

Encuesta Nacional sobre Uso del Tiempo 2019

ENUT

Diseño muestral



Segunda edición

Instituto Nacional de Estadística y Geografía

**Encuesta Nacional sobre
Uso del Tiempo 2019**

ENUT

Diseño muestral



Índice	Página
1. Objetivo de la encuesta	1
2. Población objetivo	1
3. Cobertura geográfica	1
4. Desglose geográfico	1
5. Marco de la encuesta	1
5.1 Formación de las unidades primarias de muestreo (UPM)	1
5.1.1 En urbano alto	1
5.1.2 En complemento urbano	2
5.1.3 En rural	2
5.2 Estratificación	2
5.3 Selección de las UPM de la muestra maestra	2
6. Tamaño de la muestra	3
7. Distribución de la muestra	3
8. Selección de la muestra	4
8.1 En urbano alto	4
8.2 En complemento urbano	5
8.3 En rural	6
9. Ajuste a los factores de expansión	7
9.1 Ajuste por no respuesta	8
9.1.1 Ajuste por no respuesta a nivel vivienda	8
9.1.2 Ajuste por no respuesta a nivel hogar	8
9.1.3 Ajuste por no respuesta a nivel persona	8
9.2 Ajuste por proyección	9
9.3 Ajuste por no respuesta a nivel persona para informantes de 12 años o más de edad	9
9.4 Ajuste para la muestra de población hablante de lengua indígena	10
10. Estimadores	10
11. Estimación de los errores de muestreo	11
12. Homologación de la Semaforización para los Umbrales de Indicadores de precisión estadística	12
Anexos	
- Indicadores empleados en la estratificación de los conglomerados de la Muestra Maestra (Cuadro 1).	14
- Distribución de la muestra en viviendas para la ENUT 2019, por entidad y dominio (Cuadro 2).	15
- Distribución de la muestra adicional en viviendas por entidad y dominio ENUT 2019 (Cuadro 3).	16

1. Objetivo de la encuesta

Generar información estadística sobre los patrones de uso del tiempo de la población mexicana, con el fin de ayudar a la generación y evaluación de políticas públicas.

2. Población objetivo

Población de 12 años o más de edad que residen en viviendas particulares ubicadas dentro del territorio nacional.

3. Cobertura geográfica

El proyecto cubre todo el territorio nacional, incluyendo áreas urbanas y rurales.

4. Desglose geográfico

Nacional, tamaño de localidad (menor a 10 000 habitantes e igual o mayor de 10 000 habitantes) y entidad federativa.

5. Marco de la encuesta

El diseño muestral de la ENUT 2019 se caracteriza por ser probabilístico, por lo cual los resultados obtenidos de la encuesta se generalizan a toda la población, a su vez es bietápico, estratificado y por conglomerados, donde las unidades de observación son los hogares y las personas de 12 o más de edad.

La submuestra para la ENUT 2019 se seleccionó a partir de la muestra maestra 2012 del INEGI, esta muestra maestra se diseñó y seleccionó del Marco Maestro de Muestreo 2012 (MMM) el cual se conformó de conglomerados de viviendas llamados Unidades Primarias de Muestreo (UPM), construidos a partir de la información cartográfica y demográfica que se obtuvo del Censo de Población y Vivienda 2010. La muestra maestra permite la selección de submuestras para todas las encuestas en viviendas que realiza el INEGI; su diseño es probabilístico, estratificado, unietápico y por conglomerados, pues es en ellos donde se seleccionaron, en una segunda etapa, las viviendas que integran las submuestras de las diferentes encuestas. En diseño del MMM se construyó de la siguiente manera:

5.1 Formación de las unidades primarias de muestreo (UPM)

Primeramente, 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:

5.1.1 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.

¹ Área Geoestadística Básica.

5.1.2 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.

5.1.3 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.

5.2 Estratificación

Una vez construido el conjunto de UPM, se agrupan aquellas con características similares, o sea se estratifican.

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	Localidades de 50 000 a 99 999 habitantes
	35	Localidades de 15 000 a 49 999 habitantes
	45	Localidades de 5 000 a 14 999 habitantes
	55	Localidades 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 estas, 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.

5.3 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:

* La descripción de estos indicadores se presenta en el cuadro 1.

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

Donde:

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

S_m = la 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 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 en viviendas de las encuestas en hogares.

6. Tamaño de la muestra

Para calcular el tamaño de muestra se empleó la siguiente expresión:

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

Donde:

n = tamaño de la muestra.

z = valor asentado en las tablas estadísticas de la distribución normal estándar para una confianza prefijada.

p = es la estimación de la proporción de interés.

r = error relativo máximo aceptable.

q = $1-p$

tnr = tasa de no respuesta máxima esperada.

$DEFF$ = 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.

Considerando una confianza del 90%, un efecto de diseño de 2.12, un error relativo máximo esperado de 15%, una tasa de no respuesta máxima esperada de 15%, y una proporción de 1%, se obtuvo un tamaño de muestra de 29 686 viviendas, que se ajustó a 30 018 viviendas a nivel nacional. Este tamaño de muestra alcanza para cubrir proporciones iguales o mayores de 1%, obteniendo errores esperados menores de 15%; para proporciones menores de 1% se obtendrán errores esperados mayores de 15%. En el cuadro 2 se presenta la distribución de la muestra.

Con el objeto de disminuir los errores en la estimación de las variables de uso del tiempo entre la población de 12 años o más de edad hablante de lengua indígena, se seleccionó una muestra de 2 000 viviendas en localidades con 90% o más de población hablante de lengua indígena. La distribución de la muestra se presenta en el cuadro 3. A la muestra anterior se le agregaron las observaciones de población hablante de lengua indígena obtenidas en la muestra nacional, a fin de generar resultados sobre el uso del tiempo para esta subpoblación en particular.

7. Distribución de la muestra

La distribución de la muestra se realiza dentro de cada entidad federativa, entre los diferentes estratos, de manera proporcional a su tamaño, para lo cual se emplea 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.

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

b = número de viviendas a seleccionar por UPM.

8. Selección de la muestra

La selección de la muestra fue bietápica, donde las unidades de selección de segunda etapa fueron las viviendas y la unidad de observación el hogar. Se hizo una selección independiente para cada entidad y estrato, el procedimiento varió dependiendo del ámbito, como se especifica a continuación:

8.1 En urbano alto

1. 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 la ENUT 2019.
2. En cada UPM se seleccionaron cinco viviendas con igual probabilidad para la ENUT 2019.

La probabilidad de seleccionar la i-ésima UPM en el h-ésimo estrato, en la e-ésima entidad con probabilidad proporcional al tamaño para la muestra maestra es:

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

La probabilidad de selección de la i-ésima UPM_{ehi} , del h-ésimo estrato, de la e-ésima entidad para la submuestra de la ENUT 2019 es:

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

La probabilidad de seleccionar a cada vivienda de la i-ésima UPM_{ehi} , del h-ésimo estrato, de la e-ésima entidad es:

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

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:

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

Su factor de expansión² está dado por:

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

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 ENUT 2019, 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, 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 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 al momento de la actualización del listado de viviendas, previo al levantamiento de la ENUT 2019.

8.2 En complemento urbano

1. 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 la ENUT 2019.
2. En cada UPM seleccionada, se eligieron 20 viviendas con igual probabilidad.

La probabilidad de seleccionar la i-ésima UPM en el h-ésimo estrato, en la e-ésima entidad con probabilidad proporcional al tamaño para la muestra maestra es:

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

La probabilidad de selección de la i-ésima UPM_{ehi}, del h-ésimo estrato, de la e-ésima entidad para la submuestra de la ENUT 2019 es:

² 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.

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

La probabilidad de seleccionar a cada vivienda de la i-ésima UPM_{ehi}, del h-ésimo estrato, de la e-ésima entidad es:

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

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:

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

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

Su factor de expansión está dado por:

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

Donde:

- k_{eh} = número de UPM seleccionadas para la muestra maestra en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi} = número de viviendas en la i-ésima UPM, 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, según Censo de Población y Vivienda 2010.
- k_{eh}^* = número de UPM seleccionadas para la ENUT 2019, en el h-ésimo estrato, en la e-ésima entidad.
- m_{ehi}^* = número total de viviendas en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad al momento de la actualización de los listados, previo al levantamiento de la ENUT 2019.

8.3 En rural

1. 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 la ENUT 2019.
2. En cada UPM seleccionada, se eligieron, 4 segmentos de 5 viviendas cada uno con igual probabilidad.

La probabilidad de seleccionar la i-ésima UPM del h-ésimo estrato, de la e-ésima entidad con probabilidad proporcional al tamaño para la muestra maestra es:

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

La probabilidad de selección de la i -ésima UPM_{ehi} , del h -ésimo estrato, de la e -ésima entidad para la submuestra de la ENUT 2019 es:

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

La probabilidad de seleccionar a cada vivienda de la i -ésima UPM_{ehi} , del h -ésimo estrato, de la e -ésima entidad es:

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

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:

$$\begin{aligned} P_{ehi} &= P_{1ehi} * P_{2ehi} * P_{3ehij} \\ &= \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{4 * 5}{m_{ehi}^*} = \frac{20 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*} \end{aligned}$$

En consecuencia, su factor de expansión está dado por:

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

Donde:

- k_{eh} = número de UPM seleccionadas para la muestra maestra en el h -ésimo estrato, en la e -ésima entidad.
- 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_{eh} = número total de viviendas en el h -ésimo estrato, en la e -ésima entidad, según Censo de Población y Vivienda 2010.
- k_{eh}^* = número de UPM seleccionadas para la ENUT 2019, en el h -ésimo estrato, en la e -ésima entidad.
- m_{ehi}^* = número total de viviendas en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad al momento de la actualización de los listados, previo al levantamiento de la ENUT 2019.

9. Ajuste a los factores de expansión

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

9.1 Ajuste por no respuesta

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

9.1.1 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 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.
- I_{ehij} = función indicadora con valor 1 ó 0; esta función toma el valor uno si la j-ésima vivienda seleccionada, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad proporcionó respuesta completa, y toma el valor de cero en caso contrario.

9.1.2 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:

$$F''_{ehijk} = F'_{ehij} \frac{\sum_h \sum_i \sum_j F'_{ehij}}{\sum_h \sum_i \sum_j F'_{ehij} I_{ehijk}}$$

Donde:

- F''_{ehijk} = 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 corregido por no respuesta de la j-ésima vivienda seleccionada, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.
- I_{ehijk} = función indicadora con valor 1 ó 0; esta función toma el valor uno si en el k-ésimo hogar en muestra, 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 en caso contrario.

9.1.3 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:

$$F''_{ehijk\ell} = F''_{ehijk} \frac{\sum_h \sum_i \sum_j F''_{ehijk}}{\sum_h \sum_i \sum_j F''_{ehij} I_{ehijk\ell}}$$

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} = 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.
- $I_{ehijk\ell}$ = función indicadora con valor 1 ó 0; esta función toma el valor uno si la ℓ -ésima persona seleccionada, 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 en caso contrario.

9.2 Ajuste por proyección

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

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

Donde:

- F''_D = factor de expansión corregido por proyección en el dominio D.
- F'_D = factor de expansión corregido por no respuesta en el dominio D.
- $PROy_D$ = población en el dominio D, según proyección.
- $PEXP_D$ = población total a la que expande la encuesta en el dominio D.
- D = nivel de desagregación de la población al que se realiza el ajuste por proyección.

9.3 Ajuste por no respuesta a nivel persona para informantes de 12 años o más de edad

El factor persona se ajusta para los informantes de 12 años o más de edad a nivel estrato, mediante la siguiente expresión:

$$F''_{ehi} = F'_D \frac{Q_{eh} + Q^*_{eh}}{Q_{eh}}$$

Donde:

- F'_D = factor vivienda ajustado por proyección de población.
- Q_{eh} = número de informantes de 12 años o más de edad con respuesta en el h-ésimo estrato, en la e-ésima entidad.
- Q^*_{eh} = número de informantes de 12 años o más de edad sin respuesta en el h-ésimo estrato, en la e-ésima entidad.

9.4 Ajuste para la muestra de población hablante de lengua indígena

El cálculo de los factores de expansión para la población de 12 años o más de edad hablante de lengua indígena se elaboró conforme al procedimiento descrito antes con el fin de dar resultados del uso del tiempo para esa población. Los factores de expansión se ajustaron de acuerdo con el siguiente procedimiento:

$$F_{ehi}^i = \alpha F_{ehi}^*$$

$$F_{ehi}^{ii} = (1 - \alpha) F_{ehi}^{**}$$

$$\alpha = \frac{n_1}{n_1 + n_2}, \quad \alpha + (1 - \alpha) = 1$$

Donde:

- F_{ehi}^* = factor de expansión de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad a nivel persona de 12 años o más de edad de la población hablante de lengua indígena, en la muestra tradicional.
- F_{ehi}^{**} = factor de expansión de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad a nivel persona de 12 años o más de edad de la población hablante de lengua indígena, en la muestra adicional.
- F_{ehi}^i = factor de expansión ajustado por el ponderador muestral de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad a nivel persona de 12 años o más de edad de la población hablante de lengua indígena, en la muestra tradicional.
- F_{ehi}^{ii} = factor de expansión ajustado por el ponderador muestral de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad a nivel persona de 12 años o más de edad de la población hablante de lengua indígena, en la muestra adicional.
- n_1 = número de personas de 12 años o más de edad que hablan lengua indígena en la muestra tradicional.
- n_2 = número de personas de 12 años o más de edad que hablan lengua indígena en la muestra adicional.
- α = factor de ponderación muestral para unir las dos muestras de población de 12 o más años de edad que hablan lengua indígena.

10. 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_{ehi}^{UA} = factor de expansión final, 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 el ℓ -ésimo hogar, 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_{ehi}^{CU} = factor de expansión final 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 X en el ℓ -ésimo hogar, en la s -ésima vivienda, en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad del dominio complemento urbano.

X_{ehi}^R = factor de expansión final 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 X en el ℓ -ésimo hogar, 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} .

La construcción de los estimadores para muestra de la población de 12 años o más de edad hablante de lengua indígena, se hace de forma análoga al caso nacional:

$$\hat{X} = \sum_e \sum_h \sum_i F_{ehi}^i \left(\sum_j X_{ehij}^i \right) + \sum_e \sum_h \sum_i F_{ehi}^{ii} \left(\sum_j X_{ehij}^{ii} \right)$$

Donde:

F_{ehi}^i = factor de expansión final, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad de la muestra tradicional.

X_{ehij}^i = valor observado de la característica de interés X en la j-ésima persona de 12 años y más hablante de lengua indígena, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad de la muestra tradicional.

F_{ehi}^{ii} = factor de expansión final de la i-ésima UPM, del h-ésimo estrato, de la en la e-ésima entidad de la muestra adicional.

X_{ehij}^{ii} = valor observado de la característica X en la j-ésima persona de 12 años y más hablante de lengua indígena, en la i-ésima UPM, en el h-ésimo estrato, en la e-ésima entidad de la muestra adicional.

11. Estimación de errores de muestreo

Para la evaluación de los errores de muestreo de las principales estimaciones nacionales y por dominio se usó el método de Conglomerados Últimos³, 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, 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} .

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

Donde:

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

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

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

n_{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:

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

En el caso de la estimación de varianza para las estimaciones de la muestra de la población de 12 años y más de edad hablante de lengua indígena, los cálculos se realizan utilizando la misma fórmula para estimar su precisión.

Las estimaciones de la desviación estándar (DE), efecto de diseño (DEFF) y coeficiente de variación (CV) se calculan mediante las siguientes expresiones:

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

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

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

Donde:

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

$\hat{V}(\hat{\theta})_{mas}$ = estimador de la varianza bajo 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:

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

Donde α es el nivel de significancia.

12. Homologación de la Semaforización para los Umbrales de Indicadores de precisión estadística⁴

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 CV, así como su semaforización de estos.

⁴ La fuente de esta información está basada en el documento del comité de Aseguramiento de la calidad depositado en el siguiente sitio http://intranet.ineqi.org.mx/calidad/wp-content/uploads/2017/02/Homologacion_de_umbrales.pdf

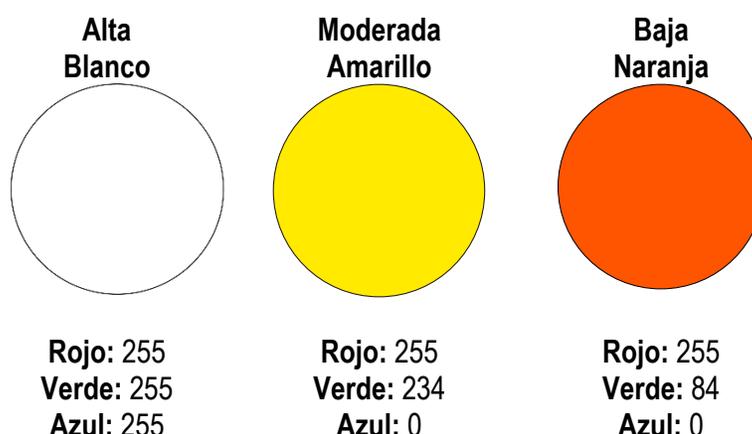
Umbrales aprobados para la cobertura del CV

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, intervalo de confianza y coeficiente de variación (CV). 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

Mnemónico	Descripción
Proporción de Población	
PPSSNOSP	Que tiene derecho a recibir servicios médicos en alguna institución de salud pública o privada excepto seguro popular.
PPDER_SS	Derechohabiente a servicios de salud.
PDP3A14A	De 3 a 14 años de edad que asiste a la escuela.
PDP15A24A	De 15 a 24 años de edad que asiste a la escuela.
PDP8A14ALF	De 8 a 14 años de edad que saben leer o escribir.
PDP15YM_SE	De 15 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.
Proporción de Viviendas Particulares Habitadas	
PVIVSINH	Que no tienen hacinamiento.
PVPH_PISDT	Que tienen piso de cemento o firme, madera, mosaico u otro material.
PVPH2YMASD	Que usan para dormir entre 2 y 25 cuartos.
PVPH_2MASC	Que tienen más de un cuarto.
PVPH3YMASC	Que tienen entre 3 y 25 cuartos.
PVPH_C_ELE	Que disponen de luz eléctrica.
PVPHAGUADV	Que tienen disponibilidad de agua entubada dentro de la vivienda, o fuera de 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.
PVDRERED	Que disponen de drenaje conectado a la red pública.
PVEXCAGU	Que disponen de excusado con descarga directa de agua.
PVPH_CSERV	Que disponen de luz eléctrica, agua entubada dentro o fuera de la vivienda, pero dentro del terreno, así como drenaje.
PSIN_HASIN	Que no se encuentran en situación de hacinamiento a nivel manzana.
Proporción de Viviendas Particulares Habitadas que disponen de:	
PVPH_TV	Televisor.
PVPH_AUTOM	Automóvil o camioneta.
PVPH_CEL	Teléfono celular.
PVCELFIJ	Teléfono celular y teléfono fijo.
PV4ELEC	Radio, televisor, refrigerador y lavadora.
PVRADTEL	Radio y televisor.
PVPHCBIEN	Todos los bienes.

Clave	Entidad		Urbano alto	Complemento urbano	Rural	Total
	Nombre					
01	Aguascalientes		635	140	160	935
02	Baja California		700	160	83	943
03	Baja California Sur		335	480	118	933
04	Campeche		475	240	220	935
05	Coahuila de Zaragoza		680	160	101	941
06	Colima		575	260	100	935
07	Chiapas		220	280	437	937
08	Chihuahua		680	120	140	940
09	Ciudad de México		920	0	20	940
10	Durango		515	140	278	933
11	Guanajuato		420	240	278	938
12	Guerrero		300	260	379	939
13	Hidalgo		195	320	421	936
14	Jalisco		520	300	121	941
15	México		700	120	124	944
16	Michoacán de Ocampo		240	400	299	939
17	Morelos		515	280	139	934
18	Nayarit		335	320	282	937
19	Nuevo León		740	140	60	940
20	Oaxaca		160	300	477	937
21	Puebla		380	300	260	940
22	Querétaro		515	160	266	941
23	Quintana Roo		695	140	102	937
24	San Luis Potosí		415	200	322	937
25	Sinaloa		460	240	246	946
26	Sonora		595	220	119	934
27	Tabasco		195	360	380	935
28	Tamaulipas		680	140	121	941
29	Tlaxcala		555	220	161	936
30	Veracruz de Ignacio de la Llave		320	280	343	943
31	Yucatán		495	300	140	935
32	Zacatecas		255	300	381	936
Total			15 420	7 520	7 078	30 018

Distribución de la muestra adicional en viviendas por entidad y dominio ENUT 2019 **Cuadro 3**

Entidad		Dominio			
Cve.	Nombre	Urbano alto	Complemento urbano	Rural	Total general
07	Chiapas		75	275	350
08	Chihuahua			50	50
12	Guerrero		55	145	200
13	Hidalgo		20	80	100
15	México	50			50
16	Michoacán de Ocampo		35	15	50
20	Oaxaca		110	230	340
21	Puebla	100		100	200
24	San Luis Potosí		10	90	100
30	Veracruz de Ignacio de la Llave	150		200	350
31	Yucatán		145	65	210
Total		300	450	1250	2 000