

ENVIPE

Encuesta Nacional de
Victimización y Percepción
sobre Seguridad Pública
2021

Diseño muestral



 INEGI

Instituto Nacional de Estadística y Geografía

Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2021

ENVIPE

Diseño muestral



Obras complementarias publicadas por el INEGI sobre el tema:

Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2011-2014. ENVIPE. Manual del entrevistador
Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2011-2014. ENVIPE. Manual del jefe de entrevistadores
Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2011-2014. ENVIPE. Síntesis metodológica
Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2011-2021. ENVIPE. Marco conceptual
Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2011-2021. ENVIPE. Informe operativo
Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2015-2020. ENVIPE. Diseño muestral.

Catalogación en la fuente INEGI:

364.101 Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (2021).
Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2021
: ENVIPE : diseño muestral / Instituto Nacional de Estadística y Geografía.-- México :
INEGI, c2021.

vii, 20 p.

1. México - Seguridad pública - Metodología. 2. Delitos y delincuentes - Estadísticas - Metodología. 3. Estadística criminal - 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

DR © 2021, **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)**, 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 **Diseño Muestral de la Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE) 2021** 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
Diseño muestral ENVIPE 2021	1
1. Objetivo general	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. Distribución 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	8
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	10
8.1.3 Ajuste por no respuesta a nivel persona	10
8.2 Ajuste por estimación de población	10
8.3 Ajuste por los delitos	11
9. Estimadores	13
10. Estimaciones de errores de muestreo	14
11. Homologación de la semaforización para los umbrales de indicadores de precisión estadística	15

Anexo	17
A. Indicadores empleados en la estratificación de la muestra maestra	19
B. Distribución de viviendas seleccionadas por entidad y dominio de la muestra para la Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2021 (ENVIPE - 2021)	20

Introducción

La **Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE) 2021** es un proyecto estadístico destinado a enriquecer la oferta de información de interés nacional vinculada al Subsistema Nacional de Información de Gobierno, Seguridad Pública e Impartición de Justicia.

La encuesta se llevó a cabo con la finalidad de obtener información que permita estimar los niveles de victimización a nivel nacional y por entidad federativa, así como los hechos delictivos denunciados y no denunciados. De igual forma, se estudia la percepción que los habitantes tienen respecto al lugar donde viven, y su evaluación de las principales autoridades en materia de seguridad pública y procuración de justicia.

Con el propósito de enmarcar las principales características metodológicas del proyecto, se realiza el Diseño Muestral de la ENVIPE 2021, 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.

Diseño muestral ENVIPE 2021

1. Objetivo general

Obtener información con representatividad a nivel nacional y estatal que permita realizar estimaciones de la prevalencia delictiva que afectó a los hogares durante 2020, los niveles de incidencia delictiva y cifra negra. Asimismo, se busca obtener información sobre la percepción de la seguridad pública, así como el desempeño de las instituciones a cargo de la seguridad pública y la justicia. También se recaba información sobre las características del delito, el contexto de la victimización y el impacto económico y social del delito. Esto con el fin de proveer información al público en general y generar elementos para la toma de decisiones de política pública en estas materias.

2. Población objetivo

La encuesta está dirigida a la población de 18 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
- Entidad federativa
- Ciudad de México cuatro regiones (Norte, Sur, Oriente y Poniente)

4. Marco de la encuesta

El diseño de la muestra para la ENVIPE 2021 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 los hogares y las personas de 18 años cumplidos o más.

Para la selección de la muestra la ENVIPE 2021, 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.

² INEGI. Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2021. SNIEG. Información de Interés Nacional

4.2 Estratificación

Una vez construido el conjunto de UPM, se agrupan aquellas con características similares, de acuerdo con 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 ENVIPE 2021, está calculado para la proporción de personas o viviendas que sufrió algún delito, 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 2.078 observado en encuestas anteriores, así como una tasa de no respuesta máxima esperada del 15%, un error relativo máximo esperado del 7.625%, para una proporción de 1.10%, se obtuvo un tamaño de muestra a nivel nacional de 102 290 viviendas. Este tamaño de muestra se ajustó, por cuestiones operativas, a 102 297 viviendas.

6. Distribución de la muestra

La distribución 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:

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.

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

b = número de viviendas a seleccionar por UPM.

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 ENVIPE-2021, 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 ENVIPE-2021.
2. En cada UPM seleccionada, se seleccionaron cinco viviendas con igual probabilidad.
3. En cada vivienda seleccionada, se seleccionó una persona de 18 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 ENVIPE-2021 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:

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

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

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 ENVIPE-2021 con igual probabilidad de selección, en el h-ésimo estrato, en la e-ésima entidad.
- Q_{ehij} = número de personas de 18 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 ENVIPE-2021.
2. En cada UPM seleccionada, se seleccionaron 20 viviendas con igual probabilidad.
3. En cada vivienda seleccionada se seleccionó una persona de 18 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:

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

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

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 ENVIPE-2021 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:

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

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

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 ENVIPE-2021 con igual probabilidad de selección, en el h-ésimo estrato, en la e-ésima entidad.
- Q_{ehij} = número de personas de 18 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 ENVIPE-2021 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 18 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 ENVIPE-2021 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 ENVIPE-2021, en el h-ésimo estrato, en la e-ésima entidad.
 Q_{ehij} = número de personas de 18 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 repuesta 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:

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

Donde:

$F'''_{ehijk\ell}$ = factor de expansión corregido por no respuesta de la ℓ -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.

$F''_{ehijk\ell}$ = factor de expansión para la ℓ -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.

$I_{ehijk\ell}$ = función indicadora con valor 1 o 0; esta función toma el valor 1 si la ℓ -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 estimación de población

Los factores de expansión ajustados por la no respuesta se corrigieron, con el fin de asegurar que en cada dominio de interés de la encuesta se obtenga el monto de la estimación de población generada por el INEGI³ a partir del primer trimestre del año 2021, referida al punto medio del levantamiento, mediante la siguiente expresión:

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

³ El INEGI estará generando una estimación de población por entidad federativa, con base en la propia actualización del Marco de Muestreo de Viviendas del INEGI y de la información del CPV 2020, a través de una muestra aleatoria de viviendas que serán visitadas y contabilizadas su número de personas. La estimación de población se comenzó a implementar a partir del primer trimestre de 2021.

¹⁰ INEGI. Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2021. SNIEG. Información de Interés Nacional

Donde:

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

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

$PEST_D$ = población en el dominio D, según estimación de población INEGI.

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

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

8.3 Ajuste por los delitos

El factor de expansión de delitos se construyó a partir del factor de expansión corregido por estimación poblacional para las dos unidades de medición: personas elegidas y hogares. La construcción se diferencia a partir de si el delito afecta al hogar en general o solo a la persona elegida. A continuación, se presenta la descripción de los delitos:

- **Grupo A. Delitos del Hogar**

01. Robo total de vehículos: robo total de un automóvil, camioneta o camión propiedad de alguno de los integrantes del hogar.

02. Robo parcial de vehículo: robo de accesorios, refacciones o herramientas de un automóvil, camioneta o camión propiedad de alguno de los integrantes del hogar.

03. Vandalismo: pinta de barda o grafiti en su casa, rayones o daños intencionales en su vehículo u otro tipo de vandalismo.

04. Robo a casa habitación: alguien entró a su casa o departamento sin permiso mediante el uso de la fuerza o por engaños y robó o intentó robar algo.

- **Grupo B. Delitos de la Persona**

05. Robo a transeúnte: robo o asalto en la calle o en el transporte público e incluye robo en banco o cajero automático.

06. Otros robos: cualquier otro tipo de robo que no sea de vehículo, de casa habitación o a transeúnte.

07. Fraude bancario: alguien usó su chequera, número de tarjeta o cuenta bancaria sin su permiso para realizar cargos o para extraer dinero de sus cuentas o le dio dinero falso.

08. Fraude al consumidor: entregó dinero por un producto o un servicio que no recibió conforme a lo acordado.

09. Extorsión: amenazas, presiones o engaños para exigirle dinero o bienes o para que hiciera algo o dejara de hacerlo.

10. Amenazas: amenazas verbales de alguien plenamente identificado o por escrito hacia su persona diciendo que le va a causar un daño a usted, a su familia, a sus bienes o su trabajo.

11. Lesiones: alguien por actitud abusiva o por una discusión lo(a) golpeó generándole una lesión física (moretones, fracturas, cortadas, etcétera).

12. Secuestro y secuestro exprés: lo secuestraron para exigir dinero o bienes.

13. Delitos sexuales: alguien en contra de su voluntad lo(a) agredió mediante hostigamiento sexual, manoseo, exhibicionismo o intento de violación.

14. Violación sexual: fue obligado(a) mediante violencia física o amenaza por alguien conocido o desconocido a tener una actividad sexual no deseada.

15. Otros delitos que reporte el informante: todo tipo de delito que no corresponda a las opciones proporcionadas previamente.

La construcción del factor de expansión se realizó en los siguientes dos casos:

Caso 1:

Los factores de expansión del delito se ajustaron para las personas elegidas y para los hogares que fueron objeto de un mismo tipo de delito en **más de cinco ocasiones**. Del total de delitos reportados por el informante, se seleccionaron cinco de estos, ya que el 99.59% de las personas en promedio, han padecido a lo más cinco delitos en un año de referencia. Este ajuste se realizó considerando dos situaciones:

a) Los factores de expansión del delito se ajustaron a partir del factor de expansión del hogar, para los delitos tipificados con códigos 01, 02, 03 y 04 (ver Grupo A), que se consideran conceptualmente delitos del hogar, este ajuste se realizó mediante la siguiente expresión:

$$F_{ehij}^{Del} = F_{ehij}'' \frac{ND_{M5}}{5}$$

Donde:

F_{ehij}^{Del} = factor de expansión a nivel persona corregido por el número de delitos declarados por el informante cuando son más de cinco delitos de un mismo tipo, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

F_{ehij}'' = factor de expansión corregido por no respuesta y proyección poblacional a nivel hogar de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

ND_{M5} = número de delitos de un mismo tipo declarados por el informante cuando es mayor a cinco ocasiones.

b) Para los delitos tipificados con códigos: 05, 06, 07, 08, 09, 10, 11, 12, 13, 14 y 15 (ver Grupo B), que se consideran conceptualmente de victimización personal, los factores de expansión se ajustaron para las personas elegidas, mediante la siguiente expresión:

$$F_{ehij}^{Del} = F_{ehij}''' \frac{ND_{M5}}{5}$$

Donde:

F_{ehij}^{Del} = factor de expansión a nivel persona corregido por el número de delitos declarados por el informante cuando son más de cinco delitos de un mismo tipo, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

F_{ehij}''' = factor de expansión corregido por no respuesta a nivel persona elegida de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

ND_{M5} = número de delitos de un mismo tipo declarados por el informante cuando es mayor a cinco ocasiones.

Caso 2:

Los factores de expansión del delito para los hogares y para las personas elegidas, que fueron objeto de un mismo tipo de delito **a lo más en cinco ocasiones**, no requieren ajuste y simplemente se iguala al factor de expansión del hogar o al factor de expansión del elegido, según el código del delito, es decir:

- a) Para los delitos tipificados con códigos 01, 02, 03 y 04 (ver Grupo A), que se consideran conceptualmente delitos del hogar, el factor de expansión del delito es igual al factor de expansión del hogar, es decir:

$$F_{ehij}^{Del} = F_{ehij}''$$

F_{ehij}'' = factor de expansión corregido por no respuesta y proyección poblacional a nivel hogar de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

- b) Para los delitos tipificados con códigos: 05, 06, 07, 08, 09, 10, 11, 12, 13, 14 y 15 (ver Grupo B), que se consideran conceptualmente de delitos de victimización personal, el factor de expansión del delito es igual al factor de expansión personal, es decir:

$$F_{ehij}^{Del} = F_{ehij}'''$$

Donde:

F_{ehij}''' = factor de expansión corregido por no respuesta y proyección poblacional a nivel persona elegida de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

9. Estimadores

El estimador del total de la característica X es:

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

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.

$X_{ehis\ell}^{UA}$ = 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 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.

$X_{ehis\ell}^{CU}$ = 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 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.

$X_{ehis\ell}^R$ = valor observado de la característica de interés 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, 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,⁴ 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} :

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

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.

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:

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

⁴ Véase Hansen, M. H. Horwitz, W.N. y Madow, W.G., Sample Survey Methods and Theory, (3) Vol. 1 pág. 242.

¹⁴ INEGI. Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública 2021. SNIEG. Información de Interés Nacional

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

$$EE = \sqrt{\hat{V}(\hat{\theta})}$$

$$DEFF = \frac{\hat{V}(\hat{\theta})}{\hat{V}(\hat{\theta})_{MAS}}$$

$$CV = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}}$$

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.

11. Homologación de la semaforización para los umbrales de indicadores de precisión estadística

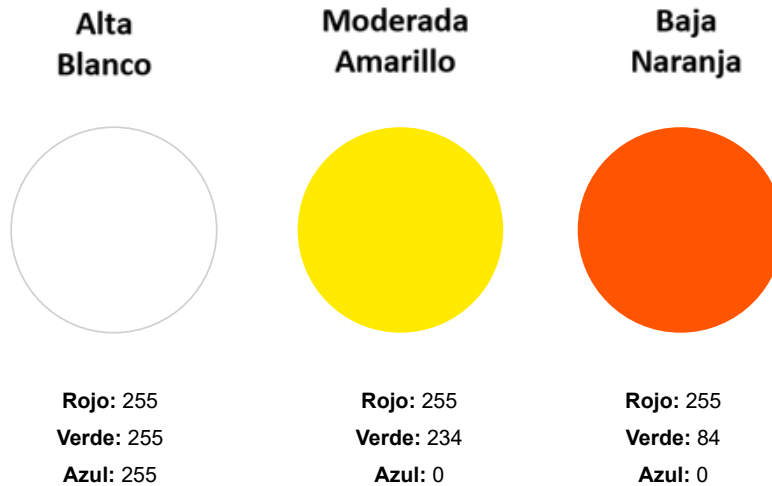
Para facilitar la interpretación de las precisiones estadísticas de la información pública en tabulados, el Comité de Aseguramiento de la Calidad, en la cuarta sesión celebrada el 1 de noviembre de 2018, aprobaron los siguientes umbrales y especificaciones para la publicación en los tabulados los coeficientes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

Umbrales aprobados para la cobertura del CV		
Interpretación	Semaforización	Viviendas/Hogares/Otras unidades diferentes a las económicas DGES/DGEGSPJ
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, coeficiente de variación (CV) e intervalo de confianza. Adicionalmente, se estandariza la coloración en los tabulados para indicar el nivel de precisión de las estimaciones con base en el CV. A continuación, se presenta el código RGB de los colores utilizados en la semaforización:

Parámetros RGB para la semaforización del coeficiente de variación.



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 coeficiente de variación 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

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 Victimización y Percepción sobre Seguridad Pública 2021 (ENVIPE - 2021)

Entidad		Distribución de viviendas seleccionadas			
Clave	Nombre	Urbano alto	Complemento urbano	Rural	Total
01	Aguascalientes	1 360	280	362	2 002
02	Baja California	2 280	500	220	3 000
03	Baja California Sur	1 040	1 560	401	3 001
04	Campeche	1 280	640	594	2 514
05	Coahuila de Zaragoza	2 340	420	242	3 002
06	Colima	2 140	980	383	3 503
07	Chiapas	900	1 240	1 914	4 054
08	Chihuahua	2 980	380	443	3 803
09	Ciudad de México	6 360		40	6 400
10	Durango	1 920	480	912	3 312
11	Guanajuato	1 100	660	750	2 510
12	Guerrero	940	880	1 183	3 003
13	Hidalgo	600	1 060	1 341	3 001
14	Jalisco	1 640	960	396	2 996
15	México	2 240	400	361	3 001
16	Michoacán de Ocampo	1 020	1 740	1 273	4 033
17	Morelos	1 620	920	468	3 008
18	Nayarit	680	720	605	2 005
19	Nuevo León	2 400	440	160	3 000
20	Oaxaca	500	960	1 570	3 030
21	Puebla	1 600	1 300	1 115	4 015
22	Querétaro	1 620	540	845	3 005
23	Quintana Roo	2 200	480	316	2 996
24	San Luis Potosí	880	420	716	2 016
25	Sinaloa	1 980	980	1 063	4 023
26	Sonora	1 240	500	261	2 001
27	Tabasco	660	1 400	1 435	3 495
28	Tamaulipas	3 400	600	508	4 508
29	Tlaxcala	1 440	600	462	2 502
30	Veracruz de Ignacio de la Llave	1 360	1 160	1 523	4 043
31	Yucatán	1 820	1 160	526	3 506
32	Zacatecas	520	680	809	2 009
Nacional		54 060	25 040	23 197	102 297