

Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares 2021 ENDUTIH

Documento preliminar del diseño
de la muestra



DISEÑO MUESTRAL

ENCUESTA NACIONAL SOBRE DISPONIBILIDAD Y USO DE
TECNOLOGÍAS DE LA INFORMACIÓN EN LOS HOGARES
2021 (ENDUTIH 2021)

Índice	Página
1. Objetivo de la encuesta	1
2. Población objetivo	1
3. Cobertura geográfica	1
4. Marco de la encuesta	1
4.1 Formación de las unidades primarias de muestreo (UPM)	1
a) En urbano alto	1
b) En complemento urbano	2
c) En rural	2
4.2 Estratificación	2
4.3 Selección de las UPM de la muestra maestra	3
5. Tamaño de la muestra	3
6. Distribución de la muestra	4
7. Selección de la muestra	4
7.1 En urbano alto	4
7.2 En complemento urbano	6
7.3 En rural	7
8. Ajuste a los factores de expansión	9
8.1 Ajuste por no respuesta	9
8.1.1 Ajuste por no respuesta a nivel vivienda	9
8.1.2 Ajuste por no respuesta a nivel hogar	9
8.1.3 Ajuste por no repuesta a nivel persona	10
8.2 Ajuste por proyección	10
9. Estimadores	11
10. Estimación de los errores de muestreo	11
 Anexos	
- Indicadores empleados en la estratificación de UPM de la muestra maestra. (Cuadro 1).	13
- Distribución de la muestra en viviendas seleccionadas para la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares 2021 (ENDUTIH 2021) por entidad (Cuadro 2).	14
- Disponibilidad y Uso de Tecnologías de la Información en los Hogares 2021 (ENDUTIH 2021) por entidad según ámbito urbano y rural (Cuadro 3).	15

1. Objetivo de la encuesta

Generar información estadística que permita conocer la disponibilidad y el uso de las tecnologías de la información y las comunicaciones en los hogares y por los individuos de seis o más años que viven en los dominios de interés.

2. Población objetivo

La encuesta está dirigida a las personas de seis o más años que residen permanentemente en viviendas particulares ubicadas en el territorio nacional a la fecha del levantamiento.

3. Cobertura geográfica

La encuesta está diseñada para dar resultados para los siguientes dominios de estudio:

- Nacional
- Nacional urbano
- Nacional rural
- Entidad

4. Marco de la encuesta

El diseño de la muestra para la ENDUTIH 2021 se caracteriza por ser probabilístico; en consecuencia, los resultados obtenidos de la encuesta se generalizan a toda la población. A la vez, el diseño es trietápico, estratificado y por conglomerados, donde la unidad última de selección es la persona que al momento de la entrevista tenga seis años cumplidos o más de edad.

Para la selección de la muestra de la encuesta se utilizó la Muestra Maestra también conocida como Marco Nacional de Viviendas 2012, este último, construido a partir de la información cartográfica y demográfica que se obtuvo del Censo de **Población y Vivienda 2010. A partir de la “Muestra Maestra” se seleccionan las submuestras para todas las encuestas** en viviendas que realiza el INEGI; el diseño de la muestra maestra es probabilístico, estratificado, unietápico y por conglomerados; estos últimos también se consideran unidades primarias de muestreo, pues es en ellos donde se seleccionan, en una segunda etapa, las viviendas que integran las muestras de las diferentes encuestas. Las Unidades Primarias de Muestreo (UPM) se forman de la siguiente manera:

4.1. Formación de las unidades primarias de muestreo (UPM)

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:

a) En urbano alto

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

- Una manzana.
- La unión de dos o más manzanas contiguas de la misma AGEB.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.

b) En complemento urbano

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

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

c) En rural

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

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

4.2. Estratificación

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

La división política del país y la conformación de localidades diferenciadas por su tamaño, forman de manera natural una estratificación geográfica. En cada entidad federativa se distinguen tres ámbitos, divididos a su vez en zonas, como se indica en el siguiente cuadro:

Ámbito	Descripción
Urbano alto	Ciudades con 100 000 o más habitantes
Complemento urbano	De 2 500 a 99 999 habitantes
Rural	Localidades menores de 2 500 habitantes

De manera paralela, se formaron cuatro estratos socioeconómicos en los que se agruparon todas las UPM del país, esta estratificación considera las características sociodemográficas de los habitantes de las viviendas, así como las características físicas y el equipamiento de 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.

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

4.3. Selección de las UPM de la muestra maestra

Las UPM de la muestra maestra fueron seleccionadas con probabilidad proporcional al tamaño, mediante la siguiente expresión:

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

Donde:

U_{ehi} = 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 captada en el Censo de Población y Vivienda 2010.

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

Las UPM seleccionadas forman la muestra maestra, de las cuales son seleccionadas las submuestras de las encuestas en hogares.

5. Tamaño de la muestra

Para el cálculo del tamaño de muestra de la ENDUTIH 2021 se empleó la siguiente expresión:

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

Donde:

n = tamaño de la muestra.

p = estimación de la proporción de interés.

q = 1-p.

r = error relativo máximo aceptable.

z = valor asentado en las tablas estadísticas, de la distribución normal estándar, que garantiza realizar las estimaciones con una confianza prefijada.

DEFF = efecto de diseño definido como el cociente de la varianza en la estimación del diseño utilizado, entre la varianza obtenida considerando un muestreo aleatorio simple para un mismo tamaño de muestra.

tnr = tasa de no respuesta máxima esperada.

PHV = promedio de hogares por vivienda.

Considerando un nivel de confianza del 90%, un efecto de diseño de 4.28 observado en experiencias anteriores, un error relativo máximo esperado de 11.14%, una tasa de no respuesta de 15%, un promedio de hogares por vivienda de 1.02 y una proporción de 1.63%. Con estos parámetros se obtiene un tamaño de muestra mínimo de 65 000 viviendas a nivel nacional.

6. Distribución de la muestra

La distribución de la muestra se realizará 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.
- 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.

En el cuadro 2 se presenta la distribución de la muestra en viviendas para la ENDUTIH 2021 en las 32 entidades.

7. Selección de la muestra

La selección de la muestra se realizó de manera independiente por entidad, dominio y estrato, el procedimiento de selección varió de acuerdo con el dominio.

7.1. En urbano alto

El procedimiento probabilístico general, de selección, se describe en tres etapas:

1. De las k_{eh} UPM seleccionadas en el h-ésimo estrato, en la e-ésima entidad para la muestra maestra², se seleccionaron k_{eh}^* , con igual probabilidad, para la ENDUTIH 2021.
2. En cada UPM seleccionada, se seleccionaron cinco viviendas con igual probabilidad.
3. En cada uno de los hogares registrados dentro de la vivienda seleccionada se seleccionó una persona de seis o más años.

En forma detallada la probabilidad de selección de las UPM, las viviendas y las personas se calculó de la siguiente manera:

- a) La probabilidad de selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra maestra está definida por:

² Del total de UPM que integran el marco de propósitos múltiples se seleccionaron con probabilidad proporcional a su tamaño k_{eh} UPM para localidades de 100 000 y más habitantes.

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

b) La probabilidad de seleccionar, a partir de la muestra maestra, la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de ENDUTIH 2021 está definida por:

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

c) La probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad está definida por:

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

d) La probabilidad de seleccionar el k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

$$P_{4ehijk} = 1,$$

e) La probabilidad de seleccionar la ℓ -ésima persona, del k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

$$P_{5ehijk\ell} = \frac{1}{Q_{ehijk\ell}},$$

Por lo tanto, la probabilidad de selección de una persona en el k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es el producto de las probabilidades de selección:

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

Su factor de expansión³ está dado por:

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

Donde:

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

³ 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.

- 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 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 listado de viviendas actualizado.
- k_{eh}^* = número de UPM a seleccionar para la ENDUTIH 2021 con igual probabilidad de selección, en el h-ésimo estrato, en la e-ésima entidad.
- $Q_{ehijk\ell}$ = número de personas de seis o más años 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.

7.2 En complemento urbano

El procedimiento probabilístico general, de selección, se describe en tres etapas:

1. De las k_{eh} UPM seleccionadas en el h-ésimo estrato, en la e-ésima entidad para la muestra maestra⁴, se seleccionaron k_{eh}^* , con igual probabilidad, para la ENDUTIH 2021.
2. En cada UPM seleccionada, se seleccionaron 20 viviendas con igual probabilidad.
3. En cada uno de los hogares registrados dentro de la vivienda seleccionada se seleccionó una persona de seis o más años.

En forma detallada la probabilidad de selección de las UPM, las viviendas y las personas se calculó de la siguiente manera:

- a) La probabilidad de selección de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra maestra está definida por:

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

- b) La probabilidad de seleccionar, a partir de la muestra maestra, la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad para la muestra de ENDUTIH 2021 está definida por:

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

- c) La probabilidad de seleccionar la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad está definida por:

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

- d) La probabilidad de seleccionar el k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad es:

$$P_{4ehijk} = 1,$$

⁴ Del total de UPM que integran el marco de propósitos múltiples se seleccionaron con probabilidad proporcional a su tamaño k_{eh} UPM para localidades de 100 000 y más habitantes.

e) La probabilidad de seleccionar la ℓ -ésima persona, del k -ésimo hogar, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad es:

$$P_{5ehijk\ell} = \frac{1}{Q_{ehijk\ell}},$$

Por lo tanto, la probabilidad de selección de una persona del k -ésimo hogar, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad es el producto de las probabilidades de selección:

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

Su factor de expansión está dado por:

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

Donde:

- k_{eh} = número de UPM seleccionadas en el h -ésimo estrato, en la e -ésima entidad, en la muestra maestra.
- 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 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 listado de viviendas actualizado.
- k_{eh}^* = número de UPM a seleccionar para la ENDUTIH 2021 con igual probabilidad de selección, en el h -ésimo estrato, en la e -ésima entidad.
- $Q_{ehijk\ell}$ = número de personas de seis o más años 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.

7.3 En rural

El procedimiento probabilístico general, de selección, se describe en tres etapas:

1. De las k_{eh} UPM seleccionadas en el h -ésimo estrato, en la e -ésima entidad para la muestra maestra, se seleccionaron k_{eh}^* , con igual probabilidad, para la ENDUTIH 2021.
2. En cada UPM seleccionadas, se seleccionaron cuatro segmentos de cinco viviendas con igual probabilidad.
3. En cada uno de los hogares registrados dentro de la vivienda seleccionada se seleccionó una persona de seis o más años.

En forma detallada la probabilidad de selección de las UPM, las viviendas y las personas se calculó de la siguiente manera:

- a) La probabilidad de selección de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad para la muestra maestra está definida por:

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

- b) La probabilidad de seleccionar, a partir de la muestra maestra, la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad para la muestra de ENDUTIH 2021 está definida por:

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

- c) La probabilidad de seleccionar la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad está definida por:

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

- d) La probabilidad de seleccionar el k -ésimo hogar, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad es:

$$P_{4ehijk} = 1,$$

- e) La probabilidad de seleccionar la ℓ -ésima persona, del k -ésimo hogar, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad es:

$$P_{5ehijk\ell} = \frac{1}{Q_{ehijk\ell}},$$

Por lo tanto, la probabilidad de selección de una persona del k -ésimo hogar, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato, de la e -ésima entidad es el producto de las probabilidades de selección:

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

Su factor de expansión está dado por:

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

Donde:

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

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 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 listado de viviendas actualizado.

K_{eh}^* = número de UPM a seleccionar para la ENDUTIH 2021 con igual probabilidad de selección, en el h-ésimo estrato, en la e-ésima entidad.

$Q_{ehijk\ell}$ = número de personas de seis o más años 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.

8. Ajuste a los factores de expansión

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

8.1. Ajuste por no respuesta

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

8.1.1 Ajuste por no respuesta a nivel vivienda

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

$$F'_{ehij} = F_{ehij} \frac{\sum_{i \in h} \sum_{j \in i} F_{ehij}}{\sum_{i \in h} \sum_{j \in i} F_{ehij} I_{ehij}}$$

Donde:

F'_{ehij} = factor de expansión corregido por no respuesta de la j-ésima vivienda 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.

8.1.2 Ajuste por no respuesta a nivel hogar

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

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

Donde:

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

F'_{ehijk} = factor de expansión del k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

I_{ehijk} = función indicadora con valor 1 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.

8.1.3 Ajuste por no repuesta a nivel persona

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

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

Donde:

$F'''_{ehijk\ell}$ = factor de expansión corregido por no respuesta de la ℓ -ésima persona seleccionada, del k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

$F''_{ehijk\ell}$ = factor de expansión para la ℓ -ésima persona seleccionada, del k-ésimo hogar, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la e-ésima entidad.

$I_{ehijk\ell}$ = función indicadora con valor 1 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.

8.2 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 estimación de población referida al punto medio del levantamiento, mediante la siguiente expresión:

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

Donde:

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

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

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

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

D = nivel de desagregación de la población al que se realiza la estimación de población.

9. Estimadores

El estimador del total de la característica X es:

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

Donde:

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

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

$X_{ehis\ell}^{CU}$ = valor observado de la característica de interés X en la persona elegida, 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 complemento urbano.

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

$X_{ehis\ell}^R$ = valor observado de la característica de interés X en la persona elegida, 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 utilizará 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. Estimación de los errores de muestreo

Para la evaluación de los errores de muestreo de las principales estimaciones del agregado nacional se usará el método de Conglomerados Últimos⁵, basado en que la mayor contribución a la varianza de un estimador, en un diseño multietápico 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 los errores de muestreo de los estimadores de razón, conjuntamente al método de Conglomerados Últimos se aplicará el método de *Series de Taylor*, obteniéndose la siguiente fórmula para estimar la precisión de \hat{R} :

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

Donde:

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

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

⁵ Véase Hansen, M. H. Horwitz, W.N. y Madow, W.G., *Sample Survey Methods and Theory*, (1953) Vol. 1 pág. 242.

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

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

\hat{Y}_{NAL}^2 = 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:

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

Las estimaciones de la desviación estándar (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 muestreo aleatorio simple.

$\hat{V}(\hat{\theta})$ = estimador de la varianza bajo el diseño de muestreo descrito en este documento.

Finalmente, el intervalo de confianza $I_{1-\alpha}$ al 100(1- α)%, se calcula de la siguiente forma:

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

Donde α es el nivel de significancia.

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.
PVPH3YMASD	Que tienen entre 3 y 25 cuartos.
PVPH_C_ELE	Que disponen de luz eléctrica.
PVPHAGUADV	Que tienen disponibilidad de agua entubada dentro de la vivienda, o fuera de la vivienda pero dentro del terreno.
PVPH_EXCSA	Que tienen excusado, retrete, sanitario, letrina u hoyo negro.
PVPHDRENAJ	Que tienen drenaje conectado a la red pública, fosa séptica, barranca, grieta, río, lago o mar.
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.

Distribución de la muestra en viviendas seleccionadas para la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares 2021 (ENDUTIH 2021) por entidad

Cuadro 2

Clave	Entidad	UPM	Viviendas
01	Aguascalientes	307	2 031
02	Baja california	331	2 031
03	Baja california Sur	208	2 031
04	Campeche	256	2 033
05	Coahuila de Zaragoza	325	2 033
06	Colima	289	2 031
07	Chiapas	173	2 057
08	Chihuahua	322	2 030
09	Ciudad de México	405	2 035
10	Durango	268	2 037
11	Guanajuato	239	2 040
12	Guerrero	199	2 041
13	Hidalgo	163	2 045
14	Jalisco	269	2 038
15	México	329	2 036
16	Michoacán de Ocampo	181	2 040
17	Morelos	265	2 032
18	Nayarit	208	2 030
19	Nuevo León	344	2 036
20	Oaxaca	154	2 044
21	Puebla	227	2 044
22	Querétaro	265	2 043
23	Quintana Roo	325	2 032
24	San Luis Potosí	238	2 040
25	Sinaloa	253	2 039
26	Sonora	289	2 035
27	Tabasco	160	2 046
28	Tamaulipas	322	2 032
29	Tlaxcala	277	2 035
30	Veracruz de Ignacio de la Llave	206	2 044
31	Yucatán	259	2 031
32	Zacatecas	181	2 038
Estados Unidos Mexicanos		8 237	65 190

Distribución de la muestra en viviendas seleccionadas para la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares 2021 (ENDUTIH 2021) por entidad según ámbito urbano y rural

Cuadro 3

Clave	Entidad	Ámbito		Total
		Urbano	Rural	
01	Aguascalientes	1 670	361	2 031
02	Baja California	1 870	161	2 031
03	Baja California Sur	1 750	281	2 031
04	Campeche	1 550	483	2 033
05	Coahuila de Zaragoza	1 830	203	2 033
06	Colima	1 810	221	2 031
07	Chiapas	1 095	962	2 057
08	Chihuahua	1 730	300	2 030
09	Ciudad de México	2 015	20	2 035
10	Durango	1 430	607	2 037
11	Guanajuato	1 435	605	2 040
12	Guerrero	1 230	811	2 041
13	Hidalgo	1 130	915	2 045
14	Jalisco	1 755	283	2 038
15	México	1 795	241	2 036
16	Michoacán de Ocampo	1 410	630	2 040
17	Morelos	1 710	322	2 032
18	Nayarit	1 430	600	2 030
19	Nuevo León	1 915	121	2 036
20	Oaxaca	990	1054	2 044
21	Puebla	1 495	549	2 044
22	Querétaro	1 470	573	2 043
23	Quintana Roo	1 810	222	2 032
24	San Luis Potosí	1 330	710	2 040
25	Sinaloa	1 510	529	2 039
26	Sonora	1 750	285	2 035
27	Tabasco	1 210	836	2 046
28	Tamaulipas	1 770	262	2 032
29	Tlaxcala	1 670	365	2 035
30	Veracruz de Ignacio de la Llave	1 295	749	2 044
31	Yucatán	1 730	301	2 031
32	Zacatecas	1 210	828	2 038
	Estados Unidos Mexicanos	49 800	15 390	65 190