

ENSAFI

Encuesta Nacional
sobre Salud Financiera

2023

Diseño muestral



 INEGI

Instituto Nacional de Estadística y Geografía

Encuesta Nacional sobre Salud Financiera 2023

ENSAFI

Diseño muestral



Catalogación en la fuente INEGI:

640.4201 Encuesta Nacional sobre Salud Financiera.
 Encuesta Nacional sobre Salud Financiera 2023 : ENSAFI : diseño muestral /
 Instituto Nacional de Estadística y Geografía.-- México : INEGI, c2024.

vii, 18 p.

1. Hogar - Finanzas - Encuestas - México. 2. Economía doméstica - Estadísticas
- México. 3. Hogar - Encuestas - Metodología. I. Instituto Nacional de Estadística y
Geografía (México).

Conociendo México

800 111 4634
www.inegi.org.mx
atencion.usuarios@inegi.org.mx



Registro en trámite

2024, **Instituto Nacional de Estadística y Geografía**

Edificio Sede
Avenida Héroe de Nacozari Sur 2301
Fraccionamiento Jardines del Parque, 20276, Aguascalientes,
Aguascalientes, Aguascalientes, entre la calle INEGI,
Avenida del Lago y Avenida Paseo de las Garzas.

Presentación

El **Instituto Nacional de Estadística y Geografía (INEGI)**, cumpliendo con las atribuciones que le confiere la Ley del Sistema Nacional de Información Estadística y Geográfica, desarrolla y aplica estrategias orientadas a consolidar el Sistema Nacional de Información Estadística y Geográfica (SNIEG), entre las cuales se encuentra la documentación de los diferentes proyectos que lleva a cabo.

En tal contexto, presenta el **Diseño muestral de la Encuesta Nacional sobre Salud Financiera (ENSAFI) 2023**, con el propósito de ofrecer un panorama general sobre la metodología empleada para el diseño y ejecución del proyecto.

De esta forma, el **INEGI** ofrece a las personas usuarias un documento en el que se muestran, en forma detallada, los diferentes aspectos del diseño estadístico del proyecto.

Asimismo, a través de este tipo de documentos es posible contribuir a transparentar el proceso de generación de información estadística y disponer de un testimonio que permita fortalecer el diseño y la operación de futuros proyectos en la materia.

Índice

Introducción	VII
1. Bases metodológicas	1
1.1 Objetivos	1
1.2 Referencias metodológicas	1
2. Diseño estadístico	3
2.1 Marco de la encuesta	3
2.2 Formación de las UPM	3
2.3 Estratificación	4
2.4 Selección de las UPM de la Muestra Maestra	4
2.5 Tamaño de la muestra	5
2.6 Afijación de la muestra	5
2.7 Selección de la muestra	5
2.7.1 En urbano alto	6
2.7.2 En complemento urbano	7
2.7.3 En rural	8
2.8 Ajuste a los factores de expansión	9
2.8.1 Ajuste por no respuesta	9
2.8.2 Ajuste por estimación de población	10
2.9 Estimadores	10
2.10 Estimación de errores de muestreo	11
2.11 Homologación de la semaforización para los umbrales de indicadores de precisión estadística	12
Anexos	15
A. Indicadores empleados en la estratificación de la Muestra Maestra	17
B. Distribución de la muestra en viviendas seleccionadas por entidad, según tamaño de localidad	18
C. Distribución de la muestra en viviendas seleccionadas por estrato, según tamaño de localidad	18

Introducción

La **Encuesta Nacional sobre Salud Financiera (ENSAFI) 2023** es un proyecto que desarrolló el **Instituto Nacional de Estadística y Geografía (INEGI)** en conjunto con la **Comisión Nacional para la Protección y Defensa de los Usuarios de Servicios Financieros (CONDUSEF)** con la finalidad de generar información estadística a nivel nacional y por entidad federativa sobre los aspectos que pueden definir la salud financiera de la población de 18 años y más que reside en México. Asimismo, permitirá conocer el tipo de preocupaciones más comunes de las personas adultas y que pudieran relacionarse con el estrés financiero.

El diseño muestral que se presenta a continuación está organizado en dos capítulos; en el primero, se abordan los objetivos y las bases metodológicas, mientras que el segundo está dedicado al diseño estadístico.

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

1. Bases metodológicas

En este apartado se presentan los objetivos del proyecto y las referencias metodológicas que lo distinguen.

1.1 Objetivos

General

Generar información estadística a nivel nacional y por entidad federativa sobre los aspectos que definen la salud financiera de la población de 18 años y más en el país, asimismo conocer sobre las preocupaciones más comunes y el estrés que la población pueda sentir a consecuencia de su situación financiera, e impactos en su vida.

Específicos

1. Generar información estadística sobre la percepción de la población de 18 años y más respecto de su seguridad financiera para cumplir con los compromisos financieros que tiene a corto plazo.
2. Generar información sobre la resiliencia financiera que tiene la población de 18 y más años para hacer frente a acontecimientos inesperados.
3. Recabar información sobre la confianza que tiene la población de 18 y más años en sus propias finanzas presentes y futuras, esto es, sobre su control financiero.
4. Obtener información sobre la libertad financiera que tiene la población de 18 años y más para cumplir con los objetivos y deseos de largo plazo en materia financiera.
5. Generar información estadística sobre las principales metas u objetivos financieros de la población de 18 años y las estrategias que considere para alcanzarlas.
6. Generar información estadística que permita conocer las principales preocupaciones financieras en la población de 18 años y más y que le puede generar ansiedad o complicaciones en aspectos de su vida como relaciones sociales, laborales, personales o familiares, y tener posibles impactos en su salud.
7. Obtener información estadística sobre las características sociodemográficas, económicas y de acceso a servicios financieros de la población de 18 años y más, y de su relación con la salud y el estrés financiero, que permita, entre otras cosas, encontrar los principales determinantes de la salud financiera en el país, identificar brechas de género y para otros grupos específicos de población.

1.2 Referencias metodológicas

Unidades de observación. La vivienda seleccionada, los hogares de la vivienda seleccionada y la persona elegida de 18 años y más.

Periodo de levantamiento. Del 25 de septiembre al 17 de noviembre de 2023.

Cobertura geográfica. La muestra se encontró distribuida en todo el país, tanto en localidades urbanas como rurales, de tal manera que generó información a nivel nacional.

Método de recolección. Se procedió a recolectar la información a través de una entrevista directa con la persona informante adecuada y las personas elegidas de 18 o más años integrantes del hogar, a través de un cuestionario electrónico en equipo Mitsui y autollenado para la captación de ingresos en el hogar y por trabajo. Solamente en caso necesario se utilizó cuestionario impreso para, posteriormente, capturar la información.

Periodo de referencia. La mayor parte de las preguntas se refirieron a la situación al momento de la entrevista; semana anterior a la entrevista respecto a la condición y verificación de actividad, y posición en el trabajo; mes anterior a la entrevista en transferencias y otros ingresos diferentes al trabajo y condición de deuda. Último mes

respecto a la privación económica, suficiencia de ingresos, medios para afrontar gastos regulares, y estrategias económicas personales. De octubre de 2022 a la fecha de la entrevista respecto a ahorro en el periodo, equivalencia del monto ahorrado, tipo de crédito informal y formal, mensualidades vencidas y emociones derivadas por el estrés.

Informantes:

- *Persona informante del hogar.* Jefa o jefe del hogar, o una persona de 18 años o más, residente del hogar y que conoce la información de las personas que lo integran y sus características sociodemográficas.
- *Persona elegida.* Para las entrevistas que se levanten en papel, es la persona integrante del hogar de 18 años y más, cuya fecha de cumpleaños es la inmediata posterior a la fecha de la entrevista. En el caso de las entrevistas que se realizan con el dispositivo electrónico, la persona elegida es seleccionada aleatoriamente por el mismo dispositivo electrónico.

2. Diseño estadístico

Comprende el conjunto de actividades relacionadas con la selección de la muestra, el tamaño suficiente que permita realizar estimaciones para la población objeto de estudio, el marco de muestreo, y la construcción y evaluación de las estimaciones planteadas para la expansión de la información a partir de los datos obtenidos en campo.

2.1 Marco de la encuesta

El diseño de la muestra para la ENSAFI 2023 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 en varias etapas, estratificado y por conglomerados, donde la unidad última de selección es la persona.

Para la selección de la muestra se utilizó el Marco Nacional de viviendas 2012 del INEGI, construido a partir de la información cartográfica y demográfica que se obtuvo del Censo de Población y Vivienda 2010. A partir de este Marco se seleccionan las Unidades Primarias de Muestreo (UPM), al interior de cada UPM se seleccionan las viviendas de la muestra y al interior de cada vivienda se selecciona una persona de 18 y más años de edad. El marco se construyó de la siguiente manera:

2.2 Formación de las UPM

Inicialmente, se conforma el conjunto de UPM que cubren el territorio nacional.

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

a) 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 AGEB¹
- 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

b) 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

¹ Área Geoestadística Básica.

c) 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 quedó conformada por un total de 240 912 UPM a nivel nacional.

2.3 Estratificación

Una vez construido el conjunto de UPM, se agruparon aquellas con características similares de acuerdo con los siguientes criterios:

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

Estratificación por tamaño de localidad según ámbito geográfico

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 indicadores² 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 obtuvo un total de 683 estratos en todo el territorio nacional.

2.4 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 = la 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.

² La descripción de estos indicadores se presenta en el anexo A.

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

2.5 Tamaño de la muestra

La expresión para calcular el tamaño de muestra fue:

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

Donde:

- n = tamaño de la muestra.
p = estimación de la proporción de interés.
q = 1-p.
r = error relativo máximo aceptable.
z = valor asentado en las tablas estadísticas de la distribución normal estándar para 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 1.77, un error relativo máximo esperado del 15 %, una tasa de no respuesta máxima esperada del 15 %, para una proporción mínima de 1.1 % (de los fenómenos objeto de estudio) se obtuvo un tamaño de muestra mínimo de 22 513 viviendas, mismo que se ajustó a 22 982 viviendas.

2.6 Distribución de la muestra

La distribución de la muestra se realizó dentro de cada entidad federativa, por estrato de manera proporcional, para lo cual se empleó la siguiente expresión:

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

Donde:

- n_{eh} = número de viviendas en muestra en el h-ésimo estrato, en la e-ésima entidad.
 n_e = número total de viviendas en muestra en la e-ésima entidad.
 N_{eh} = número total de viviendas en el h-ésimo estrato, en la e-ésima entidad.
 N_e = número total de viviendas en la e-ésima entidad.

En el anexo B se presenta la distribución de la muestra de viviendas por entidad según tamaño de localidad, mientras que el C contiene la distribución de la muestra de viviendas según tamaño de localidad por estrato.

2.7 Selección de la muestra

Se realizó de manera independiente por entidad, tamaño de localidad y estrato; el procedimiento de selección varió de acuerdo con el dominio.

2.7.1 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 eligieron k_{eh}^* UPM con igual probabilidad.
2. En cada UPM seleccionada, se eligieron cinco viviendas con igual probabilidad.
3. En cada vivienda seleccionada se eligió a una persona de 18 años y más de edad 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 ENSAFI 2023 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:

$$\begin{aligned} P_{ehij} &= P_{1ehi} * P_{2ehi} * P_{3ehij} * P_{4ehij} \\ &= \frac{k_{eh} m_{ehi}}{m_{eh}} \cdot \frac{k_{eh}^*}{k_{eh}} \cdot \frac{5}{m_{ehi}^*} \cdot \frac{1}{Q_{ehij}} = \frac{5 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^* Q_{ehij}} \end{aligned}$$

Su factor de expansión³ está dado por

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

³ El factor de expansión se define como el inverso de la probabilidad de selección. En la *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, el factor de expansión se nombra “ponderador”.

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 cifras del 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 ENSAFI 2023 con igual probabilidad, en el h-ésimo estrato, en la e-ésima entidad.
- Q_{ehij} = número de personas de 18 años y más de edad en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

2.7.2 En complemento urbano

1. De las k_{eh} UPM que integran la Muestra Maestra se seleccionaron k_{eh}^* UPM con igual probabilidad.
2. En cada UPM seleccionada, se eligieron 20 viviendas con igual probabilidad.
3. En cada vivienda seleccionada se eligió a una persona de 18 años y más de edad 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 ENSAFI 2023 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:

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

- 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:

$$\begin{aligned} P_{ehij} &= P_{1ehi} * P_{2ehi} * P_{3ehij} * P_{4ehij} \\ &= \frac{k_{eh} m_{ehi}}{m_{eh}} \cdot \frac{k_{eh}^*}{k_{eh}} \cdot \frac{20}{m_{ehi}^*} \cdot \frac{1}{Q_{ehij}} = \frac{20 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^* Q_{ehij}} \end{aligned}$$

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 cifras del 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 ENSAFI 2023 con igual probabilidad, en el h–ésimo estrato, en la e–ésima entidad.
- Q_{ehij} = número de personas de 18 años y más de edad en la j–ésima vivienda, en la i–ésima UPM, en el h–ésimo estrato, en la e–ésima entidad.

2.7.3 En rural

1. De las k_{eh} UPM, se seleccionaron k_{eh}^* UPM con igual probabilidad.
2. En cada UPM seleccionada, se eligieron cuatro segmentos de cinco viviendas aproximadamente, con igual probabilidad.
3. En cada vivienda seleccionada se eligió a una persona de 18 años y más de edad con igual probabilidad.

La probabilidad de selección de 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 ENSAFI 2023 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:

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

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

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 cifras del 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 ENSAFI 2023 con igual probabilidad, en el h -ésimo estrato, en la e -ésima entidad.
- Q_{ehij} = número de personas de 18 años y más de edad en la j -ésima vivienda, en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad.

2.8 Ajuste a los factores de expansión

Los factores de expansión elaborados conforme al procedimiento antes descrito, se ajustaron con base en los siguientes conceptos.

2.8.1 Ajuste por no respuesta

El ajuste por no respuesta se realizó a nivel de las viviendas y las personas seleccionadas a nivel UPM, en cada uno de los dominios.

A NIVEL VIVIENDA

El ajuste por no respuesta en la vivienda 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 cero; esta función toma el valor 1 si la j -ésima vivienda seleccionada, en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad tiene respuesta completa, y toma el valor de cero en caso contrario.

A NIVEL PERSONA

El ajuste por no respuesta atribuible al informante se realizó 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 de la k-ésima persona seleccionada, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.
- F'_{ehijk} = factor de expansión para la k-ésima persona seleccionada, 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 cero; esta función toma el valor 1 si la k-ésima persona seleccionada, en la j-ésima vivienda seleccionada, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad tiene respuesta completa, y toma el valor de cero en caso contrario.

2.8.2 Ajuste por estimación de población

Los factores ajustados por 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 estimación de población referida al punto medio del levantamiento, mediante la siguiente expresión:

$$F''_D = F'_D \frac{PESTy_D}{PEXP_D}$$

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.
- $PESTy_D$ = población en el dominio D, según estimación de población.
- $PEXP_D$ = población total a la que expande la encuesta en el dominio D.
- D = nivel de desagregación de la subpoblación en que se hace el ajuste por proyección.

2.9 Estimadores

El estimador del total de la característica X es:

$$X = \sum_e \sum_h \sum_i \left(\sum_s \sum_{\ell} F_{ehis}^{15+} \sum_{\ell} X_{ehis\ell}^{15+} \right) + \sum_e \sum_h \sum_i \left(\sum_s \sum_{\ell} F_{ehis}^{15-} \sum_{\ell} X_{ehis\ell}^{15-} \right)$$

Donde:

- F_{ehi}^{15+} = factor de expansión final de la s–ésima vivienda, de la i–ésima UPM, del h–ésimo estrato, de la e–ésima entidad, en las localidades de 15 000 y más habitantes.
- $X_{ehis\ell}^{15+}$ = valor observado de la característica de interés X en la ℓ –ésima persona, en la s–ésima vivienda, en la i–ésima UPM, en el h–ésimo estrato, en la e–ésima entidad, en las localidades de 15 000 y más habitantes.
- F_{ehi}^{15-} = factor de expansión final de la s–ésima vivienda, de la i–ésima UPM, del h–ésimo estrato, de la e–ésima entidad, en las localidades menores de 15 000 habitantes.
- $X_{ehis\ell}^{15-}$ = valor observado de la característica de interés X en la ℓ –ésima persona, en la s–ésima vivienda, en la i–ésima UPM, en el h–ésimo estrato, en la e–ésima entidad, en las localidades menores de 15 000 habitantes.

Para la estimación de proporciones, tasas y promedios se utiliza el estimador de razón: $\hat{R} = \frac{\hat{X}}{\hat{Y}}$

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

2.10 Estimación de errores de muestreo

Para la evaluación de los errores de muestreo de las principales estimaciones, se usó el método de *Conglomerados Últimos*,⁴ basado en que la mayor contribución a la varianza de un estimador, en un diseño multietápico, es la que se presenta entre las UPM. El término *Conglomerados Últimos* se utiliza para denotar el total de unidades en muestra de una UPM.

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} :

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

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

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:

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

⁴ Véase Hansen, M. H., Horwitz, W. N. y Madow, W.G., *Sample Survey Methods and Theory*, vol 1, 1953, pág. 242.

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

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

Donde:

$\hat{\theta}$ = estimador del parámetro poblacional θ .

$\hat{V}(\hat{\theta})_{MAS}$ = estimador de la varianza, bajo un 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 $(1 - \alpha)\%$, se construye 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.

2.11 Homologación de la semaforización para los umbrales de indicadores de precisión estadística⁵

Para facilitar la interpretación de las precisiones estadísticas de la información pública en tabulados, el Comité de Aseguramiento de la Calidad, en la cuarta sesión celebrada el 1.^o de noviembre de 2018, aprobó los siguientes umbrales y especificaciones para la publicación de tabulados de los coeficientes de variación, así como la semaforización de estos.

Umbrales aprobados para la cobertura del coeficiente de variación

Interpretación	Semaforización	Nivel de precisión
Alta	Blanco	(0 %, 15 %)
Moderada	Amarillo	[15 %, 30 %)
Baja	Naranja oscuro	>=30 %

Umbrales aprobados para el reporte de la precisión de acuerdo con el coeficiente de variación en los tabulados de resultados de los proyectos con muestreo probabilístico (acuerdo CAC-007/01/2018).

A partir del segundo trimestre de 2018, se publican los siguientes indicadores de precisión estadística en la presentación de resultados en tabulados de todas las encuestas con muestreo probabilístico del INEGI: error estándar, intervalo de confianza y coeficiente de variación. Adicionalmente, se estandariza la coloración en los tabulados para señalar el nivel de precisión de las estimaciones con base en el CV. A continuación, se presenta el código RGB (en inglés de *red, green, blue*; en español: rojo, verde y azul) de los colores utilizados en la semaforización.

⁵ La fuente de esta información está basada en el documento del Comité de Aseguramiento de la Calidad, depositado en el siguiente sitio http://intranet.inegi.org.mx/calidad/wp-content/uploads/2017/02/Homologacion_de_umbrales.pdf.

Parámetros RGB para la semaforización del coeficiente de variación

Código	Interpretación		
	Alta	Moderada	Baja
Rojo	255	255	255
Verde	255	234	84
Azul	255	0	0

El siguiente texto explicativo aparece en cada uno de los tabulados publicados de encuestas por muestreo probabilístico:

Las estimaciones que aparecen en este cuadro están coloreadas, de acuerdo con su nivel de precisión, en *Alta*, *Moderada* y *Baja*, tomando como referencia el CV (%). Una precisión baja requiere un uso cauteloso de la estimación en el que se analicen las causas de la alta variabilidad y se consideren otros indicadores de precisión y confiabilidad, como el intervalo de confianza.

Nivel de precisión de las estimaciones:

Alta, CV en el rango de (0, 15)

Moderada, CV en el rango de [15, 30]

Baja, CV de 30 % en adelante

Anexos

A. Indicadores empleados en la estratificación de la Muestra Maestra

Mnemónico	Descripción
Proporción de 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 o más años de edad que aprobaron algún grado de escolaridad diferente al nivel preescolar.
PP15PRI_CO	De 15 o más años de edad que tienen como máxima escolaridad 6 grados aprobados en primaria.
PP15SEC_CO	De 15 o más años de edad que tienen como máxima escolaridad 3 grados aprobados en secundaria.
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.
TOCU12A17	No ocupada de 12 a 17 años entre la población total de este rango.
PPOMAYED	Ocupada de 18 años y más entre la población total ocupada.
PGDO_ESC	Grado promedio de escolaridad.
PTASAOCUPA	Tasa de ocupación.
Proporción de 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.
PVPH3YMASC	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 ella, 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.
PVDRERED	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.
Proporción de 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 la muestra en viviendas seleccionadas por entidad, según tamaño de localidad

Entidad federativa	Muestra	Tamaño de localidad	
		15 000 y más habitantes	Menos de 15 000 habitantes
Nacional	22 982	14 065	8 917
Aguascalientes	715	550	165
Baja California	720	600	120
Baja California Sur	715	495	220
Campeche	714	410	304
Coahuila de Zaragoza	720	620	100
Colima	712	535	177
Chiapas	719	260	459
Chihuahua	719	580	139
Ciudad de México	719	695	24
Durango	719	420	299
Guanajuato	720	440	280
Guerrero	724	320	404
Hidalgo	718	205	513
Jalisco	720	540	180
México	721	530	191
Michoacán de Ocampo	717	335	382
Morelos	719	440	279
Nayarit	715	355	360
Nuevo León	720	660	60
Oaxaca	722	180	542
Puebla	721	365	356
Querétaro	718	375	343
Quintana Roo	714	590	124
San Luis Potosí	722	400	322
Sinaloa	718	460	258
Sonora	720	540	180
Tabasco	716	245	471
Tamaulipas	721	600	121
Tlaxcala	715	255	460
Veracruz de Ignacio de la Llave	716	300	416
Yucatán	720	470	250
Zacatecas	713	295	418

C. Distribución de la muestra en viviendas seleccionadas por estrato, según tamaño de localidad

Estrato	Muestra	Tamaño de localidad	
		15 000 y más habitantes	Menos de 15 000 habitantes
Total	22 982	14 065	8 917
Bajo	4 074	250	3 824
Medio bajo	11 873	6 940	4 933
Medio alto	4 895	4 780	115
Alto	2 140	2 095	45