

# Encuesta Nacional de la Dinámica Demográfica 2018

## ENADID



Diseño muestral



**INSTITUTO NACIONAL  
DE ESTADÍSTICA Y GEOGRAFÍA**

Instituto Nacional de Estadística y Geografía

**Encuesta Nacional  
de la Dinámica Demográfica 2018**

**ENADID**

**Diseño muestral**



**Obras complementarias publicadas por el INEGI sobre el tema:**

Encuesta Nacional de la Dinámica Demográfica 2014. ENADID. Marco de referencia.

**Catalogación en la fuente INEGI:**

304.6021072 Encuesta Nacional de la Dinámica Demográfica (2019).  
Encuesta Nacional de la Dinámica Demográfica 2019 : ENADID : diseño muestral  
/ Instituto Nacional de Estadística y Geografía.-- México : INEGI, c2019.

vii, 14 p.

1. México - Población - Encuestas. I. Instituto Nacional de Estadística y Geografía (México).

**Conociendo México**

**01 800 111 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**

DR © 2019, **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 de la Dinámica Demográfica (ENADID) 2018**, 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 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	1
<b>2. Diseño estadístico</b>	3
2.1 Marco de la encuesta	3
2.2 Formación de las Unidades Primarias de Muestreo (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	6
2.7.1 En urbano alto	6
2.7.2 En complemento urbano	6
2.7.3 En rural	7
2.8 Ajuste a los factores de expansión	8
2.9 Estimadores	9
2.10 Estimación de las precisiones	10
<b>Anexo</b>	11
A. Indicadores empleados en la estratificación de la Muestra Maestra	13
B. Distribución de la muestra en viviendas, por entidad federativa, según tamaño de localidad	14

## Introducción

---

La **Encuesta Nacional de la Dinámica Demográfica (ENADID) 2018** es un proyecto estadístico que responde a la necesidad de dar continuidad a una serie de encuestas sobre la dinámica demográfica en nuestro país, conocer los niveles de fecundidad, mortalidad infantil y movimientos migratorios, así como abordar temas que permiten predecir la dinámica de crecimiento de la población, como preferencias reproductivas, sexualidad, uso de métodos anticonceptivos y características del último embarazo de las mujeres.

El diseño muestral que se presenta a continuación está organizado en dos capítulos; en el primero se abordan los objetivos y 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

Con base en la necesidad de profundizar en el tema y generar conocimiento sobre la dinámica de aspectos demográficos en nuestro país, se establecieron los siguientes objetivos:

### General

Actualizar la información estadística relacionada con el nivel y comportamiento de los componentes de la dinámica demográfica: fecundidad, mortalidad y migración (interna e internacional), aunados a otros temas relacionados con la dinámica de crecimiento de la población, como preferencias reproductivas, sexualidad, uso de métodos anticonceptivos, salud materno infantil y nupcialidad.

### Específicos

La nueva edición de la encuesta pretende mantener la comparabilidad y dar continuidad a los indicadores obtenidos con las encuestas que le anteceden, para conocer y actualizar:

- La composición por edad y sexo de la población.
- Los niveles y estructuras de la fecundidad.
- La prevalencia de la anticoncepción.
- Los niveles y tendencias de la mortalidad infantil.
- La información sobre migración interestatal e internacional.
- El monto de la expedición de certificados y registros de nacimiento, certificados y registros de defunción para la población menor de 5 años.
- Causa y edad del abandono escolar de las mujeres de 15 a 34 años que no asisten a la escuela.
- Aspectos de la salud reproductiva de las mujeres durante el último embarazo, parto y posparto, así como del desarrollo del feto o del recién nacido(a).
- Relacionar el comportamiento de las variables demográficas con las características socioeconómicas de la población y de sus viviendas.

## 1.2 Referencias metodológicas

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

**Unidad de selección.** Vivienda particular.

**Unidad de observación.** Vivienda particular habitada, hogar, residente habitual, migrante internacional, mujer elegible de 15 a 54 años.

**Población objetivo.** Personas residentes habituales de las viviendas, así como las mujeres de 15 a 54 años de edad, con la misma condición de habitación.

**Cobertura geográfica.** La muestra se encuentra distribuida en todo el país, tanto en localidades urbanas como rurales, de tal manera que será posible generar información a nivel nacional, nacional por tamaño de localidad y entidad federativa.

**Periodo de levantamiento.** Del 13 de agosto al 5 de octubre.

**Método de recolección.** La información se obtiene mediante entrevista directa con cuestionario electrónico en un dispositivo móvil; en los casos donde no fue posible aplicar un cuestionario electrónico, se utilizaron los correspondientes en papel. También se hizo uso de una tarjeta como apoyo para aplicar la pregunta 3.11A.

**Informante adecuado,** según instrumento de captación:

- *Cuestionario para el hogar.* Integrante del hogar de 15 años o más, residente habitual de la vivienda, que conozca la información solicitada de todos los integrantes del hogar.
- *Módulo de la mujer.* Cada una de las mujeres de 15 a 54 años, residente habitual de la vivienda seleccionada.

**Tamaño de muestra.** Fue de 119 800 viviendas particulares localizadas en el territorio 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

Se empleó el Marco Nacional de Viviendas 2012 del INEGI, construido a partir de la información cartográfica y demográfica obtenida del Censo de Población y Vivienda 2010. Este marco es en realidad una Muestra Maestra de la que a su vez se seleccionan las muestras para todas las encuestas en viviendas que realiza el INEGI; como tal, su diseño es probabilístico, estratificado, unietápico y por conglomerados. A estos últimos se les denominó unidades primarias de muestreo, y es donde son seleccionadas, en una segunda etapa, las viviendas integrantes de las muestras de las diferentes encuestas.

El diseño de la muestra para la ENADID 2018 es probabilístico, lo que permite generalizar los resultados a toda la población; por conglomerados, estratificado y bietápico, porque la unidad última de muestreo (vivienda) es seleccionada en dos etapas.

### 2.2 Formación de las Unidades Primarias de Muestreo (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 AGE<sup>1</sup>.
- La unión de dos o más manzanas contiguas de diferentes AGE<sup>1</sup> de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes localidades, pero del 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 AGE<sup>1</sup>.
- La unión de dos o más manzanas contiguas de diferentes AGE<sup>1</sup> de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes AGE<sup>1</sup> y localidades del mismo municipio.

---

<sup>1</sup> 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.

La Muestra Maestra quedó conformada por un total de 240 912 UPM.

## 2.3 Estratificación

Una vez construido el conjunto de UPM, se agruparon aquellas con características similares (se estratificaron). 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 ... 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, en una primera etapa se formaron cuatro estratos 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<sup>2</sup>, para lo cual se emplearon métodos estadísticos multivariados.

En una segunda etapa, cada UPM clasificada con su estrato sociodemográfico fue asignada a su estrato geográfico. 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, mediante la expresión:

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

<sup>2</sup> La descripción de estos indicadores se presenta en el cuadro A, al final de este capítulo.

Donde:

$U_{ehi}$  = la  $i$ -ésima UPM, del  $h$ -ésimo estrato, de la  $e$ -ésima entidad.

$S$  = 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 captadas en el Censo de Población y Vivienda 2010.

$m_{eh}$  = número de viviendas en el  $h$ -ésimo estrato, en la  $e$ -ésima entidad captadas 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.

## 2.5 Tamaño de la muestra

Se calculó para una tasa de fecundidad general, considerada como una de las principales variables de la encuesta, y la cual requiere los tamaños de muestra mayores. Esto garantiza que las estimaciones del resto de las variables de interés queden cubiertas con este tamaño.

La expresión empleada para el cálculo es la siguiente:

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

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.

$\text{DEFF}$  = 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.

$\text{tnr}$  = tasa de no respuesta máxima esperada.

$\text{PMV}$  = promedio de años de exposición de las mujeres de 15 a 54 años de edad por vivienda.

Considerado una confianza de 90%, un efecto del diseño de 1.20, un error relativo máximo esperado de 15%, un promedio de años de exposición de las mujeres de 15 a 54 años de edad por vivienda de 0.49, una tasa de no respuesta máxima esperada de 15%, y para una tasa de fecundidad específica a nivel estatal de 77.04 hijos por cada 1 000 mujeres, se obtuvo un tamaño de muestra mínimo de 3 495 viviendas por entidad. Mismo que se ajustó a 3 500 viviendas. En el anexo B se presenta la distribución de la muestra en viviendas, por entidad federativa, según tamaño de localidad.

## 2.6 Afijación de la muestra

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

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

El número de UPM a seleccionar 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.
- $b$  = número de viviendas a seleccionar por UPM.

Al final del documento, en el cuadro B, se presenta la distribución de la muestra por entidad y tamaño de localidad.

## 2.7 Selección de la muestra

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

### 2.7.1 En urbano alto

1. Se seleccionaron  $K_{eh}$  UPM, con probabilidad proporcional al número de viviendas del estrato.
2. En cada UPM seleccionada, se eligieron cinco viviendas con igual probabilidad.

Por lo tanto, la probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

$$P\{V_{ehj}\} = \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{5}{m_{ehi}^*} = \frac{5k_{eh} m_{ehi}}{m_{eh} m_{ehi}^*}$$

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

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

Donde:

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

### 2.7.2 En complemento urbano

1. Se seleccionaron  $K_{eh}$  UPM, con probabilidad proporcional al total de viviendas del estrato.
2. En cada UPM seleccionada, se eligieron 20 viviendas con igual probabilidad.

<sup>3</sup> El factor de expansión se define como el inverso de la probabilidad de selección.

Por lo tanto, la probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

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

Su factor de expansión está dado por:

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

Donde:

- $k_{eh}$  = número de UPM seleccionadas en el h-ésimo estrato, en la e-ésima entidad para el marco de 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 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 listado de viviendas actualizado.

### 2.7.3 En rural

1. Se seleccionaron  $K_{eh}$  UPM, con probabilidad proporcional al total de viviendas del estrato.
2. En cada UPM seleccionada, se eligieron cuatro segmentos de cinco viviendas aproximadamente con igual probabilidad.

Por lo tanto, la probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

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

Su factor de expansión está dado por:

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

Donde:

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

### Ajuste por no respuesta

El ajuste por no respuesta atribuida al informante se realizó a nivel estrato, en cada uno de los dominios para cada una de las unidades de observación.

- Ajuste por no respuesta de las viviendas:

$$F'_{ehi} = F_{ehi} \frac{\sum_{ieh} F_{ehi} V_{ehi}}{\sum_{ieh} F_{ehi} V_{ehi}^*}$$

- Ajuste por no respuesta de las personas:

$$F''_{ehij} = F'_{ehi} \frac{\sum_{ieh} F'_{ehij} Q_{ehij}}{\sum_{ieh} F'_{ehij} Q_{ehij}^*}$$

Donde:

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

$F_{ehi}$  = factor de expansión 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.

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

$Q_{ehij}^*$  = número de personas elegidas que proporcionaron respuesta en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad.

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

### Ajuste por proyección

Los factores de expansión ajustados por la no respuesta se corrigen, con el propósito de asegurar que en cada dominio de interés de la encuesta se obtenga la población total de acuerdo a la proyección de población establecida por el Consejo Nacional de Población, la cual está referida al punto medio del levantamiento, mediante la siguiente expresión:

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

Donde:

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

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

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

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

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

## Ajuste a nivel persona para las mujeres de 15 a 54 años de edad

Los factores de expansión corregidos por proyección de población, se ajustan a nivel UPM para la población de mujeres de 15 a 54 años de edad, mediante la siguiente expresión:

$$F_{ehj}''' = F_{ehj}'' \frac{M_{15-54}}{M_{15-54}^*}$$

Donde:

$F_{ehj}'''$  = factor de expansión ajustado a nivel persona de las mujeres de 15 a 54 años de edad para la  $j$ -ésima vivienda, de la  $i$ -ésima UPM, del  $h$ -ésimo estrato, de la  $e$ -ésima entidad.

$F_{ehj}''$  = factor de expansión corregido por la proyección de la población para la  $j$ -ésima vivienda, de la  $i$ -ésima UPM, del  $h$ -ésimo estrato, de la  $e$ -ésima entidad.

$M_{15-54}$  = total de mujeres de 15 a 54 años de edad en la UPM seleccionada.

$M_{15-54}^*$  = total de mujeres de 15 a 54 años de edad que contestaron el cuestionario del módulo de fecundidad en la UPM seleccionada.

## 2.9 Estimadores

El estimador del total de la característica  $X$  es:

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

Donde:

$F_{ehis}^{UA}$  = factor de expansión final, de la  $s$ -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  $\ell$ -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.

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

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

$F_{ehis}^R$  = factor de expansión final de la  $s$ -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  $X$  en la  $\ell$ -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, la variable  $\hat{Y}$  es definida en forma análoga a  $\hat{X}$ .

## 2.10 Estimación de las precisiones

Para la evaluación de los errores de muestreo de las principales estimaciones estatales y nacionales, 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 bietápico, es la que se presenta entre las 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, se aplicó el método de *Conglomerados Últimos* conjuntamente con el método de *Series de Taylor*, obteniéndose la siguiente fórmula para estimar la precisión de  $\hat{R}$ :

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

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.

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

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

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

Estas definiciones son análogas para la variable de estudio  $\hat{Y}$ .

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

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

Las estimaciones de la desviación estándar (D.E.), efecto de diseño (DEFF) y coeficiente de variación (C.V.) se calculan mediante las siguientes expresiones:

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

Donde:

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

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

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



# Anexo

## A. Indicadores empleados en la estratificación de la Muestra Maestra

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.
PVDRENERD	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 en viviendas<sup>5</sup>, por entidad federativa, según tamaño de localidad

Entidad federativa	Total	Tamaño de localidad			
		1	2	3	4
<b>Nacional</b>	<b>119 941</b>	<b>54 915</b>	<b>18 805</b>	<b>18 354</b>	<b>27 867</b>
Aguascalientes	3 504	2 220	370	270	644
Baja California	4 002	2 910	500	290	302
Baja California Sur	4 004	1 400	1 420	640	544
Campeche	3 509	1 745	260	675	829
Coahuila de Zaragoza	3 499	2 315	670	150	364
Colima	3 994	2 440	600	520	434
Chiapas	3 508	750	560	570	1 628
Chihuahua	3 508	2 540	220	220	528
Ciudad de México	3 999	3 875	70	35	19
Durango	3 498	1 695	355	405	1 043
Guanajuato	3 507	1 390	760	330	1 027
Guerrero	3 501	1 090	440	570	1 401
Hidalgo	3 707	520	585	865	1 737
Jalisco	4 000	2 075	865	540	520
México	3 507	2 140	430	510	427
Michoacán de Ocampo	3 498	825	845	770	1 058
Morelos	4 002	1 520	850	1 010	622
Nayarit	3 503	1 085	575	800	1 043
Nuevo León	4 002	3 075	495	210	222
Oaxaca	3 513	350	585	785	1 793
Puebla	3 502	1 095	695	770	942
Querétaro	4 009	1 770	355	755	1 129
Quintana Roo	4 002	2 920	360	300	422
San Luis Potosí	4 011	1 780	400	440	1 391
Sinaloa	3 788	1 880	440	480	988
Sonora	4 006	2 475	600	405	526
Tabasco	3 814	635	700	904	1 575
Tamaulipas	4 009	2 820	440	260	489
Tlaxcala	3 504	0	1 200	1 595	709
Veracruz de Ignacio de la Llave	4 015	985	785	750	1 495
Yucatán	4 004	1 695	835	870	604
Zacatecas	3 512	900	540	660	1 412

<sup>5</sup> El tamaño de la muestra calculada fue de 119 800 viviendas; sin embargo, al cierre del operativo de campo se captó información de 119 941. Tal diferencia se debe a que en el dominio rural se seleccionan segmentos conformados por conjuntos de cuatro a seis viviendas cercanas entre sí, de las que se obtiene información, de tal manera que se recolectan datos de un número de viviendas que puede ser marginalmente inferior o superior al calculado, el cual no influye en las precisiones estadísticas.