

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA



Módulo de Condiciones Socioeconómicas

Encuesta Nacional de Ingresos y Gastos de los Hogares 2010

Diseño muestral



INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA

Obras complementarias publicadas por el INEGI sobre el tema:

Módulo de condiciones socioeconómicas de la ENIGH 2008 conociendo la base de datos 2009.
Módulo de condiciones socioeconómicas de la ENIGH 2008 Diseño muestral 2009. Niveles de vida y bienestar. Niveles de vida de la sociedad.

Catalogación en la fuente INEGI:

640.4201 Encuesta Nacional de Ingresos y Gastos de los Hogares (2010).
 Módulo de Condiciones Socioeconómicas : Encuesta Nacional de Ingresos y
 Gastos de los Hogares 2010 : diseño muestral / Instituto Nacional de Estadística y
 Geografía.-- México : INEGI, c2011.

vii, 24 p.

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

Si requiere más información sobre esta obra, favor de contactarnos a través de:

Centros de consulta y comercialización (consulte el domicilio en Internet)

Larga distancia sin costo: 01 800 111 4634

www.inegi.org.mx

atencion.usuarios@inegi.org.mx

Presentación

El **Instituto Nacional de Estadística y Geografía (INEGI)** presenta cuatro documentos metodológicos en los que se exponen las principales características del **Módulo de Condiciones Socioeconómicas de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2010 (MCS 2010)**.

En el documento de **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.

Nota

El Módulo de Condiciones Socioeconómicas de la ENIGH 2010 (MCS 2010), incorpora datos sobre ingresos, salud, educación, seguridad social, calidad de espacios de la vivienda, servicios básicos alimentación y redes sociales recolectados del 21 de agosto al 28 de noviembre de 2010.

El MCS 2010 ofrece resultados a nivel nacional con corte urbano y rural y para cada una de las entidades federativas.

Es importante mencionar que la muestra total seleccionada se compone de 30 169 viviendas provenientes de la ENIGH 2010 y 38 336 viviendas cuyo levantamiento fue financiado por el CONEVAL para recolectar información de ingresos y características sociodemográficas de los hogares; en estas viviendas no se recabó información acerca del gasto de los hogares.

Índice

Introducción	VII
1. Población objetivo, cobertura y diseño	1
1.1 Población objetivo	1
1.2 Cobertura geográfica	1
1.3 Diseño de la muestra	1
2. Marco de la encuesta	3
2.1 Formación de las Unidades Primarias de Muestreo (UPM)	3
2.2 Estratificación	4
3. Esquema de muestreo	5
4. Tamaño de la muestra	7
5. Afijación de la muestra	9
6. Selección de la muestra	11
6.1 En urbano alto	11
6.2 En complemento urbano	12
6.3 En rural	12
7. Ajuste a los factores de expansión	15
7.1 Ajuste por No respuesta	15
7.2 Ajuste por proyección	15
8. Estimadores	17
9. Estimación de las precisiones	19
Anexo	21
A. Indicadores empleados en la estratificación de la muestra maestra por ámbito de estudio	23
B. Distribución de la muestra en viviendas por entidad según dominio de estudio, MCS 2010	24

Introducción

El Módulo de Condiciones Socioeconómicas de la ENIGH 2010 (MCS 2010) constituye un esfuerzo conjunto entre el **Instituto Nacional de Estadística y Geografía (INEGI)** y el **Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL)**, para proporcionar un panorama estadístico de las variables necesarias para la medición multidimensional de la pobreza establecidas en la Ley General de Desarrollo Social.

En el primer capítulo se especifican los criterios de referencia sobre los cuales se formuló el diseño, en el segundo se describe el marco de muestreo, que constituye la infraestructura elaborada por el **INEGI** para la selección de las muestras de las encuestas.

Los capítulos tercero al sexto muestran la metodología utilizada para el cálculo del tamaño de muestra en función de la confianza fijada y el error esperado para la estimación de la variable de referencia, la distribución de la muestra que asegura la representación en la misma de los distintos grupos en los que se divide la población y los métodos de selección de las unidades muestrales empleados en cada ámbito geográfico, con las formulas para el cálculo de las probabilidades de selección, a partir de las cuales se obtienen los factores de expansión, que constituyen el peso que se le da a cada unidad muestral para generalizar los resultados de la muestra a la población.

En el capítulo siete se muestra la manera en la que se ajustan los factores de expansión para corregir el efecto de la No respuesta presentada durante el levantamiento, así como el uso de proyecciones para suavizar los efectos de la aleatoriedad de las muestras en la estimación de la población total.

En los dos últimos capítulos se presentan las expresiones empleadas para el cálculo de los principales estimadores y la medición de su calidad estadística.

1. Población objetivo, cobertura y diseño

En este capítulo se establece de manera precisa cuál es la población objetivo, la cobertura geográfica y el diseño de la muestra para el MCS 2010.

1.1 POBLACIÓN OBJETIVO

El módulo de la ENIGH está dirigido a los hogares del territorio nacional.

1.2 COBERTURA GEOGRÁFICA

El MCS 2010 está diseñado para dar resultados a nivel nacional, con corte urbano y rural, y las 32 entidades del país.

1.3 DISEÑO DE LA MUESTRA

El diseño de la muestra se caracteriza por ser probabilístico; en consecuencia, los resultados obtenidos del módulo 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.

2. Marco de la encuesta

El marco de muestreo que se empleó para el MCS 2010 es el Marco Nacional de Viviendas 2002 del INEGI, construido a partir de la información cartográfica y demográfica que se obtuvo del XII Censo General de Población y Vivienda 2000. Este marco es en realidad una muestra maestra a partir de la cual se seleccionan las muestras para todas las encuestas en viviendas que realiza el Instituto. El diseño del presente marco 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.

2.1 FORMACIÓN DE LAS UNIDADES PRIMARIAS DE MUESTREO (UPM)

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

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.

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.

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.

¹ Área Geoestadística Básica.

2.2 ESTRATIFICACIÓN

La división política del país y la conformación de localidades diferenciadas por su tamaño, forman de manera natural una primera estratificación geográfica.

En cada entidad federativa se distinguen tres ámbitos, divididos a su vez en siete zonas, como se indica en el siguiente cuadro:

Ámbito	Zona	Tamaño de localidad
Urbano alto	01	32 ciudades autorrepresentadas (CA) con 100 000 o más habitantes.
	02	Resto de las ciudades con 100 000 o más habitantes.
Complemento urbano	25	De 50 000 a 99 999 habitantes.
	35	De 15 000 a 49 999 habitantes.
	45	De 5 000 a 14 999 habitantes.
	55	De 2 500 a 4 999 habitantes.
Rural	60	Localidades menores de 2 500 habitantes.

De manera paralela, en una primera etapa se formaron cuatro estratos en los que se agruparon todas las UPM del país. Esta estratificación considera las características sociodemográficas de los habitantes de las viviendas, así como las características físicas y el equipamiento de las mismas, expresadas por medio de 24 indicadores contruidos con información del XII Censo General de Población y Vivienda 2000, para lo cual se emplearon métodos estadísticos multivariados.

En una segunda etapa, cada UPM fue asignada de acuerdo con su estrato geográfico (entidad-ámbito-zona).

En una tercera etapa, al interior de cada zona y estrato (sociodemográfico), algunas de las UPM se sometieron a un nuevo proceso de estratificación con el propósito de tener una mayor diferenciación a ese nivel, para esta estratificación se utilizaron indicadores diferenciados por ámbito.* Como resultado se tiene un total de 888 subestratos en todo el ámbito nacional.

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

3. Esquema de muestreo

El MCS 2010 fué diseñado bajo un esquema de muestreo probabilístico, estratificado, bietápico y por conglomerados.

A continuación se mencionan las definiciones de cada uno de estos métodos:

Probabilístico: Las unidades de muestreo tienen una probabilidad conocida y distinta de cero de ser seleccionadas.

Estratificado: Las unidades de muestreo se clasifican de acuerdo con sus características socioeconómicas.

Bietápico: La unidad última de muestreo (vivienda) es seleccionada en dos etapas.

Por conglomerados: Las unidades de muestreo son conjuntos de unidades elementales, con características heterogéneas en su interior y homogéneas hacia su exterior.

4. Tamaño de la muestra

Para el cálculo del tamaño de muestra del MCS 2010 se consideró como variable de referencia el promedio del ingreso corriente total por hogar. La expresión utilizada fue la siguiente:

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

Donde:

n = tamaño de la muestra.

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

s² = 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.

Fijando un nivel de confianza de 90%, un efecto de diseño de 3.3,** una varianza poblacional de 2 322 211 250** un error relativo máximo aceptable de 3.5%, un promedio de ingreso corriente total por hogar de 44 088.3,** una tasa de No respuesta máxima esperada de 15% y un promedio de hogares por vivienda de 1.02, se determinó una muestra a nivel nacional de 10 045 viviendas, el cual se ajustó a 10 000.

Con el objeto de satisfacer los requerimientos adicionales de las instituciones externas que aportaron recursos para el levantamiento, la muestra final del MCS 2010 fue de 68 505 viviendas. (Ver anexo B).

** Estos parámetros fueron obtenidos de la ENIGH 2008.

5. Afijación de la muestra

Para asignar la muestra en los diferentes dominios de estudio se consideró una muestra básica de 10 000 viviendas distribuidas aproximadamente igual entre las 32 entidades federativas. Al interior de éstas, se asignó de manera proporcional al tamaño de los estratos del Marco Nacional de Viviendas 2002. Con esto se aseguró cubrir la distribución de la muestra ENIGH tradicional.

Con la asignación de la muestra anterior se revisó el tamaño de muestra resultante a nivel nacional para cada una de las 5 regiones derivadas del índice de marginación propuesto por el Consejo Nacional de Población (CONAPO) y se asignó la muestra adicional necesaria para completar en todas las regiones un mínimo de 1774 viviendas. La muestra adicional en cada región, se distribuyó de manera proporcional al tamaño del estrato geográfico entidad-zona.

Con esta segunda asignación de muestra se revisaron los tamaños de muestra resultantes a nivel de las áreas urbanas que requerían cobertura adicional para el cálculo del índice de precios al consumidor que realiza el Banco de México y se incrementó la muestra en el número de viviendas necesarias para obtener los montos fijados para tal efecto.

Es importante mencionar que la muestra total seleccionada se compone de 30 169 viviendas provenientes de la ENIGH 2010 y 38 336 viviendas cuyo levantamiento fue financiado por el CONEVAL para recolectar información de ingresos y características sociodemográficas de los hogares; en estas viviendas no se recabó información acerca del gasto de los hogares.

En el anexo B se presenta la distribución de la muestra en viviendas por entidad según dominio de estudio para el MCS 2010.

6. Selección de la muestra

La selección de la muestra para el Módulo de Condiciones Socioeconómicas de la ENIGH 2010 se realizó en forma independiente para cada entidad y estrato: el procedimiento varió dependiendo de la zona.

6.1 EN URBANO ALTO

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

1. De las n_{ech} UPM que integran el marco de la muestra maestra, se eligieron n_{ech}^* UPM con igual probabilidad para el MCS 2010.
2. En cada UPM se seleccionaron 5 viviendas con igual probabilidad para el MCS 2010.

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

$$P\{V_{echi}\} = \frac{n_{ech} m_{echi}}{m_{ech}} \frac{n_{ech}^*}{n_{ech}} \frac{5}{m_{echi}^*} = \frac{5 n_{ech}^* m_{echi}}{m_{ech} m_{echi}^*}$$

Su factor de expansión² está dado por:

$$F_{echi} = \frac{m_{ech} m_{echi}^*}{5 n_{ech}^* m_{echi}}$$

Donde:

- n_{ech} = número de UPM seleccionadas, en el h -ésimo estrato, en la c -ésima ciudad, en la e -ésima entidad, para el marco de la muestra maestra.
- n_{ech}^* = número de UPM seleccionadas para la ENIGH 2010, en el h -ésimo estrato, en la c -ésima ciudad, en la e -ésima entidad.
- m_{ech} = número de viviendas en el h -ésimo estrato, en la c -ésima ciudad, en la e -ésima entidad.
- m_{echi} = número de viviendas en la i -ésima UPM, en el h -ésimo estrato, en la c -ésima ciudad, en la e -ésima entidad.
- m_{echi}^* = número de viviendas en la i -ésima UPM, en el h -ésimo estrato, en la c -ésima ciudad, en la e -ésima entidad al momento de la actualización del listado de viviendas, previo al levantamiento del MCS 2010.

² El factor de expansión se define como el inverso de la probabilidad de selección.

6.2 EN COMPLEMENTO URBANO

1. De las n_{eh} UPM que se seleccionaron para el marco de la muestra maestra, se eligieron n_{eh}^* UPM con igual probabilidad para el MCS 2010.
2. En cada UPM seleccionada, se eligieron 20 viviendas con igual probabilidad.

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

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

Su factor de expansión está dado por:

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

Donde:

- n_{eh} = número de UPM seleccionadas para el marco de 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.
- n_{eh}^* = número de UPM seleccionadas para el MCS 2010, 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 levantamiento del MCS 2010.

6.3 EN RURAL

De las n_{eh} UPM que se seleccionaron para el marco de la muestra maestra, se eligieron n_{eh}^* UPM con igual probabilidad para el MCS 2010.

En cada UPM seleccionada se eligieron 2 segmentos de 10 viviendas aproximadamente, con igual probabilidad.

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

$$P\{V_{ehi}\} = \frac{n_{eh}}{m_{eh}} \frac{m_{ehi}}{n_{eh}} \frac{n_{eh}^*}{m_{ehi}^*} \frac{2 \cdot 10}{m_{ehi}^*} = \frac{20 n_{eh}^* m_{ehi}}{m_{eh} m_{ehi}^*}$$

En consecuencia, su factor de expansión está dado por:

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

Donde:

- n_{eh} = número de UPM seleccionadas del marco de 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 total de viviendas en el h-ésimo estrato, en la e-ésima entidad.
- n_{eh}^* = número de UPM seleccionadas para el MCS 2010, 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 del levantamiento del MCS 2010.

7. Ajuste a los factores de expansión

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

7.1 AJUSTE POR NO RESPUESTA

El ajuste por No respuesta atribuida al informante se realiza a nivel UPM, mediante las siguientes expresiones:

$$F'_{echi} = F_{echi} \frac{nv_{h_{echi}}}{nv_{hcr_{echi}}}$$

Donde:

F'_{echi} = factor de expansión corregido por No respuesta para las viviendas de la i-ésima UPM, del h-ésimo estrato, de la c-ésima ciudad, de la e-ésima entidad.

F_{echi} = factor de expansión de la i-ésima UPM, del h-ésimo estrato, de la c-ésima ciudad, de la e-ésima entidad.

$nv_{h_{echi}}$ = número de viviendas habitadas en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad, en la e-ésima entidad.

$nv_{hcr_{echi}}$ = número de viviendas habitadas con respuesta en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad, en la e-ésima entidad.

7.2 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 generada por INEGI 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.

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

8. Estimadores

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

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

Donde:

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

$X_{echis\ell}^{UA}$ = 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 c-ésima ciudad, en la e-ésima entidad en el dominio urbano alto.

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

$X_{ehis\ell}^R$ = 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.

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.

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} .

9. Estimación de las precisiones

Para la evaluación de los errores de muestreo de las principales estimaciones estatales y nacionales se usó el método de “Conglomerados Últimos”,³ 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.

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:

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

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

$$EE = \sqrt{\hat{V}(\hat{\theta})} \quad CV = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}} \quad DEFF = \frac{\hat{V}(\hat{\theta})}{\hat{V}(\hat{\theta})_{mas}}$$

Donde:

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

$\hat{V}(\hat{\theta})_{mas}$ = estimador de la varianza, bajo muestreo aleatorio simple.

Finalmente, el intervalo de confianza al $100(1 - \alpha)\%$, se construye de la siguiente forma:

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

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

Anexo

A. INDICADORES EMPLEADOS EN LA ESTRATIFICACIÓN DE LA MUESTRA MAESTRA POR ÁMBITO DE ESTUDIO

Descripción del indicador	Ámbito de estudio			
	Nacional	Urbano alto	Complemento urbano	Rural
% De vivienda				
Que disponen de agua entubada dentro de la vivienda		X		
Con drenaje	X	X		X
Con electricidad				X
Que disponen de agua, luz y drenaje	X	X	X	X
Con piso diferente de tierra	X			X
Con paredes de material sólido				X
Con cocina exclusiva	X	X	X	X
Sin hacinamiento	X	X	X	X
Con servicio sanitario exclusivo con conexión de agua		X		
Con servicio sanitario exclusivo con admisión de agua	X		X	
Que utilizan gas para cocinar	X			X
Con radio o radiograbadora	X			X
Con televisión	X			
Con refrigerador	X	X	X	
Con licuadora	X			X
Con automóvil o camioneta propios	X	X	X	
Con video casetera			X	
Con lavadora	X		X	
Con teléfono		X	X	
Con calentador de agua		X	X	
Con cuatro bienes (teléfono, refrigerador, lavadora y boiler)		X		
Con cuatro bienes (radio, televisión, licuadora y refrigerador)	X		X	
Con el mínimo equipamiento (radio o televisión y licuadora)				X
% De población				
Derechohabiente a servicio de salud		X	X	
De 6 a 17 años que asiste a la escuela	X	X	X	
De 6 a 14 años que asiste a la escuela				X
De 15 años y más alfabeto	X			X
De 15 años y más con postprimaria	X	X	X	X
Grado promedio de escolaridad	X	X	X	X
Ocupada que gana más de 2.5 salarios mínimos	X	X	X	X
Ocupada que gana más de 5 salarios mínimos	X	X	X	
Femenina de 12 años y más económicamente activa	X	X	X	
Económicamente activa de 20 a 49 años	X	X	X	
% Otros				
De hogares de los deciles 8, 9 y 10 a nivel nacional	X	X	X	
Relación de dependencia económica	X	X	X	
Total de indicadores	24	21	21	16

B. DISTRIBUCIÓN DE LA MUESTRA EN VIVIENDAS POR ENTIDAD SEGÚN DOMINIO DE ESTUDIO, MCS 2010

Entidad		Dominio de estudio		
		Urbano	Rural	Total
01	Aguascalientes	1 650	353	2 003
02	Baja California	1 800	199	1 999
03	Baja California Sur	1 510	494	2 004
04	Campeche	1 500	511	2 011
05	Coahuila de Zaragoza	1 770	237	2 007
06	Colima	1 710	289	1 999
07	Chiapas	2 200	818	3 018
08	Chihuahua	1 660	360	2 020
09	Distrito Federal	2 900	98	2 998
10	Durango	1 510	507	2 017
11	Guanajuato	1 420	588	2 008
12	Guerrero	1 220	800	2 020
13	Hidalgo	1 410	601	2 011
14	Jalisco	1 660	349	2 009
15	México	2 620	386	3 006
16	Michoacán de Ocampo	1 320	690	2 010
17	Morelos	1 720	279	1 999
18	Nayarit	1 340	672	2 012
19	Nuevo León	1 840	161	2 001
20	Oaxaca	1 190	825	2 015
21	Puebla	1 445	578	2 023
22	Querétaro	1 390	619	2 009
23	Quintana Roo	1 660	344	2 004
24	San Luis Potosí	1 340	703	2 043
25	Sinaloa	1 430	581	2 011
26	Sonora	1 470	537	2 007
27	Tabasco	1 300	721	2 021
28	Tamaulipas	1 700	304	2 004
29	Tlaxcala	1 620	390	2 010
30	Veracruz de Ignacio de la Llave	1 305	756	2 061
31	Yucatán	2 585	543	3 128
32	Zacatecas	1 220	797	2 017
Total		52 415	16 090	68 505