



LECTURA

SELECCIONAR UNA MUESTRA APROPIADA PARA LA
INVESTIGACIÓN



Seleccionar una muestra apropiada para la investigación

Proceso de investigación séptimo paso

Síntesis

En el capítulo se analizan los conceptos de muestra, población o universo, tamaño de muestra, representatividad de la muestra y procedimiento de selección. También presenta una tipología de muestras: probabilísticas y no probabilísticas. Explica cómo definir los sujetos que van a ser medidos, cómo determinar el tamaño adecuado de muestra y cómo proceder a obtener la muestra dependiendo del tipo de selección elegido.

Aquí el interés se centra en "quiénes", es decir, en los sujetos u objetos de estudio. Esto depende del planteamiento inicial de la investigación. Así, si el objetivo es describir el uso que hacen los niños de la televisión, lo más factible es que deberemos interrogar a un grupo de niños. También serviría entrevistar a las mamás de los niños. Escoger entre los niños o sus mamás, o ambos, dependería no sólo del objetivo de la investigación sino del diseño de la misma. En el caso de la investigación de Fernández Collado, Baptista y Elkes (1986) donde el objetivo básico del estudio era describir la relación niño-televisión, se determinó que los sujetos seleccionados para el estudio fueran niños que respondieron sobre sus conductas y percepciones relacionadas con este medio de comunicación.

En otro estudio de Greenberg, Ericson y Vlahos (1972) el objetivo de análisis era investigar las discrepancias o semejanzas en las opiniones de madres e hijos con respecto al uso de la televisión. Aquí el objetivo del estudio supuso la selección de mamás y niños, para entrevistarlos por separado, correlacionando posteriormente la respuesta de cada par madre-hijo.

Lo anterior puede parecer muy obvio, pues los objetivos de los dos ejemplos mencionados son claros. En la práctica esto no parece ser tan simple para muchos estudiantes que en propuestas de investigación y de tesis no logran una coherencia entre los objetivos de la investigación y la unidad de análisis de la misma. Algunos errores comunes se encuentran en la tabla 8.1.

Por tanto, para seleccionar una muestra, lo primero es definir la unidad de análisis (personas, organizaciones, periódicos, etcétera). El "quiénes van a ser medidos", depende de precisar claramente el problema a investigar y los objetivos de la investigación. Estas acciones nos llevarán al siguiente paso, que es delimitar una población.



En el estudio del SIDA que ya hemos mencionado, la unidad de análisis fueron los receptores directos de sangre o derivados contaminados con VIH que sobreviven y los contactos de receptores directos (cónyuge e hijos nacidos durante el periodo 1984-1987) tanto si sobreviven como si no.

8.2 ¿Cómo se delimita una población?

Una vez que se ha definido cuál será la unidad de análisis, se procede a delimitar la población que va a ser estudiada y sobre la cual se pretende generalizar los resultados. Así, una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones (Selltiz, 1974).

La muestra suele ser definida como un subgrupo de la población (Sudman, 1976)." Para seleccionar la muestra deben delimitarse las características de la población. Muchos investigadores no describen lo suficiente las características de la población o asumen que la muestra la representa automáticamente. Es frecuente que muchos estudios que únicamente se basan en muestras de estudiantes universitarios (porque es fácil aplicar en ellos el instrumento de medición, pues están a la mano) hagan generalizaciones temerarias sobre jóvenes que probablemente posean otras características sociales. Es preferible entonces, establecer claramente las características de la población, a fin de delimitar cuáles serán los parámetros muestra/es.

* Algunos investigadores usan el término universo, pero los autores preferimos utilizar el término población, ya que como Kisch (1974), consideramos que universo es más bien un término descriptivo de un conjunto infinito de datos, lo que se aplica a la población.

TABLA 8.1 ¿Quiénes van a ser medidos?: errores y soluciones.

PREGUNTA DE INVESTIGACIÓN	UNIDAD DE ANÁLISIS ERRÓNEA	UNIDAD DE ANÁLISIS CORRECTO
Discriminan a las mujeres en los anuncios de la televisión?	Mujeres que aparecen en los anuncios de televisión. Error: no hay grupo de comparación.	Mujeres y hombres que aparecen en los anuncios de televisión para comparar si categorías de análisis difieren entre los dos grupos.



¿Están los obreros del área metropolitana satisfechos con su trabajo?	Computar el número de conflictos sindicales registrados en Conciliación y Arbitraje durante los últimos 5 años. Error: la pregunta propone indagar sobre actitudes individuales y esta unidad de análisis denota datos agregados en una estadística laboral y macrosocial.	Muestra de obreros que trabajan en el área metropolitana cada uno de los cuales contestará a las preguntas de un cuestionario.
¿Hay problemas de comunicación entre padres e hijos?	Grupo de adolescentes, aplicarles cuestionario. Error: Se procedería a describir únicamente cómo perciben los adolescentes la relación con sus padres.	Grupo de padres e hijos. A ambas partes se les aplicará el cuestionario.

Lo anterior puede ilustrarse con el ejemplo de la investigación sobre el uso de la televisión por los niños. Está claro que en dicha investigación la unidad de análisis son los niños. Pero, ¿de qué población se trata?, ¿de todos los niños del mundo?, ¿de todos los niños de la República Mexicana? Sería muy ambicioso y prácticamente imposible referirnos a poblaciones tan grandes. Así en nuestro ejemplo la población fue delimitada de la siguiente manera:

FIGURA 8.1 Población niños

Limites de población

Todos los niños del área metropolitana, que cursen 4°, 5° Y 6° de primaria, en escuelas privadas y públicas del turno matutino.

Esta definición eliminó entonces a niños mexicanos que no vivieran en el área metropolitana del D.F., a los que no van a la escuela y a los menores de 9 años. Pero por otra parte permitió hacer una investigación costeable, con cuestionarios contestados por niños que ya sabían escribir y un control sobre la inclusión de niños de todas las zonas de la metrópolis, al usar la ubicación de las



escuelas como puntos de referencia y de selección. En éste y otros casos, la delimitación de las características de la población no sólo depende de los objetivos del estudio, sino de otras razones prácticas. Un estudio no será mejor por tener una población más grande; la calidad de un trabajo estriba en delimitar claramente la población con base en los objetivos del estudio.

Las poblaciones deben situarse claramente en torno a sus características de contenido, lugar y en el tiempo. Por ejemplo, en un estudio sobre los directivos de empresa en México (Baptista, 1983) y con base en las consideraciones teóricas del estudio que describe el comportamiento gerencial de los individuos y la relación de éste con otras variables de tipo organizacional se procedió a definir la población de la siguiente manera:

Nuestra población comprende a todos aquellos directores generales de empresas industriales y comerciales que en 1983 tienen un capital social superior a 30 millones de pesos, con ventas superiores a los 100 millones de pesos y/o con más de 300 personas empleadas.

En este ejemplo se delimita claramente la población, excluyendo a personas que no son directores generales, a empresas que no pertenezcan al giro industrial y comercial. Se establece también claramente que se trata de empresas medianas y grandes con base en criterios de capital y de recursos humanos. Finalmente se indica que estos criterios operaron en 1983.

Los criterios que cada investigador cumpla dependen de sus objetivos de estudio, lo importante es establecerlos claramente. Toda investigación debe ser transparente, sujeta a crítica y a réplica, y este ejercicio no es posible si al examinar los resultados, el lector no puede referirlos a la población utilizada en un estudio.

8.3 ¿Cómo seleccionar la muestra?

Hasta este momento hemos visto que se debe definir cuál será la unidad de análisis y cuáles son las características de la población. En este inciso hablaremos de la muestra o mejor dicho de los tipos de muestra, a fin de poder elegir la más conveniente para un estudio.

La muestra es, en esencia, un subgrupo de la población. Digamos que es un subconjunto de elementos que pertenecen a ese conjunto definido en sus características al que llamamos población. Esto se representa en la figura 8.2. Con frecuencia leemos y oímos hablar de "muestra representativa", "muestra al azar", "muestra aleatoria" como si con los simples términos se pudiera dar más seriedad a los resultados. En realidad, pocas veces se puede medir a toda la población, por lo que obtenemos o seleccionamos una muestra y desde luego se pretende que este



subconjunto sea un reflejo fiel del conjunto de la población. Todas las muestras deben ser representativas, por tanto el uso de este término es por demás inútil. Los términos al azar y aleatorio denotan un tipo de procedimiento mecánico relacionado con la probabilidad y con la selección de elementos, pero no logra esclarecer tampoco el tipo de muestra y el procedimiento de muestreo. Hablemos entonces de esto en los siguientes apartados.

8.3.1 Tipos de muestra

Figura 8.2

Básicamente categorizamos las muestras en dos grandes ramas: las muestras no probabilísticas y las muestras probabilísticas. En estas últimas todos los elementos de la población tienen la misma posibilidad de ser escogidos. Esto se obtiene definiendo las características de la población, el tamaño de la muestra y a través de una selección aleatoria y/o mecánica de las unidades de análisis. Imagínese el procedimiento para obtener el número premiado en un sorteo de lotería. Este número se va formando en el momento del sorteo, a partir de las esferas con un dígito que se van sacando (después de revolverlas mecánicamente) hasta formar el número, de manera que todos los números tienen la misma probabilidad de ser elegidos.

En las muestras no probabilísticas, la elección de los elementos no depende de la probabilidad, sino de causas relacionadas con las características del investigador o del que hace la muestra. Aquí el procedimiento no es mecánico, ni con base en fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de una persona o grupo de personas, y desde luego las muestras seleccionadas por decisiones subjetivas tienden a estar sesgadas. Elegir entre una muestra probabilística o una no probabilística, depende de los objetivos del estudio, del esquema de investigación y de la contribución que se piensa hacer con ella. Para ilustrar lo anterior mencionaremos varios ejemplos que toman en cuenta dichas consideraciones.

En un primer ejemplo tenemos una investigación sobre inmigrantes extranjeros en México (Baptista et al., 1988). El objetivo de la investigación es documentar las experiencias de viaje, de vida y de trabajo. Para cumplir dicho propósito se seleccionó una muestra no probabilística de personas extranjeras que por diversas razones (económicas, políticas, fortuitas) hubieran llegado a México entre 1900 y 1960. Las personas se seleccionaron a través de conocidos, de asilos, de referencias. De esta manera se entrevistó a 40 inmigrantes con entrevistas semiestructuradas que permitieron al sujeto hablar libremente sobre sus experiencias.

Comentario. En este caso una muestra no probabilística es adecuada pues se trata de un estudio con un diseño de investigación exploratorio; es decir, no es concluyente, sino que su objetivo es



documentar ciertas experiencias. Este tipo de estudio pretende generar datos e hipótesis que constituyan la materia prima para investigaciones más precisas.

Como segundo caso mencionaremos el de una investigación para saber cuántos niños han sido vacunados y cuántos no, y variables asociadas (nivel socio económico, lugar donde viven, educación) con esta conducta y sus motivaciones. Se hizo una muestra probabilística nacional de 1 600 personas y de los datos se tomaron decisiones para formular estrategias de vacunación y mensajes dirigidos a persuadir la pronta y oportuna vacunación de los niños.

Comentario. Este tipo de estudio, donde se hace una asociación entre variables, cuyos resultados servirán de información para tomar decisiones políticas que afectarán a una población, se logra por medio de una investigación por encuestas y definitivamente a través de una muestra probabilística diseñada de tal manera que los datos pueden ser generalizados a la población con una estimación precisa del error que pudiera cometerse al realizar tales generalizaciones.

Se diseñó un experimento para medir si los contenidos violentos de la televisión generan conductas antisociales en los niños. Para lograr tal objetivo se seleccionó en un colegio a 60 niños de 5 años de edad de igual nivel socioeconómico e igual nivel intelectual y se asignan aleatoriamente a 2 grupos o condiciones. 30 niños verán caricaturas pro-sociales y otros 30 verán caricaturas muy violentas. Inmediatamente después de la exposición a dichos contenidos violentos, los niños serán observados en un contexto de grupo y se medirán sus conductas violentas y prosociales.

Comentario. Ésta es una muestra no probabilística. Aunque se asignen los niños de manera aleatoria a las dos condiciones experimentales, para generalizar a la población se necesitarían repetidos experimentos. Un estudio así es valioso en cuanto a que el nivel causa-efecto es más preciso al aislar otras variables, sin embargo los datos no pueden generalizarse a todos los niños, sino a un grupo de niños con las mencionadas características. Se trata de una muestra dirigida y "clásica" de un estudio de este tipo. La selección de la muestra no es al azar, aunque la asignación de los niños a los grupos sí lo es.

En el estudio del SIDA, se encuestó a 482 personas de las 2 847 que recibieron transfusiones de sangre o derivados provenientes de "Transfusiones y Hematología, S.A."



Comentario. La muestra se obtuvo de manera fortuita, rastreando a las personas a través de archivos médicos que se encuentran en los hospitales donde los pacientes recibieron los productos hemáticos.

8.4 ¿Cómo se hace una muestra probabilística?

Resumiremos diciendo que la elección entre la muestra probabilística y una no probabilística se determina con base en los objetivos del estudio, el esquema de la investigación y el alcance de sus contribuciones. Las muestras probabilísticas tienen muchas ventajas, quizá la principal es que puede medirse el tamaño de error en nuestras predicciones. Puede decirse incluso que el principal objetivo en el diseño de una muestra probabilística es reducir al mínimo este error al que se le llama error estándar (Kish, 1965).

Las muestras probabilísticas son esenciales en los diseños de investigación por encuestas en las que se pretende hacer estimaciones de variables en la población, estas variables se miden con instrumentos de medición (capítulo 9) y se analizan con pruebas estadísticas para el análisis de datos, donde se presupone que la muestra es probabilística, donde todos los elementos de la población tienen una misma probabilidad de ser elegidos. Los elementos muestrales tendrán valores muy parecidos a los de la población, de manera que las mediciones en el subconjunto, nos darán estimados precisos del conjunto mayor. La precisión de dichos estimados depende del error en el muestreo, que se puede calcular pues hay errores que dependen de la medición y estos errores no pueden ser calculados matemáticamente.

Para hacer una muestra probabilística es necesario entender los siguientes términos y sus definiciones:

8.4.1

La *población* a la que llamaremos N , es un conjunto de elementos.

La *muestra* a la que denominaremos n , es un subconjunto de la población N .

En una *población* N (previamente delimitada por los objetivos de la investigación) nos interesa establecer expresiones numéricas de las características de los elementos de N .

Nos interesa conocer valores promedio en la población, el cual se expresa como:

$y =$ es decir se refiere al valor de una variable determinada (Y) que nos interesa conocer.

También nos interesa conocer:



V = es decir la varianza de la población con respecto a determinadas variables.

Como los valores de la población no se conocen, seleccionamos una muestra n y a través de estimados «La muestra, inferimos valores en la población. Y será el valor de Y el cual desconocemos. En la muestra y es un estimado promedio que podemos determinar, Sabemos que en nuestra estimación habrá una diferencia ($Y - y = 1$) es decir, habrá un error, el cual dependerá del número de elementos muestreados. A dicho error le llamaremos error estándar = Se

Se = es la desviación estándar de la distribución muestral y representa la fluctuación de y .

$(SE)^2$ = el error estándar al cuadrado es la fórmula que nos servirá para calcular la varianza (V) de la población (N). Y la varianza de la muestra (n) será la expresión S^2 .

S^2 = varianza de la muestra, la cual podrá determinarse en términos de probabilidad donde

$$S^2 = p(1-p)$$

Para una muestra probabilística necesitamos principalmente dos cosas: determinar el tamaño de la muestra (n) y seleccionar los elementos muestrales, de manera que todos tengan la misma posibilidad de ser elegidos. Para lo primero, daremos una fórmula que contiene las expresiones ya descritas. Para lo segundo, necesitamos un marco de selección adecuado y un procedimiento que permita la aleatoriedad en la selección. Hablaremos de ambas cosas en los siguientes apartados.

El tamaño de la muestra

Cuando se hace una muestra probabilística, uno debe preguntarse: dado que una población es de N , ¿cuál es el menor número de unidades muestrales (personas, organizaciones, capítulos de telenovelas, etcétera), que necesito para conformar una muestra (n) que me asegure un error estándar menor de .01?

$$1. n' = \frac{S^2}{V^2} = \text{Tamaño provisional de la muestra} = \frac{\text{varianza de la muestra}}{\text{varianza de la población}}$$

Se corrige después con otros datos, ajustándose si se conoce el tamaño de la población.

La solución a esta pregunta pretende encontrar la probabilidad de ocurrencia de y que mi estimado de y se acerque a Y , el valor real de la población. Si nosotros establecemos el error estándar y fijamos .01, sugerimos que esta fluctuación promedio de nuestro estimado y con respecto a los



valores reales de la población y , no sea $> .01$, es decir que de 100 casos, 99 veces mi predicción sea correcta y que el valor de y se sitúe en un intervalo de confianza que comprenda el valor de Y .

Resumiendo, para una determinada varianza (V) de y , ¿qué tan grande debe ser mi muestra? Esto puede determinarse en dos pasos:

$$n' = \frac{S^2}{V^2} \cdot (1 + \frac{n'}{N})$$

Pongamos el siguiente ejemplo. En el ejemplo que ya habíamos dado en este capítulo, delimitamos una población diciendo que para un estudio de directores generales consideramos a "todos aquellos directores generales de empresas industriales y comerciales que en 1983 tienen un capital social superior a 30 millones de pesos, con ventas superiores a los 100 millones de pesos y/o con más de 300 personas empleadas". Con estas características se precisó que la población era de $N = 1176$ directores generales ya que 1 176 empresas reunían las mencionadas características. ¿Cuál es entonces el número de directores generales n que se tiene que entrevistar, para tener un error estándar menor de $.015$, y dado que la población total es de 1176?

N = tamaño de la población de 1 176 empresas.

y = valor promedio de una variable = 1, un director general por empresa. Se = error estándar = $.015$, determinado por nosotros.

V^2 = varianza de la población. Su definición (Se) cuadrado del error estándar. S^2 = varianza de la muestra expresada como la probabilidad de ocurrencia de y .

n' = tamaño de la muestra sin ajustar. n = tamaño de la muestra

Sustituyendo tenemos que:

$$n' = \frac{S^2}{V^2} \cdot (1 + \frac{n'}{N})$$
$$S^2 = p(1 - p) = .9(1 - .9) = .09 \quad V = (.015)^2 = .000225$$
$$n' = .09 \cdot .000225$$



$$= 400$$

$$2$$

$$n' n = \dots =$$

$$1 + n'/N$$

$$400 = 298$$

$$1 + \frac{400}{1176}$$

$$n = 298$$

Es decir, para nuestra investigación, necesitaremos una muestra de 298 directores generales. Esto es el primer procedimiento para obtener la muestra probabilística: determinar su tamaño, con base en estimados de la población. El segundo procedimiento estriba en cómo y de dónde seleccionar a esos 298 sujetos.

8.4.2 Muestra probabilística estratificada

El ejemplo anterior corresponde a una muestra probabilística simple. Determinamos en este caso que el tamaño de la muestra sería de $n = 298$ directivos de empresa. Pero supongamos que la situación se complica y que deberemos estratificar esta n a fin de que los elementos muestrales o unidades de análisis posean un determinado atributo. En nuestro ejemplo este atributo es el giro de la empresa. Es decir, cuando no basta que cada uno de los elementos muestrales tengan la misma probabilidad de ser escogidos, sino que además es necesario estratificar la muestra en relación a estratos o categorías que se presentan en la población y que aparte son relevantes para los objetivos del estudio, se diseña una muestra probabilística estratificada. Lo que aquí se hace es dividir a la población en subpoblaciones o estratos y se selecciona una muestra para cada estrato.

La estratificación aumenta la precisión de la muestra e implica el uso deliberado de diferentes tamaños de muestra para cada estrato, "a fin de lograr reducir la varianza de cada unidad de la media muestral" (Kish, 1965). Dice Kish (p. 92) en su libro de muestreo que, en un número determinado de elementos muestrales $n = \sum h$, la varianza de la media muestral y puede reducirse al mínimo si el tamaño de la muestra para cada estrato es proporcional a la desviación estándar dentro del estrato.

Esto es,

$$n$$

$$f_h = \frac{\quad}{N} = KSh$$



En donde la muestra n será igual a la suma de los elementos muestrales nh.

Es decir, el tamaño de n y la varianza de V, pueden minimizarse si calculamos "submuestras" proporcionales a la desviación estándar de cada estrato. Esto es:

$$n_h = k \cdot S_h$$

N_h

En donde n_h y N_h son muestra y población de cada estrato y S_h es la desviación estándar de cada elemento en un determinado estrato. Entonces tenemos que:

$$k = \frac{n}{N}$$

Siguiendo con nuestro ejemplo de los directores de empresa, la población es de 1 176 directores de empresa y el tamaño de muestra es $n = 298$. ¿Qué muestra necesitaremos para cada estrato?

$$k = \frac{n}{N} = \frac{298}{1176} = 0.2534$$

De manera que el total de la subpoblación se multiplicará por esta fracción constante a fin de obtener el tamaño de muestra para el estrato. Sustituyendo tenemos que:

$$n_h = k \cdot N_h$$

En algunos casos donde el investigador se ve limitado por recursos financieros, por tiempo, por distancias geográficas o por una combinación de éstos y otros obstáculos, se recurre al muestreo por racimos. En este tipo de muestreo se reducen costos, tiempo y energía al considerar que muchas veces las unidades de análisis se encuentran encapsuladas o encerradas en determinados lugares físicos o geográficos a los que se denomina racimos. Para dar algunos ejemplos tenemos la tabla 8.3. En la primera columna se encuentran unidades de análisis

TABLA 8.2 Muestra probabilística estratificada de directores de empresa.

TOTAL POBLACIÓN* (FH) = 1176

n (FH) = 298

ESTRATO POR GIRO

DIRECTORES GENERALES DE EMPRESA DEL GIRO

MUESTRA

1 2 3 4 5 6 7 8 9

10

Extractivo y siderúrgico Metal mecánicas Alimentos. bebidas. tabaco Papel y artes gráficas Textiles

Eléctricas y electrónicas Automotriz Químico-farmacéutica



Otras empresas transformación Comerciales

53 109 215 87 98 110

81 221 151 51 N = I 176

13 28 55 22 25 28 20 56 38 13

n = 298

por ejemplo:

Nh = 53 directores de empresas extractivas corresponde a la población total de este giro.

fh = .2534 es la fracción constante.

nh = 13 es el número redondeado de directores de empresa del giro extractivo que tendrán que entrevistarse.

* Fuente de Industridata. 1982.

Tabla 8.3 ejemplos de racimos.

Unidad de análisis

Posibles racimos

Adolescentes Obreros Amas de casa Niños

Personajes de televisión

Preparatorias Industrias Mercados Colegios

Programas de televisión

Que frecuentemente vamos a estudiar en ciencias sociales. En la segunda columna, sugerimos posibles racimos donde se encuentran dichos elementos.

Muestrear por racimos implica diferenciar entre la unidad de análisis y la unidad muestral. La unidad de análisis indica quiénes van a ser medidos, o sea, el sujeto o sujetos a quienes en última instancia vamos a aplicar el instrumento de medición. La unidad muestral (en este tipo de muestra) se refiere al racimo a través del cual se logra el acceso a la unidad de análisis. El muestreo por racimos supone una selección en dos etapas, ambas con procedimientos probabilísticos. En la primera, se seleccionan los racimos, siguiendo los ya señalados pasos de una muestra probabilística simple o estratificada. En la segunda, y dentro de estos racimos se seleccionan a los sujetos u objetos que van a ser medidos. Para ello se hace una selección que asegure que todos los elementos del racimo tienen la misma probabilidad de ser elegidos. A continuación daremos un ejemplo que comprenda varios de los procedimientos descritos hasta ahora y que ilustra la manera como frecuentemente se hace una muestra probabilística en varias etapas.

Problema de investigación: Una estación de radio local necesita saber con precisión a fin de planear sus estrategias cómo usan la radio los adultos de una ciudad de 2 500 000 habitantes. Es



decir, qué tanto radio escuchan, a qué horas, qué contenidos prefieren y sus opiniones con respecto a los programas noticiosos. Procedimientos: Se diseñará un cuestionario que indague estas áreas sobre uso del radio. Los cuestionarios se aplicarán por entrevistadores a una muestra de sujetos adultos.

Población: Todos aquellos sujetos hombres o mujeres de más de 21 años de edad, y que vivan en una casa o departamento propio o rentado de la ciudad X.

Diseño por racimos: Los directivos de la estación de radio desconocen el número total de sujetos con las características arriba señaladas. Sin embargo, nos piden que diseñemos una muestra que abarque a todos los sujetos adultos de la ciudad, adultos por edad cronológica y por ser jefes de familia, es decir, excluye a los adultos dependientes.

Se recurre entonces a la estrategia de seleccionar racimos y se considera el uso de un mapa actualizado de la ciudad, el cual muestra que en dicha ciudad hay 5 000 cuadras. Las cuadras se utilizarán como racimos, es decir como unidades muestrales a partir de las cuales obtendremos en última instancia a nuestros sujetos adultos. Lo primero entonces es determinar: ¿cuántas cuadras necesitaremos muestrear, de una población total de 5 000 cuadras, si queremos que nuestro error estándar sea no mayor de 0.15 y con una probabilidad de ocurrencia del 50%?

Tenemos entonces que $n' = .2 \dots$ para una muestra probabilística simple.

$$n = \frac{S^2}{V^2} = \frac{p(1-p) \cdot .5(1-.5) = .25}{(.015)^2 \text{ estándar } .00025} = 1111.11$$

$$n = \frac{n'}{1 + n'/N} = \frac{1111.11}{1 + 1111.11/5000} = 909.0902$$

$$n = 909$$

Necesitaremos una muestra de 909 cuadras para estimar los valores de la población con una probabilidad de error menor a .01.

* Sabemos que la población $N = 5\ 000$ cuadras está dividida por previos estudios de acuerdo con 4 estratos socioeconómicos, que categorizan esa población según el ingreso mensual promedio de sus habitantes, de manera que se distribuyen como sigue:



<i>Estrato</i>	<i>No. de cuadras</i>
1	270
2	1940
3	2000
4	790
	T=5000

¿Cómo distribuiremos los 909 elementos muestrales de n, para optimizar la muestra, de acuerdo con la distribución de la población en los 4 estratos socio económicos?

En principio tenemos que de 5 000 cuadras se seleccionarán 50 del estrato 1 1353 del estrato 2 363 del estrato 3 y 143 del estrato 4. Esta selección comprende la selección de los racimos, los cuales se pueden numerar y elegir aleatoriamente hasta completar el número de cada estrato (ver sección 8.4.2). En una última etapa, se seleccionan los sujetos dentro de cada racimo. Este procedimiento también se hace de manera aleatoria, hasta lograr un número de sujetos determinados en cada racimo. En el próximo inciso describiremos dicho procedimiento.

* Estratificación de la muestra: $f_h = \frac{N_h}{N} \cdot n = k \cdot n$

N

$f_h = \frac{N_h}{N} \cdot n = \frac{1818}{5000} \cdot 909$

<i>Estrato</i>	<i>No. de cuadras</i>	$f_h = \frac{N_h}{N} \cdot n$	n_h
1	270	$(\frac{1}{5} \cdot 909)$	50
2	1940	$(\frac{1}{5} \cdot 909)$	353
3	2000	$(\frac{1}{5} \cdot 909)$	363
4	790	$(\frac{1}{5} \cdot 909)$	143

<i>Estrato</i>	<i>No cuadras</i>	n_h	Número de hogares sujeto en cada cuadra	<i>rota/ de hogares por estado</i>
1	270	50	20	1000
2	1940	353	20	7060
3	2000	363	20	7220
4	790	143	20	2 860
	N=5000	n=909		11 840



8.5 ¿Cómo se lleva a cabo el procedimiento de selección?

Cuando iniciamos nuestra discusión de la muestra probabilística, señalamos que los tipos de muestra dependen de dos cosas: del tamaño de la muestra y del procedimiento de selección.

De lo primero, hemos hablado con todo detalle, de lo segundo hablaremos ahora. Se determina el tamaño de la muestra n , pero ¿cómo seleccionar los elementos muestrales? Se precisa el número de racimos necesario y ¿cómo se seleccionan los sujetos dentro de cada racimo? Hasta el momento sólo hemos dicho que los elementos se eligen aleatoriamente, pero ¿cómo se hace esto?

Las unidades de análisis o los elementos muestrales se eligen siempre aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido. Pueden usarse 3 procedimientos de selección:

8.5.1 Tómbola

Muy simple y no muy rápido, consiste en numerar todos los elementos muestrales del 1 al n . Hacer fichas, una por cada elemento, revolverlas en una caja, e ir sacando n fichas, según el tamaño de la muestra. Los números elegidos al azar conformarán la muestra.

Así en la tabla 8.2, tenemos que, de una población $N = 53$ empresas extractivas y siderúrgicas, se necesita una muestra $n = 13$ de directivos generales de dichas empresas. En una lista se puede numerar cada una de estas empresas. En fichas aparte se sortea cada uno de los 53 números. Los números obtenidos se verifican con los nombres y direcciones de nuestra lista, para precisar los que serán sujetos de análisis.

8.5.2 Números random o números aleatorios

El uso de números random no significa la selección azarosa o fortuita, sino la utilización de una tabla de números que implica un mecanismo de probabilidad muy bien diseñado. Los números random de la Corporación Rand, fueron generados con una especie de ruleta electrónica. Existe una tabla de un millón de dígitos, publicada por esta corporación; partes de dicha tabla se encuentran en los apéndices de muchos libros de estadística. Estas tablas son como]0 muestra la tabla 8.4 y el apéndice 5.



TABLA 8.4 Números random

26804			45610	22879	72538 701		67942	52846	
90720			94756	18124				35943	67290
85027			41416	48521	15720 902		10822	93074	
09362			96702	20772	12069 499		12510	64899	
64590			79237	82158	04553 930		72279	01916	
06432			20507	92817	39800 988		81860	68065	
02101			88949	89312	82716 347		58424	69700	
19337			62194	08574	81896 003		66220	16494	
75277			35832	41655	27155 951		06649	53040	
59535			88202	63899				06641	97291
76310			27804	48889	80070 646		04232	84008	
12805			67060	88413	31883 792		68989	80233	
32242			67123	40637	14102 555		80593	64642	
16212			13252	78974				36042	75492
75362			24828	59345		42613			
				53717		11230			
				39025	16688 695	81885		13098	
			16663	39489	18400 531	92087			
				99358		61696			
				00440		62235			

Tabla 8.4 (Continuación).

97021	23763	18491	65056	95283	92232 8669	78699	79666	88574	
25469	63708	78718	35014	40387	15921	58080	03936	15953	59658
40337	48522	11418	00090	41779	54499 0862	49092	65431	11390	
33491	98685	92536	51626	85787	47841	95787	70139	42383	44187
44764	14986	16642	19429	01960	22833	80055	39851	47350	70337
<i>Fuente:</i> Rand									

Siguiendo el ejemplo del inciso anterior, determinaremos una muestra de 909 manzanas o cuadras, y a partir de este número se determinó una submuestra para cada estrato. Véase que para el estrato 1, la población es de 270 manzanas. Numeramos en nuestro listado o mapa las 270 cuadras y seleccionamos (a partir de la tabla de números random) los 50 casos que constituirán nuestra muestra.

Se eligen aquellos casos que se dictaminen en la tabla de números random, hasta completar el tamaño de la muestra. Los números pueden recorrerse hacia arriba, hacia abajo, horizontalmente.



Al fin siempre se logra que cada elemento muestral tenga la misma probabilidad de ser elegido. Se eligen aquellos números que contenga el listado. Si en nuestro ejemplo la población es de 270, se escogen los 3 últimos dígitos y se procede de la siguiente manera a seleccionar los casos hasta completar el número de elementos muestrales.

»<		/	.•..		
78986	45961	28281	82933	24786	55586
83830	59025	40379	99989	63822	99974
(1)30226	19863	(5)95039	08909	(7)48 197	(8)23270
(2)02073	(4)59042	26440	(6)16/61	14496	24786
(3)05250	47552	95659	92356	13334	23471

8.5.3 Selección sistemática de elementos muestrales

Este procedimiento de selección es muy útil y fácil de aplicar e implica seleccionar dentro de una población N un número n de elementos a partir de un intervalo K.

K es un intervalo que va a estar determinado por el tamaño de la población y el tamaño de la muestra. De manera que tenemos que $K = N/n$, en donde K = es un intervalo de selección sistemática N = es la población.

n = es la muestra

Ilustramos los conceptos anteriores con un ejemplo. Supongamos que se quiere hacer un estudio sobre varios aspectos de la publicidad en México. Específicamente se pretende medir qué número de mensajes informativos y qué número de mensajes motivacionales tienen los comerciales en la televisión mexicana. Para tal efecto supongamos que los investigadores consiguen videocasetes con todos los comerciales que han pasado al aire en los diferentes canales de televisión, durante un periodo de tres años. Quitando los comerciales repetidos, se tiene una población de $N = 1\ 548$ comerciales. Se procede con este dato a determinar qué número de comerciales necesitamos analizar para generalizar a toda la población nuestros resultados con un error estándar no mayor de .015. Con la fórmula que ya hemos dado en la lección 8.4.3 de este capítulo tenemos que si $p = .5$, entonces $S^2 = p(1 - p) = .5 - (.5) = .25$

$n' = \sim \sqrt{2}$

Si necesitamos una muestra de $n = 647$ comerciales, podemos utilizar para la selección al intervalo K donde:



$K = \sim = 1\ 548 = 2.39 = 3'''$ (redondeando) $n = 647$

El intervalo $l/K = 3$ indica que cada tercer comercial l/K será seleccionado hasta completar $n = 647$.

"2.39 se redondea para que sea un íntegro. Véase Kísh (1969), pp. 115-117.

8.6 Los listados y otros marcos muestrales

La selección sistemática de elementos muestrales l/K se puede utilizar al elegir los elementos de 11 para cada estrato y/o para cada racimo. La regla de probabilidad, según la cual cada elemento de la población debe tener la misma probabilidad de ser elegido, se cumple empezando la selección de l/K al azar. Siguiendo nuestro ejemplo, no empezamos a elegir de los 1 548 comerciales grabados, el 1, 3, 6, 9 ... sino que procuramos que el comienzo sea determinado por el azar. Así, en este caso, podemos tirar unos dados y si en sus caras muestran 1, 6, 9, empezaremos en el comercial 169 y seguiremos: 169,172,175 ... l/K ... volver a empezar por los primeros si es necesario. Este procedimiento de selección es poco complicado y tiene varias ventajas: cualquier tipo de estratos en una población X , se verán reflejados en la muestra. Asimismo, la selección sistemática logra una muestra proporcionada, pues por ejemplo tenemos que el procedimiento de selección l/K nos dará una muestra con nombres que inician con las letras del abecedario en forma proporcional a la letra inicial de los nombres de la población.

Las muestras probabilísticas requieren la determinación del tamaño de la muestra y de un proceso de selección aleatoria que asegure que todos los elementos de la población tengan la misma probabilidad de ser elegidos. Todo esto lo hemos visto, pero nos falta discutir sobre algo esencial que precede a la selección de una muestra: el listado, el marco muestra!. El listado se refiere a una lista existente o a una lista que es necesario confeccionar "ad hoc", con los elementos de la población y a partir de la cual se seleccionarán los elementos muestrales. El segundo término se refiere a un marco de referencia que nos permita identificar físicamente los elementos de la población, la posibilidad de enumerarlos y por ende, proceder a la selección de los elementos muestrales.

Los listados con base en listas existentes sobre una población pueden ser variados: el directorio telefónico, la lista de miembros de una asociación, directorios especializados, las listas oficiales de escuelas de la zona, las listas de las canciones de éxito publicadas por una revista, la lista de alumnos de una universidad, etcétera. En todo caso hay que tener en cuenta lo completo de una lista, su exactitud, veracidad, su calidad y su nivel de cobertura en relación con el problema a investigar y la población que va a medirse, ya que todos estos aspectos influyen en la selección de la muestra.



Por ejemplo, para algunas encuestas se considera que el directorio telefónico es muy útil. Sin embargo hay que tomar en cuenta que muchos teléfonos no aparecerán porque son privados o que hay hogares que no tienen teléfono. La lista de socios de una asociación como Canacintra (Cámara Nacional de la Industria de la Transformación) puede servirnos si el propósito del estudio es, por ejemplo, conocer la opinión de los asociados con respecto a una medida gubernamental. Pero si el objetivo de la investigación es el análisis de opinión del sector patronal del país, el listado de una asociación no será adecuado por varias razones, como: hay otras asociaciones patronales, la Canacintra representa solamente el sector de la Industria de Transformación, las asociaciones son voluntarias y no todo patrón o empresa pertenece a ellas. Lo correcto en este caso, sería construir una nueva lista, con base en los listados existentes de las asociaciones patronales, eliminando de dicha lista los casos duplicados, suponiendo que una o más empresas pudieran pertenecer a dos asociaciones al mismo tiempo: como director a la Coparmex (Confederación Patronal de la República Mexicana) y como empresa a la ANIQ (Asociación Nacional de Ingenieros Químicos).

Hay listas que proporcionan una gran ayuda al investigador. Pensamos en directorios especializados como el Industridata que enlista a las empresas mexicanas medianas y grandes, el directorio de la ciudad de México por calles, el directorio de medios, que enlista casas productoras, estaciones de radio y televisión, periódicos y revistas. Este tipo de directorios realizados por profesionales son útiles al investigador pues representan una compilación (sujetos, empresas, instituciones), resultado de horas de trabajo e inversión de recursos. Recomendamos pues utilizarlos cuando sea pertinente, tomando en cuenta las consideraciones que estos directorios hacen en su introducción y que revelan a qué año pertenecen los datos, cómo los obtuvieron (exhaustivamente, por cuestionarios, por voluntarios) y muy importante, quiénes y por qué quedan excluidos del directorio.

En México se cuenta también con directorios de anunciantes en publicidad y mercadotecnia como el publicado por Mercamétrica Ediciones, S. A. Frecuentemente es necesario construir listas ad hoc, a partir de las cuales se seleccionarán los elementos que constituirán las unidades de análisis en una determinada investigación. Por ejemplo, en la investigación de La televisión y el niño (Fernández Collado et al., 1986) se hizo una muestra probabilística estratificada por racimo, donde en una primera etapa se relacionaron escuelas para en última instancia llegar a los niños. Pues bien, para tal efecto se consiguió una lista de las escuelas primarias del Distrito Federal. Cada escuela tenía un código identificable por medio del cual se eliminaron las escuelas para niños



atípicos. Este listado contenía además información sobre cada escuela, sobre su ubicación (calle y colonia), sobre su propiedad (pública o privada).

Con ayuda de otro estudio que catalogaba en diferentes estratos socioeconómicos a las colonias del Distrito Federal con base en el ingreso promedio de la zona, se hicieron 8 listas:

- 1 Escuela pública clase A
- 2 Escuelas privadas clase A 3
- Escuelas públicas clase B 4
- Escuelas privadas clase B 5
- Escuelas públicas clase C 6
- Escuelas privadas clase C 7
- Escuelas públicas clase D 8
- Escuelas privadas clase D

Cada lista representaba un estrato de la población y de cada una de ellas se seleccionó una muestra de escuelas.

No siempre existen listas que permitan identificar nuestra población. Será necesario recurrir a otros marcos de referencia que contengan descripciones del material, organizaciones o sujetos que serán seleccionados como unidades de análisis. Algunos de estos marcos de referencia son los archivos, los mapas, volúmenes de periódicos empastados en una biblioteca o las horas de transmisión de varios canales de televisión. De cada una de estas instancias daremos ejemplos con más detalles.

8.6.1 Archivos

Un jefe de reclutamiento y selección de una institución quiere precisar si algunos datos que se dan en una solicitud de trabajo están correlacionados con el ausentismo del empleado. Es decir, si a partir de datos como edad, sexo, estado civil y duración en otro trabajo, puede predecirse que alguien tenderá a ausentarse. Para establecer correlaciones se considerarán como población a todos los sujetos contratados durante 10 años. Se relacionan sus datos en la solicitud de empleo con los registros de faltas.

Como no hay una lista elaborada de estos sujetos, el investigador decide acudir a los archivos de las solicitudes de empleo. Estos archivos constituyen su marco muestral a partir del cual obtendrá la muestra. Calcula el tamaño de la población, obtiene el tamaño de la muestra y selecciona



sistemáticamente cada elemento I/K (ver sección 8.5.3), cada solicitud que será analizada. Aquí el problema que surge es que en el archivo hay solicitudes de gente que no fue contratada, y por tanto, no pueden ser consideradas en el estudio.

En este caso y en otros donde no todos los elementos del marco de referencia o de una lista (por ejemplo nombres en el directorio que no corresponden a una persona física) los especialistas en muestreo (Kish, 1965; Sutman, 1976) no aconsejan el reemplazo, con el siguiente elemento, sino simplemente no tomar en cuenta ese elemento, es decir como si no existiera, continuándose con el intervalo de selección sistemática.

8.6.2 Mapas

Los mapas son muy útiles como marco de referencia en muestras de racimo. Por ejemplo, un investigador quiere saber qué motiva a los compradores en una determinada tienda de autoservicio. A partir de un lista de tiendas de cada cadena competidora marca, sobre un mapa de la ciudad, todas las tiendas de autoservicios, las cuales constituyen una población de racimos, pues en cada tienda seleccionada, entrevistará a un número n de clientes. El mapa le permite ver la población (tiendas autoservicio) y su situación geográfica, de manera que eligió zonas donde coexistan tiendas de la competencia, para asegurarse de que el consumidor de la zona tenga todas las posibles alternativas.

8.6.3 Volúmenes

En este ejemplo Supongamos que un estudioso del periodismo quiere hacer un análisis de contenido de los editoriales de los tres principales diarios de la ciudad durante el porfiriato. El investigador va a la Hemeroteca Nacional y encuentra que los diarios son encuadernados por trimestre y año, lo cual le proporciona un marco de referencia ideal a partir del cual seleccionará n volúmenes para su análisis. Supongamos, sin embargo, que encuentra que el volumen X que contiene el periódico El Hijo del Ahuizote (enero-marzo 1899), falta en la Hemeroteca. ¿Qué hace? Pues redefine la población, manifestando explícitamente que de N volúmenes tiene 99% de los elementos y a partir de este nuevo número de N calculó su muestra n y la seleccionó.

8.6.4 Horas de transmisión

En un estudio de Portilla y Solórzano (1982), los investigadores querían hacer un análisis de anuncios en la Televisión mexicana. Las emisoras no proporcionan una lista de anuncios ni sus horas de transmisión. Por otra parte, saldría muy caro grabar todos los anuncios a todas horas y sería imposible estar frente al televisor para hacerlo.



Ante la imposibilidad de tener un listado de comerciales, se hicieron listados que identificaron cada media hora de transmisión televisiva en cada canal 2, 4, 5, 9 Y 13 de las 7 a las 24 horas durante siete días de una semana de octubre de 1982. La población estaba constituida del número total de medias horas de transmisión televisiva, $N = 1\ 190$ horas. Esta población se dividió en estratos: mañana, mediodía, tarde y noche, y se procedió a calcular el tamaño de la muestra tomando en consideración que por cada media hora de transmisión hay 6 minutos de comerciales (De Noriega, 1979).

Se calculó el número de medias horas que se seleccionarían para obtener una muestra n . Una vez obtenido el tamaño de la muestra, se seleccionaron aleatoriamente n medias horas y, por último, se grabaron y analizaron únicamente aquellos comerciales contenidos en las medias horas seleccionadas al azar Y que representaron diferentes canales y segmentos del día. El punto en este ejemplo es la construcción concreta de un marco muestral que permitiera el análisis de una muestra probabilística de comerciales.

8.7 Tamaño óptimo de una muestra y el teorema del límite central

Las muestras probabilísticas requieren dos procedimientos básicos: 1. la determinación del tamaño de la muestra y 2. la selección aleatoria de los elementos muestrales. El primer procedimiento, lo hemos descrito en su modalidad más simple, en la sección 8.4.1 de este capítulo. Precisar adecuadamente el tamaño de la muestra puede tornarse muy complejo dependiendo del problema de investigación y la población a estudiar. Para el alumno y el lector en general, será muy útil comparar qué tamaño de muestra han utilizado otros investigadores en ciencias sociales. Para tal efecto reproducimos las siguientes tablas separadas por Sudman (1976), que indican el tamaño de la muestra más utilizada por los investigadores, según sus poblaciones (nacionales o regionales) y según los subgrupos que quieren estudiarse en ellas.

TABLA 8.6 Muestras frecuentemente utilizadas en investigaciones nacionales y regionales según área de estudio.

TIPO DE ESTUDIO

NACIONALES

REGIONALES

Económicos Médicos Conductas Actitudes Experimentos de laboratorio

1000+ 1000+ 1000+ 1000+

100 500 700 -300 700- 400 100



En esta tabla vemos que el tipo de estudio en poco determina el tamaño de la muestra. Más bien interviene en la decisión de que sean muestras nacionales o regionales. Las muestras nacionales, es decir, muestras que representan a la población de un país son por lo común de más de 1 000 sujetos. La muestra del estudio "¿Cómo somos los mexicanos?" (Hernández Medina, Harro et al., 1987) consta de 1 837 sujetos repartidos de la siguiente manera:

Frontera y norte Centro (sin D.F.) Sur-sureste Distrito Federal

696 426 316 299

1837

Las muestras regionales (por ejemplo las que representen al área metropolitana) algún estado del país o algún municipio o región son típicamente más pequeñas con rangos de 700 a 400 sujetos.

El tamaño de una muestra tiende más a depender del número de subgrupos que nos interesan en una población. Por ejemplo, podemos subdividirla en hombres y mujeres de 4 grupos de edad o, aún más, en hombres y mujeres de 4 grupos de edad en cada uno de 5 niveles socioeconómicos. Si éste fuera el caso estaríamos hablando de 40 subgrupos y por ende de una muestra mayor. En la siguiente tabla se describen típicas muestras según los subgrupos bajo estudio, según su alcance (estudios nacionales o estudios especiales o regionales) y según su unidad de análisis, es decir se trata de sujetos o de organizaciones, en esta última instancia el número de la muestra se reduce, ya que éstas representan casi siempre una gran fracción de la población total.

TABLA 8.7 Muestras típicas de estudios sobre poblaciones humanas y organizaciones.

FIGURA 8.3

NÚMERO DE SUBGRUPOS

POBLACIÓN DE SUJETOS U HOGARES

POBLACIONES DE ORGANIZACIONES

Ninguno-pocos Promedio Muchos

Nacionales 1000-1500 1500-2500 2500- +

200-500 500-1000 1000 +

Regionales 200-500 500-1000 1000 +

Nacionales

Regionales 50-200 200-500 500+

Estas tablas (Sudman 1976: 86-87) fueron construidas con base en artículos de investigación publicados en revistas especializadas y nos dan una idea de las muestras que utilizan otros investigadores, de manera que pueden ayudar al investigador a precisar el tamaño de su muestra.



Recordemos que lo óptimo de una muestra depende de cuánto se aproxima su distribución a la distribución de las características de la población. Esta aproximación mejora al incrementarse el tamaño de la muestra. La "normalidad" de la distribución en muestras grandes no obedece a la normalidad de la distribución de una población. Al contrario, la distribución de las variables en estudios de ciencias sociales están lejos de ser normales. Sin embargo, la distribución de muestras de 100 o más elementos tienden a ser normales y esto sirve para el propósito de hacer estadística inferencial sobre los valores de una población. A esto se le llama teorema de límite central.

Distribución normal: esta distribución en forma de campana se logra generalmente con muestras de 100 o + Unidades muestrales y es útil y necesaria cuando se hacen inferencias de tipo estadístico.

Ésta es la distribución de una población; es anormal y, sin embargo, la distribución de una muestra de esta población de más de 100 casos tenderá a distribuirse normalmente. Esta tendencia (teorema del límite central) permite estimar los valores de la población, a partir de la inferencia estadística.

8.8 ¿Cómo son las muestras no probabilísticas?

Las muestras no probabilísticas, también llamadas muestras dirigidas, suponen un procedimiento de selección informal y un poco arbitrario. Aún así se utilizan en muchas investigaciones y a partir de ellas se hacen inferencias sobre la población. Es como si juzgásemos el sabor de un cargamento de limones, solamente probando alguno, como si para "muestra bastase un botón". La muestra dirigida selecciona sujetos "típicos" con la vaga esperanza de que serán casos representativos de una población determinada.

La verdad es que las muestras dirigidas tienen muchas desventajas. La primera es que, al no ser probabilísticas, no podemos calcular con precisión el error estándar, es decir, no podemos calcular con qué nivel de confianza hacemos una estimación. Esto es un grave inconveniente si consideramos que la estadística inferencial se basa en teoría de la probabilidad, por lo que pruebas estadísticas en muestras no probabilísticas tienen un valor limitado y relativo a la muestra en sí, mas no a la población. Es decir, los datos no pueden generalizarse a una población, que no se consideró ni en sus parámetros, ni en sus elementos para obtener la muestra. Recordemos que, en las muestras de este tipo, la elección de los sujetos no depende de que todos tengan la misma probabilidad de ser elegidos, sino de la decisión de un investigador o grupo de encuestadores.



La ventaja de una muestra no probabilística es su utilidad para determinado diseño de estudio que requiere no tanto una "representatividad de elementos de una población, sino una cuidadosa y controlada elección de sujetos con ciertas Características especificadas previamente en el planteamiento del problema". Hay varias clases de muestras dirigidas y se definirán a continuación.

8.8.1 La muestra de sujetos voluntarios

Las muestras de sujetos voluntarios son frecuentes en ciencias sociales y ciencias de la conducta. Se trata de muestras fortuitas, utilizadas también en la medicina y la arqueología donde el investigador elabora conclusiones sobre especímenes que llegan a sus manos de manera casual. Pensemos, por ejemplo, en los sujetos que voluntariamente acceden a participar en un estudio que monitorea los efectos de un medicamento o en el investigador que anuncia en una clase que está haciendo un estudio sobre motivación en el universitario e invita a aquellos que acepten someterse a una prueba proyectiva TAT. En estos casos la elección de los individuos que serán sujetos a análisis depende de circunstancias fortuitas.

Este tipo de muestra se usa en estudios de laboratorio donde se procura que los sujetos sean homogéneos en variables tales como edad, sexo, inteligencia, de manera que los resultados o efectos no obedezcan a diferencias individuales, sino a las condiciones a las que fueron sometidos.

8.8.2 La muestra de expertos

En ciertos estudios es necesaria la opinión de sujetos expertos en un tema. Estas muestras son frecuentes en estudios cualitativos y exploratorios que para generar hipótesis más precisas o la materia prima del diseño de cuestionarios. Por ejemplo, en un estudio sobre el perfil de la mujer periodista en México (Barrera et al., 1989) se recurrió a una muestra de 11 = 227 mujeres periodistas pues se consideró que eran los sujetos idóneos para hablar de contratación, sueldos y desempeño de las mujeres periodistas. Estas muestras son válidas y útiles cuando los objetivos del estudio así lo requieren.

8.8.3 Los sujetos-tipo

También esta muestra se utiliza en estudios exploratorios y en investigaciones de tipo cualitativo, donde el objetivo es la riqueza, profundidad y calidad de la información, no la cantidad ni la estandarización. En estudios de perspectiva fenomenológica donde el objetivo es analizar los valores, ritos y significados de un determinado grupo social, el uso tanto de expertos como de sujetos-tipo es frecuente. Por ejemplo, pensamos en los trabajos de Howard Becker (El músico de jazz, Los muchachos de blanco) que se basan en grupos de típicos músicos de jazz y típicos



estudiantes de medicina para adentrarse en el análisis de los patrones de identificación y socialización de estas dos profesiones: la de músico y la de médico.

Los estudios motivacionales, los cuales se hacen para el análisis de las actitudes y conductas del consumidor, también utilizan muestras de sujeto-tipo.

Aquí se definen los grupos a los que va dirigido un determinado producto (por ejemplo Jóvenes clase socioeconómica A y R, amas de casa, clase B, ejecutivos clase A-B) y se construyen grupos de 8 o 10 personas, cuyos integrantes tengan las características sociales y demográficas de dicho sub grupo.

Con ese grupo se efectúa una sesión: un facilitador o moderador dirigirá una conversación donde los miembros del grupo expresen sus actitudes, valores, medios, expectativas, motivaciones hacia las características de determinado producto o servicio.