

Encuesta Nacional sobre Discriminación

Nota metodológica sobre el diseño de la muestra

12 de Mayo de 2005

Introducción

La exclusión social es un fenómeno complejo en el que se entretajan al menos dos dimensiones estructurales. Una de ellas es *objetiva* y está directamente asociada a condiciones de desigualdad social y la otra, *subjetiva*, tiene su origen en representaciones socioculturales, estereotipos y estigmas, así como en manifestaciones simbólicas de inequidad que no necesariamente tienen un vínculo directo con las condiciones materiales de vida de las personas.

La manifestación directa del mecanismo objetivo es la pobreza, mientras que la manifestación más clara del mecanismo subjetivo de exclusión social es la discriminación, cuyo componente principal es el desprecio sistemático a determinados grupos sociales. La discriminación implica restricciones y anulaciones de derechos fundamentales y libertades básicas. El objeto o el resultado de la discriminación es la limitación de acceso a derechos y oportunidades fundamentales de bienestar, a partir de conductas sistemáticas y culturalmente estandarizadas de desprecio social a determinados grupos vulnerables. A pesar de ser el desprecio una condición subjetiva, implica privación, limitación y, en definitiva, violación de los derechos humanos.

Acorde con sus responsabilidades y atribuciones, la Secretaría de Desarrollo Social (SEDESOL) planteó el interés por conocer los niveles y las manifestaciones de este fenómeno social en la vida cotidiana de los mexicanos y las mexicanas que la ejercen o la padecen. Es por ello que se consideró importante analizar la problemática tanto desde el punto de vista de la población general como de algunas poblaciones específicas hipotéticamente expuestas a este fenómeno. En particular, se buscó analizar las manifestaciones de la discriminación hacia y desde la perspectiva del género, la etnia, las preferencias

sexuales, la edad, la religión y la discapacidad. En términos generales el estudio se propuso, entre otros objetivos, indagar los patrones de discriminación entre la ciudadanía mexicana, establecer las diferencias de grado y de forma entre las distintas poblaciones de estudio susceptibles a la discriminación, conocer las principales manifestaciones de la discriminación en distintas esferas sociales y evaluar los factores relacionados a la discriminación entre los grupos seleccionados.

Estos objetivos de la investigación dieron lugar a considerar poblaciones de estudio diferentes. Por un lado, una población de estudio fue la población de 18 años y más que reside en el país, para poder indagar algunas percepciones, actitudes y valores hacia la discriminación entre la ciudadanía mexicana. Por otro lado, fue necesario definir poblaciones de estudio específicas relativas a los grupos sociales entre los que se presume mayor vulnerabilidad en relación con la discriminación. Así, se constituyeron cinco poblaciones de estudio específicas, a saber: a) población de personas de adultos mayores,¹ b) población indígena,² c) población con creencias religiosas no católicas, d) población del sexo femenino y e) personas con discapacidad. El estudio también abarcó a personas con preferencias no heterosexuales, pero ante la dificultad de cubrir este grupo social como población objetivo, se optó por que esta parte del estudio fuera considerada como un estudio de caso.

Lo anterior obligó a diseñar siete instrumentos de captación (cuestionarios) distintos. Por una parte, se diseñó un cuestionario, que se denominó *cuestionario global*, que fue aplicado a la población ciudadana del país. Además, se formularon seis instrumentos específicos para cada una de las poblaciones concretas referidas, pero buscando que se mantuvieran elementos de comparabilidad entre estas poblaciones de interés.

¹ El límite de edad para este grupo fue de 60 años y más.

² Ya fuera por ser hablantes de lengua indígena o por auto adscripción a este grupo social.

El marco de muestreo

Se buscó que la muestra fuera probabilística con la finalidad de que los resultados que se generaran mediante la encuesta pudieran ser generalizables a las distintas poblaciones de estudio. Para tal fin, se necesitaba que todos los elementos de las poblaciones de estudio tuvieran de antemano una probabilidad positiva y conocida de ser seleccionados.

Dado que no era posible disponer de marcos de muestreo de la población ciudadana actualizados a nivel individual, ni tampoco se disponía de marcos de muestreo específicos para cada población de los grupos mayormente expuestos a actos discriminatorios, se optó por utilizar los marcos de características geográficas y sociodemográficas disponibles a partir del XII Censo de Población y Vivienda del año 2000. Con ellos, se realizaron distintas etapas de selección hasta llegar a las viviendas y elegir ahí en la muestra a los ciudadanos que permitieran abarcar a las poblaciones de interés.

Diseño de muestreo

La muestra se eligió por separado para la parte urbana y la parte rural. Para garantizar que en la muestra quedaran contempladas unidades de diferentes zonas geográficas del país, se utilizó una clasificación que la SEDESOL ha utilizado para otras investigaciones de las entidades federativas en tres regiones.³

Se distinguieron primeramente dos grandes segmentos de población (estratos), constituidos según la condición de residencia de la población mexicana: 1) población residente en áreas urbanas y 2) población residente en áreas rurales.

³ La región norte estuvo conformada por las entidades de Baja California, Baja California Sur, Coahuila, Chihuahua, Durango, Nayarit, Nuevo León, San Luis Potosí, Sinaloa, Sonora, Tamaulipas y Zacatecas. La región Centro se integró por las entidades de Aguascalientes, Colima, Distrito Federal, Guanajuato, Guerrero, Hidalgo, Jalisco, Estado de México, Michoacán, Morelos, Puebla, Querétaro y Tlaxcala. Por último, la región Sur se compone de los estados de Campeche, Chiapas, Oaxaca, Quintana Roo, Tabasco, Veracruz y Yucatán.

³ Es decir, que si una UPM tiene cien habitantes y otra tiene trescientos habitantes, la mayor de estas dos UPM's tiene una probabilidad tres veces mayor que la pequeña de resultar elegida en la muestra.

En cada uno de los estratos se realizó una selección independiente de las muestras respectivas.

Para el estrato urbano las unidades primarias de muestreo (UPM's) estuvieron constituidas por las Áreas GeoEstadísticas Básicas (AGEB's) que ha conformado el INEGI. Las unidades secundarias de muestreo (USM's) se conformaron mediante la agrupación de viviendas identificadas a partir de los amezanamientos reconocibles en las zonas urbanas; las unidades terciarias de muestreo (UTM's) fueron las viviendas que conformaban las manzanas de las USM's seleccionadas. Esta consideración de los distintos tipos de unidades de muestreo permitió elegir a las viviendas de manera aleatoria en etapas sucesivas de selección. Para el estrato urbano se eligieron primeramente AGEB's con probabilidad proporcional al tamaño,⁴ y en dentro de esta muestra de AGEB's se seleccionaron dos manzanas de manera aleatoria. Una vez en la manzana seleccionada se eligieron al azar cuatro viviendas.

Para el estrato rural las UPM's fueron las localidades de menos de 2,500 habitantes y las USM's fueron las propias viviendas de las localidades rurales. Las localidades fueron elegidas también tomando en cuenta su tamaño poblacional, y una vez en las localidades en muestra se seleccionaron aleatoriamente siete viviendas al azar.

Dada la necesidad que se tenía de cubrir poblaciones de grupos sociales específicos y la falta de marcos de muestreo respectivos, tanto en el estrato urbano como en el rural se optó por hacer la selección de una muestra probabilística nacional de viviendas, las cuales fueron tomadas como puntos base aleatorios para elegir otros puntos de muestreo. Cada vivienda base se tomó como punto de partida para conformar un segmento de cuatro viviendas contiguas; en esos segmentos se intentaba localizar tanto a personas de las poblaciones específicas de interés, como aplicar un cuestionario correspondiente a las percepciones de la población ciudadana en torno al fenómeno de estudio.

⁴ Es decir, que si una UPM tiene cien habitantes y otra tiene trescientos habitantes, la mayor de estas dos UPM's tiene una probabilidad tres veces mayor que la pequeña de resultar elegida en la muestra.

Por otro lado, con la finalidad de potenciar la posibilidad de localizar a personas de las poblaciones específicas de interés, a partir de la información censal disponible se calculó en cada UPM a) la proporción de personas mayores de sesenta años, b) la de personas que hablan lengua indígena, c) la proporción de personas no católicas, d) la proporción de mujeres y e) la de personas con discapacidad. Esto permitió que cada UPM se clasificara en uno de dos estratos según si la proporción correspondiente a la UPM era superior o inferior a la proporción nacional correspondiente a los diferentes tipos de población específica de interés.

De esta manera, la combinación de la estratificación urbana-rural, junto con la regionalización geográfica y los tipos de estrato según superaran o no la proporción preestablecida de las poblaciones de interés, permitió conformar doce estratos que fueron tomados en cuenta en la selección de la muestra.

Con la intención de que en la muestra estuvieran contempladas unidades de los distintos niveles de marginación, siempre que fue posible, la base de datos se ordenó según el grado de marginación de las UPM's, lo que determinó una estratificación implícita según este criterio.

De lo descrito en este apartado, se puede sintetizar que el esquema de muestreo de la Encuesta Nacional sobre Discriminación fue estratificado, de conglomerados, polietápico, con probabilidad proporcional al tamaño y selección aleatoria de las unidades en las distintas etapas de muestreo.

Tamaño y precisión de la muestra

Para calcular el tamaño de muestra se consideró que a partir de la encuesta se derivarían estimaciones de proporciones y dado el esquema de muestreo descrito, era necesario tomar en cuenta un efecto de diseño debido a la correlación intraclase esperada en las unidades pertenecientes a una misma unidad de muestreo conglomerada. Al tamaño de muestra derivado de estos criterios se le aplicó un ajuste por el nivel no respuesta esperado en el levantamiento. En

consecuencia, la expresión utilizada para la determinación del tamaño de muestra base fue:

$$n = \frac{z^2 * q * efd}{er^2 * p * (1 - NR)}$$

donde:

<i>n</i>	=	Es el tamaño de muestra.
<i>z</i>	=	Percentil asociado a la confianza deseada para la muestra.
<i>p</i>	=	Estimación de la proporción.
<i>q</i>	=	(1-p)
<i>efd</i>	=	efecto de diseño resultante del cociente de las varianzas del esquema particular respecto del muestreo aleatorio simple.
<i>er</i>	=	Error relativo máximo aceptable.
<i>NR</i>	=	Tasa de No respuesta esperada.

Para una muestra con una confianza del 95 por ciento, para proporciones que varían entre 0.15 y 0.85, un efecto de diseño establecido de 1.8, un error relativo que variaba entre 0.03 y 0.167, equivalente a errores absolutos de entre 3 y 4 puntos y una tasa de no respuesta esperada del 10 por ciento, se obtuvo que el tamaño de muestra resultante para una estimación a nivel nacional fue de alrededor de 1,400 entrevistas efectivas.

Los tamaños de muestra para las poblaciones específicas fueron:

➤ Adultos mayores	700
➤ Indígenas	700
➤ Personas con discapacidad	600
➤ Minorías	700
➤ Mujeres	900

Estos tamaños de muestra permiten obtener estimaciones nacionales con una confianza del 90 por ciento, con errores relativos que variaban entre 0.042 y 0.167, equivalentes a errores absolutos de cuatro puntos.

Adicionalmente, la SEDESOL solicitó levantar una muestra de 200 personas no heterosexuales. Con esta muestra sólo se buscaba explorar algunos elementos relacionados con la potencial discriminación hacia este grupo poblacional, pero sin tener la pretensión de derivar medidas estadísticas con una confianza o precisión estadística específica.

Cabe señalar que el tamaño de muestra permite obtener estimaciones a nivel nacional con la confianza y precisión señaladas de la Encuesta Nacional sobre Discriminación, pero que la comparación por regiones sólo podrán realizarse como análisis de subdases, pues las regiones no fueron establecidas como dominios de estudio.

Finalmente, cabe señalar que los ponderadores de la encuesta fueron calculados de la manera usual, es decir, a partir del inverso de las probabilidades de selección de las unidades maestras- Estos ponderadores deberán ser utilizados para generar las estimaciones de la Encuesta Nacional sobre Discriminación, con la excepción, ya mencionada, de la población de personas no heterosexuales .que será considerada como un estudio de caso y, por lo tanto, no se ponderarán sus resultados ni se deberá intentar hacer inferencias a partir de los datos recolectados para esta población específica.