



Instituto Nacional de Estadística y Geografía

**Encuesta Nacional  
sobre Discriminación 2022**

**ENADIS**

**Diseño muestral**



**Obras complementarias publicadas por el INEGI sobre el tema:**  
Informe operativo de la Encuesta Nacional sobre Discriminación (ENADIS) 2017.

**Catalogación en la fuente INEGI:**

305.021 Encuesta Nacional sobre Discriminación (2022).  
Encuesta Nacional sobre Discriminación 2022 : ENADIS : diseño muestral /  
Instituto Nacional de Estadística y Geografía.-- México : INEGI, c2023.  
  
vii, 18 p.  
  
1. Discriminación - Estadísticas – Metodología. I. Instituto Nacional de Estadística  
y Geografía (México).

**Conociendo México**

**800 111 4634**

**[www.inegi.org.mx](http://www.inegi.org.mx)**

**[atencion.usuarios@inegi.org.mx](mailto:atencion.usuarios@inegi.org.mx)**

 **INEGI Informa**  **@INEGI\_INFORMA**

Registro en trámite

2023, **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, 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 Discriminación (ENADIS) 2022**, 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 los usuarios 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

---

<b>Introducción</b>	VII
<b>1. Bases metodológicas</b>	1
1.1 Objetivos	1
1.2 Referencias metodológicas	2
<b>2. Diseño estadístico</b>	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 Distribución de la muestra	5
2.7 Selección de la muestra	6
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 de las viviendas	9
2.8.2 Ajuste por estimación de población	9
2.8.3 Ajuste por no respuesta de los hogares	10
2.8.4 Factor de expansión a nivel persona	10
2.8.5 Factor de expansión por módulo	11
2.9 Estimadores	11
2.10 Estimación de errores de muestreo	12
2.11 Homologación de la semaforización para los umbrales de indicadores de precisión estadística	13
<b>Anexos</b>	
A. Indicadores empleados en la estratificación de los conglomerados del Marco Maestro de Muestreo	17
B. Distribución de la muestra de UPM y viviendas seleccionadas por entidad federativa	18

## Introducción

---

La **Encuesta Nacional sobre Discriminación (ENADIS) 2022** es un proyecto estadístico realizado por el **Instituto Nacional de Estadística y Geografía (INEGI)**, en colaboración con el **Consejo Nacional para Prevenir la Discriminación (CONAPRED)**, y la **Comisión Nacional de los Derechos Humanos (CNDH)**.

La encuesta tiene el propósito de generar información estadística actual que permita medir la magnitud, causas y expresiones de la discriminación, así como conocer su prevalencia en México.

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

En virtud de la necesidad de conocer las actitudes, prejuicios, percepciones y experiencias discriminatorias en México, así como los factores que se le relacionan, entre otros aspectos, se plantean los siguientes objetivos:

### General

Reconocer la magnitud de la discriminación y sus diversas manifestaciones en la vida cotidiana, profundizando en el conocimiento sobre quién o quiénes discriminan, en qué ámbitos de la vida se presenta este problema con mayor frecuencia y los factores socioculturales que se le relacionan. Asimismo, conocer las percepciones de la discriminación entre la población en general y desde los distintos grupos de la población, que por sus características son discriminados.

### Específicos

- Captar la opinión de las personas de 18 años o más, sobre la apertura que tienen en un trato igualitario hacia distintos grupos de la población que son objeto de discriminación por motivos raciales, étnicos, de salud, discapacidad, de sexualidad, entre otros. Asimismo, identificar las experiencias de discriminación a causa de sus creencias religiosas, edad, preferencias sexuales, o rasgos físicos y culturales.
- Conocer los distintos aspectos que evidencien la situación de discriminación y desigualdad social que enfrentan actualmente en México las personas por su condición: indígena, afrodescendiente, de discapacidad, de migrante o refugiada, de diversidad religiosa, de adulta mayor, de adolescente o joven, de niña o niño, de mujer, y por ser trabajadora del hogar remunerada.
- Identificar, para cada uno de estos grupos de la población, el reconocimiento que tienen sobre el respeto a sus derechos, las principales problemáticas que perciben como grupo, la discriminación a través de prejuicios que perciben hacia su propia población, y sus experiencias de discriminación en los distintos ámbitos de su vida, ocurridas en los últimos doce meses, así como en los últimos cinco años.
- Medir la percepción que tienen las personas de 12 años o más, indígenas, sobre el respeto a los derechos hacia sus pueblos, así como de las principales problemáticas a las que se enfrentan actualmente como pueblo étnico, y las barreras de acceso que enfrentan cuando quieren acceder a algún tipo de información.
- Medir el impacto que tiene la discriminación en la población de 15 años o más, identificando los motivos por los que se vio obligada a cambiar su lugar de residencia, así como las situaciones que vive o vivió a causa de su condición de migrante.
- Conocer el grado de autonomía de las personas de 12 años o más que tienen alguna condición de discapacidad, así como de las de las personas de 60 años o más, y de las mujeres de 18 años o más, a partir de la libertad que sienten para realizar ciertas actividades o tomar sus propias decisiones.
- Captar la opinión que tienen las niñas, niños y adolescentes sobre los roles de género, así como las situaciones de violencia y discriminación que experimentan en su entorno familiar, en la escuela y en su comunidad.
- Identificar las principales razones por las que las y los jóvenes que actualmente no estudian, desertaron de la escuela. Así mismo, captar el motivo principal por el que esta población no se ha ocupado laboralmente.

- Conocer el motivo principal por el que las mujeres de 18 años o más, no ocupadas, actualmente no se encuentran insertas en el mercado laboral.
- Captar las condiciones laborales de las trabajadoras del hogar con remuneración, respecto al tipo de prestaciones con las que cuentan, el respeto a sus derechos laborales, el ambiente en el que se desenvuelven, y la cobertura en los gastos de salud.

## 1.2 Referencias metodológicas

Las principales referencias metodológicas asociadas a los objetivos y metas del proyecto son las siguientes:

**Unidad de observación.** Las viviendas seleccionadas y sus hogares.

**Periodo de referencia.** La encuesta indaga sobre condiciones, características, opiniones, hechos y situaciones ocurridas en diferentes periodos, entre los que destacan:

1. Al momento de la entrevista, que se considera en prácticamente todas las secciones del *Cuestionario General*, el *Cuadernillo de Módulos* y el *Cuestionario de Opinión y Experiencias*.
2. La semana pasada, en preguntas sobre condición y actividades laborales del *Cuestionario General*.
3. Últimos 12 meses, es decir, de julio de 2021 a la fecha de la entrevista, en preguntas específicas del *Cuadernillo de Módulos* y del *Cuestionario de Opinión y Experiencias*.
4. Últimos 5 años, de julio de 2017 a la fecha de la entrevista, en temas específicos del *Cuadernillo de Módulos* y del *Cuestionario de Opinión y Experiencias*.
5. Hace cinco años, referido al mes de julio de 2017, en preguntas sobre lugar y cambio de residencia en el *Cuestionario General* y *Cuadernillo de Módulos*.

**Periodo de levantamiento.** Del 18 de julio al 9 de septiembre de 2022.

**Método de recolección.** La información se obtiene mediante entrevista directa en la que se aplicó un cuestionario en Dispositivo Móvil o impreso en papel cuando las condiciones no permitieron el uso del electrónico, estructurado con preguntas ordenadas y secuenciadas, algunas con el apoyo de tarjetas que facilitaron al informante proporcionar su respuesta sin sesgos u omisiones.

**Informante adecuado.** Para las secciones I a la III del *Cuestionario General* fue una persona residente de la vivienda de 15 años o más que conociera los datos de todos los integrantes de su hogar. Posteriormente, y de acuerdo con las características sociodemográficas obtenidas, se entrevistó a la población de estudio de 12 años y más.

**Persona de interés.** Está dirigida a la población que reside permanentemente en viviendas particulares dentro del territorio nacional, así como a los siguientes grupos sociales vulnerables a la discriminación según su edad:

- Personas indígenas o afrodescendientes de 12 años o más de edad.
- Personas con discapacidad de 12 años o más de edad.
- Personas migrantes y desplazadas de 15 años o más de edad.
- Personas con diversidad religiosa de 12 años o más de edad.
- Personas de 60 años o más de edad.
- Niñas y niños de 9 a 11 años de edad.
- Adolescentes y jóvenes de 12 a 29 años de edad.
- Mujeres de 18 años o más de edad.

**Persona elegida.** Integrante del hogar de 18 años o más, cuya fecha de cumpleaños es la inmediata posterior a la fecha de la entrevista cuando el levantamiento se realiza en papel. En el caso de entrevistas electrónicas, su selección se realizó a través de un proceso aleatorio. Dio respuesta al *Cuestionario de Opinión y Experiencias* de acuerdo con el criterio de elegibilidad.

**Cobertura geográfica.** La encuesta está diseñada para dar resultados a nivel nacional.



## 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

Para la selección de la muestra se utilizó la Muestra Maestra, también conocida como el Marco Nacional de Viviendas 2012 del INEGI a partir del Marco Maestro de Muestreo 2012, el cual fue construido con base en la información cartográfica y demográfica que se obtuvo del Censo de Población y Vivienda 2010. Su diseño es probabilístico, estratificado, unietápico y por conglomerados; estos últimos también se consideran Unidades Primarias de Muestreo (UPM), pues una muestra de ellos forma la Muestra Maestra y a partir de esta se seleccionan, en una segunda etapa, las viviendas que integran las submuestras de las diferentes encuestas.

### 2.2 Formación de las UPM

Primeramente, se construye el conjunto de UPM que cubrirá 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 AGEB<sup>1</sup>.
- 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.

**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.

Por tanto, la Muestra Maestra quedó conformada por un total de 245 279 UPM.

---

<sup>1</sup> Área Geoestadística Básica.

## 2.3 Estratificación

Una vez construido el conjunto de UPM, se agruparon aquellas con características similares, lo cual se denomina estratificación.

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	
	-	
	-	Ciudades con 100 000 o más habitantes
	-	
Complemento urbano	09	
	25	De 50 000 a 99 999 habitantes
	35	De 15 000 a 49 999 habitantes
	45	De 5 000 a 14 999 habitantes
Rural	55	De 2 500 a 4 999 habitantes
	60	Localidades menores de 2 500 habitantes

De manera paralela, se formaron cuatro estratos socioeconómicos 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<sup>2</sup> construidos con información del Censo de Población y Vivienda 2010, para lo cual se emplearon métodos estadísticos multivariados.

La combinación de los estratos geográficos y los estratos sociodemográficos dio como resultado un total de 683 estratos de diseño en todo el territorio nacional.

De esta forma cada UPM del Marco Maestro de Muestreo resultó clasificada en un único estrato de diseño.

## 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:

$$P\{U_{ehi} \in S_m\} = \frac{k_{eh} m_{ehi}}{m_{eh}}$$

Donde:

$S_m$  = la  $m$ -ésima Muestra Maestra.

$k_{eh}$  = 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 captada en el Censo de Población y Vivienda 2010.

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

Las UPM seleccionadas forman la Muestra Maestra, que permite elegir las muestras de las encuestas en hogares.

<sup>2</sup> La descripción de estos indicadores se presenta en el anexo A.

## 2.5 Tamaño de la muestra

Para el cálculo del tamaño de muestra de la ENADIS 2022 se empleó la proporción de personas adultas mayores que declararon abandono, debido a que es una variable de menor frecuencia y las demás variables de estudio quedan cubiertas con el tamaño de muestra calculado. La expresión utilizada fue la siguiente:

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

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, 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.
- PHV = número promedio de personas del grupo de interés por vivienda.

Considerando una confianza del 90 %, un efecto del diseño del 3.65, un error máximo relativo esperado del 12.5 %, una tasa máxima de no respuesta esperada del 15 %, para una proporción mínima de 5.5 % y un promedio de adultos mayores (personas de 60 años o más) por vivienda de 0.43, se obtuvo un tamaño de muestra aproximado de 42 184 viviendas, el cual se ajustó a 42 302 viviendas a nivel nacional.

## 2.6 Distribución de la muestra

La distribución de la muestra  $n_e$  se realizó dentro de cada entidad federativa, de manera proporcional al número de viviendas de los estratos  $N_{eh}$ , 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 obtienen 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.
- b = número de viviendas a seleccionar por UPM.
- $k_{eh}^*$  = número de UPM a seleccionar en el h-ésimo estrato, en la e-ésima entidad.

En el anexo B se presenta la distribución de la muestra de UPM y viviendas seleccionadas por entidad federativa.

## 2.7 Selección de la muestra

La selección de la muestra se realizó de manera independiente por entidad, ámbito y estrato, el procedimiento de selección varió de acuerdo con el dominio, como se especifica a continuación:

### 2.7.1 En urbano alto

En el ámbito urbano alto la selección de la muestra se realizó en forma independiente para cada entidad y estrato de diseño mediante el siguiente procedimiento:

1. De las UPM que forman el Marco Maestro de Muestreo en el h-ésimo estrato, en la e-ésima entidad se eligieron  $k_{eh}$  UPM con probabilidad proporcional al tamaño para formar la Muestra Maestra.
2. De las  $k_{eh}$  UPM que integran la Muestra Maestra, en el h-ésimo estrato, en la e-ésima entidad se eligieron  $k_{eh}^*$  UPM con igual probabilidad para formar la muestra de UPM de la ENADIS 2022.
3. En cada UPM seleccionada, se eligieron cinco viviendas con igual probabilidad.
  - 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 selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de la ENADIS 2022 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:

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

Por lo tanto, la probabilidad total de selección de cada una de las viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad es el producto de las probabilidades de selección:

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

Su factor de expansión<sup>3</sup> está dado por:

$$F_{ehij} = \left\{ P_{ehij} \right\}^{-1} = \frac{m_{eh} m_{ehi}^*}{5 k_{eh}^* m_{ehi}}$$

<sup>3</sup> 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 para la Muestra Maestra en el h-ésimo estrato, en la e-ésima entidad.
- $k_{eh}^*$  = número de UPM seleccionadas para la muestra de la ENADIS 2022, en el h-ésimo estrato, en la e-ésima entidad.
- $m_{eh}$  = 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 el 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 el listado de viviendas actualizado.

## 2.7.2 En complemento urbano

En el ámbito complemento urbano la selección de la muestra se realizó en forma independiente para cada entidad y estrato de diseño mediante el siguiente procedimiento:

1. De las UPM que forman el Marco Maestro de Muestreo en el h-ésimo estrato, en la e-ésima entidad se eligieron  $k_{eh}$  UPM con probabilidad proporcional al tamaño, para formar la Muestra Maestra.
2. De las  $k_{eh}$  UPM que integran la Muestra Maestra, en el h-ésimo estrato, en la e-ésima entidad se eligieron  $k_{eh}^*$  UPM con igual probabilidad para formar la muestra de UPM de la ENADIS 2022.
3. En cada UPM seleccionada, se eligieron 20 viviendas con igual probabilidad.
  - 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 selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de la ENADIS 2022 es:

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

- 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:

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

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

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

Su factor de expansión está dado por:

$$F_{ehij} = \left\{ P_{ehij} \right\}^{-1} = \frac{m_{eh} m_{ehi}^*}{20 k_{eh}^* m_{ehi}}$$

Donde:

- $k_{eh}$  = número de UPM seleccionadas para la Muestra Maestra en el h-ésimo estrato, en la e-ésima entidad.
- $k_{eh}^*$  = número de UPM seleccionadas para la muestra de la ENADIS 2022 en el h-ésimo estrato, en la e-ésima entidad.
- $m_{eh}$  = 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 el 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 el listado de viviendas actualizado.

### 2.7.3 En rural

En el ámbito rural se realizó en forma independiente para cada entidad y estrato de diseño mediante el siguiente procedimiento:

1. De las UPM que forman el Marco Maestro de Muestreo en el h-ésimo estrato, en la e-ésima entidad se eligieron  $k_{eh}$  UPM con probabilidad proporcional al tamaño para formar la Muestra Maestra.
2. De las  $k_{eh}$  UPM que integran la Muestra Maestra, en el h-ésimo estrato, en la e-ésima entidad se eligieron  $k_{eh}^*$  UPM con igual probabilidad para formar la muestra de UPM de la ENADIS 2022.
3. En cada UPM seleccionada, se eligieron cuatro segmentos de cinco viviendas aproximadamente, con igual probabilidad.

- a. La probabilidad de selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la Muestra Maestra es:

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

- b. La probabilidad de selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de la ENADIS 2022 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:

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

Por lo tanto, la probabilidad total de selección de cada una de las viviendas de la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad es el producto de las probabilidades de selección:

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

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

Su factor de expansión está dado por:

$$F_{ehij} = \left\{ P_{ehij} \right\}^{-1} = \frac{m_{eh} m_{ehi}^*}{20 k_{eh}^* m_{ehi}}$$

Donde:

- $k_{eh}$  = número de UPM seleccionadas para la Muestra Maestra en el h-ésimo estrato, en la e-ésima entidad.
- $k_{eh}^*$  = número de UPM seleccionadas para la muestra de la ENADIS 2022, en el h-ésimo estrato, en la e-ésima entidad.
- $m_{eh}$  = 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 el 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 el listado de viviendas actualizado.

## 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 de las viviendas

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

$$F'_{ehij} = F_{ehij} \frac{\sum_{i \in h} F_{ehij} V_{ehi}}{\sum_{i \in h} F_{ehij} V_{ehi}^*}$$

Donde:

- $F'_{ehij}$  = factor de expansión corregido por no respuesta de la j-ésima vivienda seleccionada, 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.
- $V_{ehi}$  = número de viviendas seleccionadas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.
- $V_{ehi}^*$  = número de viviendas con respuesta en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

### 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:

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

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.
- $PEXP_D$  = población total a la que expande la encuesta en el dominio D.
- D = nivel de desagregación en que se efectúa el ajuste.

### 2.8.3 Ajuste por no respuesta de los hogares

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

$$F'''_{ehij} = F''_{ehij} \frac{\sum F''_{ehij} Q_{ehij}}{\sum F''_{ehij} Q^*_{ehij}}$$

Donde:

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

$F''_{ehij}$  = 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.

$Q_{ehij}$  = número de hogares en la j-ésima vivienda seleccionada, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

$Q^*_{ehij}$  = número de hogares con respuesta en la j-ésima vivienda seleccionada, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

### 2.8.4 Factor de expansión a nivel persona

Debido a que dentro de las viviendas no hubo submuestreo de hogares, estos conservaron la misma probabilidad de selección de la vivienda a la que pertenecen, por lo que la probabilidad de elegir a una persona en forma aleatoria dentro t-ésimo hogar, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad está definida por:

$$P_{ehijt} = P_{ehi} \frac{1}{T_{ehijt}}$$

Su factor de expansión está dado por:

$$F^*_{ehijt} = F'''_{ehijt} T_{ehijt}$$

El factor persona del *Cuestionario de Opinión y Experiencias* se ajusta por no respuesta considerando que fue elegido para este módulo, a nivel entidad, empleando la siguiente expresión:

$$F^{**}_{ehijt} = F^*_{ehijt} \frac{\sum F^*_{ehijt} T_{ehijt}}{\sum F^*_{ehijt} T^*_{ehijt}}$$

Donde:

$T_{ehijt}$  = número de personas de 18 o más años de edad, en el t-ésimo hogar, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

$T^*_{ehijt}$  = número de personas de 18 o más años de edad con respuesta, en el t-ésimo hogar, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

$F^{**}_{ehijt}$  = factor de expansión ajustado por no respuesta para la persona elegida del t-ésimo hogar, de la j-ésima vivienda seleccionada, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

$F^*_{ehijt}$  = factor de expansión de la persona elegida del t-ésimo hogar, de la j-ésima vivienda seleccionada, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.



## 2.8.5 Factor de expansión por módulo

El factor de expansión de las personas sujetas de estudio para un módulo determinado se tomó directamente del factor hogar ajustado por no respuesta.

El factor persona para un módulo determinado se ajusta por no respuesta por ser sujeto de estudio, a nivel entidad empleando la siguiente expresión:

$$F_{ehijt}^r = F_{ehij}''' \frac{\sum F_{ehij}''' T_{ehijt}^r}{\sum F_{ehij}''' T_{ehijt}^{*r}}$$

Donde:

$r$  = es el identificador del módulo de interés.

$T_{ehijt}^r$  = número de personas, sujetas de estudio para el  $r$ -ésimo módulo, en el  $t$ -ésimo hogar, en la  $j$ -ésima vivienda, en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato, en la  $e$ -ésima entidad.

$T_{ehijt}^{*r}$  = número de personas, sujetas de estudio para el  $r$ -ésimo módulo con respuesta, en el  $t$ -ésimo hogar, en la  $j$ -ésima vivienda, en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato, en la  $e$ -ésima entidad.

$F_{ehijt}^r$  = factor de expansión, ajustado por no respuesta, para el sujeto de estudio para el  $r$ -ésimo módulo, del  $t$ -ésimo hogar, 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 los sujetos de estudio a nivel hogar, de la  $j$ -ésima vivienda, de la  $i$ -ésima UPM, del  $h$ -ésimo estrato, de la  $e$ -ésima entidad.

### Identificador del módulo de interés y su descripción

Identificador $r$	Descripción del módulo
1	Personas indígenas o afrodescendientes de 12 años o más de edad.
2	Personas con discapacidad de 12 años o más de edad.
3	Personas migrantes y desplazadas de 15 años o más de edad
4	Personas con diversidad religiosa de 12 años o más de edad.
5	Personas de 60 años o más de edad.
6	Niñas y niños de 9 a 11 años de edad.
7	Adolescentes y jóvenes de 12 a 29 años de edad.
8	Mujeres de 18 años o más de edad.

## 2.9 Estimadores

El estimador del total de la característica  $X$  a nivel nacional es:

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

Donde:

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

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

$F_{ehij}^{CU}$  = factor de expansión final de la  $i$ -ésima UPM, del  $h$ -ésimo estrato, de la  $e$ -ésima entidad, del ámbito complemento urbano.

$X_{ehis\ell}^{CU}$  = valor observado de la característica  $X$  en el  $\ell$ -ésimo hogar, en la  $s$ -ésima vivienda, en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato, en la  $e$ -ésima entidad del ámbito complemento urbano.

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

$X_{ehis\ell}^R$  = valor observado de la característica  $X$  en el  $\ell$ -ésimo hogar, en la  $s$ -ésima vivienda, en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato, en la  $e$ -ésima entidad del ámbito rural.

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

Donde, la variable  $\hat{Y}$  es definida 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 del agregado nacional, se usó el método de *Conglomerados Últimos*<sup>4</sup>, basado en que la mayor contribución a la varianza de un estimador, en un diseño polietápico, 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 los errores de muestreo de los estimadores de razón, conjuntamente al método de *Conglomerados Últimos* se aplicó el método de *Serie de Taylor*, obteniéndose la siguiente fórmula para estimar la precisión de  $\hat{R}$  :

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

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.

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

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

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

Estas definiciones son análogas para la variable de estudio  $Y$ .

<sup>4</sup> Véase Hansen, M. H., Horwitz, W. N. y Madow, W.G., *Sample Survey Methods and Theory*, (1953), Vol. 1, pág. 242.

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 de la desviación estándar (D.E.), coeficiente de variación o error relativo del estimador (C.V.) y el efecto de diseño (DEFF) se calculan mediante las siguientes expresiones:

$$D.E. = \sqrt{\hat{V}(\hat{\theta})} \quad C.V. = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}} \quad DEFF = \frac{\hat{V}(\hat{\theta})}{\hat{V}(\hat{\theta})_{mas}}$$

Donde:

- $\hat{\theta}$  = estimador del parámetro poblacional  $\theta$ .
- $\hat{V}(\hat{\theta})_{mas}$  = 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  $\alpha$  es el nivel de significancia.

## 2.11 Homologación de la semaforización para los umbrales de indicadores de precisión estadística<sup>5</sup>

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

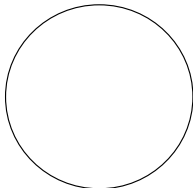
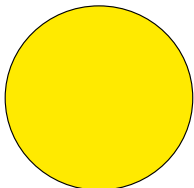
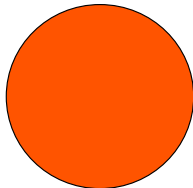
Interpretación	Semaforización	Nivel de precisión
Alta	Blanco	(0%, 15%)
Moderada	Amarillo	[15%, 30%)
Baja	Naranja oscuro	>=30%

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 indicar el nivel de precisión de las estimaciones con base en el CV. A continuación, se presenta el código RGB (sigla en inglés de *red*, *green*, *blue*; en español: rojo, verde y azul) de los colores utilizados en la semaforización.

<sup>5</sup> 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](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 los conglomerados del Marco Maestro de Muestreo

Mnemónico	Descripción
<b>Proporción de población</b>	
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.
<b>Proporción de viviendas particulares habitadas</b>	
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 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.
<b>Proporción de viviendas particulares habitadas que disponen de</b>	
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 de UPM y viviendas seleccionadas por entidad federativa

Entidad federativa	UPM	Viviendas
<b>Nacional</b>	<b>5 588</b>	<b>42 302</b>
Aguascalientes	203	1 318
Baja California	219	1 321
Baja California Sur	146	1 314
Campeche	176	1 324
Coahuila de Zaragoza	216	1 324
Colima	191	1 315
Chiapas	123	1 321
Chihuahua	213	1 323
Ciudad de México	262	1 320
Durango	180	1 325
Guanajuato	162	1 327
Guerrero	138	1 323
Hidalgo	117	1 324
Jalisco	183	1 322
México	219	1 327
Michoacán de Ocampo	129	1 321
Morelos	180	1 320
Nayarit	146	1 327
Nuevo León	228	1 320
Oaxaca	114	1 325
Puebla	156	1 322
Querétaro	180	1 322
Quintana Roo	215	1 316
San Luis Potosí	162	1 322
Sinaloa	171	1 325
Sonora	195	1 325
Tabasco	114	1 325
Tamaulipas	216	1 323
Tlaxcala	185	1 316
Veracruz de Ignacio de la Llave	144	1 324
Yucatán	177	1 322
Zacatecas	128	1 319