

Encuesta Nacional de Ingresos y Gastos de los Hogares Estacional 2020

ENIGH

Primer trimestre

Documento diseño muestral



Instituto Nacional de Estadística y Geografía

Encuesta Nacional de Ingresos y Gastos de los Hogares Estacional 2020

ENIGH

Primer trimestre

Documento diseño muestral



Obras complementarias publicadas por el INEGI sobre el tema:

Encuesta Nacional de Ingresos y Gastos de los Hogares 2020 (ENIGH 2018) Documento Diseño Muestral.

Catalogación en la fuente INEGI:

640.4201 Encuesta Nacional de Ingresos y Gastos de los Hogares Estacional (2020).
Encuesta Nacional de Ingresos y Gastos de los Hogares Estacional 2020 :
ENIGH : primer trimestre : documento diseño muestral / Instituto Nacional de
Estadística y Geografía.-- México : INEGI, c2021.

iii, 12 p.

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

Conociendo México

800 111 4634

www.inegi.org.mx

atencion.usuarios@inegi.org.mx



INEGI Informa



@INEGI_INFORMA

DR © 2021, **Instituto Nacional de Estadística y Geografía**

Edificio Sede

Avenida Héroe de Nacozari Sur 2301

Fraccionamiento Jardines del Parque, 20276 Aguascalientes,

Aguascalientes, Aguascalientes, entre la calle INEGI,

Avenida del Lago y Avenida Paseo de las Garzas

Presentación

El **Instituto Nacional de Estadística y Geografía (INEGI)** presenta diversos documentos metodológicos en los que se exponen las principales características de la **Encuesta Nacional de Ingresos y Gastos de los Hogares Estacional (ENIGH Estacional) 2020**.

En este documento, correspondiente al **Diseño muestral**, se describen las actividades que constituyen el proceso de obtención y tratamiento estadístico de la muestra, que inicia con la determinación de la cobertura y dominios de interés, para los cuales se pueden obtener estimaciones estadísticamente confiables, en función de los cuales se calcula el tamaño de la muestra y se determina su distribución, seguida de la selección de las unidades de muestreo.

Como parte del tratamiento, se describe el cálculo de las probabilidades de selección y su empleo para asignar el peso que debe tener cada unidad seleccionada. Finalmente se presentan las expresiones empleadas para obtener las estimaciones y sus medidas de calidad.

Índice

Diseño muestral	1
1. Objetivo de la encuesta	1
2. Población objetivo	1
3. Cobertura geográfica	1
a) Trimestral	1
b) Anual	1
4. Marco de la encuesta	1
4.1 Formación de las unidades primarias de muestreo (UPM)	1
4.1.1 En urbano alto	1
4.1.2 En complemento urbano	2
4.1.3 En rural	2
4.2 Estratificación	2
4.3 Selección de las UPM de la muestra maestra	2
5. Tamaño de la muestra	3
6. Afijación de la muestra	3
7. Selección de la muestra	4
7.1 En urbano alto	4
7.2 En complemento urbano	5
7.3 En rural	6
8. Ajuste a los factores de expansión	7
8.1 Ajuste por no respuesta	7
8.2 Ajuste por proyección	7
9. Estimadores	8
10. Estimaciones de errores de muestreo	8
11. Homologación de la Semaforización para los Umbrales de Indicadores de precisión estadística	9
Anexo	11
A. Indicadores empleados en la estratificación de UPM de la muestra maestra.	13
B. Distribución de la muestra en viviendas de la ENIGH Estacional 2020 por entidad federativa y decena.	14

Diseño muestral

1. Objetivo de la encuesta

El objetivo de la ENIGH Estacional 2020 es obtener información estadística de la distribución del ingreso y del gasto de los hogares en cuanto a su monto, procedencia y distribución a lo largo de un año.

2. Población objetivo

La encuesta está dirigida a los hogares de las viviendas particulares del territorio nacional.

3. Cobertura geográfica

La encuesta está diseñada para dar resultados por periodo y dominio de estudio:

a) Trimestral:

- Nacional, para el gasto trimestral e ingreso corriente del hogar.

b) Anual:

- Nacional, para los rubros de gasto e ingreso corriente del hogar.
- Por entidad, para el agregado del gasto e ingreso corriente del hogar.

4. Marco de la encuesta

El diseño de la submuestra para la ENIGH Estacional 2020 se caracteriza por ser probabilístico; en consecuencia, los resultados obtenidos de la encuesta se generalizan a toda la población. A la vez, el diseño es bietápico, estratificado y por conglomerados, donde la unidad última de selección es la vivienda y la unidad de observación es el hogar.

La submuestra de ENIGH Estacional 2020 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. El diseño del MMM se construyó de la siguiente manera:

4.1 Formación de las unidades primarias de muestreo (UPM)

Inicialmente, se construye el conjunto de UPM que cubrirá el territorio nacional.

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

4.1.1 En urbano alto

El tamaño mínimo de una UPM es de 80 viviendas habitadas y el máximo es de 160. Pueden estar formadas por:

- 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.

¹ Área Geoestadística Básica.

- La unión de dos o más manzanas contiguas de diferentes localidades, que pertenezcan al mismo tamaño de localidad.

4.1.2 En complemento urbano

El tamaño mínimo de una UPM es de 160 viviendas habitadas y el máximo es de 300. Pueden estar formadas por:

- Una manzana.
- La unión de dos o más manzanas contiguas de la misma AGEB.
- La unión de dos o más manzanas contiguas de diferentes AGEB de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes AGEB y localidades, pero del mismo municipio.

4.1.3 En rural

El tamaño mínimo de una UPM es de 160 viviendas habitadas y el máximo es de 300. Pueden estar formadas por:

- Una AGEB.
- Parte de una AGEB.
- La unión de dos o más AGEB colindantes del mismo municipio.
- La unión de una AGEB con una parte de otra AGEB colindante del mismo municipio.

4.2 Estratificación

Una vez construido el conjunto de UPM, se agrupan aquellas con características similares, en términos estadísticos 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 ²	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 las mismas, expresadas por medio de 34 indicadores* construidos con información del Censo de Población y Vivienda 2010, para lo cual se emplearon métodos estadísticos multivariados.

De esta forma, cada UPM fue clasificada en un único estrato geográfico y uno sociodemográfico. Como resultado, se obtuvo un total de 683 estratos en todo el ámbito nacional.

4.3 Selección de las UPM de la muestra maestra

Las UPM de la muestra maestra fueron seleccionadas por medio de un muestreo con probabilidad proporcional al tamaño, esto es:

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

² La descripción de zonas del ámbito urbano alto se muestra en el anexo B.

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

Donde:

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

S_m = 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 de las encuestas en hogares.

5. Tamaño de la muestra

Para el cálculo del tamaño de la muestra de la ENIGH Estacional 2020 se consideró como variable de referencia el promedio del ingreso corriente por hogar. La expresión utilizada fue la siguiente:

$$n = \frac{z^2 s^2 DEFF}{r^2 \bar{X}^2 (1 - tnr) PHV}$$

Donde:

n = tamaño de la muestra.

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

s^2 = estimación de la varianza poblacional de la variable de interés.

\bar{X} = estimación del promedio de la variable de interés.

$DEFF$ = 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.

r = error relativo máximo aceptable.

tnr = tasa de no respuesta máxima esperada.

PHV = promedio de hogares por vivienda.

El tamaño de muestra para la ENIGH Estacional 2020 se determinó considerando el promedio por hogar del ingreso corriente nacional de 49 610.29,** una varianza poblacional de 4 261 792 938.69 para esta variable, un nivel de confianza de 90%, un error relativo de 3.9%, un efecto de diseño de 3.5, una tasa de no respuesta de 20% y un promedio de hogares por vivienda de 1.017, obteniendo un tamaño de muestra de 13 253 viviendas, el cual por cuestiones operativas se ajustó a 13 822 viviendas para levantar en el primer trimestre del 2020.

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:

** Estos parámetros fueron obtenidos de la ENIGH-2018.

$$k_{eh} = \frac{n_{eh}}{b}$$

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 promedio de viviendas a seleccionar por UPM.

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

7. 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 realizó una selección independiente para cada entidad y estrato, el procedimiento varió dependiendo del ámbito, como se especifica a continuación:

7.1 En urbano alto

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

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 ENIGH Estacional 2020.
2. En cada UPM se seleccionaron seis viviendas con igual probabilidad para la ENIGH Estacional 2020.

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 ENIGH Estacional 2020 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:

$$P_{3ehi} = \frac{6}{m_{ehi}^*}$$

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:

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

$$= \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{6}{m_{ehi}^*} = \frac{6 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*}$$

Su factor de expansión³ está dado por:

$$F_{ehi} = \frac{m_{eh} m_{ehi}^*}{6 k_{eh}^* m_{ehi}}$$

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 ENIGH Estacional 2020, 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 ENIGH Estacional 2020.

7.2 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 ENIGH Estacional 2020.
2. En cada UPM seleccionada, se eligieron seis viviendas 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 ENIGH Estacional 2020 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:

$$P_{3ehi} = \frac{6}{m_{ehi}^*}$$

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:

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

³ El factor de expansión se define como el inverso de la probabilidad de selección. En la Norma Técnica del Proceso de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía, el factor de expansión se denomina Ponderador.

$$= \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{6}{m_{ehi}^*} = \frac{6 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*}$$

Su factor de expansión está dado por:

$$F_{ehi} = \frac{m_{eh} m_{ehi}^*}{6 k_{eh}^* m_{ehi}}$$

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 ENIGH Estacional 2020, 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 ENIGH Estacional 2020.

7.3 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 ENIGH Estacional 2020.
2. En cada UPM seleccionada, se eligieron, seis viviendas 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 ENIGH Estacional 2020 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:

$$P_{3ehi} = \frac{6}{m_{ehi}^*}$$

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

$$\begin{aligned} P_{ehi} &= P_{1ehi} * P_{2ehi} * P_{3ehi} \\ &= \frac{k_{eh} m_{ehi}}{m_{eh}} \frac{k_{eh}^*}{k_{eh}} \frac{6}{m_{ehi}^*} = \frac{6 k_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*} \end{aligned}$$

En consecuencia, su factor de expansión está definido por:

$$F_{ehi} = \frac{m_{eh} m_{ehi}^*}{6k_{eh}^* m_{ehi}}$$

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 ENIGH Estacional 2020, 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 ENIGH Estacional 2020.

8. Ajuste a los factores de expansión

Los factores de expansión elaborados conforme al procedimiento antes descrito se ajustan para los siguientes conceptos:

8.1 Ajuste por no respuesta

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

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

Donde:

- F'_{ehi} = factor de expansión corregido por no respuesta 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.

8.2 Ajuste por proyección

Los factores ajustados por no respuesta se corrigen, a fin de asegurar que en cada dominio de interés de la encuesta se obtenga la población total determinada por la proyección de población determinada por el Consejo Nacional de Población (CONAPO), 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 la proyección de CONAPO.
- $PEXP_D$ = población total a la que expande la encuesta en el dominio D.
- D = es el nivel de desagregación de la población al que se realiza el ajuste por proyección

9. Estimadores

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

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

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_{ehisl}^{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_{ehisl}^{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.
 F_{ehi}^R = factor de expansión final de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad del dominio rural.
 X_{ehisl}^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} .

10. Estimaciones de errores de muestreo

Para la evaluación de los errores de muestreo de las principales estimaciones estatales y nacionales se usó el método de "Conglomerados Últimos",⁴ basado en que la mayor contribución a la varianza de un estimador, en un diseño polietápico, es la que se presenta entre las unidades primarias de muestreo (UPM). El término "Conglomerados Últimos" se utiliza para denotar el total de unidades en muestra de una unidad primaria de muestreo.

Para obtener las precisiones de los estimadores de razón, 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} .

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

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

Estas definiciones son análogas para la variable de estudio Y.

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

⁴ Vease Hansen, M.H. Horwitz, W.N. y Madow, W.G, *Sample Survey Methods and Theory*, (1953), Vol. 1 página 242.

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

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

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

Donde:

- $\hat{\theta}$ = estimador del parámetro poblacional θ .
- $\hat{V}(\hat{\theta})_{mas}$ = estimador de la varianza bajo muestreo aleatorio simple.
- $\hat{V}(\hat{\theta})$ = estimador de la varianza obtenida bajo el diseño 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.

11. Homologación de la Semaforización para los Umbrales de Indicadores de precisión estadística⁵

Para facilitar la interpretación de las precisiones estadísticas de la información pública en tabulados, el comité de Aseguramiento de la calidad, en la cuarta sección 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.

Umbrales aprobados para la cobertura del CV

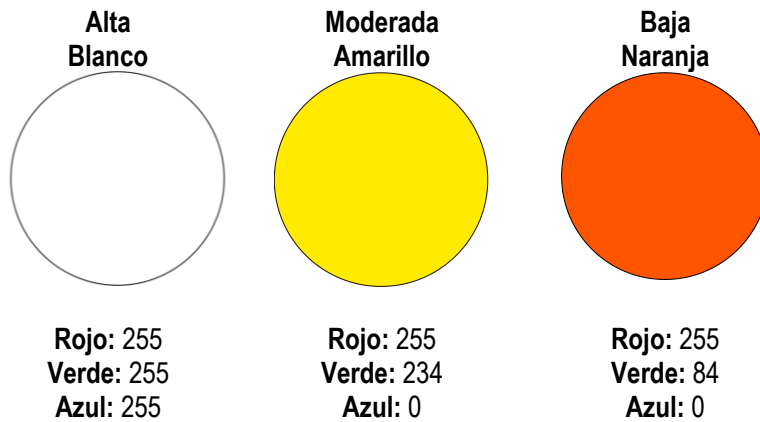
		Viviendas/Hogares/Otras unidades diferentes a las económicas
Interpretación	Semaforización	DGES/DGEGSPJ
Alta	Blanco	[0%, 15%)
Moderada	Amarillo	[15%, 30%)
Baja	Naranja oscuro	>=30%

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:

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

Parámetros RGB para la semaforización del coeficiente de variación.



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

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

Nivel de precisión de las estimaciones:

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

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

Baja, CV de 30% en adelante

Anexo

MNEMÓNICO	DESCRIPCIÓN
Población	
PPSSNOSP	Que tiene derecho a recibir servicios médicos en alguna institución de salud pública o privada excepto seguro popular.
PPDER_SS	Derechohabiente a servicios de salud.
PDP3A14A	De 3 a 14 años de edad que asiste a la escuela.
PDP15A24A	De 15 a 24 años de edad que asiste a la escuela.
PDP8A14ALF	De 8 a 14 años de edad que saben leer o escribir.
PDP15YM_SE	De 15 años o más de edad que aprobaron algún grado de escolaridad diferente al nivel preescolar.
PP15PRI_CO	De 15 años o más de edad que tienen como máxima escolaridad 6 grados aprobados en primaria.
PP15SEC_CO	De 15 años o más de edad que tienen como máxima escolaridad 3 grados aprobados en secundaria.
PGDO_ESC	Grado promedio de escolaridad.
PPEA	De 12 años y más que trabajaron; tenían trabajo pero no trabajaron o; buscaron trabajo en la semana de referencia.
PPEA_F	Femenina de 12 años y más que trabajaron; tenían trabajo pero no trabajaron o; buscaron trabajo en la semana de referencia.
PTASAOcupA	Tasa de ocupación.
TOCU12A17	No ocupada de 12 a 17 años de edad entre la población de 12 a 17 años de edad.
PPOMAYED	Ocupada de 18 y más años de edad entre la población ocupada.
Viviendas particulares habitadas	
PVIVSINH	Que no tienen hacinamiento.
PVPH_PISDT	Que tienen piso de cemento o firme, madera, mosaico u otro material.
PVPH2YMASD	Que usan para dormir entre 2 y 25 cuartos.
PVPH_2MASC	Que tienen más de un cuarto.
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	Total de manzanas sin hacinamiento.
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.

Distribución de la muestra en viviendas de la ENIGH Estacional 2020 por entidad federativa y decena*

Anexo B

Clave	Entidad Nombre	Decenas									Total
		01	02	03	04	05	06	07	08	09	
00	Estados Unidos Mexicanos	1 527	1 543	1 537	1 531	1 556	1 543	1 554	1 513	1 518	13 822
01	Aguascalientes	48	47	47	47	48	47	52	48	48	432
02	Baja california	48	48	42	47	48	47	53	52	48	433
03	Baja california sur	48	47	48	48	47	51	47	42	47	425
04	Campeche	48	46	48	42	56	52	47	47	47	433
05	Coahuila de Zaragoza	47	53	47	48	48	42	48	47	48	428
06	Colima	48	42	48	47	48	48	47	48	47	423
07	Chiapas	51	52	61	47	47	51	50	45	44	448
08	Chihuahua	47	47	54	47	46	47	47	48	48	431
09	Ciudad de México	48	48	48	48	48	48	48	48	42	426
10	Durango	47	47	48	47	46	48	54	48	46	431
11	Guanajuato	53	47	47	57	48	47	47	45	47	438
12	Guerrero	42	47	47	46	47	48	45	51	56	429
13	Hidalgo	50	51	40	45	51	47	53	52	50	439
14	Jalisco	48	48	48	47	48	48	52	41	48	428
15	México	48	48	48	47	47	48	52	48	47	433
16	Michoacán de Ocampo	46	46	52	45	51	47	55	47	46	435
17	Morelos	48	47	48	48	48	41	54	48	48	430
18	Nayarit	47	47	47	48	47	48	47	47	51	429
19	Nuevo león	47	48	48	48	48	48	48	48	48	431
20	Oaxaca	52	50	45	50	56	52	38	45	47	435
21	Puebla	47	56	52	47	47	47	46	47	42	431
22	Querétaro Arteaga	47	53	57	48	41	46	48	47	47	434
23	Quintana roo	47	48	48	48	48	53	47	48	47	434
24	San Luis potosí	40	47	47	51	56	52	46	47	46	432
25	Sinaloa	51	52	41	47	47	52	46	47	46	429
26	Sonora	48	47	47	52	47	47	47	41	52	428
27	Tabasco	46	46	48	50	55	55	51	46	45	442
28	Tamaulipas	47	48	47	42	47	48	48	53	48	428
29	Tlaxcala	48	47	48	48	48	42	48	48	53	430
30	Veracruz de Ignacio de la Llave	46	42	47	56	55	47	46	46	46	431
31	Yucatán	47	48	48	47	48	52	52	47	47	436
32	Zacatecas	52	53	46	46	44	47	45	51	46	430

*Por la pandemia solo se lograron levantar 13 822 viviendas en el primer trimestre.