiento proporcionará un valor de idioma probable, ya que las personas que hablan el mismo idioma o idiomas similares suelen estar o vivir muy cerca las unas de las otras.

181. En el gráfico 14 la variable correspondiente al “idioma” no contiene información en el caso del jefe del hogar. Por el motivo que sea, la anotación o el código correspondientes al idioma tal vez no han sido captados por el lector óptico (escáner) o por la persona encargada de introducir los datos, o puede haberse producido cualquier

Gráfico 14

Ejemplo de un jefe del hogar y de su progenitor a los que no se ha asignado un valor de idioma



otro error. Sin embargo, puesto que el cónyuge y los hijos hablan swahili, ese idioma puede asignarse al jefe del hogar y al progenitor del jefe del hogar, cuyo valor de idioma también se ha omitido. Téngase en cuenta que el jefe de hogar, en el gráfico 14, es del sexo masculino.

182. Cuando no conste el idioma de ninguna persona del hogar, el programa de revisión deberá seguir otro procedimiento. En primer lugar, buscará otras variables que proporcionen una estimación indirecta del idioma utilizado. A veces la raza, el grupo étnico o el lugar de nacimiento dan una indicación del idioma que habrá que atribuir. Si se dispone de este medio de identificación, el equipo de revisión podrá decidir utilizarlo para determinar el idioma del jefe del hogar. De no ser así, pueden utilizarse las variables de edad y de sexo para la atribución. La matriz de atribución sería semejante a la del gráfico 15.

Gráfico 15

Valores iniciales de una matriz de atribución dinámica para la variable de idioma

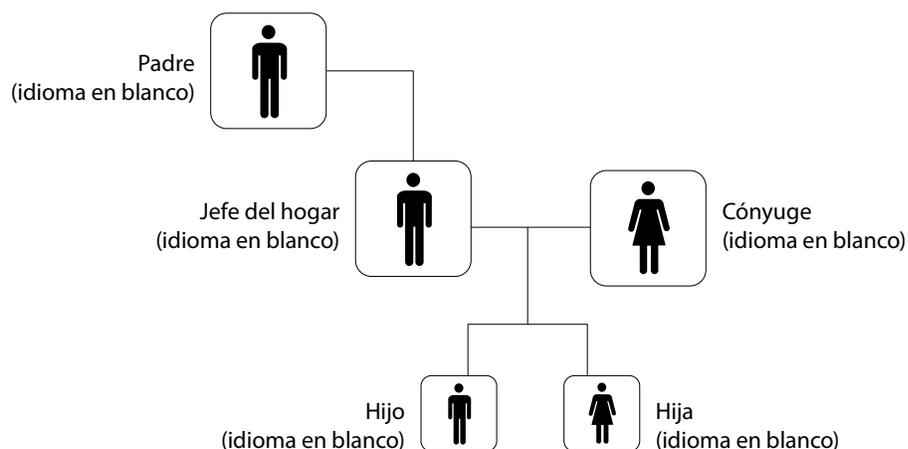
	Edad					
	Menos de 15 años	15-29 años	30- 44 años	45-59 años	60-74 años	75 años y más
Masculino	Idioma 1	Idioma 1	Idioma 1	Idioma 1	Idioma 1	Idioma 2
Femenino	Idioma 1	Idioma 1	Idioma 1	Idioma 1	Idioma 1	Idioma 2

183. Si se decide efectuar la atribución, el programa asigna al jefe del hogar un valor de idioma basado en las variables de grupo de edad y sexo. En tal caso, los valores de la matriz de atribución serán sólo los correspondientes a anteriores jefes del hogar, ya que todas las demás personas de un hogar determinado recibirán el mismo código de idioma que el jefe del hogar.

184. En este punto, si en el hogar no hay todavía una persona que haya declarado hablar un idioma determinado, el programa de revisión recurrirá a la matriz de atribución para asignar al jefe del hogar un valor de idioma basado en sus valores de edad y de sexo. El idioma asignado será el más reciente de los que figuren en el archivo de datos de idioma hablado por otro jefe del hogar de la misma edad y el mismo sexo. Puesto que la matriz de atribución se “actualiza” constantemente a medida que se

Gráfico 16

Ejemplo de miembros de un hogar a los que no se han asignado valores en la variable de idioma



encuentran casos aceptables, es probable que el idioma asignado sea un idioma que se habla en la comunidad en general (véase la figura 16).

185. Las excepciones a las normas de revisión se producirán ya al inicio de una serie de revisiones. El personal debe prestar mucha atención a los cambios de idioma que pueden producirse cuando pasan de una zona geográfica a otra. En algunos países también hay que prestar atención a grupos de personas que hablan idiomas distintos en un mismo lugar. Sin embargo, incluso así, a menos que la presencia de determinados idiomas no se declare en todos los casos, el porcentaje de valores asignados y no asignados resultantes de la atribución debería ser aproximadamente el mismo.

186. En otra revisión podría examinarse la religión. También en este caso los valores de esa variable pueden atribuirse sobre la base de la edad y del sexo. El programa de revisión seguirá actualizándose cuando se disponga de toda la información y tomará respuestas de la matriz de atribución para sustituir la información “desconocida”. Esta matriz de atribución tendrá un aspecto semejante a la del idioma, aunque en las casillas constará la religión, en vez del idioma.

187. Esta explicación se basa en un procedimiento secuencial. Los equipos de revisión que utilicen planteamientos más complejos, como el método de Fellegi-Holt y la metodología de atribución basada en el vecino más próximo (MAVP) (véase el anexo V), efectúan simultáneamente todas las revisiones relacionadas. El procedimiento que se examina también supone la existencia de un orden apropiado en el proceso de revisión.

188. Con respecto a muchas de las características económicas, como la participación en la fuerza de trabajo, el tiempo trabajado en la última semana o en las últimas semanas y el tiempo trabajado en el último año, pueden atribuirse valores utilizando características similares. Utilizando matrices de atribución similares, el programa de revisión puede comprobar rápidamente el valor de las características de las variables, y el proceso de revisión en general debería avanzar con rapidez.

189. Cabe observar que a veces es difícil obtener características adecuadamente revisadas para las primeras matrices de atribución de una serie. Normalmente la oficina de estadística no quiere incluir elementos no depurados como dimensiones en una matriz de atribución; el sexo y la edad no se utilizarán como dimensiones de la matriz de atribución antes de que los datos se hayan depurado. Por tanto, las primeras matrices de atribución usarán distintas variables que no requieran revisión o cuyo valor no pueda cambiar. Para la primera matriz de atribución correspondiente a los elementos de población puede utilizarse el número de personas de la unidad de habitación, asignando un cero a las unidades desocupadas.

190. Para las revisiones de datos de vivienda en general, la primera matriz de atribución también podría utilizar el número de personas de las unidades de habitación como dimensión inicial, pero el equipo de revisión podría modificar las operaciones correspondientes a los elementos de vivienda para tener en cuenta las unidades desocupadas. Por ejemplo, si la primera revisión correspondiente a la vivienda se refiere al “material de construcción de las paredes exteriores” o al “tipo de paredes”, los valores iniciales podrían basarse en el número de personas de la unidad de habitación, incluido un valor para el caso de que la unidad esté desocupada.

191. Cuando la unidad esté desocupada pero el valor del “tipo de paredes” sea válido, se actualizará la primera casilla con el tipo de paredes exteriores. Cuando se conozca este valor, la casilla correspondiente al número de personas de la unidad se actualizará respecto de una unidad ocupada. Sin embargo, cuando se desconozca el material de construcción de las paredes exteriores, la matriz de atribución suministrará un valor para la variable del material de construcción de las paredes exteriores, basado en el número de personas de la unidad.

192. Después de la utilización inicial de esta matriz de atribución, el equipo de revisión tal vez desee pasar a alguna otra característica de la vivienda, como el “tipo de cubierta” o la “tenencia”. Independientemente de la variable que se elija, hay que distinguir claramente entre las unidades y disponer de diversidad suficiente para que no se seleccione repetidamente el mismo atributo. La selección repetida del mismo atributo puede producir valores más parecidos a los de la atribución estática que a los de la atribución dinámica. Por ejemplo, la utilización de una atribución dinámica en un cuartel militar (“local de habitación colectivo”) podría dar lugar a que se aplicara repetidamente el mismo valor si las únicas características seleccionadas son la edad y el sexo. En este caso todos los residentes serían probablemente de sexo masculino, y la mayoría estarían dentro de un intervalo de edades limitado. Por tanto, esa matriz particular tal vez no produciría resultados óptimos. Si la “tenencia” presenta bastante diversidad, con porcentajes suficientes de propietarios e inquilinos, esta variable podría ser útil. De otro modo, el país podría utilizar los distintos tipos de cubierta de la vivienda.

193. En general, muchos equipos de revisión consideran que si utilizan dimensiones comparables para las matrices de atribución necesitarán hacer menos comprobaciones, obtendrán sus resultados más rápidamente y estos serán probablemente más exactos.

g) Casos en que no se utiliza la atribución dinámica

194. Aunque el equipo de revisión decida no utilizar en absoluto la atribución dinámica, la secuencia de revisión seguirá siendo importante. La edad, por ejemplo, está asociada con muchos elementos, en particular la relación con el jefe del hogar, el nivel de escolarización, el empleo y la fecundidad (en el caso de las mujeres). Véase el caso de los miembros del hogar identificados en el gráfico 17 *infra*.

Gráfico 17

Ejemplo de jefe del hogar y un hijo, en el que no consta la edad del hijo

Persona	Relación	Edad	Escolarización	Situación laboral	Ocupación	Hijos nacidos vivos
1	1	40	12	1	33	EN BLANCO
3	3	X	7	EN BLANCO	EN BLANCO	EN BLANCO

Nota: X = No consta la edad.

195. El registro correspondiente a la persona 3 tiene como código de relación 3 (hijo), pero la edad no se ha declarado. Para determinar la edad, el programa de revisión puede utilizar la diferencia de edades entre el jefe del hogar y el hijo (que puede ser un valor de atribución estática o un valor tomado por atribución de una unidad anterior). Si la diferencia es 25, por ejemplo, la edad del hijo será de 15 años (la edad del jefe del hogar, 40 años, menos los 25 años de la diferencia de edades).

196. También se conoce el número de años de escolarización, que en este caso es 7. La edad de 15 años puede corresponder perfectamente a este nivel escolar. Puesto que el intervalo de años de escolarización apropiado para una edad determinada es más breve que el intervalo de años de la diferencia de edades entre el jefe del hogar y el hijo, es mejor comprobar primero si el nivel de escolarización es apropiado. Si se ha registrado ese nivel, puede utilizarse una diferencia de edades determinada por atribución estática o por atribución dinámica para obtener una edad apropiada. Si el nivel de escolarización se desconoce, podrá utilizarse la diferencia de edades entre el jefe del hogar y el hijo para asignar un valor a la variable de edad de este.

197. No obstante, también puede faltar la información relativa a la diferencia de edades. Efectivamente, en muchos países es más probable que se omita el nivel de educación que la edad. El ejemplo siguiente indica las medidas que puede adoptar el equipo de revisión si faltan las variables correspondientes a la edad y al nivel de escolarización.

Gráfico 18

Ejemplo de un jefe de hogar y un hijo en el que se desconocen la edad y el nivel de escolarización del hijo

Persona	Relación	Edad	Escolarización	Situación laboral	Ocupación	Hijos nacidos vivos
1	1	40	12	1	33	EN BLANCO
3	3	X	X	EN BLANCO	EN BLANCO	EN BLANCO

198. En el caso del gráfico 18 no constan ni la edad ni el nivel de escolarización, pero hay otra información. La persona 3 no tiene edad suficiente para tener un empleo y es demasiado joven para haber tenido hijos (o es de sexo masculino). Utilizando la información sobre empleo, puede generarse una edad con un conjunto de valores de atribución estática, pero será una edad inferior a la edad mínima aceptable para tener un empleo. Otra posibilidad es que, si el equipo de revisión utiliza la atribución dinámica, el valor de la matriz de atribución proporcione un valor para la variable de la edad. El valor seleccionado probablemente tendrá en cuenta la edad del jefe del hogar como una de las variables para mantener la coherencia. Por ejemplo, si el jefe del hogar tiene una edad de 20 años y no de 40, obviamente no será apropiado asignar una edad de 14 años a la persona 3. Cuando se fije la edad también podrá determinarse el nivel de escolarización, que deberá ser coherente tanto con la edad como con la variable de la situación laboral.

199. Si el equipo de revisión decide atribuir valores a todos los elementos, o a la mayoría de ellos, deberá formular una estrategia para llevar a cabo las revisiones de manera lógica. Con respecto a los elementos de población, la revisión debería comenzar considerando todos los elementos que pueden tener valores desconocidos. Los equipos de revisión deberían utilizar los datos de las encuestas y registros administrativos, los censos anteriores, el proyecto piloto del censo sometido a consideración y otras informaciones disponibles para ayudar a determinar la inclusión de cada elemento en la primera matriz de atribución, y en las que se elaboren más adelante. Si bien es verdad que los detalles de las matrices de atribución dependen mucho de los países, todas las oficinas nacionales de estadística o del censo tendrán probablemente a su disposición cierta información para estos fines. La comprobación de diversos conjuntos de variables en los procedimientos de atribución dinámica contribuirá a conseguir el conjunto más adecuado para un país dado.

200. Muchos programas informáticos de revisión mantienen constancia del número de personas de la unidad de habitación a medida que se va avanzando en la revisión. Una matriz de atribución para una variable desconocida de sexo, por ejemplo, podría permitir asignar un valor de sexo masculino o sexo femenino según el número de ocupantes de la unidad de habitación. Así, el valor inicial que se elija para una persona con un valor desconocido o inválido en la variable de sexo en un hogar con una sola persona podría ser masculino. En un hogar de dos personas, el valor inicial podría ser femenino. En un hogar de tres personas, el valor sería masculino, y así sucesivamente. La matriz solo se utilizaría como último recurso cuando se hubieran comprobado y resuelto todas las revisiones de coherencia, como el sexo del jefe del hogar y el cónyuge y la existencia de información sobre fecundidad.

h) Tamaño apropiado de las matrices de atribución

201. La mayoría de programas informáticos pueden aceptar matrices de atribución pluridimensionales. Antes de crear las matrices de atribución habría que tener en cuenta los aspectos siguientes:

i) *Problemas que se plantean cuando la matriz de atribución es demasiado grande*

202. Uno de los mayores problemas que tienen algunas oficinas nacionales de estadística o del censo cuando colaboran los especialistas temáticos y los especialistas en procesamiento de datos es el del exceso de celo de los revisores. Es muy fácil extralimitarse al elaborar programas de revisión. En esos casos, lo que puede suceder es que la programación lleve mucho más tiempo del necesario y que se frene el procesamiento del censo o de la encuesta. El equipo de revisión puede decidir, por ejemplo, que para determinar la edad, además de las variables de “sexo”, “nivel de instrucción” y “participación en la fuerza de trabajo”, también haya que incluir, en el caso de las mujeres, el “número de hijos nacidos vivos”. La adición de la variable “número de hijos nacidos vivos” puede permitir una estimación de la edad ligeramente más precisa, pero ello quizá no justifique la mayor complejidad de la programación. Los equipos de revisión tienen que decidir qué número de dimensiones de la matriz de atribución proporcionará los mejores resultados, tanto en lo que respecta a exactitud como a eficiencia. Las matrices de atribución de un tamaño excesivo (con demasiadas casillas) no pueden actualizarse a fondo, y es posible que en su lugar se utilicen de manera inadecuada valores de atribución estática.

ii) *Comprensión del funcionamiento de la matriz de atribución*

203. Además del tamaño excesivo de las matrices, otro posible factor de confusión está asociado con las trayectorias que deben seguirse. Es importante cerciorarse de que tanto el personal temático como el de procesamiento de datos puede seguir bien toda la secuencia de la revisión. Ambos deben velar por que la matriz de atribución realice la tarea prevista. También en este caso los especialistas temáticos y los especialistas en procesamiento de datos deben colaborar para verificar que cada variable o dimensión de la matriz de atribución se aplica adecuadamente. Además, deben asegurarse de que todas las combinaciones funcionan de manera apropiada.

iii) *Problemas que se plantean cuando la matriz de atribución es demasiado pequeña*

204. La matriz de atribución es de tamaño insuficiente si no tiene todas las dimensiones necesarias o si, debido a los agrupamientos (por ejemplo, número insuficiente de grupos de edad o niveles de educación), el mismo valor de la matriz de atribución se utiliza repetidamente antes de ser actualizado. Por ejemplo, si en una matriz de edades no se incluye una dimensión de sexo, es más probable que se asigne a todos los hijos de una familia la misma edad cuando este valor se desconozca. El personal temático deberá colaborar con los especialistas en procesamiento de datos para poner a prueba las matrices de atribución respecto de todas las combinaciones distintas y asegurarse de que ninguna de ellas se produce con demasiada frecuencia.

iv) *Elementos que presentan dificultades en las matrices de atribución*

205. Algunos elementos, como la “ocupación” y la “rama de actividad económica”, han resultado especialmente difíciles de revisar. Mientras que unas matrices de

atribución distintas para la ocupación y la rama de actividad económica pueden producir resultados incongruentes, la comprobación cruzada de todos los pares de valores relativos a la ocupación y la rama de actividad económica puede resultar costosa y difícil. Por ejemplo, si se encuentran barberos o peluqueros trabajando en plantas para la elaboración de pescado, se necesitará alguna otra norma de revisión. Además, el gran número de categorías de ocupación y de ramas de actividad económica puede dificultar mucho la atribución dinámica. Respecto de algunos elementos, el equipo de revisión puede decidir que la revisión es contraproducente y optar por consignar “no contesta” o “no declarado”. Por otra parte, puede bastar con utilizar un procedimiento de atribución estática.

4. Comprobación de las matrices de atribución

206. La estructura básica de la matriz de atribución de un programa informático de revisión debe tener un aspecto parecido al gráfico 19. Las especificaciones de revisión deben identificar las matrices utilizadas para la atribución y usar valores de atribución estática para el conjunto inicial de valores.

a) Establecimiento de la matriz estática inicial

207. El procedimiento que se indica a continuación actualiza la matriz de atribución cada vez que encuentra a una persona con valores válidos en los tres elementos: en este caso, la “relación”, el “sexo” y la “edad”. Sin embargo, cuando en el proceso de revisión se encuentre un valor inválido (o en blanco) para la variable de sexo, la matriz de atribución seleccionará un valor basado en códigos válidos de relación y de sexo (variables ya revisadas).

b) Mensajes de error

208. Los programas de revisión deben contemplar varios métodos para asegurarse de que las revisiones y las atribuciones de valores sean correctas. A continuación se examinan dos de ellas: los comandos de mensaje y los comandos de escritura.

Gráfico 19

Ejemplo de un conjunto de valores para una matriz de atribución estática y modelo de código de atribución

Valores iniciales para la matriz de atribución

A01-EDAD-FM-SEXOREL (2,6)

Jefe del hogar	Cónyuge	Hijo	Otro familiar	Progenitor	No declarado	Sexo
40	40	10	20	65	20	Masculino
40	40	10	20	65	20	Femenino

.

Código de atribución

si EDAD = 0:98

A01-EDAD-FM-SEXOREL (SEXO,RELACIÓN) = EDAD

de no ser así

mensaje “Edad desconocida, por tanto, atribuida” EDAD

escribir “Edad desconocida, por tanto atribuida, Edad =” EDAD

atribuir EDAD = A01-EDAD-FM-SEXOREL(SEXO,RELACIÓN)

mensaje “EDAD conocida ahora” EDAD

final - si

.

.

.

209. Una fuente de información es la aparición de un mensaje, como el que figura en el gráfico 19. Este comando genera mensajes concretos y los recuentos correspondientes (el número total de veces que se produce el mensaje) para niveles geográficos determinados (por ejemplo, zona de empadronamiento, división administrativa pequeña o división administrativa grande), así como para cada cuestionario. Para el conjunto de cuestionarios, un informe resumido podría ser semejante al gráfico 20.

Gráfico 20

Ejemplo de informe resumido sobre el número de atribuciones por error

Recuento	Número de error	Mensaje	Número de línea
-	14-1	Demasiados hijos por mujer	2629
-	14-2	Demasiados hijos por mujer	2645
2	14-3	No constan los hijos varones	2669
2	14-4	No constan las hijas	2678
33	14-5	No consta el mes del último nacimiento	2723
7	15-6	No constan los hijos nacidos vivos; la diferencia de edad entre madre e hijo es correcta	2892

Nota: En ese caso "14" se refiere simplemente al elemento 14 de una determinada serie; los errores se numeran secuencialmente.

210. Un informe organizado por cuestionarios (gráfico 21) podría indicar el número del cuestionario e incluir todos los códigos geográficos especificados. A continuación, el informe podría enumerar los errores descubiertos en el programa, por elemento (en este caso la edad) y por número de línea del programa informático, que se consigna en la parte derecha. En este ejemplo la edad se dejó en blanco, pero la matriz de atribución suministró un valor de 48 años, basado en la relación y el sexo de esa persona. En este caso la edad concreta era desconocida, pero el comando de mensaje puede escribir también esa información, si se considerase conveniente.

211. Naturalmente, si bien tiene sentido enumerar todos los errores individuales de las pruebas aleatorias o pequeños conjuntos de datos seleccionados, en las series de producción el tamaño del producto sería sumamente grande y engorroso (y acabaría perdiendo todo significado). Para superar este inconveniente se establecería un nivel de activación que cancelaría todos o algunos de los problemas del cuestionario individual para el censo completo. Permanecerían, obviamente, las estadísticas resumidas.

Gráfico 21

Modelo de informe sobre los errores de un cuestionario

	Cuestionario: 01 01 017	Número de línea
EDAD (1) =	La edad se desconoce, y se le ha atribuido un valor	#46
EDAD (1) = 48	Ahora se conoce la edad	

c) Enumeración de errores adaptada a necesidades concretas

212. El programa informático también podría incluir otro comando que permitiera un análisis más detallado de las especificaciones de revisión y de la secuencia de revisión. El comando puede utilizarse para mostrar la información antes de que se efectúe un cambio y después de todos los cambios efectuados. Por último, muestra de nuevo el registro o los registros, con todos los cambios realizados. De esta forma el analista puede asegurarse de que en la revisión se siguen correctamente todos los procedimientos. Los resultados pueden verse en el gráfico 22. La primera fila indica las

variables (por ejemplo, provincia, relación, sexo, edad). A continuación se presentan los datos introducidos, seguidos por el error (en este caso no consta la edad), y después los datos una vez efectuado el cambio.

Gráfico 22

Ejemplo de enumeración suplementaria de error por cuestionario, con múltiples variables

	Provincia	Distrito	Jefe del hogar	Relación	Sexo	Edad
Datos introducidos	01	01	17	1	1	
Error	Edad desconocida, por lo que se ha atribuido un valor edad = EN BLANCO					
Datos revisados	01	01	17	1	1	48

213. Este procedimiento ayuda al equipo de revisión a determinar si la revisión sigue los pasos apropiados.

214. El método siguiente representa una forma posible de someter a prueba los procedimientos de revisión. El proceso podría comenzar haciendo que los especialistas realizaran sistemáticamente el análisis creando un hogar “perfecto”. Un hogar perfecto es un hogar completo —jefe del hogar, cónyuge, hijos, otros familiares y personas no emparentadas— con todas sus características. El hogar perfecto deberá superar todas las normas de revisión sin indicaciones de error. El proceso procede como se describe a continuación:

- a) Los encargados del procesamiento de datos introducen un único error en cada hogar, de acuerdo con la secuencia de las especificaciones de revisión y el programa de revisión;
- b) El analista comprueba luego todas las etapas de la secuencia al principio del proceso de revisión;
- c) Cuando la revisión sigue adecuadamente todas las etapas, los encargados del procesamiento de datos aplican una revisión completa a una muestra del conjunto de datos, buscando rasgos específicos del conjunto de datos real y haciendo las modificaciones necesarias;
- d) Por último, los encargados del procesamiento de datos aplican el programa a todo el conjunto de datos.

215. Cuando está convencido de que los mensajes funcionan adecuadamente y de que se han hecho las modificaciones apropiadas, el encargado del procesamiento de datos puede decidir no aplicarlos a niveles más bajos (por ejemplo, en cada cuestionario). Si en un país grande hubiera que aplicar este procedimiento a todos los conjuntos de datos, incluyendo los mensajes para cada cuestionario, la cantidad resultante de líneas y de papel sería prohibitiva. Sin embargo, la práctica de incluir un informe resumido de esos mensajes debería continuar, porque brinda información útil para los distintos niveles geográficos. El resultado será parecido al que muestra el gráfico 22.

216. Los programas informáticos de revisión suelen incluir un procedimiento de salvaguardia. El rastro de verificación muestra todas las modificaciones de datos y contiene una enumeración de los cambios y los valores atribuidos. Una consulta del rastro de verificación determinará si el número de modificaciones es suficientemente bajo para que pueda aceptarse el grupo de registros.

217. Si un elemento particular contiene demasiados errores, tal vez la razón es que el elemento no ha sido bien comprobado previamente, ya sea de forma independiente o en relación con los demás elementos, lo cual indicaría que los empadrona-

dores o los declarantes no comprendieron el elemento. Algunos empadronadores se confunden y, por ejemplo, sólo recogen información sobre la fecundidad en el caso de los adultos de sexo masculino y no de los de sexo femenino. Si los datos se han reunido sistemáticamente de esta forma, el equipo de revisión podría pedir a los programadores que trasladaran los datos de fecundidad de los hombres a las mujeres de una pareja casada. Por lo demás, no es mucho más lo que podría hacer el equipo de revisión en esta etapa para corregir el error.

218. Normalmente, el programa de revisión debe considerar varios archivos distintos para abarcar todas las situaciones. Además, los encargados del procesamiento de datos deberán efectuar cambios debido a que se han encontrado errores de sintaxis o de lógica. Incluso los especialistas más experimentados en procesamiento de datos a veces escriben el signo de “mayor que” en vez del de “menor que”, y el error sólo se descubre después de haber efectuado varias revisiones, puesto que el problema concreto tal vez no se ponga de manifiesto inmediatamente. Del mismo modo, es posible que inicialmente no se pongan de manifiesto los pequeños defectos de lógica. También en este caso los especialistas temáticos y los especialistas en procesamiento de datos deben colaborar para resolver estos problemas en una etapa temprana del proceso de revisión, de ser posible.

d) ¿Cuántas veces debe aplicarse el programa de revisión?

219. Como ya se ha observado, tan pronto como se fija el cuestionario debe comenzar la elaboración y comprobación de las especificaciones y programas de revisión. Cuando se utiliza un enfoque secuencial deben formularse elementos individuales por separado, pero incluso cuando hay que aplicar múltiples variables al mismo tiempo, las revisiones de los elementos individuales deberán comprobarse con pequeñas partes de todo el conjunto de datos. Las especificaciones de revisión deberán ser elaboradas por los especialistas temáticos, y luego los programadores deberán aplicar los programas de revisión individual. Más adelante se puede establecer y aplicar el conjunto de la revisión a partes cada vez mayores del conjunto de datos, al mismo tiempo que continúa el proceso de perfeccionamiento.

220. En general, conviene hacer tres aplicaciones del programa de revisión: para ambas partes del programa y para el conjunto, como se explica a continuación.

221. En la primera aplicación se suministran a las matrices de atribución valores reales en vez de los valores creados en la matriz estática inicial. Algunos países utilizan datos de otras fuentes —bien de censos o encuestas anteriores o bien de registros administrativos— para suministrar valores estáticos a una matriz. El encargado del procesamiento de datos efectúa una aplicación al conjunto completo de datos, o a una parte importante de este, para suministrar valores a la matriz de atribución. Unos valores de atribución estática tomados del conjunto de datos reales tendrán más probabilidades de ser correctos y vigentes. Para la revisión se utilizará sólo alrededor del 2% de esta matriz estática inicial; el resto serán valores de atribución dinámica.

222. En la segunda aplicación del programa se procede a la revisión propiamente dicha. Esta segunda revisión requiere varias aplicaciones para abarcar todas las situaciones. En ese momento los encargados del procesamiento de datos deberán hacer modificaciones para corregir los errores causados por faltas de sintaxis o de lógica. Además, como hasta los especialistas en procesamiento de datos más experimentados pueden cometer errores, es posible que el error sólo se descubra después de varias aplicaciones, dado que el problema concreto tal vez no se manifieste inmediatamente. También puede suceder que inicialmente los pequeños defectos de lógica no se pongan de manifiesto.

223. La tercera aplicación del programa de revisión sirve para comprobar que no quedan errores en el conjunto de datos y que el programa de revisión no ha introducido nuevos errores. Cuando los encargados del procesamiento de datos apliquen el programa por última vez no deberían aparecer indicaciones de la existencia de errores. Si eso sucede es probable que la lógica del programa sea defectuosa, y los encargados del procesamiento de datos deberán modificarla. Esta aplicación suele indicar también al encargado del procesamiento de datos si el programa ha introducido accidentalmente nuevos errores atribuibles a la lógica del programa.

5. Señales de alerta de atribución de valores

224. Las señales de alerta de atribución de valores son un método que se utiliza para conservar información acerca de los datos no depurados. Como se ha mencionado, a muchos equipos de revisión les preocupa que se pierda información potencial cuando se modifican las respuestas no depuradas. En los casos en que se modifica un valor a causa de una incongruencia, es posible que los equipos de revisión deseen conservar el valor o los valores originales a fin de llevar a cabo nuevos análisis demográficos o de errores después del censo. Tanto los especialistas temáticos como los programadores pueden querer analizar distintos aspectos de los datos que no constan, que no son válidos o que presentan incongruencias. Los miembros del equipo de revisión deben asegurarse de que las distribuciones atribuidas y no atribuidas sean coherentes, a fin de observar si aparece algún error sistemático en el plan de revisión y atribución. Por ejemplo, a veces los especialistas en procesamiento de datos utilizan accidentalmente sólo valores de una matriz estática porque el programa no actualiza la matriz de atribución. Si el país llevó a cabo un ensayo previo del censo, es posible que el equipo de revisión necesite investigar las relaciones entre algunas de las variables después del ensayo previo a fin de dar por finalizado el cuestionario. Antes de que fueran habituales los microordenadores con discos duros de gran capacidad, muchas oficinas de estadística no poseían en sus cintas u otros soportes de almacenamiento el espacio necesario para conservar datos adicionales; sin embargo, conservar información acerca de los datos no depurados ya no representa un problema para la mayor parte de los países.

225. Algunos países optan por mantener una variable binaria simple como señal de alerta para cada elemento. Es un método sencillo y sólo requiere un *byte* para cada variable. Por ejemplo, la Oficina del Censo de los Estados Unidos incluye en los registros de vivienda y en los de población señales de alerta de atribución de valores para cada variable al final de cada registro. Para cada elemento de vivienda, por ejemplo, la variable de alerta es inicialmente “0”, pero se cambia por “1” si el elemento original se ha modificado de alguna forma. El programa no conserva el valor original, aunque a veces las oficinas recopilan esos valores para cada registro o para el conjunto.

226. Existen otros métodos para conservar las respuestas sin depurar. En el ejemplo del gráfico 23 la oficina nacional de estadística o del censo ha cambiado la edad del cónyuge de 70 a 40 años utilizando una matriz de atribución. La oficina nacional de estadística o del censo puede consignar fácilmente el valor previo a la atribución —en este caso 70— en el espacio reservado para las alertas de atribución, y reservar la variable utilizada en las tabulaciones publicadas para el valor asignado: 40, en este caso. Para examinar las modificaciones en el conjunto de datos, la oficina de estadística puede realizar distribuciones de frecuencias o tabulaciones cruzadas de los valores asignados y no asignados. Si después de este análisis de los efectos de la revisión en el conjunto de datos las tabulaciones basadas en los datos revisados parecen sospechosas o anómalas, los equipos de revisión tal vez deseen modificar el programa de revisión

o parte de la secuencia del proceso de revisión y almacenar (gracias al aumento de la capacidad de memoria de los discos duros eso no sería un problema) en los registros todos los valores iniciales, para su utilización posterior. Es probable que las oficinas de estadística deseen mantener al menos dos archivos, ya que un archivo que contenga únicamente los datos revisados se tabulará probablemente con mayor rapidez.

Gráfico 23

Ejemplo de registros de población con señales de los valores atribuidos

Persona	Sexo	Edad	Hijos nacidos vivos (HNV)	Alerta (sexo)	Alerta (edad)	Alerta (HNV)
1	1	40	EN BLANCO			1
2	2	40	7		70	

227. El gráfico 24 presenta el caso de una mujer de 13 años de edad cuya información indica que ha tenido un hijo (el valor de la variable de los hijos nacidos es 1). No obstante, el equipo de revisión ha decidido que la edad mínima para el primer nacimiento será de 14 años, y que lo más probable es que los nacimientos atribuidos a mujeres menores de 14 años sean un error. Como siempre, se plantea la cuestión de si este caso representa una situación de “ruido” en el conjunto de datos y no un valor real.

Gráfico 24

Ejemplo de alerta en el caso de una mujer joven con el valor de fecundidad en blanco y una señal de alerta añadida

Persona	Sexo	Edad	Hijos nacidos vivos (HNV)	Alerta (sexo)	Alerta (edad)	Alerta (HNV)
Fecundidad en blanco						
4	2	13	1			
Fecundidad en blanco y señal de alerta añadida						
4	2	13	EN BLANCO			1

228. Con arreglo a las normas de revisión, la atribución anula (deja en blanco) la información correspondiente a los hijos nacidos vivos. Cabe advertir que la señal de alerta correspondiente a los hijos nacidos es un poco más complicada, ya que debe dejar constancia de un valor en blanco que fue atribuido, así como de los valores numéricos introducidos. Cabe suponer que el personal temático desea estudiar el número y las características de las mujeres de 13 años de edad que hayan tenido un hijo. Los encargados del procesamiento de datos pueden hacer constar la información original en la parte del registro reservada a las alertas, que normalmente está al final de la hoja. Entonces la información sobre los hijos nacidos se excluirá del conjunto de cuadros publicados en el caso de esta mujer, pero esa información seguirá estando disponible para investigaciones ulteriores. En algún momento posterior, particularmente cuando se planifique una encuesta complementaria o el censo siguiente, los equipos de revisión podrán utilizar la información sobre hijos de mujeres de 13 años de edad para decidir si es necesario reducir la edad, de forma que se incluya ese dato.

229. Un problema de la utilización de señales de alerta de atribución de valores es que el procedimiento que se acaba de describir ocupa bastante espacio en el ordenador. Cuando las alertas repiten cada variable, el tamaño del conjunto de datos revisados será aproximadamente el doble del conjunto de datos no depurados. Para muchos países esto sería inaceptable si se piensa en su conservación a largo plazo. Sin embargo, los datos originales y las normas de revisión podrían almacenarse para su posterior reconstrucción.

230. Los países con muchos habitantes tal vez prefieran utilizar las señales de alerta de atribución de valores aplicando técnicas de muestreo para fines de investigación. Por ejemplo, un país puede querer crear un conjunto de datos que contenga una de cada 100 unidades de vivienda. En este caso el programa de revisión utilizaría las señales de alerta de atribución de valores en este conjunto de menor tamaño, lo cual contribuiría a evaluar de qué manera la revisión afecta a la calidad de los datos y determinaría las diferencias existentes entre los datos no depurados y los depurados.

F. Otros sistemas de revisión de datos

231. El presente *Manual* describe en buena parte la utilización del método secuencial para la revisión por ordenador de los datos de censos y encuestas. Algunos países aplican para la revisión por ordenador otro método más complicado: la revisión con variables múltiples (véase la sección D.2 *supra*). Fellegi y Holt (1976) fueron los primeros en formular ese procedimiento, que suele aplicarse a las variables más importantes de un censo o encuesta: la edad, el sexo, la relación y el estado civil. Sin embargo, el método puede aplicarse a cualquier grupo de variables o a todas las variables del cuestionario de un censo o encuesta. En dicho método el programa de revisión examina simultáneamente las respuestas a esos elementos con respecto a una persona o a todas las personas de un hogar, a fin de encontrar respuestas que faltan o son incongruentes. Cuando aparecen elementos desconocidos (en blanco), inválidos o contradictorios, una serie de pruebas determina cuál de los elementos seleccionados es más erróneo, y ese es el primero que se cambia. Luego, se repiten los ensayos para comprobar que no quedan respuestas inválidas ni incongruencias; si continuara habiéndolas, la revisión modifica el elemento con el mayor número de problemas pendientes. El procedimiento se repite hasta que no quede ningún error.

232. La Oficina de Estadística del Canadá desarrolló el enfoque de Fellegi-Holt y lo aplicó en los censos canadienses entre 1976 y 1991. Este procedimiento se perfeccionó para el censo del Canadá de 1996 y recibió el nombre de metodología de atribución del vecino más próximo (MAVP). Ello permitió por primera vez la atribución simultánea, con modificaciones mínimas, de variables numéricas y cualitativas para grandes problemas [de revisión y atribución de datos] (Bankier, Houle y Luc, sin fecha).

233. Si el proceso de revisión se lleva a cabo aplicando un método tradicional de atribución dinámica, la información usada para atribuir valores a una serie de elementos de un cuestionario puede proceder de muchas personas distintas, según la información que se utilice para actualizar la matriz de atribución. Por ejemplo, si el sexo, la relación y el estado civil de la persona A son correctos, esos valores servirán para actualizar las matrices de atribución pertinentes. Si el valor correspondiente a la edad de A no consta o es inválido, naturalmente no se utilizará para actualizar las matrices de atribución. En la práctica, se utilizarán otros elementos para actualizar ese valor. De manera que si en el elemento de sexo de la próxima persona consta un valor incongruente y hay que atribuir un nuevo valor a esa variable, la persona A “donará” ese valor. Si también se desconoce la edad, el programa de revisión utilizará la edad de alguna otra persona.

234. La Nueva Metodología de Atribución recurre a donantes de valores para ciertos elementos, con la esperanza de que toda la información que falta o es incongruente pueda proceder de un único donante o de unos pocos donantes. A fin de obtener la totalidad o la mayor parte de información de un único donante, hay que almacenar en la memoria del ordenador registros de datos completos. Entonces, cuando se desconozcan o sean inválidos los valores de edad y los de sexo, la misma variable almacenada proporcionará valores para ambos elementos.

235. Una metodología de atribución dinámica automatizada debería tener los siguientes objetivos:

- a) Un hogar con valores atribuidos debería ser muy similar al hogar en que la revisión ha hallado errores;
- b) Es preferible que, de ser posible, los datos atribuidos a un hogar procedan de un único donante, y no de dos o más. Además, el hogar con valores atribuidos debería ser muy similar a ese donante único;
- c) Unas operaciones de atribución igualmente apropiadas, basadas en los donantes disponibles, deberían tener una probabilidad similar de ser elegidas, de modo que se evite aumentar falsamente el tamaño de grupos pequeños, aunque importantes, de la población (Bankier, Houle y Luc, sin fecha).

236. Con arreglo a la Nueva Metodología de Atribución, los anteriores objetivos pueden lograrse determinando en primer lugar los hogares revisados aceptables que presenten la máxima similitud con el hogar en que la revisión ha detectado un error. Esto significa que debería haber una correspondencia entre los dos hogares respecto del mayor número posible de variables cualitativas, y tan sólo pequeñas diferencias entre las variables numéricas. Los hogares con esas características se conocen como “vecinos más próximos”. El paso siguiente consiste en determinar para cada uno de los vecinos más próximos los conjuntos mínimos de variables no coincidentes (tanto numéricas como cualitativas) que, al atribuirse a otro hogar, le permitirán superar la revisión. Entonces se selecciona al azar una de esas operaciones de atribución que permita superar la revisión y sea tan parecida al hogar que no ha pasado la revisión como a los hogares que la han superado (Bankier, Houle y Luc, sin fecha).

237. En este capítulo se han examinado los procedimientos generales de revisión y tabulación. En el capítulo III se consideran las revisiones estructurales, que abarcan la primera tarea de edición con ordenador, que es también la más importante, en tanto en cuanto permite comprobar que cada una de las unidades de habitación está recogida, está recogida una sola vez y lo está en el lugar adecuado dentro de la jerarquía del país.

Capítulo III

Revisiones estructurales

238. En las revisiones estructurales se comprueba la cobertura general y se determina la correspondencia entre los distintos registros. Estas revisiones deben permitir determinar: *a)* que todos los registros de hogares y locales de habitación colectivos de una zona de empadronamiento están presentes y siguen el orden apropiado; *b)* que hay registros personales para todas las unidades de habitación ocupadas pero que no los hay para las unidades desocupadas; *c)* que en los hogares no hay registros personales duplicados y no faltan registros personales; y *d)* que en las zonas de empadronamiento no hay registros de vivienda omitidos ni duplicados. Así, en las revisiones estructurales se comprueba que, en general, los cuestionarios estén completos.

Recuadro 3

Directrices para las revisiones estructurales

En las revisiones estructurales deberían realizarse las tareas siguientes:

- ☞ Asegurarse de que cada lote procedente de una zona de empadronamiento contiene los códigos geográficos apropiados (provincia, distrito, zona de empadronamiento, etcétera) y que se sigue una práctica común para designar los lotes;
- ☞ Asegurarse de que se han incluido todas las unidades de habitación y de que constan todos los hogares de una zona de empadronamiento;
- ☞ Combinar los hogares en las zonas de empadronamiento que les corresponden y combinar las zonas de empadronamiento en el nivel geográfico superior apropiado;
- ☞ Ayudar a decidir entre la utilización de páginas personales y páginas sobre los hogares, incluidas o no en un cuaderno de empadronamiento, sobre la base del tamaño de la población y del formato del cuestionario;
- ☞ Asignar cada registro a su tipo de registro válido;
- ☞ Tratar de manera separada tanto los registros de locales de habitación colectivos o de grupo como las unidades de habitación;
- ☞ Asegurarse de que existe una correspondencia entre los diversos tipos de registros (por ejemplo, las unidades desocupadas no deben contener ninguna persona, mientras que las unidades ocupadas deben contener por lo menos una persona). Asegurarse de que el número de registros personales de cada hogar corresponda al número total de personas del hogar que figura en el registro de vivienda. Asegurarse de que exista el número apropiado de cuestionarios cuando se utilicen documentos múltiples para un solo hogar y de que estén vinculados adecuadamente;
- ☞ Eliminar los registros duplicados dentro de un hogar (personas duplicadas) y entre hogares (hogares, o partes de hogares, duplicados) para evitar una cobertura excesiva;
- ☞ Resolver la cuestión de los registros en blanco dentro de un tipo de registros;
- ☞ Resolver la cuestión de las unidades de habitación omitidas.

239. Es posible que las revisiones estructurales concretas que se aplican a un censo o a una encuesta deban modificarse con el tiempo debido a los rápidos cambios de la tecnología utilizados para determinar y corregir errores estructurales. Por tanto, en este capítulo se examina la cuestión más general de la validez de los elementos y las relaciones que presentan los elementos entre los registros y dentro de un mismo registro. Los capítulos IV y V tratan de elementos concretos de población y de vivienda.

A. Revisión de datos geográficos

1. Ubicación del local de habitación (localidad)

240. Según los *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008, párr. 2.78), una **localidad** se define como “toda aglomeración de población... cuyos habitantes vivan en unidades de habitación próximas y que tenga un nombre o una situación jurídica reconocida localmente”. Puede encontrarse más información sobre la ubicación del local de habitación en las definiciones de “localidad” y de “población urbana y rural” que figuran en los párrafos 2.78 a 2.88 de *Principios y recomendaciones*. Es esencial que quienes se ocupan del levantamiento de censos de vivienda estudien esta información, dado que los conceptos geográficos utilizados para describir la ubicación de los locales de habitación al levantar un censo son sumamente importantes tanto para la ejecución de este como para la ulterior tabulación de sus resultados (Naciones Unidas, 2008, párr. 2.455).

241. Cuando se efectúa una revisión del valor de localidad, los códigos geográficos deben ser perfectamente correctos. Lograr códigos completos y precisos de la jerarquía geográfica para fines de procesamiento de datos es una de las tareas más difíciles de todo el censo. Si los códigos geográficos son erróneos, los encargados de introducir los datos podrían asignar la unidad o las unidades de habitación a algún otro lugar del país. A menudo resulta muy difícil corregir este tipo de error.

2. Población urbana y rural

242. La distinción tradicional entre las zonas urbanas y rurales de un país se ha basado en la hipótesis de que las primeras, cualquiera que sea la forma en que se definan, tienen un tipo de vida distinto y, en general, un nivel de vida más alto que las zonas rurales. Esta distinción se ha atenuado en muchos países industrializados, y la principal diferencia entre las zonas urbanas y rurales en lo tocante a las circunstancias de la vida tiende a ser el grado de concentración de la población. Si bien en los países en desarrollo continúan siendo significativas las diferencias de género y nivel de vida entre la ciudad y el campo, la rápida urbanización que se produce en ellos ha hecho muy necesario disponer de información sobre zonas urbanas de distintos tamaños (Naciones Unidas, 2008, párr. 2.82).

243. Antes de llevar a cabo el censo, la mayor parte de los países clasifican las zonas geográficas en dos categorías: zonas “urbanas” y zonas “rurales”, y efectúan los ajustes necesarios una vez reunidos los datos censales. Si el país asigna códigos a las zonas urbanas y rurales (por ejemplo, 1 para las zonas urbanas y 2 para las zonas rurales), esos códigos pueden consignarse en la etapa de introducción de datos o pueden determinarse durante la revisión, sobre la base de los criterios que prescriba el equipo de revisión. Cuando el equipo de revisión suministra una lista de unidades geográficas que indica las que son urbanas y las que son rurales, los encargados del procesamiento de datos pueden asignar fácilmente los códigos apropiados a los registros de vivienda.

244. Hay que procurar asegurarse de que, en general, las características de la población sean coherentes con la zona de empadronamiento. Por ejemplo, en algunos países se encontrarán pocos profesionales en las zonas rurales, salvo en el caso de los médicos, los maestros y personas con ocupaciones parecidas, mientras que deberían encontrarse pocos trabajadores agrícolas en zonas urbanas. El equipo de revisión deberá comprobar que la zona geográfica se haya clasificado de manera correcta.

B. Comprobaciones de la cobertura del censo

1. Empadronamiento *de facto* y *de jure*

245. *Definición de residencia habitual.* En general, a los efectos del censo, por “residencia habitual” se entiende el lugar en que la persona vive en el momento del censo y en el que ha estado o tiene intención de permanecer por algún tiempo. En términos generales, la mayor parte de las personas empadronadas no se han desplazado desde hace algún tiempo, y por ello la definición de su lugar de residencia habitual es clara. En otros casos, la aplicación de la definición puede dar lugar a numerosas interpretaciones, en particular si la persona ha realizado frecuentes traslados. Se recomienda que los países apliquen un umbral de 12 meses al considerar el lugar de residencia habitual, de acuerdo con uno de los dos criterios siguientes: *a)* el lugar donde la persona ha vivido de forma ininterrumpida durante la mayor parte de los 12 últimos meses (es decir, al menos seis meses y un día), sin contar las ausencias temporales por vacaciones o motivos laborales, o donde tiene intención de vivir durante al menos seis meses; y *b)* el lugar donde la persona ha vivido de forma ininterrumpida durante al menos los 12 últimos meses, sin contar las ausencias temporales por vacaciones o motivos laborales, o tiene intención de vivir durante al menos 12 meses (Naciones Unidas, 2008, párrs. 1.461 a 1.463).

246. Habida cuenta de esta definición de la residencia habitual, las oficinas nacionales de estadística o del censo suelen proceder a un empadronamiento *de facto* (que indica dónde se encuentran las personas en la noche del censo) o *de jure* (que indica dónde se encuentran habitualmente). En la revisión que se efectúa para comprobar la relación entre los registros de vivienda, particularmente el recuento de las personas del local de habitación y cada uno de los registros personales, debe tenerse en cuenta el tipo de censo de que se trata. A veces los países reúnen tanto información *de facto* como *de jure*. Para cada persona puede haber un elemento que indique: 1) si dicha persona es un residente permanente, 2) si está de visita pero tiene una residencia habitual en otro lugar, o 3) si suele residir en ese hogar pero está temporalmente ausente. Cuando se encuentren los tres elementos, las tabulaciones hechas sobre una base *de facto* utilizan sólo los valores de 1) y 2), y las tabulaciones hechas sobre una base *de jure* utilizan únicamente los valores 1) y 3).

247. Las organizaciones nacionales de estadística o del censo que se encarguen de la aplicación de estas categorías deben tener mucho cuidado en todas las etapas, es decir, durante la reunión y el procesamiento de datos y durante la difusión y el análisis. En particular, las tabulaciones del conjunto de datos deben realizarse teniendo en cuenta la categoría de población deseada. Los usuarios de esas tres categorías deben tener un buen conocimiento de la población seleccionada, ya que el análisis de los datos supondrá la inclusión por duplicado de algunas personas. Si se requiere una población *de facto*, la tabulación debe excluir la categoría 3 (personas temporalmente ausentes); si se requiere una población *de jure*, la tabulación debe excluir la categoría 2. Durante las tabulaciones iniciales y en las tabulaciones para los informes impresos y

los medios complementarios, el equipo de revisión tal vez desee construir un subgrupo del conjunto total de datos para su procesamiento. Para las tabulaciones posteriores la documentación de los archivos debería indicar explícitamente la forma de tratar las distintas posibilidades. Pueden ser más aconsejables los ficheros múltiples.

248. El programa de revisión debe estar diseñado de forma que cuando estén presentes los tres tipos de registros las correspondencias sean apropiadas. El hecho de que los registros *de facto* tengan pocas respuestas puede indicar que se trata en realidad de residentes ausentes o que hay algún otro problema de empadronamiento que puede exigir un trato especial.

2. Jerarquía de los hogares y de las unidades de habitación

249. En el capítulo V se examinan las relaciones entre los hogares, las unidades de habitación y los locales de habitación. La aplicación de esos conceptos dependerá de cada organización nacional de estadística o del censo. Sin embargo, antes de proceder a revisiones concretas de registros de vivienda, el equipo de revisión deberá elaborar métodos de comprobación para asegurarse de que se respete la jerarquía durante la reunión y la introducción de los datos.

3. Fragmentos de cuestionarios

250. Antes de llevar a cabo a una revisión elemento por elemento, el programa informático deberá comprobar los registros válidos, los registros que faltan y los números de línea duplicados, como parte de la revisión estructural. El programa también deberá determinar si los registros que se están revisando corresponden a personas que viven en locales de habitación de grupo. Los encargados de introducir los datos pueden cometer una equivocación al efectuar la consignación de un registro, y a veces acaso dejarán de suprimir información fragmentaria (partes de registros). Una función de las revisiones preliminares debería examinarse si el archivo contiene registros fragmentarios, a fin de suprimirlos. El caso más habitual será un registro que contenga códigos geográficos pero no elementos de población o de vivienda.

C. Estructura de los registros de vivienda

251. Uno de los temas que pueden incluirse en la información recogida en las encuestas o censos nacionales de vivienda es el número de viviendas de un edificio. En este caso la unidad de empadronamiento es un edificio, y la información que se reúne es sobre el número de viviendas de tipo corriente y básicas que hay en él (véase Naciones Unidas, 2008, párr. 2.524).

252. La expresión “revisión general” se refiere a la práctica de asegurarse de que el número de unidades de habitación que constan como partes del edificio corresponde al número total de unidades de habitación del registro de vivienda. En los casos en que existe discrepancia, el número de unidades de habitación consignado como característica del edificio debe corregirse para que corresponda al número de registros de unidades de habitación. Si el edificio en cuestión se ha codificado como un edificio de cinco unidades de habitación pero el recuento efectivo de los registros de unidades de habitación de ese edificio indica cuatro, el equipo de revisión deberá decidir qué ajuste realizará: o cambiar la primera cifra sobre la base del recuento de los distintos registros (que sería lo más aceptable en la mayoría de los casos), o introducir otro registro utilizando información acerca de los registros existentes (lo cual debería evitarse).

D. Correspondencia entre los registros de vivienda y de población

253. Si el censo o la encuesta incluyen registros de vivienda y de población, se necesitará una revisión estructural para asegurarse de que coincidan los dos tipos de registros.

1. Viviendas desocupadas y ocupadas

254. Una unidad de habitación desocupada no debe tener registros de población, mientras que una unidad de habitación ocupada deberá tenerlos. Cuando existen registros de población pero se indica que la vivienda está desocupada, el valor de la variable se modificará de forma que la vivienda conste como ocupada. A veces el formato del registro incluye en el mismo elemento la información sobre ocupación y sobre tenencia, de forma que al hacer la determinación habrá que tener en cuenta también esta información. Además, si existe una respuesta para la variable “valor de la unidad”, en el caso de unidades ocupadas por sus propietarios, o de “renta pagada”, en el caso de unidades ocupadas por inquilinos, el programa de revisión utilizará esta información para su determinación; de otro modo, puede necesitarse una matriz de atribución.

255. En los casos en los que no aparecen registros de población para lo que se supone que es una unidad ocupada, el equipo de revisión deberá decidir si la considerará una unidad desocupada o si le asignará personas de otra unidad. Si la unidad está desocupada, la variable de ocupación puede modificarse fácilmente mediante atribución. Sin embargo, si la unidad está ocupada, el equipo de revisión deberá decidir si asignará personas de otra unidad con el mismo número de personas y características similares, de ser posible, y la forma de hacerlo. Puesto que es imposible conocer las características de las personas que faltan, este método sólo deberá utilizarse cuando el equipo de revisión decida que no tiene otra opción. A continuación se describen tres opciones posibles.

a) Decidir que la unidad de habitación constará como desocupada

256. En este caso el equipo de revisión decide que las unidades de habitación que se han consignado como desocupadas en el recuento sobre el terreno deben seguir como desocupadas, de forma que no se les atribuirán valores. Las normas de revisión de los registros de vivienda para las unidades desocupadas se describen en el capítulo V.

b) Volver a visitar la unidad de habitación en varias ocasiones para completar los cuestionarios

257. La oficina nacional de estadística o la del censo puede decidir aplicar el procedimiento según el cual se requerirá a los empadronadores seguir haciendo averiguaciones acerca de los datos sobre unidades desocupadas hasta tener la certidumbre de que esas unidades habitacionales están desocupadas u ocupadas, o hasta que los empadronadores hayan reunido por lo menos un mínimo de características fiables. En este caso el equipo de revisión deberá formular procedimientos de revisión para comprobar si la unidad está desocupada o si presenta suficientes características que permitan considerarla “ocupada”. Dependiendo de lo que el equipo de revisión considere información “mínima”, se aplicará el proceso ordinario de revisión que se describe en

el capítulo IV, o bien se suministrarán datos de registros donantes para asignarlos a las personas que faltan, como se ha descrito *supra*.

c) Utilizar otra unidad de habitación para sustituir a las personas que faltan

258. Los procedimientos para sustituir hogares enteros o diversas personas que no se hayan consignado se describen en otro lugar del presente capítulo. Esos procedimientos requieren suponer que las personas que faltan tienen las mismas características que las personas utilizadas para sustituirlas, que es algo que muy raras veces sucede; los propios procedimientos son muy complicados. Con todo, sin esos procedimientos el recuento del número de personas, y el de personas por características, podrían disminuir.

2. Duplicación de hogares y de unidades de habitación

259. Las duplicaciones de unidades de habitación se producen por diversas razones. A veces un encargado de introducir los datos consigna dos veces la misma unidad de habitación. Otras veces distintos encargados vuelven a introducir accidentalmente las mismas unidades de habitación o incluso zonas de empadronamiento enteras debido a un control de calidad insuficiente en la oficina nacional de estadística o del censo. En tercer lugar, es posible que un empadronador se equivoque al registrar el código geográfico de una unidad de habitación, creando así información duplicada, al asignarle la misma identidad geográfica que a otra unidad de habitación.

260. Si los encargados de la supervisión en las oficinas se encargaran de introducir los lotes, no se producirían duplicaciones. No obstante, hay que elaborar un programa de revisión que garantice que no se producirán duplicaciones de hogares como consecuencia de que los operadores han introducido dos veces el mismo hogar o los mismos hogares. Los países no deberían clasificar sus datos hasta haber concluido las comprobaciones estructurales y hasta percatarse de que se han eliminado los problemas relacionados con los registros duplicados. Antes de la clasificación, el personal puede corregir manualmente los lotes; después de ello probablemente no podrá descubrir el problema. Una vez clasificados los datos, un programa de revisión puede comprobar los hogares duplicados y utilizar técnicas de atribución para eliminar la consiguiente duplicación de registros.

3. Omisión de hogares y de unidades de habitación

261. Análogamente al caso de la duplicación, después de la clasificación puede ponerse de manifiesto que hay hogares que faltan. Por ejemplo, el programa de revisión prevé una secuencia de hogares al nivel geográfico más bajo, como 1,2,3,4, pero obtiene sólo 1,2,4. Entonces hay que tomar la decisión de reenumerar las unidades o de encontrar algún otro método “aceptable” para que otra unidad sustituya a la unidad 3. Son varios los medios disponibles para agregar los hogares ausentes cuando esté claro que faltan y que deben suministrarse. Un método esencial consiste en duplicar el hogar anterior. No obstante, si el número de personas del hogar es conocido, como suele ocurrir (aun cuando no se conozcan sus características), es posible retroceder y duplicar la unidad previa con el mismo número de personas. De la misma manera, si se conoce la edad y el sexo de los miembros del hogar, esa información puede ser útil para obtener una tasa sustitutiva. No es recomendable tratar de utilizar la atribución dinámica para generar información sobre los miembros del hogar, ya que este método produce muchas veces variables que son incongruentes entre sí.

4. Correspondencia entre el número de ocupantes y la suma de los ocupantes

262. El número de ocupantes que consta en el registro de vivienda deberá ser exactamente igual a la suma de personas del hogar. El programa de revisión suma el número de personas y a continuación compara ese valor con el número de ocupantes que consta en el registro de vivienda. Si la suma difiere del valor correspondiente al número de ocupantes, habrá que ajustar el valor del número de ocupantes para que sea igual a la suma de personas, o ajustar cada uno de los elementos. En el capítulo V se trata de los procedimientos de revisión de los registros de vivienda respecto del número de ocupantes.

a) Cuando el número de ocupantes es mayor que la suma de los ocupantes

263. Si el valor de una variable específica correspondiente al “número de ocupantes” que consta en el registro de vivienda es mayor que la suma de los registros de las personas individuales, se plantea un verdadero problema para el equipo de revisión. Nadie puede conocer las características de las personas que faltan. Por lo tanto, los equipos de revisión que decidan atribuir a las personas omitidas características basadas en las características de hogares similares o sustituirlas por personas de hogares similares, se enfrentan a un dilema. Las personas omitidas no deberían ser reemplazadas. Sin embargo, si se acepta el valor del número de ocupantes, la otra opción supondrá reducir el tamaño de la población empadronada. El equipo de revisión deberá analizar la situación de conjunto y decidir entonces el camino que seguirá.

264. Existen diversos métodos para localizar y sustituir los registros que faltan, pero ninguno de ellos es completamente satisfactorio. Pueden conservarse hogares enteros con características importantes diferentes. Cuando se encuentra un hogar con algunas personas pero no todas, puede buscarse en el archivo un hogar en el que la totalidad o la mayoría de las características conocidas coincidan con las del hogar de que se trata, y entonces podrá efectuarse un ajuste de las personas omitidas utilizando los valores de las otras personas del hogar donante. Programar estas operaciones, sin embargo, es muy complicado, de manera que las oficinas nacionales de estadística o del censo que utilicen este método deberían comenzar a planificar esta operación con mucha anticipación.

265. Una variante de este procedimiento consiste en señalar todos los hogares cuyos registros sean incompletos y proceder al resto de las revisiones. Al final del proceso de revisión, una vez que se hayan corregido todos los elementos, el equipo de revisión podrá decidir que los especialistas en procesamiento de datos revisen el archivo efectuando adiciones y cambios sobre la base del conjunto de datos totalmente depurado. Si se aplica este enfoque secuencial, es posible que el equipo de revisión encuentre donantes aceptables.

b) Comprobación del número de personas por sexo

266. A veces, en el registro de vivienda constará el número de ocupantes desglosado por sexo. En ese caso, en la revisión habrá que sumar por separado el número de personas de cada sexo. También en este caso, si hay una discrepancia entre las sumas y el número de ocupantes, habrá que ajustar uno de los valores. Normalmente se ajustan los totales de los registros de vivienda en vez de añadir registros “omitidos” o de suprimir registros que contienen información útil, porque es probable que el empadronador haya cometido un error en el cuestionario de vivienda.

c) Numeración secuencial

267. Hay que establecer una secuencia para los registros de población y numerarlos siguiendo ese orden. Esa numeración deberá aparecer en el cuestionario como variable, como un número de línea o un número de secuencia. Además, los números de secuencia deben seguir un orden numérico. A veces pueden producirse errores del tipo de cuestionarios o impresos personales que no siguen un orden porque los empadronadores no han reunido la información en la secuencia apropiada o porque se han saltado páginas, dejando accidentalmente en blanco algunas páginas del conjunto de datos. Aunque la revisión y la tabulación no suelen verse afectadas porque no se siga una secuencia, muchas oficinas nacionales de estadística o del censo prefieren volver a fijar una secuencia apropiada para las personas. Por lo tanto, el programa de revisión deberá poder localizar a las personas que están fuera de orden y volver a colocarlas en la secuencia pertinente. Dado que en algunas ocasiones el nuevo orden puede afectar a la relación con el jefe del hogar, esto debe tenerse en cuenta en las especificaciones de revisión. El establecimiento de una nueva secuencia afectará sin duda a variables como el número de línea de la madre o el número de línea del marido.

5. Correspondencia entre los ocupantes y el tipo de edificio/hogar

268. El tipo de relación entre los miembros del hogar debe ser congruente con el tipo de unidad de habitación. A veces los miembros de un hogar aparecen en una casa declarada como local de habitación colectivo, o viceversa. En tales casos el tipo de relación o el tipo de unidad de habitación deberán tener en cuenta el tamaño del hogar y otras variables.

E. Duplicación de registros

269. No es probable que aparezcan números de serie duplicados en los cuestionarios sometidos a lectura óptica o a otro tipo de lectura automática. Para los impresos cuyos datos se introducen manualmente, la oficina nacional de estadística o del censo puede decidir comprobar manualmente que haya una correspondencia entre la lista del hogar y los números de línea que habrá que introducir. Esta comprobación manual puede mejorar la calidad de los datos introducidos, particularmente al compararse los nombres de las personas que aparecen en una página donde se enumeran todas las personas del hogar con los datos que figuran en las columnas, renglones o páginas personales. Dos personas que inicialmente parecieran ser un duplicado pueden ser en realidad mellizos cuando se tienen en cuenta sus nombres.

270. Los impresos que se rellenan manualmente no deberían presentar números de línea duplicados si se han establecido correctamente los métodos de escrutinio de datos y las pautas de omisión. La mayor parte de los programas informáticos actuales crean automáticamente números de secuencia como parte del proceso de introducción de datos. Es posible que se genere un error cuando el personal introduce registros duplicados para una persona, o bien que un número de línea equivocado produzca registros duplicados. A medida que se procesa cada registro, el programa de revisión lo compara con los registros de población anteriores de la unidad de habitación. La revisión deberá determinar que cada número de línea se haya captado correctamente. Los números de línea duplicados son errores, y deben modificarse.

271. Los países pueden decidir elaborar sus propios planes de introducción de datos en vez de utilizar un programa informático ya existente. En ese caso el equipo

de revisión deberá tomar una decisión sobre el nivel aceptable de errores. Para adoptar esas decisiones existen varias alternativas. Una de ellas consistiría en seguir las directrices que figuran a continuación:

- a) Si el número de línea de dos registros distintos es idéntico y el número de características que difieren es 2 o inferior, en la revisión se eliminará uno de los registros, ya que probablemente se trate de un duplicado;
- b) Si tres o más características son distintas, se modificará el número de línea⁷.

⁷ Tradicionalmente, los registros duplicados se localizaban y se corregían manualmente, pero cada vez es más frecuente utilizar un proceso automatizado, parcial o totalmente. Un estudio reciente (Winkler, 2006) ya considera la posibilidad de automatizar en conjunto las revisiones estructurales y las de contenido.

F. Poblaciones especiales

1. Personas que viven en locales colectivos

272. La revisión estructural debe tratar de manera diferente a las personas que viven en locales colectivos —instituciones, cuarteles o clínicas— que a las que viven en unidades de habitación ordinarias. Puesto que los colectivos normalmente no tienen un jefe del hogar, los países deberán determinar la mejor forma de distinguir entre los tipos de unidades. Un método consiste en que haya un tipo de registro diferente para los locales colectivos. Otro método consiste en asignar a la relación un código particular que corresponda a locales “de grupo” o “colectivos”.

a) Cuando los locales colectivos constituyen un tipo de registro diferente

273. Cuando la oficina nacional de estadística o del censo decida utilizar un tipo de registro separado, el equipo de revisión no tendrá dificultad alguna para determinar qué registros corresponden a colectivos. Las tabulaciones correspondientes podrán hacerse con facilidad consultando directamente sólo esos registros. Las variables que sean exclusivas de los registros colectivos, como el tipo de colectivo, pueden ser revisadas por separado, y los valores necesarios también pueden atribuirse por separado. Las variables que se hayan excluido de los registros colectivos pueden ser comprobadas fácilmente para asegurarse de que han quedado efectivamente en blanco. No obstante, el resultado es un archivo más voluminoso, ya que es más probable que esos registros sean más breves que los registros ordinarios de población, aunque ocuparán el mismo lugar en un archivo rectangular. Además, es posible que en el proceso de revisión y atribución algunos programas tengan que comprobar ciertos elementos tanto en los registros de población como en los registros colectivos.

b) Cuando una variable permite distinguir los registros colectivos de otros registros

274. Cuando se utiliza una variable separada en vez de un tipo de registro separado, el equipo de revisión puede tener más dificultades para determinar qué registros corresponden a colectivos. En esas circunstancias las tabulaciones aplicables exclusivamente a los colectivos sólo pueden seguir produciéndose con facilidad consultando la propia variable, la cual indica qué registros se refieren a personas de colectivos. Las variables exclusivas de los colectivos, como la del tipo de colectivo, pueden seguir siendo sometidas a procedimientos de revisión y atribución de valores por separado. Las variables excluidas de los registros colectivos pueden comprobarse con facilidad para asegurarse de que en efecto están en blanco, consultando el código correspondiente a los colectivos. El resultado es un archivo más reducido, ya que los registros adicionales para las personas de los colectivos no son necesarios, sino que simplemente se incluyen como registros de población con un código distinto para la variable de

hogar/colectivo. Cuando se produzca el proceso de revisión y atribución de valores, el programa sólo tendrá que comprobar ciertos elementos en los registros de población, pero no en los registros de población y en los registros colectivos.

c) Cuando se ha omitido el código correspondiente a “tipo de colectivo”

275. El código que indica que se trata de un colectivo puede haberse omitido o ser inválido, o bien puede producirse una discrepancia entre el código de colectivo y los códigos de relación. La solución que se sugiere cuando falta el código correspondiente al colectivo pero los códigos de relación indican un colectivo, consiste en cambiar, en consecuencia, el código de colectivo. Si el código de colectivo está presente pero falta la relación, el código de relación podría determinarse a partir del tipo de colectivo.

d) Cuando aparece el código correspondiente a un colectivo pero todas las personas están emparentadas

276. Cuando aparece un código correspondiente a un colectivo pero los códigos de relación ponen de manifiesto que todas las personas de la unidad de habitación están emparentadas, el código deberá modificarse de forma que denote una unidad de habitación. Por otra parte, si la unidad se ha codificado como hogar pero en ella no hay dos personas emparentadas, podría ser necesario modificar el código para indicar una vivienda de grupo o colectiva. En un hogar puede haber cinco o seis personas no emparentadas y no tratarse de un colectivo. Como se ha subrayado, la consulta entre los miembros del equipo de revisión puede ser necesaria si se quiere resolver casos concretos poco habituales.

e) Distinción entre varios tipos de colectivos

277. En la mayoría de los países se distingue entre varios tipos de colectivos. A veces la información se desglosa aún más para indicar tipos específicos de locales colectivos. Esta información puede codificarse separadamente como un elemento propio (“tipo de local colectivo”) o incluirse en forma de múltiples posibilidades en los códigos de relación del hogar.

2. Grupos de difícil empadronamiento

a) Migrantes estacionales

278. En algunos países en los que se produce una migración estacional, el entrevistador necesitará saber si la unidad está desocupada u ocupada a causa del presente momento de referencia. Así, aunque haya información completa sobre el hogar, la familia también puede haberse empadronado en otro lugar. Por supuesto, también puede darse el caso contrario. Si no se presta la atención necesaria es posible que se omita por completo una familia que tenga dos viviendas en lugares distintos (en algunos casos, para pasar los meses del invierno en un lugar más cálido).

279. En algunas ocasiones hay familias enteras que viven de manera bastante regular en un lugar durante parte del año y en otro lugar durante el resto del año. La oficina nacional de estadística o del censo y el equipo de revisión deberán decidir de qué manera tratar cada tipo de situación. Por ejemplo, algunas personas pasan una parte del año en otra casa, como las que viven en una parte más fría de un país durante el estío y en una parte más calurosa del país durante el invierno. Otro caso es el de los nómadas que viajan parte del año pero son sedentarios el resto del año, que tal vez coincida con la temporada en la que el país decide levantar su censo.

b) Personas sin hogar

280. Por definición, el registro de una persona sin hogar no contendrá información sobre la vivienda. Sin embargo, la creación de un registro “ficticio” (un nuevo registro que inicialmente incluye valores en blanco para algunas variables) facilitará la comprobación estructural y permitirá que el registro sea congruente con la estructura de las demás unidades de habitación. El equipo de revisión deberá decidir si crea este registro de vivienda ficticio para facilitar los procedimientos de procesamiento y tabulación de datos.

c) Nómadas y personas que viven en zonas de difícil acceso

281. En este caso, lo mismo que en el de las personas sin hogar, puede resultar difícil una revisión estructural. Algunos países recopilan cierta información sobre la “vivienda”, que puede utilizarse para ayudar a revisar la estructura de la “unidad”. Por ello, las revisiones de la vivienda serán distintas de las utilizadas en las unidades estándar. La información sobre la población deberá recopilarse lo mismo que en el caso de las personas que viven en unidades de habitación estándar, y revisarse de la forma habitual, siguiendo las orientaciones especificadas a continuación.

d) Residentes civiles temporalmente ausentes del país

282. En los censos *de jure*, los residentes civiles temporalmente ausentes del país pero que viven en hogares que pueden registrarlos deben incluirse en las revisiones de población estándar. En cuanto a los censos *de facto*, un indicador debe identificar a las personas que están temporalmente ausentes, de modo que sea posible determinar tanto las poblaciones *de jure* como *de facto*. La revisión de la vivienda no diferirá debido a la población ausente. No obstante, es evidente que estas personas no se incluirán en un censo *de facto*; por ello, no aparecerán en las revisiones de la población.

e) Extranjeros civiles que no cruzan la frontera diariamente y se hallan en el país temporalmente, con inclusión de las personas no documentadas o los transeúntes que se encuentran en barcos fondeados en puertos en el momento del censo

283. En tanto en cuanto en un censo *de facto* debe estar incluido todo residente del país en el momento del censo, las personas a las que se refiere este apartado también deberían estar incluidas. Los individuos deben estar incluidos en su lugar de residencia en el momento del censo, y revisarse utilizando los sistemas habituales para los elementos relativos a la población. Si no se registra la vivienda de una unidad colectiva, u otra unidad de habitación de tipo especial, tampoco se lleva a cabo la revisión correspondiente a esos individuos. Si los barcos fondeados en puertos se consideran unidades de habitación, deberán describirse, y revisarse, las características de la vivienda, utilizando para ello la información correspondiente a otros barcos para la atribución dinámica.

284. Las personas extranjeras que se encuentran en el país sólo temporalmente en principio no se incluyen en los censos *de jure*. Las personas no documentadas se incluirían, sobre todo en los países que no distinguen en el censo entre personas documentadas y no documentadas (aunque tal distinción produciría mejores resultados). Los transeúntes no se incluirían en el censo *de jure* después de la revisión, a no ser que sean transeúntes en la localidad concreta pero residan habitualmente en el país. Si un barco está habitualmente fondeado en el país, las personas del barco deberían incluirse normalmente como residentes habituales, y revisarse como tales.

f) Refugiados

285. Los refugiados pueden estar en una vivienda provisional. A causa de eso tal vez se necesite una indicación en una variable específica, un tipo de registro separado o un registro de vivienda ficticio para registrar su condición. El equipo de revisión deberá formular y aplicar los procedimientos apropiados. Normalmente, los elementos relativos a la habitación y población utilizarán la revisión estándar, con atribución dinámica en la que se incluya “vivienda de refugiados” como indicador.

g) Personal militar, naval y diplomático y sus familiares que están fuera del país, y personal militar, naval y diplomático extranjero y sus familiares que se encuentran en el país

286. En un censo *de jure*, normalmente se incluiría el personal militar, naval y diplomático y sus familiares, independientemente de que se encuentren dentro o fuera del país. En muchos países la información sobre el personal militar no se obtiene en un censo, por lo que la oficina de estadística del país debe conformarse con recuentos simples o recuentos con información adicional mínima. Cuando la información es limitada, el uso de la atribución dinámica será difícil, y es probable que se produzcan errores en el conjunto de datos; por ello, normalmente es preferible no incluir en el censo los hogares de militares registrados bajo esta categoría. El personal diplomático puede presentar problemas semejantes. De todas formas, el empadronamiento dentro de un país puede dar buenos resultados cuando se utilizan cuestionarios y procedimientos estándar; por ello, estas unidades de habitación deben incluirse en la revisión regular, aunque debería haber un indicador de la condición especial de esa unidad de habitación. Como las unidades de habitación fuera del país quizá no se enumeran de la forma habitual, deberá procederse con cautela al determinar si se incluyen o no en las revisiones (no obstante, podrían incluirse en algunas de las tabulaciones).

287. En un censo *de facto*, normalmente sólo se incluirían las unidades de habitación que se encuentran dentro del país. Por lo general, los hogares del personal militar, naval y diplomático que viven fuera del país no se incluirían. Sus viviendas serían registradas por quienes viven en esas unidades en el país de origen y se incluiría la población de los residentes actuales.

h) Civiles extranjeros que cruzan la frontera a diario para trabajar en el país

288. Los civiles extranjeros que cruzan la frontera diariamente para trabajar en el país normalmente no se incluirían en los censos ni *de jure* ni *de facto*, ya que no residían en el país en la fecha de referencia ni residen habitualmente en el país. Tanto en los censos *de jure* como en los *de facto* lo común es que se registren en el país de origen.

i) Civiles residentes que cruzan la frontera a diario para trabajar en otro país

289. Los civiles residentes que cruzan la frontera diariamente para trabajar en otro país son residentes del país que realizó el censo, y deben incluirse en los recuentos *de jure* y *de facto*. En estos casos las partidas correspondientes a la vivienda y población se revisarían en la forma habitual.

j) Marineros mercantes y pescadores residentes en el país que se encuentran en el mar en el momento del censo (incluso aquellos que no tienen otro lugar de residencia que sus camarotes)

290. Los marinos mercantes deberían empadronarse en un censo *de jure* en sentido estricto, y también en un censo *de jure* modificado (censo ajustado para incluir

a las personas que no tienen otra residencia), pero no en un censo *de facto*. Cuando se incluyan, las revisiones de la vivienda deben hacer referencia al tipo especial de residencia, pero debería ser posible utilizar revisiones estándar para los elementos relativos a la población cuando para los barcos se utilice el cuestionario normal del país.

G. Determinación del jefe del hogar y del cónyuge

1. Revisión de la variable de jefe del hogar

291. Al identificar a los miembros del hogar, lo tradicional es comenzar con el jefe o persona de referencia del hogar, y seguir con el resto de los miembros, según su relación con él. El jefe del hogar es la persona que los demás miembros de ese hogar reconocen como tal. Los países pueden utilizar el término que consideren más apropiado para identificar a esta persona (jefe del hogar, cabeza del hogar, persona de referencia del hogar) siempre que se utilice solo esa persona para determinar las relaciones entre los miembros del hogar. Se recomienda que en los informes publicados el país aclare los conceptos y definiciones utilizados (Naciones Unidas, 2008, párr. 2.114).

a) El orden de las relaciones

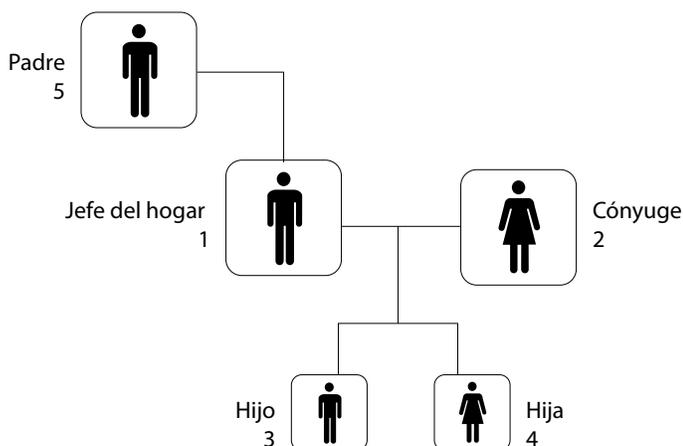
292. El orden de las relaciones de la unidad afectará al procedimiento de revisión, ya que en muchos de ellos se supone que el jefe del hogar es la primera persona y que sus datos serán revisados primero. Por ejemplo, en la revisión correspondiente al jefe del hogar se comprueban primero variables como el idioma, el grupo étnico y la religión. Si el jefe del hogar posee información válida para cualquiera de esas variables, esa información se atribuirá a cualquier otra persona del hogar cuando se haya omitido, tiene un código erróneo o se ha introducido erróneamente (véase el capítulo IV). La información sobre el jefe del hogar debe revisarse en primer lugar porque sus características se utilizarán para asignar o atribuir valores a otros miembros del hogar.

b) Cuando el jefe del hogar no es la primera persona

293. Las operaciones que los empadronadores realizan sobre el terreno, según los diferentes tipos de situaciones con que se encuentran respecto de la designación del jefe del hogar, tendrán efectos en el proceso de revisión. Para comprender mejor este aspecto puede considerarse el hogar que figura como ejemplo en el gráfico 25.

Gráfico 25

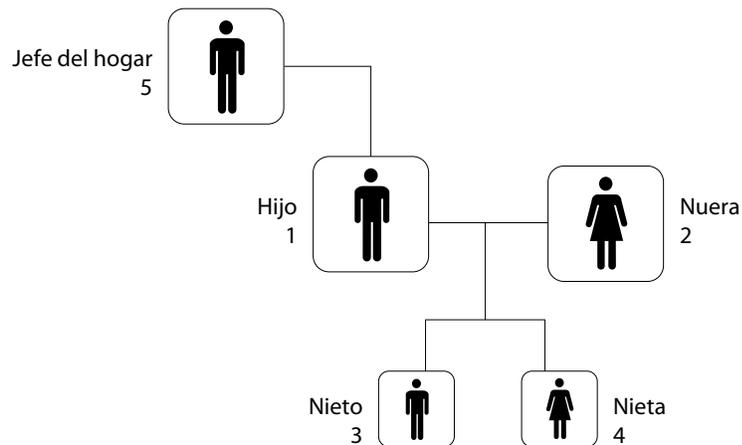
Ejemplo de hogar cuyo jefe ha sido empadronado como primera persona



294. El hogar que representa la figura 25 presenta la situación típica que suele encontrarse sobre el terreno: un jefe del hogar y un cónyuge, sus hijos y el padre del jefe del hogar. Si el empadronador reúne la información de esta manera, un proceso de revisión basado en la premisa de que el jefe del hogar ocupará la primera posición del hogar funcionará perfectamente.

295. Sin embargo, si el empadronamiento se realiza de manera que el abuelo es designado jefe del hogar, las relaciones se reconfiguran, como se observa en la del gráfico 26. Esta situación se produciría cuando un empadronador fuera a una casa, encontrara una familia nuclear compuesta de marido y mujer y sus hijos y durante la entrevista el padre del jefe del hogar entrara en la habitación donde se realiza la encuesta y afirmara que él es el jefe del hogar. Si el presunto jefe del hogar estuviera de acuerdo, la persona 5 se convertiría en el jefe del hogar, y la persona 1 se convertiría en el hijo, la persona 2 en la nuera, y así sucesivamente.

Gráfico 26

Ejemplo de hogar cuyo jefe ha sido empadronado como quinta persona

296. Obviamente, como se pone de manifiesto en los hogares que ilustran los gráficos 25 y 26, los procesos de revisión basados en diferentes designaciones del jefe del hogar serían distintos. Existen tres posibilidades para determinar el jefe del hogar efectivo con miras a las revisiones y las tabulaciones restantes: i) puede utilizarse un indicador para señalar qué persona es el jefe del hogar, y dicho indicador puede aplicarse en todas las revisiones y tabulaciones; ii) si el jefe del hogar no ha sido empadronado como primera persona, se le puede trasladar a la primera posición, y trasladar consecutivamente a una posición posterior a las personas que ocupan un lugar anterior de la lista; y iii) pueden modificarse los códigos de relación para indicar que la primera persona es el jefe del hogar, independientemente de las demás relaciones (véanse los párrafos 297 a 301).

i) *Asignación de un indicador al registro del jefe del hogar*

297. En los procedimientos de revisión relativos al jefe del hogar se utiliza un indicador para determinar el número de línea del jefe del hogar de la unidad. Si el jefe permanece en la posición establecida al reunir la información, puede utilizarse un indicador para esa posición y siempre será fácil encontrar al jefe del hogar cuando haya que realizar una revisión o tabulación determinada. Puede haber una variable (“indicador del jefe del hogar”) en el número de línea del jefe del hogar, que se utilizará durante la revisión para asignar o atribuir a otras personas de la unidad características

omitidas o inválidas. Si el jefe del hogar es la primera persona del hogar, el valor de la variable “indicador del jefe del hogar” será “1”.

ii) *Determinación de la primera persona como jefe del hogar*

298. El equipo de revisión puede decidir trasladar al jefe del hogar a la primera posición del hogar. En este caso la programación correspondiente es algo más compleja que la que se necesita en el caso i) anterior. El especialista en procesamiento de datos deberá elaborar un programa que traslade al jefe del hogar a la primera posición de la lista, seguido por la persona que anteriormente se hallaba en la posición 1, seguida después por la persona que se encontraba en la posición 2, y así sucesivamente hasta llegar a la persona inmediatamente anterior a la que era el jefe del hogar. Así, si el jefe del hogar se encuentra en la posición 5, el orden del resto de las personas del hogar pasará de 1, 2, 3, 4, 5 a 5, 1, 2, 3, 4. Una vez hecha esta modificación, el jefe del hogar ocupará la posición 1, lo que facilitará el resto de las revisiones, ya que él siempre estará en esa posición. No obstante, si se efectúa esta operación se causará cierto “perjuicio” a la integridad del conjunto de datos. Puesto que se ha modificado el orden de las personas, los analistas pueden tener dificultades para determinar el orden real de las personas obtenido sobre el terreno y el efecto potencial de este orden al interpretar los resultados.

iii) *Reasignación de los códigos de relación de modo que la primera persona sea el jefe del hogar*

299. Si el equipo de revisión decide que la primera persona empadronada sea el jefe del hogar, habrá que seguir en la revisión los procedimientos siguientes:

- a) Se asigna a la primera persona el valor de jefe del hogar;
- b) Se aplica un procedimiento que reasigna valores a otras personas del hogar para hacer el ajuste correspondiente.

300. Tómese como ejemplo el gráfico 26, donde el progenitor es inicialmente el jefe del hogar. Cuando se convierte a la persona 1 en jefe del hogar, habrá que asignar el valor de “cónyuge” a la persona 2, se asignará el valor de “hijos” a las personas 3 y 4, y se atribuirá el valor de “progenitor” a la persona 5 (como se indica en el gráfico 25). El programa deberá contener una matriz que conserve tanto los valores iniciales como los valores modificados.

301. La integridad del conjunto de datos se ve aún más afectada con este procedimiento. El orden de las personas no se modifica como en el ejemplo anterior, y los analistas no tendrán dificultad alguna para determinar el orden real de las personas obtenido sobre el terreno. No obstante, todas las relaciones cambiarán y los analistas no sabrán qué persona había sido seleccionada inicialmente como jefe del hogar. Asimismo, si en el censo o en la encuesta se recopila el número de persona de la madre, el del padre o el del cónyuge, esos números deben tenerse en cuenta al cambiar la numeración. Por otra parte, las tabulaciones pueden ser nominalmente más fáciles con el jefe del hogar en la primera posición. A diferencia del ejemplo anterior, para este procedimiento los programadores no deberán trasladar físicamente los registros.

c) *Hay más de un jefe del hogar*

302. Cuando se encuentra más de un jefe del hogar, el equipo de revisión deberá determinar a quién se designará jefe del hogar. La revisión deberá efectuarse sobre la base de las características establecidas por los especialistas temáticos y por las

secuencias de revisión. El programa de revisión deberá reasignar entonces la relación de la otra u otras personas designadas como jefe del hogar.

303. Un caso especial es el de los países que permiten “jefes conjuntos”, ya sea como consecuencia de razones socioeconómicas (por ejemplo, ausencia frecuente de los jefes varones que se dedican a la minería o a otras actividades que obligan a dejar a la esposa como jefa) o bien porque los declarantes insisten en la “igualdad”. Por tradición, a los efectos de revisión es importante designar un solo jefe de hogar, manteniendo en ese caso en el registro los datos originales. No obstante, en Naciones Unidas 2008, párr. 2.117, se incluye una disposición para los jefes conjuntos. Si un país decide incluirlos, deben mantenerse en la revisión; de lo contrario, deberían modificarse muchas de las revisiones posteriores propuestas en este *Manual*: cuando los jefes conjuntos practican diferentes religiones, pertenecen a tribus distintas o tienen otras diversas características sociales y demográficas no se puede utilizar una sola persona en los procedimientos de atribución.

d) No se ha consignado un jefe del hogar

304. De la misma manera que sucede en el caso de que haya más de un jefe de hogar, cuando acontece que no se ha encontrado ningún jefe del hogar, en el proceso de revisión habrá que determinar a quién se designa como tal. En este caso es probable que en la revisión haya que ajustar la relación entre otras personas del hogar. Al determinar el jefe del hogar de esta manera deberán tenerse en cuenta variables como la edad, el nivel de instrucción y la actividad económica, a fin de determinar el jefe del hogar más probable. En el anexo IV puede verse un diagrama de secuencia que puede utilizarse con tal fin.

2. Revisión de los datos del cónyuge

a) Cuando se encuentra exactamente un cónyuge en sociedades monógamas

305. Si se encuentra exactamente un cónyuge, la variable “indicador de cónyuge” señala el número de línea del cónyuge para revisiones ulteriores. Las normas de revisión podrían incluir la búsqueda de valores de sexo opuesto para el jefe del hogar y el cónyuge, diferencias de edad apropiadas u otras características pertinentes. (En los países con cónyuges del mismo sexo, la revisión debería adaptarse.)

b) Cuando se encuentra más de un cónyuge en sociedades monógamas

306. Si en una sociedad monógama se encuentra más de un cónyuge en el conjunto de datos, en la revisión habrá que determinar quién es el cónyuge y reasignar las relaciones de las demás personas que se identificaron como cónyuges. También en este caso los especialistas temáticos deberán determinar cuáles serán las características y la secuencia de las revisiones.

c) Los cónyuges en sociedades polígamas

307. Si en una sociedad polígama se encuentra más de un cónyuge, puede suceder que el equipo de revisión tal vez opte por dejar la información tal como estaba, o hacer alguna comprobación de coherencia. Por ejemplo, como mínimo, cada uno de los cónyuges polígamos debería ser del sexo opuesto al del jefe del hogar. Si se encuentran cónyuges del mismo sexo habría que aplicar el proceso de revisión para cónyuges del mismo sexo del que trata el párrafo anterior.

d) Otras características de los jefes de hogar y los cónyuges

308. Las buenas prácticas de revisión implican la atribución de otras informaciones importantes correspondientes al jefe de hogar y al cónyuge cuando se identifican en esta parte de la revisión global. Esas informaciones corresponden a la edad del jefe de hogar y del cónyuge y el estado civil, que quizá sean necesarias posteriormente en los archivos de atribución y para otros objetivos de la revisión. Asimismo, es también una buena idea obtener al comienzo informaciones de contenido “social”, como la religión, el grupo étnico y el idioma del jefe de hogar, sobre todo si este no es la primera persona de la lista. Dado que la mayoría de los programas comienzan con la primera persona de la lista y continúan en orden descendente, es importante que la información del jefe del hogar esté recogida debidamente antes de revisar la información correspondiente a las otras personas de la unidad. En el anexo IV puede verse una muestra de diagramas de secuencia que podría utilizarse para la revisión relativa al cónyuge.

H. Edad y fecha de nacimiento

1. Cuando consta la fecha de nacimiento pero no la edad

309. Cuando se ha recogido la fecha de nacimiento pero no la edad, este dato puede obtenerse restando la fecha de nacimiento de la fecha del censo o de la encuesta. Algunas oficinas nacionales de estadística o del censo prefieren obtener la edad basándose sólo en el año del censo y en el año de nacimiento, lo cual produce un valor con una posible desviación. Si se toman el año y el mes, la edad será más precisa, pero los resultados más exactos se obtendrán con el día, el mes y el año.

2. Cuando hay discrepancia entre la edad y la fecha de nacimiento

310. Cuando en el censo o en la encuesta se obtiene tanto la edad como la fecha de nacimiento se produce una edad “calculada” restando la fecha de nacimiento de la fecha de referencia. Si este valor difiere en más de un año de la edad que se ha indicado, el equipo de revisión tal vez desee adoptar alguna medida correctiva. Normalmente la fecha de nacimiento tiene preferencia sobre la edad que se ha comunicado, y la edad “calculada” reemplaza a la edad declarada.

I. Recuento de asientos inválidos

311. Algunos equipos de revisión tal vez decidan aplicar procedimientos para contar el número de asientos inválidos o incongruentes con respecto a las principales variables (o a todas ellas), como la edad y el sexo, antes de iniciar la revisión propiamente dicha. Si el equipo de revisión se prepara de antemano o realiza encuestas periódicas utilizando los mismos elementos, puede disponer de varias matrices diferentes de atribución dinámica. Si el porcentaje de asientos inválidos o incongruentes es reducido, el equipo de revisión puede decidir emplear sólo unas cuantas variables para la atribución de valores, pero si el porcentaje es considerable tal vez necesite utilizar más variables para tener en cuenta el gran número de atribuciones necesarias.

312. Las matrices de atribución de menor tamaño suelen funcionar mejor porque son más fáciles de comprobar a medida que se van ejecutando las revisiones y las atribuciones de valores y porque se pueden utilizar más fácilmente durante el proceso de revisión. Sin embargo, si los valores se utilizan repetidamente, se necesitará una matriz de atribución de mayor tamaño y con más variedad.

Capítulo IV

Revisión de los elementos de población

313. Este capítulo trata de los procedimientos de revisión de los elementos de población, incluidos los relacionados con las características demográficas, de migración, sociales y económicas. Las especificaciones de revisión correspondientes tienen en cuenta la validez de cada uno de los elementos y la congruencia entre los elementos de población y entre los elementos de población y los de vivienda. Si se cuenta con algún conocimiento sobre las relaciones entre los elementos será posible planear normas de coherencia a fin de obtener datos de más calidad para la tabulación. Por ejemplo, los registros de población no deben incluir mujeres de 15 años de edad con diez hijos ni niños de siete años de edad matriculados en enseñanza preuniversitaria.

314. Al asignar valores a los elementos de población, el equipo de revisión deberá decidir si asigna un valor “no declarado”; un valor de atribución estática para un elemento “desconocido” u otro valor; o bien un valor de atribución dinámica basado en las características de otras personas o unidades de habitación.

315. En muchos casos se prefiere la atribución dinámica, porque elimina la revisión en la etapa de tabulación, cuando solo se dispone de la información de las propias tabulaciones para adoptar decisiones sobre los valores desconocidos. Las matrices de atribución proporcionan valores para los elementos en blanco, los asientos inválidos o las incongruencias cuando no existen otros elementos conexos con respuestas válidas. Algunos países presentan cierta variación en las características de población de ciertas partes del territorio, pero muy poca variación en la mayoría de las localidades concretas. Otros presentan una variación considerable entre localidades, en particular por lo que se refiere a la población urbana y rural. Esta variación deberá tenerse en cuenta al elaborar matrices de atribución, y particularmente al establecer los valores iniciales de atribución estática. Los equipos de revisión deben especificar las circunstancias en que habrá que suministrar un valor para sustituir una variable en blanco. Este valor debería proceder de una unidad de habitación anterior con características similares.

316. Todos los registros de población deberían tener números de serie para facilitar el procesamiento de los datos. Las revisiones estructurales descritas en el capítulo III sirven para comprobar la correspondencia entre el número de secuencia y el orden de los números de serie.

317. En cada registro de población, el equipo de revisión sólo deberá comprobar los elementos pertinentes, que pueden diferir según se trate de zonas urbanas o rurales o en función de las condiciones climáticas u otras circunstancias. Es conveniente hacer una revisión selectiva atendiendo a esas condiciones, pero en la práctica pocos países tienen tiempo o conocimientos técnicos para elaborar y aplicar matrices múltiples que permitan modificar los datos omitidos o incongruentes. Aún es menor el número de países que aplica realmente este procedimiento adicional.

318. La información recopilada en el cuestionario se aplica algunas veces únicamente a determinados grupos de población. Por ejemplo, las preguntas sobre fecundidad sólo se hacen a las mujeres, y las de actividad económica sólo a los adultos.

319. En algunas ocasiones el equipo de revisión debe permitir que conste un valor “no declarado” para ciertos elementos. Es posible que el equipo de revisión carezca de una buena base para atribuir respuestas a ciertas características. La decisión de dejar respuestas “no declaradas” deberá equilibrarse con la necesidad de producir características tabulares apropiadas para la planificación y la formulación de políticas. En la medida en que los casos “no declarados” tengan la misma distribución que los casos declarados, asignar valores a los casos “no declarados” cuando los planificadores necesitan una determinada información no debería plantear ningún problema. Sin embargo, si los casos “no declarados” presentan algún sesgo, la atribución posterior a la compilación puede resultar dificultosa, en especial para zonas pequeñas o en los casos de tipos de condiciones particulares. Por ejemplo, si las declarantes adolescentes se niegan a revelar su información de fecundidad, ni habrá datos sobre esta variable ni el proceso de revisión podrá ayudar a obtener esta información.

320. Los procedimientos de revisión de las variables de población tienden a ser más complejos que los relativos a la vivienda porque las tabulaciones cruzadas suelen ser mucho más complicadas. La mayor parte de los países compilan determinadas características de vivienda únicamente por diversos niveles geográficos, pero es posible que haya muchos niveles de tabulaciones cruzadas para los elementos de población. Como ya se ha explicado, los países que opten por no utilizar la atribución dinámica deberían establecer un identificador para los valores “desconocidos”, que se utilizará cuando se produzcan respuestas inválidas o incongruentes.

321. En los países que apliquen técnicas de atribución dinámica, el equipo de revisión deberá elaborar matrices de atribución simples con dimensiones que permitan distinguir las características de población. Para la mayoría de los países, el grupo de edad y el sexo son las mejores variables primarias para la atribución dinámica y las que deben revisarse en primer lugar. Las oficinas nacionales de estadística o del censo que utilicen un procedimiento de revisión de variables múltiples deberían revisar simultáneamente la edad, el sexo y otras variables, como la relación y el estado civil. Otros elementos que pueden resultar útiles en la atribución dinámica son el nivel de instrucción y la categoría en el empleo.

322. Los equipos de revisión deben prestar mucha atención a no crear sesgos en los datos durante la atribución. No deben suponer que los datos atribuidos y los no atribuidos tendrán necesariamente la misma distribución. A menudo los propios datos desconocidos presentan algún sesgo. Por ejemplo, los ancianos tienen menos tendencia a comunicar su edad que las personas más jóvenes.

A. Características demográficas

323. Los datos sobre la relación, el sexo, la edad y el estado civil de cada persona son básicos para cualquier censo, y probablemente deberían revisarse conjuntamente. Las estructuras por edad y por sexo de las poblaciones y subpoblaciones son fundamentales para prácticamente cualquier tipo de planificación que se base en los censos de población. Esos elementos también son esenciales para producir tabulaciones significativas, ya que virtualmente todos los demás análisis se basan en tabulaciones cruzadas de otras variables por edad y por sexo.

324. El método de variables múltiples (Fellegi-Holt) para la revisión de los datos de población y de vivienda se ha presentado en el capítulo II del presente *Manual*. Puesto que las variables demográficas son parte integrante de toda planificación censal, debería utilizarse este método si el tiempo y los conocimientos disponibles lo permiten. Es prácticamente seguro que la calidad del conjunto de datos general

se beneficiará de un proceso de revisión que examine prioritariamente la edad y el sexo y otras variables seleccionadas para determinar los errores o las incongruencias. Los elementos en los que suelen encontrarse más errores se revisarán en primer lugar, seguidos por los elementos con menos errores o incongruencias.

1. Relación

325. El elemento de la relación se utiliza para ayudar a determinar la estructura del hogar y de la familia. Aparece casi al principio en la mayor parte de los cuestionarios de los censos y encuestas y contribuye a asegurar que se haya contado a todas las personas de la unidad de habitación. El empadronador y el declarante utilizan la información sobre las relaciones entre los miembros del hogar para asegurarse de que no se omita a nadie. El elemento de la relación también ayuda a comprobar la congruencia de las variables de sexo y edad entre los miembros del hogar. La determinación de un único jefe del hogar y de no más de un cónyuge (en las sociedades no polígamas) se trata en las revisiones estructurales.

a) Revisión de la variable de relación

326. Puesto que las estadísticas sobre la relación están cobrando más importancia, hay que procurar elaborar procedimientos de revisión que contemplen formaciones familiares y subfamiliares para diversos tipos de tabulaciones. Evidentemente, para ello será útil establecer en primer lugar códigos de relación apropiados (véase en el anexo I el “Núcleo familiar” y “Composición de la familia” y las recodificaciones relativas al número en la subfamilia y la relación en la subfamilia).

327. Cuando no es posible asignar un valor a la relación y no se utiliza la atribución dinámica habrá que hacer constar un valor “desconocido” para las respuestas inválidas o incongruentes. Aplicando técnicas de atribución dinámica, la relación puede asignarse a partir de una matriz de atribución por edades y por sexos u otras características apropiadas. Las matrices de atribución no deben atribuir relaciones que entren en conflicto con relaciones ya establecidas en el hogar. Por ejemplo, no deben atribuirse segundos y terceros cónyuges, ni siquiera en hogares polígamos, a menos que el grupo de revisión decida aplicar este tipo de procedimiento.

b) Cuando el jefe del hogar debe constar en primer lugar

328. Si el jefe del hogar no aparece como primera persona puede utilizarse un indicador para señalar la posición del jefe del hogar al aplicar las revisiones estructurales presentadas en el capítulo III. Si el equipo de revisión quiere que el jefe del hogar sea la primera persona, este puede ser colocado en la primera posición modificando el orden de las personas o dejando el hogar tal como está y reordenando las relaciones, como se ha indicado en el capítulo sobre las revisiones estructurales. El primer método requiere considerables conocimientos de programación; el segundo puede ser perjudicial para el conjunto de datos si no se actúa con mucho cuidado.

c) Cuando las relaciones se codifican en sentido inverso

329. Algunos empadronadores compilan las relaciones en sentido inverso: en vez de consignar la relación de cada persona del hogar con el jefe, consignan la relación del jefe con cada persona. Así, la relación de la tercera persona consta como “progenitor” en vez de “hijo”. El resultado puede ser un hogar con cuatro progenitores en vez de cuatro hijos. Cuando el equipo de revisión encuentre un problema sistemático de este tipo deberá formular una solución que no cause demasiados trastornos en el hogar.

330. El procedimiento para invertir las relaciones implica normalmente la aplicación de un archivo de “consulta” que contenga las relaciones originales y las relaciones invertidas, teniendo en cuenta el sexo de los declarantes.

d) Cuando aparecen cónyuges polígamos

331. En las revisiones estructurales aplicadas como se indica en el capítulo III ya se habrá comprobado que en los hogares monógamos el jefe sea “uno y sólo uno” y que no haya “más que un cónyuge”. En los hogares polígamos, el equipo de revisión debe decidir cuándo se aceptarán o no las relaciones polígamas. Las relaciones polígamas que parecen presentar algunos hogares son en realidad errores.

332. Por ejemplo, en un hogar puede haberse identificado a un jefe y un cónyuge, pero haberse descrito a otra pareja como “cónyuges” entre sí, lo cual arroja un total de tres cónyuges. En la revisión habría que comprobar que en el caso de la segunda pareja no se trate en realidad de padre y madre, hijo y nuera, hermana y cuñado, o alguna otra combinación. Unas veces esas relaciones pueden determinarse con alguna certeza y otras veces no. Cuando las relaciones mencionadas se codifican, el equipo de revisión esperará que se hayan hecho atribuciones apropiadas. Cuando los cónyuges adicionales son realmente cónyuges, en hogares polígamos, en la revisión debería comprobarse el sexo y tal vez la edad.

e) Cuando aparecen varios progenitores

333. En un hogar no deberían constar más de dos “progenitores”, y estos deberían ser de sexo opuesto. Cuando aparecen más de dos, los progenitores adicionales deberían convertirse probablemente en “otro familiar”. A veces los censos o las encuestas tienen un código para “padre/madre” o “suegro/suegra”, lo cual permitiría que hubiera hasta cuatro “progenitores” en vez de dos, y no más de dos progenitores de cada sexo.

f) Cuando en los censos se compilan las relaciones con especificación de sexo

334. En algunos censos o encuestas las relaciones se compilan con especificación de sexo, como “marido” y “esposa” separadamente, en vez de “cónyuge”, o “hijo” e “hija” en vez de “hijo”, etcétera. Si estas respuestas no se revisan, en las tabulaciones pueden aparecer datos con hijas de “sexo masculino” o maridos de “sexo femenino”. El equipo de revisión deberá decidir la prioridad de las revisiones, es decir, si se dará preferencia a la relación o al sexo. La revisión es más importante en algunos casos —en el del marido y la esposa, por ejemplo— que en otros; por ejemplo, el de un niño pequeño. Téngase en cuenta que no es conveniente utilizar las relaciones con especificación de sexo, ya que la redundancia no sólo no aclara sino que más bien oscurece la naturaleza de las relaciones y requiere revisión adicional.

g) Cuando hay una discrepancia entre la relación y el estado civil

335. Cuando hay una superposición entre la relación y el estado civil no debe haber discrepancia entre las variables: las personas que declaran tener una relación de “cónyuge” deberán constar como “casadas” en el elemento de estado civil. Cuando haya una discrepancia entre esos elementos, el equipo de revisión deberá decidir qué variable modificará. A veces las relaciones son ambiguas, cosa que hay que tener presente al formular especificaciones de revisión. Por ejemplo, en muchos países un cuñado puede ser tanto el hermano de un cónyuge (y no tiene por qué estar casado) como el cónyuge de un hermano o hermana (y tendría que estar casado).

336. En el mundo de hoy también pueden surgir otros problemas en cuanto a las relaciones declaradas. Cuando dos personas no casadas de sexo opuesto viven juntas fuera del matrimonio, el código de relación podría ser “miembro de una pareja no casada” o “cónyuge”. Si el censo o la encuesta tienen un código especial para el miembro de una pareja no casada, el estado civil correcto no debería ser “casado”, a menos que la persona encuestada esté casada con una persona distinta de aquella con la que vive.

337. Del mismo modo, actualmente hay personas del mismo sexo que conviven en relaciones de tipo sentimental o simplemente de conveniencia. Las personas que conviven en una relación no sentimental podrían calificarse como “compañeros de vivienda” o “personas no emparentadas”. Para las personas que conviven en una relación sentimental, la categoría “pareja no casada” podría ser apropiada en algunos países. En este caso el equipo de revisión también deberá decidir el estado civil que más convenga. Los censos no pueden distinguir entre relaciones platónicas y románticas.

2. Sexo

338. El sexo es una de las características más fáciles de compilar, pero hay que prestar cierta atención a su revisión. Es una de las variables más importantes, ya que la mayoría de características de la población se analizan sobre la base del sexo. La atribución de un valor a la variable de sexo exige cierta comparación con otras variables. En algunos casos el sexo se basará en las diferencias entre los sexos de personas relacionadas, el jefe del hogar y el cónyuge normalmente, pero también entre progenitores y familiares políticos. Probablemente el valor del sexo no debería quedar como “inválido” o “desconocido”, por tratarse de una variable de tanta importancia. Por ello habría que prestar cierta atención a la mejor forma de obtener resultados comparables a la situación real de un país. El anexo IV contiene un ejemplo de diagrama de secuencia para revisar la información sobre el jefe del hogar y el cónyuge.

339. Si una persona no es el jefe del hogar ni el cónyuge de este, no existe ninguna otra persona que pueda usarse como referencia; por lo tanto, habrá que comprobar otros elementos del registro de esa persona. Si hay suficientes elementos de fecundidad, habrá que asignar el código correspondiente a sexo femenino. Sin embargo, si el sexo de esa persona se ha omitido o es inválido, por ejemplo, pero existe un cónyuge para el cual consta un valor en la variable del sexo, en la revisión puede atribuirse el sexo opuesto a la persona en cuestión.

- a) Cuando el código de sexo es válido pero se ha consignado el mismo valor en la variable de sexo para el jefe del hogar y el cónyuge

340. En los casos en que parece haber pruebas contradictorias firmes, el código de sexo deberá cambiarse aunque exista un código válido. Por ejemplo, el registro indica que en el hogar hay una segunda pareja casada cuando ya existen en él un jefe del hogar y un cónyuge o una pareja casada en una subfamilia. Si ambas personas de la segunda pareja comunican el mismo sexo, la información sobre fecundidad y otros elementos pueden utilizarse para determinar cuál es de sexo masculino y cuál de sexo femenino. Entonces podrá cambiarse el registro erróneo.

- b) Cuando hay información sobre fecundidad para una persona de sexo masculino o cuando no la hay para una mujer adulta

341. Es posible que en una revisión aparezca en el hogar una persona de sexo masculino con información sobre fecundidad y/o con hijos, un error que puede estar

causado por el número de persona de la madre o una variable similar. Si no hay un cónyuge, puede consignarse el valor de sexo como femenino, en vez de suprimir la información sobre fecundidad. Análogamente, una mujer adulta sin información de fecundidad y sin hijos podrá consignarse como persona de sexo masculino en determinadas circunstancias fijadas por el equipo de revisión.

c) Cuando el código de sexo es inválido y hay un cónyuge

342. Si la variable de sexo está en blanco o es inválida, el programa de revisión debería utilizar los elementos correspondientes a la relación con el jefe del hogar y al sexo del cónyuge, si este valor es válido, para determinar el código correcto. Si la relación con el jefe del hogar es “jefe del hogar”, el programa verificará que haya un cónyuge presente (comprobando si en el hogar hay otra persona cuya relación sea cónyuge). Una vez que se ha determinado el código de sexo del cónyuge, se asigna el código de sexo opuesto al jefe del hogar.

d) Cuando el código de sexo del cónyuge es inválido

343. Si la relación de la persona con el jefe del hogar es “cónyuge” y se indica el sexo del jefe del hogar, el programa asigna a esta persona el sexo opuesto al del jefe del hogar.

e) Cuando el código de sexo es inválido y existe información que denota sexo femenino

344. En el cuestionario hay numerosas pistas que indican si un declarante es mujer. Si el programa todavía no ha determinado el sexo de la persona y existen indicadores que denotan sexo femenino, habría que asignar un valor de sexo femenino al registro de esa persona. Por ejemplo, si para la persona que se está revisando hay suficientes elementos de fecundidad, podrá asignarse un valor de sexo femenino. Los elementos de fecundidad incluyen los hijos nacidos vivos, los hijos que viven en ese hogar, los hijos que viven en otros lugares, los hijos fallecidos y los hijos nacidos vivos en los últimos 12 meses. También es posible que esta persona sea la madre de otra persona del hogar, es decir, que el número de línea de esa persona sea igual al número de línea de la madre de otra persona del hogar.

f) Cuando el código de sexo es inválido y la persona es marido del cónyuge

345. Si con arreglo a un elemento que indica el número de línea del marido la persona es el marido de otra persona del hogar, en la variable de sexo debería consignarse sexo masculino.

g) Cuando el código de sexo es inválido y no hay información suficiente para determinar el sexo

346. Si el equipo de revisión no aplica ninguna técnica de atribución dinámica, habrá que asignar un valor desconocido al elemento de sexo. Lamentablemente, esto significa que todas las tabulaciones deberán incluir una columna adicional o un renglón adicional, o series de columnas o renglones, para personas de sexo desconocido. Puesto que el sexo es una variable binaria, los valores pueden asignarse alternativamente, comenzando con cualquiera de ellos, asignando el sexo opuesto al segundo asiento inválido, y así sucesivamente.

h) Nota sobre las proporciones de los valores atribuidos a la variable de sexo

347. Cuando se aplica una técnica de atribución estática es probable que el sexo femenino se asigne con mayor frecuencia. Las mujeres adultas son las únicas que tienen valores para la fecundidad, y su selección no es totalmente aleatoria. Así, si no se dispone de información suficiente, será más probable que una persona sin ninguna información sea de sexo masculino que femenino. Por ello es importante considerar la posibilidad de elaborar matrices de atribución que tengan en cuenta las proporciones generales entre los sexos.

3. Fecha de nacimiento y edad

348. La edad es una de las características más difíciles de recopilar y de revisar. Sin embargo, probablemente se trata de la variable más importante, ya que virtualmente todas las características de población se analizan sobre la base de la edad. Revisar los valores de edad requiere una amplia comparación con otras variables y otras personas del hogar. En la mayoría de los casos la edad atribuida debería basarse en las diferencias acumuladas entre las edades de personas relacionadas. Si no puede atribuirse de esta forma un valor para la edad, habrá que utilizar otras características del registro de la persona. El procedimiento requerirá probablemente una serie de matrices de atribución que incluyan la edad por sexo, estado civil, relación y asistencia escolar; la diferencia de edad entre madre e hijo; la diferencia de edad entre marido y mujer; y la diferencia de edad entre el jefe del hogar y el cónyuge.

a) Edad y fecha de nacimiento

349. En la revisión estructural se calcula la edad a partir de la fecha de nacimiento. Sin embargo, es útil examinar en primer lugar la diferencia entre la edad y la fecha de nacimiento. Como se recoge en *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008, párr. 2.136), la información sobre la edad se puede obtener preguntando el año, el mes y el día del nacimiento, o directamente la edad de la persona en el último cumpleaños.

350. La fecha de nacimiento proporciona información más precisa y debe utilizarse siempre que lo permitan las circunstancias. Si se desconoce el día exacto, e incluso el mes de nacimiento, podrá sustituirse por una indicación de la estación del año. La pregunta sobre el nacimiento es pertinente cuando las personas conocen la fecha, que puede establecerse según el calendario solar o lunar o contarse por ciclos regulares asociados a ciertas denominaciones, a la manera como en algunas culturas populares tradicionales se enumeran o identifican los años.

351. Sin embargo, es importantísimo que exista un entendimiento entre el empadronador y el declarante respecto del calendario en que se basa la fecha de nacimiento. Si hay posibilidad de que algunos declarantes contesten refiriéndose a un calendario distinto del utilizado por otros, habrá que prever la necesidad de señalar en el cuestionario el calendario que se utiliza. No conviene que el empadronador trate de convertir la fecha de un calendario a otro, pues eso puede hacerse mejor como parte del trabajo de revisión por ordenador (Naciones Unidas, 2008, párr. 2.137).

352. Es probable que la pregunta directa sobre la edad dé lugar a respuestas menos exactas, por varios motivos. Incluso si todas las respuestas se basan en el mismo método de calcular la edad, siempre existe el riesgo de que el declarante no entienda si la edad que se le pide es la del último aniversario, la del aniversario siguiente o la del

aniversario más próximo. Además, puede haber con relativa facilidad un redondeo a la edad más inmediata terminada en cero o en cinco, que se hagan conjeturas no declaradas como tales y declaraciones deliberadamente inexactas (Naciones Unidas, 2008, párr. 2.138).

353. Muchas oficinas nacionales de estadística o del censo compilan la fecha de nacimiento o la edad, pero no ambas cosas. Como se dice en *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008), la edad en años cumplidos es muy importante: sirve para muchas de las comprobaciones y como dimensión para muchas de las matrices de atribución. Más importante aún es que muchas políticas nacionales se basan en la edad, de manera que hay que hacer todo lo posible para lograr que los valores de edad sean de la máxima calidad. Sin embargo, en la realidad puede suceder que la edad no se comunique. Por tanto, hay que procurar que la edad se haya calculado correctamente y que sea congruente con otras respuestas de los distintos miembros del hogar.

b) Relación entre la fecha de nacimiento y la edad

354. Durante la revisión estructural habrá que calcular la edad si esta no se ha compilado independientemente de la fecha de nacimiento. La revisión del valor de edad durante las comprobaciones individuales será una buena prueba de coherencia dentro de un mismo registro y entre registros, pero el primer paso sería calcular la edad a partir de la fecha de nacimiento y de la fecha del censo. Es importante someter a prueba la edad que se ha calculado sobre la base de la fecha de nacimiento para determinar si corresponde a los límites de la fecha del censo.

355. La edad de los niños nacidos durante el año del censo pero después de la fecha del censo se calculará como -1 y deberá rectificarse. Los niños empadronados después de la fecha del censo probablemente deberían quedar eliminados de él. Sin embargo, si después del examen se descubre que la fecha de nacimiento estaba equivocada por un error de empadronamiento o de procesamiento, habrá que utilizar otras variables para obtener una mejor estimación de la edad.

c) Cuando la edad calculada queda por encima del límite superior

356. Para los censos del año 2000 y posteriores, la mayoría de los países consignarán las cuatro cifras del año de nacimiento. Para los realizados en torno a 2010, los límites aceptables se fijarán desde el decenio de 1900 ó 2000 hasta el año del censo. Aunque basta con tres cifras para que el ordenador haga su trabajo, la utilización de años de tres cifras podría confundir tanto a los empadronadores como a los funcionarios de las oficinas. A veces la edad calculada queda por encima del límite superior

Recuadro 4.

Revisión y atribución de un valor para la edad

Para efectuar la revisión y atribuir un valor a la edad se requerirá lo siguiente:

- ☞ Asignar un valor a la edad en los casos en que la variable esté en blanco;
- ☞ Comprobar que las personas que han estado casadas alguna vez tengan una edad mínima;
- ☞ Comprobar la edad mínima del jefe del hogar;
- ☞ Comprobar la edad mínima de los progenitores; y
- ☞ Llevar a cabo cualesquiera otras comprobaciones específicas para un país.

de las edades definidas por el censo y se necesita un ajuste. Si el censo se realiza en el año 2010 y una persona comunica haber nacido en 1860 es probable que la edad calculada de 150 años quede fuera de los límites aceptables, y será necesario cambiarla.

d) Revisión del valor de la edad

357. Los programas de revisión deberían comprobar la congruencia entre la edad declarada de la persona y la edad declarada de la madre, el padre o el hijo de esa persona. La revisión debería fijar una diferencia mínima de años entre la edad de la madre o la del padre y la edad del hijo. Cuando la edad sea un valor atribuido, habrá que hacer comprobaciones de congruencia valiéndose de elementos como los años que se ha vivido en el distrito (duración de la residencia) y el máximo grado de escolarización terminado (nivel de instrucción). Todas esas comprobaciones deberían efectuarse antes de modificar la edad o antes de asignarle un valor atribuido.

358. El procedimiento de revisión debería comenzar con una comprobación de validez. Si la edad es válida, los especialistas tal vez deseen comprobar si la edad de esa persona es congruente con la edad de su madre (si la madre de la persona está en el hogar) y con la edad de los hijos de esa persona (si esa persona es una mujer que tiene hijos en el hogar). Si las edades no son congruentes habría que tomar nota de la edad de la persona, que sería modificada más adelante.

e) Revisión del valor de la edad cuando constan el jefe del hogar y el cónyuge

359. El próximo paso del proceso de revisión consiste en determinar si hay un cónyuge presente. De ser así, habrá que comprobar la validez de la edad del cónyuge (por lo menos X años de edad, según la edad mínima para el matrimonio definida en el país). Si la edad no es congruente y si se aplican técnicas de atribución dinámica, el programa utilizará un valor de atribución especial derivado de la diferencia entre la edad del marido y la edad de la esposa. Las diferencias de edad varían menos que las propias edades, de forma que una matriz de atribución del programa almacenará la diferencia de edades (de registros anteriores) de un marido y una esposa. Este valor se suma o se resta de la edad del cónyuge de esa persona para calcular una edad.

360. Para asegurar que esa edad calculada sea congruente con otras características, la matriz de atribución deberá incluir también el estado civil, el tiempo de residencia en el lugar y el nivel de instrucción. La exclusión de esas variables puede dar lugar a una edad calculada inferior al número de años que la persona ha vivido en el lugar o inferior a lo que entraña el nivel de instrucción. Por ejemplo, la matriz de atribución puede generar una edad de ocho años, pero la persona puede haber comunicado que ha vivido en el lugar durante 10 años. Sin las otras variables, cuando el programa de revisión compruebe los años que se ha vivido en un lugar, otra matriz de atribución sustituirá el valor correcto de los años de residencia por un valor incorrecto.

f) Revisión del valor de la edad del jefe del hogar cuando no consta el cónyuge pero consta un hijo

361. Cuando para determinar la edad del jefe del hogar no es posible la comparación con la edad del cónyuge, el programa de revisión puede comprobar la relación. Si la relación es “jefe del hogar”, se pueden comprobar los otros registros del hogar (de haberlos), en busca de un hijo o una hija de una edad que se sepa que es correcta. El programa comprueba la edad del hijo o de la hija y calcula una edad para la persona utilizando una atribución dinámica basada en la “diferencia de edades”, similar a la

técnica descrita respecto del marido y la esposa. Como antes, la edad calculada tendrá en cuenta la duración de la residencia y el nivel de instrucción. La edad así calculada será congruente con esas variables; se evitarán errores evidentes si como parte de la matriz de atribución se incluyen los años que la persona ha vivido en el distrito y el nivel de instrucción.

g) Revisión del valor de la edad del jefe del hogar cuando consta un progenitor

362. Cuando una persona no se ajuste a ninguna de las categorías descritas, el programa puede buscar a un progenitor de esa persona en el hogar. Si se encuentra al progenitor es posible calcular una edad con una matriz de atribución que utilice la diferencia de edades. La diferencia de edades entre un hijo y un progenitor suele variar mucho más que la que existe entre el marido y la esposa. Por este motivo el programa aplica este procedimiento sólo cuando ha fracasado la técnica de la diferencia de edad entre el marido y la esposa. La edad calculada deberá tener en cuenta las características educacionales, incluido el nivel de instrucción, los años vividos en el distrito, el estado civil, la fecundidad y la actividad económica. El programa deberá suponer que una persona tiene por lo menos la edad mínima aceptable si se ha casado alguna vez, si tiene hijos o si declara una actividad económica de cualquier tipo.

h) Revisión del valor de la edad del jefe del hogar cuando consta un nieto

363. Cuando una persona no se ajuste a ninguna de las categorías descritas, el programa puede buscar a un nieto de esa persona en el hogar. Si se encuentra al nieto, es posible calcular una edad con una matriz de atribución que utilice la diferencia de edades. La diferencia de edades entre el jefe del hogar y el nieto suele variar mucho más que la que existe entre el marido y la esposa o entre el jefe del hogar y un hijo. Por este motivo el programa aplica este procedimiento sólo cuando ha fracasado la técnica de la diferencia de edad entre el marido y la esposa y entre el jefe del hogar y un hijo. La edad calculada deberá tener en cuenta las características educacionales, incluido el nivel de instrucción, los años vividos en el distrito, el estado civil, la fecundidad y la actividad económica. El programa deberá suponer que una persona tiene por lo menos la edad mínima aceptable si se ha casado alguna vez, si tiene hijos o si participa en una actividad económica de cualquier tipo.

i) Revisión del valor de la edad del jefe del hogar cuando no se dispone de ninguna otra edad

364. Cuando una persona no se ajuste a ninguna de las categorías descritas, el programa puede buscar a otro familiar o a otra persona no emparentada con el jefe del hogar. Si se encuentra a esa persona y la persona ha declarado una edad, el equipo de revisión deberá decidir si utiliza la información de que disponga con una matriz de atribución que emplee la diferencia de edades. Sin embargo, esas diferencias de edad entre el jefe del hogar y otros familiares o personas no emparentadas varían tanto que el equipo de revisión puede decidir abandonar del todo ese método o simplemente utilizar otras variables para la atribución dinámica de la edad del jefe del hogar. En cualquier caso, el programa aplica este procedimiento sólo cuando ha fracasado la técnica de la diferencia de edad entre el marido y la esposa, el jefe del hogar y un hijo, el jefe del hogar y un progenitor y el jefe del hogar y un nieto. Cualquiera que sea la forma en que se calcule la edad, habrá que tener en cuenta las características educacionales, incluido el nivel de instrucción, los años vividos en el distrito, el estado civil, la fecundidad y la

actividad económica. El programa deberá suponer que una persona tiene por lo menos la edad mínima aceptable si se ha casado alguna vez, si tiene hijos o si participa en una actividad económica de cualquier tipo.

j) Revisión del valor de la edad del cónyuge cuando ya se ha determinado la edad del jefe del hogar

365. La revisión del valor de la edad del cónyuge suele efectuarse al mismo tiempo que la revisión del valor de la edad del jefe del hogar, ya que se necesita la información sobre ambas personas para comprobarla conjuntamente. Sin embargo, si la revisión se hace por separado, cuando la edad del cónyuge es inválida o no es congruente con otras variables, habrá que aplicar una matriz de atribución dinámica que utilice la diferencia de edades entre el jefe del hogar y otras variables para estimar de la mejor forma posible la edad del cónyuge. Como en los casos anteriores, la edad calculada deberá tener en cuenta las características educacionales, incluido el nivel de instrucción, los años que se ha vivido en el distrito, el estado civil, la fecundidad y la actividad económica. El programa deberá suponer que una persona tiene por lo menos la edad mínima aceptable si se ha casado alguna vez, si tiene hijos o si participa en una actividad económica de cualquier tipo.

k) Revisión del valor de la edad de otras parejas casadas del hogar cuando se conoce la edad de una de las personas

366. El proceso de revisión debería determinar en primer lugar si ese registro corresponde a una persona casada. De ser así, el programa puede buscar entre los otros registros del hogar al cónyuge de esa persona. Si no se encuentra un cónyuge, el programa pasa a la siguiente parte de la revisión. Si se encuentra un cónyuge, habrá que comprobar la validez de la edad del cónyuge (por lo menos X años de edad, según la edad mínima para el matrimonio definida en el país). Si la edad no es congruente y si se aplican técnicas de atribución dinámica, el programa utilizará en ese momento un valor de atribución especial derivado de la diferencia entre la edad del marido y la edad de la esposa. Las diferencias de edad varían menos que las propias edades, de forma que una matriz de atribución del programa almacenará las diferencias de edades (de registros anteriores) de un marido y una esposa. Este valor se suma o se resta de la edad del cónyuge de esa persona para calcular una edad.

367. Para asegurar que esa edad calculada sea congruente con otras características, la matriz de atribución deberá incluir también el estado civil, la duración de la residencia y el nivel de instrucción. La exclusión de esas variables puede dar lugar a una edad calculada inferior al número de años que la persona ha vivido en el lugar, o inferior a lo que entraña el nivel de instrucción.

j) Revisión del valor de la edad de un hijo cuando ya se ha determinado la edad del jefe del hogar

368. Si se trata del hijo o de la hija del jefe del hogar, puede obtenerse una edad calculada utilizando la edad del jefe del hogar, la diferencia de edades, la duración de la residencia y el nivel de instrucción. También en este caso la edad calculada deberá tener en cuenta las características educacionales, incluido el nivel de instrucción, y los años vividos en el distrito, el estado civil, la fecundidad y la actividad económica. El programa deberá suponer que una persona tiene por lo menos la edad mínima aceptable si se ha casado alguna vez, si tiene hijos o si participa en una actividad económica de cualquier tipo.

m) Revisión del valor de la edad de un progenitor cuando ya se ha determinado la edad del jefe del hogar

369. Si se trata de un progenitor del jefe del hogar, puede obtenerse una edad calculada utilizando la edad del jefe del hogar, la diferencia de edades, la duración de la residencia y el nivel de instrucción. La edad calculada deberá tener en cuenta las características educacionales, incluido el nivel de instrucción, y los años que ha vivido en el distrito, el estado civil, la fecundidad y la actividad económica. El programa deberá suponer que una persona tiene por lo menos la edad mínima aceptable si se ha casado alguna vez, si tiene hijos o si participa en una actividad económica de cualquier tipo.

n) Revisión del valor de la edad de un nieto cuando ya se ha determinado la edad del jefe del hogar

370. Si se trata de un nieto del jefe del hogar, puede obtenerse una edad calculada a partir de la edad del jefe del hogar, la diferencia de edades, la duración de la residencia y el nivel de instrucción. También en este caso la edad calculada deberá tener en cuenta las características educacionales, incluido el nivel de instrucción, y los años vividos en el distrito, el estado civil, la fecundidad y la actividad económica. El programa deberá suponer que una persona tiene por lo menos 12 años de edad si ha estado casada alguna vez, si tiene hijos o si participa en una actividad económica de cualquier tipo.

o) Revisión del valor de la edad de todas las demás personas

371. El equipo de revisión deberá definir matrices de atribución pertinentes para todos los demás familiares y personas no emparentadas del hogar. Las directrices dependerán del censo o de la encuesta de que se trate y de las características sociales y económicas del país. Por ejemplo, una persona que ha estado casada alguna vez, ha tenido hijos o ha participado en alguna actividad económica tendrá probablemente por lo menos cierta edad mínima definida en el país. Sobre la base de esa información, si se aplican técnicas de atribución dinámica, el valor que se obtenga de la matriz de atribución no deberá estar por debajo de la edad mínima. Del mismo modo, si una persona va a la escuela, tiene cierto nivel de instrucción o puede leer y escribir pero no es el jefe del hogar, nunca ha estado casada y no tiene actividad económica, esa persona deberá ser colocada en un grupo cuya edad sea inferior a la edad mínima para ser considerada adulta pero superior o igual a la edad mínima para asistir a la escuela. Entonces podrá encontrarse un valor de la matriz de atribución para las personas de edad inferior a la edad mínima para asistir a la escuela. Aunque no es perfecta, esta técnica limita la gama de valores que puede tomar la matriz de atribución.

4. Estado civil

372. En *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (párrs. 2.144 a 2.151), el estado civil se define como la situación de cada persona en relación con las leyes o costumbres relativas al matrimonio que existen en el país. Las categorías de estado civil que se pueden identificar son, como mínimo: a) solteros, es decir, personas que nunca han contraído matrimonio; b) casados; c) viudos que no han vuelto a contraer matrimonio; d) divorciados que no han vuelto a contraer matrimonio; y e) casados pero separados. En algunos países la categoría b) puede exigir el establecimiento de una subcategoría de personas con contrato matrimonial pero que no viven todavía como marido y mujer. En todos los países la categoría e) abarcará

a las personas separadas legalmente y separadas de hecho, que pueden clasificarse en subcategorías separadas si se desea. Con independencia de que los matrimonios separados todavía puedan considerarse como parejas casadas (porque no tienen libertad para volver a contraer matrimonio), ninguna de las subcategorías de *e*) se incluirá en la categoría *b*). En algunos países será preciso tener en cuenta las uniones consuetudinarias (que son legales y vinculantes en virtud del derecho consuetudinario) y las extralegales, llamadas también uniones de hecho (consensuales).

a) Revisión del valor del estado civil

373. El equipo de revisión deberá decidir la edad mínima apropiada para el primer matrimonio a los efectos del censo o la encuesta. La edad mínima en el primer matrimonio (una edad *X*) puede ser distinta en diferentes partes de un país o para diferentes grupos étnicos. Si, por ejemplo, la población rural suele contraer matrimonio antes que la población urbana, las normas de revisión deberán contemplar este hecho. Normalmente, la oficina nacional de estadística o del censo determinará la edad del primer matrimonio antes del empadronamiento, de forma que sólo se hace la pregunta a las personas mayores de la edad indicada y se asigna automáticamente a las personas más jóvenes a la categoría de “nunca se ha casado”. Sin embargo, si la pregunta sobre el estado civil se hace a todos, el equipo de revisión deberá elaborar un procedimiento de revisión para toda la población.

b) Asignación del valor del estado civil cuando no se utiliza la atribución dinámica

374. Aunque la variable del estado civil sólo debería tabularse para las personas de *X* años de edad o más, siendo *X* la edad mínima para el primer matrimonio, los equipos de revisión deberán determinar si hay que efectuar alguna comprobación y, de ser así, cuál será su alcance. Si el país utiliza solamente los valores “no declarado” o “desconocido” para las respuestas inválidas o incongruentes, cuando se encuentren valores inválidos o incongruentes la respuesta inapropiada deberá ser sustituida por el código correspondiente a “no declarado”. Si en el caso de las personas menores de *X* años de edad falta la respuesta “nunca se ha casado”, habrá que atribuirle ese valor; dado que las oficinas de estadística presentan muestras de datos al público, es importante que elementos como el estado civil contengan siempre información.

c) Asignación del valor del estado civil cuando se utiliza la atribución dinámica

375. Si se utiliza la atribución dinámica, el procedimiento de revisión del estado civil deberá: *a)* atribuir un valor cuando un asiento esté fuera de los límites, y *b)* comprobar la congruencia entre el valor de estado civil declarado y los de relación y edad.

d) El cónyuge debe estar casado

376. Todas las personas que consten como “cónyuge” en la categoría de relación deberán ser codificadas como casadas.

e) El cónyuge en una pareja casada

377. Si el número de línea del cónyuge de la persona *A* (persona *b*) es una variable, en la revisión habrá que dar a la persona *A* el valor de cónyuge de la persona *B*; *A* y *B* deberán estar casados y tener valores distintos en la variable de sexo.

f) Si hay un cónyuge, el jefe del hogar debe estar casado

378. Si no se ha consignado un valor de estado civil pero en la variable de relación con el jefe del hogar consta “jefe”, el programa deberá comprobar si hay un cónyuge (comprobando los valores de relación de los otros miembros del hogar). Si hay un cónyuge, el programa asignará el valor de “casado” al estado civil del jefe del hogar.

g) Jefe del hogar sin cónyuge y sin hijos

379. Si no hay un cónyuge presente y la persona es de sexo masculino, con hijos, el programa atribuirá un valor de estado civil por la edad con presencia de hijos. En el caso de que no haya hijos, el programa podrá atribuir el estado civil por la edad con ausencia de hijos. Un jefe del hogar de sexo masculino cuya esposa no esté en el hogar, probablemente estará divorciado o separado, o será viudo.

h) La atribución se utilizará como último recurso

380. En el caso de las personas cuyos códigos estén fuera de los límites y a las que no sea posible asignar códigos sobre la base de las pruebas anteriores habrá que hacer una comprobación de la edad. Si para esa variable hay un asiento válido inferior a X años, habrá que asignar el valor “nunca se ha casado”. En todos los demás casos habrá que asignar un valor tomado de una matriz de atribución. Deberá establecerse una matriz de atribución por sexo y por edad (bidimensional); por sexo, edad y relación (tridimensional); o por sexo, edad, relación y número de hijos nacidos vivos (cuatridimensional). También en este caso los equipos de revisión deberán haber determinado el orden que seguirá el proceso, de manera que al elaborar las matrices de atribución es importante recordar los elementos que se han revisado y los que no se han revisado. Si solamente se han revisado el sexo y la relación antes del estado civil, la matriz de atribución deberá permitir que los demás elementos figuren como “no declarados”.

i) Relación entre la edad y el estado civil en el caso de personas jóvenes

381. Para todas las personas que tengan un estado civil válido distinto de “nunca se ha casado” habrá que hacer una comprobación de congruencia con la edad. Todas las personas que se han casado alguna vez deberán tener X años de edad o más, siendo X la edad mínima en el país para que pueda considerarse que una persona ha estado casada alguna vez. Si la edad es inferior a X o está en blanco, habrá que hacer nuevas comprobaciones de congruencia basadas en otras variables pertinentes (como el número de hijos nacidos vivos o la actividad económica). Si los asientos correspondientes a estos elementos no son válidos, habrá que asignar al estado civil el valor de “nunca se ha casado”; en todos los demás casos no deberá modificarse el estado civil.

5. Edad del primer matrimonio

382. Con arreglo a los *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (párr. 2.192), la “fecha del primer matrimonio” es el día, mes y año en que se celebró el primer matrimonio. En los países en los que resulte difícil obtener la fecha del primer matrimonio es aconsejable reunir información sobre la edad en el momento del matrimonio o sobre cuántos años han pasado desde que se celebró el matrimonio (duración del matrimonio). Es preciso incluir no sólo los primeros matrimonios contractuales y las uniones de hecho, sino también los matrimonios consuetudinarios y los matrimonios religiosos. En el caso de las mujeres viudas, separadas o divorciadas en el momento del censo, debe obtenerse información sobre la “fecha/edad/

número de años desde la disolución del primer matrimonio”. La información sobre la disolución del primer matrimonio (si es pertinente) proporciona datos necesarios para calcular la “duración del primer matrimonio” en la etapa de elaboración como tema derivado. En aquellos países en los que la reunión de información sobre la duración del matrimonio puede ser más fiable que la edad de la mujer de que se trate, las tabulaciones de la información sobre los hijos nacidos vivos por la duración del matrimonio ofrecen estimaciones de la fecundidad más ajustadas que las que se basan en los hijos nacidos vivos clasificados por la edad de la mujer. Los datos sobre la duración del matrimonio pueden obtenerse sustrayendo la edad en el momento del matrimonio de la edad actual, o directamente del número de años transcurridos desde que se celebró el matrimonio.

383. La fecha del primer matrimonio ha de consignarse para todas las personas que han estado casadas. El programa deberá comprobar una correspondencia: las personas que nunca se han casado no deben tener información, pero las personas que se han casado alguna vez deben tener un día, mes y año válidos. Los equipos de revisión deberán decidir si es necesario que el día y el mes sean válidos: los países que no utilicen la atribución dinámica pueden asignar un valor “desconocido” al día y al mes; los países que utilicen la atribución dinámica pueden atribuir valores al día y el mes cuando estos falten.

a) La edad al contraer matrimonio debe quedar en blanco para las personas que nunca se han casado

384. Las personas que nunca han estado casadas no deben comunicar una edad en el momento del primer matrimonio. Si aparece un asiento válido para una persona que nunca se ha casado, el equipo de revisión deberá decidir si modifica el estado civil o deja en blanco la edad de la persona. Si se modifica el estado civil, los países que utilicen sólo el valor “no declarado” aplicarán ese código. Los países que utilicen la atribución dinámica probablemente deberán utilizar los valores de edad y de sexo para obtener una respuesta apropiada para el estado civil.

b) Debe constar un valor para las personas que han estado casadas alguna vez

385. Los países que no utilicen la atribución dinámica pueden asignar un valor “no declarado” o “desconocido” al año del primer matrimonio. Los países que utilicen la atribución dinámica pueden emplear otras variables, como la edad del cónyuge o las diferencias de edad entre los cónyuges, el número de hijos y los hijos nacidos en el último año, para determinar un valor apropiado para el año del primer matrimonio.

6. Fecundidad: hijos nacidos vivos e hijos vivos

386. “Hijos nacidos vivos” quiere decir el número total de hijos que nacieron con vida, lo cual excluye los nacidos muertos y los abortos espontáneos o provocados. En demografía se utiliza a veces la expresión “hijos nacidos”, pero en el presente documento se utiliza el término “hijos nacidos vivos”.

387. El sector de la población en el que deben reunirse datos sobre cada uno de los temas incluidos en la presente sección lo integran las mujeres de 15 años de edad o más (u otra edad mínima aceptable), independientemente de su estado civil o de subcategorías particulares tales como la de las mujeres que han estado casadas. En los países que no recopilan o tabulan los datos correspondientes a las mujeres de 50 años de edad o más, la actividad deberá concentrarse sólo en reunir datos sobre las mujeres de entre 15 y 50 años. En la investigación de la fecundidad reciente, en algunos países

puede ser apropiado reducir en algunos años el límite de edad inferior (Naciones Unidas, 2008, párr. 2.170).

a) Elementos de fecundidad que se compilan

388. En *Principios y recomendaciones para los censos de población y habitación*, las Naciones Unidas (2008) se recomienda obtener información sobre tres elementos relativos a la fecundidad: número total de hijos nacidos, fecha de nacimiento del último hijo nacido vivo y edad de la madre en el momento del nacimiento del primer hijo nacido vivo. Las respuestas a los elementos de edad y fecha o duración del matrimonio pueden mejorar las estimaciones de fecundidad basadas en el número de hijos nacidos vivos. Además, muchos países siguen recopilando información sobre hijos vivos, lo cual es de ayuda, particularmente en análisis de fecundidad retrospectivos.

389. En los censos y encuestas se reúne información sobre fecundidad respecto de todas las mujeres, aplicando una edad mínima, y a veces también una edad máxima, definidas por el país.

b) Normas generales para la revisión del valor de la fecundidad

390. Para la inclusión de información sobre la fecundidad debe comprobarse las mujeres que no han cumplido la edad mínima, y de todos los varones, y se dejará en blanco toda la información que pudiera haber sobre la fecundidad.

391. La finalidad de la revisión del valor de fecundidad es hacer que los valores sean congruentes entre sí y con la edad:

- a) El número total de hijos nacidos vivos no puede ser mayor que la edad de la persona sumada a una edad mínima definida en el país, multiplicada por un determinado factor. Este factor será 1 cuando se permita un intervalo de un año por nacimiento, de 1,5 para un intervalo de año y medio entre hijos adyacentes, y así sucesivamente. Véase *infra* la sección 8, sobre la “edad de la madre en el momento del nacimiento del primer hijo nacido vivo” (página 102) en relación con la revisión para determinar la diferencia mínima de edad entre la madre y el primer hijo nacido vivo;
- b) El número total de hijos nacidos vivos no puede ser superior a la suma del número de hijos que viven en la unidad de habitación, que viven en otro lugar y que han fallecido. Cuando el número total es superior a la suma de las partes, los equipos de revisión deberán decidir qué valor tiene preferencia, para poder hacer el ajuste correspondiente;
- c) Si se han recogido datos sobre los hijos vivos y los hijos que han fallecido, el número total de esos hijos no puede ser superior al número de hijos nacidos vivos;
- d) El número de hijos nacidos vivos no puede ser inferior al que consta en el asiento correspondiente a “hijos nacidos en los últimos 12 meses”;
- e) Según el país de que se trate y el número efectivo de hijos nacidos vivos e hijos que siguen vivos, podría utilizarse una matriz de atribución para el elemento relativo a los hijos nacidos en los últimos 12 meses, a fin de asignar una respuesta basada en los valores de edad y de hijos nacidos vivos. Sin embargo, hay que tener mucho cuidado al asignar un valor a los hijos nacidos en los últimos 12 meses cuando esta variable haya quedado en blanco. En la mayoría de los países el hecho de que este elemento esté en blanco significa que no ha nacido ningún hijo. Los valores asignados podrían generar un sesgo en los datos;

- f) A veces los países recopilan datos desglosados por sexo sobre los hijos nacidos vivos, los hijos vivos y otros elementos de fecundidad. En tales casos los procedimientos de revisión que aquí se indican serán válidos en conjunto, aunque tal vez los países deseen añadir otras comprobaciones para dar cuenta de la información adicional de que se dispone. Esas comprobaciones adicionales pueden incluir asegurarse de que el número de hijos varones nacidos vivos sea equivalente a la suma de los hijos varones vivos y los hijos varones fallecidos, y de que el número de hijas nacidas vivas sea equivalente a la suma de hijas vivas e hijas fallecidas. En cuanto a la revisión de los valores correspondientes a los hijos sin distinción de sexo, habrá que adoptar medidas apropiadas cuando las sumas no sean equivalentes a las partes.

c) **Relación entre los hijos nacidos vivos y los hijos vivos**

392. Los datos sobre hijos nacidos vivos e hijos vivos se utilizan para efectuar estimaciones indirectas de la fecundidad y la mortalidad. Los resultados del censo o de la encuesta se organizan por grupos de edad de mujeres, con intervalos de un año o de cinco años. Hay distintos algoritmos que permiten obtener estimaciones constantes o variables de la mortalidad. Sin embargo, para obtener resultados óptimos los equipos de revisión deben procurar determinar un procedimiento de revisión apropiado para los datos disponibles.

393. Parte del problema de formular un procedimiento general de revisión es que los países solicitan tipos diferentes de información. Por ejemplo, en según qué países se recopilan las siguientes combinaciones de información:

- a) Hijos nacidos vivos solamente;
- b) Hijos nacidos vivos e hijos vivos (combinando ambos sexos o por separado);
- c) Hijos nacidos vivos, hijos vivos e hijos fallecidos (combinando ambos sexos o por separado);
- d) Hijos nacidos vivos, hijos que viven en el hogar, hijos que viven en otro lugar e hijos fallecidos (combinando ambos sexos o por separado).

d) **Procedimiento de revisión cuando sólo se han declarado valores para los hijos nacidos vivos**

394. Si el país no aplica técnicas de atribución dinámica, habría que asignar un valor “desconocido” a un asiento inválido u omitido en la variable de “hijos nacidos vivos”. En los países que aplican la atribución dinámica, los especialistas deberán decidir si quieren utilizar esta técnica para todos los elementos. De ser así, el valor de hijos nacidos vivos puede obtenerse sobre la base de la edad de la mujer, con intervalos de un año, y por lo menos otra característica. También es posible utilizar una matriz unidimensional para la edad de la mujer, con intervalos de un año. Las otras características podrían ser elementos como el nivel de instrucción o la religión, ya que se sabe que en muchos países las diferencias de fecundidad responden a distintos niveles de instrucción o varían según las diferentes confesiones religiosas.

e) **Procedimiento de revisión cuando se han comunicado valores para los hijos nacidos vivos y los hijos vivos**

395. Si se han dado respuestas a las variables de “hijos nacidos vivos” e “hijos vivos”, el programa deberá determinar lo siguiente:

- a) Si los elementos tienen coherencia interna (si el número de hijos nacidos vivos es igual o superior al número de hijos vivos);
- b) Si cada uno de los elementos es congruente con la edad de la mujer;
- c) Si el valor de “hijos nacidos vivos” concuerda con el de “hijos nacidos en el último año” (o el último hijo que ha nacido), en caso de que se haya recopilado esa información.

396. Los demógrafos utilizan los valores de los hijos nacidos vivos y de los hijos vivos para obtener estimaciones indirectas de la mortalidad. Por ello hay que mantener la relación entre los dos elementos en la revisión. A veces sólo se declara uno de ellos y el otro se desconoce. Una revisión fácil consistiría en suponer que no ha habido fallecimientos de los hijos nacidos vivos y equiparar ambos elementos. Sin embargo, si se equiparan los dos elementos, la estimación indirecta de la mortalidad no tendrá en cuenta a los niños fallecidos después del nacimiento, con lo cual se subestimarán la mortalidad y se sobreestimarán la esperanza de vida. Si en el censo o la encuesta aparecen pocos casos de este tipo, el perjuicio será de poca importancia. Sin embargo, si ocurren con cierta frecuencia, como cabría esperar en los países que aplican el método indirecto, los efectos podrían ser considerables. Un ejemplo es el que se presenta en el gráfico 27.

Gráfico 27

Ejemplo de un hogar con información sobre la fecundidad

Persona	Relación	Sexo	Edad	Hijos nacidos vivos	Hijos vivos
1	Jefe del hogar	1	60		
2	Cónyuge	2	60	5	99
3	Hija	2	40	3	3
4	Nieta	2	20	1	1
5	Nieta	2	18	0	0
6	Nieta	2	1		

Nota: 99 = datos omitidos o inválidos

397. En este caso se han consignado para el cónyuge cinco hijos nacidos vivos, pero, por el motivo que sea, no hay constancia del número de hijos vivos. El declarante o el empadronador no comunicaron el valor, o bien el encargado de introducir los datos se equivocó al consignar la información. En muchos países se aplica un procedimiento de revisión consistente en asignar el valor “5” a los hijos supervivientes, sobre la base del número de hijos nacidos vivos. Sin embargo, si se actúa de esta manera puede introducirse un sesgo en los datos.

398. En realidad, no es necesario modificar el valor. Los países que no utilicen técnicas de atribución dinámica pueden decidir dejar el valor como “desconocido”. Por supuesto, esta decisión también crea un sesgo, ya que la revisión asigna la misma distribución en las tabulaciones a las respuestas “desconocidas” y “conocidas”. Si en un país se necesitan datos sobre los hijos nacidos vivos y los hijos vivos para efectuar estimaciones indirectas de la mortalidad, probablemente se tratará también de un país con problemas de comunicación de los datos. En tal caso es probable que mantener valores desconocidos en los datos introduzca un sesgo en el análisis final. Los datos de las mujeres con valores desconocidos para los hijos nacidos vivos o los hijos vivos no podrán utilizarse en estimaciones de la mortalidad, ya que no puede determinarse la diferencia entre los hijos nacidos vivos y los hijos vivos.

399. Los países que aplican técnicas de atribución dinámica deberían considerar la posibilidad de determinar la información que falta sobre la base de otro elemento

de fecundidad y la edad de la mujer, como mínimo. Las matrices de atribución pueden ser actualizadas cuando exista información válida para la edad de la mujer, los hijos nacidos vivos y los hijos vivos, y esta información podrá utilizarse cuando falte el elemento. Cuando falten los hijos nacidos vivos, la matriz de atribución tendrá la edad de la mujer y el número de hijos vivos. Cuando falten los hijos vivos, la matriz de atribución tendrá la edad de la mujer y el número de hijos nacidos vivos.

400. Además, al elaborar las matrices de atribución es importante recordar que el número de hijos nacidos vivos y el número de hijos vivos deben concordar con la diferencia de edad entre la madre y el hijo mayor (si existe tal información) y el número total de hijos nacidos vivos de una madre de una edad determinada.

401. Por ejemplo, la diferencia entre el valor atribuido del número de hijos nacidos vivos y la edad de la madre podría ser como mínimo 12. En el caso de que así fuera, una matriz de atribución que utilizara grupos de edad de mujeres con intervalos de cinco años generaría con toda probabilidad información incompatible en algunos casos.

402. La matriz de atribución que se presenta en el gráfico 28 indica edades de mujeres en la parte superior y el número de hijos nacidos vivos en la columna de la izquierda. Los asientos son los valores atribuidos a la variable de los hijos vivos. A veces las respuestas serán apropiadas, pero otras veces no. Si el programa encuentra una mujer de 19 años con cinco hijos nacidos vivos, el valor de cinco hijos vivos probablemente superaría los criterios de diferencia de edad (una diferencia de edad de 15, sobre la base de los hijos vivos y la edad declarada). No obstante, para una mujer de 15 años de edad, ni los cinco hijos nacidos vivos (diferencia de edad de 10) ni los cuatro hijos vivos (diferencia de edad de 11) serían aceptables.

Gráfico 28

Valores iniciales para determinar los hijos vivos cuando los asientos de edad de la madre e hijos nacidos vivos son válidos

Hijos nacidos vivos	Edad												
	15	16	17	18	19	20	21	22	23	24	25-29	30-34	35+
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	0	0	0
2		2	2	2	2	2	2	2	2	2	1	1	1
3			3	3	3	3	3	3	3	3	2	2	2
4					4	4	4	4	4	4	3	3	3
5					5	5	5	5	5	5	4	4	4

403. La matriz de atribución será mejor cuando la edad de las mujeres jóvenes se consigne con intervalos de un año. En este caso sólo se incluirían en la matriz de atribución las respuestas de diferencia de edad válidas para esa edad concreta y sólo podrían tomarse de dicha matriz respuestas válidas.

f) Procedimiento de revisión cuando se han declarado valores para los hijos nacidos vivos, los hijos vivos y los hijos fallecidos

404. El valor de “hijos nacidos vivos” deberá ser la suma de los “hijos vivos” y los “hijos fallecidos”. Toda incongruencia podrá resolverse como se explica a continuación.

i) *Cuando se han declarado los tres elementos*

405. Si constan los tres elementos de información, el programa deberá determinar:

- a) Si los tres elementos tienen coherencia interna, es decir, si el número de hijos nacidos vivos es igual a la suma de hijos vivos y de hijos fallecidos;
- b) Si cada uno de los tres elementos es congruente con la edad de la mujer;
- c) Si el número de hijos nacidos vivos es congruente con el número de hijos nacidos en el último año (o el último hijo nacido), si se ha recopilado este dato.

406. Si todos esos valores son congruentes, la revisión habrá concluido. Sin embargo, si hay alguna incongruencia habrá que resolverla en la revisión. Es posible que los tres elementos no presenten coherencia interna: por ejemplo, puede constar que una mujer ha tenido cinco hijos nacidos vivos pero sólo dos hijos vivos y dos fallecidos. El equipo de revisión deberá decidir qué variable tiene preferencia sobre el resto. En muchos casos puede que la mujer recuerde todos los hijos que ha tenido, aunque tal vez olvide el número exacto de los que han fallecido. En este caso el equipo de revisión puede decidir aceptar el número de hijos nacidos vivos y de hijos vivos y restar esos valores para obtener un nuevo valor congruente para los hijos fallecidos.

ii) *Cuando se han declarado dos elementos*

407. Puesto que la categoría de hijos nacidos vivos (HNV) es la suma de los hijos vivos (HV) y de los hijos fallecidos (HF), si se dispone de dos de los tres elementos de información el programa informático podrá determinar la tercera variable:

- a) Si se conocen los valores de HNV y HV $\rightarrow HF = HNV - HV$
- b) Si se conocen los valores de HV y HF, $\rightarrow HNV = HV + HF$
- c) Si se conocen los valores de HNV y HF, $\rightarrow HV = HNV - HF$

Normalmente esas pruebas se efectuarán en primer lugar. Una vez que el programa haya determinado que los tres elementos de información son válidos y congruentes, la revisión habrá terminado.

iii) *Cuando se ha declarado un elemento*

408. Cuando sólo se conoce uno de los tres elementos, y si el país no utiliza la atribución dinámica, los valores de los otros dos elementos deberán consignarse como “desconocidos”. Si el país aplica técnicas de atribución dinámica, el equipo de revisión deberá determinar un método para obtener como mínimo un elemento más, y el tercero podrá obtenerse entonces mediante sustracción o adición. Puede utilizarse una matriz bidimensional para obtener el segundo valor de fecundidad, sobre la base del primer elemento y de una edad para la mujer, con intervalos de un año. Si se conoce el valor de los hijos nacidos vivos, por ejemplo, los hijos vivos pueden tomarse de la matriz de atribución, como se ha descrito, y por sustracción podrá obtenerse el valor de los hijos fallecidos. Análogamente, si se conoce el valor de los hijos vivos, el de los hijos nacidos vivos se obtendrá de una matriz de atribución en la que figuren la edad de la mujer con intervalos de un año y los hijos vivos, y el número de hijos fallecidos se obtendrá por sustracción.

iv) *Cuando no se ha declarado ningún elemento*

409. Cuando no se disponga de ninguno de los tres elementos, el equipo de revisión deberá decidir la forma de proceder. Si el país no aplica técnicas de atribución

dinámica, todos los valores deberán consignarse como “desconocidos”, y no deberán utilizarse en los métodos indirectos para estimar la mortalidad o la fecundidad. En los países que aplican técnicas de atribución dinámica, los especialistas deberán decidir si desean utilizarla para todos los elementos.

410. Si los especialistas deciden utilizar la atribución dinámica, el valor de los hijos nacidos vivos puede obtenerse sobre la base de la edad de la mujer, con intervalos de un año, y por lo menos una característica más. También es posible utilizar una matriz unidimensional que contenga la edad de la mujer, con intervalos de un año. Las otras características podrían ser elementos como el nivel de instrucción o la religión.

411. Una vez que se haya determinado el primer elemento, para obtener el segundo elemento de fecundidad la revisión puede seguir los pasos indicados para el caso en que se ha declarado un solo elemento. Entonces el tercer elemento podrá obtenerse a partir de los dos primeros. Los tres elementos deberían ser compatibles, dado que las matrices de atribución deberían actualizarse sólo cuando todos los elementos fueran compatibles. El valor de fecundidad obtenido también debería ser compatible con el de otras mujeres de la zona geográfica, ya que la información procedente de esas mujeres se utiliza para actualizar la matriz de atribución.

g) Procedimiento de revisión cuando se han declarado valores para los hijos nacidos vivos, los hijos que viven en el hogar, los hijos que viven fuera del hogar y los hijos fallecidos

i) *Cuando se han declarado los cuatro elementos*

412. Si se conocen los cuatro elementos de información, el programa deberá determinar:

- a) Si los cuatro elementos tienen coherencia interna, de manera que el número de hijos nacidos vivos sea igual a la suma de hijos que viven en el hogar, los hijos que viven fuera del hogar y los hijos fallecidos;
- b) Si cada uno de los cuatro elementos es congruente con la edad de la mujer;
- c) Si el valor de los “hijos nacidos vivos” es congruente con el de los “hijos nacidos en el último año” (o el último hijo nacido), si se ha recopilado este dato.

413. Si todos esos valores son congruentes, la revisión habrá concluido. Sin embargo, si hay alguna incongruencia habrá que resolverla en la revisión. Como en el caso de los tres elementos descritos anteriormente, es posible que los cuatro elementos no presenten coherencia interna. También en este caso el equipo de revisión deberá decidir qué variable tiene preferencia sobre las demás. En muchos casos es probable que la mujer declarante recuerde todos los hijos que ha tenido, aunque tal vez olvide algunos de los que se han ido del hogar o el número exacto de los que han fallecido. En este caso el equipo de revisión puede decidir aceptar el número de hijos nacidos vivos y de hijos vivos (la suma de los hijos que viven fuera del hogar y de los hijos que viven en el hogar) y restar esos valores para obtener nuevos valores congruentes para otras variables. Es posible que el equipo de revisión deba construir algoritmos para diversas combinaciones de casos.

ii) *Cuando se han declarado tres de los cuatro elementos*

414. El valor correspondiente a los hijos nacidos vivos (HNV) es la suma de los hijos que viven en el hogar (HVH) y de los hijos que viven fuera del hogar (HVFH) y de los hijos fallecidos (HF). Si se dispone de tres de los cuatro elementos de información, el programa informático podrá determinar la cuarta variable:

Si se conocen los valores de HNV, HVH y HVFH \rightarrow HF = HNV-HVH-HVFH

Si se conocen los valores de HVH, HVFH y HF \rightarrow HNV = HVH+HVFH+HF

Si se conocen los valores de HNV, HVH y HF \rightarrow HVFH = HNV-HVH-HF

Si se conocen los valores de HNV, HVFH y HF \rightarrow HVH = HNV-HVFH-HF.

iii) Cuando se han declarado dos de los cuatro elementos

415. Si sólo se conocen dos de los elementos, el equipo de revisión deberá decidir la forma de proceder. Por ejemplo, en muchos países las mujeres no declaran el número exacto de hijos que han fallecido. El otro elemento que con mayor probabilidad puede haberse omitido es la información sobre los hijos que viven fuera de la unidad de habitación, que es un valor que normalmente tampoco puede obtenerse directamente. Por lo tanto, al formular el cuestionario y al realizar el empadronamiento y el procesamiento hay que procurar obtener datos de la máxima calidad para todos los elementos de fecundidad.

416. Los datos sobre los hijos que viven en el hogar (HVH) pueden obtenerse sumando el número de hijos presentes en la unidad de habitación. En tanto que sólo una mujer de la unidad tenga la relación apropiada, un simple cálculo proporcionará el número de hijos que viven en la unidad. Si más de una mujer tiene esa relación, todavía podría utilizarse el programa de revisión, suponiendo que al recopilar los datos los hijos seguirán inmediatamente a la madre. Cuando no haya más remedio, los países que aplican técnicas de atribución dinámica podrán asignar un valor al número de hijos que viven en el hogar a partir de la edad de la madre y de una de las demás variables conocidas. (Véanse las normas generales que figuran más adelante para atribuir valores a distintos elementos de fecundidad a partir de otros elementos y de la edad de la madre.) Siempre que sea posible es importante que conste la edad de la madre con intervalos de un año, así como un número único de niños nacidos vivos, que viven en el hogar, que viven fuera del hogar o fallecidos.

417. Por ejemplo, los valores de hijos nacidos vivos y de hijos fallecidos pueden ser válidos, pero los correspondientes a hijos que viven en el hogar e hijos que viven fuera de él pueden ser inválidos. En ese caso el número de hijos que viven en el hogar puede determinarse sumando los hijos que tengan la relación apropiada con la madre (suponiendo que el jefe del hogar sea la madre). Así se dispondrá de tres de los cuatro elementos; el cuarto (los hijos que viven fuera del hogar) podrá sacarse por sustracción: $HV FH = HNV - HVH - HF$.

418. Sin embargo, cuando sólo se conocen dos de los elementos es más probable que sea necesario recodificar el valor de los hijos nacidos vivos y los hijos que viven en el hogar. Normalmente las mujeres declaran sin problemas el número de hijos nacidos vivos, y la información sobre los hijos que viven en el hogar puede obtenerse por observación o colaborando con los declarantes al realizar el empadronamiento, si bien esas opciones no existen en el caso de los hijos que viven fuera del hogar o los hijos fallecidos. Entonces el programa de revisión puede utilizar una matriz de atribución con la edad de la mujer y los hijos nacidos vivos (HNV) o, aún mejor, la edad de la mujer, los hijos nacidos vivos (HNV) y los hijos que viven en el hogar (HVH). Las variables permitirán obtener información de una mujer similar con las mismas características respecto de los hijos que viven fuera del hogar (HV FH).

419. Los países que utilizan sólo una matriz bidimensional para la edad de la mujer y los hijos nacidos vivos (HNV) sin incluir también la tercera dimensión —los hijos que viven en el hogar (HVH)— pueden obtener para los hijos que viven fuera del hogar (HV FH) un valor que no sea compatible con los otros dos. Por ejemplo, si la

edad de la mujer es 25 y el valor de HNV es 5, podría obtenerse de la matriz de atribución un valor de 3 para los hijos que viven fuera del hogar. Si el valor correspondiente a los hijos que viven en el hogar es 2, la revisión no presentará problemas. El valor para los hijos fallecidos debería ser 0, y los elementos de fecundidad deberían ser: HNV = 5, HVH = 2, HVFH = 3, HF = 0.

420. Sin embargo, el valor de los hijos que viven en el hogar podría ser en realidad 4, y utilizarse únicamente la edad de la mujer y los hijos nacidos vivos para determinar el valor de los hijos que viven fuera del hogar. En ese caso el valor de 3 para los hijos que viven fuera del hogar produciría una incompatibilidad entre los elementos. El valor de los hijos nacidos vivos (5) sería inferior a la suma de los hijos vivos (4 en el hogar y 3 fuera del hogar, es decir, un total de 7). Por lo tanto, habría que utilizar una matriz tridimensional: para 5 HNV y 4 HVH, el valor de la matriz de atribución podría ser de 1 para los hijos que viven fuera del hogar (para los hijos fallecidos habría que determinar por sustracción el valor de 0). O, por el contrario, el valor de la matriz de atribución debería ser 0 para los hijos que viven fuera del hogar (y para los hijos fallecidos habría que determinar por sustracción el valor de 1). Es necesario elaborar matrices de atribución similares para los otros pares de información conocidos, como en el gráfico 29.

Gráfico 29

Ejemplos de matrices de atribución que pueden elaborarse para pares de elementos de información conocidos

Si se conocen los elementos correspondientes a:		Puede utilizarse la atribución dinámica para uno de los siguientes (y a continuación sumar o restar)	
Hijos nacidos vivos	Hijos que viven en el hogar	Hijos que viven fuera del hogar	Hijos fallecidos
Hijos nacidos vivos	Hijos que viven fuera del hogar	Hijos que viven en el hogar	Hijos fallecidos
Hijos nacidos vivos	Hijos fallecidos	Hijos que viven en el hogar	Hijos que viven fuera del hogar
Hijos que viven en el hogar	Hijos que viven fuera del hogar	Hijos nacidos vivos	Hijos fallecidos
Hijos que viven en el hogar	Hijos fallecidos	Hijos nacidos vivos	Hijos que viven fuera del hogar
Hijos que viven fuera del hogar	Hijos fallecidos	Hijos nacidos vivos	Hijos que viven en el hogar

421. En cada caso se dispone de dos de los cuatro elementos. El tercero se obtiene por atribución dinámica, y el cuarto por adición o sustracción. Los equipos de revisión deberán decidir la mejor forma de proceder atendiendo a las circunstancias culturales.

iv) Cuando sólo se ha declarado un elemento

422. Cuando sólo se conoce uno de los cuatro elementos, la situación es aún más difícil. Los países deberán decidir la forma de proceder cuando falta tanta información. Si se aplican técnicas de atribución dinámica, como ya se ha indicado, la primera matriz de atribución utilizaría un elemento como la edad de la mujer, con intervalos de un año, y el otro elemento conocido para crear una matriz bidimensional que permita atribuir valores a cualquiera de los demás elementos. Una vez que se hayan determinado dos elementos, los otros dos seguirán siendo desconocidos, por definición. Por lo tanto, seguir utilizando la atribución dinámica para el tercer elemento no debería crear

una incompatibilidad con los demás elementos, ya que son desconocidos. El método que se ha indicado para el caso en que se conocen dos elementos y se desconocen otros dos se utilizará para obtener un tercer elemento. Entonces el cuarto elemento puede obtenerse por sustracción. Los cuatro elementos deberían ser compatibles.

v) *Cuando no se ha declarado ningún elemento*

423. Cuando no se dispone de ninguno de esos cuatro elementos, el equipo de revisión deberá decidir la forma de proceder sin conocer ninguno de los elementos. Si el país no aplica técnicas de atribución dinámica, todos los valores deberán consignarse como “desconocidos”, y no deberán utilizarse en métodos indirectos para estimar la mortalidad o la fecundidad. En los países que aplican técnicas de atribución dinámica, los especialistas deberán decidir si desean utilizarla para todos los elementos.

424. Si los especialistas deciden utilizar la atribución dinámica, los valores de los hijos nacidos vivos pueden obtenerse sobre la base de la edad de la mujer, con intervalos de un año, y por lo menos una característica más. También es posible utilizar una matriz unidimensional que contenga la edad de la mujer, con intervalos de un año. Las otras características podrían ser elementos como el nivel de instrucción o la religión, ya que se sabe que en muchos países la fecundidad presenta diferencias según los distintos niveles de instrucción o la confesión religiosa.

425. Una vez que se haya determinado el primer elemento, el método que se ha descrito para el caso en que sólo se conoce un elemento podrá aplicarse para obtener el segundo elemento de fecundidad. Entonces el tercer elemento podrá obtenerse a partir de los dos primeros, y el cuarto podrá obtenerse por sustracción. Los cuatro elementos deberían ser compatibles, dado que las matrices de atribución deberían actualizarse sólo cuando todos los elementos fueran compatibles. El valor de fecundidad obtenido también debería ser compatible con el de otras mujeres de la zona geográfica, ya que la información procedente de esas mujeres se utiliza para actualizar la matriz de atribución.

h) *Caso especial en que hay cinco o más elementos*

426. Dado que la migración ha adquirido una importancia creciente en algunos pequeños países, se está recopilando información adicional sobre los hijos. Cuando la variable “hijos que viven fuera del hogar” se divide en “hijos que viven fuera del hogar pero en el mismo país” e “hijos que viven fuera del hogar en otro país”, los procedimientos para cuatro variables —en el hogar, fuera del hogar, fallecidos y total— deben ampliarse a fin de tener en cuenta esa información adicional. Además, como ya se ha señalado, conviene tener una única matriz por cada edad de una mujer, con información completa sobre la fecundidad cuando todos los elementos sean válidos, internamente congruentes y en consonancia con su edad; luego, en los casos en que la información sobre la fecundidad sea incongruente (en particular con la edad), se puede recurrir a la atribución utilizando la línea completa pertinente.

i) *Importancia de que un único donante sea la fuente de todos los elementos de fecundidad*

427. De ser posible, es muy importante atribuir a todos los elementos valores procedentes de una sola mujer cuando no se conozca ninguno de ellos. Para asegurarse de que toda la información procede de la misma fuente, tal vez sea necesario elaborar matrices de atribución que utilicen toda la información de fecundidad. En tal caso, las matrices de atribución sólo podrían actualizarse cuando el programa de revisión determinara que todos los elementos de fecundidad concuerdan. Como se ha

observado en el párrafo 426, es mejor no hacer una atribución de valores elemento por elemento, pero, cuando varios elementos presentan problemas, debe utilizarse toda la información de otra mujer.

j) **Relación de los hijos propios con los hijos que se encuentran en el hogar y los hijos supervivientes**

428. Cuando los países utilizan el método de los hijos propios para ayudar a comprobar la revisión de la fecundidad a medida que se elabora y aplica, la información relativa a los hijos en el hogar y en la matriz madre-hijo puede ayudar a valorar la fiabilidad de los resultados de la revisión. Dado que son muy pocos los países que utilizan este método como ayuda para la revisión, continúa siendo todavía meramente experimental; no obstante, los resultados parecen prometedores.

7. Fecundidad: fecha de nacimiento del último hijo nacido vivo y nacimientos en los 12 meses anteriores al censo

429. La información sobre los últimos nacimientos ayuda a presentar estimaciones de la fecundidad inmediatamente antes de un censo o encuesta. Hay un método consistente en obtener información sobre la fecha de nacimiento (día, mes y año) del último hijo nacido vivo y el sexo del hijo y determinar si el hijo está vivo o no. El segundo método trata de recopilar información sobre los nacimientos en los 12 meses anteriores al censo; es más fácil para los empadronadores y declarantes recurrir a este segundo método, ya que se necesita únicamente un “sí” o un “no”, en vez de una fecha exacta.

430. En la etapa de elaboración puede realizarse una estimación del número de hijos nacidos vivos en los 12 meses inmediatamente anteriores a la fecha del censo partiendo de la información sobre la fecha de nacimiento del último hijo nacido vivo (que luego se mantendría como una recodificación). Esta forma de proceder aporta datos más precisos que la información sobre el número de nacimientos de una mujer en los 12 meses inmediatamente anteriores a la fecha del censo para estimar las tasas actuales de fecundidad por edades y otras medidas de la fecundidad (Naciones Unidas, 2008, párrs. 2.188 a 2.191).

431. Cabe señalar que la información sobre la fecha de nacimiento del último hijo nacido vivo no produce datos sobre el número total de hijos nacidos vivos durante el período de 12 meses. Aun cuando no hubiera errores en la comunicación de los datos sobre el último hijo nacido vivo, este elemento permite averiguar el número de mujeres que han tenido por lo menos un hijo nacido vivo en el período de 12 meses, pero no el número de nacimientos, pues una pequeña proporción de mujeres habrá tenido más de un hijo en un año (Naciones Unidas, 2008, párr. 2.189).

432. Esta información sólo debe reunirse para las mujeres de edades comprendidas entre los 15 y los 50 años que hayan manifestado haber tenido por lo menos un hijo nacido vivo durante su vida. Además, la información debe reunirse para todas las categorías de estado civil de las mujeres respecto de las cuales se reúnan datos sobre hijos nacidos vivos, por sexo. Si los datos sobre fecundidad durante toda la vida se obtienen únicamente para una muestra de mujeres, la información sobre la fecundidad actual también debe obtenerse de la misma muestra (Naciones Unidas, 2008, párr. 2.190).

433. En el programa de revisión deberían incluirse las comprobaciones siguientes. Debería consignarse la fecha de nacimiento del último hijo en el caso de todas las mujeres cuya edad esté entre un mínimo y un máximo definidos por el país. El programa debería comprobar que hubiera una correspondencia. Por ejemplo, no debería haber información sobre personas de sexo masculino ni sobre personas de sexo feme-

nino que no pertenecieran al grupo de edad seleccionado. Además, las mujeres del grupo de edad seleccionado con un valor superior a cero deberían tener un valor válido para el día, mes y año del último hijo nacido vivo (o una indicación de si tuvo lugar un nacimiento en los 12 últimos meses, si se utiliza esa pregunta).

434. El equipo de revisión deberá decidir si el día y el mes deben ser válidos; los equipos de revisión que apliquen técnicas de atribución dinámica pueden atribuir el día y el mes cuando estos falten; los que no utilicen ese método asignarán un valor “desconocido” al día y al mes. Si los especialistas temáticos, normalmente demógrafos, necesitan la edad efectiva de las madres en el momento del nacimiento de sus hijos como recodificación para los análisis de fecundidad, deberá atribuirse como mínimo el mes del último nacimiento, si no está presente. Luego puede obtenerse la recodificación.

435. De la misma manera, algunos demógrafos desean analizar los meses desde el último nacimiento. La revisión del año y del mes del último nacimiento ofrece la información necesaria para obtener los meses transcurridos desde el último nacimiento. Cuando se recopila también el día del último nacimiento se puede utilizar para determinar la recodificación de los meses transcurridos desde el último nacimiento (véase en el apartado 5 de la sección B del anexo I un método para obtener una recodificación de los meses desde el último nacimiento).

436. Si no hay información sobre el año de nacimiento del último hijo nacido vivo, o si esa información es inválida, los países que no aplican técnicas de atribución dinámica pueden asignar un valor “no declarado” o “desconocido”. Los países que aplican la atribución dinámica pueden utilizar otras variables como la edad y el número de hijos nacidos vivos para obtener la fecha de nacimiento del último hijo.

437. Dada la importancia del uso de la fecha del último nacimiento para tener una orientación sobre la fecundidad nacional, regional y local reciente, deberán considerarse comprobaciones adicionales. Una revisión útil implica la comprobación dentro del hogar de uno o varios hijos de cero años de edad, y la utilización de las relaciones de la madre y ese hijo o esos hijos (o el número de persona de la madre correspondiente al hijo, si se recopila) para determinar si el hijo está registrado como último nacimiento de la madre. Las comprobaciones deberán hacerse en doble sentido: los hijos de cero años deberán comprobarse con relación a las madres, y los últimos nacimientos, con respecto a la enumeración del hogar.

438. Los países que recopilan también los fallecimientos en el año anterior al censo o la encuesta quizá decidan incluir una comprobación de los fallecimientos de hijos de cero años de edad en el año anterior al censo con respecto a los últimos nacimientos, cuando el último nacimiento se registra como “fallecido” o bajo alguna otra fórmula semejante. Si bien esta comprobación no será posible si la madre no está en la casa por fallecimiento o traslado o si, por alguna razón, el hijo no está registrado, sí sería posible comprobar de esa manera una proporción de fallecimientos de niños.

8. Fecundidad: edad de la madre en el momento del nacimiento del primer hijo nacido vivo

439. La edad de la madre en el momento del nacimiento del primer hijo nacido vivo se utiliza para estimar indirectamente la fecundidad sobre la base de los primeros nacimientos y para suministrar información sobre el comienzo de la reproducción. Si el tema se incluye en el censo, debe obtenerse información sobre cada mujer que haya tenido por lo menos un hijo nacido vivo (Naciones Unidas, 2008, párr. 2.193).

440. La edad de la madre en el momento del nacimiento del primer hijo nacido vivo se determina directamente con un elemento concreto, “edad al nacimiento del pri-

mer hijo nacido vivo”, o por la diferencia entre la edad actual de la madre y la edad del hijo mayor, si la edad de él se conoce. La edad más temprana para tener hijos definida por el país no corresponde a la edad biológica más temprana. Si, por ejemplo, la edad más temprana aceptable en un país para el primer nacimiento es de 13 años, es posible que los declarantes comuniquen o que los empadronadores registren para una persona determinada una edad de 11 o 12 años. Entonces el equipo de revisión deberá decidir si hay que modificar la edad más temprana aceptable, suprimir el nacimiento o modificar la edad de la madre o la edad en el momento del nacimiento del primer hijo (utilizando la edad del hijo o la edad de la madre, según las variables utilizadas para determinar la diferencia de edad). Análogamente, el equipo de revisión deberá decidir qué “edad máxima” se aceptará como edad de la madre en el momento del nacimiento del primer hijo. Aunque las mujeres pueden tener hijos después de haber cumplido los 50 años, esto no sucede con mucha frecuencia, de manera que, a fin de corregir las equivocaciones, en la revisión deberá determinarse si los casos que están fuera de los límites son auténticos.

441. Es importante recordar que la edad más temprana o la edad más tardía en el momento del nacimiento del primer hijo (y la diferencia de edades entre la madre y el hijo mayor presente en el hogar) deben ser acordes con las costumbres y tradiciones del país. Los especialistas temáticos deberán decidir cuándo un valor determinado se considerará “ruido” y no una edad correcta de la madre en el momento del primer nacimiento. Cuando se establezcan las normas, los especialistas deberán decidir cómo resolver el problema. Si no se utiliza la atribución dinámica, el programa debería asignar un valor “desconocido”. Cuando se utiliza la atribución dinámica, la edad de la madre en el momento del nacimiento del primer hijo puede determinarse recurriendo a otras mujeres de edad similar y con un número similar de hijos nacidos vivos. Los especialistas que construyan la matriz de atribución tal vez deseen tener en cuenta factores como la residencia en zona urbana o rural (si hay diferencias en cuanto a fecundidad entre las dos zonas), la presencia de la mujer en la fuerza de trabajo (aunque la situación actual en cuanto al empleo no sea necesariamente la misma que la situación en el momento de tener el primer hijo) y el nivel de instrucción.

9. Mortalidad

442. La información sobre las defunciones en los últimos 12 meses se usa para estimar el nivel y la estructura de la mortalidad por sexo y por edad en los países sin estadísticas continuas satisfactorias sobre defunciones a partir del registro civil. Para que la estimación derivada de ese elemento resulte fiable es importante que las defunciones en los últimos 12 meses se registren lo más completa y exactamente posible por sexo y por edad. La inclusión generalizada de preguntas sobre la mortalidad en el cuestionario del censo ha permitido en los últimos decenios mejorar la utilización de procedimientos de estimación indirecta para averiguar la mortalidad adulta (Naciones Unidas, 2008, párr. 2.194).

443. Lo ideal sería determinar la mortalidad en cada hogar por el número total de defunciones en el último período de 12 meses antes de la fecha del censo. Cuando no sea posible obtener información sobre las defunciones de los últimos 12 meses es aconsejable obtener al menos información sobre defunciones de niños menores de un año. Respecto de cada defunción que se haya comunicado también debe reunirse información sobre el nombre, la edad, el sexo y la fecha (día, mes y año) de fallecimiento de la persona. Hay que indicar claramente el período de referencia al declarante para evitar errores por una mala interpretación. Por ejemplo, en cada país podría definirse un período de referencia preciso en relación con una festividad o una fecha histórica pertinente (Naciones Unidas, 2008, párr. 2.195).

a) Edad y sexo de la persona fallecida

444. En *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008) se sugiere obtener el nombre, la edad y el sexo y el día, mes y año de defunción de las personas que fallecieron en el año anterior al censo. Los países que no utilicen la atribución dinámica pueden asignar un valor “desconocido” a cada una de esas variables cuando sean inválidas. Los países que apliquen técnicas de atribución dinámica pueden utilizar la edad (en grupos de edad), el sexo y el año de defunción como dimensiones de las matrices de atribución para las demás variables. En la práctica, es probable que las matrices de atribución sean específicas para cada país, y el equipo de revisión deberá colaborar para construir matrices de atribución apropiadas. La estructura de población del país, o los niveles geográficos subnacionales, pueden ser utilizados para efectuar la revisión más indicada.

b) Causa del fallecimiento

445. Algunos países están recopilando información sobre la causa del fallecimiento en las defunciones ocurridas en los 12 meses anteriores al censo. Por tratarse de una cuestión delicada, y debido a la dificultad existente en algunos casos para obtener información sobre el terreno, es posible que los países hagan la siguiente pregunta: ¿Se debió la muerte a un accidente o a actos de violencia?, con el fin de obtener información indirecta sobre el VIH/SIDA en relación con grupos de edad determinados. La revisión de este elemento implicará normalmente el supuesto de que, si la información no se ha recopilado o es inválida, normalmente se atribuirá un valor “desconocido”. Si un país opta por recurrir a la atribución, sería pertinente una atribución dinámica que utilizara el sexo y grupos de edad de 0, 1 a 4 y, luego, 5 años.

c) Mortalidad materna

446. En la ronda actual de censos, cada vez son más los países que preguntan también si la persona fallecida era una mujer, y si estaba embarazada en el momento de su fallecimiento. Esta información contribuye a determinar la mortalidad materna tanto en el plano nacional como en el regional. La revisión de este elemento podría suponer la atribución del valor “desconocido” en las entradas inválidas o en blanco. No obstante, si un país opta por la atribución, la atribución dinámica se aplicaría sólo a las mujeres, obviamente, y sólo en las edades probables de embarazo —normalmente entre 12 y 54 años— y, probablemente, por años concretos, en vez de por grupos de edad de cinco años.

d) Mortalidad infantil

447. Finalmente, en *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (párr. 2.191) se propone la recopilación de información sobre los fallecimientos de los niños nacidos “en los 12 últimos meses”. Normalmente, esta pregunta se formularía sólo conjuntamente con la relativa a los nacimientos en los 12 meses anteriores al censo. Si se utilizara el otro elemento relacionado con la fecundidad, la fecha del último nacimiento, es probable que no se emplea este elemento.

448. Los datos sobre la relación entre fallecimientos y nacimientos en el año anterior al censo ayudan a los países con registros civiles de calidad a comprobar sus tasas de mortalidad infantil, tanto en el plano nacional como en el regional. Los países que no tienen un registro civil de buena calidad pueden utilizar esa información para obtener estimaciones de la mortalidad infantil. También en este caso una comproba-

ción de los últimos nacimientos que fallecieron en el año anterior al censo en comparación con los fallecimientos de los niños de cero años de edad en ese año es un método adecuado y podría facilitar información útil sobre la mortalidad infantil.

449. Las revisiones de este elemento requieren cierta reflexión y suelen ser específicas de cada país. En teoría, la información sobre los niños nacidos vivos y sobrevivientes puede utilizarse para comprobar la información registrada; cuando hay una sola mujer adulta en el hogar, la comprobación es relativamente fácil. Cuando hay varias mujeres, habrá que tener cuidado en asociar a cada niño con la mujer pertinente.

10. Orfandad de madre o de padre y número de línea de la madre

450. Para reunir información sobre la situación de orfandad de los miembros de la población deben formularse dos preguntas directas: *a)* si la madre natural de la persona empadronada en el hogar todavía vive en el momento del censo y *b)* si el padre natural de la persona empadronada en el hogar todavía vive en el momento del censo, independientemente de que la madre y el padre estén empadronados en el mismo hogar. La investigación debe aportar información sobre los padres biológicos. Por lo tanto, hay que poner cuidado en excluir a los padres adoptivos. Dado que suele haber más de un hijo superviviente que responderá a esas preguntas, es necesario que las preguntas estén formuladas de manera que eviten las duplicaciones que supondrán las comunicaciones de hermanos. Para este fin deben hacerse otras dos preguntas: *c)* si el encuestado es el hijo mayor superviviente de su madre y *d)* si el encuestado es el hijo mayor superviviente de su padre. Las tabulaciones deberán realizarse refiriéndose sólo al hijo mayor superviviente (Naciones Unidas, 2008, párr. 2.199).

451. Los procedimientos de revisión respecto de los elementos “madre con vida” y “número de línea de la madre” están interrelacionados y deben aplicarse conjuntamente. Respecto de las personas que no hayan respondido “sí” a la pregunta de si la madre vive, habrá que comprobar si el número de línea de la madre es un asiento válido; de ser así, habrá que asignar el código correspondiente a “sí” al elemento “madre con vida”. Respecto de las personas que no hayan respondido “sí” a la pregunta de si la madre vive, habrá que comprobar el número de línea de la madre para ver si es 00 o si equivale al número de línea de una mujer de una edad superior o igual a 12 años. Si se da cualquiera de esas situaciones, el programa supondrá que la persona tiene una madre y asignará el código correspondiente a “sí” al elemento de “madre con vida”. Si el asiento que figura en el número de línea de la madre no es válido y el elemento correspondiente a “madre con vida” tiene el código correspondiente a “no” o “no sabe”, habrá que eliminar el asiento del número de línea de la madre. En los demás casos habrá que asignar el código correspondiente a “no sabe” al elemento “madre con vida” y habrá que eliminar todo asiento en el número de línea.

452. El país puede decidir no revisar el número de línea de la madre en el caso de las personas que hayan contestado “no” o “no sabe” a la pregunta de si la madre está con vida. En todos los demás casos podría comprobarse el número de línea para ver si es congruente, o podría atribuirse un valor utilizando la relación de la persona y el número de línea, el sexo, la relación y la edad de la persona que consta como madre. Cuando haya discrepancias o no pueda determinarse si una persona es la madre, podría asignarse el código correspondiente a “vive fuera del hogar”. Conviene señalar que en las revisiones estructurales, si el jefe no era la primera persona y se le coloca ahora en esa posición, el número de línea de esa madre quizá deba ajustarse para una o más personas.

B. Características de migración

453. Las características demográficas de un país cambian con el tiempo como consecuencia del incremento natural (fécundidad y mortalidad) y de la migración neta. Puede tratarse de una migración de largo plazo (desde el nacimiento) o una migración de corto plazo, medida sobre la base de la residencia anterior y su duración o bien en un momento previo especificado. Puesto que estos elementos a menudo están interrelacionados, en algunos países podría ser apropiado un procedimiento de revisión conjunto similar al descrito para las variables demográficas básicas. Si se aplica un método secuencial, el orden de las revisiones será importante, ya que ciertos elementos deben ser revisados antes que otros.

454. A menudo los elementos de migración requieren códigos más detallados que otros elementos, ya que pueden requerirse unidades geográficas más pequeñas para la planificación y la formulación de políticas. Es posible que el personal encargado de planificar una nueva escuela o un centro de salud necesite información detallada sobre una zona pequeña. Por otra parte, pueden requerirse diferentes sistemas de codificación y diferentes procedimientos de revisión para lugares que están dentro o fuera del país.

455. Tradicionalmente, la mayoría de los países no tuvieron una inmigración internacional muy numerosa, por lo que hicieron hincapié en la migración interna, que continúa siendo la preocupación principal. No obstante, en un mundo cada vez más globalizado, cada vez se insiste más también en la migración internacional.

456. Por lo que se refiere a la migración dentro del país (migración interna), debe comprobarse la congruencia de los datos sobre el país de nacimiento y los años que se ha vivido en el distrito, ya que existen relaciones obvias entre esos elementos. Por otra parte, existen ciertas relaciones razonables entre las respuestas de distintos miembros del hogar. Por ejemplo, si no hay respuesta en relación con el número de años que se ha vivido en el distrito en el caso de un hijo, el valor puede atribuirse a partir de la respuesta de la madre, y el programa de revisión deberá comprobar que el valor atribuido no sea superior a la edad del hijo.

457. Desde el punto de vista de la migración internacional son importantes el país de nacimiento y el año de entrada en el país.

1. Lugar de nacimiento

458. En primera instancia, se entiende por lugar de nacimiento la división administrativa en que nació la persona. Sin embargo, hay que señalar que el país de nacimiento no está relacionado necesariamente con la nacionalidad legal, que es algo separado (véase Naciones Unidas, 2008, párrs. 2.92 a 2.102). Para las personas nacidas en el país en el que se levanta el censo (nativas), el concepto de lugar de nacimiento incluye también la unidad geográfica específica del país en que residía la madre de la persona en el momento del nacimiento de esta. Sin embargo, en algunos países el lugar de nacimiento de los nativos se define como la unidad geográfica en que se produjo realmente el nacimiento. Cada país deberá explicar qué definición ha utilizado en el censo (Naciones Unidas, 2008, párr. 2.57).

a) Relación entre los asientos correspondientes al país de nacimiento y a los años que se ha vivido en el distrito

459. Puede comprobarse la congruencia entre los asientos correspondientes al lugar de nacimiento y a la duración de la residencia, ya que existen fuertes relaciones

entre los dos elementos. También existen relaciones entre los distintos miembros de un hogar y pueden hacerse suposiciones a partir de otros miembros de la familia sobre si la persona en cuestión ha emigrado.

b) Asignación de un valor “desconocido” a los asientos inválidos del lugar de nacimiento

460. Si un país decide no aplicar técnicas de atribución dinámica, todo asiento inválido en la variable del lugar de nacimiento deberá convertirse en “desconocido”. Normalmente el país no debería revisar las respuestas incongruentes entre miembros de una familia o en relación con zonas geográficas a menos que la codificación esté equivocada.

c) Utilización de la atribución estática para el valor del lugar de nacimiento

461. El asiento correspondiente al país de nacimiento sólo debería modificarse si quedara fuera de los límites. Si el código correspondiente a la duración de la residencia es “siempre”, habría que asignar el código del país. Si el valor es distinto de “siempre”, puede utilizarse la información de una persona anterior. Por ejemplo, si una persona anterior es la madre, el número de años que la madre ha vivido en el distrito podría compararse con la edad de la persona. Si el asiento de la madre tiene un valor superior o igual a la edad de la persona, habría que asignar el código correspondiente a “este país”; de otro modo, habría que asignar el valor del “país de nacimiento de la madre”. Si el país de nacimiento no puede asignarse sobre la base de los datos de la madre, pueden utilizarse del mismo modo los de otros familiares. Si después de esas pruebas todavía no es posible asignar un valor, el país de nacimiento puede consignarse como “desconocido”.

462. Dado que los países presentan ahora al público muestras de sus datos, es importante facilitar un código específico durante la revisión de los elementos que aparecen en blanco ya que son eludidos por los empadronadores que están siguiendo las pautas de omisión. Es decir, en muchos casos el cuestionario indica al empadronador que eluda la pregunta sobre el lugar de nacimiento si la persona vivió siempre en ese lugar. Durante la revisión, deberá asignarse el código correspondiente al lugar específico para ayudar a los usuarios, de manera que luego no tengan que mirar en dos lugares al hacer sus propias tabulaciones cruzadas.

d) Utilización de la atribución dinámica para el valor del lugar de nacimiento

463. Como antes, el asiento correspondiente al país de nacimiento sólo debería modificarse si queda fuera de los límites. Si el asiento correspondiente a la duración de la residencia es “siempre”, habrá que asignar al país de nacimiento el código correspondiente a “este país”. Si el valor es distinto de “siempre”, habrá que estudiar la información de otras personas del hogar para obtener pistas sobre el país de nacimiento de la persona.

e) Asignación de un valor del lugar de nacimiento cuando consta la madre de la persona

464. Si el asiento del país de nacimiento ha quedado en blanco o es inválido y la duración de la residencia es distinta de “siempre”, puede hacerse una búsqueda de la madre de la persona. Si se encuentra a la madre en el hogar, se examina el elemento

de la duración de la residencia de la madre. Si el valor correspondiente es “siempre”, puede asignarse al país de nacimiento de la persona el valor “este país”. Si la madre de la persona no ha vivido siempre en el distrito pero si la edad de la persona es inferior o igual al número de años que la madre ha vivido en el distrito, el programa también puede asignar a la variable del país de nacimiento el valor “este país”. Si la edad de la persona es superior al número de años que la madre ha vivido en el distrito y el país de nacimiento de la madre es válido, se asignará a la variable del país de nacimiento de la persona el mismo país de nacimiento que la madre.

f) Asignación de un valor del lugar de nacimiento al hijo del jefe del hogar

465. Si la madre de la persona no está en el hogar pero esa persona es un hijo o una hija del jefe del hogar, para obtener el lugar de nacimiento pueden efectuarse varias comprobaciones utilizando la información del registro del jefe del hogar. Si el valor correspondiente a los años que el jefe del hogar ha vivido en el distrito es “siempre”, el programa deberá asignar a la variable del país de nacimiento del registro de la persona el valor “este país”. Si la duración de la residencia del jefe del hogar en el distrito no es “siempre” pero la edad de la persona es inferior o igual al número de años que el jefe del hogar ha pasado en ese distrito, el programa deberá asignar el valor “este país” a la variable del país de nacimiento de la persona. Sin embargo, si la edad de esa persona es superior al número de años que el jefe del hogar ha pasado en este distrito, el programa deberá asignar el valor del país de nacimiento del jefe del hogar si este tiene un código válido para esa variable.

g) Asignación de un valor del lugar de nacimiento al hijo, pero no al jefe del hogar

466. Pueden atribuirse valores muy distintos según que una persona esté o no por encima o por debajo de una edad concreta (edad X) fijada por el equipo de revisión del país. Si la edad de una persona es inferior a X, el país de nacimiento deberá tomarse del primer registro previo de un hijo menor de X años, por edad y sexo.

h) Asignación de un valor del lugar de nacimiento a mujeres adultas que tienen marido

467. Si la persona tiene X años de edad o más y es de sexo femenino, el programa deberá comprobar si tiene un marido en el hogar. Si la mujer tiene un marido y este tiene un código válido para la variable del país de nacimiento, el programa deberá asignar al registro de la mujer el código del país de nacimiento del marido. Si el marido no tiene un código válido para el país de nacimiento, habrá que examinar su asiento correspondiente a la duración de la residencia. Si la duración de la residencia del marido en el distrito se ha codificado como “siempre”, habrá que asignar a la variable del país de nacimiento de la mujer el código de “este país”. Si la duración de la residencia del marido en el distrito no es “siempre”, habrá que atribuir un valor al país de nacimiento de la mujer, por edad y sexo.

i) Asignación de un valor del lugar de nacimiento a mujeres adultas que no tienen marido

468. Aunque una mujer mayor de una edad mínima fijada por el equipo de revisión no tenga un marido en el hogar, es posible que sea la madre de un hijo del hogar. En este caso el programa deberá buscar a su hijo mayor. Si no puede encontrar ese hijo, el programa podrá atribuir el valor del país de nacimiento por edad y sexo.

Si el hijo tiene un código válido para el país de nacimiento y la duración de la residencia de la madre en el distrito es mayor que la edad del hijo, el programa deberá atribuir el valor del país de nacimiento por edad y sexo. Pero si la duración de la residencia de la madre en el distrito es inferior o igual a la edad del hijo, el programa deberá asignar a la madre el país de nacimiento del hijo.

j) Asignación de un valor del lugar de nacimiento a personas de sexo masculino

469. Para obtener el lugar de nacimiento de una persona de sexo masculino, el programa de revisión puede intentar encontrar a su esposa; si se trata del jefe del hogar, el programa deberá intentar encontrar a sus hijos. Primero, el programa intenta encontrar a la esposa del hombre. Si la encuentra, y la duración de la residencia del hombre en el distrito es inferior o igual a la de ella, se asignará al registro del hombre el país de nacimiento de la esposa. Si la duración de la residencia del hombre en el distrito es superior a la de su esposa, habrá que asignar un valor al país de nacimiento utilizando una matriz de atribución por edad y sexo. Si el hombre es el jefe del hogar, tiene un hijo o una hija presentes en él y ha vivido en el distrito por un período igual o inferior a la edad del hijo, el programa deberá asignar el mismo país de nacimiento del hijo. Si el tiempo que ha vivido en el distrito es mayor que la edad del hijo, el programa deberá efectuar una atribución por edad.

2. Ciudadanía

470. Deben reunirse datos sobre la ciudadanía que permitan clasificar a la población del país en tres categorías: a) ciudadanos de nacimiento; b) ciudadanos por naturalización (por declaración, opción, matrimonio o de otra forma); y c) extranjeros. Además, habrá que reunir datos acerca del país de donde son ciudadanos los extranjeros. Es importante registrar el país de la ciudadanía como tal y no como gentilicio, ya que algunas de esas designaciones son las mismas que se utilizan para referirse a un grupo étnico.

471. Es esencial que la codificación de la información sobre el país de ciudadanía se efectúe con un detalle suficiente que permita la identificación individual de todos los países de ciudadanía que están representados entre la población extranjera del país. Se recomienda que con fines de codificación los países utilicen el sistema de códigos numéricos que se expone en los *Códigos uniformes de país o de zona para uso estadístico* (Naciones Unidas, 1999). La utilización de códigos normalizados para la clasificación de la población extranjera por país de ciudadanía mejorará la utilidad de esos datos y permitirá un intercambio internacional de información sobre la población extranjera entre los países. Si un país decide combinar los países de ciudadanía en grupos más amplios, se recomienda que se adopten las clasificaciones regionales y subregionales normalizadas que figuran en la publicación indicada arriba (Naciones Unidas, 2008, párr. 2.97).

a) Revisión del valor de la ciudadanía

472. La ciudadanía depende de las definiciones de cada país. En la mayoría de países, aunque no en todos, las personas nacidas en el país son automáticamente ciudadanos de nacimiento. Por tanto, en la revisión habrá que buscar la relación entre el lugar de nacimiento y la ciudadanía, y es posible que haya que asignar el valor de “ciudadanos de nacimiento” a las personas nacidas en el país.

b) Relación entre el grupo étnico/raza y la ciudadanía

473. Algunos países también reúnen datos sobre el “grupo étnico” o la “raza”, lo cual tal vez brinde información adicional que pueda usarse para determinar la ciudadanía, particularmente cuando la respuesta obtenida sea inválida. En muchos países, los migrantes de primera y de segunda generación deberían presentar una congruencia casi perfecta entre su origen étnico y su ciudadanía. En los países con una larga historia de inmigración internacional esta característica puede ser menos valiosa, pero todavía podría tenerse en cuenta junto con otras variables.

c) Relación entre la naturalización y la ciudadanía

474. En los países en los que existe la naturalización, los requisitos para esta pueden o no estar cubiertos en las preguntas del censo. Si, por ejemplo, se requiere un período de residencia, el elemento relativo a la “duración de la residencia” podrá utilizarse como prueba del cumplimiento del período de naturalización. En este caso, si una persona ha nacido en otro país y su registro presenta una respuesta inválida o incongruente con respecto a la ciudadanía, el equipo de revisión puede decidir asignar a esa variable el valor “naturalizado”. Otras personas que no cumplan los requisitos de duración de la residencia para la naturalización deberán recibir un valor de “extranjeros”, por aplicación del método de atribución estática.

d) Relación entre la duración de la residencia y la ciudadanía

475. Es posible que el elemento de “duración de la residencia” no aparezca en el cuestionario, que sea ambiguo para determinar la ciudadanía o que el equipo de revisión decida no utilizarlo. Si el valor de la ciudadanía es inválido o incongruente con el lugar de nacimiento, habrá que asignar un valor “desconocido” cuando no se utilice un método de atribución dinámica. Los países que apliquen la atribución dinámica para asignar valores a los elementos inválidos deberían utilizar por lo menos dos características (una de ellas debería ser el lugar de nacimiento) para obtener información “conocida” procedente de personas similares de la misma zona geográfica.

3. Duración de la residencia

476. La duración de la residencia es el intervalo de tiempo hasta la fecha del censo, expresado en años completos, durante el cual cada persona ha vivido en: *a)* la localidad que es su residencia habitual en el momento del censo, y *b)* la división administrativa mayor o división administrativa menor en que está situada esa localidad (Naciones Unidas, 2008, párr. 2.64).

a) Revisión del valor de la duración de la residencia

477. Al igual que el país de nacimiento, la duración de la residencia tiene importancia al compilar estadísticas sobre la movilidad de la población. En algunos casos un subgrupo de la población puede tener una movilidad mucho mayor que el conjunto de la nación. El procedimiento de revisión de este elemento tiene en cuenta el lugar de nacimiento de la persona y las respuestas de otros miembros de los hogares. La “duración de la residencia” deberá revisarse teniendo en cuenta el “lugar de residencia anterior” o el “lugar de residencia en una fecha especificada del pasado”.

b) Residencia *de facto/de jure* y duración

478. El procedimiento de revisión puede variar según se trate de un censo *de facto* o *de jure*. Debido a que el censo *de jure* reúne información en la residencia

habitual, es posible que la duración de la residencia no facilite la misma información que en un censo *de facto*, en el cual las personas son empadronadas en el lugar donde estaban en la noche del censo. Además, los códigos y las revisiones deberán tener en cuenta a las personas que hayan vivido “siempre” en el lugar o que “nunca lo han abandonado”. En el caso de esas personas, el programa deberá omitir las pruebas de coherencia y otras comprobaciones.

c) Relación entre la edad y la duración de la residencia

479. En la primera parte de la revisión deberá comprobarse que la relación entre la edad y el lugar de nacimiento es congruente, así como la existencia de un valor válido en el elemento relativo a los años que se ha vivido en la localidad o en la división administrativa. El número de años que una persona ha vivido en una localidad o en una división administrativa no puede ser superior a su edad, ni una persona que nació fuera del país no puede haber vivido siempre en la localidad o en la división administrativa. El programa debería asignar el valor “siempre” a los años que se ha vivido en la localidad o en la división administrativa si el valor correspondiente es superior a la edad y si el valor correspondiente a la variable del país de nacimiento es “este país”. Si el valor asignado a los años en la localidad o en la división administrativa es superior pero el país de nacimiento no es “este país”, el valor de la edad de la persona debería asignarse a los años de residencia en la localidad o en la división administrativa. En tal caso se supone que, aunque ha nacido fuera de este país, la persona se trasladó a la localidad o a la división administrativa cuando tenía menos de un año de edad.

d) Relación entre el lugar de nacimiento y la duración de la residencia

480. En el caso de los valores que queden fuera de los límites habrá que aplicar las mismas pruebas que en el caso del lugar de nacimiento. Puede hacerse una búsqueda de personas anteriores relacionadas (madre, jefe del hogar, marido, hijo). Las atribuciones deberían basarse en la información que se haya encontrado. Sin embargo, antes de asignar un valor hay que asegurarse de que este sea congruente con la edad y con el lugar de nacimiento de la persona cuyo registro se está modificando.

e) Personas que siempre han vivido en el lugar

481. Si la respuesta a la variable del número de años que una persona ha vivido en la localidad o en la división administrativa es “siempre” pero el país de nacimiento no es “este país”, el equipo de revisión tal vez desee asignar el valor de la edad de la persona a la duración de la residencia en la localidad o en la división administrativa. Los especialistas supondrán que, aunque nació fuera del país, la persona se trasladó a la localidad o a la división administrativa cuando tenía menos de un año de edad. En la parte siguiente de la revisión se comprobará la validez del asiento de los años de residencia en la localidad o en la división administrativa. Puesto que el período que una persona ha vivido en la localidad o en la división administrativa no puede ser mayor que la edad de esa persona, el valor de la edad se asignará a los años que se ha vivido en la localidad o en la división administrativa para esta situación.

f) Duración de la residencia de la persona a partir de la duración de la residencia de la madre

482. Si la categoría no tiene un código válido, el programa puede efectuar una comprobación entre registros buscando en el hogar a la madre de la persona. Si la encuentra, el registro de la madre puede suministrar información útil para asignar

los valores que faltan. Si la madre de la persona siempre ha vivido en la localidad o en la división administrativa y su país de nacimiento es “este país” (como debería ser), el programa asignará el valor “siempre” a la categoría de los años que ha pasado esta persona en la localidad o en la división administrativa. Si el país de nacimiento de la madre no es “este país”, aun cuando el valor de los años que se ha vivido en la localidad o en la división administrativa sea “siempre”, esto indica que hay alguna equivocación en los valores de la madre. En este caso el programa hará caso omiso del país de nacimiento de la madre y asignará el valor de la edad a la variable de la duración de la residencia en la localidad o en la división administrativa. Si en el asiento correspondiente a los años que la madre ha vivido en la localidad o en la división administrativa no consta “siempre” pero el código es válido, y si la edad de la persona es inferior al número de años que la madre ha vivido en la localidad o en la división administrativa, habrá que retroceder a un punto anterior del proceso de revisión y comprobar el país de nacimiento de la madre. Si el país de nacimiento de la madre es “este país”, el programa asignará el valor de la edad de la persona a los años vividos en la localidad o en la división administrativa. Sin embargo, si la edad de una persona es igual o superior a los años que la madre ha vivido en la localidad o en la división administrativa, el programa asignará el valor de los años que la madre ha vivido en la localidad o en la división administrativa a la variable de años que esa persona ha vivido en la localidad o en la división administrativa.

**g) Duración de la residencia de la persona
a partir de la duración de la residencia de un hijo**

483. Si la persona en cuestión es un hijo o una hija, el programa de revisión deberá comprobar el registro del jefe del hogar para encontrar alguna información que ayude a asignar valores a los datos que faltan respecto de la duración de la residencia. Cuando el jefe del hogar nació en “este país” y ha vivido siempre en la localidad o en la división administrativa, el programa asignará el valor de “siempre” a los años que el hijo ha vivido en la localidad o en la división administrativa. Cuando el jefe del hogar siempre haya vivido en la localidad o en la división administrativa pero no haya nacido en “este país”, se asignará el valor de la edad del hijo a los años que ha vivido en la localidad o en la división administrativa. Cuando en el asiento correspondiente a los años que el jefe del hogar ha vivido en la localidad o en la división administrativa no conste el valor “siempre” pero el código sea válido, esta información podrá utilizarse si es congruente con la edad que figura en el registro del hijo que se está revisando. Si la edad del hijo es igual o superior al número de años que el jefe del hogar ha vivido en la localidad o en la división administrativa, el programa utilizará los años que el jefe del hogar ha vivido en la localidad o en la división administrativa como valor de los años que el hijo ha vivido en la localidad o en la división administrativa. Si la edad del hijo es inferior al número de años que el jefe del hogar ha vivido en la localidad o en la división administrativa, el programa asignará un valor con arreglo al país de nacimiento del jefe del hogar. Este valor será “siempre” si el jefe del hogar nació en “este país”; de no ser así, el programa asignará el valor de la edad del hijo o de la hija a la variable del número de años en la localidad o en la división administrativa.

**h) Duración de la residencia de una persona
cuando no se disponga de otra información**

484. Cuando todos los métodos indicados no hayan podido producir un asiento válido, el programa puede asignar un valor “no declarado” o “desconocido” a los años que la persona ha vivido en la localidad o en la división administrativa. Si el

asiento sigue siendo inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características para obtener información “conocida” de personas similares de la misma zona geográfica.

4. Lugar de residencia anterior

485. Por lugar de residencia anterior se entiende la división administrativa mayor o menor, o el país extranjero, en que la persona residía inmediatamente antes de emigrar a la división administrativa en que ahora reside habitualmente (Naciones Unidas, 2008, párr. 2.67).

a) Revisión del valor del lugar de residencia anterior

486. El elemento “lugar de residencia anterior” deberá revisarse teniendo en cuenta la “duración de la residencia”. Si la persona nació en este lugar (país, localidad o división administrativa, según el elemento del censo) y nunca emigró, tal elemento debería dejarse en blanco o habría que asignar un código específico para indicar que “nunca emigró”. Sin embargo, los elementos en blanco pueden causar problemas durante la tabulación, por lo que el equipo de revisión deberá decidir cuál es la mejor forma de proceder en esta situación.

b) Lugar de residencia anterior cuando las fronteras han cambiado

487. Como las fronteras de los países cambian con el tiempo, hay que asegurarse de que las correspondencias pertinentes se reflejen en los planes de codificación. Además, los códigos deben formularse de manera que permitan agrupamientos lógicos. Por ejemplo, como se ha dicho, en un código de tres cifras la primera podría representar el continente de residencia, la segunda una región del continente y la tercera el país de la región.

c) Cuando la persona no se ha trasladado desde que nació

488. Los encargados de procesar los datos construyen cuadros sobre ciertos elementos concretos. Por lo tanto, los especialistas deberían asegurarse de que se utiliza un código especial para “nació aquí” además de los otros códigos de lugar. De esta forma el programa podrá distinguir entre las personas nacidas en un lugar y las que nacieron en un lugar pero se trasladaron a otro lugar de la misma zona geográfica.

d) Utilización de otras personas de la unidad

489. Cuando el valor de la variable “lugar de residencia anterior” sea inválido o incongruente, normalmente se aplicarán procedimientos de revisión similares a los relativos a la “duración de la residencia”. El programa de revisión puede examinar la residencia anterior de la madre si esta está en la unidad de habitación. A continuación el programa puede examinar la residencia anterior del jefe del hogar en relación con los hijos, y con los adultos, en aquellos países donde los adultos no emigran con frecuencia.

e) No hay otra persona apropiada para determinar la residencia anterior

490. Si ninguno de los métodos anteriores produce un asiento válido, el programa puede asignar un valor “no declarado” o “desconocido” a los años que la persona ha vivido en la residencia anterior. Si el asiento sigue siendo inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por

lo menos dos características para obtener información “conocida” de personas similares de la misma zona geográfica.

5. Lugar de residencia en una fecha especificada del pasado

491. Por lugar de residencia en una fecha especificada del pasado se entiende la división administrativa mayor o menor, o el país extranjero, en que la persona residía en una fecha especificada anterior al censo. La fecha de referencia elegida será la más útil para los fines nacionales; en la mayoría de los casos se ha considerado que es de uno o cinco años antes de la fecha de levantamiento del censo. La primera fecha de referencia proporciona estadísticas corrientes de la migración durante un solo año, mientras que la segunda puede resultar más apropiada para reunir datos con que analizar la migración interna, si bien quizá sea menos adecuada para analizar la migración interna corriente. Al elegir la fecha de referencia también hay que tener en cuenta la capacidad de las personas para recordar con precisión su residencia habitual uno o cinco años antes de la fecha del censo. En los países que levantan censos cada cinco años, para la mayoría de los empadronados será fácil relacionar la fecha pertinente con el momento del censo anterior. En otros casos, es más probable que se recuerde la situación un año atrás que cinco años atrás.

492. Sin embargo, algunos países tal vez tengan que usar una referencia cronológica distinta de uno o cinco años antes del censo, pues ambos lapsos pueden plantear dificultades de memoria. Las circunstancias del país quizá hagan que la referencia cronológica pueda asociarse con algún acontecimiento importante, que la mayoría de las personas recordarán bien. Además, la información sobre el año de llegada al país puede ser útil en el caso de migrantes internacionales (Naciones Unidas, 2008, párr. 2.69).

493. El procedimiento de revisión del “lugar de residencia en una fecha especificada del pasado” es similar al de la residencia anterior. Normalmente los países preguntarán acerca de la “duración de la residencia” y el “lugar de residencia anterior”, o simplemente el “lugar de residencia en una fecha especificada”. Si la persona nació en el lugar de empadronamiento (país, localidad o división administrativa, según como se haya formulado la pregunta del censo) y nunca ha emigrado, este elemento podrá dejarse en blanco, o bien puede existir un código específico para indicar que “nunca emigró”. Como se ha mencionado, los elementos en blanco pueden provocar problemas durante la tabulación. En este caso se aplicarán los mismos procedimientos descritos en los párrafos anteriores con respecto al lugar de residencia anterior.

6. Año de llegada

494. En *Principios y recomendaciones para los censos de población y habitación, Revisión 2*, las variables de migración se dividen en migración interna y migración internacional. El año de llegada hace normalmente referencia al año en que se llegó desde el exterior al país en cuestión. Por ello, el año de llegada es una información que suele solicitarse junto con su complemento, es decir, el lugar de residencia antes de la llegada al país (Naciones Unidas, 2008, párr. 2.103).

a) Relación de la edad con el año de llegada

495. La primera parte de la revisión podría comprobar la congruencia entre la edad y el lugar de nacimiento y una entrada válida en el año de llegada a la localidad o división administrativa. El número de años que una persona ha vivido desde que llegó a una localidad o división administrativa no puede ser mayor que la edad de la

persona. Además, una persona que haya nacido fuera del país no puede haber vivido siempre en esta localidad o división administrativa. El programa debe asignar “siempre” al año de llegada a la localidad o división administrativa si el número de años en esa localidad o división administrativa es mayor que la edad y el país de nacimiento es ese país. Si el número de años desde la llegada a la localidad o división administrativa es mayor pero el país de nacimiento no es ese país, una manera de resolver este caso sería atribuir la edad de la persona como años transcurridos en la localidad o división administrativa. En esas circunstancias, se supone que la persona, aunque haya nacido fuera de este país, llegó a la localidad o división administrativa cuando tenía menos de un año de edad.

496. Para ayudar a los usuarios de muestras de uso público, las oficinas de estadística deben facilitar códigos correspondientes a “menos de uno” y “siempre” en este elemento informativo. El código “siempre” debería corresponder normalmente al lugar efectivo de residencia actual a fin de ayudar a realizar los cuadros directamente. El código “menos de uno” permitirá a los usuarios tener la seguridad de que en su tabulación cruzada han incluido a todos los miembros de la población.

b) Relación del lugar de nacimiento con el año de llegada

497. En el caso de entradas que se salgan de los límites previstos, deberían utilizarse las mismas pruebas que en el caso del lugar de nacimiento. Debería realizarse una búsqueda de las personas emparentadas anteriores (madre, jefe de hogar, esposo, hijo). La atribución debería basarse en la información encontrada. No obstante, antes de que se atribuya un valor, este debe estar en consonancia con la edad y el lugar de nacimiento de la persona cuyo registro se está revisando.

c) Personas que siempre han vivido en el lugar

498. Si la respuesta sobre el número de años desde la llegada que una persona ha vivido en la localidad o división administrativa es “siempre vivió aquí” pero el país de nacimiento no es “este país”, el equipo de revisión podría desear utilizar la edad de la persona para asignar el año de llegada a la localidad o división administrativa. Los especialistas supondrán que, aunque haya nacido fuera de este país, la persona se trasladó a la localidad o división administrativa cuando tenía menos de un año de edad. La fase siguiente de la revisión tratará de determinar una entrada válida para el año de llegada a la localidad o división administrativa. Como el tiempo que una persona vivió en la localidad o división administrativa no puede ser mayor que la edad de esa persona, en esta situación la edad se atribuirá considerando los años transcurridos en la localidad o división administrativa.

d) Año de llegada de la persona desde el año de llegada de la madre

499. Si la categoría no tiene un código válido, el programa puede realizar una comprobación entre diferentes registros buscando a la madre de esa persona dentro del hogar. Si se encuentra, su registro puede facilitar información útil para atribuir los valores no disponibles. Si la madre de la persona ha vivido siempre en la localidad o división administrativa y su país de nacimiento es “este país” (como es lógico), el programa atribuirá “siempre” a los años de esta persona en la localidad o división administrativa. Si el país de nacimiento de la madre no es “este país”, aun cuando la entrada correspondiente a sus años en la localidad o división administrativa sea “siempre”, ello indica la presencia de un error relacionado con las categorías de la madre. En ese caso, el programa no tendrá en cuenta el país de nacimiento de la madre y atribuirá la edad

tomando como base el año de llegada a la localidad o división administrativa. Si la entrada del año de llegada de la madre a la localidad o división administrativa no es “siempre”, pero es un código válido, y la edad de la persona es inferior al número de años transcurridos desde que la madre llegó a la localidad o división administrativa, la revisión retrocederá y comprobará el país de nacimiento de la madre. Si el país de nacimiento de la madre es “este país”, el programa atribuirá la edad de esta persona a los años en la localidad o división administrativa. No obstante, si la edad de una persona es igual o mayor al número de años transcurridos desde la llegada de la madre a la localidad o división administrativa, el programa atribuirá “año de llegada de la madre a la localidad o división administrativa” al año de llegada de esta persona a la localidad o división administrativa.

e) Año de llegada del hijo con respecto al año de llegada del jefe de hogar

500. Si la persona en cuestión es un niño (hijo o hija), el programa de revisión debería comprobar el registro del jefe de hogar para obtener información que pudiera ayudarlo a atribuir los valores correspondientes a los datos no disponibles sobre el año de llegada. Cuando el jefe de hogar nació en “este país” y ha vivido siempre en esta localidad o división administrativa, el programa atribuirá “siempre” a los años del niño en la localidad o división administrativa. Cuando el jefe de hogar ha vivido siempre en la localidad o división administrativa pero no nació en “este país”, se atribuye la edad del hijo a la localidad o división administrativa. Cuando la entrada del jefe de hogar correspondiente al año de llegada a la localidad o división administrativa no es “siempre” pero es un código válido, esa información puede utilizarse si está en consonancia con la edad del registro del niño que se está revisando. Si la edad del niño es algo mayor que la determinada por el año de llegada a la localidad o división administrativa del jefe de hogar, el programa utilizará el año de llegada del jefe a la localidad o división administrativa como año de llegada de ese niño a la localidad o división administrativa. Si la edad del niño es menos de la determinada por el año de llegada del jefe de hogar a la localidad o división administrativa, el programa atribuirá un valor según el país de nacimiento del jefe de hogar. Este valor será “siempre” si el jefe de hogar nació en “este país”; de lo contrario, el programa atribuirá la edad del hijo o la hija a los años en la localidad o división administrativa.

f) Año de llegada de la persona cuando no se dispone de otra información

501. Cuando los esfuerzos anteriores no consiguen arrojar un valor válido, el programa puede atribuir los valores “no registrado” o “desconocido” al año de llegada a la localidad o división administrativa de esa persona. Si el valor continúa siendo inválido, se indicará “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica en el caso de valores inválidos deberían utilizar un número adecuado de características para obtener información “conocida” de personas semejantes de la zona geográfica.

7. Relación entre duración de la residencia y año de llegada

502. Es importante observar que algunos países se concentran en la migración interna e incluyen la información sobre la duración de la residencia (muchas veces junto con la residencia anterior). Otros países que se interesan por la migración internacional incluirán un elemento referente al año de llegada (muchas veces, junto con la residencia antes del traslado). La mayoría de los países tienen o una considerable migración interna y escasa migración internacional o una considerable inmigración

y poca migración interna. En cambio, otros países tienen al mismo tiempo migración interna e internacional, y por lo tanto incluirán ambos elementos.

503. Cuando se incluyen dos elementos, el personal de la oficina de estadística debe proceder con gran cautela para realizar revisiones que terminen presentando incongruencias internas. Es decir, las variables de edad, duración de la residencia y año de llegada deben considerarse conjuntamente para comprobar que la suma de la duración de residencia y el tiempo transcurrido desde la llegada no es superior a la edad. Cuando ello ocurra, los programadores deberán considerar todas las variables al mismo tiempo.

504. Cuando se utiliza la atribución dinámica, el personal de la oficina de estadística debe utilizar una atribución dinámica que incluya matrices multidimensionales para tener en cuenta las diferentes edades y años. Asimismo, cuando la duración de la residencia y el año de entrada sean un solo año, la atribución dinámica deberá utilizar también un solo año, pues la actualización correspondiente al grupo de cinco años, por ejemplo, puede provocar un conflicto durante la atribución de valores.

505. Asimismo, es preciso proceder con gran cautela cuando los datos agrupados correspondientes a la duración de la residencia o al año de entrada o ambos se recojan durante este proceso de comprobación, y al preparar y aplicar las imputaciones dinámicas. Los datos agrupados provocan problemas de superposición. Los países quizá decidan que la mejor solución en este caso es la atribución del valor “desconocido”.

8. Residencia habitual

506. En general, a los efectos del censo, por “residencia habitual” se entiende el lugar en que la persona vive en el momento del censo y en el que ha permanecido o tiene intención de permanecer por algún tiempo (Naciones Unidas, 2008, párrs. 1.461 a 1.463). Se recomienda que los países apliquen un umbral de 12 meses al considerar el lugar de residencia habitual, de acuerdo con uno de los dos criterios siguientes:

- a) El lugar en que la persona ha vivido de forma ininterrumpida durante la mayor parte de los 12 últimos meses (es decir, al menos seis meses y un día), sin contar las ausencias temporales por vacaciones o motivos laborales, o donde tiene intención de vivir durante al menos seis meses;
- b) El lugar en que la persona ha vivido de forma ininterrumpida durante al menos los 12 últimos meses, sin contar las ausencias temporales por vacaciones o motivos laborales, o tiene intención de vivir durante al menos 12 meses.

507. No obstante, los países que realizan censos *de facto* pueden incluir un elemento adicional relativo a la “residencia habitual” para obtener información tanto *de jure* como *de facto*. Las revisiones de este elemento varían según la situación del país en cuestión. En el caso de las personas que no han cambiado nunca de residencia, la residencia habitual será la misma que la residencia actual; por ello, la información no disponible puede introducirse directamente.

508. En cambio, cuando los datos revelan la existencia de movimiento, la situación es más complicada. Normalmente, cuando esta información se deja en blanco, los países suponen que la residencia habitual es también la residencia actual y que el empadronador y/o el declarante sencillamente omitieron esa información.

509. Cuando, con respecto a la duración de la residencia o del año de llegada, los datos contienen pruebas más o menos fehacientes de un cambio de residencia, el personal de estadística quizá desee tratar de elaborar métodos que le permitan deducir

con mayor precisión las zonas geográficas concretas o el conjunto del país. Aunque la revisión específica dependerá de la situación del país en cuestión, probablemente se utilizará como solución de último recurso el valor “desconocido”.

510. Si el empadronador ha recibido instrucciones de dejar en blanco la información cuando la residencia habitual es la misma que el lugar de empadronamiento, durante la revisión deberá colocarse el código correspondiente al lugar de empadronamiento en el elemento relativo a la residencia habitual. Otra variable debería indicar que los revisores han introducido este cambio. La disponibilidad de un conjunto completo de códigos ayudará a los usuarios de la muestra de uso público a realizar tabulaciones completas de sus datos.

C. Características sociales

511. Las características sociales varían según los países, pero en general son elementos que describen distintos aspectos de las condiciones socioculturales del país. Los elementos relativos a la educación —como el grado de alfabetización, la asistencia escolar y el nivel de instrucción, así como la especialidad y los títulos académicos— pueden clasificarse con arreglo a las categorías de la revisión de 1997 de la *Clasificación Internacional Normalizada de la Educación* (CINE) formulada por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) (Naciones Unidas, 2008, párrs. 2.202 a 2.230).

1. Grado de alfabetización

512. Deberían obtenerse datos sobre alfabetización de todas las personas mayores de 10 años. Sin embargo, en varios países algunas personas de edades comprendidas entre los 10 y los 14 años pueden estar a punto de ser alfabetizadas mediante la escolarización, y la tasa de alfabetización de este grupo de edad puede dar lugar a errores. Por consiguiente, en una comparación internacional los datos deben tabularse para todas las personas mayores de 15 años. Cuando en los países se reúnan datos sobre personas más jóvenes, en las tabulaciones sobre alfabetismo debe distinguirse por lo menos entre las personas menores de 15 años y las de 15 años o más (Naciones Unidas, 2008, párr. 2.202).

513. Cada país deberá establecer la edad mínima para las tabulaciones del grado de alfabetización; análogamente, los equipos de revisión deberán decidir la edad mínima para comprobar este elemento, ya que pueden necesitarse tabulaciones adicionales para uso interno. Cuando se elabore el cuestionario, el equipo de revisión deberá decidir la edad mínima para la reunión de datos y el nivel de instrucción a partir del cual ya no será necesario formular la pregunta. Así pues, si el declarante ya ha alcanzado cierto nivel de instrucción, no es necesario que el empadronador formule la pregunta acerca del grado de alfabetización. No obstante, este elemento debería rellenarse durante la revisión para ayudar a los investigadores y otros usuarios de los datos de uso público.

514. En el procedimiento de revisión de la variable de alfabetización se comprueba en primer lugar el nivel escolar alcanzado; si el nivel máximo tiene un valor de “alfabetizado” con arreglo a las especificaciones, habrá que asignar el código correspondiente a “sí”. Las personas que hayan alcanzado un nivel definido de escolarización deberán considerarse alfabetizadas. En los casos en que se encuentre un código inválido en la variable del grado de alfabetización habrá que asignar un valor. En el asiento deberá consignarse “no declarado” o bien determinar un valor utilizando una

matriz de atribución basada en variables especificadas, como el nivel de instrucción alcanzado y el sexo. El “nivel más alto” dependerá de la definición de “alfabetizado” en cada país.

2. Asistencia escolar

515. En principio, la información sobre la asistencia escolar debería obtenerse con respecto a las personas de todas las edades. Se refiere en particular a la población en edad escolar oficial, que en general abarca desde los 5 hasta los 29 años, pero que puede variar según los países con arreglo a la estructura educacional nacional. En el caso de que se amplíe la reunión de datos de manera que abarque la asistencia a la enseñanza preprimaria y/o a otros programas sistemáticos de educación y formación organizados para adultos en empresas productivas y de servicios, organizaciones comunitarias y otras instituciones no docentes, la gama de edades puede ajustarse según convenga (Naciones Unidas, 2008, párr. 2.209).

a) Revisión del valor de la asistencia escolar

516. El equipo de revisión de cada país deberá decidir qué edades son apropiadas para reunir datos sobre asistencia escolar. Puesto que la mayor parte de los países también dividen la escolarización en varios niveles, si esos niveles van a compilarse por edades, los especialistas deberán decidir también qué grupos de edad son apropiados para los distintos niveles de escolarización. Los asientos correspondientes a todas las demás personas deberán modificarse. Si el programa de revisión produce respuestas incongruentes para esta categoría, habrá que cambiar la edad o bien la asistencia escolar. Normalmente la edad ya se ha fijado en el momento en que se lleva a cabo esta revisión, de manera que lo que se cambiará será la asistencia escolar. Hay que instruir a los empadronadores para que omitan la asistencia escolar con respecto a las personas mayores de una edad predeterminada, si ello es apropiado para el país de que se trate. En los casos en que las personas continúen la escolarización secundaria o terciaria hasta la mediana edad tal vez no sea apropiado fijar límites superiores para la asistencia escolar. Presumiblemente, las respuestas y combinaciones de respuestas se someterán a prueba antes del censo mediante ensayos previos, de forma que esas decisiones podrán adoptarse antes del censo propiamente dicho.

b) Matrícula a tiempo completo o a tiempo parcial

517. Algunos países tal vez deseen obtener información sobre la asistencia escolar a tiempo parcial o a tiempo completo. Si se incluye ese elemento, podrá formar parte de la revisión del valor de la asistencia escolar o podrá constituir un procedimiento separado.

c) Congruencia entre la asistencia escolar y la actividad económica

518. Las comprobaciones para determinar la congruencia con otros elementos importantes, como la actividad económica principal, son las primeras que tendrán que realizarse. Si la asistencia escolar es uno de los elementos de la actividad económica principal y una persona declara que su actividad principal es asistir a la escuela, habrá que asignar a la variable de la asistencia escolar el código correspondiente a “sí”, y el código para la actividad económica principal deberá ser “estudiante”. Es decir, las respuestas deberán ser congruentes. En todos los demás casos habrá que aceptar todas las respuestas válidas.

d) **Asignación de valores a asientos inválidos o incongruentes de la variable de asistencia escolar**

519. Si hay un valor que está fuera de los límites y el asiento correspondiente al máximo nivel de instrucción alcanzado es válido, habrá que asignar un valor utilizando una matriz de atribución basada en la edad, el sexo y el máximo nivel de instrucción alcanzado. Si este no tiene un código válido, deberá utilizarse el valor de la alfabetización para asignar el valor de la asistencia escolar. Si no hay un código válido para el grado de alfabetización, habrá que asignar un valor a la asistencia escolar únicamente sobre la base de la edad y el sexo.

520. Pueden necesitarse matrices de atribución para reflejar las distintas estructuras de asistencia escolar por sexo y edad (a veces con intervalos de un año o por pequeños grupos de edad).

3. Nivel de instrucción (grado o nivel máximo terminado)

a) **Revisión del valor del nivel de instrucción**

521. El procedimiento de revisión del nivel de instrucción (grado o nivel máximo terminado) debería consistir en: *a)* una comprobación de la congruencia entre un asiento válido y la edad, y *b)* la atribución de un valor cuando el asiento original esté fuera de los límites. Como se ha mencionado, en los países que no utilizan la atribución dinámica el valor debería consignarse como “no declarado”. En los países que utilizan la atribución dinámica se necesitará el sexo y la edad con intervalos de un año para los jóvenes y el sexo y pequeños grupos de edad para los niños un poco mayores. En los países cuyos datos incluyan el grado y el nivel máximo tal vez sean necesarias matrices de atribución múltiples (Naciones Unidas, 2008, párr. 2.215). En el punto 4 de la sección B del anexo I (página 165) se trata de la oportunidad de establecer un nuevo código para el “nivel de instrucción actual” tomando como base la frecuencia de asistencia a la escuela y el más alto nivel de estudios obtenido.

b) **Edad mínima para el nivel de instrucción**

522. El equipo de revisión de cada país deberá decidir cuál es la edad mínima para ingresar en la escuela. Una vez fijada la edad mínima, el máximo nivel terminado, ordinariamente, no deberá ser superior a la edad de una persona más una constante (que representa esa edad mínima para ingresar en la escuela). También en este caso es importante utilizar la edad de los niños con intervalos de un año, ya que en la actualización de las matrices de atribución pueden introducirse errores si los grupos de edad son muy amplios.

c) **Relación entre la edad y el nivel de instrucción**

523. El equipo de revisión también deberá decidir qué grado de “ruido” se permitirá en el conjunto de datos. Normalmente es mejor modificar unos pocos casos excepcionales cuando la edad y el nivel de instrucción entran en conflicto en vez de aceptar un gran número de respuestas que son realmente incongruentes. Por consiguiente, en los casos en que el asiento original esté fuera de los límites o sea incongruente con la edad podrá asignarse un valor. En los países que no utilicen la atribución dinámica puede consignarse un valor “no declarado”. En los países que utilicen la atribución dinámica puede obtenerse un valor basado en la edad (incluida la edad con intervalos de un año para las personas en edad escolar), el sexo y la asistencia escolar.

La UNESCO reconoce que la alfabetización es distinta del nivel de instrucción, de forma que la “capacidad de leer y escribir” probablemente no debería utilizarse como valor en la matriz de atribución.

4. Especialidad y titulación académica

524. La información sobre las personas por nivel educacional y especialidad académica es importante para examinar la correspondencia entre la oferta y la demanda de mano de obra cualificada con determinadas especializaciones en el mercado de trabajo. Es igualmente importante para planificar y regular la capacidad de producción de diferentes niveles, tipos y ramas de las instituciones educacionales y los programas de formación (Naciones Unidas, 2008, párr. 2.223).

525. Para las personas menores de 15 años (o de otra edad predeterminada) no debería constar información acerca de la especialidad y/o de los títulos académicos. Para una persona de 15 años o más debería existir una relación entre el nivel de instrucción y la especialidad y/o los títulos académicos. En cada caso, cuando se encuentren asientos inválidos, los países que no utilicen la atribución dinámica pueden consignar un valor “desconocido”. Los países que utilicen la atribución dinámica tal vez deseen considerar la posibilidad de utilizar la edad, el sexo, el nivel de instrucción y, posiblemente, la ocupación para asignar el valor de la especialidad y/o la titulación académicas.

5. Religión

526. A los efectos censales, la religión puede definirse como: *a)* la creencia o preferencia religiosa o espiritual, sin tener en cuenta si está representada o no por un grupo organizado, o *b)* la afiliación a un grupo organizado que tiene determinados principios religiosos o espirituales. Los países que incluyan en su censo una pregunta sobre religión deberán emplear la definición que resulte más adecuada para sus necesidades y hacer constar en la publicación censal la definición que han adoptado (Naciones Unidas, 2008, párr. 2.152).

Revisión del valor de la religión

527. La religión es una de las variables que se ajustan a los ejemplos presentados en el capítulo II. En lo que respecta a los elementos relativos a la religión, a diferencia de lo que ocurre en otros casos, la ausencia de respuesta es significativa y debe tenerse en cuenta: algunas personas se resisten a declarar su religión. Un valor válido para una persona puede obtenerse directamente de otro miembro del hogar, si hay un valor disponible, o bien de otro jefe del hogar con características similares. El equipo de revisión deberá determinar el plan de revisión lógico que utilizará para las demás variables sociales. El jefe del hogar deberá ser designado y su información revisada en primer lugar, tanto si se trata de la primera persona de la unidad de habitación como si no. Si para el jefe del hogar se ha consignado un valor inválido o desconocido en la variable de religión, deberán adoptarse las siguientes medidas:

- a) No consta la religión del jefe del hogar pero consta la religión de otra persona de la unidad de habitación

528. El primer paso consiste en determinar si alguna otra persona de la unidad de habitación tiene un asiento válido en la variable de la religión, y se asignará el primer valor válido.

b) No consta la religión del jefe del hogar ni de ninguna otra persona de la unidad de habitación

529. Si no consta la religión de ninguna persona del hogar habrá que asignar un valor “desconocido” (si el país no utiliza la atribución dinámica) o bien atribuir un valor de la variable de religión tomada del jefe del hogar más reciente con características similares, entre ellas la edad y el sexo, así como el idioma, el lugar de nacimiento y otras variables que resulten apropiadas teniendo en cuenta las circunstancias.

c) No consta la religión de alguna persona distinta del jefe del hogar

530. Si la persona no es el jefe del hogar y no se ha consignado para ella un valor en la variable de la religión, el equipo de revisión podría atribuir la religión del jefe del hogar.

6. Idioma

531. Los censos permiten reunir tres tipos de datos sobre el idioma (Naciones Unidas, 2008, párr. 2.156):

- La lengua materna, que se define como el idioma hablado habitualmente en el hogar de la persona durante su primera infancia;
- El idioma habitual, que se define como el que la persona habla corrientemente o más a menudo en su hogar actual; y
- La capacidad de hablar uno o más idiomas determinados.

a) Revisión del valor del idioma

532. De las tres variables de idioma que pueden aparecer en el cuestionario, las dos primeras (lengua materna e idioma habitual) están relacionadas. Cuando ambas estén presentes en un cuestionario, los equipos de revisión deberán considerar la posibilidad de revisarlas conjuntamente. Si una es inválida, podrá utilizarse la otra para suministrar un valor.

b) Revisión del valor del idioma: el jefe del hogar

533. El idioma es otra variable que se ajusta a los ejemplos presentados en el capítulo II. El equipo de revisión deberá establecer el plan de revisión lógico utilizado para las demás variables sociales, comprobando en primer lugar el registro del jefe del hogar. Si la persona que tiene un asiento inválido o desconocido en la variable del idioma (lengua materna o idioma habitual) es el jefe del hogar, habrá que determinar primero si alguna otra persona de la unidad de habitación tiene un asiento válido para el idioma y asignar el primer valor válido. De no ser así, o bien se asignará un valor “desconocido”, si no se utiliza la atribución dinámica, o bien se atribuirá un idioma tomado del jefe del hogar más reciente con características similares, incluidas la edad y el sexo, así como otras variables de idioma, lugar de nacimiento y otras variables pertinentes en esas circunstancias.

c) Revisión del valor del idioma: personas distintas del jefe del hogar

534. Si la persona cuyo asiento en la variable del idioma es inválido no es el jefe del hogar, habrá que asignarle el idioma del jefe del hogar.

d) Revisión del valor del idioma: utilización del origen étnico o del lugar de nacimiento

535. El idioma y el origen étnico, y a veces el lugar de nacimiento, están estrechamente relacionados, y en algunos países podrán revisarse conjuntamente. Además, el equipo de revisión deberá considerar la posibilidad de organizar los códigos de forma que reflejen las relaciones entre esas variables. Con arreglo al número de cifras del código y a la distribución de los idiomas y los grupos étnicos en el país, podrán establecerse correspondencias que ayuden a asignar valores a las respuestas desconocidas o incongruentes.

e) Revisión del valor del idioma: lengua materna

536. Si la lengua materna es desconocida pero la persona es filipina y nació en Filipinas, puede asignarse un idioma equivalente apropiado, como el tagalo, el ilokano u otro idioma de Filipinas. Normalmente el idioma sólo se asignará de esta forma al jefe del hogar, y el código correspondiente a ese idioma se asignará a los demás miembros del hogar, pero el equipo de revisión de cada país deberá considerar las circunstancias particulares, incluidas la geografía (como la residencia en zona urbana o rural), la edad u otros elementos.

f) Revisión del valor del idioma: capacidad de hablar un idioma determinado

537. La capacidad de hablar un idioma determinado es una tercera variable que se ajusta a los ejemplos presentados en el capítulo II. También en este caso habrá que revisar en primer lugar el registro del jefe del hogar. Si el asiento correspondiente al idioma del jefe del hogar es inválido o desconocido, el primer paso consistirá en averiguar si alguna otra persona de la unidad de habitación tiene un asiento válido respecto de la capacidad de hablar un idioma determinado, y se asignará el primer valor válido. Si no se encuentra esa persona, se asignará un valor “desconocido”, si el país no utiliza la atribución dinámica, o bien se atribuirá el valor de la capacidad de hablar un idioma determinado del jefe del hogar más reciente con características similares (por ejemplo, edad y sexo, pero también lugar de nacimiento y otras variables apropiadas según las circunstancias). Si la persona no es el jefe del hogar y el asiento de la capacidad de hablar un idioma determinado es inválido, se asignará el valor correspondiente del jefe del hogar.

7. Grupo étnico y pueblos indígenas

538. La decisión de recopilar y difundir información sobre los grupos étnicos o nacionales de una población en un censo depende de diversas consideraciones y circunstancias nacionales; por ejemplo, de las necesidades de dichos datos en el país y de la conveniencia y oportunidad de incluir en el censo preguntas que pueden resultar comprometidas. La determinación de las características étnicas y culturales de la población del país tiene una importancia creciente en el contexto de la migración, la integración y las políticas que afectan a los grupos minoritarios (Naciones Unidas, 2008, párr. 2.160).

539. En *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (párr. 2.163) se recomienda proceder con especial esmero al identificar a los pueblos indígenas, normalmente como subconjunto de la identidad étnica. Al elaborar las listas de códigos debe procurarse que la condición de indígena se especifique de forma específica para que puedan realizarse las revisiones y tabulaciones pertinentes a

fin de facilitar la planificación y la formulación de políticas para los pueblos indígenas. Por ejemplo, quizá se necesiten códigos independientes para el mismo grupo cuando son nómadas, a diferencia de cuando se han asentado en zonas residenciales. Pueden realizarse revisiones especiales, parcialmente a través de archivos de “búsqueda”, para grupos concretos de pueblos indígenas, a fin de garantizar que se identifiquen correcta y plenamente para los cuadros posteriores. Pueden establecerse procedimientos de atribución especiales para estos grupos, o utilizarse categorías adicionales en el marco de las actividades de atribución dinámica existentes.

a) **Revisión del valor del grupo étnico**

540. Varias otras variables, si se han recopilado, podrán ayudar a “determinar” el grupo étnico cuando este valor sea inválido o desconocido. En muchos países existe una relación entre el lugar de nacimiento, tanto dentro del país como en el extranjero, y el grupo étnico. De manera análoga, la “lengua materna” es a menudo un buen indicador del grupo étnico en muchos países, ya que las categorías, y por tanto los códigos, serán similares, si no idénticos.

b) **Revisión del valor del grupo étnico: el jefe del hogar**

541. El origen étnico también se ajusta a los ejemplos presentados en el capítulo II. El equipo de revisión deberá seguir el plan ya descrito para las demás variables sociales. El registro del jefe del hogar deberá ser revisado en primer lugar. Si el asiento inválido o desconocido para la variable del origen étnico es el del jefe del hogar, habrá que buscar en primer lugar un grupo étnico válido para cualquier otra persona de la unidad de habitación y asignar este primer grupo étnico válido. De no encontrarse esa persona, el próximo paso será asignar un valor “desconocido” o, si el país no utiliza la atribución dinámica, atribuir el valor del grupo étnico del jefe del hogar más reciente con características similares (edad y sexo, así como idioma, lugar de nacimiento y otras variables que puedan ser apropiadas según las circunstancias).

c) **Revisión del valor del grupo étnico:
personas distintas del jefe del hogar**

542. Si el asiento inválido en la variable del origen étnico no es el del jefe del hogar, se asignará el origen étnico del jefe del hogar.

d) **Revisión del valor del grupo étnico:
utilización del idioma y del lugar de nacimiento**

543. El origen étnico, el idioma e incluso el lugar de nacimiento están estrechamente relacionados, y algunos países podrán revisarlos en conjunto. Los equipos de revisión deberán considerar además la posibilidad de organizar sus códigos para que reflejen las relaciones entre esas variables. Con arreglo al número de cifras del código y a la distribución de los grupos étnicos y los idiomas en el país podrán establecerse correspondencias que ayuden a asignar valores a las respuestas desconocidas o incongruentes.

544. Por ejemplo, si el origen étnico es desconocido pero la persona habla uno de los idiomas de Filipinas y nació en Filipinas, podría asignarse un origen étnico equivalente apropiado, es decir, filipino. Normalmente sólo se asignará el grupo étnico de esta forma al jefe del hogar (y se asignará el mismo código a los demás miembros), pero el equipo de revisión de cada país deberá examinar las circunstancias particulares, incluidas la geografía (como la residencia en zonas urbanas o rurales), la edad u otros elementos.

8. Discapacidad

545. La situación en materia de discapacidad divide a la población entre quienes tienen y no tienen una discapacidad. Una persona con discapacidad es una persona que tiene un riesgo mayor que la población general de sufrir limitaciones para la realización de determinadas tareas o para participar en actividades relacionadas con algunas funciones. Las Naciones Unidas recomiendan la inclusión de cuatro tipos al evaluar la situación en materia de discapacidad: 1) movilidad, 2) visión, 3) audición y 4) cognitiva (Naciones Unidas, 2008, párrs. 2.351, 2.352 y 2.367 a 2.371).

546. La pregunta para identificar a las personas con discapacidad debe enumerar las categorías generales de discapacidad, de modo que cada persona pueda comprobar la presencia o ausencia de cada uno de los tipos. La discapacidad puede supervisarse siguiendo los siguientes elementos basados en la Clasificación Internacional de Deficiencias, Incapacidades y Minusvalías (ICIDH): 1) funcionamiento y discapacidad, con inclusión de las funciones y estructura corporales (minusvalías) y las actividades (limitaciones) y la participación (restricciones) y 2) factores contextuales, con inclusión de los factores ambientales y personales (Naciones Unidas, 2008, párr. 2.354).

a) Preguntas del censo relativas a la discapacidad

547. Se recomienda especial cautela al formular las preguntas del censo encaminadas a medir la discapacidad. Los términos y la redacción de las preguntas influyen enormemente en la precisión a la hora de localizar a las personas con discapacidades. Cada ámbito debería de ser objeto de una pregunta diferente⁸. La terminología utilizada debería ser clara, inequívoca y sencilla. Deberían evitarse los términos negativos. Las preguntas sobre la discapacidad deberían formularse a cada uno de los miembros del hogar, y deberían evitarse las preguntas generales sobre la presencia de personas con discapacidades en el hogar. En caso necesario, podría recurrirse a una tercera persona para que informe, en sustitución del miembro de la familia discapacitado. Lo más importante es tener en cuenta a cada uno de los miembros de la familia en forma individual, en vez de formular preguntas de alcance general. La inclusión de categorías con respuestas graduadas puede mejorar también la notificación de la discapacidad (Naciones Unidas, 2008, párr. 2.373).

⁸ Cuando se combinan los ámbitos, por ejemplo, formulando preguntas sobre la visión o la audición, los declarantes muchas veces se desconciertan, suponiendo que deben tener dificultades en los dos ámbitos para poder responder afirmativamente. Asimismo, la disponibilidad de cifras sobre las personas con limitaciones específicas es útil para la planificación interna y para las comparaciones transnacionales.

b) Revisión del valor de la discapacidad

548. Cuando alguien no responde a esta pregunta es difícil determinar si el elemento ha quedado en blanco porque no existe discapacidad o porque, por el motivo que sea, el declarante no desea responder. El equipo de revisión del país deberá decidir si desea revisar el elemento de la forma habitual, asignando un valor desconocido cuando no se aplique la atribución dinámica, o utilizando las respuestas de otras personas cuando se aplique este tipo de atribución. Otra posibilidad es que los especialistas decidan que sólo se aceptarán las respuestas en las que se indique que existe discapacidad y que toda respuesta inválida deberá consignarse como “no hay discapacidad”. En este último caso no se utilizaría la atribución dinámica.

c) Discapacidades múltiples

549. Los países que reúnan información sobre discapacidades múltiples deberán modificar el procedimiento de revisión. El programa de revisión deberá efectuar un seguimiento del número total de discapacidades posible y de la duplicación y distribución de esas discapacidades. Como en casos anteriores, la mayoría de los países considerarán inadecuado utilizar datos de otras personas para asignar valores de dis-

capacidad, de forma que tal vez sea necesario consignar el valor como “desconocido” o incluso como “se desconoce si hay discapacidad” en los casos inválidos.

d) Revisión del valor de las causas de la discapacidad

550. El equipo de revisión del país deberá decidir si desea revisar el elemento siguiendo la manera habitual, asignando un valor desconocido cuando no se aplique la atribución dinámica, o utilizando las respuestas de otras personas cuando se aplique este tipo de atribución. Otra posibilidad es que los especialistas decidan que sólo se aceptarán las respuestas en las que se indique que hay una causa de la discapacidad, y que no se utilizará una matriz de atribución.

D. Características económicas

551. La información sobre la actividad económica debería abarcar en principio a toda la población, pero en la práctica se reúne sobre las personas que hayan cumplido una edad mínima determinada fijada con arreglo a las condiciones de cada país. No debe tomarse automáticamente la edad mínima hasta la cual es obligatoria la escolarización como límite de edad mínimo para la reunión de información sobre la situación laboral. Los países en los que normalmente muchos niños participan en la agricultura o en otros tipos de actividad económica (por ejemplo, la minería, la industria textil o el pequeño comercio) deberán establecer una edad mínima más baja que los países en que el empleo de niños es poco habitual.

552. En las tabulaciones de las características económicas debería distinguirse por lo menos entre las personas menores de 15 años y las de 15 años o más; y los países donde la escolarización sigue siendo obligatoria después de los 15 años y en los que hay niños económicamente activos menores de esta edad deberían esforzarse en obtener datos sobre las características económicas de esos niños con miras a lograr la comparabilidad internacional, por lo menos para las personas mayores de 15 años. A menudo también se pasa por alto la participación en actividades económicas de hombres y mujeres ancianos después de la edad normal de jubilación. Esto exige prestar mucha atención al medir la población económicamente activa. En general no debería fijarse un límite de edad máximo para medir la población económicamente activa, ya que es posible que un número considerable de ancianos que hayan superado la edad de jubilación se dediquen a actividades económicas de manera habitual u ocasional (Naciones Unidas, 2008, párr. 2.241).

553. Cada país deberá determinar una edad mínima para la participación en la actividad económica. Los países interesados en reunir datos sobre el trabajo infantil pueden necesitar fijar una edad mínima baja, pero deberán tener presente que se producirá cierto grado de “ruido” si por equívoco se empadrona como miembros de la fuerza de trabajo a niños que no participan en ella. Una vez establecida la edad mínima, los elementos de la actividad económica se revisan para poder ser tabulados respecto de las personas de X años de edad o mayores; la comprobación de los registros de niños menores de X años, pues, sólo será necesaria para asegurarse de que todos los elementos hayan quedado en blanco. Para facilitar las tabulaciones habrá que eliminar todas las respuestas que puedan haberse consignado para niños menores de X años.

1. Población activa

554. La actividad económica consta de varias variables económicas; algunas se describen a continuación. Esas variables son satisfactorias para la reunión de datos,

pero tal vez sea necesario establecer nuevas categorías para el procesamiento y el análisis de los datos.

555. La “situación laboral actual” es la relación de una persona con la actividad económica, basada en un período de referencia breve, como una semana. La utilización del concepto de actividad actual se considera más apropiada en los países en donde la actividad económica de las personas no está muy influida por factores estacionales o de otro tipo que causen variaciones a lo largo del año. El período de referencia de una semana que se utilice puede ser una semana reciente especificada, la última semana civil completa o los últimos siete días anteriores al empadronamiento (Naciones Unidas, 2008, párr. 2.248).

556. Según Naciones Unidas, 2008 (párr. 2.253), son personas **empleadas** todas las personas por encima de una edad determinada que, durante un breve período de referencia de una semana o un día, *a*) realizaron algún trabajo por un salario o un sueldo, un beneficio o prestación familiar, en dinero o en especie, o *b*) estuvieron temporalmente ausentes de un empleo —en el que habían trabajado ya y con el que mantienen un vínculo formal— o de un trabajo por cuenta propia, como una empresa, una explotación agrícola o de una actividad relacionada con los servicios.

557. La población *no actualmente activa* o, en otras palabras, la población *no comprendida en la fuerza de trabajo*, incluye todas las personas que durante el breve período de referencia utilizado para medir la actividad actual no estuvieron *empleadas* ni *desempleadas*, incluidas las personas por debajo de una determinada edad especificada para la medición de la población económicamente activa (Naciones Unidas, 2008, párrs. 2.278 y 2.279).

a) Categorías relativas a la actividad económica

i) Población desempleada

558. Según Naciones Unidas 2008 (párr. 2.271), son personas **desempleadas** todas las personas mayores de una edad determinada que durante el período de referencia estaban:

- a) *Sin empleo*: No tenían un empleo remunerado ni un empleo independiente;
- b) *Disponibles actualmente para trabajar*: Estaban disponibles para trabajar en un empleo remunerado o un empleo independiente durante el período de referencia;
- c) *En busca de empleo*: Habían tomado medidas concretas en un período reciente especificado para buscar un empleo remunerado o un empleo independiente. Las medidas concretas pueden ser, por ejemplo, el registro en una agencia de colocación pública o privada; la presentación de ofertas a empleadores; la realización de averiguaciones en obras, explotaciones agrícolas, entradas de fábricas, mercados u otros lugares de oferta de trabajo; la colocación de anuncios en los periódicos o el envío de respuestas a ese tipo de anuncios; la formulación de peticiones de ayuda a amigos y parientes; la búsqueda de terrenos, edificios, maquinaria o equipo para establecer una empresa propia; la realización de gestiones para conseguir recursos financieros; y la solicitud de permisos y licencias. Sería útil distinguir en la clasificación de los desempleados entre las personas que buscan trabajo por primera vez y otras personas que buscan trabajo.

559. En general, para estar clasificada como desempleada una persona deberá satisfacer los tres criterios indicados. Sin embargo, en las situaciones en que los medios convencionales de búsqueda de trabajo tienen una pertinencia limitada, en que el mer-

cado laboral está muy poco organizado o cubre un ámbito limitado, en que la asimilación de los trabajadores resulta transitoriamente insuficiente o en que gran parte de la fuerza de trabajo se dedica a tareas por cuenta propia, la definición ordinaria de desempleo puede aplicarse relajando el criterio relativo a la búsqueda de trabajo; con ello se tiene en cuenta la situación de aquellos países en desarrollo donde el criterio de búsqueda de trabajo no permite reflejar todo el alcance del desempleo. Con la aplicación laxa de ese criterio, que en casos extremos puede llegar a su supresión completa, los dos criterios básicos que siguen siendo aplicables son “sin empleo” y “disponibles actualmente para trabajar” (Naciones Unidas, 2008, párr. 2.272).

560. Los procedimientos de revisión del valor consignado en la variable de desempleo —“desocupado”, “en busca de trabajo”, si la persona está en condiciones de aceptar un empleo, y “último año en el que trabajó” (si existen esos elementos)— deberían aplicarse de manera conjunta. Además, es necesario que sean compatibles con la respuesta relativa a la actividad económica, y en la mayoría de los casos no deberían constar si se han llenado los elementos correspondientes a tiempo trabajado, rama de actividad económica, ocupación, categoría en el empleo y lugar de trabajo. Si los especialistas temáticos determinan que se necesita un elemento que corresponda a “desocupado” cuando la respuesta haya quedado en blanco o sea inválida, podría aplicarse una matriz de atribución que utilizara la edad y el sexo y tal vez el nivel de instrucción de la persona.

ii) *En busca de empleo*

561. La revisión de la variable “en busca de empleo” deberá hacerse conjuntamente con las correspondientes a “desocupado” y “por qué no busca empleo”. El personal temático deberá elaborar procedimientos de revisión en que los asientos correspondientes a esos elementos se utilicen para atribuir los demás valores. El procedimiento de revisión deberá tener en cuenta las condiciones locales y regionales y las variables del censo o de la encuesta.

iii) *No actualmente activa*

562. La población que consta como “no actualmente activa”, es decir, las personas que “no están en la fuerza de trabajo”, abarca a todas las personas que no estaban “empleadas” ni “desempleadas” durante el breve período de referencia utilizado para medir la actividad actual (Naciones Unidas, 2008, párr. 2.278). Según los motivos que aduzcan para no ser “actualmente activas”, esas personas podrán clasificarse en cualquiera de los grupos siguientes:

- a) Asisten a un establecimiento docente;
- b) Realizan trabajos en el hogar;
- c) Viven de una pensión o de rentas del capital;
- d) Otros motivos.

Los procedimientos de revisión del valor “no actualmente activa” se han integrado en los procedimientos de revisión indicados para la actividad económica.

iv) *Por qué no busca trabajo*

563. Este elemento sólo deberá revisarse en el caso de las personas que consten como “no busca trabajo”; en todos los demás casos el asiento deberá estar en blanco. Por el contrario, si aparece un asiento válido en relación con la ocupación, la rama de actividad económica y la categoría en el empleo, habrá que consignar el código correspondiente a “empleado, pero no trabajando”. Este código designa a las personas econó-

micamente activas que estaban empleadas pero que no trabajaron durante el período de referencia. En el resto de los casos, si no se utiliza la atribución dinámica, puede asignarse un valor “desconocido”. Los países que apliquen la atribución dinámica pueden asignar un valor utilizando la edad, el sexo y la actividad principal.

b) Revisión del valor de la actividad económica

564. Generalmente, la actividad económica tiene las categorías siguientes:

Persona económicamente activa

- 1) Empleado
- 2) Desempleado

Persona no económicamente activa

- 3) Estudiante
- 4) Trabajador del hogar
- 5) Jubilado o persona que recibe rentas del capital
- 6) Otros.

i) Personas empleadas

565. Si se ha seleccionado la categoría 1 (“Empleado”), habrá que llenar las variables de tiempo trabajado, ocupación, rama de actividad económica, categoría en el empleo y lugar de trabajo. Si esas variables no se han llenado, deberán someterse a revisión y consignarse un valor desconocido o bien valores de atribución estática o dinámica. Si se ha seleccionado la categoría 1, deberán quedar en blanco las variables correspondientes a desocupado, en busca de empleo y último año en que trabajó. Si esas variables se han llenado, el asiento deberá modificarse de modo que quede en BLANCO.

ii) Actividad económica de las personas desempleadas

566. Si se ha seleccionado la categoría 2, habrá que llenar las variables de “desocupado”, “en busca de empleo” y “último año en que trabajó”. Si no se han llenado esos elementos con valores válidos, deberán someterse a revisión y consignarse un valor desconocido, o bien valores de atribución estática o dinámica. Si se ha seleccionado una categoría de 2 a 6, deberán quedar en blanco las variables correspondientes a tiempo trabajado, ocupación, rama de actividad económica, categoría en el empleo y lugar de trabajo. Si se han llenado, el asiento deberá modificarse de manera que quede en blanco.

iii) Actividad económica de los estudiantes y de los jubilados

567. Si se ha seleccionado la categoría 3, estudiante, el personal temático deberá decidir si el asiento correspondiente a la variable de asistencia escolar debe ser “sí, en la escuela”. Si se ha seleccionado la categoría 5, jubilado, el personal temático deberá decidir si las personas tienen que haber alcanzado una edad determinada para poder figurar como jubiladas.

iv) Cuando la actividad económica no tiene un asiento válido y se han declarado variables de empleo

568. Si el asiento correspondiente a la actividad económica no es válido y se han comunicado algunas de las variables de tiempo trabajado, ocupación, rama de actividad económica y lugar de trabajo, la actividad económica del declarante deberá codificarse con el valor 1. Probablemente se necesitará una matriz de atribución para seleccionar la respuesta apropiada.

v) *Cuando la actividad económica no tiene un asiento válido y se han declarado variables de desempleo*

569. Si se han comunicado valores para algunas de las variables de “desocupado”, “en busca de empleo” y “último año en que trabajó”, el asiento correspondiente a la actividad económica deberá codificarse con un valor de 2 a 6. Si la persona asiste a la escuela, el valor probablemente será 3. Si se trata de un anciano, el valor deberá ser probablemente 5. De otro modo, los especialistas temáticos podrán decidir aplicar una matriz de atribución para asignar una respuesta apropiada.

vi) *Cuando la actividad económica no tiene un asiento válido y no se ha declarado ninguna de las variables económicas*

570. Si no hay respuesta para ninguno de los elementos de actividad económica, los especialistas temáticos probablemente desearán utilizar matrices de atribución para determinar la respuesta más apropiada, y a continuación asignar valores por atribución a los demás elementos económicos.

2. Tiempo trabajado

571. El tiempo trabajado es el tiempo total dedicado a producir bienes y servicios durante el período de referencia adoptado respecto de la actividad económica en el censo. Se recomienda que si el período de referencia es breve (por ejemplo, la semana anterior al censo), el tiempo trabajado se mida en horas. Si el período de referencia es largo (por ejemplo, los 12 meses anteriores al censo), el tiempo trabajado deberá medirse en unidades de semanas, o en días cuando sea posible, o en intervalos de tiempo más largos. El tiempo trabajado también deberá incluir actividades que, si bien no conducen directamente a la producción de bienes o servicios, se definen igualmente como parte de las tareas y funciones del empleo, tal que el tiempo dedicado a preparar, reparar o mantener el lugar de trabajo o los instrumentos de trabajo. En la práctica, el tiempo trabajado incluirá también el tiempo de inactividad durante la realización de esas actividades, tal que el tiempo pasado esperando o en reserva, u otras interrupciones breves. Deben excluirse interrupciones más largas para comidas o el tiempo que no se ha trabajado por vacaciones, fiestas, enfermedad o por conflictos laborales (por ejemplo, huelgas y cierres patronales) (Naciones Unidas, 2008, párr. 2.323).

572. Este elemento sólo deberá revisarse en el caso de las personas cuya respuesta para el elemento de la actividad económica fue “empleado, trabajando” o “con un empleo independiente, trabajando”. En algunos países también debería incluirse el tiempo trabajado por los trabajadores del hogar. Habría que aceptar las categorías predeterminadas por el equipo de revisión. Si no se utiliza la atribución dinámica, los códigos correspondientes a respuestas en blanco, cero o códigos no numéricos deberían sustituirse por “no declarado”, y los especialistas temáticos podrían querer sustituir el valor de la variable de la actividad económica por “no trabaja” si las horas declaradas son equivalentes a cero.

573. Si se aplican técnicas de atribución dinámica, las variables mínimas para la matriz de atribución incluyen los grupos de edad y el sexo, pero también pueden utilizarse otras variables como el nivel de instrucción, la ocupación o las principales categorías económicas.

3. Ocupación

574. Por ocupación se entiende el tipo de trabajo que durante el período de referencia efectuó una persona empleada (o el que efectuaba anteriormente, si se trata

de una persona desempleada), cualquiera que sea la rama de actividad económica de que forme parte o la categoría en el empleo en que deba clasificarse (Naciones Unidas, 2008, párr. 2.301).

575. Este elemento sólo deberá revisarse en el caso de las personas cuya actividad económica conste como “empleado”. Si no se utiliza la atribución dinámica, las respuestas en blanco, inválidas o con un código cero deberán sustituirse por un valor “no declarado”.

576. Los códigos de ocupación suelen estar elaborados de manera que diferentes cifras representan los códigos de ocupación principales y secundarios. Las descripciones escritas, que son casi inevitables en el caso de la ocupación, complicarán la tarea de codificación.

577. Si se aplican técnicas de atribución dinámica, las variables mínimas para la matriz de atribución serán los grupos de edad y el sexo, pero también pueden utilizarse otras variables, como el nivel de instrucción o las principales categorías económicas.

4. Rama de actividad económica

578. Según *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (párr. 2.306), por “rama de actividad económica” se entiende la actividad del establecimiento en que una persona empleada trabajó durante el período de referencia adoptado para los datos correspondientes a las características económicas (o donde trabajó por última vez si se trata de una persona desempleada). Para obtener orientación sobre la selección del empleo o la actividad que deberá clasificarse, véase el párrafo 2.307 de *Principios y recomendaciones*.

579. Este elemento sólo deberá revisarse en el caso de personas cuya actividad económica conste como “empleado”. Si no se utiliza la atribución dinámica, las respuestas en blanco, el código cero o las respuestas inválidas deberán sustituirse por un valor “no declarado”.

580. Los códigos de rama de actividad económica tienden a estar elaborados de manera que diferentes cifras representan los códigos de rama de actividad económica principal o secundaria. Las descripciones escritas, que son casi inevitables en el caso de este elemento, complicarán la tarea de codificación.

581. Si se aplican técnicas de atribución dinámica, las variables mínimas para la matriz de atribución serán los grupos de edad y el sexo, pero también pueden utilizarse otras variables, como el nivel de instrucción o las principales categorías de actividad económica.

5. Categoría en el empleo

582. La categoría en el empleo denota la relación entre una persona económicamente activa y su empleo, es decir, el tipo de contrato explícito o implícito de trabajo con otras personas u organizaciones que esa persona tiene en su empleo. Los criterios básicos aplicados para definir los grupos de la clasificación son el tipo de riesgo económico, un elemento del cual es la solidez de la relación entre la persona y el empleo, y el tipo de autoridad sobre los establecimientos y otros trabajadores que la persona tiene o tendrá en el empleo. Hay que prestar atención a asegurar que una persona económicamente activa sea clasificada por categoría en el empleo en relación con el mismo puesto o puestos utilizados para clasificar a la persona por ocupación, rama de actividad económica y sector (Naciones Unidas, 2008, párr. 2.310).

583. La población económicamente activa deberá clasificarse por categoría en el empleo de la manera siguiente (Naciones Unidas, 2008, párr. 2.311):

- a) Empleados, entre los cuales es posible distinguir entre empleados con contratos estables (incluidos los empleados regulares) y otros empleados;
- b) Empleadores;
- c) Trabajadores por cuenta propia;
- d) Trabajadores familiares no remunerados;
- e) Miembros de una cooperativa de producción;
- f) Personas no clasificables por categoría.

584. Deberá identificarse por separado a los propietarios-gerentes de empresas constituidas en sociedad, que normalmente se clasificarán entre los empleados pero a los que, para determinados fines descriptivos y analíticos, puede ser preferible agrupar con los empleadores.

585. Este elemento deberá revisarse sólo en el caso de las personas cuya actividad económica se declare como “empleado”. Si no se utiliza la atribución dinámica, las respuestas en blanco, los valores equivalentes a cero o las respuestas inválidas pueden sustituirse por “no declarado”. Si se aplican técnicas de atribución dinámica, las variables mínimas de la matriz de atribución deberán ser los grupos de edad y el sexo, pero también pueden utilizarse otras variables, como el nivel de instrucción o las principales categorías de actividad económica.

6. Ingresos

586. Los temas del censo relativos a las características económicas de la población que se presentan en *Principios y recomendaciones para los censos de población y habitación* se centran en la población económicamente activa, definida según las recomendaciones de la Organización Internacional del Trabajo (OIT), en las que el concepto de producción económica se define con respecto al Sistema de Cuentas Nacionales (Naciones Unidas, 2008, párr. 2.331). La población económicamente activa abarca a todas las personas de uno u otro sexo que aportan o son capaces de aportar su trabajo para producir bienes y servicios económicos, según la definición del Sistema de Cuentas Nacionales durante un período de referencia especificado (Naciones Unidas, 2008, párr. 2.237).

587. En este contexto, los ingresos pueden definirse como: a) los ingresos mensuales en dinero y/o en especie provenientes del trabajo realizado por cada persona económicamente activa, o b) los ingresos anuales totales en dinero y/o en especie de los hogares, cualquiera que sea su fuente. La reunión de datos sobre el ingreso, en particular el proveniente del empleo independiente y de la propiedad, es sumamente difícil en las investigaciones generales sobre el terreno, particularmente en los censos de población. La inclusión de ingresos que no sean en dinero acrecienta aún más las dificultades. La reunión de datos sobre los ingresos en los censos de población, aun cuando se limita a los ingresos en dinero, presenta problemas adicionales en cuanto a mayor volumen de trabajo, errores en las respuestas y otros aspectos. Por ello este tema, incluida la definición más general de ingresos, suele considerarse más apropiado para una encuesta por muestreo. No obstante, según las necesidades nacionales, los países tal vez deseen obtener alguna información sobre los ingresos en dinero. Si se parte de esta definición, la información que se reúna podrá aportar datos para las estadísticas sobre la distribución del ingreso, el consumo y la acumulación de los hogares, además de las finalidades inmediatas que tiene para el censo.

588. En *Principios y recomendaciones para los censos de población y habitación, Revisión 2* se indican dos tipos de ingresos: los ingresos personales y los ingre-

del hogar. Ambos elementos exigen comprobaciones similares. En el caso de los ingresos personales, si no se utiliza la atribución dinámica habrá que asignar un valor “no declarado” o “desconocido” a las respuestas inválidas. Si se aplica la atribución dinámica podrían utilizarse, entre otras, las variables de la edad, el sexo, el nivel de instrucción, la rama de actividad económica o la ocupación para configurar la matriz de atribución de valores de ingresos.

589. En esta variable los ingresos del hogar corresponden a la suma de todos los ingresos obtenidos por los miembros del hogar, y se consignan en el registro de vivienda. No obstante, la revisión mediante atribución dinámica es aproximadamente la misma, pero utiliza los valores de la edad, el sexo y el nivel de instrucción del jefe del hogar, y no los de cada persona. Véase una exposición más detallada de las recodificaciones de los ingresos del hogar y de la familia en el anexo I.

7. Sector institucional del empleo

590. Por sector institucional del empleo se entiende la organización jurídica y las principales funciones, comportamiento y objetivos de la empresa a la que corresponde un empleo (Naciones Unidas, 2008, párr. 2.335).

591. Existe una relación entre algunas de las posibles ramas de actividad económica y ocupaciones y el sector institucional de empleo (empresas, sector público, instituciones sin fines de lucro, hogares, etcétera). Algunos países pueden decidir comprobar esas relaciones entre las variables para asegurarse de que no haya incongruencias en ciertas tabulaciones cuando se efectúen tabulaciones cruzadas de esas variables.

592. En el proceso de revisión, los países que no utilicen la atribución dinámica deberán asignar un valor “desconocido” al sector institucional de empleo cuando este no se conozca. Los países que apliquen técnicas de atribución dinámica deberían considerar la posibilidad de utilizar la edad y el sexo y tal vez la rama de actividad económica principal o la ocupación de personas similares de la misma zona geográfica.

8. Empleo en el sector informal

593. En los casos en que las actividades del sector informal desempeñan un papel importante en la creación de empleo y la generación de ingresos, algunos países utilizan la información sobre la actividad laboral y otros aspectos económicos para identificar el sector informal. Otros países formulan preguntas específicas sobre la participación en el sector informal (Naciones Unidas, 2008, párr. 2.337).

594. La revisión correspondiente al sector informal debe ser sencilla. Si la participación en el sector informal es independiente de la participación en el sector formal, puede atribuirse el valor “desconocido” a las entradas en blanco o inválidas, o realizar una atribución dinámica basada en la edad y el sexo. Si la participación en el sector informal no es independiente de la participación en el sector formal, podría incluirse una variable adicional en la matriz de atribución dinámica para indicar si la persona se encontraba también en el sector formal.

9. Lugar de trabajo

595. El “lugar de trabajo” es el lugar en el cual una persona actualmente empleada realiza su trabajo y donde una persona habitualmente empleada tiene el empleo principal utilizado para determinar sus otras características económicas, como

la ocupación, la rama de actividad económica y la categoría en el empleo. Mientras que la información sobre el lugar de trabajo puede utilizarse para establecer perfiles de la zona en cuanto a la fuerza de trabajo empleada (distintos de los perfiles demográficos por lugar de residencia), el objetivo principal es vincular la información sobre el lugar de trabajo con el lugar de residencia (Naciones Unidas, 2008, párr. 2.346).

596. Puesto que el “lugar de trabajo” se utiliza para las estadísticas sobre desplazamientos diarios de la población, es importante que todas las modificaciones en la información comunicada correspondan a las zonas geográficas concretas que se están estudiando. Por tanto, los equipos de revisión de los países tal vez deseen considerar la posibilidad de asignar un valor “desconocido” a los casos inválidos y analizar únicamente los casos “conocidos”.

597. El tiempo necesario para las operaciones de codificación de este elemento y su complejidad aumentarán si se aceptan las descripciones por escrito y hay que codificarlas. Si se determina una jerarquía para las cifras del código (por ejemplo, que la primera cifra represente la provincia, la segunda el distrito, etcétera), es probable que la operación de codificación sea más eficiente y más exacta.

598. En las matrices de atribución, los encargados del procesamiento de datos deberán asegurarse de que sólo se asignen a las matrices zonas geográficas probables. Puede ser prudente iniciar una nueva matriz estática para cada división administrativa u otra zona geográfica con objeto de asegurarse de que no puedan seleccionarse valores anteriores. En las propias matrices de atribución podrá incluirse la edad y el sexo, y tal vez categorías modificadas de ocupación principal o rama de actividad económica. También pueden requerirse matrices de atribución distintas para el trabajo en el país y fuera de él.

599. En este capítulo se han considerado las variables de población recomendadas en *Principios y recomendaciones*. Ningún país debería utilizar todas estas variables, y las variables seleccionadas, y sus relaciones especiales con otras variables deberían comprobarse a fondo en condiciones de invernadero y en encuestas previas al censo para determinar la existencia de respuestas fiables y completas. Dado que las informaciones relativas a la población, a diferencia de las variables de habitación, normalmente son objeto de tabulaciones cruzadas en muchas combinaciones diferentes, se necesita una comprobación a fondo.

Capítulo V

Revisión de los elementos de vivienda

600. Las especificaciones de revisión de las variables de vivienda toman en consideración tanto la validez de cada uno de los elementos como la coherencia entre ellos. Si en un país determinado se conocen las relaciones específicas entre elementos será posible planear comprobaciones de la coherencia y obtener datos de más calidad para la tabulación. Una unidad de habitación, por ejemplo, no deberá tener un techo de cemento cuando las paredes estén construidas de bambú. De modo análogo, las unidades deberán tener agua corriente en el interior de la casa para que pueda haber un retrete o una bañera o ducha dentro de la estructura.

601. Al igual que sucede con las variables de población, cuando falten elementos o estos sean inválidos el equipo de revisión deberá decidir si asigna un valor “no declarado”, un valor de atribución estática para un elemento “desconocido” u otro valor, o bien un valor de atribución dinámica basado en las características de otras unidades de habitación. Como antes, en muchos casos se prefiere la atribución dinámica, porque elimina el tipo de atribución que se necesita en la etapa de tabulación, cuando sólo se dispone de la información que figura en las propias tabulaciones para adoptar decisiones sobre los valores desconocidos. Las matrices de atribución construidas de este modo suministran valores para los elementos en blanco, los asientos inválidos o las incongruencias resueltas cuando no existen otros elementos conexos con respuestas válidas. Es posible que en algunos países haya cierta variación en las características de vivienda según las zonas de la nación pero muy poca dentro de las localidades mismas. En otros países puede haber variaciones considerables respecto de elementos particulares entre las localidades, particularmente entre las zonas urbanas y rurales. Esta variación deberá tenerse en cuenta al elaborar matrices de atribución y, particularmente, al establecer los valores iniciales de atribución estática. El equipo de revisión tal vez desee especificar las circunstancias en que habrá que suministrar un valor para un asiento en blanco a partir de una unidad de habitación previa con otras características similares.

602. Excepto en los casos en que en un país se carezca de información sobre las unidades de habitación colectivas (de grupo), habrá que asignar un registro de vivienda (y sólo uno) a cada número de serie (véase el capítulo III, “Revisiones estructurales”, pág. 59 y ss.). En ese capítulo se enumeran varios procedimientos de garantía de calidad. Con arreglo a las decisiones que adopte el equipo de revisión, el programa podrá crear un registro de vivienda si este falta. Del mismo modo, el programa podrá eliminar uno o más registros cuando se encuentren registros duplicados o múltiples.

603. Lo ideal sería que cada registro de vivienda fuera sometido a revisión selectivamente, sólo para comprobar los elementos pertinentes. Las variables revisadas pueden ser distintas según se trate de zonas urbanas o rurales, o en función de las condiciones climáticas y otras circunstancias. Sin embargo, en la práctica pocos países tienen tiempo o conocimientos técnicos para elaborar y aplicar matrices múltiples que permitan modificar los datos omitidos o incongruentes. Aún es menor el número de países que realizan efectivamente una labor de revisión selectiva.

604. No obstante, más por razones estéticas que técnicas, y en particular por lo que se refiere a la habitación, dado que la corrección es cada vez más compleja y detallada, ahora se enfatiza que las zonas geográficas seleccionadas tengan únicamente respuestas “adecuadas”. Por ejemplo, si algunas zonas geográficas del país no tienen electricidad, tampoco deberían tener aire acondicionado, refrigeradores eléctricos o cocinas eléctricas. Para resolver ese tipo de cuestiones, en determinadas zonas geográficas puede preverse una revisión, a fin de evitar que se produzcan anomalías en el conjunto de datos finales. Quizá lo más indicado sea adoptar un método aproximativo y empírico y eliminar los casos que puedan ser de hecho ajenos. Por ejemplo, aunque quizá haya casos en que individuos adinerados de una zona compren generadores de gas para utilizarlos porque no se disponga de electricidad de otro modo, el equipo de revisión puede decidir no incluir esos casos en el conjunto de datos.

605. La información que se haya reunido en el cuestionario dependerá también del tipo de vivienda (unidad de habitación o vivienda de grupo) y de si la unidad de habitación estaba desocupada u ocupada. En el caso de las viviendas colectivas o de grupo, la comprobación puede limitarse únicamente a los elementos que se reúnen en las viviendas de grupo, o incluir los que se reúnen tanto en las viviendas de grupo como en otras unidades de habitación.

606. Por definición, no suele haber registros de vivienda para las personas sin hogar. Si esos registros existen porque el país ha decidido que debe haber identificadores para ellos, el país deberá tratarlos de la misma forma que los relativos a las viviendas colectivas, o bien requerir un procedimiento de revisión completamente distinto, que puede incluir la posibilidad de no efectuar ninguna comprobación.

607. A veces habrá que permitir que para un elemento particular conste un valor “no declarado”. Esto puede ocurrir cuando el equipo de revisión del país carece de una base sólida para atribuir respuestas para una determinada característica. La decisión de dejar respuestas “no declaradas” deberá equilibrarse con la necesidad de producir características tabulares apropiadas para la planificación y la formulación de políticas. En la medida en que los casos “no declarados” tengan la misma distribución que los casos declarados, asignar valores a los casos “no declarados” cuando los planificadores necesitan determinada información no debería plantear ningún problema. Sin embargo, si los casos “no declarados” presentan algún sesgo, la atribución posterior a la compilación puede plantear problemas, especialmente para zonas pequeñas o tipos de condiciones particulares. Por ejemplo, si los declarantes que viven en un tipo de vivienda que no reúne unas condiciones mínimas definidas por el país se niegan a revelar algunas de sus características de vivienda y el empadronador tampoco consigna la situación, es posible que los planificadores no puedan establecer programas correctivos para aliviar esas condiciones deficientes.

608. La revisión de las variables de vivienda tiende a ser más sencilla que la de las variables de población porque, en general, las tabulaciones cruzadas son mucho menos complicadas. La mayoría de países compila las distintas características de vivienda únicamente por niveles geográficos. Como se ha indicado, los países que opten por no aplicar métodos de atribución dinámica deberían establecer un identificador para los valores “desconocidos”, que se utilizará cuando se produzcan respuestas inválidas o incongruentes.

609. En los países que apliquen técnicas de atribución dinámica, el equipo de revisión deberá elaborar matrices de atribución simples con dimensiones que permitan distinguir las características de vivienda. Para la mayor parte de los países, una variable sobre el “tipo de local de habitación”, tanto si se trata de una unidad de habitación como de una vivienda colectiva, incluido el tipo de unidad dentro de esas categorías, será la mejor variable primaria para la atribución dinámica.

610. En algunos países las zonas geográficas pueden utilizarse como una de las dimensiones de las matrices de atribución. También puede utilizarse el tipo de tenencia. Por ejemplo, si en el país aproximadamente la mitad de las unidades es de alquiler y la otra mitad de propiedad, la tenencia es una variable apropiada para incluirla como una de las dimensiones de la matriz de atribución. Sin embargo, si las unidades de alquiler sólo representan el 5% del total, sería más apropiada alguna otra característica. A menudo la tenencia es una variable útil para las matrices de atribución, particularmente en países que tienen grandes porcentajes de los tipos principales de tenencia. Otras características que pueden considerarse son el tipo de paredes y la existencia de una instalación de corriente eléctrica.

611. En cada país, las variables concretas que se incluyan como dimensiones de las matrices de atribución deberán corresponder a las variables del conjunto de datos, de manera que para los elementos de vivienda hay que prestar atención a que los distintos elementos, así como las combinaciones de estos, hagan una distinción entre esas características.

A. Temas básicos y adicionales

612. Las unidades de enumeración de los censos de vivienda son: *a)* los edificios, *b)* los locales de habitación y *c)* los ocupantes de los hogares. Las Naciones Unidas han elaborado una lista de temas básicos de interés y valor general que también son importantes para efectuar una comparación exhaustiva de las estadísticas a nivel internacional. Para facilitar su utilización, a continuación figuran entre paréntesis los códigos de revisión sugeridos para esos temas y varios temas adicionales. Los temas se agrupan por tipos de unidad de numeración.

1. Edificio: descripción del edificio (tema básico)

613. La clasificación que se esboza a continuación describe un sistema de códigos de tres cifras (Naciones Unidas, 2008, párrs. 2.412 a 2.454) para reunir en grupos amplios las unidades de habitación y los locales de habitación colectivos con características estructurales similares. La distribución de los ocupantes (población) entre los diversos grupos proporciona datos valiosos acerca del tipo de alojamiento disponible en el momento de levantarse el censo. La clasificación proporciona además una base adecuada de estratificación para las encuestas por muestreo. Los locales de habitación pueden dividirse en las siguientes categorías:

1. Unidades de habitación
 - 1.1. Viviendas de tipo corriente
 - 1.1.1. Tiene todas las instalaciones básicas
 - 1.1.2. No tiene todas las instalaciones básicas
 - 1.2. Otras unidades de habitación
 - 1.2.1. Unidades de habitación semipermanentes
 - 1.2.2. Unidades de habitación móviles
 - 1.2.3. Unidades de habitación improvisadas
 - 1.2.4. Unidades de habitación en edificios permanentes no destinados a habitación humana
 - 1.2.5. Otros locales no destinados a habitación humana
2. Locales de habitación colectivos
 - 2.1. Hoteles, pensiones y otras casas de huéspedes

- 2.2. Instituciones
 - 2.2.1. Hospitales
 - 2.2.2. Instituciones correccionales (prisiones, centros penitenciarios)
 - 2.2.3. Instituciones militares
 - 2.2.4. Instituciones religiosas (monasterios, conventos, etcétera)
 - 2.2.5. Residencias de ancianos
 - 2.2.6. Dormitorios para estudiantes, y semejantes
 - 2.2.7. Locales para el personal (por ejemplo, albergues y residencias para personal de enfermería)
 - 2.2.8. Orfanatos
 - 2.2.9. Otros
- 2.3. Campamentos y alojamientos para trabajadores
 - 2.3.1. Campamentos militares
 - 2.3.2. Campamentos para trabajadores
 - 2.3.3. Campamentos de refugiados
 - 2.3.4. Campamentos para personas internamente desplazadas
 - 2.3.5. Otros
- 2.4. Otros

614. Los equipos de revisión deberán elaborar procedimientos para asegurarse de que todos los locales colectivos y unidades de habitación contengan información congruente. Si el asiento del tipo de local de habitación es desconocido o inválido, el equipo de revisión tal vez desee elaborar un procedimiento que examine otras variables con objeto de asignar un valor al tipo de local de habitación. De otro modo, si el asiento es inválido habrá que asignar un valor “desconocido” si no se utiliza la atribución dinámica. Las oficinas nacionales de estadística o del censo que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de edificio, la tenencia, el número de cuartos, la superficie, o el estado de ocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

2. Ubicación del local de habitación (tema básico)

615. La ubicación del local de habitación es una variable geográfica y se describe junto con las revisiones estructurales en el capítulo III.

3. Ocupación (o desocupación) (tema básico)

616. La decisión de registrar los locales de habitación cuyos ocupantes están temporalmente ausentes o presentes como “ocupados” o “desocupados” dependerá de que se levante un censo de población *de jure* o *de facto*. En cualquiera de los dos casos parece conveniente que, en lo posible, se distingan los locales de habitación empleados como primer lugar de residencia de los locales empleados como segundo lugar de residencia. Este dato sería especialmente importante cuando el segundo lugar de residencia tuviera características que lo diferenciaran notablemente del primero. Así sucede, por ejemplo, cuando personas de hogares del sector agrícola abandonan su lugar de residencia permanente en una aldea y pasan a ocupar estructuras rudimentarias en explotaciones agrícolas durante ciertas épocas del año (Naciones Unidas, 2008, párr. 2.466). La clasificación recomendada para las viviendas corrientes y básicas es la siguiente:

1. Ocupada
2. Desocupada
 - 2.1. Desocupada con carácter estacional
 - 2.1.1. Finca de recreo
 - 2.1.2. Alojamientos para trabajadores estacionales
 - 2.1.3. Otras situaciones
 - 2.2. Desocupada con carácter no estacional
 - 2.2.1. Residencia secundaria
 - 2.2.2. En alquiler
 - 2.2.3. En venta
 - 2.2.3. Destinada a demolición
 - 2.2.4. Otras situaciones

617. En aquellos casos en los que la unidad de habitación está ocupada, el número de ocupantes y el recuento de los registros de población no deben tener un valor de cero. Cuando no aparezcan personas registradas, la unidad se considerará que se encuentra desocupada, o que faltarán los datos sobre las personas. Como se ha señalado anteriormente con respecto a la revisión estructural, los especialistas deberán crear procedimientos a fin de determinar si la unidad se encuentra desocupada. Si la unidad figura como ocupada pero en realidad está desocupada, habrá que elaborar un método para determinar el tipo de desocupación, bien sea consignando un valor “desconocido” o bien utilizando la atribución dinámica. En el caso de que la unidad figure como desocupada pero pueda determinarse que en realidad está ocupada debido a que existe información sobre el número de ocupantes o porque existe un recuento en los registros de población, la variable de ocupación deberá modificarse de modo que el valor sea “ocupado”.

618. Si el asiento es inválido, el valor del número de ocupantes es cero y no hay registros de población, habrá que asignar el valor correspondiente a “desocupada y desconocida” cuando no se utilice la atribución dinámica. Si el asiento es inválido pero el número de ocupantes no es cero o existen registros de población, habrá que asignar el valor correspondiente a “ocupada”. Aquellos países que apliquen técnicas de atribución dinámica para los asientos inválidos (para asignar un valor al tipo de ocupación) deberán utilizar por lo menos dos características para obtener información “conocida” de unidades de habitación similares de la misma zona geográfica. Otra posibilidad sería asignar a la unidad de habitación el valor correspondiente a “desocupada y desconocida”.

4. Tipo de propiedad (tema básico)

619. Este tema se refiere al tipo de propiedad de la unidad de habitación en sí, y no al del terreno en que se levante (Naciones Unidas, 2008, párr. 2.467). No debe confundirse el tipo de propiedad con la tenencia o título de ocupación. La información sobre el tipo de propiedad debe indicar si el local de habitación pertenece al sector público (gobierno central, administración local, empresas del Estado) o si la unidad de habitación es de propiedad privada (si pertenece a hogares, empresas privadas, cooperativas, asociaciones inmobiliarias u otras entidades privadas). A veces se amplía esta pregunta para que se indique si la unidad de habitación se ha pagado totalmente, se está pagando a plazos o está hipotecado. La clasificación de las unidades de habitación por tipo de propiedad es la siguiente:

1. Ocupada por el propietario
2. No ocupada por el propietario
 - 2.1. De propiedad pública
 - 2.2. De propiedad privada
 - 2.3. De propiedad comunal
 - 2.4. De propiedad cooperativa
 - 2.5. Otros tipos

620. Si la propiedad está relacionada con la tenencia, deberá tenerse en cuenta al elaborar el procedimiento de revisión; si esos dos elementos no guardan relación, el tipo de propiedad probablemente será independiente de otras variables de vivienda. Si el asiento correspondiente al “tipo de propiedad” es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el material de construcción de las paredes, la tenencia, el tipo de unidad de habitación y el número de cuartos, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

5. Número de cuartos (tema básico)

621. Se considera cuarto al espacio situado en una unidad de habitación u otro local de habitación cerrado por paredes que se eleven desde el suelo hasta el techo, o por lo menos hasta una altura de dos metros a partir del suelo, y con una superficie suficiente para dar cabida a una cama para una persona adulta; o sea, un espacio de por lo menos cuatro metros cuadrados. Por tanto, en el total de cuartos se incluyen los dormitorios, comedores, salas, estudios, áticos habitables, cuartos de sirvientes, cocinas, cuartos que se emplean para fines profesionales o comerciales y demás espacios separados usados para alojar personas o destinados a ello, siempre que se ajusten a los criterios relativos a superficie cubierta y paredes. Los pasillos, galerías abiertas, vestíbulos, cuartos de baño y retretes no deben considerarse como cuartos aunque se ajusten a los criterios descritos. Para fines nacionales, se puede obtener información por separado sobre los espacios de menos de cuatro metros cuadrados que se ajustan en otros aspectos a la definición de cuarto, si se considera que su número justifica este procedimiento (Naciones Unidas, 2008, párr. 2.472).

622. Puesto que el número de cuartos puede ser independiente de las otras variables de vivienda, si el asiento es inválido habría que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de unidad de habitación, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

6. Número de dormitorios (tema adicional)

623. Además de enumerar la cantidad de cuartos, en varios censos nacionales se reúne información sobre el número de dormitorios de la unidad de habitación, que es la unidad de enumeración para este tema. Un dormitorio se define como un cuarto equipado con una cama y utilizado para el descanso nocturno (Naciones Unidas, 2008, párr. 2.475).

624. A veces los empadronadores consignarán para el número de dormitorios un valor superior al del número de cuartos⁹. Si esto ocurre y el país sólo utiliza el valor

⁹ Si se indican los cuartos y los dormitorios, esos asientos deberían revisarse de manera conjunta, y el número de dormitorios no debería sobrepasar al de cuartos. Puesto que el número de dormitorios es un tema “adicional”, la revisión sólo se efectuará cuando existan ambos elementos.

“no declarado” para las respuestas inválidas o incongruentes, el número de dormitorios deberá figurar como “no declarado”. Si se utiliza la atribución dinámica, habrá que estimar el número de dormitorios a partir de una matriz de atribución que tenga como uno de sus elementos el número de cuartos. De esta forma el número de dormitorios no será superior al número de cuartos, porque el valor de los dormitorios sólo se actualizará cuando concuerden los valores correspondientes a los cuartos y a los dormitorios. El caso más simple sería una matriz lineal en la que se utilizarán como casillas el número de cuartos, y el valor correspondiente a los dormitorios se consignará en las casillas. Una matriz de atribución más compleja podría incluir el número de personas de la unidad de habitación y el tipo de estructura.

625. De lo contrario, si el asiento correspondiente al número de dormitorios es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características (una de las cuales deberá ser el número de dormitorios) a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

7. Superficie útil (tema adicional)

626. Este punto se refiere a la superficie útil de los locales de habitación, es decir, la superficie medida dentro de las paredes exteriores de los locales de habitación, excluidos los sótanos y áticos no habitables. En los edificios de viviendas múltiples deben excluirse todos los espacios comunes. Los criterios aplicados a las unidades de habitación y a los locales de habitación colectivos deben ser distintos (Naciones Unidas, 2008, párr. 2.476).

627. La superficie puede estar relacionada con el número de cuartos y/o de dormitorios, de manera que los equipos de revisión de los países tal vez deseen tener esto en cuenta al formular los procedimientos de revisión. Otros elementos útiles para la atribución dinámica son el número de ocupantes y los ocupantes por cuarto. Por lo demás, la superficie es independiente de los otros elementos de vivienda. Puede ser necesario especificar una unidad de medida, como los metros cuadrados. Si el asiento es inválido, habrá que asignar un valor “desconocido” si no se utiliza la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de unidad de habitación, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

8. Sistema de suministro de agua (tema básico)¹⁰

628. Según Naciones Unidas, 2008 (párr. 2.479), los datos básicos que se obtengan en el censo deben permitir saber si el local de habitación cuenta con un sistema de suministro de agua corriente. Esta información indicará si el agua es traída al local de habitación por tuberías desde una red comunitaria o por instalaciones individuales, como un depósito a presión o una bomba. La unidad de enumeración para este tema es la unidad de habitación. Se debe indicar también si el local de habitación tiene grifo en el interior o, en caso negativo, si hay un grifo a cierta distancia de la puerta. La distancia recomendada es 200 metros, suponiendo que el acceso al agua corriente en este radio permita a los ocupantes de la unidad de habitación obtener agua para las necesidades del hogar sin un esfuerzo desmesurado. Además de la ubicación de un grifo, también es de especial interés la fuente del agua disponible. La clasificación

¹⁰ Respecto de las variables siguientes, la unidad de enumeración es la unidad de habitación: sistema de suministro de agua, instalaciones de retrete y alcantarillado, instalaciones de baño, instalaciones para cocinar, alumbrado y eliminación de residuos sólidos.

recomendada de las unidades de habitación por sistema de suministro de agua es, pues, la siguiente:

1. Agua corriente dentro de la unidad
 - 1.1. De la red comunitaria
 - 1.2. De una fuente privada
2. Agua corriente fuera de la unidad pero en un radio de 200 metros
 - 2.1. De la red comunitaria
 - 2.1.1. De uso exclusivo
 - 2.1.2. Compartida
 - 2.2. De una fuente privada
 - 2.2.1. De uso exclusivo
 - 2.2.2. Compartida
3. Otras posibilidades

629. Una red comunitaria es una red sometida a inspección y control por las autoridades públicas. Estas redes suelen estar a cargo de un organismo público, pero en algunos casos es una cooperativa la que se encarga de su gestión, o una empresa privada.

630. Los elementos relativos a las instalaciones de agua —sistema de suministro de agua, instalaciones de retrete y alcantarillado, instalaciones de baño y disponibilidad de agua caliente— probablemente deberán revisarse juntos. Puesto que están estrechamente relacionados, cuando hay un asiento que falta o es inválido los demás podrán utilizarse para generar un valor. En las zonas geográficas que carecen de agua corriente es posible que los especialistas necesiten comprobaciones especializadas para las unidades. Por lo demás, es probable que las otras unidades de la zona tengan características similares, y esos elementos son idóneos en caso de que se aplique la atribución dinámica.

631. Si el asiento correspondiente al sistema de suministro de agua es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características. Estas podrían incluir, por lo general, el tipo de unidad de habitación y las instalaciones de retrete y alcantarillado y las instalaciones de baño, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

9. Agua potable; fuente principal de (tema básico)

632. La revisión de la información relativa al agua potable debe realizarse juntamente con la del sistema de suministro de agua. Muchos de los criterios antes descritos se aplican también aquí. Dado que normalmente se considerarán en el cuestionario el agua embotellada y otras fuentes no tradicionales de agua potable, deberán incluirse también en la revisión.

633. Si el asiento correspondiente al sistema de suministro de agua es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características. Estas podrían incluir, por lo general, el tipo de unidad de habitación y las instalaciones de retrete y alcantarillado y las instalaciones de baño, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

10. Tipo de retrete (tema básico) y

11. Instalaciones de alcantarillado (tema básico)

634. Los tipos de retrete y las instalaciones de alcantarillado deben revisarse conjuntamente con otras variables relacionadas con la fontanería a fin de obtener los resultados más coherentes. Aunque en *Principios y recomendaciones para los censos de población de habitación, Revisión 2* se combinaban las dos variables, para 2010 aparecen por separado. No obstante, estos elementos deben revisarse conjuntamente y, si es posible, deberían utilizarse las mismas matrices de atribución dinámica.

635. Algunos países han observado que es útil ampliar la clasificación de los retretes sin salida de agua para distinguir ciertos tipos cuyo uso está muy generalizado y que indican cierto nivel de saneamiento. La clasificación de la unidad de habitación por tipo de instalaciones de retrete recomendada en Naciones Unidas, 2008 (párr. 2.487) es la siguiente:

1. Con retrete en la unidad de habitación
 - 1.1. Retrete con descarga de agua/de sifón
 - 1.2. Otros
2. Con retrete fuera de la unidad de habitación
 - 2.1. De uso exclusivo
 - 2.1.1. Retrete con descarga de agua/de sifón
 - 2.1.2. Letrina mejorada de pozo con ventilación
 - 2.1.3. Letrina de pozo sin ventilación con cobertura
 - 2.1.4. Perforaciones o pozos excavados con cobertura temporal o sin protección
 - 2.1.5. Otros
 - 2.2. Compartido
 - 2.2.1. Retrete con descarga de agua/de sifón
 - 2.2.2. Letrina mejorada de pozo con ventilación
 - 2.2.3. Letrina de pozo sin ventilación con cobertura
 - 2.2.4. Perforaciones o pozos excavados con cobertura temporal o sin protección
 - 2.2.5. Otros
3. Sin retrete
 - 3.1. Servicio o uso de cubos (las excretas se eliminan manualmente)
 - 3.2. Uso del medio ambiente natural, por ejemplo, arbusto, río, arroyo, etc.

636. El tipo de instalaciones de retrete y alcantarillado es otro elemento de vivienda relacionado con el agua y debe formar parte de un procedimiento de revisión conjunto con otros elementos de este tipo. Los valores como “privado”, “compartido”, “de uso exclusivo”, etcétera, pueden utilizarse para determinar si los valores son congruentes y, de no ser así, para establecer los procedimientos de revisión que habrá que seguir para resolver el problema. Cuando haya una o más variables relacionadas con el agua, puede llegar a estimarse la información desconocida o incongruente sin tener que recurrir al uso de un valor “desconocido” o a la atribución dinámica. Sin embargo, si con ello no se obtiene un asiento válido, habrá que asignar un valor “desconocido” si no se utiliza la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, incluido el tipo de unidad de habitación, por lo general, así como el sistema de suministro de agua, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

12. Instalaciones de baño (tema básico)

637. Según Naciones Unidas, 2008 (párr. 2.490), se debe averiguar si existen un baño o ducha fijos dentro de cada local de habitación. La unidad de enumeración para este tema es también la unidad de habitación. Se puede obtener información adicional que indique si estos servicios son para uso exclusivo de los ocupantes de los locales de habitación y si hay agua caliente para bañarse o solamente fría. Sin embargo, es posible que en algunas zonas del mundo esta distinción no resulte la más adecuada para las necesidades nacionales. Al contrario, es posible que, por ejemplo, revista importancia saber si se dispone de un cuarto separado para bañarse en el local de habitación, un cuarto separado para bañarse en el edificio o un cubículo abierto para bañarse en el edificio, o si se puede hacer uso de una casa de baños pública. La clasificación recomendada de las unidades de habitación por disponibilidad y tipo de instalaciones de baño es la siguiente:

1. Con baño o ducha fijos en la unidad de habitación
2. Sin baño ni ducha fijos en la unidad de habitación
 - 2.1. Baño o ducha fijos disponibles fuera de la unidad de habitación
 - 2.1.1. De uso exclusivo
 - 2.1.2. Compartidos
 - 2.2. Sin baño ni ducha fijos

638. El tipo de instalaciones de baño debería ser objeto de una revisión conjunta con otros elementos relacionados con el agua. Los valores como “privado”, “compartido” o “de uso exclusivo” pueden utilizarse para determinar si los valores son congruentes y, de no ser así, para establecer los procedimientos de revisión que habrá que seguir para resolver el problema. Cuando haya una o más variables distintas relacionadas con el agua, puede llegar a estimarse la información desconocida o incongruente sin tener que recurrir al uso de un valor “desconocido” o a la atribución dinámica. Sin embargo, si no hay otro remedio y el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características. Por regla general, estas podrán ser el tipo de unidad de habitación y el sistema de suministro de agua, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

13. Cocina, disponibilidad de (tema básico)

639. Según los *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008, párr. 2.494), la reunión de datos sobre si se dispone o no de una cocina puede proporcionar una oportunidad para recoger información sobre el tipo de equipo utilizado para cocinar, como estufa, hornillo o fogón, y la disponibilidad en la cocina de una pila y de espacio para almacenar alimentos de manera que no se echen a perder. La clasificación recomendada de las unidades de habitación con arreglo a la disponibilidad de cocina u otro espacio dedicado a cocinar es la siguiente:

1. Con cocina dentro de la unidad de habitación
 - 1.1. De uso exclusivo
 - 1.2. Compartido
2. Con otro espacio para cocinar dentro de la unidad de habitación
 - 2.1. De uso exclusivo
 - 2.2. Compartido

3. Sin cocina ni otro espacio para cocinar dentro de la unidad de habitación
 - 3.1. Cocina u otro espacio para cocinar disponible fuera de la unidad de habitación
 - 3.1.1. De uso exclusivo
 - 3.1.2. Compartido
 - 3.2. Sin cocina ni otro espacio para cocinar

640. Con respecto a las instalaciones para cocinar, pueden utilizarse valores como “privado”, “compartido”, “de uso exclusivo”, etcétera, para determinar si los asientos son congruentes y, en el caso de hallar incongruencia en ellos, para establecer los procedimientos de revisión que habrá que seguir para resolver el problema. Cuando haya una o ambas variables relativas a las instalaciones para cocinar, puede llegar a estimarse la información desconocida o incongruente sin recurrir al uso de un valor “desconocido” o a la atribución dinámica. Sin embargo, si el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se aplique la atribución dinámica. Aquellos países que apliquen la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, que por lo general serán el tipo de unidad de habitación, además de las instalaciones de suministro de agua, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

14. Combustible utilizado para cocinar (tema básico)

641. En el contexto de la necesidad de vigilar de cerca el aprovechamiento de los recursos naturales, el tema del combustible utilizado para cocinar se incluye en varios censos nacionales de vivienda. La unidad de enumeración es la unidad de habitación; por “combustible utilizado para cocinar” se entiende el combustible utilizado predominantemente para la preparación de las comidas principales. Si se utilizan dos tipos de combustible (por ejemplo, electricidad y gas) hay que enumerar el que se utiliza más frecuentemente. La clasificación de los combustibles utilizados para cocinar depende de las circunstancias nacionales y puede abarcar la electricidad, el gas, el petróleo, el carbón, la leña y los residuos animales (estiércol). También resulta útil reunir esta información para los locales de habitación colectivos, en particular si el número de este tipo de locales de habitación es significativo en el país (Naciones Unidas, 2008, párr. 2.496).

642. Las respuestas a la pregunta sobre el combustible utilizado para cocinar deberán revisarse conjuntamente con las relativas a las instalaciones para cocinar. El equipo de revisión determinará la relación existente entre las dos variables y formulará un procedimiento para comprobar que sean congruentes. Probablemente se utilizarán valores como “privado”, “compartido”, “de uso exclusivo”, etcétera, para determinar si los valores son congruentes y, de no ser así, para establecer los procedimientos de revisión que habrá que seguir para resolver el problema. Cuando existan asientos para una o ambas variables relacionadas con la cocina podrá llegarse a estimar la información desconocida o incongruente sin recurrir a usar un valor “desconocido” ni a la atribución dinámica. Sin embargo, si el asiento es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como las instalaciones de cocina, el tipo de edificio, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información de otras unidades de habitación similares de la misma zona geográfica.

15. Tipo de alumbrado y/o electricidad (tema básico)

643. Se debe obtener información que indique el tipo de alumbrado que hay en la unidad de habitación (por ejemplo, eléctrico, de gas, con lámparas de petróleo o de alguna otra fuente). Si el alumbrado consta como eléctrico es posible que algunos países deseen recoger información que les indique si la electricidad procede de un servicio de suministro comunitario, de una instalación generadora o de alguna otra fuente, como podría ser una planta industrial o una mina. Además del tipo de alumbrado, los países deberían evaluar la información relativa a la disponibilidad de electricidad para fines distintos del alumbrado (como serían cocinar, calentar agua o calentar los locales). Si las condiciones de la vivienda en el país permiten derivar esta información del tipo de alumbrado, no se necesitará una pregunta adicional (Naciones Unidas, 2008, párr. 2.497).

644. Si el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, una de las cuales será por lo general el tipo de unidad de habitación, además de la electricidad, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

16. Eliminación de residuos sólidos, tipo principal de (tema básico)

645. Según los *Principios y recomendaciones para los censos de población y habitación, Revisión 2* (Naciones Unidas, 2008, párr. 2.500), este tema se refiere a la recogida y eliminación de los residuos sólidos generados por los ocupantes de la unidad de habitación. La unidad de enumeración es la unidad de habitación. A continuación se presentan las directrices para clasificar las unidades de habitación por sistema de eliminación de residuos sólidos:

1. Recogida de residuos sólidos de manera regular por personas autorizadas
2. Recogida de residuos sólidos de manera irregular por personas autorizadas
3. Recogida de residuos sólidos a cargo de personas independientes
4. Los ocupantes vierten los residuos sólidos en un vertedero local supervisado por las autoridades
5. Los ocupantes vierten los residuos sólidos en un vertedero local no supervisado por las autoridades
6. Los ocupantes queman sus residuos sólidos
7. Los ocupantes entierran sus residuos sólidos
8. Los ocupantes vierten los residuos sólidos en un río/mar/arroyo/estanque
9. Los ocupantes producen abono con los residuos sólidos
10. Otros

646. La eliminación de residuos sólidos es independiente de las demás variables de vivienda. Si el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, que por regla general serán el tipo de unidad de habitación, además del material de construcción de las paredes, la tenencia, la ocupación o desocupación y las instalaciones para cocinar, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

17. Tipo de calefacción, y energía utilizada (tema adicional)

647. Este tema se refiere al tipo de calefacción de los locales de habitación y a la energía utilizada para ese fin. Las unidades de enumeración son los locales de habitación. Este tema no es pertinente para varios países en los que, debido a la situación geográfica y al clima, no hay necesidad de calefacción en los locales de habitación. El “tipo de calefacción” se refiere a la clase de sistema utilizado para calentar la mayor parte de la superficie. Puede ser una calefacción central exclusiva para un local de habitación o para un conjunto de locales de habitación, o puede no ser central, en cuyo caso la calefacción funciona por separado en el local de habitación mediante una estufa, un hogar o algún otro artefacto que despiden calor. La “energía utilizada para la calefacción” está estrechamente relacionada con el tipo de calefacción y se refiere a la fuente de energía predominante, como combustibles sólidos (carbón, lignito y productos del carbón y el lignito, leña), petróleo y derivados, combustibles gaseosos (gas natural o licuado) y electricidad (Naciones Unidas, 2008, párr. 2.501).

648. El tipo de calefacción y la energía utilizada para calefacción están relacionados entre sí, con la disponibilidad de agua caliente y otros servicios de la unidad de habitación como la electricidad y las instalaciones de gas. El equipo de revisión deberá tener en cuenta la disponibilidad de esos elementos al realizar las especificaciones de revisión correspondientes al tipo de calefacción y a la energía utilizada. El tipo de calefacción puede ser independiente de otros elementos de vivienda, por lo que tal vez haya que revisarlo por separado. Sin embargo, cuando el asiento correspondiente a “energía utilizada para la calefacción” sea desconocido o incongruente, el programa podrá comprobar qué energía se usa para el alumbrado. Por último, si el asiento es inválido y no se utiliza la atribución dinámica habrá que asignar un valor “desconocido”. Los países que utilicen la atribución dinámica para los asientos inválidos deberán tomar dos características, como el tipo de unidad de habitación, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, para obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

18. Agua caliente, disponibilidad de (tema adicional)

649. Este tema se refiere a la disponibilidad de agua caliente en los locales de habitación. Agua caliente significa agua calentada hasta una temperatura determinada y transportada por tuberías y grifos hasta los ocupantes. La información reunida puede referirse a si hay o no agua caliente disponible en el interior del local de habitación, o disponible fuera del local de habitación para uso exclusivo o compartido, o si no se dispone de agua caliente (Naciones Unidas, 2008, párr. 2.502).

650. La disponibilidad de agua caliente puede estar relacionada con el medio de calentar el agua, aunque la utilización de la energía solar para calentarla tal vez no guarde relación con los demás elementos de vivienda. El equipo de revisión deberá decidir el procedimiento apropiado, en función de otros elementos de vivienda y de la ubicación geográfica. En definitiva, si el asiento es inválido habrá que asignar un valor “desconocido” si no se utiliza la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como las relativas a la instalación de agua corriente, para obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

19. Gas por tubería, disponibilidad de (tema adicional)

651. Este tema se refiere a si el local de habitación dispone o no de suministro de gas por tubería. El gas por tubería suele definirse como el gas natural o elaborado

que se distribuye por conducciones y cuyo consumo queda registrado. Este tema puede no ser pertinente para varios países en los que no existe un sistema desarrollado de conducciones de gas o sin fuentes de gas natural (Naciones Unidas, 2008, párr. 2.503).

652. El gas por tubería no está relacionado con los demás elementos de vivienda, excepto el tipo de alumbrado y el combustible utilizado para cocinar. Los equipos de revisión deberán determinar el procedimiento apropiado, así como las comprobaciones de coherencia. Si el asiento sigue siendo inválido o incongruente, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como la energía utilizada para la calefacción, el tipo de edificio, el tipo de unidad de habitación, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

20. Uso de la unidad de habitación (tema adicional)

653. El “uso de la unidad de habitación” indica si dicha unidad se utiliza totalmente para fines de habitación (residenciales) o no. La unidad de habitación puede utilizarse para fines de habitación y para usos comerciales o industriales o de algún otro tipo (Naciones Unidas, 2008, párr. 2.504).

654. El “uso de la unidad de habitación” es independiente de los demás elementos de vivienda. Si el asiento es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de unidad de habitación, el material de construcción de las paredes, la tenencia y el título de propiedad, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

21. Ocupación por uno o más hogares (tema básico)

655. La ocupación por uno o más hogares es independiente de los demás elementos de vivienda. Si el asiento es inválido, el país deberá contar el número de jefes del hogar y utilizar ese número. Es importante señalar que esta revisión debe realizarse después de que la revisión estructural determine el jefe del hogar.

22. Número de ocupantes (tema básico)

656. Se considerará ocupante de una unidad de habitación u otro local de habitación a toda persona que resida habitualmente en él. Por lo tanto, las unidades de enumeración para este tema son los locales de habitación. Sin embargo, como los censos de habitación se suelen levantar al mismo tiempo que los de población, la aplicabilidad de esta definición depende de que la información recogida y registrada sobre cada persona en el censo de población indique dónde se encontraba el día del censo o se refiera a su residencia habitual. Respecto de las personas que se alojan en unidades de habitación móviles, tales como barcos, coches-vivienda y remolques, hay que procurar distinguir entre quienes las utilizan como alojamiento y quienes las utilizan como medio de transporte (Naciones Unidas, 2008, párr. 2.510).

657. El “número de ocupantes” está relacionado con el número de registros de población, y ambas variables deberían ser idénticas. Si no, habrá que adoptar medidas para corregir el número de ocupantes o el de registros de población. Normalmente, el número de ocupantes se modificará para ajustarlo al número de personas presentes en la unidad de habitación. Este tema no debería tener un valor “desconocido” ni atribuido.

23. Tipo de edificio (tema básico)

658. Se recomienda la siguiente clasificación por tipos para los edificios en los que se utiliza algún espacio con fines residenciales:

1. Edificios que constituyen una unidad de habitación
 - 1.1. Físicamente separados
 - 1.2. No físicamente separados
2. Edificios que contienen más de una unidad de habitación
 - 2.1. Hasta dos pisos
 - 2.2. De 3 a 4 pisos
 - 2.3. De 5 a 10 pisos
 - 2.4. Once pisos o más
3. Edificios destinados a personas que viven en instituciones
4. Todos los demás.

659. Si el asiento es inválido, habrá que asignar un valor “desconocido” si no se utiliza la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar al menos dos características, que podrían ser el material de construcción de las paredes exteriores, el período de construcción y/o el tipo de unidades de habitación del edificio, para tener información “conocida” a partir de unidades de habitación similares de la misma zona geográfica.

24. Año o período de construcción (tema adicional)

660. Este tema se refiere a la edad del edificio en que está situado el local de habitación. Se recomienda que se indique el año exacto para los edificios construidos durante el período intercensal inmediatamente precedente si este no pasa de diez años. En los países en que no se haya levantado previamente un censo, o cuando el período intercensal sea de más de 10 años, debe preguntarse el año exacto de construcción para los edificios levantados durante los últimos diez años. Para los edificios construidos antes, la información debería recogerse con referencia a períodos que proporcionen un medio útil de determinar la edad de los locales de habitación existentes. Pueden surgir dificultades al recoger datos sobre este tema porque en algunos casos los ocupantes quizá no sepan cuál es la fecha de construcción (Naciones Unidas, 2008, párr. 2.519).

661. Algunos países, incluso los que utilizan la atribución dinámica, aceptan un valor “desconocido” para el elemento relativo al año o período de construcción. Cuando esto sucede, el país puede decidir no aplicar la atribución dinámica a este elemento, aun cuando utilice matrices de atribución para otras variables. En esta situación, si el asiento es inválido, probablemente debería asignarse un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de edificio, el material de construcción de las paredes exteriores y/o el tipo de unidades de habitación del edificio, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

25. Viviendas en el edificio, número de (tema adicional)

662. La revisión del valor del número de unidades de habitación en un edificio se explica en el capítulo III, como parte de las revisiones estructurales.

26. Materiales de construcción de las paredes exteriores (tema básico)

663. Este tema se refiere a los materiales con que están construidas las paredes exteriores del edificio en que se encuentra el local de habitación. Si las paredes están construidas con más de un material, debe indicarse el material predominante. Los tipos que se distingan (ladrillos, hormigón, madera, adobe) dependerán de los materiales que se usen con más frecuencia en el país y de la importancia que tengan desde el punto de vista de la permanencia de la construcción o para la apreciación de su durabilidad (Naciones Unidas, 2008, párr. 2.525).

664. Si el asiento es inválido, habrá que asignar un valor “desconocido” si no se utiliza la atribución dinámica. Los países que apliquen técnicas de atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el período de construcción y/o el tipo de unidades de habitación del edificio, a fin de obtener información “conocida” a partir de unidades de habitación similares de la misma zona geográfica.

27. Material de construcción del piso y el tejado (tema adicional)

665. En algunos casos el material utilizado para la construcción de los tejados y pisos puede ser de interés especial y puede utilizarse para evaluar mejor la calidad de las viviendas del edificio. Este tema se refiere al material utilizado para el tejado y/o el piso (aunque, según las necesidades concretas de un país, puede referirse también a otras partes del edificio, como la estructura o los cimientos). Solamente se enumera el material predominante, que en el caso del tejado, puede ser tejas, hormigón, planchas metálicas, palma, paja, bambú o algún otro material vegetal, barro o láminas de plástico, entre otros (Naciones Unidas, 2008, párr. 2.528).

666. A veces la respuesta a la pregunta sobre el material de construcción de las paredes exteriores no concuerda con la respuesta a la pregunta sobre el material de construcción del tejado; esto podría suceder, por ejemplo, si el material de construcción indicado para las paredes no es bastante fuerte para soportar el tejado. Como se ha señalado, cuando esto ocurre, los especialistas deberán decidir si habrá que modificar una de las dos variables o utilizar un valor “desconocido”. Si el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de edificio, el tipo de unidad de habitación, el material de construcción de las paredes, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

667. El material de construcción del piso que se ha comunicado puede o no concordar con la construcción del tejado y de las paredes. Si el equipo de revisión del país encuentra combinaciones incongruentes o inválidas, deberá decidir si asigna un valor “desconocido” o si utiliza matrices de atribución para cambiar una o más respuestas. Si el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de edificio, el material de construcción de las paredes exteriores, el tipo de unidad de habitación, la tenencia y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

28. Ascensor, disponibilidad de (tema adicional)

668. Este tema se refiere a la disponibilidad de ascensor (una plataforma cerrada que se eleva y desciende para transportar personas y carga) en un edificio de varios pisos. Se reúne información sobre la disponibilidad de ascensor la mayor parte del tiempo, es decir, que funcione la mayor parte del tiempo, con un mantenimiento ordinario (Naciones Unidas, 2008, párr. 2.529).

669. Si el edificio sólo tiene un piso o se trata de una unidad separada, no deberá haber ascensor. Si lo hay, el equipo de revisión deberá decidir si tiene preferencia el número de pisos o el hecho de que haya ascensor. Si se otorga preferencia al ascensor, habrá que cambiar de número de pisos, asignando a la variable un valor “desconocido” o bien utilizando la atribución dinámica para obtener otro valor. Si se otorga preferencia al número de pisos, y el edificio sólo tiene un piso, la respuesta a la pregunta “disponibilidad de ascensor” deberá cambiarse de manera que sea “no”.

670. Si hay un ascensor eléctrico, habrá que comprobar que haya electricidad en el edificio.

671. Por último, si el asiento correspondiente al ascensor es inválido, habrá que asignar un valor “desconocido” si no se utiliza la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de edificio y el material de construcción de las paredes exteriores, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

29. Edificio de tipo agrícola (tema adicional)

672. Algunos países han considerado necesario especificar si el edificio enumerado era de tipo agrícola o no. Un edificio agrícola es el que forma parte de una explotación agropecuaria y se utiliza para fines agrícolas y/o de vivienda (Naciones Unidas, 2008, párr. 2.531).

673. Los edificios agrícolas son independientes de los demás elementos de vivienda. Los países pueden decidir comprobar si hay correspondencias con los elementos de población referentes a ocupación y rama de actividad económica. De otro modo, si el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

30. Estado de conservación (tema adicional)

674. Este tema se refiere a si el edificio necesita reparaciones y el tipo de ellas que se necesita. La unidad de enumeración es el edificio. La clasificación de los edificios con arreglo al estado de conservación puede incluir los siguientes apartados: “sin necesidad de reparaciones”, “con necesidad de reparaciones menores”, “con necesidad de reparaciones medianas”, “con necesidad de reparaciones mayores” e “irreparable”. Las reparaciones menores se refieren principalmente al mantenimiento habitual del edificio y sus componentes, como una ventana rota. Por reparaciones medianas se entiende la corrección de defectos de mediana importancia, como la falta de desagües en el tejado, grandes superficies de revestimiento desconchado o escaleras sin unas barandillas seguras. Las reparaciones mayores son necesarias en caso de graves defectos estructurales del edificio, como la falta de tejas de madera o de barro en el tejado, la presencia de grietas

y agujeros en las paredes exteriores o la falta de escaleras. Irreparable es un edificio tan necesitado de reparaciones importantes, es decir, con tantos defectos estructurales graves, que se considera más apropiado derribarlo que efectuar reparaciones. Este término se utiliza en la mayoría de los casos para edificios que tienen en pie tan sólo la estructura, sin paredes exteriores enteras y/o sin tejado (Naciones Unidas, 2008, párr. 2.532).

675. El estado de conservación del edificio es independiente de las demás variables de vivienda. Por tanto, si el asiento es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de edificio, el material de construcción de las paredes exteriores y el tipo de unidad de habitación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

31. Características del jefe del hogar u otro miembro de referencia del hogar (tema básico)

676. Las características del jefe del hogar suelen obtenerse de los registros de población para ayudar a elaborar información en tabulaciones cruzadas para fines de planificación y análisis. Estos elementos, que incluyen la etnicidad, la religión o los ingresos, ayudan a determinar diferencias sociales o situaciones de necesidad. Puesto que esas características ya habrán sido revisadas en relación con los elementos de población, en este caso no se requeriría una nueva revisión.

32. Tenencia o título de ocupación (tema básico)

677. Según Naciones Unidas (2008, párr. 2.536), el régimen de tenencia consiste en los arreglos en virtud de los cuales el hogar ocupa toda una unidad de habitación o parte de ella. La unidad de enumeración es un hogar que ocupa una unidad de habitación. La clasificación de los hogares conforme a la tenencia es:

1. Un miembro del hogar es propietario de una unidad de habitación
2. Un miembro del hogar alquila la totalidad o parte de una unidad de habitación
 - 2.1. Un miembro del hogar alquila la totalidad o parte de una unidad de habitación como arrendatario principal
 - 2.2. Un miembro del hogar alquila una parte de una unidad de habitación en calidad de subarrendatario
3. Ocupación sin alquiler
4. Otros arreglos

678. La tenencia puede estar relacionada con el tipo de propiedad, de manera que es posible que el equipo de revisión tenga que examinar la relación entre ambos elementos al elaborar los procedimientos de revisión. Por lo demás, si el asiento correspondiente a la tenencia es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que decidan aplicar este tipo de atribución para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de unidad de habitación, el alquiler y la ocupación o desocupación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

33. Costos de alquiler o de la vivienda ocupada por su propietario (tema adicional)

679. El elemento relativo a los costos de alquiler o de la vivienda ocupada por su propietario es independiente de las demás variables de vivienda excepto en

cuanto que, obviamente, los costos de alquiler deberían producirse solamente en el caso de unidades alquiladas, y los costos de la vivienda ocupada por su propietario sólo deberían producirse en el caso de este tipo de unidades. El equipo de revisión deberá examinar cada caso y determinar las relaciones más apropiadas entre esas variables. Si el asiento es inválido, habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que apliquen este tipo de atribución para los asientos inválidos deberán utilizar por lo menos dos características a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

34. Amueblado o no amueblado (tema adicional)

680. El elemento para indicar si la unidad está o no amueblada es nuevo. Los equipos de revisión deberían examinar la posibilidad de comprobar el elemento, si se incluye, para determinar cuáles convendría utilizar en la atribución dinámica, si se emplea ese método para resolver las respuestas inválidas o incongruentes.

35. Sistemas de tecnología de la información y las comunicaciones, disponibilidad de (tema básico)

681. La importancia de la disponibilidad de sistemas de tecnología de la información y las comunicaciones (TIC) está aumentando de forma significativa en la sociedad actual. Estos sistemas prestan una serie de servicios que están transformando la estructura y la pauta de los grandes fenómenos sociales y económicos. El censo de habitación ofrece una oportunidad singular de evaluar la disponibilidad de estos sistemas en los hogares. La gama de temas incluidos debería ser lo bastante amplia para que se pueda comprender el lugar de la TIC en el hogar, para que esa información se pueda utilizar en las actividades de planificación del gobierno y del sector privado de manera que permita una prestación más amplia y mejorada de servicios, y para evaluar su impacto en la sociedad (Naciones Unidas, 2008, párr. 2.543). La clasificación recomendada es la siguiente:

1. Hogares que tienen radio
2. Hogares que tienen televisor
3. Hogares que tienen teléfono de línea fija
4. Hogares que tienen teléfono(s) celular(es)
5. Hogares que tienen computadora(s) personal(es)
6. Hogares que tienen acceso a Internet desde casa
7. Hogares que tienen acceso a Internet desde otro lugar
8. Hogares que no tienen acceso a Internet

682. Los sistemas de tecnología de la información y las comunicaciones son elementos nuevos. Los artículos que requieren electricidad sólo estarán disponibles en los lugares donde llegue la electricidad. No obstante, a medida que se generaliza el uso de la energía solar y de la eólica y de otras fuentes de energía “renovable”, deberá tenerse en cuenta este factor al preparar las revisiones de este elemento. Los equipos de revisión de los países deben comprobar a fondo esta información y sus matrices de atribución antes de realizar el censo o la encuesta. Entre las informaciones útiles para la atribución dinámica se incluyen el nivel social de los hogares (determinado por el índice de riqueza, por ejemplo) y la edad del jefe del hogar.

683. Este tema se refiere a la existencia de un teléfono en el local de habitación. Por ejemplo, un teléfono denota una línea telefónica más que un teléfono físico, ya que es posible conectar más de un teléfono con una sola línea telefónica (Naciones Unidas, 2008, párr. 2.547 y 2.548). Los teléfonos no guardan relación con los demás elementos de vivienda durante la revisión, aunque si ciertas zonas geográficas no tienen teléfonos, el equipo de revisión deberá tenerlo en cuenta al formular los procedimientos del caso. Si el asiento correspondiente a “teléfono” es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de unidad de habitación, el material de construcción de las paredes y la tenencia, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

36. Automóviles, número de (tema adicional)

684. El “número de automóviles” se refiere a la cantidad de automóviles y furgonetas de que disponen normalmente los miembros del hogar para su uso. “Normalmente disponibles” significa los automóviles y furgonetas que son propiedad de los ocupantes o que estos utilizan en virtud de un acuerdo más o menos permanente, como el arriendo, y los suministrados por un empleador para uso del hogar, pero con exclusión de las furgonetas utilizadas exclusivamente para el transporte de mercancías (Naciones Unidas, 2008, párr. 2.551).

685. El número de vehículos es independiente de las demás variables de vivienda. Si en el país hay zonas sin vehículos, los especialistas tal vez deseen estudiar procedimientos especiales de revisión para determinadas zonas geográficas. De otro modo, si el asiento es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de unidad de habitación, el material de construcción de las paredes, la tenencia y el título de propiedad, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

37. Aparatos electrodomésticos, disponibilidad de (tema adicional)

686. Se reúne información sobre la disponibilidad de aparatos electrodomésticos, como lavadoras, lavavajillas, refrigerador, congelador, televisor y ordenador personal, según las circunstancias nacionales (Naciones Unidas, 2008, párr. 2.552).

687. En la mayoría de los casos, en la unidad de habitación deberá haber electricidad para que funcionen los aparatos. Cuando existan estos elementos, el equipo de revisión deberá considerar un procedimiento para comprobar si existe electricidad (con las posibles excepciones de un refrigerador que pueda funcionar con gas, o de una nevera de hielo). Además, si en el país de que se trate se requiere agua corriente para hacer funcionar una lavadora o un lavavajillas, el procedimiento de revisión deberá tenerlo también en cuenta. Pueden utilizarse procedimientos para determinar si debe existir un elemento particular, según se disponga de electricidad y de agua, y habrá que adoptar las medidas apropiadas cuando aparezcan incongruencias. También es posible que ciertas partes de un país no tengan electricidad o agua corriente, y tal vez los especialistas necesiten reconocer este hecho al formular sus procedimientos de revisión. Si el asiento es inválido o incongruente habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos

características, como el tipo de unidad de habitación, la electricidad, el material de construcción de las paredes y la tenencia, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica (pues los niveles sociales de los hogares deben ser similares).

38. Espacio exterior disponible (tema adicional)

688. Este punto se refiere a la existencia de un espacio exterior para actividades recreativas de los miembros del hogar que ocupan una unidad de habitación. La clasificación se refiere al espacio exterior disponible como parte de la unidad de habitación (por ejemplo, el patio trasero en el caso de una casa separada), al espacio exterior adyacente al edificio (como patios traseros y patios de juegos situados junto a un edificio de apartamentos), al espacio exterior disponible como parte de zonas recreativas comunes en un radio de diez minutos a pie desde la unidad de habitación (por ejemplo, parques, centros deportivos, etcétera) o a la no existencia de un espacio exterior utilizable en un radio de diez minutos a pie (Naciones Unidas, 2008, párr. 2.553).

689. El espacio exterior disponible para uso del hogar es independiente de otros elementos de vivienda. Sin embargo, es posible que en ciertas zonas geográficas o en ciertos tipos de edificios no se disponga de espacio exterior. Los equipos de revisión deberán estudiar las circunstancias concretas al formular sus procedimientos. Si el asiento es inválido habrá que asignar un valor “desconocido” cuando no se utilice la atribución dinámica. Los países que opten por la atribución dinámica para los asientos inválidos deberán utilizar por lo menos dos características, como el tipo de edificio y el tipo de unidad de habitación, a fin de obtener información “conocida” de unidades de habitación similares de la misma zona geográfica.

B. Unidades de habitación ocupadas y desocupadas

690. Los procedimientos de revisión descritos se refieren a las unidades de habitación ocupadas. Sin embargo, las unidades de habitación desocupadas y las ocupadas presentan a menudo características distintas y no requieren las mismas comprobaciones. El equipo de revisión de la oficina nacional de estadística o del censo deberá formular diferentes procedimientos de revisión para cada tipo de unidad cuando, como suele suceder, no se recopilen todos los elementos de vivienda para las unidades de habitación desocupadas. El equipo de revisión deberá prestar especial atención a las variables de la matriz de atribución, ya que estas son las que más probablemente diferirán.

691. En este capítulo se han considerado las variables de población recomendadas en *Principios y recomendaciones para los censos de población y habitación, Revisión 2*. Ningún país debería utilizar todas estas variables, y las variables seleccionadas y sus relaciones especiales con otras variables deberían comprobarse a fondo en condiciones de invernadero y en encuestas previas al censo, para determinar la existencia de respuestas fiables y completas. Las variables relativas a la vivienda son importantes para la elaboración de parte del índice de riqueza, que permite evaluar el bienestar en todas o en algunas partes de un país.

Anexo I

Variables derivadas

1. Para sacar el máximo provecho de los datos de sus censos o encuestas, los países necesitan a menudo variables que son combinaciones de otras variables. Por ejemplo, el elemento relativo a la actividad económica (véase el capítulo IV, sección D.1) ya es una combinación de distintas variables recogidas en el censo. En lugar de tener que elaborar un programa para volver a codificar la información cada vez que la oficina nacional de estadística o del censo necesita una tabulación especial, los especialistas en procesamiento de datos pueden escribir un programa que efectúe la recodificación una sola vez, almacene la información recodificada en el registro de la persona y la utilice para nuevas tabulaciones cuando sea menester. Al determinar si deberán o no producir y almacenar la información, las oficinas nacionales de estadística o del censo deberán decidir con qué frecuencia se utilizarán las recodificaciones en las tabulaciones y qué pertinencia tendrá una recodificación particular. Es importante recordar que las recodificaciones también ocupan espacio en los registros personales. Cuanto mayor sea el tamaño de la población, más espacio se utilizará.

2. De esta forma pueden crearse muchas variables. Por ejemplo, si se comunica una fecha de nacimiento pero no la edad, esta puede determinarse sustrayendo la fecha de nacimiento de la fecha de referencia del censo, y esa información se almacenará en el registro. Del mismo modo, los ingresos del hogar pueden obtenerse sumando los ingresos de cada persona e incluyendo la suma en el registro de vivienda, para su uso posterior.

3. A veces las variables derivadas proceden de una combinación de uno o varios asientos de un único registro; otras veces provienen de varios registros. Por ejemplo, para clasificar a una persona como “no económicamente activa—asiste a un establecimiento docente” puede ser necesario examinar las respuestas de hasta cuatro elementos. Al elaborar los formatos de los cuadros o al planear cuadros complementarios, la utilización de variables derivadas hará que la programación sea más fácil y más eficiente y ayudará también a que los datos sean comparables en el tiempo. A continuación se ofrecen algunos ejemplos de posibles registros derivados.

A. Variables derivadas para los registros de vivienda

1. Ingresos del hogar

4. La variable derivada de los ingresos del hogar es la suma de los ingresos obtenidos en todas las categorías de ingresos por todas las personas de un hogar. Las categorías de información sobre los ingresos pueden incluir los sueldos, los ingresos por negocios propios, los intereses y dividendos, los ingresos por pagos de seguridad social y de jubilación, las remesas, las regalías y los alquileres. Si también se recopilan datos sobre los ingresos totales, durante la revisión habrá que comprobar los ingresos totales de cada persona sumando cada una de las categorías. A continuación el total se

compara con los ingresos totales registrados. Si los ingresos sumados no son iguales a los totales declarados, el equipo de revisión deberá formular un plan de corrección, según el cual habrá que modificar el total de forma que refleje la suma de las partes, o bien habrá que modificar una o más de las categorías de ingresos de las personas. Cuando se determinan los ingresos totales de todas las personas de un hogar, la variable correspondiente a ingresos del hogar se obtiene sumando los ingresos personales.

5. El equipo de revisión deberá tener en cuenta la situación en que una o más personas del hogar tengan ingresos negativos a causa de la quiebra de un negocio o por otros motivos. En tal caso, los ingresos totales del hogar se verán reducidos, y no aumentados, a causa de los ingresos negativos de esa persona particular.

2. Ingresos familiares

6. La variable derivada de los ingresos familiares es la suma de los ingresos obtenidos en todas las categorías de ingresos por todas las personas de una familia. A diferencia de los hogares, las familias suelen estar integradas sólo por personas emparentadas, pero esta definición dependerá de la situación particular del país. En algunos países los hogares y las familias serán lo mismo, por lo que no se necesitará una variable derivada para los ingresos familiares. Las categorías de información sobre los ingresos familiares pueden incluir los sueldos, los ingresos por negocios propios, los intereses y dividendos, los ingresos por pagos de seguridad social y jubilación, las remesas, las regalías o los alquileres. Si también se recopilan datos sobre los ingresos totales, durante la revisión habrá que comprobar los ingresos totales de cada persona sumando cada una de las categorías. A continuación el total se compara con los ingresos totales registrados. Si los ingresos sumados no son iguales a los ingresos totales declarados, el equipo de revisión deberá formular un plan de corrección. O bien habrá que modificar el total de forma que refleje la suma de las partes, o modificar una o más de las categorías de ingresos de las personas. Cuando se determinan los ingresos totales de todas las personas, los ingresos de la familia se obtienen sumando los ingresos personales dentro de la familia.

7. El equipo de revisión deberá tener en cuenta la situación en que una o más personas de la familia tengan ingresos negativos por quiebra de un negocio o por otros motivos. En tal caso, los ingresos totales de la familia se verán reducidos, y no aumentados, a causa de los ingresos negativos de esa persona particular.

3. Núcleo familiar

8. En cuanto a la composición de los hogares, en *Principios y recomendaciones para los censos de población y habitación, Revisión 2*, se elaboró un código para el núcleo familiar, decidido según una de las siguientes formas (la reclasificación se indica entre paréntesis en itálica):

1. **Pareja casada** (o pareja que vive en unión consensual) **sin hijos** (cabeza del hogar y cónyuge o jefes conjuntos del hogar o una pareja que vive en unión consensual)
2. **Pareja casada** (o pareja que vive en unión consensual) **con uno o más hijos solteros** (igual que supra, pero, mediante una búsqueda del hogar o una recodificación del número de hijos casados en la unidad de habitación, al menos un hijo soltero)
3. **Un padre con uno o más hijos solteros** (cabeza del hogar varón, sin esposa presente, con al menos un hijo no casado, determinado como supra)

4. **Una madre con uno o más hijos casados** (*cabeza del hogar femenino, sin esposo presente, con al menos un hijo soltero determinado como supra*)

4. Tipo de hogar

9. En *Principios y recomendaciones para los censos de población de habitación, Revisión 2*, se incluye una característica general de varios tipos de hogares para ayudar a realizar una recodificación de la composición de los hogares. Los países pueden optar por realizar una sola recodificación, o una serie de recodificaciones, según el uso potencial de los datos.

10. Una primera recodificación podría identificar el tipo de hogar, representado por los siguientes elementos, incluidas las definiciones. Las recodificaciones propuestas se presentan en la siguiente sección.

1. Hogar unipersonal
2. Hogar nuclear: Una sola familia nuclear, es decir, un matrimonio o una pareja en unión consensual, con o sin hijos, o progenitor solo con un hijo o hijos.
3. Hogar extenso: Un solo núcleo familiar y otras personas emparentadas con el cabeza del hogar; dos o más núcleos familiares; o dos o más personas emparentadas entre sí pero ninguna de las cuales forma un núcleo familiar
4. Hogar compuesto (otros tipos de hogares)

5. Composición de los hogares

11. *Los hogares unipersonales* son hogares, más que familias, y por lo tanto deberían incluirse como categoría independiente en la recodificación de la composición de los hogares.

12. *Hogares de familias nucleares*. Los hogares de familias nucleares pueden dividirse en (y pueden recibir códigos individuales relativos a): 1) matrimonios con hijos, 2) matrimonios sin hijos, 3) parejas en unión consensual con hijos, 4) parejas en unión consensual sin hijos, 5) padres con hijos y 6) madres con hijos. Para determinar el código adecuado, se determina el sexo del cabeza del hogar, seguido de búsquedas de un cónyuge y de hijos en el hogar. El tipo de hogar nuclear podría ser un código de dos dígitos con el valor 2 utilizado para el primero de los dos dígitos (el código 1 se reservaría para los hogares unipersonales); así, el código 21 representaría un matrimonio con hijos.

13. *Hogares de familias extensas*. Las familias extensas pueden dividirse también en categorías, en las que se incluirían las siguientes (de acuerdo con las designaciones precedentes): (31) un solo núcleo familiar y otras personas emparentadas con él; (32) dos o más núcleos familiares emparentados entre sí, sin ninguna otra persona; (33) dos o más núcleos familiares emparentados entre sí, más otras personas emparentadas con al menos uno de los núcleos, y (34) dos o más personas emparentadas entre sí pero ninguna de las cuales forma un núcleo familiar. Los códigos reales se determinarían buscando en el hogar las cifras de núcleos y relaciones entre las personas del mismo. Si un hogar está ya codificado como nuclear, el procedimiento no se llevaría a cabo.

14. *Hogares compuestos*. Todos los demás hogares serían hogares compuestos. Valiéndose del sistema anterior se obtendrían los siguientes resultados: (41) un solo núcleo familiar más otras personas, algunas de las cuales están emparentadas con el núcleo, mientras que otras no lo están; (42) un solo núcleo familiar más otras personas,

ninguna de las cuales está emparentada con el núcleo; (43) dos o más núcleos familiares emparentados entre sí más otras personas, algunas de las cuales están emparentadas por lo menos con uno de los núcleos y otras no están emparentadas con ninguno de los núcleos; (44) dos o más núcleos familiares emparentados entre sí más otras personas, ninguna de las cuales está emparentada con ninguno de los núcleos; (45) dos o más núcleos familiares no emparentados entre sí, con o sin otra persona; (46) dos o más personas emparentadas entre sí, pero ninguna de las cuales forma un núcleo familiar, más otras personas no emparentadas, y (47) personas no emparentadas. También en ese caso una serie de búsquedas y de resúmenes darán lugar a la designación adecuada de cada tipo de hogar.

6. Composición de la familia

15. Las familias son un subconjunto de hogares; por ello, la recodificación de la composición de la familia incluirá las categorías pertinentes para las familias antes descritas. En la medida en que un hogar unipersonal no constituye una familia, no se incluirá en la recodificación de la composición de la familia. Asimismo, los hogares compuestos son hogares pero no familias, por lo que tampoco se incluirán. Cada país decide luego si desea incluir una única recodificación para todas las familias (familias nucleares y extensas) o recodificaciones independientes para las familias nucleares y las familias extensas, dándose por supuesto que estas recodificaciones no deben superponerse (aunque podría argumentarse la conveniencia de incluir los hogares de familias nucleares junto con las familias extensas en el caso de todas las familias).

7. Posición en el hogar y en la familia

16. La posición en el hogar y en la familia indica la relación de una persona con otros miembros del hogar o de la familia. El criterio para determinar la posición en el hogar y en la familia se distingue del método tradicional de clasificar a los miembros del hogar solo por su relación con el jefe del hogar o con la persona de referencia.

17. En *Principios y recomendaciones para los censos de población de habitación, Revisión 2*, se incluye el siguiente código de codificación propuesto para la posición en el hogar. El primer conjunto de códigos corresponde a las personas de los hogares con al menos un núcleo familiar (es decir, hogares que son también familias). Entre las determinaciones propuestas de la recodificación se incluyen las siguientes:

- 1.1. Esposo (hombre jefe de hogar o cónyuge varón)
- 1.2. Esposa (mujer jefe de hogar o cónyuge femenina)
- 1.3. Pareja en unión consensual o pareja cohabitante (a partir de los códigos de parentesco, en su caso, o de la combinación de códigos de parentesco y estado civil)
- 1.4. Madre sola (sobre la base de un marido ausente para una mujer, pero con hijos presentes)
- 1.5. Padre solo (sobre la base de una esposa ausente para un varón, pero con hijos presentes)
- 1.6. Hijo que vive con ambos progenitores (hijo del cabeza del hogar, con ambos progenitores presentes en la casa)
- 1.7. Hijo que vive con la madre sola (hijo del cabeza del hogar, pero el padre del hijo no está presente)
- 1.8. Hijo que vive con el padre solo (hijo del cabeza del hogar, pero la madre del hijo no está presente)

- 1.9. Persona que no es miembro de un núcleo familiar (cualquier otro familiar). En *Principios y recomendaciones para los censos de población y habitación, Revisión 2*, este elemento se divide en dos grupos, según que se trate de personas que 1) viven con parientes o 2) viven con personas no emparentadas.

18. El segundo conjunto para la recodificación corresponde a las personas de los hogares sin un núcleo familiar: personas que viven solas y personas que viven con otros parientes y/o personas no emparentadas, sin incluir el cónyuge o el hijo del cabeza del hogar. Las categorías son las siguientes:

- 2.1. Personas que viven solas (hogar unipersonal)
- 2.2. Personas que viven con otros (persona que vive en una unidad de habitación sin el cónyuge o un hijo del cabeza del hogar). Esta categoría se divide a su vez en personas que viven 1) con hermanos, 2) con otros parientes distintos de los hermanos o 3) con personas no emparentadas.

19. Debería elaborarse una variable única a partir de estas categorías, ya que se excluyen mutuamente. La variable debería tener dos dígitos. Algunos organismos de estadística quizá deseen que el primer dígito sea independiente del segundo; en este caso, el primer dígito indicará si el hogar es un núcleo familiar o no, y el segundo especificará el tipo de posición de una persona dentro del hogar.

20. En *Principios y recomendaciones para los censos de población de habitación, Revisión 2*, se incluyen también categorías para la clasificación de una persona de acuerdo con la situación familiar. Entre ellas se incluyen las siguientes: 1) hombre o mujer de una pareja formada por un cabeza del hogar y un cónyuge con o sin hijos; 2) progenitor sólo, por sexo; 3) hijo del cabeza del hogar, hijo de una pareja casada o hijo de un progenitor solo, por sexo del progenitor, y 4) persona que no es miembro del núcleo familiar (no emparentada o emparentada y, en este último caso, de qué manera). Las determinaciones antes señaladas con respecto a la posición en el hogar pueden utilizarse también para la situación familiar.

8. Efectos del VIH/SIDA en la estructura de los hogares

21. Dados los efectos de la epidemia del VIH/SIDA en la estructura de los hogares de tantos países, una recodificación puede ayudar a describir los diferentes tipos de unidades de habitación. Por ejemplo, una recodificación que describa los hogares con una generación perdida (solo abuelos y nietos), los hogares con jefes de menos de 18 años de edad, los hogares presididos por viudas, etcétera, pueden utilizarse para evaluar el efecto social y económico de la epidemia, aunque sea en forma indirecta. Los niños que entran y salen de la fuerza de trabajo, la estructura de la fuerza de trabajo dentro de esos hogares, etcétera, pueden ayudar a los encargados gubernamentales de la planificación a describir detalladamente los efectos del VIH/SIDA.

9. Familiares

22. Los familiares son las personas que tienen algún grado de parentesco con el jefe del hogar. La variable derivada para los familiares es la suma de todas las personas emparentadas con el jefe del hogar. Este valor es particularmente importante en situaciones en que hay un gran número de personas no emparentadas que viven juntas en una misma unidad de habitación. Cuando muchas personas no emparentadas viven juntas de esta manera, a menudo se indica que viven en “locales colectivos” o “locales de grupo”.

23. Al elaborar los conjuntos de datos, las oficinas nacionales de estadística a menudo crean variables derivadas para diferentes conjuntos de personas emparentadas, por edad. Por ejemplo, podrían elaborarse variables derivadas para niños emparentados de 0 a 5 años de edad, niños emparentados de 5 a 17 años de edad, niños emparentados de 6 a 17 años de edad, niños emparentados de 0 a 17 años de edad, personas emparentadas de 65 años de edad y más, y personas emparentadas de 75 años de edad y más.

24. Los “niños emparentados” de una familia podrían comprender, por ejemplo, los hijos propios del jefe del hogar y otras personas menores de 18 años de edad que vivan en el hogar, independientemente de su estado civil, y que estén emparentadas con el jefe del hogar, excepto el cónyuge de la persona jefe del hogar. Los niños emparentados pueden o no incluir los hijos adoptivos, ya que estos no están emparentados con el jefe del hogar; pero esta decisión dependerá de la situación del país.

10. Trabajadores de la familia

25. A veces los países desean comparar ciertas variables del hogar por número de trabajadores, como las distribuciones de ingresos por tamaño del hogar y los trabajadores por persona a cargo. El país podría obtener la variable derivada para el número de trabajadores de la familia sumando el número de personas que trabajaron por lo menos una hora durante un período de referencia, como una semana o un año (o bien un año civil o bien los últimos 12 meses). El país podría utilizar el número de personas que trabajaron “la semana pasada” si sólo se reúnen datos para ese período.

11. Instalación completa de agua corriente

26. Varios elementos del cuestionario del censo pueden servir para obtener datos sobre el agua corriente. Suelen estar relacionados con la presencia de agua traída por conducciones, un retrete con descarga y una bañera o ducha, y normalmente tanto se obtienen en las unidades de habitación ocupadas o en las desocupadas. Una variable derivada sobre la instalación completa de agua corriente puede ser de ayuda para comparar las condiciones socioeconómicas entre zonas o grupos en un momento determinado o en el tiempo. La variable derivada sobre instalación completa de agua corriente puede obtenerse, por ejemplo, cuando existan (dentro de la unidad o fuera del edificio en que está ubicada la unidad) tres instalaciones: conducciones de agua (caliente o fría), retrete con descarga y baño o ducha. El equipo de revisión deberá determinar el conjunto de variables más apropiado para “instalación completa de agua corriente”.

27. En este ejemplo la variable derivada puede obtenerse cuando se pregunta por separado sobre los tres elementos, y durante la operación de revisión podrá determinarse la suma de la presencia de los tres elementos. Si la unidad de habitación cuenta con conducciones de agua, un retrete con descarga y un baño o ducha, tendrá una “instalación completa de agua corriente”. Sin la totalidad de los tres elementos, la unidad “carece de instalación completa de agua corriente”.

12. Cocina completa

28. Los censos pueden utilizarse para obtener datos sobre instalaciones para cocinar a partir de los elementos del cuestionario relacionados con el equipo de cocina, el refrigerador y la pila; los elementos se reúnen tanto para las unidades de habitación ocupadas como para las desocupadas. Puede considerarse que una unidad tiene “instalaciones de cocina completa” cuando las instalaciones para cocinar (una estufa eléc-

trica, de petróleo o de gas, un horno de microondas y fogones no portátiles o bien un hornillo para cocinar), un refrigerador y una pila con agua corriente están ubicados en el mismo edificio que la unidad de habitación que se está enumerando. No es necesario que estén en el mismo cuarto.

29. La variable derivada se obtiene cuando se pregunta por separado sobre los tres elementos y durante la operación de revisión se determina la suma de la presencia de los tres elementos. “Carece de instalaciones de cocina completa” denota los casos en que están presentes tres instalaciones de cocina específicas pero el equipo está ubicado en un edificio distinto; cuando están presentes algunas de las instalaciones pero no todas, o cuando ninguna de las tres instalaciones de cocina especificadas está presente en el mismo edificio que el local de habitación que se está enumerando.

13. Alquiler bruto

30. Los países suelen recoger datos sobre el alquiler en efectivo, es decir, el que consta en el contrato. El alquiler en efectivo suele excluir el costo de los servicios. A veces los países también necesitan información sobre el alquiler bruto. Este es el alquiler en efectivo que figura en el contrato más la suma media mensual estimada del costo de los servicios (electricidad, gas y agua) y los combustibles (como el petróleo, el carbón, el queroseno y la leña) si su pago corre a cargo del inquilino. La información sobre la renta bruta tiene el propósito de eliminar las diferencias resultantes de las distintas prácticas en cuanto a la inclusión de los servicios y los combustibles como parte del pago del alquiler. Las unidades de alquiler ocupadas sin que medie el pago de un alquiler en efectivo pueden presentarse por separado en las tabulaciones con el título “sin alquiler en efectivo”.

31. La variable derivada para el alquiler bruto se obtiene sumando el monto del alquiler pagado y el pagado por concepto de servicios, si estos se cobran por separado.

14. Índice de riqueza

32. El índice de riqueza mide el bienestar de un país o de determinadas partes del país. En la mayoría de los casos, el índice se elabora utilizando los activos de los hogares. Muchas veces se recurre al análisis sectorial para obtener el mejor conjunto de elementos y sus variantes. En general se atribuye a los elementos valores binarios —1 para “presente” y 0 para “ausente”—, y luego se suman los valores. Cuanto mayor sea el valor, mayor será la riqueza. Por ejemplo, cuando se tiene un televisor, el código sería 1; cuando no se tiene, el valor sería 0. Por otro lado, en relación con el tipo de retrete, el código sería 1, 2 o 3 según que se tratara de retretes externos, de retretes de descarga por gravedad o de descarga (lo que significaría que hay tres conjuntos de variables binarias). Los distintos elementos podrían ponderarse en la suma final.

33. Pueden crearse quintiles tomando cada quinta parte de la distribución de los valores del índice de riqueza. El quintil más bajo estaría integrado por los hogares más pobres; el quintil más alto correspondería a los hogares más ricos.

B. Variables derivadas para los registros de población

1. Situación laboral

34. Una variable derivada sobre la situación laboral puede ser muy útil para las tabulaciones, pero para ello se necesita información de distintas variables. Si se

siguen las categorías de los *Principios y recomendaciones para los censos de población y habitación, Revisión 2*, se necesitará reconfigurar diversas variables. La variable derivada para los registros de población según la situación laboral podría constar de dos categorías, divididas en seis subcategorías:

1. Económicamente activo
 - 1.1. Empleado
 - 1.2. Desempleado
2. No económicamente activo
 - 2.1. Estudiante
 - 2.2. Trabajador en el hogar
 - 2.3. Persona que vive de una pensión o de rentas del capital
 - 2.4. Otras categorías.

35. Puesto que las distintas clasificaciones de la actividad económica se utilizan en muchos de los cuadros conexos, el equipo de revisión deberá considerar la posibilidad de introducir una variable derivada en los registros de datos en vez de hacer que los encargados del procesamiento de datos reclasifiquen la situación laboral durante la tabulación. La reclasificación durante la tabulación puede introducir errores, ya que distintos encargados de procesar los datos podrían efectuar la reclasificación de formas ligeramente distintas; incluso un mismo programa podría hacer la reclasificación de manera distinta, dependiendo de las necesidades particulares de la revisión o de la tabulación. Los especialistas en características económicas deberán preparar las especificaciones de la variable derivada.

2. Hijos propios

36. A veces los países desean producir información acerca de los “hijos propios”, que son los hijos biológicos del jefe del hogar y/o del cónyuge. En los cuadros podrían indicarse los “hijos propios” clasificados como hijos que viven con los dos progenitores o sólo con uno de ellos.

37. La variable derivada para “hijos propios” podría ser la suma del número de hijos propios de una persona particular, normalmente de sexo femenino, siguiendo las definiciones seleccionadas por los equipos de revisión. A veces los usuarios necesitan información más detallada sobre los hijos propios por edad. En los Estados Unidos, por ejemplo, se establecen variables derivadas para el número de hijos propios menores de seis años de edad y de entre 6 y 17 años de edad. Esos valores figuran en los registros de todas las mujeres. La información se utiliza especialmente para determinar las características de las mujeres que participan en la fuerza de trabajo y que tienen “hijos propios”.

3. Progenitores en el hogar

38. Esos datos se refieren a las características de los hijos en familias monoparentales en comparación con las unidades de habitación en las que residen ambos progenitores. En la revisión se obtiene esta variable derivada determinando cuántos progenitores de una persona particular están en el hogar, sobre la base de los códigos de relación. El programa examina el código de relación para cada hijo y utiliza esa información en combinación con la información sobre las subfamilias para determinar si uno o ambos progenitores viven en la unidad de habitación.

4. Año en curso en la escuela

39. Algunos países formulan dos preguntas sobre la educación:

- a) Si la persona asiste a la escuela actualmente;
- b)Cuál es el máximo nivel de instrucción alcanzado.

40. En esos países los equipos de revisión suelen encontrar una falta de correspondencia entre los dos elementos cuando una persona asiste efectivamente a la escuela en el momento del empadronamiento. A veces esto puede dar lugar a que el máximo nivel de instrucción alcanzado por la persona sea un año inferior al año escolar actual. Si la persona se encuentra a la mitad de una serie de grados o niveles, las estadísticas no se verán afectadas. Sin embargo, si la persona asiste al primer grado de una serie de un nivel particular, tal vez no será posible la correspondencia con datos de otras fuentes. Por ejemplo, una persona que asista a la escuela en el primer grado figurará como persona que asiste a la escuela pero sin ningún nivel de instrucción. Análogamente, una persona que esté comenzando la escuela secundaria constará como persona que está en la escuela, aunque su máximo nivel de instrucción será el grado superior de la escuela primaria.

41. Para esta combinación de elementos puede formularse una variable derivada denominada “asistencia escolar en el año en curso”. Si la persona no está asistiendo a la escuela en ese momento, el código será el mismo que el del máximo nivel de instrucción alcanzado. Si la persona está asistiendo a la escuela, la revisión agregará un grado al máximo nivel alcanzado y asignará ese valor al “año en curso en la escuela”.

42. Algunos países formulan tres preguntas sobre la educación: las dos indicadas y un tercer elemento relativo a si se ha concluido el nivel máximo. Si también se obtiene esa información, debería utilizarse para determinar el “año en curso en la escuela”.

5. Meses transcurridos desde el último nacimiento

43. Si se recoge información sobre la fecha del último nacimiento, puede realizarse una recodificación para obtener estimaciones indirectas de la fecundidad de edades específicas, año por año y de la del total. La recodificación considera la fecha de empadronamiento (normalmente el mes y el año y lo reduce todo a meses), la fecha del último nacimiento (y convierte todo ello a meses) y luego resta para obtener el número de meses desde el último nacimiento. Estas cifras se archivan en el registro de la mujer para ayudar a determinar las estimaciones de la fecundidad año por año.

Anexo II

Relación entre el formato del cuestionario y la introducción de datos

1. Los dos formatos más habituales de cuestionarios de población de un censo o una encuesta son las páginas personales y las páginas sobre el hogar.

2. Las páginas personales consisten en una o dos páginas de información sobre población, con páginas separadas para cada persona. Este método es útil porque toda la información sobre una persona figura en la misma página y es más fácil recopilarla. Además, este formato facilita la comprobación de la coherencia interna durante el empadronamiento. Las páginas personales pueden juntarse en una carpeta para facilitar su uso sobre el terreno (véase el gráfico A.II.1).

Gráfico A.II.1

Modelo de impreso de cuestionario con páginas personales

Página personal de la persona X		Página personal de la persona X+1	
Elemento 1	Elemento 10	Elemento 1	Elemento 10
Elemento 2	Elemento 11	Elemento 2	Elemento 11
.	.	.	.
.	.	.	.
.	.	.	.

3. La codificación y la introducción de datos en las páginas personales es una operación básicamente mecánica, en la que no se espera que el codificador o el encargado de introducir los datos evalúen la validez de la información suministrada, sino que asignen el código apropiado o introduzcan el valor correspondiente. El gráfico A.II.2 ilustra la secuencia de información para una persona determinada registrada en una única página. Será más fácil introducir los datos de esa persona en una sola página

Gráfico A.II.2

Ejemplo de secuencia en un cuestionario con páginas personales

Página personal	
Elemento 1	Elemento 11
Elemento 2	Elemento 12
Elemento 3	Elemento 13
Etcétera	Etcétera
.	
.	
.	

que tener que pasar las páginas. Las comprobaciones de validez se efectuarán más adelante, durante las revisiones por ordenador.

4. Las páginas sobre el hogar contienen toda la información correspondiente a un hogar en una sola página, siempre que es posible, o en una serie de páginas, con todos los miembros del hogar enumerados en cada página. Enumerar de esta forma a los miembros del hogar es útil porque no es necesario imprimir los elementos del cuestionario para cada persona, lo cual permite ahorrar espacio. Además, el empadronador puede comparar los asientos para los distintos miembros del hogar a medida que se van recopilando los datos (véase el gráfico A.II.3).

Gráfico A.II.3

Modelo de página sobre un hogar en un cuestionario con todas las personas en la misma página

Página sobre el hogar					
Persona	Elemento 1	Elemento 2	Elemento 3	Elemento 4	Etcétera
1					
2					
3					
4					
5					
.					
.					
.					

5. Un tercer método consiste en que haya impresos separados para cada persona y que el empadronador los junte en una carpeta de hojas sueltas durante el empadronamiento o después. Este método es eficiente porque el empadronador sólo junta el número exacto de impresos (páginas) necesario para el hogar. El inconveniente es que los impresos pueden separarse durante la transferencia u otras operaciones, lo cual puede hacer que se creen problemas de revisión y de cobertura si la oficina del censo no puede reunir todos los impresos que corresponden a un hogar concreto.

6. El tamaño de las páginas del cuestionario es importante tanto para el empadronamiento como para la introducción de los datos. Durante la codificación y la introducción de datos el documento debe poder estar extendido sobre el escritorio, de modo que los codificadores o los encargados de la introducción de datos puedan encontrar fácilmente los elementos en el impreso y realizar las operaciones necesarias.

7. Cuando toda la información está en la misma página, el personal puede introducir también los datos de las páginas sobre el hogar con facilidad, y obviamente con más rapidez, ya que el encargado de introducir los datos no tendrá que pasar las páginas. El gráfico A.II.4 ofrece un ejemplo de la secuencia de información en una página sobre un hogar.

8. Al introducir datos sobre población o sobre vivienda que ocupen más de una página pueden producirse problemas. Para resolverlos, la oficina nacional de estadística probablemente adoptará uno de los dos criterios que se indican a continuación.

9. Pueden introducirse de una vez los datos de una misma persona. El encargado de la introducción de los datos puede introducir la línea de información correspondiente a la persona en la primera página de la serie de páginas y después pasar

Gráfico A.II.4

Ejemplo de secuencia en un cuestionario con páginas sobre un hogar y varias personas por página

Página sobre el hogar					
Persona	Elemento 1	Elemento 2	Elemento 3	..	Etcétera
1					
2					
3					
4					
5					
.					
.					
.					

a la página segunda y siguientes. Al terminar las páginas de la primera persona, el encargado de la introducción de datos regresa a la primera de las páginas sobre el hogar e introduce los datos de la segunda persona; después pasará a la tercera persona, y así sucesivamente. Este tipo de introducción de datos funcionará bien en tanto que el operador permanezca en la línea correspondiente durante el proceso. Aunque pueden crearse programas de revisión por ordenador que aclaren la situación cuando se hayan introducido elementos personales por equivocación en la línea de otra persona, el programa mismo es muy difícil de elaborar.

10. También pueden introducirse a la vez los datos que figuran en una misma página. El operador puede introducir la información de una página entera antes de pasar a la página siguiente. En este caso introducirá toda la información que figure en la primera página, independientemente del número de personas. A continuación pasará a la página e introducirá la parte siguiente de la información relativa a todas las personas. Pueden introducirse o no pautas de omisión de determinados datos, según el método de introducción que se aplique (con o sin revisión por ordenador). En cualquier caso, durante la revisión por ordenador habrá que reunir los registros de los distintos conjuntos de datos introducidos y resolver en ese momento todo error que se haya producido en la introducción de números personales.

11. En el ejemplo del gráfico A.II.5 (véase la página siguiente) la información demográfica del hogar no plantea ningún problema de introducción de datos poco habitual, ya que el censo obtuvo respuestas para todos los elementos respecto de todas las personas.

12. Sin embargo, una segunda página correspondiente al mismo hogar (gráfico A.II.6, en la página siguiente) podría presentar ciertos problemas de introducción de datos. Por ejemplo, si el país decide reunir datos sobre uso del idioma únicamente para las personas mayores de cinco años de edad, esa información quedará en blanco para la persona 8, que tiene tres años de edad. El encargado de introducir los datos deberá dejar en blanco la casilla correspondiente a esta niña, o, si llegara el caso, se corregiría más tarde en la revisión por ordenador.

13. Otros elementos deberían quedar igualmente en blanco, como las personas que no alcanzan la edad mínima para participar en la fuerza de trabajo, las mujeres que están por debajo de la edad mínima de fecundidad y los elementos de fecundidad respecto de todas las personas de sexo masculino. Un posible error que podría aparecer en el gráfico A.II.6 es que pudiera haberse dado el caso de que el operador

Gráfico A.II.5

Ejemplo de una página sobre un hogar con varias personas sin problemas de introducción de datos

Página sobre el hogar				
Personas	Relación	Sexo	Edad	Etcétera
1	Jefe del hogar	M	40	
2	Cónyuge	F	35	
3	Hija	F	18	
4	Hijo	M	12	
5	Hermano	M	35	
6	Hermana del cónyuge	F	30	
7	Hijo de la hermana	M	5	
8	Hija de la hermana	F	3	
Etcétera				

hubiera introducido incorrectamente la información sobre los hijos nacidos vivos de la persona 6 (en este caso 4) en el lugar correspondiente a la persona 5. Entonces, en la revisión por ordenador se suprimirían los elementos de fecundidad de la persona de sexo masculino y se atribuirían datos de fecundidad a la persona de sexo femenino, si bien es posible que no se asignara el valor correcto.

Gráfico A.II.6

Ejemplo de una página sobre un hogar con varias personas con posibles problemas de introducción de datos

Página 2 sobre el hogar				
Personas	Idioma	Fuerza de trabajo	Hijos nacidos vivos	Etcétera
1	Idioma 1	Sí		
2	Idioma 1	No	3	
3	Idioma 1	No	0	
4	Idioma 1			
5	Idioma 1	Sí		
6	Idioma 1	No	4	
7	Idioma 1			
8				
Etcétera				

14. Muchas veces un país debe utilizar el impreso sobre el hogar a causa de limitaciones de costo o de espacio. Sin embargo, cuando la población sea pequeña o el país pueda permitirse el gasto adicional, en un impreso con páginas personales probablemente habrá menos errores de correspondencia atribuibles a equivocaciones de introducción de datos que en el caso de los impresos sobre el hogar.

Anexo III

Escaneado frente a introducción manual de los datos

1. Muchos países aplican técnicas de escáner utilizando equipo de lectura óptica de marcas o equipo de lectura óptica de caracteres. Cada método presenta ciertas ventajas con respecto al teclado manual cuando la operación es rápida y eficiente y los costos no son grandes. Sin embargo, muchos países, incluso algunos que aplican técnicas de escáner, tal vez no puedan permitirse los gastos iniciales ni los gastos corrientes de mantenimiento durante el empadronamiento y en las etapas posteriores. Un aspecto positivo es que muchos países utilizan los aparatos de escáner obtenidos para el censo para diversos propósitos, incluso para otras encuestas y registros administrativos, como los impresos de aduanas. Los países pueden también subcontratar las operaciones de escaneado, o arrendar los escáneres mientras dure el censo.

2. Una de las ventajas de la introducción manual es que las técnicas aprendidas durante la introducción pueden aplicarse a otras actividades de las oficinas nacionales de estadística o del censo y a otros organismos públicos. Una vez que gracias al censo los operadores han adquirido experiencia en la introducción de datos, esas personas pueden dedicarse a introducir datos para otros estudios complementarios. Puede tratarse de las encuestas posteriores al empadronamiento o de otros estudios, como las encuestas sobre fecundidad o sobre ingresos y gastos de los hogares. El mismo personal puede encargarse también de introducir datos de registros administrativos, como los del registro civil y los correspondientes al comercio, la inmigración y emigración y las aduanas.

A. Introducción de los datos

1. El escáner

3. Los países que utilizan un escáner u otros aparatos de lectura óptica para introducir los datos no suelen efectuar correcciones en los datos en el momento de la introducción, aunque puede haber cambios con arreglo a las pautas de omisión integradas en el sistema. Con todo, los países que decidan introducir manualmente sus datos pueden elegir entre varias opciones según la rapidez con que necesiten introducirlos y el grado de comprobación manual que se requiera. Cada una de las opciones dependerá de las necesidades de los equipos de revisión, los conocimientos técnicos de los encargados de introducir los datos y la complejidad del programa de revisión.

4. La cantidad y el tipo del equipo de introducción de datos necesario dependerán del método de captación de datos que se haya elegido, del tiempo de que se disponga para esta etapa del censo, del tamaño del país, del grado de descentralización de las operaciones de captación de datos y de varios otros factores. Para la introducción de datos a través del teclado, la media varía entre 5.000 y 10.000 pulsaciones por hora.

Algunos operadores de introducción de datos no alcanzan esa cota, mientras que otros la superan con creces. Otros factores que afectan a este aspecto son: *a)* los programas informáticos de apoyo; *b)* la complejidad de las tareas de los operadores; *c)* las características ergonómicas, la fiabilidad y la velocidad del equipo; *d)* el que haya o no trabajo por hacer ininterrumpidamente; *e)* la formación y la aptitud del personal contratado; y *f)* la motivación de los trabajadores (Naciones Unidas, 2007, párr. 1.309).

2. Introducción directa de datos a través del teclado

5. La introducción directa de datos a través del teclado puede adoptar dos formas. La primera consiste en introducir manualmente todos los datos que se van encontrando, sin seguir ninguna pauta de omisión. En este caso la introducción es más rápida porque los operadores no tienen que detenerse cuando se encuentran con información claramente inválida o incongruente. También puede ser más precisa, ya que la tarea realiza de forma más mecánica. La segunda forma de introducción directa entraña detenerse para comprobar resultados inválidos o incongruentes en los cuestionarios, en cuyo caso el proceso avanza con más lentitud y requiere muchos más conocimientos por parte de los operadores. Deberá tenerse muy presente el elevado costo que ello supone desde el punto de vista de la velocidad. Paradójicamente, la exactitud también puede mejorarse con este método si los operadores descubren que los datos estaban registrados correctamente pero que se les asignó un código erróneo. A veces los errores de introducción manual también pueden localizarse inmediatamente porque el programa de revisión incluye la comprobación automática.

a) Introducción directa de datos a través del teclado sin pautas de omisión

6. Cuando todos los asientos se introducen o se omiten manualmente puede mantenerse un ritmo particular y ciertas pautas de omisión no pasarán por alto información válida aunque sea temporalmente incongruente. Por ejemplo, si en un registro el sexo de una persona consta como masculino, casi todos los equipos de revisión requerirán que se omita toda la sección sobre fecundidad. En este caso el encargado de introducir los datos utilizará la barra espaciadora o la tecla de flecha de su teclado para saltarse parte del registro de una persona de sexo masculino o de una persona joven de sexo femenino, porque todos los asientos estarán en blanco. Sin embargo, esto lleva tiempo y es posible que las omisiones no sean totalmente necesarias. El encargado de la introducción de datos podría avanzar demasiado o no avanzar suficiente, y podrían introducirse equivocadamente otros elementos por no estar bien alineados. Si todos los campos se introducen de esta forma, la información podrá introducirse cuando no se sigan pautas de omisión. Por ejemplo, cuando el encargado de introducir los datos se encuentre con el caso de una mujer adulta sobre la que consta información sobre fecundidad (una mujer respecto de la cual se han recopilado y codificado datos sobre elementos como los hijos nacidos vivos, los hijos vivos o los hijos nacidos en el último año), introducirá todos los asientos. Si la información sobre fecundidad está introducida, el programa de revisión por ordenador puede determinar qué elemento o conjunto de elementos es válido y cuál deberá ser modificado. Cuando el programa de revisión determina que la persona es una mujer adulta pero la información sobre fecundidad ha quedado en blanco, habrá que aplicar técnicas de atribución dinámica u otro medio apropiado para obtener información sobre fecundidad para las tabulaciones. Si la información real se ha perdido a causa de las pautas de omisión, el equipo de revisión deberá decidir si la pérdida queda compensada por el aumento de eficiencia y de velocidad. Si se aplican pautas de omisión, los encargados de introducir los datos todavía tienen la posibilidad de regresar a la posición apropiada para efectuar

correcciones. Aunque perderán algún tiempo pasando por elementos en los que no se introducirán valores, con esta forma de introducción de datos las incongruencias entre los valores de sexo, edad y fecundidad pueden resolverse durante la revisión y no en el momento de la introducción.

b) Introducción directa de datos a través del teclado con pautas de omisión

7. Un segundo método de introducción directa supone utilizar pautas de omisión. Si el equipo de revisión requiere pautas de omisión, normalmente para representar la forma en que los empadronadores recopilan los datos, la introducción será más fácil y rápida si las pautas son fáciles de seguir y si los operadores aprenden rápidamente las secuencias de introducción. Si las pautas de omisión son muy complicadas, los operadores pueden confundirse y sistemáticamente introducir datos en los lugares equivocados. La forma más eficiente de introducción de datos con pautas de omisión es la que consiste en aplicar un número limitado de pautas que abarcan grandes porciones del registro que se está introduciendo.

8. El equipo de revisión deberá determinar las pautas de omisión apropiadas para el censo o la encuesta de su país. Algo que parece lógico omitir son los elementos relativos al empleo, en especial en el caso de los niños; es decir, las personas por debajo de la edad definida por el país para poder trabajar. Esos elementos omisibles a veces representan la mitad de los elementos de población, por lo que resulta eficiente omitirlos en el caso de los niños, excepto en casos especiales, y en los de personas de edad dudosa o cuando el país puede tener interés en conocer los datos sobre el trabajo infantil.

9. El equipo de revisión decidirá caso por caso los elementos que se incluirán dependiendo de los grupos de edad. Los elementos pueden agruparse de manera que puedan seguirse fácilmente las pautas de omisión.

10. No siempre es fácil adoptar decisiones claras acerca de las pautas de omisión. Un ejemplo sería la secuencia siguiente:

1. ¿Cuál es la ciudadanía de esta persona?
 - Nacido en este país (pásese al elemento 3)
 - Naturalizado
 - No es ciudadano
2. ¿En qué año entró esta persona en el país?
3. ELEMENTO SIGUIENTE

11. Aunque podría crearse una pauta de omisión para pasar de 1 a 3, es decir, saltarse el elemento sobre el año de entrada para las personas nacidas en el país, a veces los encargados de introducir los datos no siguen la pauta de omisión ya se deba a que el empadronador o el codificador han cometido un error o por un error de introducción manual. Entre los numerosos factores presentes figuran el nivel técnico de los encargados de introducir los datos, las circunstancias culturales, la presentación del cuestionario y el formato de las pantallas. A menudo el equipo de revisión coopera para determinar si en un caso como este la pauta de omisión es razonable.

3. Introducción interactiva de datos

12. Los métodos interactivos pueden utilizarse para la introducción de datos de un censo, pero son más apropiados para las encuestas, particularmente para las encuestas pequeñas en las que los elementos asignados podrían afectar los resulta-

dos generales. En la introducción interactiva de datos pueden efectuarse correcciones manuales o automáticas, según la información de que se disponga para efectuar los cambios o correcciones.

13. Un ejemplo sería el caso de una pequeña encuesta. En las pequeñas encuestas cada respuesta es importante. Si un país realiza una encuesta con una muestra del 1%, por ejemplo, cada respuesta representa a 100 personas, unidades de habitación o explotaciones agrícolas. Unos pocos casos inválidos o incongruentes podrían tener un efecto muy considerable en los resultados de la encuesta. En esos casos los demógrafos y otros especialistas en ciencias sociales normalmente quieren mantener un control considerable sobre el procesamiento de los datos.

14. El control puede llevarse a efecto de varias maneras. Los demógrafos y otros especialistas pueden encargarse personalmente de introducir los datos, comprobando las respuestas no pertinentes, las inválidas o las incongruentes a medida que van avanzando, basados en la información tal como aparece registrada en los impresos de recopilación de datos. A menudo los conflictos, los errores de codificación u otras incongruencias pueden resolverse inmediatamente estudiando directamente la información recogida. En otras ocasiones los responsables pueden optar por volver a enviar sobre el terreno los cuestionarios incompletos o inválidos. Este tipo de introducción interactiva de datos arroja los mejores resultados, ya que el demógrafo actúa también como encargado de la introducción, pero es con mucho el más costoso y muchos países no pueden permitirse este método.

15. Los equipos de revisión pueden formular normas pormenorizadas de revisión para determinar qué deben hacer los encargados de introducir los datos cuando se les presentan casos especiales al respecto. Pueden decidir el valor que deberá introducir el operador respecto de cada código inválido por resolver. El equipo de revisión puede resolver los casos que no hayan quedado abarcados por las normas detalladas e incluso modificar las normas (aunque con ello corre el riesgo de que haya incongruencias entre la primera parte y partes ulteriores del proceso de introducción de datos).

16. Las pautas de omisión, que tienen un papel destacado en la introducción directa, también son importantes en este caso. Como sucede con la introducción directa, los operadores deben conocer y aprender las pautas de omisión que se aplicarán. Como se ha mencionado, las pautas de omisión pueden incrementar la velocidad de la introducción manual de datos, aunque normalmente habrá cierta pérdida de calidad. En el caso de la introducción interactiva, una norma práctica habitual es que el nivel de calidad estará en relación inversa con el número de omisiones.

17. Después de formular las instrucciones para la introducción de datos, las oficinas nacionales de estadística o del censo deberán hacer que los operadores pongan a prueba esas instrucciones antes de adoptar una decisión sobre la operación real, independientemente de que se utilice o no el método de la introducción directa. Al poner a prueba las instrucciones podrán resolverse las deficiencias del sistema y lograrse un nivel óptimo de introducción de datos.

B. Verificación

18. La oficina nacional de estadística o del censo también deberá decidir cuál es el nivel de verificación apropiado. Muchos expertos aconsejan llevar a cabo una verificación absoluta en los datos introducidos manualmente. En ese caso todos los elementos vuelven a introducirse para asegurarse de que los datos recogidos son los que entrarán en el ordenador para ser procesados. Empero, en muchos casos la verificación

total no es una solución práctica, bien porque no haya tiempo para volver a introducir todos los datos o porque el país carezca de los recursos financieros o humanos necesarios. La muestra porcentual verificada debería ser mayor cuando los encargados de introducir los datos sean principiantes, y menor si tienen experiencia. Además, si las tasas de error de introducción comprobadas son muy bajas y los operadores cometen muy pocos errores, es probable que la verificación completa no sea necesaria.

19. En cualquier actividad de verificación lo primero es determinar la información necesaria. ¿Desea el país supervisar a operadores concretos o a un equipo de encargados de introducir los datos? ¿Desea determinar si se está adquiriendo, o se mantiene, la competencia necesaria? Las unidades de control podrían ser también importantes, en particular en lo que se refiere a la presentación de informes diarios, semanales y mensuales, etcétera, para decidir el flujo de trabajo y los conocimientos adquiridos.

20. Finalmente es importante que la verificación sea independiente y que desde la entrada de los datos la lleve a cabo un conjunto diferente de encargados de introducir los datos, o al menos otros miembros del mismo equipo. La utilización de conjuntos diferentes de encargados de introducir los datos permite mayor independencia en las operaciones y, por consiguiente, mejores resultados.

21. En cuanto a los datos escaneados, algunos países realizan una verificación para comprobar que sea completo y exhaustivo. Debido a que la tecnología del escaneado es aún bastante joven, e independientemente de que los sistemas se comprueben a fondo con datos piloto o con pruebas previas, los cambios de calidad del papel, la impresión de los formularios en distintas imprentas, las condiciones de almacenamiento, etcétera, pueden afectar al proceso de escaneado y provocar problemas que tendrán que ser resueltos mediante la verificación.

22. Si los errores son sistemáticos y pueden eliminarse mediante el programa de revisión, los encargados de la introducción manual de los datos y los de la verificación no tendrían nada que decir sobre la corrección. No obstante, ambos son responsables de encontrar los errores. Estos pueden darse porque no se ha comprobado debidamente el equipo de escaneado, lo que provocaría errores sistemáticos en determinados elementos o en la combinación de varios de ellos, confusión en la lectura de algunos dígitos (por ejemplo, confundir el 2 o el 3 con el 8 o el 9), lecturas erróneas de los recuadros de comprobación de continuación, etcétera.

23. La lectura incorrecta ha sido un problema constante en los últimos años, y sólo puede resolverse en la fase de revisión. Si los formularios no son contiguos, se necesitarán otros procedimientos —muy probablemente durante las revisiones estructurales— para resolver los problemas. Como ya se ha señalado, antes de comenzar la revisión de contenido se necesita un archivo completamente válido y estructurado.

1. Verificación dependiente

24. Las técnicas de verificación pueden ser de dos tipos: dependientes o independientes. En la verificación dependiente los operadores vuelven a introducir los datos previamente introducidos por otras personas. Cuando los valores difieren, el programa informático informa al operador. En ese caso, según el programa de que se trate, el operador anulará los datos anteriores o bien se tomará nota de la discrepancia. Puesto que los datos se introducen a partir de los cuestionarios originales, normalmente el propio operador puede adoptar una decisión fundamentada acerca de si el dato introducido originalmente estaba equivocado.

2. Verificación independiente

25. En la verificación independiente los operadores vuelven a introducir los datos desde el principio y crean un archivo completamente independiente de los datos introducidos, utilizando los cuestionarios originales. A continuación se utiliza un programa informático para comparar los dos archivos resultantes (los datos introducidos originalmente y el conjunto de datos de verificación) a fin de observar las discrepancias. Lo habitual es que se efectúe algún tipo de operación manual para corregir los datos inválidos e incongruentes.

C. Revisión cuando se trata de datos escaneados

26. Cada vez son más los países que escanean sus datos. A comienzos del decenio de 2000 muchos de esos países se sorprendieron al comprobar que el escaneado introducía tipos de errores distintos de los asociados con la introducción manual de los datos. Parte del problema de la revisión de los datos escaneados es la falta de control de calidad durante el proceso inicial. Hace una década, la tecnología era tan nueva que muchas oficinas de estadística no tenían la experiencia ni los medios necesarios para realizar un control de calidad adecuado de todos los elementos. Muchos de los países que disponían de procedimientos de control de calidad válidos terminaron no desarrollándolos para todos los elementos; por ello, algunos de los elementos que se encontraban al final de la pregunta —en particular los relacionados con la fecundidad— resultaban ser inválidos o incongruentes.

27. Naturalmente, muchas de las incongruencias encontradas en los datos introducidos manualmente se producen también en los datos escaneados. No obstante, conviene examinar algunos de los problemas especiales derivados del uso de datos escaneados. Como los cuestionarios escaneables requieren marcadores para facilitar la lectura mecánica, los elementos aparecen muchas veces de tal forma que los empadronadores y declarantes pueden tener problemas durante la recopilación de datos. Los problemas relacionados con estos elementos deben abordarse sistemáticamente. Cuando los elementos están estrechamente relacionados con otros, como la religión o el origen étnico, pueden utilizarse las revisiones habituales descritas en varias partes del *Manual*.

28. No obstante, debe procederse con cautela cuando los elementos necesarios para la planificación y la formulación de políticas pueden generar dificultades. Es común que el elemento correspondiente al **sexo** no produzca problemas, ya que implica sólo dos posibilidades. De todas formas, como ya se ha señalado, mientras que el encargado de la introducción manual de los datos normalmente sólo puede elegir entre un 1 o un 2 (o un código correspondiente a “desconocido”), en las columnas relativas al sexo puede aparecer cualquier valor, debido a la presencia de otros dígitos, caracteres alfabéticos u otros caracteres. Por ello, para tener en cuenta esos valores variopintos es preciso agregar algún tipo de revisión a lo que se hacía antes con los datos introducidos manualmente.

29. Los códigos de **relación** son un ejemplo ilustrativo. Cuando hay un solo dígito para los códigos de relación, normalmente no se plantea ningún problema. En cambio, cuando se utilizan dos dígitos se suele producir un problema de escaneado porque el primer dígito está codificado incorrectamente o ha sido leído incorrectamente por el escáner. Si se utilizan códigos numéricos del 1 al 12, el encargado de la introducción manual solamente podrá elegir uno de tales códigos, y el programa de entrada detectará la presencia de un código impropio. Con el escaneado se acepta

prácticamente todo (aunque los sistemas de escaneado pueden programarse para detectar también algunos errores). En esos casos, los códigos erróneos deben cambiarse durante la revisión; de lo contrario, provocarán todo tipo de problemas durante la fase de tabulación.

30. La **edad** es algunas veces causa de problemas, sobre todo cuando se utilizan tres columnas (para las personas de más de 100 años); por ello, para realizar una revisión adecuada quizá sea necesario un análisis dígito por dígito, lo que significa examinar por separado los dígitos correspondientes a las unidades, decenas y centenas. Una vez que se ha comprobado que la edad está debidamente reflejada, puede comenzar la revisión normal.

31. No obstante, en los casos en que aparecen tanto la **edad como la fecha de nacimiento**, la información engañosa puede provocar problemas cuando un elemento tiene precedencia sobre los otros. Normalmente, los especialistas temáticos prefieren utilizar la fecha de nacimiento con la fecha de referencia del censo o la encuesta para obtener (mediante sustracción) una edad exacta, que se comparará con la edad registrada. Cuando faltan uno o varios dígitos, debe comprobarse que todos los demás dígitos se utilicen debidamente para obtener la estimación más adecuada de la edad calculada para la comparación. Cuando el escáner no capta un único dígito, por ejemplo, la revisión puede tenerlo en cuenta a fin de realizar la estimación más plausible de lo que debería aparecer. Este tipo de problema no suele suceder en el caso de la introducción manual de los datos.

32. Los elementos que en los primeros años del decenio de 2000 creaban mayores problemas a causa del escaneado estaban relacionados con los datos sobre la **fecundidad**: afectaban tanto el número de niños nacidos y sobrevivientes como el de niños nacidos en el último año o a lo largo del tiempo. En la mayoría de los países el principal problema ha sido la falta de control de calidad durante el escaneado, lo que ha dado lugar a la presencia de elementos extraños en la captación de los datos. Cuando el valor del elemento de un determinado país era de 17, 18 ó 19 niñas muertas, por ejemplo, los datos, si no se habían revisado, no podían utilizarse para actividades de planificación.

33. La información sobre la **mortalidad** puede presentar también problemas en los datos escaneados. En los datos introducidos manualmente, cuando hay una serie de elementos correspondientes a los fallecimientos a lo largo del año anterior al censo (por ejemplo, sexo y edad de la persona fallecida, indicación de si falleció de muerte natural o si se trataba de una muerte materna), la introducción manual continúa aun cuando haya tachaduras y datos borrados. En cambio, cuando se utiliza el escáner, como las cifras borrosas o tachadas no puede leerlas, dejará un blanco y procederá a la fase de captación. El programa de revisión deberá trasladar la información a los espacios adecuados, para su tabulación y el análisis posterior. Conviene señalar que últimamente las operaciones de escaneado ya pueden realizarse durante la captación e inmediatamente después de ella.

D. Conclusión

34. Lamentablemente, los problemas de cada país dependen del programa concreto de escaneado y del funcionamiento de los escáneres utilizados, por lo que resulta difícil ofrecer orientaciones generales. No obstante, en todos los casos abordados hasta ahora los problemas de escaneado han sido sistemáticos. Cuando el personal determinó el algoritmo necesario para evitar los problemas, pudieron obtenerse conjuntos de datos plenamente revisados.

Anexo IV

Modelos de diagramas de secuencia

1. Una de las tareas del equipo de revisión es establecer una estructura relacional para las variables usadas en el proceso de revisión. Los diagramas de secuencia facilitan la identificación de interrelaciones de las variables y ayudan a formular claras y concisas especificaciones de revisión. Tales especificaciones de vinculaciones relacionales ayudarán a los especialistas temáticos y a los de procesamiento de datos a enfocar el proceso de revisión y facilitarán la comunicación entre ambos grupos.

2. En las páginas siguientes se presentan tres modelos de diagramas de secuencia:
- a) Diagrama de secuencia para determinar el jefe del hogar;
 - b) Diagrama de secuencia para determinar la existencia de un cónyuge en el hogar;
 - c) Diagrama de secuencia para revisar la variable de sexo del jefe del hogar y del cónyuge.

Estos modelos de diagramas de secuencia se presentan meramente como ejemplo y deben considerarse como tales. El equipo de revisión tal vez desee modificar el modelo de diagrama de secuencia según las necesidades que plantee la situación del país.

3. Deben elaborarse diagramas de secuencia de revisión para cada una de las variables de un censo. El equipo de revisión deberá colaborar para elaborar los diagramas de secuencia, y los especialistas en procesamiento de datos deberán utilizarlos junto con las especificaciones de revisión para crear programas informáticos que permitan la revisión de los datos del censo. Los diagramas de secuencia y las especificaciones de revisión deberán estar bien documentados, de forma que puedan utilizarse para el procesamiento de datos de futuros censos y encuestas.

Gráfico A.IV.1

Modelo de diagrama de secuencia para determinar el jefe del hogar

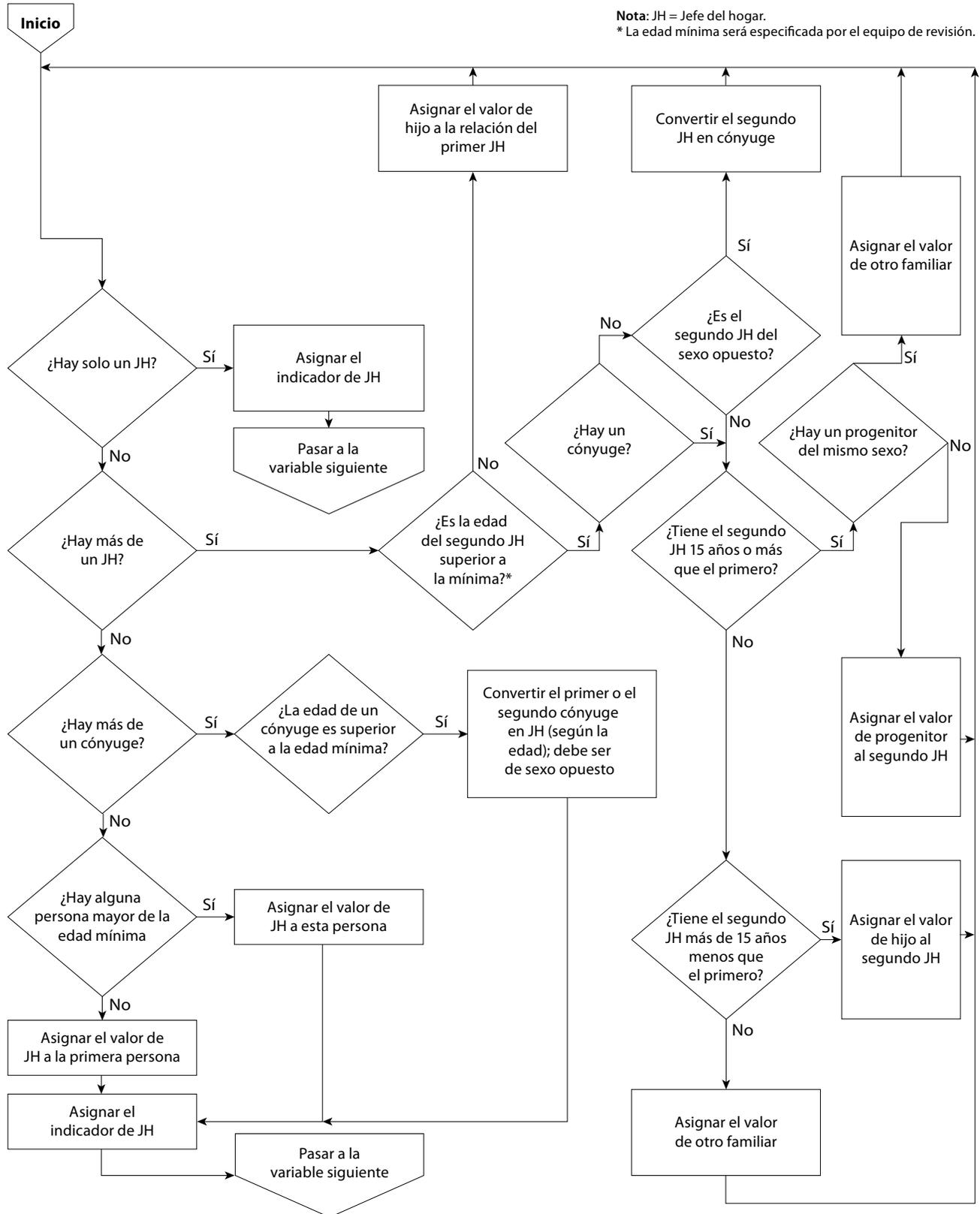


Gráfico A.IV.2
Modelo de diagrama de secuencia para determinar la existencia de un cónyuge en el hogar

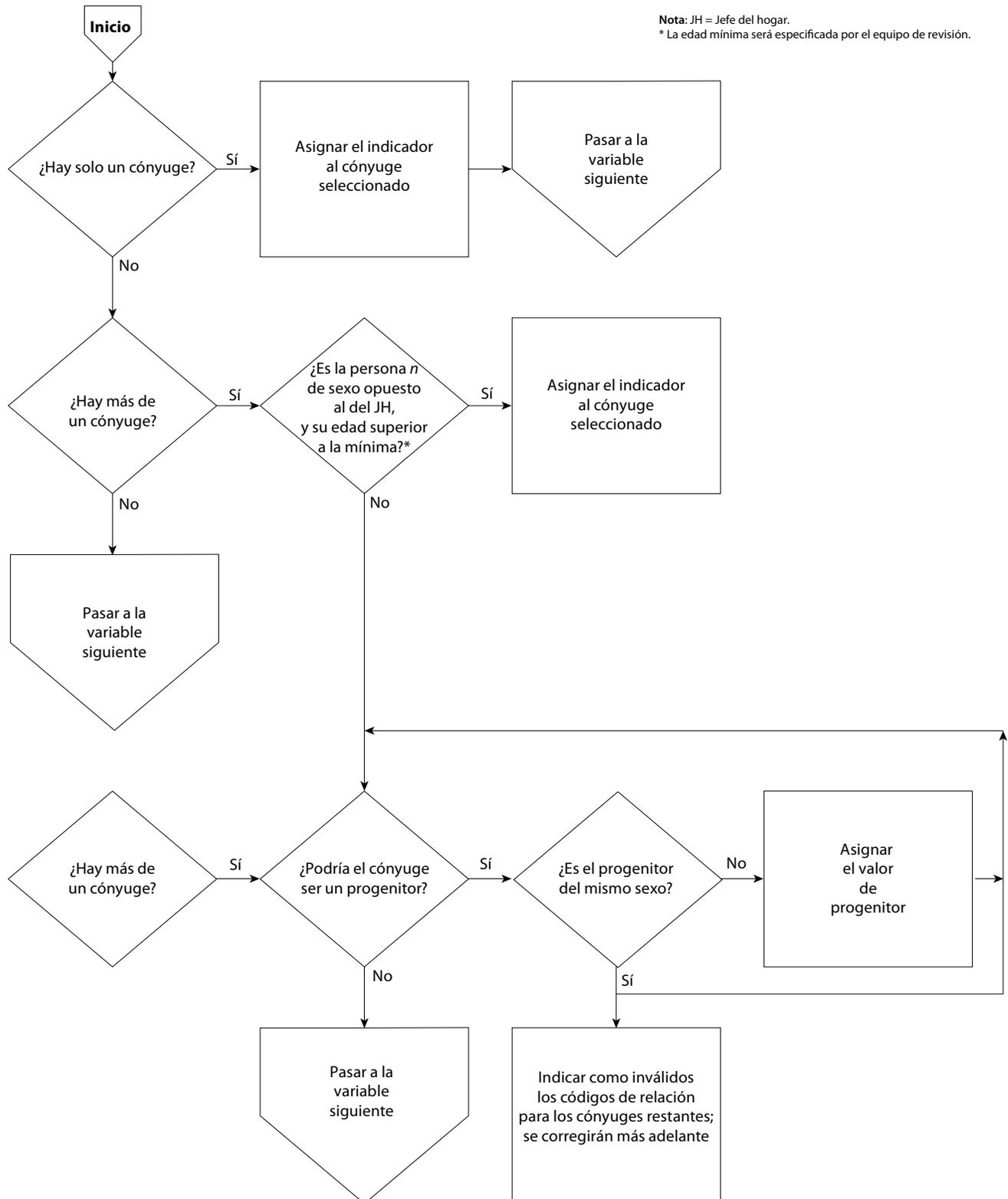


Gráfico A.IV.3

Modelo de diagrama de secuencia para revisar la variable de sexo del jefe del hogar y del cónyuge

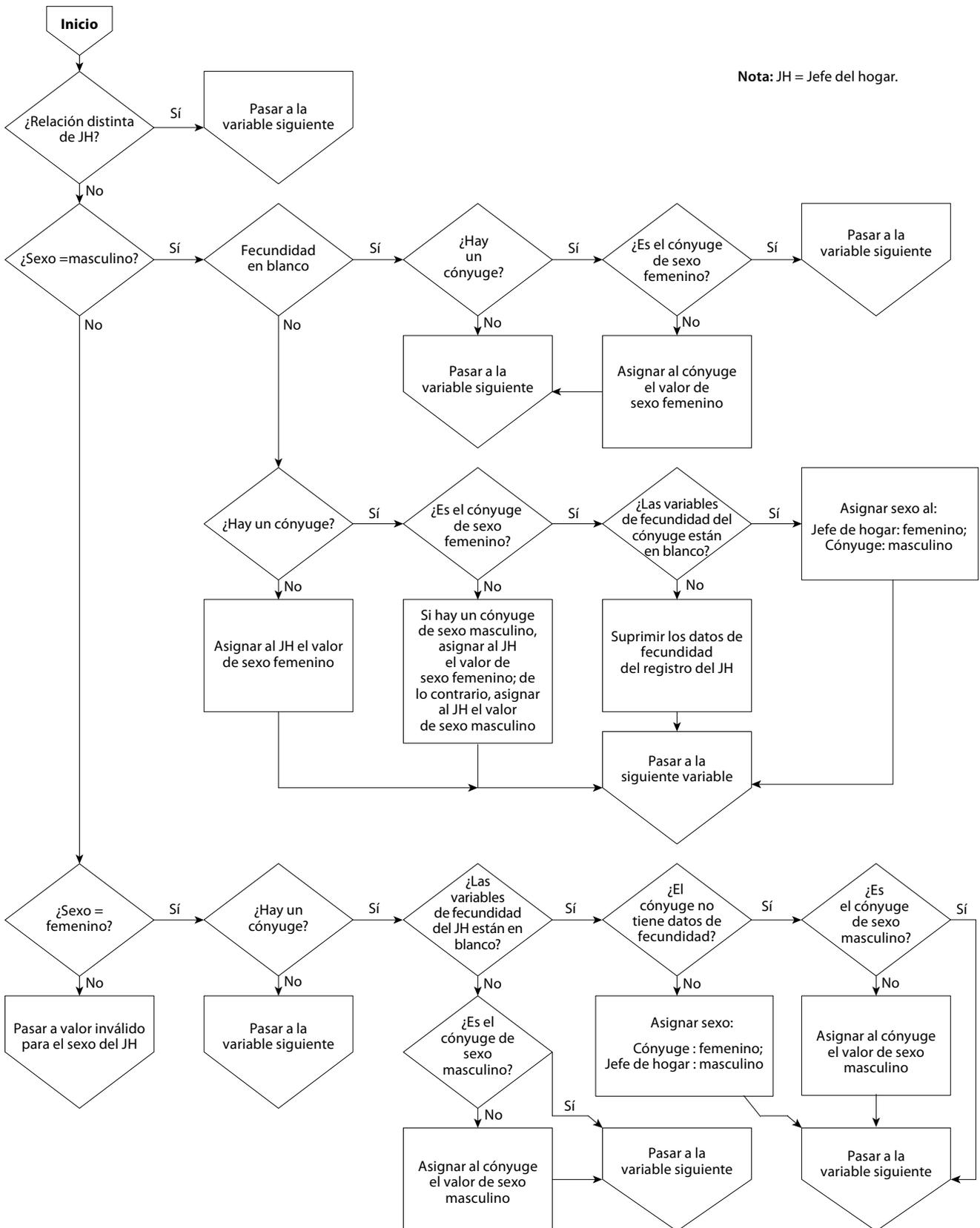
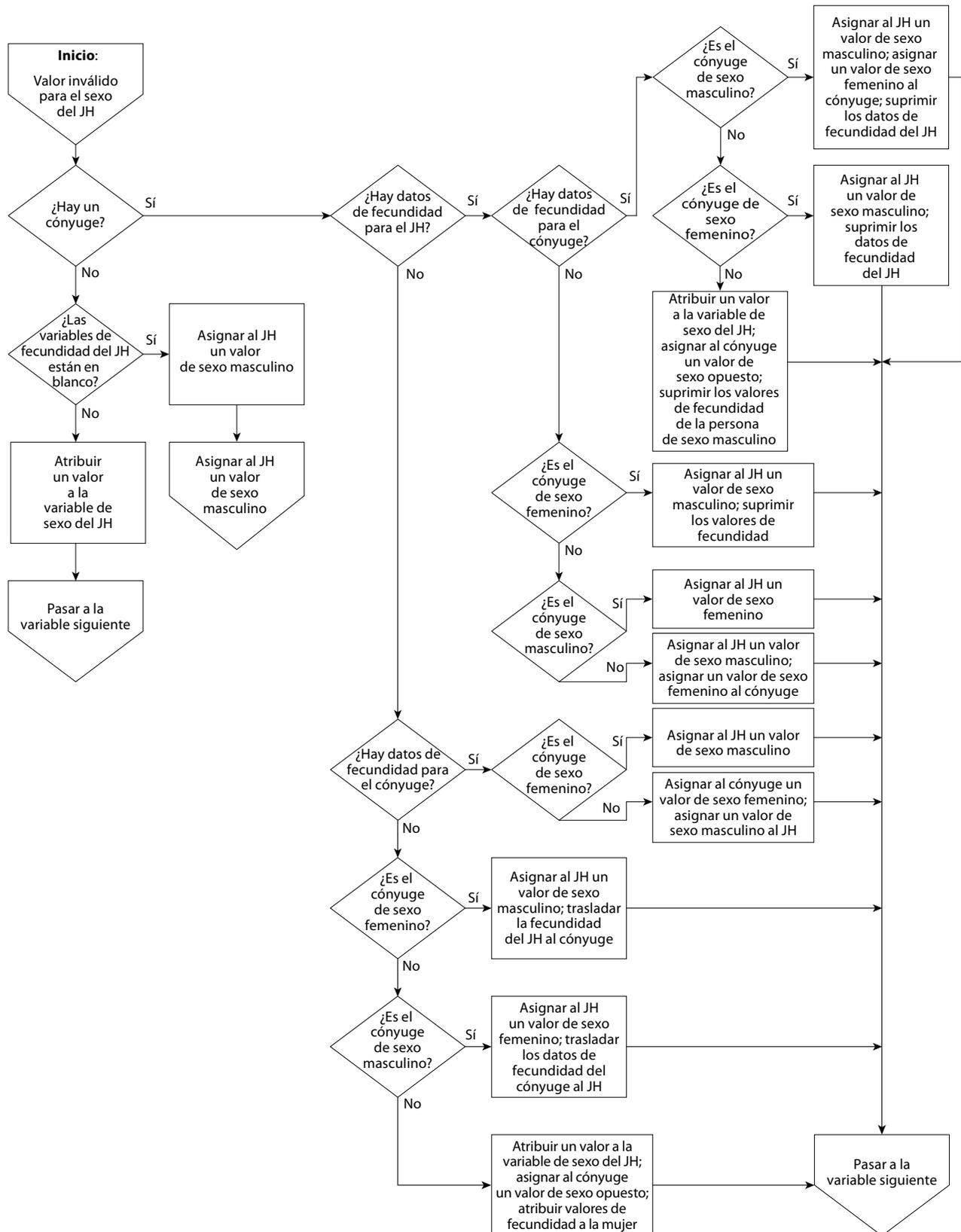


Gráfico A.IV.3 (continuacion)

Modelo de diagrama de secuencia para revisar la variable de sexo del jefe del hogar y del cónyuge



Anexo V

Métodos de atribución de valores

1. Se han elaborado varios métodos de atribución de valores. La mayoría de los que se describen a continuación se examinan en textos de Kalton y Kasprzyk (1982, 1986), Sande (1982) y Särndal, Swensson y Wretman (1992).

2. Los métodos de atribución pueden clasificarse en estocásticos o deterministas, según en qué medida intervenga el azar en la atribución de los datos.

3. Los **métodos de atribución deterministas** incluyen la atribución deductiva; los métodos de atribución basados en modelos, como la atribución basada en valores medios y la atribución mediante técnicas de regresión; y (si procede) la atribución basada en el vecino más próximo.

4. La **atribución deductiva** es un método que permite deducir con certeza un valor omitido o incongruente. A menudo se basará en la pauta de las respuestas que se hayan dado a otros elementos del cuestionario.

5. Lo más frecuente es que la técnica de atribución tenga que reemplazar un valor del que no se sabe con certeza si es el valor real. En los párrafos siguientes se presentan algunos procedimientos corrientes de atribución de valores.

6. Con la excepción de los algoritmos de atribución dinámica con un único donante, los métodos que se describen a continuación atribuyen valores elemento por elemento. Por tanto, dentro de cada categoría de atribución los elementos del registro se examinarán sucesivamente, uno por uno. Normalmente esto se hace examinando sólo los aspectos que correspondan explícitamente al elemento en cuestión o a un pequeño conjunto de variables estrechamente relacionadas con él. Dado que puede haber una relación explícita o implícita entre el elemento o los elementos en cuestión y otros elementos que se estudiarán en un momento ulterior del proceso, este procedimiento puede tener como consecuencia que un valor atribuido, a pesar de superar la revisión que se efectúa en ese momento, genere errores en otras variables que se considerarán en un momento ulterior del proceso. Sólo cuando se revise un conjunto completo de variables, incluidas todas las comprobaciones implícitas, podrá asegurarse que los valores atribuidos superan todas las normas de revisión. Una norma de revisión implícita es la que puede derivarse combinando lógicamente dos o más condiciones explícitas.

7. En las descripciones siguientes se entenderá por “registros aceptados” los que hayan superado todas las revisiones correspondientes al elemento en cuestión. Por “registros no aceptados” se entenderá aquellos que, como mínimo, no hayan superado una de las revisiones correspondientes al elemento o elementos en cuestión.

8. El **método de atribución de valores basado en la media general** asigna el valor medio de los registros aceptados a un elemento omitido o incongruente en todos los registros no aceptados. Este método puede producir estimaciones puntuales razonables, pero tiene menos utilidad si hay que aplicar estimadores estándar de la varianza. La varianza puede subestimarse considerablemente a menos que la propor-

ción de atribuciones sea muy baja o que el estimador de la varianza se modifique para tener en cuenta el hecho de que se han utilizado valores atribuidos.

9. El **método de atribución de valores basado en la media registrada dentro de una categoría** utiliza categorías de atribución definidas para crear grupos de registros que presenten cierto grado de similitud. Dentro de cada categoría, el valor medio de los registros aceptados se atribuye al elemento omitido o incongruente en todos los registros no aceptados. Es un método muy parecido al que se basa en la media general, pero es probable que los efectos en la distribución y los problemas para la estimación de la varianza sean menos graves.

10. La **atribución basada en técnicas de regresión** o, de manera más general, **la atribución basada en un modelo** utiliza datos de registros aceptados para tomar el valor que se necesita atribuir de un conjunto de variables independientes. Estas pueden ser elementos del propio cuestionario o variables auxiliares. A continuación se utiliza la ecuación de regresión para atribuir los valores de los elementos omitidos o incongruentes. Este es un caso especial de la atribución basada en un modelo. Este método se utiliza frecuentemente para las variables continuas en aplicaciones de encuestas comerciales en que los datos de otras ocasiones a menudo permiten una buena predicción de los valores de la ocasión presente.

11. En el **método de atribución basado en el vecino más próximo** se asigna un valor a un registro no aceptado utilizando el registro aceptado “más próximo”; este se define aplicando una función de distancia basada en otras variables conocidas. Este método puede aplicarse dentro de determinadas categorías de atribución. Normalmente se considera apropiado para variables continuas, pero también puede aplicarse a variables no numéricas.

12. Los **métodos de atribución estocásticos** incluyen la regresión, o cualquier otro método determinista, con la adición de elementos aleatorios, y los métodos de atribución estática y de atribución dinámica.

13. A cada método determinista le corresponde un método estocástico. Esto se consigue agregando un elemento aleatorio tomado de una distribución pertinente al valor asignado en la atribución determinista. Este procedimiento ayudará a preservar mejor la estructura de frecuencias del archivo de datos. Kalton y Kasprzyk (1986) han descrito ciertas formas de aplicar esta técnica.

14. Con las técnicas de **atribución estática** y **atribución dinámica** se intenta crear una variabilidad de los valores atribuidos más realista que la obtenida con los métodos deterministas. Los procedimientos de atribución dinámica sustituyen los valores omitidos o incongruentes con valores seleccionados (al azar) de registros de variables aceptados de la misma encuesta o censo. Los procedimientos de atribución estática asignan valores tomados de otras fuentes, que a menudo son datos históricos, tal que datos de ocasiones anteriores del mismo censo o encuesta. Hay distintas formas de atribución dinámica y estática.

15. La **atribución aleatoria general** es la forma más simple de atribución dinámica. Para cada registro no aceptado se selecciona al azar un registro aceptado del conjunto de todos los registros aprobados, y el valor declarado para el elemento en cuestión se atribuye al registro inválido.

16. En la **atribución aleatoria dentro de una categoría** también se utilizan categorías de atribución para limitar la selección aleatoria del registro donante a un conjunto que se considera que guarda cierta similitud con el registro al que hay que asignar un valor.

17. La **atribución dinámica secuencial** también utiliza categorías de atribución y tiene la ventaja de que basta con una sola pasada por el archivo de datos para llevar a cabo el proceso de atribución. El procedimiento comienza con un valor de atribución estática para cada categoría de atribución, y los registros del archivo de datos se examinan sucesivamente. Cuando se detecta un registro aceptado, su valor para el elemento en cuestión sustituye el valor almacenado para la categoría de atribución. Cuando aparece un registro no aceptado, el valor omitido o incongruente es sustituido por el valor almacenado. El número de categorías de atribución no puede ser excesivamente grande, ya que hay que asegurar que se disponga de donantes en cada categoría de atribución. Si el orden de los registros del archivo de datos es aleatorio, este método será casi equivalente a la atribución aleatoria dentro de una categoría. Un inconveniente de este procedimiento es que a menudo supone utilizar repetidamente los mismos donantes, lo cual puede tener efectos adversos para la distribución del elemento y para las estimaciones de varianza.

18. La **atribución dinámica de tipo jerárquico** es una variante perfeccionada de la atribución dinámica secuencial en la que se utiliza un gran número de categorías de atribución. Cuando no es posible encontrar un donante en la categoría de atribución inicial se van recorriendo de manera jerárquica las categorías hasta que se encuentra un donante.

19. El objetivo de los algoritmos de **atribución dinámica con un donante único** es atribuir a un registro inválido los datos de un mismo donante. Ello permite la atribución conjunta de todos los valores de variables de un registro que la revisión haya determinado como problemáticos. En la práctica se trata a menudo de utilizar a un único donante para cada sección de variables estrechamente relacionadas del registro. Este enfoque ofrece la ventaja considerable de mantener mejor no sólo las distribuciones marginales, como los métodos de atribución dinámica mencionados, sino también las distribuciones de frecuencia conjuntas. Otra ventaja de los métodos de atribución dinámica con un único donante es que reducen el problema de atribuir valores que no se aceptarían en las revisiones que se efectúen de otras secciones de variables. En el contexto de los métodos de atribución dinámica con un único donante, un registro aceptado es el que ha superado todas las revisiones pertinentes para la sección. Un registro no aceptado es el que no ha superado por lo menos una de esas revisiones.

20. El **método de revisión y atribución de Fellegi-Holt** (Fellegi y Holt, 1976) examina un valor teniendo en cuenta simultáneamente todas las normas de revisión. Una característica fundamental del método de revisión y atribución de Fellegi-Holt es que las normas de atribución proceden de las revisiones correspondientes sin especificación explícita. Para cada registro no aceptado se pasa primero por una etapa de localización de errores en la que se determina el conjunto mínimo de variables que habrá que atribuir, así como los límites aceptables de los valores que pueden atribuirse, y a continuación se procede a la atribución. En la mayoría de los casos se selecciona a un único donante de los registros aceptados buscando coincidencias sobre la base de otras variables que intervienen en el proceso de revisión pero que no requieren atribución. El método busca una única correspondencia exacta y puede ampliarse para tener en cuenta otras variables que no participen explícitamente en las revisiones. Ocasionalmente no podrá encontrarse a un único donante y habrá que emplear un método de atribución alternativo.

21. La **metodología de atribución basada en el vecino más próximo (MAVP)** (Bankier y otros, 1996; Bankier, Lachance y Poirier, 1999) es similar al método de Fellegi-Holt en tanto que tiene en cuenta simultáneamente todas las normas de revisión, no especifica explícitamente las operaciones de atribución y atribuye valores de

un único donante. Para cada registro no aceptado identifica operaciones de atribución que entrañen cambios mínimos, dependiendo de los donantes potenciales disponibles. Ello garantiza que se encontrará un donante. A diferencia del método de Fellegi-Holt, la MAVP busca primero a los donantes y determina operaciones de atribución que entrañen modificaciones mínimas. La MAVP busca donantes estableciendo correspondencias en las que se utilizan todas las variables que intervienen en la revisión (incluidas las que posiblemente deberán atribuirse) y puede aceptar correspondencias casi perfectas de variables numéricas, además de correspondencias para la mayoría de las otras variables, aunque no necesariamente para todas. Se determinan operaciones de atribución basadas en cada donante potencial y se identifican las que supondrán un mínimo de modificaciones. El método también examina operaciones de atribución que supongan un número de modificaciones casi mínimo; a veces estas operaciones pueden producir registros con valores atribuidos más plausibles. Por último, se selecciona al azar una de las operaciones de atribución que supongan cambios mínimos o irrelevantes y se realiza la atribución.

22. Aunque tanto el método de Fellegi-Holt como la MAVP plantean muchas exigencias desde el punto de vista informático, se dispone de algoritmos eficientes, de forma que su formulación y aplicación son viables con ordenadores modernos. Esto es particularmente cierto en el caso de la MAVP, que puede manejar con facilidad problemas de revisión y atribución algo mayores que los que permite abordar el método de Fellegi-Holt.

23. Todos los métodos de atribución descritos producen un único valor atribuido para cada valor omitido o incongruente. Todos ellos supondrán un cierto grado de distorsión de la distribución habitual de los valores del elemento en cuestión y pueden provocar estimaciones inapropiadas de la varianza cuando se utilicen estimadores estándar de la varianza. El grado de distorsión variará considerablemente en función del volumen de valores atribuidos y del método utilizado.

24. La **atribución múltiple** es un método propuesto por Rubin (1987) que aborda este problema atribuyendo valores m veces respecto de cada variable que requiere atribución. A continuación, para el conjunto completo de datos pueden producirse m estimaciones para el elemento. A partir de estas estimaciones se produce una única estimación combinada junto con una global de la varianza que expresará la incertidumbre acerca del valor que habrá que atribuir. Un inconveniente del método de atribución múltiple es que aumenta la labor de procesamiento de datos y de cálculo de las estimaciones.

25. En la mayoría de los sistemas de atribución se aplica una combinación de métodos; normalmente se utiliza la atribución deductiva siempre que sea posible, seguida por uno o más procedimientos distintos. La mayoría de las oficinas nacionales de estadística aplican en sus censos alguna forma de atribución dinámica para la revisión y la atribución de valores. Actualmente los más utilizados son los métodos de atribución dinámica secuencial y de Fellegi-Holt. De las oficinas nacionales de estadística que actualmente utilizan el método de Fellegi-Holt, una está pasando a la MAVP y otras están estudiando esa posibilidad. Sin embargo, teniendo en cuenta el público al que se dirige principalmente, el presente *Manual* se concentra en una forma de atribución dinámica secuencial.

Anexo VI

Programas informáticos de revisión

1. Con la aparición de microordenadores relativamente poco costosos, los países deberían estar en condiciones de revisar exhaustivamente y con rapidez los datos de los censos y encuestas*. Hasta hace muy poco, cada país tenía que escribir su propio programa de revisión para atender a sus necesidades concretas, lo cual requería mucho tiempo de elaboración y corrección de los errores y era muy costoso. Sin embargo, con la aparición de programas informáticos de revisión estándar pueden verse muy facilitadas las necesidades de revisión de un país, ya que se requieren menos conocimientos técnicos para el procesamiento de los datos.

2. Una de las ventajas de emplear un programa informático de revisión es que los datos serán congruentes y de buena calidad si se utiliza de modo apropiado, de forma que podrá tabularse con más rapidez. Muchos programas informáticos, como el SAS y el SPSS u otros lenguajes de alto nivel, pueden utilizarse para escribir programas de revisión. Otra posibilidad es que un país opte por utilizar uno de los programas informáticos creados especialmente para la revisión de datos de censos y encuestas. Para la mayor parte de los países la revisión general será más rápida con un programa estándar que con programas escritos especialmente, ya que el programa estándar no requerirá el mismo nivel de conocimientos de procesamiento de datos que los programas escritos especialmente.

3. Un buen programa informático de revisión permitirá la comunicación entre los especialistas temáticos y los programadores y deberá permitir también la colocación de descripciones o de pseudocódigo junto al código de programación, a menos que el propio código de programación sea evidente o transparente para los especialistas temáticos. Los demógrafos y otros especialistas deberían poder seguir el programa línea por línea y comprender exactamente lo que este está haciendo.

4. Cualquier programa informático de revisión existente que un país considere utilizar deberá poder crear y producir informes de las distintas comprobaciones, pruebas y atribuciones que requiere la revisión de los datos del censo. Estos requisitos también son pertinentes cuando los especialistas en procesamiento de datos producen programas de revisión específicos. El programa deberá cumplir los requisitos siguientes:

- a) Tener la capacidad de introducir datos y/o comprobar la introducción de datos. El programa debería permitir la adición de pautas de omisión. Por ejemplo, el equipo de revisión puede decidir que hay que saltarse la información sobre fecundidad en el caso de las personas de sexo masculino;
- b) Realizar revisiones estructurales, que permitirán determinar que estén realmente presentes los tipos de registros que deberían estar presentes; ello incluye, por ejemplo, que haya un registro de vivienda para cada número de serie;
- c) Generar registros omitidos y/o agregar coeficientes de ponderación a los registros existentes;
- d) Determinar que cada variable tenga un asiento válido;

* Es importante señalar que este *Manual* se centra en la revisión de los censos. Dado que las encuestas tienen un número menor de declarantes y, normalmente, más preguntas, muchas veces se preparan revisiones más detalladas. Algunos de los programas examinados en el presente anexo están destinados a encuestas aplicadas en países con población muy numerosa. No obstante, los programas serían también válidos para países más pequeños cuando levanten sus censos. Cuando los archivos son mayores, es más difícil utilizar algunos métodos estadísticos, como el análisis de regresión y de variantes múltiples. Por otro lado, con un archivo exhaustivo, el efecto de la ausencia de respuesta es menor que en una encuesta donde cada individuo incluido en la muestra "cuenta" por muchos. Por ello, cada oficina de estadística debe comprobar diversos programas para determinar el que mejor responde a sus necesidades.

- e) Almacenar la totalidad o parte de los registros ya revisados;
- f) Poner a prueba la congruencia entre dos o más características del mismo registro y entre registros. Un subconjunto de este factor es poner a prueba la congruencia dentro de un mismo hogar, comprobando las respuestas con las de anteriores miembros del hogar.
- g) Atribuir valores mediante la técnica de atribución dinámica, si el país opta por utilizarla;
- h) Utilizar varios valores de un registro o de múltiples registros para construir una variable derivada e insertar la variable derivada en el registro apropiado;
- i) Identificar y eliminar los registros duplicados;
- j) Producir un diario de los errores y las modificaciones, por zonas geográficas pequeñas.

5. Normalmente, los programas revisan los registros uno por uno, pero los más recientes permiten también la comprobación entre registros, particularmente dentro de una unidad de habitación.

6. Como se observa en el texto, hasta la introducción del método de Fellegi-Holt (1976) y sus sucesivas actualizaciones, casi todas las actividades de revisión utilizaban un enfoque secuencial; es decir, los diversos elementos se realizaban por orden, casi siempre en el mismo orden en que se habían recopilado. Por ejemplo, ya que el primer elemento asociado con la población que se revisa suele ser la “relación”, el elemento correspondiente al “sexo” se revisaría partiendo de aquel, luego se procedería con la “edad”, teniendo en cuenta tanto el sexo como la relación, y así sucesivamente.

7. En los últimos decenios se han elaborado varios sistemas de imputación de cambio mínimo basados en Fellegi-Holt. Entre ellos cabe citar CANEDIT y el General Edit and Imputation System (GEIS) de Statistics Canada y el CANadian Census Edit and Imputation System (CANCEIS) (Bankier, 2005; Chen, 2007)*. Los Estados Unidos de América elaboraron DISCRETE (Chen y otros, 2000; Winkler, 1997a; Winkler, 1997b; Winkler y Chen, 2002) y Structured Programs for Economic Editing and Referrals (SPEER) (Greenberg y Surdi, 1984; Winkler y Draper, 1997) en la Oficina del Censo de los Estados Unidos. Kovar y Winkler (2000) compararon con cierto detalle los sistemas del Canadá y de los Estados Unidos. Entre los demás programas basados en Fellegi-Holt se encuentra CHERRYPI, programa de revisión basado en Fellegi-Holt elaborado por Statistics Netherlands (De Waal y Van de Pol, 1997).

8. En el censo del Canadá de 1996 (y en otros casos) se utilizó un enfoque diferente, denominado metodología de atribución basada en el vecino más próximo (MAVP). La versión de 1996 atribuía las respuestas correspondientes a la edad, sexo, estado civil y relación para todas las personas de una casa de manera simultánea (Bankier, 1999). El método fue mejorado y ampliado para el censo del Canadá de 2001 y posteriores actividades estadísticas (Bankier, Lachance y Poirier, 2000; 2001).

9. La MAVP busca primero los vecinos más próximos y luego determina las atribuciones de cambio mínimo sobre la base de esos donantes. Mientras que el método de Fellegi-Holt implica la atribución del menor número de variables y la conservación de la integridad de las subpoblaciones, la MAVP, que invierte el orden de la operación —comienza con la búsqueda de donantes y luego modifica el número mínimo de variables— presenta una ventaja desde el punto de vista del cálculo. Por otro lado, la MAVP puede realizar la atribución utilizando solo donantes, mientras que Fellegi-Holt puede emplearse con otras metodologías (por ejemplo, el enfoque secuencial). Statistics Canada incorporó la MAVP a su Canadian Census Edit and Imputation System para los censos de 2001 y 2006.

* En un examen de CANCEIS, Banquier, Lachance y Poirier (2000, pág. 10) resumen así sus conclusiones: CANCEIS, con sus algoritmos de revisión e imputación de gran eficiencia, ofrece grandes posibilidades para la solución de los problemas de imputación de alcance muy general que impliquen gran número de normas de revisión y de variables cualitativas y cuantitativas cuando resulta aconsejable la atribución del donante de cambio mínimo. En cambio, el algoritmo de revisión y atribución de cambio mínimo de Fellegi/Holt debería ser el método preferido cuando se trata de problemas de atribución de menor magnitud si no hay donantes suficientes disponibles o si resulta más adecuado utilizar otro método para la atribución.

10. Una serie de reuniones posteriores a 2000 agruparon a muchos estadísticos europeos interesados en examinar diversos aspectos de la revisión y atribución en los censos y encuestas. Los debates se centraron en la elaboración y evaluación de nuevos métodos para la revisión y la atribución (EUREEDIT) (Chambers, 2000; Poirier, 2000). Statistics Canada ha elaborado también un método para medir y reducir la variabilidad en la atribución. Su sistema SIMPVAR, aplicable en las estadísticas, trata de procesar los cuatro métodos principales de atribución (coeficiente, media, atribución dinámica y vecino más próximo) (Rancourt y otros, 1997). El Istituto Nazionale di Statistica (ISTAT) ha descrito el sistema de revisión y atribución para el censo de 2001 en Italia, en el que se utiliza DIESIS (Data Imputation and Editors System - Italian Software) y otros métodos (Di Zio, 2002; Bianchi y otros, 2005).

11. Existen otros métodos para la atribución de los valores desconocidos, además de la utilización de casos reales. A veces se utiliza el promedio de las medidas. Algunos países emplean modelos de regresión (Federación de Rusia, Comité Estatal de Estadística, 2000). La regresión se utilizó también para la atribución de la edad en el formulario breve de los Estados Unidos de 2000 (Williams, 1998).

12. Debido a los avances en la introducción de datos interactiva, algunos sistemas integran en un solo sistema la codificación, la introducción de los datos y la revisión, en particular en el caso de las encuestas. Entre ellos se encuentra el sistema CRIPTAX del Brasil (Hanono y Rocha Barbosa, sin fecha), que implica un método de revisión en el momento de la entrada. Otros sistemas, como el Census and Survey Processing System (CSPro), tienen aspectos de revisión interactiva. Como se ha señalado en otro lugar, las oficinas de estadística de los países deben decidir la rentabilidad de las inversiones realizadas en maquinaria, persona, tiempo, etcétera.

13. En cuanto al enfoque secuencial, la Oficina del Censo de los Estados Unidos elaboró el Integrated Microcomputer Processing System (IMPS) para los censos desde 1980 hasta el 2000. Este programa es en DOS e incluye las funciones de entrada, revisión y tabulación, entre otras. Muchos países todavía lo utilizan. Durante los decenios de 1990 y 2000, la Oficina del Censo elaboró la versión Windows, conocida con el nombre de CSPro (antes mencionada), que realiza muchas de las tareas del IMPS y es compatible con Windows*. CSPro funciona satisfactoriamente para las encuestas y los censos de países pequeños; no obstante, la elaboración es tan lenta que la revisión dura más que en IMPS. De todas formas, los países que empiezan sus revisiones desde cero deberían utilizar la versión Windows. Todas las aplicaciones en IMPS y CSPro pueden adaptarse al método de Fellegi-Holt; la aplicación del método del vecino más próximo requeriría mucho más trabajo.

14. CSPro permite al usuario crear, modificar y realizar la introducción de datos, la revisión por lotes y la tabulación a partir de un único entorno de desarrollo integrado. Procesa los datos caso por caso (uno o más cuestionarios) y un caso puede estar constituido por uno o muchos registros de datos. Los datos se almacenan en archivos de texto tipo American Standard Code for Information Interchange (ASCII) descritos por diccionarios de datos. CSPro contiene un potente procedimiento común para aplicar el control de las entradas y las normas de revisión.

15. En concreto, la función de revisión por lotes de CSPro localiza y notifica los errores de estructura, valor y congruencia en los datos del cuestionario. Puede elaborar informes sinópticos o detallados de los errores y correcciones, acceder a múltiples archivos de búsqueda y leer archivos secundarios o escribir en ellos.

16. CSPro ofrece también instrumentos para ver los archivos de datos y otros archivos de texto, ver los cuadros y mapas temáticos creados por CSPro, convertir diccionarios de datos IMPS e ISSA a y desde CSPro, y convertir archivos en formato

* CSPro es un programa informático para la entrada, revisión, tabulación y difusión de datos de censos y encuestas. Combina las características del Integrated Microcomputer Processing System (IMPS) y del Integrated System for Survey Analysis (ISSA) en un entorno Windows.

shapefiles (mapas) a archivos de mapas CSPro. CSPro fue creado por la Oficina del Censo de los Estados Unidos, Macro International y Serpro, con considerable financiación de la Agencia de los Estados Unidos para el Desarrollo Internacional. CSPro es de dominio público y puede obtenerse y distribuirse gratuitamente; puede descargarse en www.census.gov/ipc/www/cspro.

Glosario

Atribución — Asignación de un valor a un campo bien porque no hay respuesta o bien para sustituir un valor registrado que, según se ha determinado, es incongruente con un conjunto de normas de revisión (Pierzchala, 1995).

Atribución aleatoria dentro de una categoría — Utiliza categorías de atribución para limitar la selección al azar del registro donante a un conjunto que parece presentar ciertas similitudes con el registro al que hay que atribuir un valor.

Atribución aleatoria general — Respecto de cada registro no aceptado se selecciona al azar un registro aceptado del conjunto del total de registros aceptados, y el valor declarado del elemento en cuestión se atribuye al registro no aceptado.

Atribución basada en donantes — Un método consistente en emparejar cada registro para el que se necesitan valores atribuidos, denominado registro receptor o candidato, con un registro de una población donante definida, como, por ejemplo, en la atribución dinámica de valores (Pierzchala, 1995).

Atribución basada en el “vecino más próximo” — Se asigna a un registro no aceptado un valor del registro aceptado “más próximo”; la “proximidad” se define utilizando una función de distancia en relación con otras variables conocidas.

Atribución basada en un modelo — Este procedimiento utiliza los datos de registros aceptados anteriores para obtener para un elemento un valor tomado de una serie de variables independientes.

Atribución basada en valores medios, por categorías — Utiliza categorías de atribución de valores definidas para crear grupos de registros que presenten cierto grado de similitud.

Atribución deductiva — Un método que permite deducir con certeza un valor omitido o incongruente, a menudo sobre la base de la pauta de respuestas dadas a otros elementos del cuestionario.

Atribución determinista — Esta situación se plantea cuando en un campo sólo hay un valor que permitirá que el registro supere todas las revisiones. Se produce en algunos casos (por ejemplo, cuando la suma de las partes no es equivalente al total). Es la primera solución que hay que comprobar en un procedimiento de revisión y atribución automatizadas de los datos de encuestas (Pierzchala, 1995).

Atribución dinámica — Un método de atribución de valores en el que se toman registros donantes de la serie corriente de datos (la atribución estática, por el contrario, se refiere al método de atribución de valores en el que el registro donante procede de datos de estudios anteriores) (Pierzchala, 1995).

Atribución dinámica con un donante único — Atribuye a un registro no aceptado los datos de un único donante, lo cual permite la atribución conjunta de todos los valores de un registro que las revisiones hayan determinado como problemático.

Atribución dinámica secuencial — Atribución que se da cuando se revisan sucesivamente diversas variables y sólo se usan los valores revisados como variables para una nueva matriz dinámica.

Atribución estadística — Un ejemplo de atribución estadística sería la utilización de un modelo de regresión en el que habrá que atribuir la variable dependiente, y los coeficientes de las variables independientes se derivarán de las respuestas presumiblemente válidas (Pierzchala, 1995).

Atribución estática — La matriz estática inicial es una base de corrección cuyos elementos se han determinado antes de la corrección y no varían mientras se realiza. La base de corrección podría ser, por ejemplo, la información de un año anterior. Una variante del método de atribución estática podría consistir en ajustar los valores de la matriz con arreglo a datos corrientes (posiblemente agregados).

Atribución estructural — Esta atribución se utiliza cuando existe una relación estructural entre varias variables. Por ejemplo, un total deberá ser igual a la suma de sus partes; así pues, en el caso de una madre el número de los hijos nacidos vivos deberá ser igual al de los hijos vivos más los hijos fallecidos (Pierzchala, 1995).

Atribución general basada en valores medios — Asigna el valor medio de los registros aceptados a las variables omitidas o incongruentes de los registros no aceptados.

Atribución múltiple — Este procedimiento consiste en dar varias atribuciones a cada elemento que lo necesita y luego producir una estimación para el elemento.

Atribución por regresión — Utiliza datos de registros aceptados para tomar de un conjunto de variables independientes un valor para un elemento al que hay que atribuir un valor.

Búsqueda — En el procedimiento de atribución dinámica, la operación de buscar un registro donante (Pierzchala, 1995).

Captación (o captura) de datos — El proceso en el cual los datos recopilados adquieren una forma que pueda leer una máquina. A menudo se realizan comprobaciones elementales en los submódulos del programa de captación de datos.

Coherencia interna — Esta expresión se refiere a las relaciones entre las variables de una determinada unidad de muestra y es la causa de las revisiones de la mayor parte de los procedimientos censales (Ford, 1983; Pierzchala, 1995).

Conjunto completo de revisiones — Suma de las normas de revisión explícitas e implícitas. Es necesario para generar regiones plausibles para la atribución de valores (de modo que los valores atribuidos superen las revisiones) (Pierzchala, 1995).

Conjunto mínimo — El conjunto mínimo de campos a los que habrá que atribuir valores para asegurarse de que se superarán todas las normas de revisión (Pierzchala, 1995).

Corrección automatizada — Corrección de los errores que presentan los datos efectuada por ordenador, sin intervención humana. Es un aspecto de la revisión automatizada de datos (Pierzchala, 1995).

Correspondencia — En el procedimiento de atribución dinámica de valores, la operación consistente en encontrar un registro donante para un determinado registro receptor (candidato) (Pierzchala, 1995).

Correspondencia de la función de distancia — Este procedimiento asigna a un registro inválido el valor de un elemento tomado del registro aceptado “más próximo”. La proximidad se define utilizando una función de distancia con respecto a otras variables conocidas.

Correspondencia estadística (en la atribución dinámica) — La operación de encontrar un registro donante para un registro receptor (candidato) con arreglo a ciertos criterios estadísticos a fin de transferir los datos del donante al receptor (Pierzchala, 1995).

Diagrama de secuencia — Descripción gráfica de todas las funciones que deberán realizarse.

Errores de calidad — Errores que pueden distorsionar la calidad de los datos; por ejemplo: errores sistemáticos que provocan sesgos (Granquist, 1984; Pierzchala, 1995).

Función de distancia — Para los datos numéricos, una función definida a partir de las variables que guardan correspondencia en el registro receptor (candidato) y el registro donante y que se utiliza para cuantificar el concepto de similitud. Se aplica para encontrar registros que guarden correspondencia en el método de atribución dinámica (Pierzchala, 1995).

Indicador — Variable utilizada para señalar un elemento u otra variable a fin de referirse a ellos posteriormente. Los indicadores se utilizan, por ejemplo, para señalar los números de línea del “jefe del hogar” y del “cónyuge”, que más tarde se examinarán para asegurarse de que los cónyuges sean de sexos opuestos y de que ambos estén casados.

Introducción directa de datos — Un estilo de introducción de datos en el que el aparato utilizado no localiza los errores en los datos a medida que se van introduciendo, lo cual permite al operador introducir los datos de manera rápida y eficiente.

Introducción interactiva de datos — Un estilo de introducción de datos en el cual el aparato utilizado localiza los errores en los datos a medida que se van introduciendo, lo cual permite al operador corregir inmediatamente los errores (Pierzchala, 1995).

Lista de códigos — Lista de todos los valores permitidos (admisibles) para un elemento.

Matriz — Una serie de valores numéricos. Una matriz puede utilizarse para almacenar datos numéricos de carácter repetitivo.

Método de Fellegi-Holt para la corrección automática — Método de corrección automática en el que se modifica el menor número posible de elementos y se utiliza el modelo de Fellegi-Holt para determinar los conjuntos aceptables de valores o los límites para los elementos atribuidos. Pueden aplicarse métodos de atribución estática o dinámica secuencial o simultánea.

Normas de revisión (definición 1) — Limitaciones lógicas impuestas a los valores que puede asumir cada variable (Pierzchala, 1995).

Normas de revisión (definición 2) — Normas que localizan las combinaciones prohibidas de respuestas (Pierzchala, 1995).

Metodología de atribución basada en el vecino más próximo (MAVP)— Es un método similar al de Fellegi-Holt en tanto que considera simultáneamente todas las normas de revisión, no especifica explícitamente operaciones de atribución y atribuye valores tomados de un único donante. En la MAVP se examina cada registro no aceptado para identificar operaciones de atribución que supongan un cambio mínimo, siempre que se disponga de donantes potenciales.

Omisión — Cuestionario incompleto o en el que faltan valores (Pierzchala, 1995).

Ponderaciones — En la escuela de revisión y atribución de datos de Fellegi-Holt se asigna un coeficiente de ponderación a los campos sobre la base de la fiabilidad. Cuanto más elevado sea el coeficiente de ponderación, más probable será que se atribuya un campo (en igualdad de condiciones). También pueden asignarse coeficientes de ponderación a las normas de revisión (Pierzchala, 1995).

Pseudocódigo — Instrucciones o especificaciones escritas de revisión.

Rastro de verificación de cambios — Un método que permite efectuar un seguimiento de las modificaciones de los valores de un campo, y de la motivación y la fuente de cada cambio. Normalmente los rastros de verificación comienzan a seguirse una vez concluida la entrevista inicial.

Rastro de verificación de revisiones — Véase “rastros de verificación de cambios”.

Registro — Una representación de los datos de un estudio almacenada en soporte magnético y que puede ser leída por un ordenador. Normalmente hay un registro para cada cuestionario, aunque es posible distribuir los datos de un solo cuestionario en más de un registro, como los de población y vivienda (Pierzchala, 1995).

Registro depurado — Un registro en el que no faltan valores y que ha superado todas las revisiones (Pierzchala, 1995).

Registros aceptados — En la revisión y la atribución de valores, los registros que han superado todas las revisiones pertinentes a los elementos en cuestión.

Registros no aceptados — En la revisión y atribución de valores, los registros que no han superado por lo menos una especificación de revisión correspondiente al elemento de que se trata.

Revisión de coherencia — Comprobación de relaciones determinantes, como las partes que deben ser equivalentes a cierto total; la “superficie de cosecha” siempre tiene que ser menor que la “superficie sembrada” (Pierzchala, 1995).

Revisión de macronivel — Detección de errores concretos mediante comprobaciones de datos agregados, o por comprobaciones aplicadas a todo el conjunto de registros. Las comprobaciones se basan en las estimaciones (Granquist, 1987; Pierzchala, 1995).

Revisión de micronivel — Revisión tradicional efectuada en los datos a nivel de los registros. Es el antónimo lógico de la revisión de macronivel (Pierzchala, 1995).

Revisión de micronivel/macronivel — Un procedimiento de revisión en el cual los procedimientos detallados de revisión de micronivel se sustituyen por una combinación de revisiones de micronivel y una revisión de macronivel sobre una base estadística. Las revisiones de micronivel del procedimiento combinado son menos detalladas que en el primer caso. Se trata de efectuar revisiones sobre la base de tener en cuenta los efectos en las estimaciones, y no de localizar todas las incongruencias de los datos (Granquist, varias fechas; Pierzchala, 1995).

Revisión dentro de un mismo registro — Sinónimo de revisión de validación (Pierzchala, 1995).

Revisión determinista — Una norma de revisión que, de ser transgredida, indica un error en los datos con una probabilidad de 1. Por ejemplo: edad = 5 y relación = madre. Contrasta con la revisión estocástica (Pierzchala, 1995).

Revisión de validación — Una revisión de comprobación realizada entre campos de un registro particular. Incluye comprobar todos los campos de cada registro para ver si contienen un asiento válido y comprobar los asientos en una combinación predeterminada de campos, para conocer si los valores son congruentes entre sí (Pierzchala, 1995).

Revisión entre registros — Revisión efectuada en campos que afectan a más de un registro de la encuesta. Las revisiones estadísticas son un ejemplo de revisiones entre registros porque se generan distribuciones a partir de conjuntos de campos que abarcan todos los registros de la encuesta (Pierzchala, 1995).

Revisión estadística — Serie de comprobaciones basadas en un análisis estadístico de los datos del declarante: por ejemplo, la proporción entre dos campos deberá

estar entre unos límites determinados por un análisis estadístico de esa proporción para declarantes, presumiblemente válidos (Greenberg y Surdi, 1984; Pierzchala, 1995).

Revisión estocástica — Una norma de revisión que, de transgredirse, indica un error en los datos con una probabilidad inferior a 1 (Pierzchala, 1995).

Revisión estructural — Comprobaciones basadas en una relación lógica entre dos o más campos revisados. Por ejemplo, un total debe ser equivalente a la suma de sus partes; o, debido a una pauta de omisión inherente a un cuestionario, dos variables no pueden ser ambas distintas de 0. La revisión estructural asegura que la estructura del cuestionario se mantenga en el registro de datos (Pierzchala, 1995).

Revisión explícita — Normas de revisión formuladas explícitamente por un especialista temático (contrasta con la revisión implícita) (Pierzchala, 1995).

Revisión implícita — Una norma de revisión no declarada derivada lógicamente de normas de revisión explícitas formuladas por un especialista temático (Pierzchala, 1995).

Revisión manual — Una revisión realizada por una persona antes de introducir los datos en el ordenador (Pierzchala, 1995).

Revisión multidimensional — Revisión estadística que usa distribuciones de múltiples variables para evaluar los datos y hallar valores que queden fuera de los límites (Pierzchala, 1995).

Revisiones cuantitativas — Revisiones aplicadas a campos medidos en una escala continua (Pierzchala, 1995).

Señal de alerta — Se trata de una variable utilizada para señalar información útil acerca de otra variable o variables. Por ejemplo, si el valor de un elemento es inválido y se modifica para que sea válido, la señal de alerta puede utilizarse para indicar la información original o para denotar que se ha modificado el valor del elemento.

Serie de producción — La operación de procesar grandes cantidades de datos una vez que se han eliminado las deficiencias iniciales del programa de revisión o de tabulación.

Similitud — En los datos numéricos, un concepto de proximidad de dos registros sobre la base de unas variables de correspondencia prescritas. Se utiliza una función de distancia para cuantificar este concepto con arreglo a ciertos criterios (Pierzchala, 1995).

Sistema de Fellegi-Holt — Se refiere a las hipótesis y objetivos de revisión y atribución de valores presentados por Fellegi y Holt en su estudio publicado en *Journal of the American Statistical Association* (1976). Una característica fundamental del modelo de Fellegi-Holt es que demuestra que las revisiones implícitas son necesarias para asegurar que un conjunto de valores en campos de datos no atribuidos produzca siempre registros finales (con valores atribuidos) que puedan ser satisfactorios en todos los casos.

Valores “fuera de los límites” — Elementos que tienen un valor que está fuera de algún límite, con arreglo a cierta determinación de los límites (Pierzchala, 1995).

Variables de correspondencia — Las variables que se utilizan para encontrar un registro donante para un registro receptor (candidato) (Pierzchala, 1995).

Referencias bibliográficas

- Banister, J. (1980). Use and abuse of census editing and imputation. *Asian and Pacific Census Forum*, vol. 6, No. 3, págs. 1 a 20.
- Bankier, M. (1999). "Experience with the New Imputation Methodology used in the 1996 Canadian Census with Extensions for Future Censuses", *Actas del Taller sobre revisión de datos estadísticos, Naciones Unidas/Comisión Económica para Europa*, Italia (Roma).
- _____ (2005). "Edit and imputation for the 2006 Canadian Census". Trabajo presentado en la Sesión de trabajo sobre la revisión de datos estadísticos, Ottawa, 16 a 18 de mayo de 2005.
- _____, A.-M. Houle y M. Luc (s. f.). "Canadian census demographic variables imputation". Manuscrito.
- Bankier, M., M. Lachance y P. Poirier (1999). A generic implementation of the new imputation methodology. En *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginia: American Statistical Association (de próxima aparición).
- _____ (2000). "2001 Canadian Census minimum change donor imputation methodology". Trabajo presentado en la Sesión de trabajo de la CEPE de las Naciones Unidas sobre revisión de datos estadísticos, Cardiff (Reino Unido), 18 a 20 de octubre de 2000.
- Bankier, M., P. Poirier y M. Lachance (2001). Efficient methodology within the Canadian Census Edit and Imputation System (CANCEIS). *Actas de la Reunión anual de la American Statistical Association*, 5 a 9 de agosto de 2001.
- Bankier, M. y otros (1996). Imputing numeric and qualitative census variables simultaneously. En *Proceedings of the Section on Survey Research Methods*. Alexandria, Virginia: American Statistical Association, págs. 287-292.
- Bianchi, G y otros (2005). "New procedures for editing and imputation of demographic variables". Trabajo presentado en la Sesión de trabajo sobre la revisión de datos estadísticos, Ottawa, 16 a 18 de mayo de 2005.
- Boucher, L. (1991). Micro-editing for the annual salary of manufacturers: what is the value added? En *Proceedings of the Annual Research Conference*. Washington, D.C.: United States Bureau of the Census, págs. 765-781.
- Chambers, Ray (2000). "Evaluation criteria for editing and imputation in EUREDIT". Trabajo presentado en la Sesión de trabajo de la CEPE de las Naciones Unidas sobre revisión de datos estadísticos, Cardiff (Reino Unido), 18 a 20 de octubre de 2000.
- Chen, Bor-Chung (2007). *CANCEIS Experiments of edit and imputation with 2006 Census test data*. Statistical Research Division Study Series, No. 2007-1. Washington, D.C.: U.S. Bureau of the Census.

- _____ y otros (2000). "Using the DISCRETE Edit System for ACS Surveys". Bureau of the Statistical Research Division Statistical Research Report Series, No. RR2000/03. Washington, D.C.: U.S. Bureau of the Census.
- Comisión de Estadística y Comisión Económica para Europa (1994). *Statistical Data Editing*, vol. 1: *Methods and Techniques*. Conference of European Statisticians, Statistical Standards and Studies Series, No. 44. No. de venta: 94.II.E.36.
- _____ (1997). *Statistical Data Editing*, vol. 2: *Methods and Techniques*. Conference of European Statisticians, Statistical Standards and Studies Series, No. 48. No. de venta: 96.II.E.30.
- De Waal, Tom y Frank van de Pol (1997). "A recipe for applying CHERRYPI in the edit process". Trabajo presentado en la Sesión de trabajo sobre la revisión de datos estadísticos, Praga, 14 a 17 de octubre de 1997.
- Di Zio, M. (2002). "Evaluating editing and imputation processes: the Italian experience". Trabajo presentado en la Sesión de trabajo de la CEPE de las Naciones Unidas sobre revisión de datos estadísticos, Helsinki, 27 a 29 de mayo de 2002.
- Federación de Rusia, Comité Estatal de Estadística (2000). Data imputation based on regression models with variations of entropy. Trabajo presentado en la Sesión de trabajo de la CEPE de las Naciones Unidas sobre revisión de datos estadísticos, Cardiff (Reino Unido), 18 a 20 de octubre de 2000.
- Fellegi, I. P. y D. Holt (1976). A systematic approach to automatic edit and imputation. *Journal of the American Statistical Association*, vol. 71, No. 353 (marzo), págs. 17 a 35.
- Ford, Barry L. (1983). An overview of hot deck procedures. En *Incomplete Data in Sample Surveys*, vol. 2, *Theory and Bibliographies*. William G. Madow. Ingram Olkin y Donald B. Rubin, eds.
- Granquist, L. (1984). "Data editing and its impact on the further processing of statistical data". Trabajo presentado en el Seminario sobre informática estadística", Budapest, 12 a 17 de noviembre de 1984.
- _____ (1987). "The short-term developing program for computer-supported editing at Statistics Sweden". Informe presentado en la reunión del Grupo mixto de revisión de datos, Madrid, 22 a 24 de abril de 1987. Estocolmo: Statistics Sweden.
- _____ (1997). The new view on editing. *International Statistical Review*, vol. 65, No. 3, Nueva York: Academic Press, págs. 381 a 387.
- _____, y J. G., Kovar (1997). Editing of survey data: how much is enough? En *Survey Measurement and Process Quality*, Lyberg y otros, eds. Nueva York: Wiley and Sons, págs. 415 a 435.
- Greenberg, Brian y Rita Surdi (1984). A flexible and interactive edit and imputation system for ratio edits. En *Proceedings of the American Statistical Association*, Section on Survey Research Methods. Alexandria, Virginia: American Statistical Association, págs. 421-426.
- Hanono, Reina Marta y Dulce Maria Rocha Barbosa (sin fecha). *Generalized environment for application development for capturing, editing and coding statistical survey's data*. Rio de Janeiro: Brazilian Institute of Geography and Statistics (IBGE).
- Ireback, H. (2000). "The impact of new information technologies on data collection at Statistics Sweden". Trabajo presentado en la Sesión de trabajo de la CEPE de

- las Naciones Unidas sobre revisión de datos estadísticos, Cardiff (Reino Unido), 18 a 20 de octubre de 2000.
- Kalton, G. y D. Kasprzyk (1982). Imputing for missing survey responses. En *Proceedings of the Section on Survey Research Method*. American Statistical Association, págs. 23 a 31.
- _____ (1986). The Treatment of Missing Survey Data. *Survey Methodology*, vol. 12, págs. 1 a 16.
- Kovar, J. y W. Winkler (2000). *Comparison of GEIS and SPEER for editing economic data*. U.S. Bureau of the Census, Statistical Research Division, Statistical Research Report Series, No. RR 2000/04. Washington, D.C.: U.S. Bureau of the Census.
- Naciones Unidas (1992a). *Manual de censos de población y habitación, parte I: Planificación, organización y administración de los censos de población y de habitación*. Estudios de métodos, Serie F, No. 54. No. de venta: S.92.XVII.8.
- _____ (1992b). *Manual de censos de población y habitación, parte II: Características demográficas y sociales*. Estudios de métodos, Serie F, No. 54. No. de venta: S.91.XVII.9.
- _____ (1999). *Códigos uniformes de país o de zona para uso estadístico*. Informes estadísticos, Serie M, No. 49/Rev. 4. No. de venta: M.98.XVII.9.
- _____ (2008). *Principios y recomendaciones para los censos de población y habitación, Revisión 2*. Informes estadísticos, Serie M, No. 67/Rev.2. No. de venta: S.07.XVII.8.
- Naus, J. I. (1975). *Data Quality Control and Editing*. Nueva York: Marcel Dekker.
- Nordbotten, S. (1963). *Automatic editing of individual statistical observations*. Conference of European Statisticians, Statistical Standards and Studies, No. 2. Nueva York: Naciones Unidas.
- Pierzchala, M. (1995). Editing systems and software. En *Business Survey Methods*. B. G. Cox, y otros, eds. Nueva York: John Wiley and Sons, págs. 425 a 441.
- Poirier, C. (2000). "A prototype knowledge base on data editing and imputation". Trabajo presentado en la Sesión de trabajo de la CEPE de las Naciones Unidas sobre revisión de datos estadísticos, Cardiff, (Reino Unido), 18 a 20 de octubre de 2000.
- Pullum, T. W., T. Harpham y N. Ozsever (1986). The machine editing of large-sample surveys: the experience of the World Fertility Survey. *International Statistical Review*, vol. 54, 311 a 326.
- Rancourt, E. y otros (1997). Estimation of variance in presence of imputation. *Proceedings of Symposium 1997: New Directions in Surveys and Censuses*. Ottawa: Statistics Canada, págs. 273 a 279.
- Rubin, D. B. (1987). *Multiple Imputation for Non-response in Surveys*. Nueva York: J. Wiley & Sons, xxix+258 páginas.
- Sande, I. G. (1982). Imputation in surveys: coping with reality. *American Statistician*, vol. 36, págs. 145 a 152.
- Särndal, C. E., B. Swensson y J. Wretman (1992). *Model Assisted Survey Sampling*. Nueva York: Springer-Verlag.

- Statistics Canada (1998). *Statistics Canada Quality Guidelines*, 3a. edición. Ottawa: Statistics Canada.
- Williams, Todd R. (1998) *Imputing person age for the 2000 Census Short Form: a model-based approach*. Washington, D.C.: U.S. Bureau of the Census, Statistical Research Division, Statistical Research Report Series, No. RR98/07.
- Winkler, W. E. (1997a). "Edit/imputation System for the U.S. Decennial Census". Trabajo presentado en la Sesión de trabajo sobre la revisión de datos estadísticos, Praga, 14 a 17 de octubre de 1997.
- _____ (1997b). *Set-covering and editing discrete data. Technical report*. Washington, D.C.: U.S. Bureau of the Census.
- _____ (2006). *Data quality: automated edit/imputation and record linkage*. U.S. Census Bureau, Statistical Research Division, Research Report Series, No. 2006-7. Washington, D.C.: U.S. Bureau of the Census.
- _____ y B. C. Chen (2002). *Extending the Fellegi-Holt Model of Statistical Data Editing*. U. S. Bureau of the Census, Statistical Research Division, Statistical Research Report Series, No. 2002-02.
- Winkler, W. E. y L. R. Draper (1997). The SPEER edit system. En Comisión de Estadística y Comisión Económica para Europa, ed., *Statistical Data Editing*, vol. 2: *Methods and Techniques*, págs. 56 a 62. Conference of European Statisticians, Statistical Standards and Studies Series, No. 48. No. de venta: E.96.II.E.30.