

Encuesta Nacional de Cultura Cívica 2020

ENCUCI

Diseño muestral



Instituto Nacional de Estadística y Geografía

Encuesta Nacional de Cultura Cívica 2020

ENCUCI

Diseño muestral



Obras complementarias publicadas por el INEGI sobre el tema:

Encuesta Nacional sobre Cultura Política y Prácticas Ciudadanas 2001. ENCUP; Encuesta Nacional sobre Cultura Política y Prácticas Ciudadanas. ENCUP 2008. Manual del entrevistador; Encuesta Nacional sobre Cultura Política y Prácticas Ciudadanas. ENCUP 2008. Documento metodológico; Encuesta Nacional sobre Cultura Política y Prácticas Ciudadanas. ENCUP 2008. Actividades del instructor supervisor; Encuesta Nacional de Cultura Cívica 2020. ENCUCI. Marco conceptual; Encuesta Nacional de Cultura Cívica 2020. ENCUCI. Informe operativo.

Catalogación en la fuente INEGI:

306.97201 Encuesta Nacional de Cultura Cívica (2020).
Encuesta Nacional de Cultura Cívica 2020 : ENCUCI : diseño muestral / Instituto Nacional de Estadística y Geografía.-- México : INEGI, c2021.

vii, 16 p.

1. México - Cultura - Metodología. 2. México - Condiciones sociales - Metodología. I. Instituto Nacional de Estadística y Geografía (México).

Conociendo México

800 111 4634

www.inegi.org.mx

atencion.usuarios@inegi.org.mx



INEGI Informa



@INEGI_INFORMA

Presentación

El **Instituto Nacional de Estadística y Geografía (INEGI)**, en el marco de las atribuciones que le confiere la Ley del Sistema Nacional de Información Estadística y Geográfica, realiza la documentación de los diferentes proyectos que lleva a cabo.

Por lo anterior presentamos el **Documento Metodológico de la Encuesta Nacional de Cultura Cívica (ENCUCI)** con la intención de dar a conocer los principales rasgos de la metodología utilizada para el diseño y desarrollo estadístico de este proyecto.

Con la finalidad de transparentar el proceso de generación de información estadística, el **INEGI** pone a disposición de los usuarios este documento y de esta manera se dan a conocer los procedimientos y lineamientos que se siguieron y que hicieron posible la realización del proyecto.

Índice

Introducción	VII
1. Objetivo de la encuesta	1
2. Población objetivo	1
3. Cobertura geográfica	1
4. Marco de la encuesta	1
4.1 Formación de las Unidades Primarias de Muestreo (UPM)	2
4.1.1 En urbano alto	2
4.1.2 En complemento urbano	2
4.1.3 En rural	2
4.2 Estratificación	3
4.3 Selección de las UPM de la muestra maestra	3
5. Tamaño de la muestra	4
6. Afijación de la muestra	4
7. Selección de la muestra	5
7.1 En urbano alto	5
7.2 En complemento urbano	6
7.3 En rural	7
8. Ajuste a los factores de expansión	9
8.1 Ajuste por no respuesta	9
8.1.1 Ajuste por no respuesta a nivel vivienda	9
8.1.2 Ajuste por no respuesta a nivel hogar	9
8.1.3 Ajuste por no respuesta a nivel persona	10
8.2 Ajuste por proyección	10
9. Estimadores	10
10. Estimaciones de errores de muestreo	11
Anexo	13
A. Indicadores empleados en la estratificación de la muestra maestra	15
B. Distribución de viviendas seleccionadas por entidad y dominio de la muestra para la Encuesta Nacional de Cultura Cívica (ENCUCI) 2020	16

Introducción

La **Encuesta Nacional de Cultura Cívica (ENCUCI) 2020** es resultado de la colaboración entre el **Instituto Nacional de Estadística y Geografía (INEGI)** y el **Instituto Nacional Electoral (INE)**, en el marco de los trabajos para la elaboración del Informe País 2020 que busca conocer los valores y prácticas de los(las) ciudadanos(as), con el objetivo de entender el involucramiento de la ciudadanía en los asuntos públicos, diseñar estrategias para promover esa participación, así como la interacción constructiva con autoridades e instituciones. Con la ENCUCI se busca conocer las características de la cultura cívica y el ejercicio de ciudadanía de las y los mexicanos.

Con el propósito de enmarcar las principales características metodológicas del proyecto, se realiza el Diseño Muestral de la ENCUCI 2020, el cual aborda las principales características del diseño estadístico empleado para el desarrollo de la encuesta, como son el marco de muestreo, el tamaño y la distribución de la muestra, así como el cálculo de las probabilidades de inclusión de las unidades de observación, los factores de expansión y los principales estimadores para la explotación de información.

De esta manera, el **INEGI** da a conocer la metodología empleada y contribuye a la transparencia del proceso de generación de información estadística.

1. Objetivo de la encuesta

El objetivo de la ENCUCI 2020 es obtener información estadística que permita generar estimaciones sobre creencias, valores, actitudes y prácticas que estructuran y modelan la relación de las personas de 15 años y más con el poder público, con diferentes formas asociativas y con otros individuos, así como generar estimaciones sobre la pertenencia y reconocimiento de derechos y obligaciones como ciudadano(a) mexicano(a).

2. Población objetivo

La encuesta está dirigida a la población de 15 años cumplidos o más, que residen permanentemente en viviendas particulares dentro del territorio nacional.

3. Cobertura geográfica

La encuesta está diseñada para dar resultados a los siguientes niveles de desagregación:

- Nacional
 - Urbano
 - Rural
- Regiones:

Región 1: Aridoamérica Occidental	Región 2: Aridoamérica Oriental	Región 3: Zona Maya
Aguascalientes Baja California Baja California Sur Chihuahua Durango Sinaloa Sonora Zacatecas	Coahuila de Zaragoza Nuevo León San Luis Potosí Tamaulipas	Campeche Chiapas Quintana Roo Tabasco Yucatán
Región 4: Mesoamérica	Región 5: Mesoamérica Occidental	Región 6: Mesoamérica Central
Guerrero Oaxaca Puebla Tlaxcala Veracruz de Ignacio de la Llave	Colima Guanajuato Jalisco Michoacán de Ocampo Nayarit Querétaro	Ciudad de México Hidalgo México Morelos

4. Marco de la encuesta

El diseño de la muestra para la ENCUCI 2020 se caracteriza por ser probabilístico; en consecuencia, los resultados obtenidos de la encuesta se generalizan a toda la población. A la vez, el diseño es trietápico, estratificado y por conglomerados, donde la unidad última de selección es la vivienda y las unidades de observación son el hogar y las personas de 15 años cumplidos o más.

Para la selección de la muestra de la ENCUCI 2020, se utilizó el Marco Maestro de Muestreo 2012 del INEGI, construido a partir de la información cartográfica y demográfica que se obtuvo del Censo de Población y Vivienda 2010. En el marco, se seleccionó una muestra maestra a partir de la cual se seleccionan las submuestras para todas las encuestas en viviendas que realiza el INEGI; su diseño es probabilístico, estratificado, unietápico y por conglomerados; estos últimos también se consideran unidades primarias de muestreo, pues es en ellos donde se seleccionan, en una segunda etapa, las viviendas que integran las muestras de las diferentes encuestas. La muestra maestra se construye de la siguiente manera:

4.1 Formación de las Unidades Primarias de Muestreo (UPM)

En primer lugar, se construye el conjunto de UPM que cubrirá el territorio nacional.

Las unidades primarias de muestreo están constituidas por agrupaciones de viviendas con características diferenciadas dependiendo del ámbito al que pertenecen, como se especifica a continuación:

4.1.1 En urbano alto

El tamaño mínimo de una UPM es de 80 viviendas habitadas y el máximo es de 160. Pueden estar formadas por:

- Una manzana.
- La unión de dos o más manzanas contiguas de la misma AGEB.¹
- La unión de dos o más manzanas contiguas de diferentes AGEB de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes localidades, que pertenezcan al mismo tamaño de localidad.

4.1.2 En complemento urbano

El tamaño mínimo de una UPM es de 160 viviendas habitadas y el máximo es de 300. Pueden estar formadas por:

- Una manzana.
- La unión de dos o más manzanas contiguas de la misma AGEB.
- La unión de dos o más manzanas contiguas de diferentes AGEB de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes AGEB y localidades, pero del mismo municipio.

4.1.3 En rural

El tamaño mínimo de una UPM es de 160 viviendas habitadas y el máximo es de 300. Pueden estar formadas por:

- Una AGEB.
- Parte de una AGEB.
- La unión de dos o más AGEB colindantes del mismo municipio.
- La unión de una AGEB con una parte de otra AGEB colindante del mismo municipio.

De esta manera, la muestra maestra está conformada con un total de 240 912 UPM a nivel nacional.

¹ Área Geoestadística Básica.

4.2 Estratificación

Una vez construido el conjunto de UPM, se agrupan aquellas con características similares, de acuerdo a lo siguiente:

La división política del país y la conformación de localidades diferenciadas por su tamaño, forman de manera natural una estratificación geográfica. En cada entidad federativa se distinguen tres ámbitos, divididos a su vez en zonas, como se indica en el siguiente cuadro:

Ámbito	Zona	Descripción
Urbano alto	01 a 09	Ciudades con 100 000 o más habitantes.
Complemento urbano	25	De 50 000 a 99 999 habitantes
	35	De 15 000 a 49 999 habitantes
	45	De 5 000 a 14 999 habitantes
	55	De 2 500 a 4 999 habitantes
Rural	60	Localidades menores de 2 500 habitantes.

De manera paralela, se formaron cuatro estratos sociodemográficos en los que se agruparon todas las UPM del país, esta estratificación considera las características sociodemográficas de los habitantes de las viviendas, así como las características físicas y el equipamiento de las mismas, expresadas por medio de 34 indicadores* construidos con información del Censo de Población y Vivienda 2010, para lo cual se emplearon métodos estadísticos multivariados.

De esta forma, cada UPM fue clasificada en un único estrato geográfico y uno sociodemográfico. Como resultado, se tienen un total de 683 estratos en todo el ámbito nacional.

4.3 Selección de las UPM de la muestra maestra

Las UPM de la muestra maestra fueron seleccionadas por medio de un muestreo con probabilidad proporcional al tamaño, esto es:

$$P\{U_{ehi} \in S\} = \frac{n_{eh} m_{ehi}}{m_{eh}}$$

Donde:

U_{ehi} = UPM i-ésima, en el h-ésimo estrato, en la e-ésima entidad.

S = muestra maestra.

n_{eh} = número de UPM en la muestra maestra, en el h-ésimo estrato, en la e-ésima entidad.

m_{ehi} = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad en el Censo de Población y Vivienda 2010.

m_{eh} = número de viviendas en el h-ésimo estrato, en la e-ésima entidad en el Censo de Población y Vivienda 2010.

Las UPM seleccionadas forman la muestra maestra que permite seleccionar las submuestras de las encuestas en hogares, en las siguientes etapas del diseño estadístico.

* La descripción de estos indicadores se presenta en el anexo A.

5. Tamaño de la muestra

El tamaño de la muestra de la ENCUCI 2020, está calculado para la proporción de personas que identifica la pertenencia a un país como ciudadanía, considerada de las variables principales de la encuesta. Esto garantiza que las estimaciones del resto de las variables de interés queden cubiertas con ese tamaño.

La expresión empleada para el cálculo es la siguiente:

$$n = \frac{z^2 q \text{ DEFF}}{r^2 p (1 - \text{tnr})}$$

Donde:

n = tamaño de la muestra.

p = estimación de la proporción de interés.

q = 1-p.

r = error relativo máximo esperado.

z = valor asentado en las tablas estadísticas que garantiza realizar las estimaciones con una confianza prefijada.

DEFF = efecto de diseño definido como el cociente de la varianza en la estimación del diseño utilizado, entre la varianza obtenida considerando un muestreo aleatorio simple para un mismo tamaño de muestra.

tnr = tasa de no respuesta máxima esperada.

Considerando una confianza del 90%, un efecto de diseño de 2.0 observado en encuestas anteriores, así como una tasa de no respuesta máxima esperada del 15%, un error relativo máximo esperado del 7%, para una proporción de 4.9%, se obtuvo un tamaño de muestra a nivel nacional de 25 113 viviendas.

Para las regiones, se consideró una confianza del 90%, un efecto de diseño de 2.0 observado en encuestas anteriores, así como una tasa de no respuesta máxima esperada del 15%, un error relativo máximo esperado del 15%, para una proporción entre 6.3 y 6.4% dependiendo la región, se obtuvo un tamaño de muestra mínima de 4 166 y máxima de 4 202.

6. Afijación de la muestra

La afijación de la muestra se realizó dentro de cada entidad federativa entre los estratos, de manera proporcional a su tamaño en viviendas, para lo cual se empleó la siguiente expresión:

$$n_{eh} = \frac{N_{eh}}{N_e} n_e$$

El número de UPM a seleccionar en el h-ésimo estrato, en la e-ésima entidad se obtiene a partir de la siguiente expresión:

$$k_{eh} = \frac{n_{eh}}{b}$$

Donde:

= número de viviendas en muestra en el h-ésimo estrato, en la e-ésima entidad.

n_{eh} = número total de viviendas en muestra en la e-ésima entidad.

n_e = número total de viviendas en el h-ésimo estrato, en la e-ésima entidad.

N_{eh} = número total de viviendas en la e-ésima entidad.

N_e = número de UPM a seleccionar en el h-ésimo estrato, en la e-ésima entidad.

k_{eh} = número de viviendas a seleccionar por UPM.

b

En el anexo B, se presenta la distribución de la muestra en viviendas respectivamente, por entidad y dominio.

7. Selección de la muestra

La selección de la muestra para la ENCUCI 2020, se realizó de manera independiente por entidad, dominio y estrato, el procedimiento de selección varió de acuerdo con el dominio.

7.1 En urbano alto

1. De las k_{eh} UPM seleccionadas en el h-ésimo estrato, en la e-ésima entidad para la muestra maestra, se seleccionaron k_{eh}^* con igual probabilidad, para la ENCUCI 2020.
2. En cada UPM seleccionada, se seleccionaron cinco viviendas con igual probabilidad.
3. En cada vivienda seleccionada, se seleccionó una persona de 15 años cumplidos o más.

La probabilidad de selección de las viviendas y personas se calculó de la siguiente manera:

- a) La probabilidad de selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra maestra está definida por:

$$P_{1ehi} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

- b) La probabilidad de seleccionar en la muestra maestra a la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de ENCUCI 2020 está definida por:

$$P_{2ehi} = \frac{k_{eh}^*}{k_{eh}}$$

- c) La probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad está definida por:

$$P_{3ehij} = \frac{5}{m_{ehi}^*}$$

- d) La probabilidad de seleccionar una persona, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

$$P_{4ehij} = \frac{1}{Q_{ehij}}$$

Por lo tanto, la probabilidad total de selección de una persona de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es el producto de las probabilidades de selección:

$$P_{ehij} = P_{1ehi} * P_{2ehi} * P_{3ehij} * P_{4ehij}$$

$$P_{ehij} = \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{5}{m_{ehi}^*} \frac{1}{Q_{ehij}} = \frac{5 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^* Q_{ehij}}$$

Su factor de expansión² está dado por:

$$F_{ehij} = \frac{Q_{ehij} m_{eh} m_{ehi}^*}{5 k_{eh}^* m_{ehi}}$$

Donde:

- k_{eh} = número de UPM seleccionadas en el h-ésimo estrato, en la e-ésima entidad para la muestra maestra.
- m_{eh} = número de viviendas en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi} = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- m_{ehi}^* = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según listado de viviendas actualizado.
- k_{eh}^* = número de UPM a seleccionar para la ENCUCI 2020 con igual probabilidad de selección, en el h-ésimo estrato, en la e-ésima entidad.
- Q_{ehij} = número de personas de 15 años cumplidos o más, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

7.2 En complemento urbano

1. De las k_{eh} UPM que integran la muestra maestra se seleccionaron k_{eh}^* UPM con igual probabilidad para la ENCUCI 2020.
2. En cada UPM seleccionada, se seleccionaron 20 viviendas con igual probabilidad.
3. En cada vivienda seleccionada se seleccionó una persona de 15 años cumplidos o más con igual probabilidad.

La probabilidad de selección de las viviendas y personas se calculó de la siguiente manera:

- a) La probabilidad de selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra maestra está definida por:

$$P_{1ehi} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

- b) La probabilidad de seleccionar en la muestra maestra a la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de ENCUCI 2020 está definida por:

$$P_{2ehi} = \frac{k_{eh}^*}{k_{eh}}$$

² El factor de expansión se define como el inverso de la probabilidad de selección.

c) La probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad está definida por:

$$P_{3ehij} = \frac{20}{m_{ehi}^*}$$

d) La probabilidad de seleccionar una persona, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

$$P_{4ehij} = \frac{1}{Q_{ehij}}$$

Por lo tanto, la probabilidad total de selección de una persona de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es el producto de las probabilidades de selección:

$$P_{ehij} = P_{1ehi} * P_{2ehi} * P_{3ehij} * P_{4ehij}$$

$$P_{ehij} = \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{20}{m_{ehi}^*} \frac{1}{Q_{ehij}} = \frac{20 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^* Q_{ehij}}$$

Su factor de expansión está dado por:

$$F_{ehij} = \frac{Q_{ehij} m_{eh} m_{ehi}^*}{20 k_{eh}^* m_{ehi}}$$

Donde:

- k_{eh} = número de UPM seleccionadas en el h-ésimo estrato, en la e-ésima entidad para la muestra maestra.
- m_{eh} = número de viviendas en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi} = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- m_{ehi}^* = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según listado de viviendas actualizado.
- k_{eh}^* = número de UPM a seleccionar para la ENCUCI 2020 con igual probabilidad de selección, en el h-ésimo estrato, en la e-ésima entidad.
- Q_{ehij} = número de personas de 15 años cumplidos o más, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

7.3 En rural

1. De las k_{eh} UPM se seleccionaron k_{eh}^* UPM para la ENCUCI 2020 con igual probabilidad.
2. En cada UPM seleccionada, se seleccionaron cuatro segmentos de cinco viviendas aproximadamente con igual probabilidad.
3. En cada vivienda seleccionada se selecciona una persona de 15 años cumplidos o más, con igual probabilidad.

La probabilidad de selección de las viviendas y personas se calculó de la siguiente manera:

- a) La probabilidad de selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra maestra está definida por:

$$P_{1ehi} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

- b) La probabilidad de seleccionar en la muestra maestra a la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de ENCUCI 2020 está definida por:

$$P_{2ehi} = \frac{k_{eh}^*}{k_{eh}}$$

- c) La probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad está definida por:

$$P_{3ehij} = \frac{4 * 5}{m_{ehi}^*}$$

- d) La probabilidad de seleccionar una persona, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

$$P_{4ehij} = \frac{1}{Q_{ehij}}$$

Por lo tanto, la probabilidad total de selección de una persona de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es el producto de las probabilidades de selección:

$$P_{ehij} = P_{1ehi} * P_{2ehi} * P_{3ehij} * P_{4ehij}$$

$$P_{ehij} = \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{20}{m_{ehi}^*} \frac{1}{Q_{ehij}} = \frac{20 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^* Q_{ehij}}$$

Su factor de expansión está dado por:

$$F_{ehij} = \frac{Q_{ehij} m_{eh} m_{ehi}^*}{20 k_{eh}^* m_{ehi}}$$

Donde:

- k_{eh} = número de UPM seleccionadas en el h-ésimo estrato, en la e-ésima entidad, para la muestra maestra.
- m_{eh} = número de viviendas en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi} = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según Censo de Población y Vivienda 2010.
- m_{ehi}^* = número de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, según listado de viviendas actualizado.
- k_{eh}^* = número de UPM seleccionadas para la ENCUCI 2020, en el h-ésimo estrato, en la e-ésima entidad.
- Q_{ehij} = número de personas de 15 años cumplidos o más, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

8. Ajuste a los factores de expansión

Los factores de expansión elaborados conforme al procedimiento antes descrito se ajustaron en base a los siguientes conceptos:

8.1 Ajuste por no respuesta

El ajuste por no respuesta, se realizó a nivel de las viviendas, los hogares y las personas seleccionadas a nivel UPM, en cada uno de los dominios mediante las siguientes expresiones:

8.1.1 Ajuste por no respuesta a nivel vivienda

El ajuste por no respuesta atribuida al informante se realizó a nivel estrato, en cada uno de los dominios de estudio, mediante la siguiente expresión:

$$F'_{ehij} = F_{ehij} \frac{\sum_{i \in h} \sum_{j \in i} F_{ehij}}{\sum_{i \in h} \sum_{j \in i} F_{ehij} I_{ehij}}$$

Donde:

F'_{ehij} = factor de expansión corregido por no respuesta de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

F_{ehij} = factor de expansión de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

I_{ehij} = función indicadora con valor 1 o 0; esta función toma el valor 1 si la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad proporcionó respuesta completa, y toma el valor de cero 0 en caso contrario.

8.1.2 Ajuste por no respuesta a nivel hogar

El ajuste por no respuesta atribuida al informante se realiza a nivel estrato, en cada uno de los dominios de estudio, mediante la siguiente expresión:

$$F''_{ehijk} = F'_{ehijk} \frac{\sum_{i \in h} \sum_{j \in i} \sum_{k \in j} F'_{ehijk}}{\sum_{i \in h} \sum_{j \in i} \sum_{k \in j} F'_{ehijk} I_{ehijk}}$$

Donde:

F''_{ehijk} = factor de expansión corregido por no respuesta del k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

F'_{ehijk} = factor de expansión del k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

I_{ehijk} = función indicadora con valor 1 o 0; esta función toma el valor 1 si en el k-ésimo hogar, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad proporcionó respuesta completa, y toma el valor de cero 0 en caso contrario.

8.1.3 Ajuste por no respuesta a nivel persona

El ajuste por no respuesta atribuida al informante se realizó a nivel estrato, en cada uno de los dominios de estudio, mediante la siguiente expresión:

$$F'''_{ehijk\lambda} = F''_{ehijk\lambda} \frac{\sum_{i \in h} \sum_{j \in i} \sum_{k \in j} \sum_{\lambda \in k} F''_{ehijk\lambda}}{\sum_{i \in h} \sum_{j \in i} \sum_{k \in j} \sum_{\lambda \in k} F''_{ehijk\lambda} I_{ehijk\lambda}}$$

Donde:

$F'''_{ehijk\lambda}$ = factor de expansión corregido por no respuesta de la λ -ésima persona seleccionada, del k -ésimo hogar, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad.

$F''_{ehijk\lambda}$ = factor de expansión para la λ -ésima persona seleccionada, del k -ésimo hogar, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad.

$I_{ehijk\lambda}$ = función indicadora con valor 1 o 0; esta función toma el valor 1 si la λ -ésima persona, en el k -ésimo hogar, en la j -ésima vivienda, en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad proporcionó respuesta completa, y toma el valor de cero 0 en caso contrario.

8.2 Ajuste por proyección

Los factores de expansión ajustados por la no respuesta se corrigieron, a fin de asegurar que en cada dominio de interés de la encuesta se obtenga la población total determinada, por la proyección de población generada por la CONAPO referida al punto medio del levantamiento, mediante la siguiente expresión:

$$F''_D = F^*_D \frac{PROy_D}{PEXP_D}$$

Donde:

F''_D = factor de expansión corregido por proyección en el dominio D.

F^*_D = factor de expansión corregido por no respuesta en el dominio D.

$PROy_D$ = población en el dominio D, según proyección.

$PEXP_D$ = población total a la que expande la encuesta en el dominio D.

D = nivel de desagregación en que se efectúa el ajuste.

9. Estimadores

El estimador del total de la característica X es:

$$\hat{X} = \sum_e \sum_h \sum_i F_{ehij}^{UA} \left(\sum_s \sum_\lambda X_{ehis\lambda}^{UA} \right) + \sum_e \sum_h \sum_i F_{ehij}^{CU} \left(\sum_s \sum_\lambda X_{ehis\lambda}^{CU} \right) + \sum_e \sum_h \sum_i F_{ehij}^R \left(\sum_s \sum_\lambda X_{ehis\lambda}^R \right)$$

Donde:

F_{ehij}^{UA} = factor de expansión final de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad en el dominio urbano alto.

$X_{ehis\lambda}^{UA}$ = valor observado de la característica de interés X en la λ -ésima persona, en la s-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, en el dominio urbano alto.

F_{ehij}^{CU} = factor de expansión final de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad, en el dominio complemento urbano.

$X_{ehis\lambda}^{CU}$ = valor observado de la característica de interés X en la λ -ésima persona, en la s-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, en el dominio complemento urbano.

F_{ehij}^R = factor de expansión final de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad del dominio rural.

$X_{ehis\lambda}^R$ = valor observado de la característica de interés en la λ -ésima persona, en la s-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad, del dominio rural.

Para la estimación de proporciones, tasas y promedios se utiliza el estimador de razón:

$$\hat{R} = \frac{\hat{X}}{\hat{Y}}$$

Donde, \hat{Y} se define en forma análoga a \hat{X} .

10. Estimaciones de errores de muestreo

Para la evaluación de los errores de muestreo de las principales estimaciones estatales y nacionales se usó el método de Conglomerados Últimos,³ basado en que la mayor contribución a la varianza de un estimador, en un diseño polietápico es la que se presenta entre las unidades primarias de muestreo (UPM). El término “Conglomerados Últimos” se utiliza para denotar el total de unidades en muestra de una unidad primaria de muestreo.

Para obtener las precisiones de los estimadores de razón, conjuntamente al método de Conglomerados Últimos se aplicó el método de series de Taylor, obteniéndose la siguiente fórmula para estimar la precisión de \hat{R} :

$$\hat{V}(\hat{R}) = \frac{1}{\hat{Y}^2} \sum_h \left\{ \sum_e \frac{k_{eh}}{k_{eh} - 1} \sum_i^{k_{eh}} \left[\left(\hat{X}_{ehi} - \frac{1}{k_{eh}} \hat{X}_{eh} \right) - \hat{R} \left(\hat{Y}_{ehi} - \frac{1}{k_{eh}} \hat{Y}_{eh} \right) \right]^2 \right\}$$

Donde:

\hat{X}_{ehi} = total ponderado de la variable de estudio X en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

\hat{X}_{eh} = total ponderado de la variable de estudio X en el h-ésimo estrato, en la e-ésima entidad.

\hat{Y}^2 = el cuadrado del estimador del total de la característica Y.

k_{eh} = número de UPM en el h-ésimo estrato, en la e-ésima entidad.

L_e = número de estratos, en la e-ésima entidad.

³ Véase Hansen, M. H. Horwitz, W.N. y Madow, W.G., Sample Survey Methods and Theory, (1953) Vol. 1 pág. 242.

Estas definiciones son análogas para la variable de estudio Y.

La estimación de la varianza del estimador de un total, se calcula con la siguiente expresión:

$$\hat{V}(\hat{X}_{NAL}) = \sum_{e=1}^{32} \sum_{h=1}^{L_e} \frac{k_{eh}}{k_{eh} - 1} \sum_{i=1}^{k_{eh}} \left(\hat{X}_{ehi} - \frac{1}{k_{eh}} \hat{X}_{eh} \right)^2$$

Las estimaciones del error estándar (EE), efecto de diseño (DEFF) y coeficiente de variación (CV) se calculan mediante las siguientes expresiones:

$$EE = \sqrt{\hat{V}(\hat{\theta})} \qquad DEFF = \frac{\hat{V}(\hat{\theta})}{\hat{V}(\hat{\theta})_{MAS}} \qquad CV = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}}$$

Donde:

- $\hat{\theta}$ = estimador del parámetro poblacional θ .
- $\hat{V}(\hat{\theta})_{MAS}$ = estimador de la varianza bajo muestreo aleatorio simple.
- $\hat{V}(\hat{\theta})$ = estimador de la varianza bajo el diseño de muestreo descrito en este documento.

Finalmente, el intervalo de confianza $I_{1-\alpha}$ al 100(1- α)%, se calcula de la siguiente forma:

$$I_{1-\alpha} = \left(\hat{\theta} - z_{1-\alpha/2} \sqrt{\hat{V}(\hat{\theta})}, \hat{\theta} + z_{1-\alpha/2} \sqrt{\hat{V}(\hat{\theta})} \right)$$

Donde α es el nivel de significancia.

Anexo

A. Indicadores empleados en la estratificación de la muestra maestra

Mnemónico	Descripción
Población	
PPSSNOSP	Que tiene derecho a recibir servicios médicos en alguna institución de salud pública o privada excepto seguro popular.
PPDER_SS	Derechohabiente a servicios de salud.
PDP3A14A	De 3 a 14 años de edad que asiste a la escuela.
PDP15A24A	De 15 a 24 años de edad que asiste a la escuela.
PDP8A14ALF	De 8 a 14 años de edad que saben leer o escribir.
PDP15YM_SE	De 15 años o más de edad que aprobaron algún grado de escolaridad diferente al nivel preescolar.
PP15PRI_CO	De 15 años o más de edad que tienen como máxima escolaridad 6 grados aprobados en primaria.
PP15SEC_CO	De 15 años o más de edad que tienen como máxima escolaridad 3 grados aprobados en secundaria.
PGDO_ESC	Grado promedio de escolaridad.
PPEA	De 12 años y más que trabajaron; tenían trabajo pero no trabajaron o; buscaron trabajo en la semana de referencia.
PPEA_F	Femenina de 12 años y más que trabajaron; tenían trabajo pero no trabajaron o; buscaron trabajo en la semana de referencia.
PTASAOcupa	Tasa de ocupación.
TOCU12A17	No ocupada de 12 a 17 años de edad entre la población de 12 a 17 años de edad.
PPOMAYED	Ocupada de 18 y más años de edad entre la población ocupada.
Viviendas particulares habitadas	
PVIVSINH	Que no tienen hacinamiento.
PVPH_PISDT	Que tienen piso de cemento o firme, madera, mosaico u otro material.
PVPH2YMASD	Que usan para dormir entre 2 y 25 cuartos.
PVPH_2MASC	Que tienen más de un cuarto.
PVPH3YMASD	Que tienen entre 3 y 25 cuartos.
PVPH_C_ELE	Que disponen de luz eléctrica.
PVPHAGUADV	Que tienen disponibilidad de agua entubada dentro de la vivienda, o fuera de la vivienda pero dentro del terreno.
PVPH_EXCSA	Que tienen excusado, retrete, sanitario, letrina u hoyo negro.
PVPHDRENAJ	Que tienen drenaje conectado a la red pública, fosa séptica, barranca, grieta, río, lago o mar.
PVDRENERED	Que disponen de drenaje conectado a la red pública.
PVEXCAGU	Que disponen de excusado con descarga directa de agua.
PVPH_CSERV	Que disponen de luz eléctrica, agua entubada dentro o fuera de la vivienda, pero dentro del terreno, así como drenaje.
PSIN_HASIN	Que no se encuentran en situación de hacinamiento a nivel manzana.
Viviendas particulares habitadas que disponen de:	
PVPH_TV	Televisor.
PVPH_AUTOM	Automóvil o camioneta.
PVPH_CEL	Teléfono celular.
PVCELFIJ	Teléfono celular y teléfono fijo.
PV4ELEC	Radio, televisor, refrigerador y lavadora.
PVRADTEL	Radio y televisor.
PVPHCBIEN	Todos los bienes.

B. Distribución de viviendas seleccionadas por entidad y dominio de la muestra para la Encuesta Nacional de Cultura Cívica (ENCUCI) 2020

Regiones	Entidad		Distribución de viviendas seleccionadas		
	Clave	Nombre	Urbano alto & Complemento Urbano	Rural	Total
1	01	Aguascalientes	400	120	520
1	02	Baja California	440	101	541
1	03	Baja California Sur	420	96	516
3	04	Campeche	640	202	842
2	05	Coahuila de Zaragoza	760	278	1 038
5	06	Colima	620	79	699
3	07	Chiapas	440	398	838
1	08	Chihuahua	420	98	518
6	09	Ciudad de México	960	100	1 060
1	10	Durango	380	141	521
5	11	Guanajuato	480	221	701
4	12	Guerrero	500	335	835
6	13	Hidalgo	580	456	1 036
5	14	Jalisco	600	100	700
6	15	México	840	201	1 041
5	16	Michoacán de Ocampo	480	219	699
6	17	Morelos	800	243	1 043
5	18	Nayarit	480	219	699
2	19	Nuevo León	800	243	1 043
4	20	Oaxaca	400	437	837
4	21	Puebla	600	243	843
5	22	Querétaro	500	199	699
3	23	Quintana Roo	740	99	839
2	24	San Luis Potosí	680	366	1 046
1	25	Sinaloa	380	136	516
1	26	Sonora	420	99	519
3	27	Tabasco	480	361	841
2	28	Tamaulipas	920	119	1 039
4	29	Tlaxcala	680	159	839
4	30	Veracruz de Ignacio de la Llave	520	322	842
3	31	Yucatán	700	142	842
1	32	Zacatecas	320	201	521
		Nacional	18 380	6 733	25 113