

DISEÑO MUESTRAL DE LA ENCUESTA NACIONAL SOBRE PERCEPCIONES ACERCA DE LA CONSTITUCIÓN

Yvón ANGULO REYES⁴⁵

El objetivo del estudio fue obtener un diagnóstico a nivel nacional sobre las percepciones que los ciudadanos tienen acerca de la Constitución, que permitiera obtener información que se pudiera desagregar a nivel temático, a nivel regional, y recoger algunas características demográficas particulares.

1. *Marco muestral*

El marco muestral que se considera ideal para toda encuesta probabilística es el listado de las personas a entrevistar, a partir del cual se pudiera realizar la selección directa de los entrevistados. En virtud de la imposibilidad de contar con un marco muestral con dichas características, se recurre a métodos de selección en etapas para poder llegar a la unidad última de muestreo, que en este caso será la población mayor de 15 años.

El Censo General de Población y Vivienda representa una alternativa viable como marco de muestreo. El Censo General de Población y Vivienda proporciona información a nivel estatal, municipal, local y por Ageb (Área Geoestadística Básica), de tal forma que para poder llegar a la población de interés pueda hacerse mediante la selección de los diferentes niveles de agrupación.

La población objeto de estudio se definió como todos aquellos ciudadanos mayores de 15 años de edad, residentes en el territorio nacional en viviendas particulares.⁴⁶ La información censal que se utilizó como marco de muestreo fue del Censo General de Población y Vivienda de 2000.⁴⁷

2. *Alcance de la encuesta*

Dada la relevancia del tema y la diversidad del país, se consideró importante poder contar con información que permitiera generar estimaciones válidas a nivel nacional y regional. Se realizó una regionalización con la finalidad de poder dar cuenta de las diferencias en cuanto a percepciones, valores y actitudes en las diferentes regiones del país.

Para la regionalización se consideraron cuatro dominios de estudio, en los cuales cada uno de las entidades del país fue clasificada de manera única. Debido al tipo de información con la que se deseaba contar, se dividió el país en cuatro regiones, la clasificación se realizó de

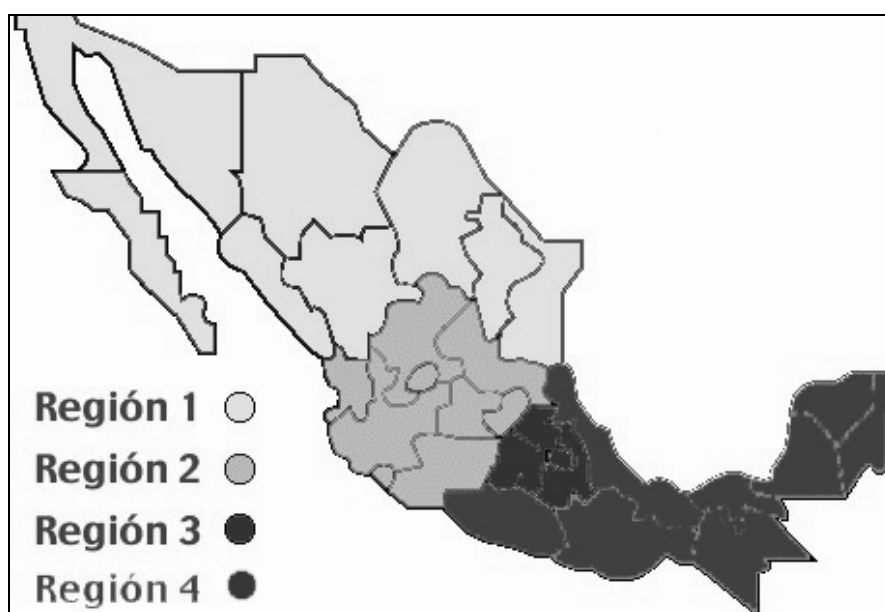
⁴⁵ Quiero expresar mi agradecimiento a Luis Enrique Calva por su valioso apoyo en esta parte de la investigación.

⁴⁶ Se excluye a los habitantes de islas y viviendas colectivas.

⁴⁷ El Instituto de Investigaciones Sociales de la UNAM cuenta con la información completa en CDs del Censo General de Población y Vivienda del 2000 (Scince 2000).

acuerdo a características sociodemográficas y culturales,⁴⁸ su conformación se presenta a continuación:

<i>Región</i>	<i>Entidades que la conforman</i>
I.	Baja California, Baja California Sur, Coahuila, Chihuahua, Durango, Nuevo León, Sinaloa y Sonora.
II.	Aguascalientes, Colima, Guanajuato, Jalisco, Nayarit, Michoacán, Querétaro, San Luis Potosí y Zacatecas.
III.	Distrito Federal, Hidalgo, México, Morelos, Puebla y Tlaxcala.
IV.	Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz y Yucatán.



3. *Diseño muestral*

Con la finalidad de obtener una muestra con representatividad de cada una de las cuatro regiones, y dadas las características del estudio, el esquema de muestreo que se propuso fue estratificado, de conglomerados, polietápico, con selección proporcional al número de habitantes mayores de 15 años en las diferentes etapas de selección (UPM, USM y UTM), sistemática para la selección de las viviendas a visitar y aleatoria al interior de las mismas para la selección de los entrevistados.

4. *Tamaño de la muestra*

El tamaño de muestra propuesto se calculó para estimación de proporciones a nivel nacional y regional. Por lo tanto, se consideró conveniente un tamaño muestra de 1,800 casos, distribuidos de manera aproximadamente proporcional, de acuerdo al tamaño de la región y el estrato.

⁴⁸ La construcción de las regiones tuvo como base regionalizaciones anteriores de investigaciones sociopolíticas.

El cálculo del tamaño de la muestra se realizó para estimar proporciones a nivel nacional y por regiones. El cálculo del tamaño de muestra para proporciones parte de la siguiente expresión probabilística:

$$\Pr(| p - P | \leq d) = 1 - \alpha ,$$

donde:

P = parámetro poblacional que se busca estimar,

p = proporción que se estima a partir de la muestra,⁴⁹

d = margen de error de la proporción estimada p (grado de precisión), y

α = riesgo que se está dispuesto a correr de que el error real sea mayor a d .

Es decir, se desea estimar p con un margen de error entre la estimación y el valor real (P) menor a d , y la aceptación de un pequeño riesgo α de que el error sea mayor a d . La expresión que toma en cuenta las consideraciones anteriores es la siguiente:

$$n = \frac{z^2 p(1-p) DEFF}{d^2 (1-TNR)} ,$$

donde z es el valor de la distribución normal asociado con la confianza deseada.

El tamaño de muestra calculado para cada una de las cuatro regiones fue de 450 casos, con un error de estimación no mayor al 7%, con 95% de confianza y tasa de no respuesta del 10%. Con un tamaño de muestra de 450 casos por región, se obtiene un tamaño de muestra de 1,800 casos a nivel nacional, con lo cual se podrán tener proporciones con un error de estimación mayor a 3.5%, con 95% de confianza y tasa de no respuesta de 10%.

5. Procedimiento de selección

La selección de los informantes se realizó de manera independiente en cada una de las cuatro regiones. Como primer paso, al interior de cada región, se seleccionaron 45 municipios (UPM) con probabilidad proporcional al número de personas mayores de 15 años. Al interior de cada uno de los municipios se realizó la selección una Ageb (UTM) con probabilidad proporcional al número de personas mayores de 15 años. Y las unidades terciarias de muestreo (UTM), conformadas en este caso por las viviendas, se seleccionaron con igual probabilidad.

6. Factores de expansión

El esquema de muestreo estratificado, de conglomerados, polietápico, con selección proporcional al número de habitantes mayores de 15 años en las diferentes etapas de selección (UPM, USM y UTM), sistemática para la selección de las viviendas a visitar y aleatoria al interior de las mismas para la selección de los entrevistados.

De acuerdo con lo anterior, para analizar los resultados fue necesario el uso de ponderadores. El uso de los ponderadores o factores de expansión permite asignar un peso a cada uno de los individuos de la muestra con la finalidad de corregir la distribución muestral, y aproximarse a la distribución real de las respuestas de la población objeto de estudio captadas a través de los

⁴⁹ Uno de los supuestos principales para la determinación de un tamaño de muestra es que p tiene una distribución normal.

cuestionarios. De igual manera, los ponderadores sirven para corregir la no respuesta. A continuación se presenta la forma en la que se obtuvieron.

La probabilidad de selección de un individuo dentro de la región k , en el municipio j , dentro del Ageb i , en la manzana l es:

$$P_{lijk} = (45)(2)(5) \frac{n_{ijk}}{M_k TV_{ijk}}$$

En donde:

M_k : Población total mayor de 15 años de la región k

m_{jk} : Población mayor de 15 años en el municipio j de la región k

n_{ijk} : Población mayor de 15 años del Ageb i , del municipio j , de la región k

Z_{ijk} : Número de manzanas del Ageb i , del municipio j , de la región k

V_{lijk} : Número de viviendas de la manzana l , del Ageb i , del municipio j , de la región k .

TV_{ijk} : Total de viviendas en Ageb i , del municipio j , de la región k .1

Por lo tanto, su factor de expansión es: $F_{lijk} = \frac{1}{P_{lijk}}$

7. Trabajo de campo

Se diseñaron dos cuestionarios. El primer cuestionario, de opinión, se dividió en dos secciones. Una primera recoge las opiniones, las actitudes y los valores de los entrevistados, consta de 83 preguntas, 73 cerradas y 10 preguntas semiabiertas abiertas, que tienen formatos que permiten expresar opiniones definidas a los encuestados, además de captar las respuestas ofrecidas espontáneamente.

El diseño hace posible codificar las respuestas en la forma más adecuada según su naturaleza, de acuerdo con categorías discretas o escalas ordinales y continuas. En una segunda sección, se contemplan 12 preguntas dedicadas al análisis de la información sociodemográfica de los entrevistados (véase el anexo III).

El segundo cuestionario, de léxico, está dirigido a recoger las asociaciones, valoraciones, descripciones y definiciones de los entrevistados acerca de la Constitución. Dicho cuestionario consta de 12 preguntas abiertas y se aplicó a 300 entrevistados.

8. Prueba piloto

Se diseñaron dos versiones del cuestionario de opinión, con el propósito de probar la redacción y el orden de las preguntas. Se aplicó una prueba piloto de 100 cuestionarios en diciembre de 2002. Se corrigieron los contenidos y se validó el funcionamiento de las preguntas.

Durante enero de 2003 se entrevistó en vivienda a 1,794 personas de 15 años y más en 29 entidades federativas, 143 municipios, 180 Ageb⁵⁰ de localidades urbanas y rurales. Adicionalmente, se levantaron 300 cuestionarios de léxico en municipios seleccionados del país.

El operativo para el levantamiento de la encuesta, implicó el trabajo de 48 personas en campo, 8 dedicadas a la codificación, 1 programador y 8 capturistas.

⁵⁰ Áreas geoestadísticas básicas: unidades en las que el INEGI ha dividido el territorio nacional para su estudio.