

# **DISEÑO MUESTRAL**

**ENCUESTA LABORAL Y DE CORRESPONSABILIDAD  
SOCIAL 2012  
(ELCOS-2012)**

## ÍNDICE

## Página

1. Objetivo de la encuesta	1
2. Población objetivo	1
3. Cobertura geográfica	1
4. Diseño de la muestra	1
5. Marco de la encuesta	1
5.1 Formación de las unidades primarias de muestreo (UPM)	1
En urbano alto	1
5.2 Estratificación	2
6. Esquema de muestreo	2
a) Probabilístico	2
b) Estratificado	2
c) Bietápico	2
d) Por conglomerados	2
7. Tamaño de la muestra	2
8. Afijación de la muestra	3
9. Selección de la muestra	4
10. Ajuste a los factores de expansión	4
10.1 Ajuste por No respuesta a nivel vivienda	4
10.2 Ajuste por proyección	5
11. Estimadores	5
12. Estimación de las precisiones	5
Anexos	
- Indicadores empleados en la estratificación del marco de la muestra maestra por ámbito de estudio (Cuadro 1)	
-Distribución de la muestra para la Encuesta Laboral y de Corresponsabilidad Social 2012 (ELCOS-2012) (Cuadro 2).	

## 1. Objetivo de la encuesta

Obtener información estadística sobre la situación de las mujeres en edad productiva en términos de su inserción en el mercado laboral y su responsabilidad en el cuidado de niñas, niños, personas enfermas y/o con limitaciones para valerse por sí mismos.

## 2. Población objetivo

La población objetivo son las mujeres en edad productiva que residen en viviendas particulares ubicadas en el agregado de 32 áreas urbanas de interés.

## 3. Cobertura geográfica

La ELCOS-2012 está diseñada para dar resultados a los siguientes niveles de desagregación:

- Agregado de las 32 áreas urbanas.
- Ciudades de México, Guadalajara y Monterrey.

## 4. Diseño de la muestra

El diseño de la muestra para la ELCOS-2012 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 son las mujeres en edad productiva.

## 5. Marco de la encuesta

El marco de muestreo que se empleó para la ELCOS-2012 es el Marco Nacional de Viviendas 2002 del INEGI, construido a partir de 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 INEGI; como tal, su diseño es probabilístico, estratificado, unietápico y por conglomerados, a los que se denominó 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.

### 5.1 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, en el caso del urbano alto es de la siguiente manera:

#### 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.<sup>1</sup>

<sup>1</sup> Área Geoestadística Básica.

- 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 pero del mismo tamaño de localidad.

## 5.2 Estratificación

En el ámbito urbano alto, se forman 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 21 indicadores\* contruidos con información del XII Censo General de Población y Vivienda 2000, para lo cual se emplearon métodos estadísticos multivariados.

## 6. Esquema de muestreo

El tipo de muestreo empleado es probabilístico, estratificado, bietápico y por conglomerados. A continuación, se mencionan las definiciones de cada uno de estos métodos.

### a) Probabilístico

Las unidades de muestreo tienen una probabilidad conocida y distinta de cero de ser seleccionadas, lo que significa que, se puede hacer inferencia estadística de la población.

### b) Estratificado

Las unidades primarias de muestreo con características socioeconómicas similares que pertenecen a un mismo tamaño de localidad se agrupan para formar estratos.

### c) Bietápico

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

### d) Por conglomerados

Previamente a la selección de la primera etapa se forman conjuntos de unidades muestrales llamadas unidades primarias de muestreo, de las cuales se obtiene la muestra.

## 7. Tamaño de la muestra

Para el cálculo del tamaño de la muestra se tomó como variable de referencia la proporción de mujeres, que cuidan niños o enfermos, la expresión del cálculo de la muestra es:

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

Donde:

- n = tamaño de la muestra.
- p = estimación de la proporción de interés.
- q = (1-p).

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

- z = valor en tablas estadísticas de la distribución normal estándar para una confianza prefijada.
- r = error relativo máximo esperado.
- 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.
- PMV = promedio de mujeres en edad productiva por vivienda.

Se calculó un tamaño de muestra de 15 000 viviendas para el agregado de 32 áreas urbanas de interés considerando una confianza del 90%, un error relativo máximo esperado de 15%, una tasa de No respuesta máxima esperada de 20%, un efecto de diseño de 3.57 y un promedio de mujeres por vivienda de 0.68 de que cuidan niños o enfermos, para una proporción de 5%.

Para dar resultados en la Ciudad de México, Guadalajara y Monterrey, se calcularon sus tamaños de muestra de manera independiente:

El tamaño de muestra de la Ciudad México fue de 3 500 viviendas, se obtuvo considerando una confianza de 90%, un efecto de diseño de 2.36, error relativo esperado de 15%, una tasa de No respuesta máxima esperada de 20% y un promedio de mujeres por vivienda que cuidan niños o enfermos de 0.67, para una proporción de 15%.

El tamaño de muestra para las ciudades de Guadalajara y Monterrey fue de 2 000 viviendas en cada una, se obtuvo considerando una confianza de 90%, un efecto de diseño de 1.81, un error relativo esperado de 15%, una tasa de No respuesta máxima esperada de 20% y un promedio de mujeres por vivienda que cuidan niños o enfermos de 0.77, para una proporción de 15%.

Estas ciudades ya tenían una proporción de la muestra que les correspondía de las 15 000 iniciales, se les sumó el complemento hasta completar la muestra calculada, al final se obtuvo una muestra total de 19 850 viviendas.

## 8. Afijación de la muestra

La afijación de la muestra se realiza dentro de cada ciudad entre los diferentes estratos de manera proporcional a su tamaño, para lo cual se emplea la siguiente expresión:

$$n_{ch} = \frac{N_{ch}}{N_c} n_c$$

Donde:

- $n_{ch}$  = número de viviendas en muestra en el h-ésimo estrato, en la c-ésima ciudad.
- $n_c$  = número total de viviendas en muestra, en la c-ésima ciudad.
- $N_{ch}$  = número total de viviendas en el h-ésimo estrato, en la c-ésima ciudad.
- $N_c$  = número total de viviendas, en la c-ésima ciudad.

En el cuadro 2, se presenta la distribución de la muestra por área urbana de interés.

## 9. Selección de la muestra

La selección de la muestra para la ELCOS-2012 se realizó en forma independiente para cada área urbana, el procedimiento es como sigue:

1. De las  $k_{ch}$  UPM que integran la muestra maestra<sup>2</sup> se eligieron  $k_{ch}^*$  UPM con igual probabilidad para la ELCOS-2012.
2. En cada UPM se seleccionaron cinco viviendas con igual probabilidad.

a) La probabilidad de seleccionar una vivienda de la  $i$ -ésima UPM, del  $h$ -ésimo estrato, de la  $c$ -ésima ciudad es:

$$P\{V_{chi}\} = \frac{k_{ch} m_{chi}}{m_{ch}} \frac{k_{ch}^*}{k_{ch}} \frac{5}{m_{chi}^*} = \frac{5 k_{ch}^* m_{chi}}{m_{ch} m_{chi}^*}$$

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

$$F_{chij} = \frac{m_{ch} m_{chi}^*}{5 k_{ch}^* m_{chi}}$$

Donde:

$k_{ch}$  = número de UPM seleccionadas para el marco de la muestra maestra en el  $h$ -ésimo estrato, en la  $c$ -ésima ciudad.

$m_{chi}$  = número de viviendas en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato, en la  $c$ -ésima ciudad.

$m_{ch}$  = total de viviendas en el  $h$ -ésimo estrato, en la  $c$ -ésima ciudad.

$m_{chi}^*$  = total de viviendas en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato, en la  $c$ -ésima ciudad al momento del levantamiento de la ELCOS-2012.

$k_{ch}^*$  = número de UPM seleccionadas, en el  $h$ -ésimo estrato, en la  $c$ -ésima ciudad, para la ELCOS-2012.

## 10. Ajuste a los factores de expansión

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

### 10.1. Ajuste por No respuesta a nivel vivienda

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

$$F'_{chi} = F_{chi} \frac{V_{chi}}{V_{chi}^*}$$

Donde:

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

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

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

- $F_{chi}$  = factor de expansión de la i-ésima UPM, del h-ésimo estrato, de la c-ésima ciudad.
- $V_{chi}$  = número de viviendas habitadas seleccionadas en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad.
- $V_{chi}^*$  = número de viviendas habitadas seleccionadas en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad para las cuales se obtuvo respuesta.

## 10.2. Ajuste por proyección

Los factores de expansión ajustados anteriormente, se corrigen a fin de asegurar que se obtenga la población determinada por la proyección de población generada por el 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 proyección.
- $PEXP_D$  = población total a la que expande la encuesta en el dominio D.

## 11. Estimadores

El estimador del total de la característica X es:

$$\hat{X} = \sum_c \sum_h \sum_i \sum_j F_{chij} X_{chij}$$

Donde:

- $F_{chij}$  = factor de expansión final de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, de la c-ésima ciudad.
- $X_{chij}$  = valor observado de la característica de interés X de personas seleccionadas, en la j-ésima vivienda, en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad.

Para la estimación de proporciones, tasas y promedios se utiliza el estimador de razón:

$$\hat{R} = \frac{\hat{X}}{\hat{Y}}$$

Donde,  $\hat{Y}$  se define en forma análoga a  $\hat{X}$ .

## 12. Estimación de las precisiones

Para la evaluación de los errores de muestreo de las principales estimaciones nacionales se usó el método de Conglomerados Últimos,<sup>4</sup> basado en que la mayor contribución a la varianza de un estimador, en un diseño bietápico es

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

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  $\hat{R}$ :

$$\hat{V}(\hat{R}) = \frac{1}{\hat{Y}^2} \sum_c \left\{ \sum_h^{L_c} \frac{k_{ch}}{k_{ch}-1} \sum_i^{k_{ch}} \left[ \left( \hat{X}_{chi} - \frac{1}{k_{ch}} \hat{X}_{ch} \right) - \hat{R} \left( \hat{Y}_{chi} - \frac{1}{k_{ch}} \hat{Y}_{ch} \right) \right]^2 \right\}$$

Donde:

$\hat{X}_{chi}$  = total ponderado de la variable de estudio X en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad.

$\hat{X}_{ch}$  = total ponderado de la variable de estudio X en el h-ésimo estrato, en la c-ésima ciudad.

$k_{ch}$  = número de UPM en el h-ésimo estrato, en la c-ésima ciudad.

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

El cálculo de la varianza del estimador de un total, se obtiene con la siguiente expresión:

$$\hat{V}(\hat{X}_{NAL}) = \sum_c \sum_{h=1}^{L_c} \frac{k_{ch}}{k_{ch}-1} \sum_{i=1}^{k_{ch}} \left( \hat{X}_{chi} - \frac{1}{k_{ch}} \hat{X}_{ch} \right)^2$$

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

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

Donde:

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

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

Finalmente, el intervalo de confianza a 90%, se calcula de la siguiente forma:

$$I_{1-\alpha} = \left( \hat{\theta} - 1.645 \sqrt{\hat{V}(\hat{\theta})}, \hat{\theta} + 1.645 \sqrt{\hat{V}(\hat{\theta})} \right)$$

donde el nivel de significancia es de  $\alpha = 0.10$ .



Indicadores empleados en la estratificación del marco de la muestra maestra por ámbito de estudio

Cuadro 1

Descripción del indicador	Ámbito de estudio			
	Nacional	Urbano alto	Complemento urbano	Rural
<b>% De vivienda</b>				
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 videocasetera			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
<b>% De población</b>				
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	
<b>% De otros</b>				
Hogares de los deciles 8, 9 y 10 a nivel nacional	X	X	X	
Relación de dependencia económica	X	X	X	
<b>Total de indicadores</b>	<b>24</b>	<b>21</b>	<b>21</b>	<b>16</b>

Distribución de la muestra para la Encuesta Laboral y de Corresponsabilidad Social 2012 (ELCOS-2012)

Cuadro 2

Clave de la entidad	Entidad	Clave del área	Áreas urbanas de interés	Muestra
01	Aguascalientes	14	Aguascalientes	350
02	Baja California	21	Frontera Tijuana	850
03	Baja California Sur	40	La Paz	350
04	Campeche	28	Campeche	350
05	Coahuila de Zaragoza	17	Saltillo	350
06	Colima	33	Colima	350
07	Chiapas	19	Tuxtla Gutiérrez	350
08	Chihuahua	09	Chihuahua	450
09	Distrito Federal	01	Cd. de México	1 750
10	Durango	26	Durango	350
11	Guanajuato	05	León	850
12	Guerrero	13	Acapulco	450
13	Hidalgo	43	Pachuca	350
14	Jalisco	02	Guadalajara	2 000
15	México	01	Cd. de México	1 750
15	México	16	Toluca	450
16	Michoacán de Ocampo	15	Morelia	350
17	Morelos	29	Cuernavaca	450
18	Nayarit	27	Tepic	350
19	Nuevo León	03	Monterrey	2 000
20	Oaxaca	31	Oaxaca	350
21	Puebla	04	Puebla	850
22	Querétaro	36	Querétaro	350
23	Quintana Roo	41	Cancún	350
24	San Luis Potosí	07	San Luis Potosí	450
25	Sinaloa	24	Culiacán	350
26	Sonora	25	Hermosillo	350
27	Tabasco	18	Villahermosa	350
28	Tamaulipas	10	Tampico	350
29	Tlaxcala	39	Tlaxcala	350
30	Veracruz de Ignacio de la Llave	12	Veracruz	450
31	Yucatán	08	Mérida	450
32	Zacatecas	32	Zacatecas	350
<b>Total</b>				<b>19 850</b>