

METODOLOGÍA
Percepciones Sociales
del INEGI 2017 Encuesta
Nacional en Vivienda

Percepciones Sociales del INEGI 2017

Índice

DISEÑO MUESTRAL.....	2
Población objetivo	2
Marco de la Encuesta	2
Formación de la Unidades Primarias de Muestreo (UPM).....	2
Estratificación	3
Esquema de muestreo.....	4
Tamaño de la Muestra.....	4
Afijación de la muestra	5
Cálculo de los factores de expansión.....	7
Ajuste de los factores de expansión	8
Ajuste por no respuesta.....	8
Ajuste por proyección.....	9
Calculo de los estimadores	9
1. Estimaciones puntuales	9
2. Medidas de dispersión.....	10
ANEXO 1.....	12
Indicadores empleados en la estratificación del marco de la muestra maestra	12

DISEÑO MUESTRAL

Percepciones Sociales del INEGI 2017

Encuesta Nacional en Vivienda

La encuesta nacional en vivienda tiene como objetivo obtener información acerca de los niveles de conocimiento, utilización y confianza que tiene la población sobre la información que recopila y produce el **Instituto Nacional de Estadística y Geografía (INEGI)**. Para ello se diseñó una encuesta que permite captar información a nivel nacional.

Población objetivo

Para esta encuesta, la población objetivo la constituyeron las personas de 15 años o más que habitan en el país.

Marco de la Encuesta

Como marco de muestreo se empleó el Marco Nacional de Viviendas 2012 del INEGI, construido a partir de la información cartográfica y demográfica obtenida del Censo Nacional de Población y Vivienda 2010.

El Marco Nacional de Viviendas 2012, es una muestra maestra elaborada por el INEGI, y de la cual se diseñan todas las muestras de las encuestas en vivienda que realiza el Instituto. Para ello, el Instituto genera Unidades Primarias de Muestreo (UPM) como se describirá en las siguientes líneas.

Formación de la Unidades Primarias de Muestreo (UPM)

Las Unidades Primarias de Muestreo (UPM) están constituidas por agrupaciones de viviendas con características diferenciadas, dependiendo del ámbito al que pertenecen, como se indica 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 del mismo AGEB¹

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 del mismo AGEB

c) En Rural

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

- Una localidad de un municipio.
- Una unión de dos o más localidades de un municipio.

Estratificación

En una primera etapa, las UPM están asociadas a un estrato geográfico conformado por tres niveles (Entidad-Ámbito-Zona).

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 de estudio: urbano alto, complemento urbano y rural, los cuales a su vez se dividen en zonas, determinadas por criterios de tamaño de la población.

En una segunda etapa se formaron cuatro estratos en donde se agruparon todas las UPM del país. Esta estratificación considera las características sociodemográficas de los habitantes de las viviendas, las características físicas y equipamiento de las mismas, a través

¹ Área Geoestadística Básica

de 34 indicadores² (véase Anexo 1) contruidos con información del Censo de Población y Vivienda 2010.

Esquema de muestreo

La selección de unidades se realizó bajo un diseño probabilístico polietápico, estratificado y por conglomerados, donde cada unidad de muestreo tuvo una probabilidad conocida y diferente de cero de ser seleccionada.

Para cada estrato, se consideraron las siguientes unidades de muestreo:

- a) *Unidades Primarias de Muestreo (UPM)*: De las n_{eh} UPM seleccionadas del Marco Nacional de Viviendas 2012, en el h-ésimo estrato, en la e-ésima entidad, se seleccionaron n^*_{eh} bajo un diseño de muestreo aleatorio simple (MAS)³.
- b) *Unidades Secundarias de Muestreo (USM)*: Se seleccionaron 7 viviendas de cada UPM, de forma sistemática, con arranque aleatorio.
- c) *Unidades Últimas de Muestreo (UUM)*: Se seleccionó de manera aleatoria (MAS) a un individuo de 15 años y más dentro de la vivienda.

Tamaño de la Muestra

El tamaño de muestra obtenido fue de 7 903 casos, los cuales fueron distribuidos entre los diferentes estratos. Para obtener el tamaño de muestra se empleó la siguiente fórmula:

$$n = \frac{z_{1-\alpha/2}^2(1-p)DEFF}{d^2p(1-TNR)}$$

donde:

n = tamaño de la muestra;

p =proporción de la variable de interés a estimar;

² El INEGI se encarga de construir los 34 indicadores empleados para la estratificación del marco de la muestra maestra. La descripción de estos indicadores se muestra como Anexo 1 de este documento.

³ La selección de UPM fue elaborada por el INEGI.

d =error relativo máximo posible;

$z_{1-\alpha/2}^2$ =valor en tablas de valores de probabilidad acumulada para la distribución normal estándar. Este valor depende del nivel de confianza asignado $(1-\alpha)*100$ por ciento;

$DEFF$ = efecto de diseño, calculado 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.

Considerando un efecto de diseño de 3.5, una proporción de 10.50%⁴ un error relativo de 15% y una tasa de no respuesta máxima esperada del 35% con un nivel de confianza del 95%, el tamaño de muestra fue de aproximadamente 7 903 casos.

Afijación de la muestra

La distribución de la muestra se realizó de forma proporcional al número de viviendas habitadas dentro de cada estrato, como se muestra en la siguiente expresión:

$$n_{e,d,h} = \frac{N_{e,d,h}}{N} n$$

donde:

n es el número de viviendas habitadas en muestra.

$n_{e,d,h}$ es el número de viviendas habitadas en muestra en el h -ésimo estrato, en el d -ésimo dominio, en la e -ésima entidad.

N es el número total de viviendas habitadas en el marco muestral.

$N_{e,d,h}$ es el número total de viviendas habitadas en el h -ésimo estrato en el d -ésimo dominio, en la e -ésima entidad.

Las 7 903 unidades de la muestra se distribuyeron de la siguiente manera en las 32 entidades federativas:

⁴ La proporción se consideró con base en los resultados de la EMECOI 2014 para la variable de *uso de la información proporcionada por el INEGI*.

Entidad Federativa	Total
Aguascalientes	245
Baja California	245
Baja California Sur	195
Campeche	195
Chiapas	245
Chihuahua	299
Ciudad de México	245
Coahuila De Zaragoza	245
Colima	196
Durango	244
Guanajuato	301
Guerrero	242
Hidalgo	244
Jalisco	301
México	299
Michoacán De Ocampo	241
Morelos	245
Nayarit	195
Nuevo León	245
Oaxaca	245
Puebla	300
Querétaro	244
Quintana Roo	244
San Luis Potosí	244
Sinaloa	244
Sonora	245
Tabasco	244
Tamaulipas	243
Tlaxcala	195
Veracruz De Ignacio De La Llave	299
Yucatán	244
Zacatecas	245
Total	7903

Cálculo de los factores de expansión

Para obtener los factores de expansión utilizados en el cálculo de los estimadores de los parámetros tanto para individuos como para viviendas, se utilizó la probabilidad de selección de cada etapa de muestreo.

La probabilidad de selección del k -ésimo individuo, de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato y de la e -ésima entidad es:

$$P[x_{e,h,i,j,k}] = \frac{n_{e,h} m_{e,h,i}}{m_{e,h}} \frac{n_{e,h}^*}{n_{e,h}} \frac{7}{m_{e,h,i}} \frac{1}{I_{e,h,i,j}} = \frac{7 n_{e,h}^*}{m_{e,h} I_{e,h,i,j}}$$

donde:

$x_{e,h,i,j,k}$ es el k -ésimo individuo en la j -ésima vivienda, en la i -ésima UPM, en el h -ésimo estrato y e -ésima entidad.

$n_{e,h}$ es el número de UPM seleccionadas en el h -ésimo estrato, en la e -ésima entidad dentro del Marco Nacional de Viviendas 2012.

$n_{e,h}^*$ es el número de UPM seleccionadas aleatoriamente, a partir de las primeras $n_{e,h}$, en el h -ésimo estrato, en la e -ésima entidad.

$m_{e,h}$ es el número de viviendas en el h -ésimo estrato, en la e -ésima entidad.

$m_{e,h,i}$ es el número de viviendas en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad.

$I_{e,h,i,j}$ es el número de personas de 15 años y más, en la j -ésima vivienda, en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad.

El recíproco de la probabilidad de selección del k -ésimo individuo es el factor de expansión correspondiente, es decir:

$$F_{e,h,i,j,k} = \frac{1}{P[x_{e,h,i,j,k}]} = \frac{m_{e,h} I_{e,h,i,j}}{7 n_{e,h}^*}$$

La probabilidad de selección de la j -ésima vivienda, de la i -ésima UPM, del h -ésimo estrato y de la e -ésima entidad es:

$$P[x_{e,h,i,j}] = \frac{n_{e,h} m_{e,h,i}}{m_{e,h}} \frac{n_{e,h}^*}{n_{e,h}} \frac{7}{m_{e,h,i}} = \frac{7 n_{e,h}^*}{m_{e,h}}$$

El factor de expansión correspondiente se obtiene de la misma manera que el anterior, es decir:

$$F_{e,h,i,j} = \frac{1}{P[x_{e,h,i,j}]} = \frac{m_{e,h}}{7 n_{e,h}^*}$$

Ajuste de los factores de expansión

Los factores de expansión descritos anteriormente fueron calibrados con base en la información obtenida durante el levantamiento, así como los datos poblacionales del Censo de Población y Vivienda 2010, y proyectados por el Consejo Nacional de Población (CONAPO), con el fin de expandir a la población en las proporciones que se presentan en el país. A continuación, se describen los procedimientos realizados.

Ajuste por no respuesta

El ajuste por no respuesta es un método de calibración que se utiliza cuando no se realizan todas las entrevistas esperadas, y se calcula de la siguiente forma:

$$F'_{e,d,h,i} = F_{e,d,h,i} \frac{v'_{e,d,h}}{v''_{e,d,h}}$$

Donde:

$F'_{e,d,h,i}$ es el factor de expansión corregido por no respuesta de la i -ésima UPM, del h -ésimo estrato, del d -ésimo dominio, de la e -ésima entidad.

$F_{e,d,h,i}$ es el factor de expansión natural de la i -ésima UPM, del h -ésimo estrato, del d -ésimo dominio, de la e -ésima entidad.

$v'_{e,d,h}$ es el total de viviendas seleccionadas en el h -ésimo estrato, en el d -ésimo dominio, en la e -ésima entidad.

$v''_{e,d,h}$ es el total de viviendas seleccionadas con respuesta en el h -ésimo estrato, en el d -ésimo dominio, en la e -ésima entidad.

Ajuste por proyección

Los factores ajustados por no respuesta se corrigen nuevamente con el fin de obtener estimaciones poblacionales en las proporciones que se presentan en el país, con base en las proyecciones generadas por el CONAPO. Para calcular el ajuste por proyección se utilizaron los datos poblacionales proyectados a mitad de año (2017), por estado, sexo y grupos decenales de edad.

El ajuste por proyección se calculó de la siguiente forma:

$$F''_{e,r,s} = F'_{e,r,s} \frac{P'_{e,r,s}}{P''_{e,r,s}}$$

donde:

$F''_{e,r,s}$ es el factor de expansión corregido por proyección en la e -ésima entidad, del grupo decenal r y sexo s .

$F'_{e,r,s}$ es el factor de expansión corregido por no respuesta en la e -ésima entidad, del grupo decenal r y sexo s .

$P'_{e,r,s}$ es la población total proyectada por CONAPO a mitad de año de la e -ésima entidad, del grupo decenal r y sexo s .

$P''_{e,r,s}$ es la población total a la que expande la encuesta en la e -ésima entidad, del grupo decenal r y sexo s .

Calculo de los estimadores

Para el cálculo de las estimaciones puntuales y sus respectivas medidas de dispersión, se utilizaron las siguientes fórmulas.

1. Estimaciones puntuales

Para el cálculo de proporciones, promedios y tasas, se utilizó el estimador de razón:

$$\hat{R} = \frac{\sum_{i=1}^{32} \sum_{h=1}^H \sum_{j=1}^J \sum_{k=1}^7 X_{ihjk} F''_{ihjk}}{\sum_{i=1}^{32} \sum_{h=1}^H \sum_{j=1}^J \sum_{k=1}^7 Y_{ihjk} F''_{ihjk}}$$

donde:

X_{ihjk} Es el valor observado de la característica de interés X en la k -ésima vivienda, de la j -ésima UPM, en el h -ésimo estrato, en la i -ésima entidad.

Y_{ihjk} Es el valor observado de la característica de interés Y en la k -ésima vivienda, de la j -ésima UPM, en el h -ésimo estrato, en la i -ésima entidad.

F''_{ihjk} Es el factor de expansión de la k -ésima vivienda, de la j -ésima UPM, en el h -ésimo estrato, en la i -ésima entidad.

2. Medidas de dispersión

Para el estimador de la Varianza se utilizó el método de Conglomerados Últimos⁵, al que se le aplicó el método de series de Taylor, obteniéndose la siguiente fórmula para estimar la precisión del estimador de razón \hat{R} .

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

donde:

\hat{X}_{ehi} es el valor total observado y ponderado de la característica de interés X en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad.

\hat{Y}_{ehi} es el valor total observado y ponderado de la característica de interés Y en la i -ésima UPM, en el h -ésimo estrato, en la e -ésima entidad.

\hat{X}_{eh} es el valor total observado y ponderado de la característica de interés X en el h -ésimo estrato, en la e -ésima entidad.

\hat{Y}_{eh} es el valor total observado y ponderado de la característica de interés Y en el h -ésimo estrato, en la e -ésima entidad.

\hat{Y} es el valor total observado y ponderado de la característica de interés Y a nivel nacional.

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

n_{eh} es el número de UPM en el h-ésimo estrato de la e-ésima entidad.

La estimación de la desviación estándar (*D.E.*), intervalo de confianza al $100(1-\alpha)\%$ (*I.C.*), el coeficiente de variación (*C.V.*) y efecto de diseño (*DEFF*) se hace a través de las siguientes expresiones:

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

$$I.C._{1-\alpha} = \hat{R} \pm Z_{\alpha/2} \sqrt{\hat{V}(\hat{R})}$$

donde:

\hat{R} es el estimador de razón

$\hat{V}(\hat{R})$ es la varianza estimada del estimador de razón.

$\hat{V}(\hat{R})_{MAS}$ es la varianza estimada del estimador de razón.

$Z_{\alpha/2}$ es el factor de confiabilidad para un nivel de significancia α .

ANEXO 1

Indicadores empleados en la estratificación del marco de la muestra maestra

Indicadores empleados en la estratificación del marco de la muestra maestra

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 o más años de edad que aprobaron algún grado de escolaridad diferente al nivel preescolar.
PP15PRI_CO	de 15 o más años de edad que tienen como máxima escolaridad 6 grados aprobados en primaria.
PP15SEC_CO	de 15 o más años 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.
TOCU12A17	no ocupada de 12 a 17 años entre la población total de este rango.
PPOMAYED	ocupada de 18 años y mas entre la población total ocupada.
PTASAOcupa	Tasa de ocupación.
Viviendas particulares habitadas:	
PVIVSINH	sin hacinamiento.
PVPH_PISDT	con piso de cemento o firme, madera, mosaico y 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 ella, 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.

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.