

---

# El muestreo

---

PID\_00270397

Albert Padró-Solanet i Grau

---

Tiempo mínimo de dedicación recomendado: 4 horas

---



**Albert Padró-Solanet i Grau**

Licenciado en Filosofía por la Universidad Autónoma de Barcelona (UAB). Másteres en Gestión pública y en Ciencia política (UAB). Profesor de los Estudios de Derecho y Ciencia Política de la Universitat Oberta de Catalunya (UOC).

El encargo y la creación de este recurso de aprendizaje UOC han sido coordinados por el profesor: Albert Padró-Solanet i Grau (2020)

Primera edición: febrero 2020  
© Albert Padró-Solanet i Grau  
Todos los derechos reservados  
© de esta edición, FUOC, 2020  
Av. Tibidabo, 39-43, 08035 Barcelona  
Realización editorial: FUOC

*Ninguna parte de esta publicación, incluido el diseño general y la cubierta, puede ser copiada, reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea este eléctrico, químico, mecánico, óptico, grabación, fotocopia, o cualquier otro, sin la previa autorización escrita de los titulares de los derechos.*

# Índice

<b>Introducción</b> .....	5
<b>Objetivos</b> .....	10
<b>1. El concepto de muestra y la utilidad del muestreo</b> .....	11
1.1. Conceptos de población y muestra representativa .....	14
<b>2. ¿Cómo funciona el muestreo probabilístico?</b> .....	20
2.1. ¿Cómo se conocen los parámetros poblacionales a partir de estadísticos muestrales con diferentes grados de confianza y de precisión? .....	21
2.2. ¿Cuál es la unidad de análisis de la distribución muestral? .....	23
2.3. ¿Qué queremos saber con las muestras? .....	26
2.4. ¿Cuál es la mecánica para una estimación de intervalo? .....	27
2.5. ¿Qué factores inciden en la deducción? .....	29
<b>3. El muestreo probabilístico</b> .....	32
3.1. El marco de muestreo .....	32
3.2. El muestreo simple aleatorio .....	33
3.3. El muestreo aleatorio sistemático .....	33
3.4. El muestreo estratificado .....	34
3.5. El muestreo estratificado no proporcional .....	36
3.6. El muestreo de conglomerados (agrupamientos o clústeres) .....	38
3.7. El muestreo multinivel (o de conglomerados con estratificación) .....	41
<b>4. Los muestreos no probabilísticos</b> .....	43
4.1. El muestreo intencional .....	43
4.2. El muestreo por cuotas .....	44
4.3. Las muestras de conveniencia .....	46
4.4. Las muestras de bola de nieve .....	47
<b>Bibliografía</b> .....	49



## Introducción

Los casos que elegimos determinan las conclusiones a las que llegamos: la importancia de la muestra y el muestreo.

El concepto de **muestra** en la ciencia empírica se refiere al tipo de información que tendrán que tener en cuenta los científicos para conocer la realidad. Las conclusiones a las que llegamos respecto a los fenómenos de la realidad dependen crucialmente de la información que aceptamos como válida y relevante para estudiar los fenómenos que nos interesan, por eso, la ciencia empírica se preocupa de determinar las reglas explícitas que identifiquen cuáles son los casos o ejemplos que se deberán tener en cuenta.

Definir la muestra es el esfuerzo de sistematizar y controlar cuáles son las cosas que se tienen que tener en cuenta si queremos tener un conocimiento válido de un fenómeno.

En nuestro conocimiento común, mucha de la información que tenemos sobre la realidad no la hemos experimentado personalmente. Pero la hay que sí. Por ejemplo, los sentidos nos permiten saber si en la habitación donde estamos estudiando hace frío, por el trato con los amigos sabemos si son simpáticos o depresivos. Pero existe toda otra parte de lo que conocemos a la que podemos llamar una «realidad acordada», en la que pensamos que las cosas son reales porque nos lo han dicho y todo el mundo parece estar de acuerdo con ello. El proceso de aprendizaje para formar parte de cualquier sociedad consiste precisamente en saber cuáles son estas cuestiones sobre las cuales hay un consenso de que son reales. Estas suposiciones son precisamente lo que caracteriza a una sociedad y son la base de muchos de los malentendidos y conflictos que aparecen entre las comunidades humanas.

### Francis Bacon y el método científico

Los científicos de la revolución científica del siglo XVI reconocieron la necesidad de una observación controlada y metódica para poder conocer sistemáticamente la realidad. El inventor del método científico (según Voltaire), Francis Bacon (1561-1626), centró su crítica de la ciencia aristotélica dominante en su época, en una idea de ciencia como conocimiento basado en la observación y la inducción. La visión inaugural de la ciencia de este filósofo nos ayuda a entender la importancia crucial del muestreo en las actuales ciencias sociales, tanto las cuantitativas como las cualitativas.

Según Bacon, el valor de las conclusiones basadas en la observación depende del carácter de esta observación. Para poder extraer conclusiones sobre la realidad, «inducir» las categorías o clasificaciones que reflejan mejor los fenómenos que queremos estudiar y saber cuáles son las relaciones que hay entre estas categorías, hemos de tener información correcta de esta realidad. Por eso, la fase crucial sobre la cual se puede construir un auténtico conocimiento de la realidad se tiene que basar en la recogida sistemática y controlada de los hechos que se deberán tener en cuenta. Definir o diseñar cómo se tiene que producir esta observación para que no afecte indebidamente a las conclusiones, las sesgue, es un momento clave de toda investigación.

Para Bacon, los científicos tienen que utilizar la observación sistemática para superar los prejuicios y las actitudes preconcebidas que actúan como barreras o filtros del conocimiento de la realidad. Bacon clasificaba los prejuicios en cuatro categorías (Copleston, 1981, p. 286-291):

#### Lectura recomendada

F. Copleston (1981). *Historia de la filosofía* (vol. 3). Barcelona: Ariel.

1) Los **ídolos de la tribu** (*idola tribus*) son los errores inherentes a la naturaleza humana. Por un lado, basarse excesivamente en la experiencia de los sentidos, sin un espíritu crítico. Por ejemplo, creer que lo que en realidad está en movimiento es fijo. Después está la tendencia a confiar en las ideas que hemos recibido y creído o que nos son agradables, y, en cambio, pasar por alto o rechazar los casos contrarios a las creencias recibidas (la psicología cognitiva actual ha confirmado empíricamente estas tendencias). Finalmente, está la idea de pensar en la realidad teleológicamente, tratando antropomórficamente las relaciones causales de la realidad, como si hubiera una finalidad en los fenómenos (la **teoría del diseño inteligente**, que comporta que la realidad natural es producto de una conciencia absoluta que lo ha diseñado, es un ejemplo de pensamiento teleológico y se contrapone a la **teoría de la evolución**, en la que el mecanismo real es la adaptación a un medio cambiante por medio de la diversificación. Ambas teorías explican la complejidad creciente de los sistemas biológicos, incluido el humano).

2) Los **ídolos de la caverna** (*idola specus*) son los errores producto de las características de cada individuo, su carácter o temperamento, educación e influencias que ha recibido.

3) Los **ídolos de la plaza del mercado** (*idola fori*) son los errores que provienen del lenguaje. A veces las palabras no son lo bastante precisas para describir o ser aplicadas a fenómenos concretos, a veces no tienen significados compartidos por toda la comunidad, y se convierten en una barrera más que en una herramienta para entender la realidad.

4) Los **ídolos del teatro** (*idola theatri*) son los sistemas filosóficos del pasado que presentan mundos irreales que han creado los hombres. Bacon utiliza esta categoría para criticar el aristotelismo que daba primacía a la lógica, pero también los estudios empíricos que se basaban en pocas observaciones y oscuras.

Este catálogo de errores es antiguo pero completamente vigente. Tendríamos que poder reconocer el origen de muchas de las falacias y sesgos que sufrimos cotidianamente en la comprensión de nuestra realidad. Pero normalmente los sufrimos de forma inconsciente. Los estudiosos de la psicología cognitiva han estudiado cómo funcionan estos sesgos y han intentado divulgar cómo funcionan los mecanismos para hacer las sociedades menos sujetas a la manipulación política (Nisbett y Ross, 1980).

Uno de los tipos de error de inducción más extendido, inconsciente y pernicioso, es el que Nisbett y Ross bautizaron como **error fundamental de atribución**. Consiste en asociar esencialmente las minorías sociales con comportamientos minoritarios o extraordinarios. Por ejemplo, la tendencia es asociar un comportamiento extraordinario, «robar», con formar parte de un grupo social minoritario, «ser gitano» (podríamos multiplicar los ejemplos al infinito: por ejemplo, «árabes/terroristas»). Estos tipos de atribuciones son la base de los estereotipos sociales y no son exclusivos de personas sin formación. De hecho, muchas de las regulaciones de las cuotas de inmigración en los Estados Unidos de América según el país de origen se basaban en ideas racistas sostenidas por científicos sociales. Las asociaciones son tan «convincientes» que los científicos pueden intentar encontrar explicaciones sofisticadas que las expliquen.

Aunque nos sorprenda, a comienzos del siglo XX, en los Estados Unidos de América los equipos más importantes del baloncesto estaban formados por muchos jugadores judíos que también eran las estrellas más brillantes. Para algunos periodistas y analistas, el hecho de que los judíos fueran históricamente perseguidos los obligó a desarrollar unas habilidades de coordinación y una confianza mutua que les daba una ventaja decisiva a la hora de jugar en equipo. Esta atribución nos resulta sorprendente o, incluso, ridícula porque vivimos en una época en la que los principales jugadores de baloncesto del mundo son negros. Y la tendencia general es la de encontrar la asociación entre el grupo social y el éxito deportivo en algún tipo de argumento biológico: la selección de los individuos físicamente más dotados en el sistema del esclavismo ha dado una ventaja insuperable a los individuos negros. La explicación podría tener alguna verosimilitud, porque las diferencias en las proporciones de ciertas características físicas son la base de la agrupación en «razas», pero nos hace pasar por alto factores que pueden estar interviniendo y que podrían explicar mejor las características físicas en el dominio de los jugadores negros en el baloncesto actual. Por ejemplo, la «población negra» americana está discriminada y no tiene las mismas oportunidades laborales y profesionales que los miembros de los grupos dominantes «blancos», esto puede explicar por qué los jóvenes de color están apostando desproporcionadamente por carreras deportivas. La apuesta por una carrera deportiva es arriesgada, puesto que la vida deportiva depende en gran medida del azar: estado de forma en el momento de las competiciones, existencia de lesiones, etc. No es lo mismo que apostar por invertir en una educación, que es uno de los predictores más claros de éxito laboral según todos los estudios. Por lo tanto, a igualdad de cualidades físicas, es más probable que un negro (sin tantos recursos para pagarse una educación y después discriminado en las oportunidades laborales) se arriesgue a una carrera deportiva que lo haga un «blanco». Este factor es seguramente más explicativo del dominio de jugadores negros que una diferencia biológica «racial» –y seguramente es lo que también explicaba el dominio de jugadores judíos a comienzos del siglo pasado, cuando grandes masas de población judía habían acabado de inmigrar a los Estados Unidos de América.

Para Nisbett y Ross (1980), que seguían el camino iniciado por los que terminaron siendo premios Nobel de Economía, Twerski y Kahnemann (Kahnemann, Slovic y Tversky, 1982), una parte importante de la educación cívica en una democracia tendría que tratar los errores de inducción en los que tendemos a caer los humanos y tener una mínima alfabetización estadística para saber cómo observar y medir la realidad, y así, evitar los sesgos de inducción: el pensamiento estereotipado y caer en manipulaciones políticas. Desgraciada y previsiblemente, los especialistas en manipulación de la comunicación (el marketing comercial y político) son los que han aprendido a fondo estos estudios sobre los sesgos de la inducción natural y sus implicaciones.

En resumen, los conceptos de **muestra** y de **muestreo** hacen referencia a los procedimientos que se han desarrollado para las ciencias con el fin de seleccionar los casos que nos permitirán estudiar la realidad de forma válida.

Dado que el objetivo es lograr una representación correcta y válida de la realidad, tanto la metodología cuantitativa como la cualitativa convergen en el punto de controlar qué datos se usan en la inducción científica.

### Muestreo y metodología cualitativa

La **metodología cuantitativa** se preocupa de conseguir un conjunto de datos en los que los fenómenos que se quieren estudiar puedan ser medidos como variables en las unidades de análisis (los casos). Podríamos decir que la metodología cuantitativa sigue una estrategia extensiva: cuanto más unidades de análisis haya para medir las variables, mejor.

En cambio, la **metodología cualitativa** se centra en fuentes de información ricas y complejas, en las que no es fácil hacer una estandarización: se quieren analizar muchas características –que no tenemos muy claro cómo se miden– simultáneamente en un único caso, en un proceso, en un texto o en las explicaciones que han dado en una entrevista un pequeño número de actores implicados en un fenómeno; por ejemplo, en los casos de acoso escolar (*bullying*) o en los abusos sexuales en centros educativos. Por lo tanto, la estrategia cualitativa es intensiva (idiográfica). Pero si tiene que mostrar la relevancia del conocimiento obtenido en un caso en otros casos más generales, será más convincente y claro conocer en qué medida podemos pensar que el conocimiento obtenido depende de los casos o ejemplos elegidos. Para hacerlo, el énfasis no está en los sistemas de construcción de la muestra (como los que detallaremos más adelante en la investigación cuantitativa), sino que se busca que las interpretaciones obtenidas sean válidas en las situaciones reales y cotidianas en las que se obtuvieron los datos utilizando la comparación y el control con otras situaciones y casos. Los casos son elegidos porque se supone que son representativos (aunque no estadísticamente) de una población.

### Sesgos cognitivos

Los efectos de los sesgos cognitivos que estudia la psicología social y cognitiva es el contenido de los episodios de la serie de divulgación científica de National Geographic *Brain Games*. Disponible en: [nationalgeographic.org/education/channel/brain-games/](http://nationalgeographic.org/education/channel/brain-games/)

### Lectura recomendada

H. E. Brady y D. Collier (2010). *Rethinking social inquiry: Diverse tools, shared standards*. Lanham: Rowman & Littlefield.

Nuestro análisis detallado de sus características y la comparación con otros casos sirven de garantía a esta representatividad asumida. Siguiendo esta lógica, las investigaciones de casos múltiples controlan la variabilidad en aspectos clave entre los casos para garantizar que se recoge el conjunto de características de los individuos o de las situaciones planteadas dentro del sistema que se quiere estudiar. Por otro lado, en el enfoque metodológico de la **teoría fundamentada** (Glaser y Strauss, 1967) –que pretende la recuperación inductiva de las interpretaciones y de las prácticas sociales en las comunidades– hay una acentuada preocupación por la muestra. Los conceptos de muestreo teórico y gradual sirven para decidir en cada momento de la investigación qué tipo de información se tendría que tener en cuenta en cada fase sucesiva, de acuerdo con los objetivos teóricos.

## Objetivos

Con el estudio del presente módulo se tendrán que alcanzar los objetivos siguientes:

- 1.** Entender el concepto de muestreo y el objetivo de la generalización de una muestra.
- 2.** Comprender las restricciones y condiciones de la muestra para la generalización: la representatividad.
- 3.** Conocer los tipos de muestreo: probabilísticos y no probabilísticos. Ventajas e inconvenientes.

## 1. El concepto de muestra y la utilidad del muestreo

Podemos entender el problema de la muestra y el muestreo estadístico como una respuesta a los problemas inevitables que como humanos limitados en el espacio, en el tiempo y en los recursos, no podemos esperar a disponer del conocimiento directo y rico del conjunto de objetos –o de todos los hechos de un tipo o de una clase– para hacer inferencias para el conjunto de objetos o de hechos de una misma categoría o clase. Por eso nos basamos en un conjunto inevitablemente más pequeño de elementos, que hemos experimentado o estudiado en detalle, para poder extrapolarlo al conjunto de objetos o hechos – como hemos visto en la introducción –, esta limitación –que nos lleva a confiar en una experiencia limitada y en las experiencias de los otros condensadas en descripciones y explicaciones de la realidad –, es una fuente de error y de manipulación. Hay que tener en cuenta que, incluso, los estudios de datos masivos<sup>1</sup> –que precisamente se definen por la gran cantidad de datos que pueden ser tratados gracias al crecimiento exponencial de la capacidad de computación– tienen que seleccionar una parte de esta información y, si esta selección está sesgada en algún sentido, puede afectar a la veracidad de las conclusiones que se extraigan de los datos.

<sup>(1)</sup>En inglés, *big data*.

El problema de la muestra es, pues, el de definir cuáles son las guías que los investigadores tendrían que seguir a la hora de definir este conjunto más reducido de casos para evitar el sesgo y el error en nuestras extrapolaciones. Una de las críticas más duras a la credibilidad de una investigación es que es la muestra la que viola estas guías.

Seguramente, el ejemplo más utilizado para explicar y justificar el concepto de muestra es el del cocinero, que para saber si la sopa que está preparando tiene suficiente sal, prueba una cucharada (una muestra). El cocinero quiere saber si la sopa en conjunto está suficientemente salada, el total de la sopa de la olla (es decir, la población o el universo). Pero para saberlo, no se la toma toda. ¡Sería absurdo!

Las soluciones de la estadística respecto de los problemas de la muestra y el muestreo requieren entender el punto de vista característico de la estadística. En estadística, la **evidencia empírica** consiste en una serie de observaciones en las que se pueden medir los fenómenos o las características que nos interesan. Esta serie de observaciones se puede referir a una **realidad física común** que permite que todas estas observaciones puedan ser comparadas y agregadas (Chatterjee, 2003, p. 17).

Este punto de vista de los estudios cuantitativos es, en parte, intuitivo y ya conocido. Por ejemplo, a nadie le extraña que se mida la capacidad de una clase con una nota individual a cada estudiante y después se calcule la media de todas las calificaciones. Pero debemos ser conscientes de la importancia de

las operaciones que se están llevando a cabo para utilizar mejor la estrategia – sacarle más jugo y evitar errores. Del mismo modo que estudiar la gramática nos ayuda a hablar y escribir mejor.

Por ejemplo, podemos tirar 100 veces una moneda y ver cuántas veces sale cara. En este caso, la realidad física común es la moneda que conecta la serie de observaciones. Pero podríamos tirar un puñado de 100 monedas y ver cuántas son cara. En este caso, la realidad común son las 100 monedas. Otro ejemplo, elegimos 50 viviendas de un municipio de 800 viviendas y medimos su tamaño. En este caso, la realidad física relevante es el conjunto de viviendas de un municipio. Aún otro ejemplo más, cogemos una muestra de pacientes con hipertensión de todos los pacientes que se han tratado durante un año en una clínica y medimos sus presiones sistólica y diastólica antes y después de haber tomado una droga (de cada observación tomamos cuatro datos). En este caso, la realidad física relevante son los pacientes de hipertensión y que han tomado la droga.

A partir de estos ejemplos se puede entender cómo los datos estadísticos básicos referidos a las observaciones (casos o individuos) y caracteres o variables pueden ser registrados en una tabla donde en las filas estén las observaciones y en las columnas estén los atributos observados en cada caso (las características o las variables). Normalmente, esta tabla de registro de datos, con observaciones y variables, se denomina *matriz de datos*.

Tabla 1. Matriz de datos

<b>Observaciones / Variables</b>	<b>Variable 1</b>	<b>Variable 2</b>	<b>...</b>	<b>Variable M</b>
<b>Observación 1</b>				
<b>Observación 2</b>				
<b>...</b>				
<b>Observación N</b>				

La estrategia del estudio cuantitativo o estadístico se fundamenta en una simplificación radical de la realidad. Implica que las unidades de análisis –o casos– son idénticas entre sí y solo se diferencian de forma relevante por las características o variables que nos interesa estudiar. Esta estrategia de fijarse en las variables que quieren ser medidas y de simplificación de las unidades en las que son medidas proviene de las ciencias naturales. En física, la masa o la velocidad puede ser medida en cualquier unidad –no importa si es una bola metálica o un elefante– y en cualquier unidad se puede calcular su momento cinético. Se asume que las diferencias que hay en los casos –y que pueden influir en los valores de las variables cuando se miden– quedan incorporadas en el denominado *error de medida*. El requisito de los análisis cuantitativos de tener en cuenta muchos casos sirve para que estos errores o las diferencias individuales se anulen entre sí para que, finalmente, podamos tener una medida aceptable de las variables y de las relaciones que hay entre ellas.

Esta simplificación radical de la realidad será más o menos verosímil según qué fenómeno se quiera estudiar. Pero siguiendo las reglas del método científico, se prefieren las explicaciones más simples a las más complejas, y solo se in-

corporará complejidad cuando sea necesario (**principio de parsimonia**). Por ejemplo, el orden como se obtienen las observaciones no es importante, pero en algunos estudios en los que las condiciones externas cambiantes pueden influir en los valores observados de nuestras variables, o el hecho de repetir la medida sobre el mismo individuo, habrá que tener en cuenta el orden temporal en las observaciones. Cuando tiramos 100 veces una moneda, no hay que tener en cuenta el orden, porque no cambia nada. Pero cuando medimos la inflación o el miedo al delito, con mediciones repetidas a unos mismos individuos, hay que tener en cuenta el orden temporal.

El método cuantitativo o estadístico es un enfoque que indica cómo se tendría que utilizar el método científico para responder a preguntas de investigación y tomar decisiones. Este método no solamente es un procedimiento para analizar datos, consiste en reglas e instrumentos que se han ido desarrollando con objeto de ayudar a conocer la realidad en diferentes fases:

- 1) ¿Cómo se tienen que diseñar las investigaciones para que sirvan para responder las preguntas planteadas?
- 2) ¿Cómo se tienen que recoger buenos datos de la realidad?
- 3) ¿Cómo pueden ser descritos los datos con números y gráficos?
- 4) ¿Cómo pueden ser analizados estos datos?
- 5) ¿Cómo se pueden extraer conclusiones a partir de estos análisis?

Una vez que se ha seleccionado el diseño de la investigación (que puede ser observacional o experimental), se tienen que elegir los sujetos (las unidades de análisis, los casos o los individuos) de los cuales se recogerán los datos. Esta fase de selección de los individuos es crucial para producir buenos datos. Un dicho famoso entre los estadísticos es «*garbage in, garbage out*» (a menudo denotado con el acrónimo GIGO): ‘si en el análisis entra basura, solo puede salir basura. Un defecto básico en los datos que entran puede provenir del hecho de que el grupo de datos que hemos obtenido esté sesgado –es decir, que haya individuos o grupos de individuos que tienen más probabilidad de formar parte de la muestra. De hecho, como veremos más adelante, para minimizar el sesgo en las muestras, tenemos que elegir a los individuos de la muestra de forma aleatoria. Con una muestra sesgada, es fácil que los resultados obtenidos con su análisis también estén sesgados.

Por ejemplo, en todas las encuestas electorales sabemos que las muestras –por mucho que queramos que sean aleatorias y representativas del electorado– siempre contienen sesgos, porque no todos los votantes tienen las mismas ganas de responder a las preguntas sobre qué votarán o qué han votado, y estas ganas o no de responder normalmente están relacionadas con las preferencias partidistas de los votantes. Naturalmente, los encuestadores conocen la existencia de estos sesgos y tienen que corregir los pesos de los diferentes votantes en la muestra para poder hacer estimaciones correctas del resultado en la población. Pero esto significa hacer suposiciones sobre cómo se han comportado los diferentes grupos y explica por qué, a pesar de la sofisticación de las técnicas empleadas, no es extraño que las encuestas electorales puedan equivocarse estrepitosamente.

Desde este punto de vista, se entiende que nuestra experiencia como individuos nos proporciona inevitablemente muestras sesgadas. Los psicólogos cognitivos han comprobado que naturalmente la mayor parte de la población

tiende a pensar que ellos trabajan más que la media de la gente. Este error es comprensible si tenemos en cuenta que nuestra muestra, respecto a nuestro trabajo y al trabajo de los demás, está sesgado: cada cual conoce perfectamente todas las veces que trabaja, pero solo esporádicamente sabe cuándo trabajan los otros. La existencia de errores de este tipo y la necesidad de controlarlos justifica las estrategias metodológicas cuantitativas o estadísticas en la selección de los casos.

### 1.1. Conceptos de población y muestra representativa

En estadística, una **población** se define como el conjunto de unidades de análisis (casos, individuos) de una misma naturaleza sobre la cual queremos conocer un tipo de información cuantificable.

El concepto de población en ciencia se distingue del normal. La población no solamente es un conjunto de personas sino de cualquier cosa que deseamos estudiar: los delitos en una ciudad, las sentencias de los tribunales o los miembros de las fuerzas de seguridad de un país pueden ser tratados como una población. Así, la población (o universo) es el conjunto de casos relevantes para un determinado tipo de problema estadístico.

Una **muestra** es un subconjunto extraído de la población que posibilita llevar a cabo análisis, cuyas conclusiones son aplicables al conjunto de la población.

Es decir, la muestra nos permite generalizar a la población, la muestra representa a la población. Por ejemplo, si queremos saber cómo responderán los habitantes de un barrio a la instalación de un centro para menores tutelados, podemos preguntar a una muestra de vecinos y suponer que las respuestas representan a la opinión del barrio. Si queremos estudiar los casos que ha resuelto un tribunal de justicia, seguramente no podremos estudiarlos a todos y tendremos que seleccionar a una muestra que represente todos los casos de este tribunal.

Si la población es suficientemente pequeña, se pueden hacer análisis estadísticos descriptivos sin necesidad de seleccionar muestras. La necesidad de construir muestras aparece cuando la población es tan grande que no puede ser efectivamente tratada en su totalidad.

Por ejemplo, en muchos casos de estudios retrospectivos del efecto de las políticas públicas o programas sociales, no es posible estudiar a toda la población que ha sido afectada por la política. Hay otros estudios prospectivos que quieren estudiar el efecto potencial de un programa o política pública, que, por definición, tampoco permite estudiar al conjunto de la población. Los estudios demográficos, como por ejemplo los censos de población, como indica su nombre, tienen el propósito de estudiar al conjunto de la población. Pero estos estudios son extremadamente costosos. Tradicionalmente, implicaban a una gran cantidad de personal especializado durante años. Por eso en países con poblaciones

muy grandes nunca se llegaron a hacer censos de población para planificar las políticas públicas, sino que se desarrollaron sistemas muestrales que permitieran realizar el mismo trabajo de manera suficientemente cuidadosa. Actualmente, las nuevas tecnologías han simplificado el trabajo de tener un registro poblacional con datos suficientemente ricos, pero estudios cruciales sobre los cuales se basan decisiones políticas importantes, como por ejemplo, la evolución del paro o los precios en los países desarrollados, se llevan a cabo mediante estudios de encuestas con muestras (en España, el Instituto Nacional de Estadística hace la Encuesta de población activa y la Encuesta de presupuestos familiares).

Las muestras son **eficientes** porque no son tan costosas como los estudios de poblaciones grandes. Al mismo tiempo que son menos costosas, estas muestras son más eficaces porque permiten conocer mejor los casos que son elegidos, ya que hay más recursos (dinero, personal, tiempo) para estudiarlos. Todavía más, los datos obtenidos podrán ser más detallados. Con más tiempo para medir, recoger, procesar y analizar los datos, habrá menos errores en el proceso y más oportunidades para controlar que se ha realizado correctamente.

La principal característica que se desea de la muestra es que sea **representativa** de la población respecto de un atributo relevante para la investigación que se quiere llevar a cabo. La distribución de esta característica en la muestra tiene que ser parecida a la distribución del atributo de la población. La idea es aparentemente simple. Pero siempre aparece alguna dificultad para resolver esta cuestión en la práctica de la investigación. Precisamente, los procedimientos de muestreo definen estrategias para obtener, de forma eficiente y detallada, subconjuntos de las poblaciones que representen las características relevantes de la población. Las características relevantes de la población dependen del tipo de estudio que queremos hacer. Una buena muestra lo es en relación con la pregunta que queremos responder, y que depende de la teoría o explicación de la realidad que queremos comprobar.

El **muestreo probabilístico** es un tipo especial de muestreo que nos permite generalizar estadísticamente poblaciones más grandes. Se define como el método de selección en el que cada miembro de una población tiene una probabilidad positiva conocida de ser elegido.

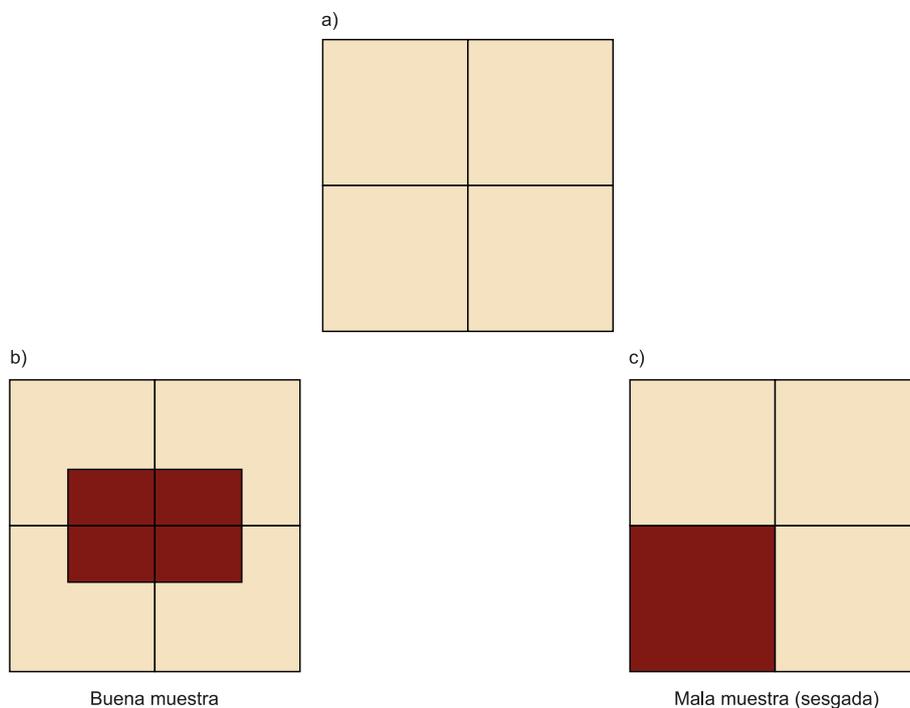
Si sabemos la probabilidad de que cualquier miembro de la muestra sea elegido, podemos garantizar que nuestra muestra representa a la población de la que es extraída y podemos hacer estimaciones precisas relativas a la población. Si los miembros de nuestra población fueran idénticos, o muy parecidos en todas las características que queremos estudiar, no nos haría falta ningún procedimiento especial de muestreo: sería suficiente con elegir un individuo aleatoriamente, por suerte, y estudiándolo para poder extrapolar sus características en toda la población. Pero sabemos que las poblaciones reales son heterogéneas y pueden variar en muchos aspectos importantes para nuestra investigación. Si en la muestra que hemos elegido hay individuos con unas ciertas características que tienen más probabilidad que otras de formar parte de

#### China e India

Por ejemplo, en países con poblaciones muy grandes, como China o la India, no se realizan censos de población tal como se realizaban en países desarrollados con poblaciones más reducidas, sino que se trabaja con muestras para poder planificar las políticas públicas.

ella, o algunas con diferentes características que no tienen ninguna probabilidad de estar presentes, entonces tendremos muestras sesgadas, es decir, no representativas, que no nos permitirán conocer correctamente a la población.

Figura 1. Esquema de población dividida en 4 características relevantes



Fuente: elaboración propia.

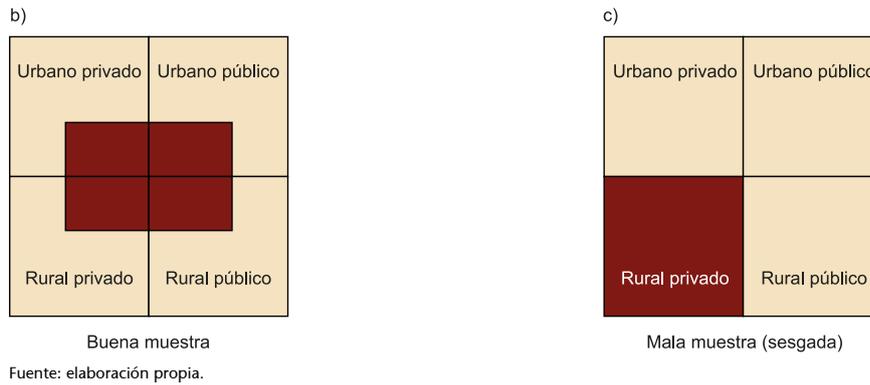
### Figura 1

En la figura a) aparece representada una población que tiene cuatro características relevantes para la investigación, que se representan con cuatro cuadrantes. La muestra (representada por el cuadrado oscuro) de la figura b) es representativa, no está sesgada, porque permite la entrada proporcional de unidades de análisis (o casos) de cada uno de los cuatro cuadrantes. En la figura c), en cambio, la muestra es sesgada porque solo contiene casos del tercer cuadrante.

## Investigación que quiere estudiar el acoso en los centros escolares de secundaria de Cataluña

Se quiere conocer qué tipo de políticas de prevención, detección y tratamiento se están utilizando y cómo reaccionan los diferentes estamentos: estudiantes, profesores, dirección, padres. La **población** que se quiere estudiar en este nivel es el conjunto de centros escolares de secundaria de Cataluña y, por lo tanto, las **unidades de análisis** son los centros escolares. No es viable desplazar a un equipo de investigación a todos los centros para pedir la información necesaria y hacer las observaciones pertinentes, por lo cual, se tiene que elegir una muestra de las escuelas que se van a observar según los recursos que tengamos. Suponemos que será suficiente con diez centros. Si esperamos que todos los centros escolares tengan las mismas políticas y modo de tratar el maltrato, podríamos listar todos los centros catalanes y elegir los diez primeros, y podríamos argumentar que es una muestra válida de la población de centros. Pero suponemos que en nuestra explicación del acoso sospechamos (basándonos en la experiencia, en la investigación previa o en nuestra intuición) que las escuelas o institutos de las zonas rurales y las de ámbitos metropolitanos difieren en la forma y prevalencia del acoso, y que las escuelas o institutos de titularidad pública difieren de las privadas y concertadas, por lo tanto, no correremos el riesgo de elegir solo escuelas públicas o solo metropolitanas para nuestra muestra. Así pues, tendríamos cuatro subpoblaciones de centros escolares que suponemos relevantes para el estudio del acoso escolar: escuelas rurales privadas, escuelas rurales públicas, escuelas metropolitanas privadas y escuelas metropolitanas públicas. Si usamos el esquema de la figura 1, mientras que la muestra de la figura b) recoge casos de las cuatro subpoblaciones, la muestra de la figura c) solo recoge casos de una de estas y, por lo tanto, estará sesgada.

Figura 2. Centros escolares y acoso



Ejemplos de muestras no representativas o sesgadas:

1) Actualmente, en la mayor parte de las encuestas preelectorales, las muestras se suelen elegir por medio de una selección aleatoria de números de teléfono fijos en las diferentes circunscripciones electorales. El método es conveniente porque con poco coste se puede acceder a una gran cantidad de población de electores con un coste razonable. Pero dado que cada vez una parte más grande de la población solo tiene teléfonos móviles, la muestra no podrá contener un tipo de población: los más jóvenes y los que tienen menos estabilidad en su vivienda (alquilan o están temporalmente en una zona) no podrán formar parte de la muestra. Si el comportamiento electoral está relacionado con las características de esta parte de la población de electores, si los más jóvenes o más dinámicos económicamente tienen unas preferencias electorales diferentes del conjunto de la población, la muestra estará sesgada y, en el análisis, este defecto se tendrá que tener en cuenta y aplicar algún tipo de corrección.

2) Los cambios tecnológicos tienen implicaciones en los sesgos que tenemos en las muestras de las encuestas preelectorales. Ahora mismo, el crecimiento en la utilización de internet y especialmente de las redes sociales por sectores más amplios de la población hace posible otras formas de encuestar, pero en este caso el sesgo de la muestra en comparación con la población objetivo todavía puede ser mayor.

#### Reflexión

¿Podrías pensar en qué tipos de diferencias imagináis que pueden afectar a la elaboración de encuestas preelectorales por medio de internet?

Un caso famoso en la historia del muestreo para encuestas preelectorales es el desastre que sufrió una revista americana, *The Literary Digest*, en los años treinta del siglo pasado. Desde el año veinte, durante las campañas electorales presidenciales, la revista hacía una encuesta por correo a una enorme muestra de más de un millón de personas preguntando a qué candidato votarían. Durante años, la encuesta fue muy precisa en sus predicciones, hasta que en las elecciones de 1936 erró estrepitosamente, dando por perdedor al presidente saliente, Franklin D. Roosevelt, con el 40 % de los votos, cuando en realidad ganó las elecciones con un 60 % de los votos. ¿Cuál fue la causa de este error? Los miembros de la muestra fueron elegidos a partir de dos fuentes: los listados telefónicos y los de matriculación de coches. En una sociedad afectada por la crisis económica del 29, los propietarios de casas con teléfono y de coches no eran exactamente representativos de la población general. La muestra sobrerrepresentaba los segmentos más ricos de la población, que tenían preferencias partidistas más cercanas a los republicanos que a los demócratas. Esta fue la causa del error en la predicción electoral (Manheim y Rich, 1988).

3) Supongamos que queremos hacer una investigación sobre el miedo al delito en un barrio conflictivo de la ciudad. Para un investigador no experimentado puede parecer que la forma más directa y válida de responder a esta pregunta

consiste en preguntar a la gente que se pasea por las calles del barrio –tal como hacen los políticos cuando quieren tomar el pulso de una población. Pero este tipo de muestreo inevitablemente incorporará un sesgo a la muestra. La gente que está en la calle a una hora concreta del día no tiene por qué parecerse al conjunto de la población por muchas razones: dependiendo de la hora que elegimos, puede haber una parte de la población que esté trabajando y que no pueda aparecer en la proporción correcta en la muestra. Hay una parte de la población con movilidad reducida que tampoco encontraremos en la calle. Habrá gente que no se dejará entrevistar y, si nuestra pregunta –el miedo al delito– está relacionada con esta actitud (si los que tienden a tener miedo al delito también tienden a evitar ser entrevistados en la calle), entonces nuestra muestra tenderá a tener encuestados con menos temor al delito que los que hay en la población en general. Además, la parte de la población que suele estar en la calle –en igualdad de condiciones: salud, edad, condición laboral, etc.– tenderá a ser menos miedosa que la que lo evita, introduciendo inevitablemente otro sesgo en nuestra muestra.

4) En otra investigación, el Ayuntamiento de un municipio quiere conocer el número de viviendas que han sufrido intentos de robo o daños por gamberismo. Estos hechos no son denunciados si no acaban con robo o si los delincuentes no son detenidos mientras están cometiendo el delito. No es viable hacer una investigación que tenga en cuenta todas las viviendas de la población porque es demasiado grande. Por eso, se tiene que seleccionar una muestra. Un sistema directo sería elegir todas las viviendas que hay en la vía principal que atraviesa la población y preguntar a sus propietarios. Pero si, tal como es normal, las calles más céntricas tienden a estar más concurridas durante más horas del día, no tenemos que pensar que sean representativas del número de delitos cometidos en el conjunto del municipio.

La **representatividad** de las muestras consiste en el hecho de que, de forma agregada, las características de la muestra se aproximan a las características de la población. No hace falta que las muestras sean representativas en todos los aspectos, solo respecto de las características de la población que queremos estudiar.

El muestreo probabilístico se basa en la idea de que todas las unidades de la población tienen la misma probabilidad de ser escogidas, pero incluso las muestras probabilísticas más detalladas difícilmente logran una representatividad perfecta de las poblaciones. De todas maneras, son las muestras preferidas para la investigación, porque evitan los sesgos de selección que pueden afectar a cualquier otra forma de elegir las muestras. Permiten generalizar desde pocas observaciones a toda la población teniendo un control del error que se comete.

#### **Técnicas de muestreo no probabilísticas**

Ciertamente, los criminólogos no pueden utilizar muestras probabilísticas para responder a muchas de las preguntas que se formulan, puesto que establecer cuáles son las pobla-

ciones de interés a menudo es obviamente difícil y problemático (por ejemplo, ¿cuál es la población de delincuentes?, ¿cuál es la población de los delitos?). No obstante, por esta misma razón, se han desarrollado **técnicas de muestreo no probabilísticas** que pueden ser útiles para la investigación en criminología a la hora de reducir los sesgos potenciales (podéis ver el apartado «Los muestreos no probabilísticos» de este módulo).

## 2. ¿Cómo funciona el muestreo probabilístico?

En este apartado se introducen los fundamentos de la estadística inferencial.

La **teoría de la probabilidad** es el instrumento que permite inferir cómo se distribuyen en la población los datos obtenidos en las muestras.

La **estadística descriptiva** propone diferentes cálculos que nos permiten identificar de forma resumida las características (variables) de la población; son instrumentos útiles que sirven para sintetizar o «describir» estas características en la población.

Por ejemplo, la renta media (variable) de los vecinos empadronados en un municipio (población) nos dice cuál es la renta más característica de esta población. La **media** es un estadístico de tendencia central, del mismo modo que la **moda** o la **mediana**, y sirve para identificar cuáles son los valores más representativos de una distribución, qué valores tienden a tener los casos de una población. Si nos encontráramos con un vecino de este municipio y nos preguntaran cuál es su renta, sin saber nada más, la mejor respuesta, con la que nos equivocaríamos menos, sería la media. Hay otros estadísticos que sirven para medir hasta qué punto los estadísticos de tendencia central se equivocan. La renta media puede ser la misma en dos municipios, pero las dos distribuciones pueden ser muy diferentes: mientras que en un municipio hay gente más rica y más pobre, en el otro todo el mundo está más cercano a la media. Los **estadísticos de dispersión** miden este tipo de diferencias. ¿Hasta qué punto la media (o el estadístico de tendencia central) es buena para describir al conjunto de la población? Cuanto mayor sea la dispersión, menor será la bondad del estadístico de tendencia central. En el municipio con poca dispersión de renta, con poca **desviación típica o estándar**, apostando por la renta media, tendría más probabilidad de ganar que en el municipio con más dispersión o con más desviación.

En la población, los estadísticos de tendencia central o de dispersión (en realidad, cualquier otro tipo de estadístico poblacional) se llaman **parámetros**.

La **estadística inferencial** consiste en estimar los parámetros poblacionales a partir de los estadísticos obtenidos en las muestras. Por ejemplo, a partir de la renta media estimada de una muestra, queremos saber cuál es la renta media de toda una población. O, a partir de la edad media de una muestra de individuos, queremos saber la media de toda la población.

Tabla 2

	<b>Población (no observada) Parámetro</b>	<b>Muestra (observada) Estadístico</b>
<b>Media</b>	$\mu$	$\bar{x}$
<b>Desviación típica</b>	$\sigma$	$s$
<b>Pendiente</b>	$\beta$	$b$
<b>Proporción</b>	$\Pi$	$p$

**Estadísticos poblacionales**

En la tabla 2 vemos cómo convencionalmente, en los estadísticos poblacionales, los parámetros se simbolizan con letras griegas, mientras que los estadísticos muestrales se simbolizan con letras latinas.

El objetivo del muestreo es elegir un conjunto de elementos de una población de forma que la descripción de estos elementos (los **estadísticos muestrales**) haga posible averiguar cuáles son los **parámetros poblacionales**. El **muestreo probabilístico** permite conseguir este objetivo y nos mide el grado de acierto a la hora de hacer estas estimaciones.

La clave del muestreo probabilístico es la selección aleatoria. Por un lado, la selección aleatoria, o por azar, evita que haya casos que acaben formando parte de la muestra por algún tipo de motivo no relacionado con formar parte de la población. Elegir aleatoriamente impone una barrera a la discrecionalidad y a la manipulación más o menos inconsciente en la selección de los casos. Fuerza al investigador a «atarse de manos» y permitir que cualquier caso que sea miembro de la población, tal como lo hemos descrito, tenga la posibilidad de formar parte de la muestra. Si los investigadores eligen los casos de manera intuitiva, pueden caer en la trampa de elegir los que tienden a confirmar sus teorías (tal como sabemos que pasa con el **sesgo de confirmación** estudiado por los psicólogos cognitivos). Por otro lado, la selección aleatoria permite utilizar la teoría de la probabilidad para estimar los parámetros poblacionales a partir de los estadísticos muestrales y saber hasta qué punto estas estimaciones son correctas.

### **2.1. ¿Cómo se conocen los parámetros poblacionales a partir de estadísticos muestrales con diferentes grados de confianza y de precisión?**

La respuesta a cómo es posible conocer los parámetros poblacionales a partir de los estadísticos muestrales es que lo podemos hacer gracias a la **distribución muestral**, que hace de puente entre la muestra y la población.

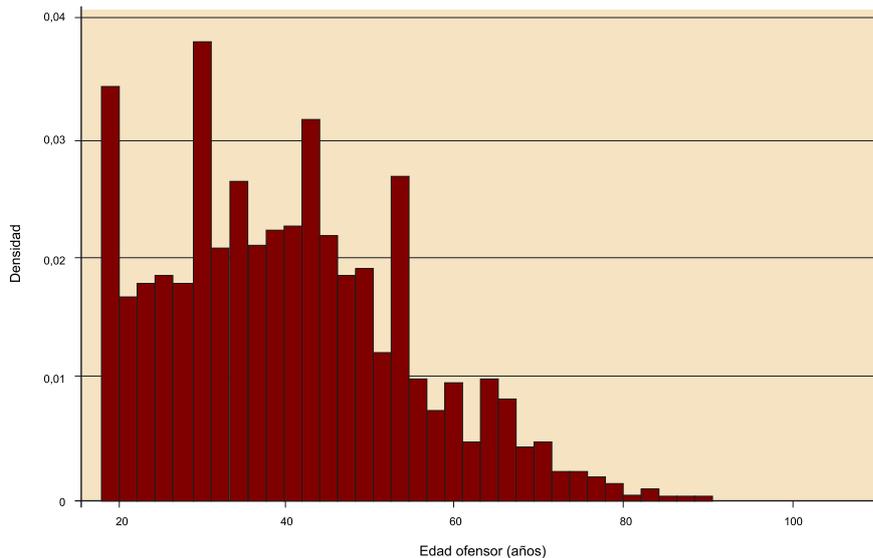
La **distribución muestral** es la distribución de los valores de los estadísticos muestrales cuando tenemos muchas muestras (infinitas) de una cierta magnitud de una determinada población.

Para entenderlo mejor, revisemos qué quiere decir una distribución de frecuencias. Una **distribución de frecuencias** es el número de veces que cada valor de una variable deviene en una muestra o en una población.

A menudo se presenta en una tabla de frecuencias, en la que para cada valor o rango de valores que puede adoptar una variable se atribuye el número de veces que aparece. Para las variables medidas numéricamente, la forma más normal de representar gráficamente la distribución de frecuencias es una **curva o histograma de frecuencia relativa** en la que se representa la proporción de veces que aparece cada valor o rango de valores en una población o en una muestra.

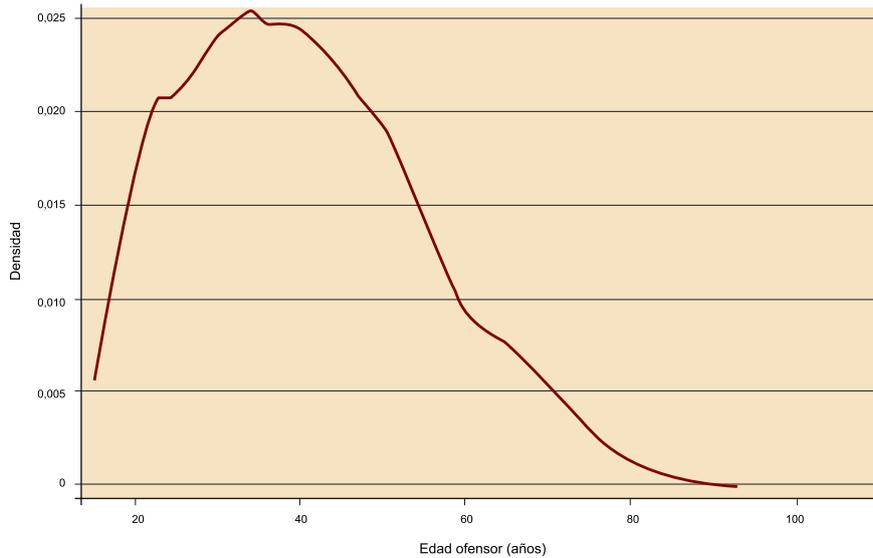
En la figura 3 vemos la distribución de frecuencias de la variable edad del ofensor en todas las sentencias sobre delitos sexuales a menores, de las audiencias provinciales españolas entre 2014 y 2016. Para diferentes rangos de edad representados en el eje de las  $x$ , tenemos la proporción de casos (en tanto por 1) en el eje de la  $y$ . Entre los 18 y los 20 años hay un 3,5 % de los casos. Si sumáramos todas las frecuencias de todos los rangos de edad, tendríamos el 100 % de los casos, 1.

Figura 3



En la figura 3, las proporciones se representan con rectángulos que corresponden a diferentes rangos de edad. Si los rangos de edad se hicieran infinitamente pequeños, entonces, en lugar de un diagrama de barras tendríamos un área debajo de la cual estaría el 100 % de los casos. En la figura 4 se representa la distribución de frecuencias suavizada.

Figura 4



## 2.2. ¿Cuál es la unidad de análisis de la distribución muestral?

Como hemos visto, en las unidades de análisis (o casos, o individuos) de los análisis cuantitativos, las filas de las matrices de datos son individuos: los ciudadanos de un Estado, los políticos de un parlamento, las viviendas en una ciudad, los robos en un barrio, los accidentes de tráfico en una región. Pero también pueden ser **agregados** de estos individuos. Por ejemplo, en lugar de tener la altura de los ciudadanos de un Estado (cada fila de la matriz es un ciudadano), podemos tener un agregado de las alturas, la media de las alturas por cada región (cada fila es una región). O podemos tener la suma de las viviendas de las ciudades de una región, el número de robos con violencia anual en una ciudad, el total de accidentes de tráfico con heridos por cada demarcación territorial de un semestre, etc.

La unidad de análisis de la distribución muestral es un agregado especial: la **muestra**.

En las investigaciones normalmente solo se realiza una muestra y solo se observa una muestra. Pero la distribución muestral es la distribución teórica de los valores de los estadísticos muestrales (media, proporción, desviación típica, etc.) en muchas muestras (infinitas) de un mismo tamaño y extraídas de una misma población. Es decir, la **distribución muestral** es el compendio de las probabilidades de los valores de un determinado estadístico calculado en una muestra de una población de un determinado tamaño.

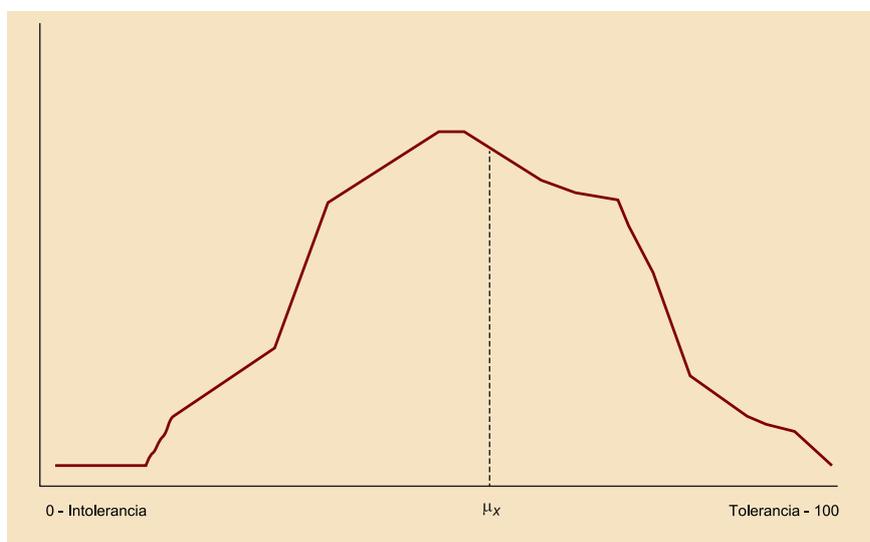
Las distribuciones muestrales de los diferentes estadísticos tienen fórmulas matemáticas que han sido estudiadas detalladamente por la estadística: por ejemplo, la curva normal, la *t* de Student, la Ji-cuadrado o la F. Como veremos en

este apartado, el hecho de conocer a fondo, con todo detalle, las distribuciones de los estadísticos es la clave para poder pasar de las muestras que obtenemos a poder inferir los valores de los parámetros poblacionales.

Veámoslo en un ejemplo.

Supongamos que la tolerancia de los habitantes de una gran ciudad (por ejemplo, Barcelona), ante la presencia de extranjeros no europeos en sus calles, puede ser descrita con un índice que va del 0 al 100. Si representáramos los valores de la variable «Tolerancia» ( $x$ ), que recoge a todos y cada uno de los ciudadanos de Barcelona, tendríamos como resultado el histograma de la figura 5. La curva tiene una forma aproximadamente normal: una forma de campana, en la que los valores medios son los más frecuentes, mientras que los valores más elevados y los valores más bajos son cada vez menos frecuentes. En la figura también está representada la media poblacional ( $\mu_x$ ), en valores centrales, pero un poco desplazada a la derecha, hacia los valores más tolerantes.

Figura 5



Si extraemos una muestra aleatoria de 500 individuos ( $n = 500$ ) y calculamos la media de la tolerancia de esta muestra ( $\bar{x}$ ), tendremos un valor que no necesariamente será el mismo que el de la media poblacional de la tolerancia ( $\mu_x$ ), pero como para crear la muestra hemos extraído aleatoriamente los casos de la población y todos tienen la misma probabilidad de aparecer, tenemos casi la misma probabilidad de que en la muestra haya casos más tolerantes que la media y casos más intolerantes que la media. Esto quiere decir que, del mismo modo que en una tirada de 500 veces de una moneda es muy difícil que siempre salga cara, en una muestra de 500 individuos es relativamente fácil que la mitad tengan valores más intolerantes que la media y la otra mitad sean más tolerantes. Por eso, siempre que calculemos la media de una muestra de 500 individuos de la ciudad, es muy probable que esta media tenga unos valores no demasiado alejados de la media poblacional. Además, en nuestra distribución hipotética de la tolerancia en Barcelona, tenemos la ventaja de que los valores más centrales, más parecidos a la media, son los más frecuentes. Así que, en este caso, todavía es más probable que la media de la muestra que hemos extraído sea muy cercana a la media poblacional.

En general, es relativamente fácil intuir que, cuanto mayor sea la muestra que extraemos de una población para calcular su media, más cercana será esta media a la media de la población. Es posible que una muestra pequeña solo tenga individuos en un lado de la distribución poblacional –solo intolerantes o solo tolerantes– del mismo modo que tirando solo tres veces una moneda es posible que las tres veces salga cruz. Pero cuanto mayor es la muestra, es más difícil que esto suceda. Del mismo modo, si la distribución poblacional está agrupada en torno a la media, como en nuestro caso, si hay poca variación en los valores poblacionales, es más fácil que los valores de las muestras sean cercanos a la media de la población.

Estas intuiciones las demuestra el **teorema central del límite**. Según este teorema:

1) La **distribución muestral de la media** es aproximadamente una curva normal (es decir, tiene forma de una campana de Gauss: una distribución simétrica, en la que los valores más frecuentes están en el centro, donde coinciden la moda, la mediana y la media de la distribución; mientras que la probabilidad se aleja a medida que nos alejamos del centro, y en los extremos la probabilidad tiende asintóticamente a cero).

2) La **media de la distribución muestral** ( $\mu_{\bar{x}}$ ) (es decir, la media poblacional de la distribución de las medias de las muestras de la población) es igual a la media poblacional ( $\mu_x$ ).

3) La **desviación típica** de la distribución muestral (que en las distribuciones muestrales adopta el nombre específico de **error típico**) es:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

donde  $n$  es el tamaño de la muestra.

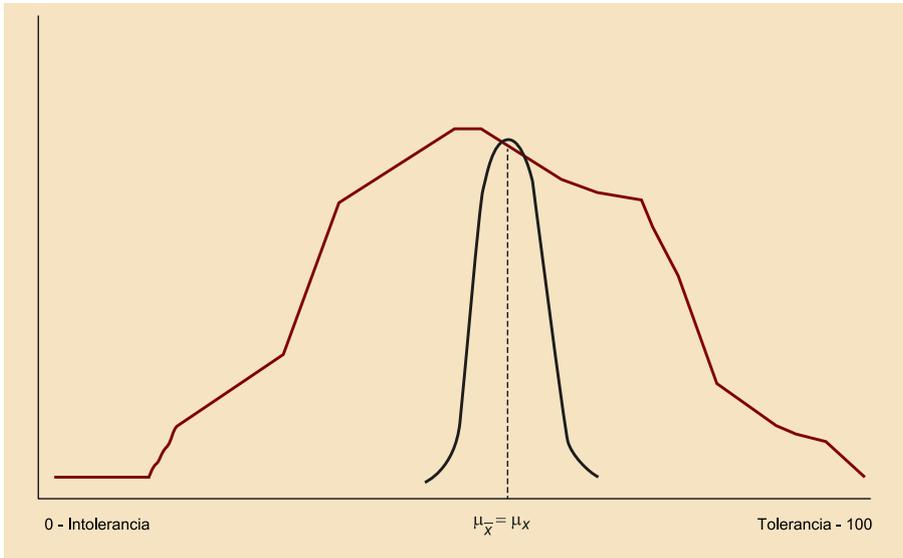
Es decir, la **dispersión** de la distribución muestral en torno a su media  $\sigma_{\bar{x}}$  (que es la misma media de la población) es igual a la dispersión de la población  $\sigma_x$  dividida por la raíz cuadrada del tamaño de la muestra  $\sqrt{n}$ .

Por lo tanto, la dispersión muestral es bastante más pequeña que la dispersión original de la población. Cuanto mayor sea la muestra  $n$ , menor será la dispersión. Si la muestra es  $n = 500$ , la dispersión muestral es  $1/\sqrt{500} = 1/22,22$  veces menor que la dispersión poblacional. Por lo tanto, hay mucha concentración de valores en torno a la media poblacional.

Pero esta dispersión no disminuye proporcionalmente en la población, puesto que es la raíz cuadrada del tamaño de la muestra. En un principio, el incremento del tamaño de la muestra tiene mucho impacto en la disminución del

error típico, pero el impacto del crecimiento de la muestra disminuye a medida que la muestra se hace grande y llega a un punto en el que el incremento de la muestra tiene un efecto muy bajo en la disminución del error.

Figura 6



En la figura 6, la distribución muestral de la media está sobrepuesta sobre la distribución de la población que hemos visto en la figura 5. Se ve cómo la distribución muestral tiene la forma de una curva normal, cómo la media muestral coincide con la media poblacional y cómo la dispersión de las dos distribuciones es muy diferente: la concentración de las medias muestrales en torno a la media de la distribución muestral es mucho más grande que la concentración de los datos poblacionales.

Estos resultados que vinculan una característica o parámetro de la población, su media, en relación con la distribución muestral de las medianas (estadísticos muestrales), los hemos ilustrado con la variable de la tolerancia, que tiene una distribución parecida a la normal con muchos valores concentrados en torno a la media. Pero hemos de tener en cuenta que estos resultados son válidos para cualquier distribución poblacional. Naturalmente, en una población mucho más dispersa, por ejemplo, una distribución uniforme, la dispersión de la distribución muestral (el error muestral) será mayor, pero la media muestral será la misma que la media poblacional y la disminución de la dispersión será igualmente proporcional al tamaño de la muestra.

### 2.3. ¿Qué queremos saber con las muestras?

Normalmente, con las muestras se quiere conocer cuál es el valor de un parámetro poblacional. En estadística inferencial no se hacen estimaciones puntuales de estos parámetros poblacionales, sino una **estimación del intervalo**.

Sabemos que los valores de los estadísticos que hemos obtenido en la muestra son una pista de en qué intervalo se encuentran los parámetros poblacionales que nos interesan. Pero aunque nuestra muestra se haya obtenido con un muestreo aleatorio, no podemos afirmar que el resultado obtenido en la muestra es el de la población. Sabemos que la distribución muestral contiene un cierto error en torno al parámetro que nos interesa. Por eso, en las estimaciones estadísticas hechas a partir de muestras, se obtiene un valor, un intervalo en torno a este valor y una probabilidad de que el valor poblacional esté dentro de este intervalo.

Por ejemplo, si en la muestra hemos observado una media de  $\bar{x} = 27,5$ , diremos que la media poblacional  $\mu_x$  está entre  $27,5 \pm 3$  (24,5, 30,5) con una cierta probabilidad, por ejemplo, 95 %.

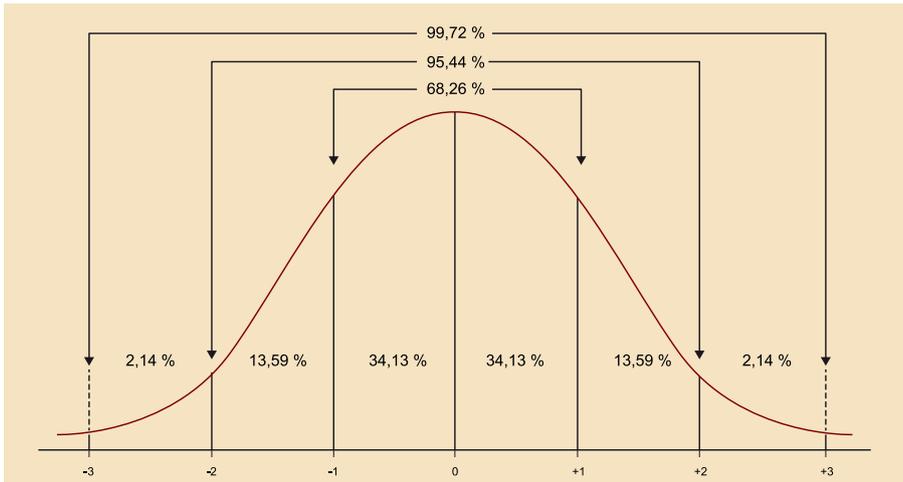
Los periodistas políticos tienden a pasar por alto estas características de los resultados obtenidos con muestras y, cuando describen los resultados de las encuestas preelectorales, pueden dedicar bastante tiempo a discutir una ventaja de pocos porcentuales a favor de un partido o candidato respecto del otro, cuando a veces no es posible afirmar con suficiente probabilidad que el candidato con mayor apoyo puntual en la muestra tiene realmente más apoyo que el otro entre los votantes, puesto que los intervalos de los resultados obtenidos están superpuestos. Pero para los periodistas estos detalles no tienen que estropear una buena historia.

#### 2.4. ¿Cuál es la mecánica para una estimación de intervalo?

En primer lugar, elegir un nivel de riesgo o de certeza que queremos aceptar. Convencionalmente se elige el 95 % de probabilidad, que quiere decir que estamos dispuestos a equivocarnos en un 5 % de las veces cuando decimos que un parámetro poblacional se encuentra dentro de un intervalo de valores.

El origen de este nivel de confianza proviene de los estudios de la agricultura científica y se basa en la distribución normal, puesto que es precisamente en este valor donde la curva de probabilidad deja de bajar pronunciadamente. En la figura 7 se representa la **curva normal tipificada** con una media 0 y una desviación típica de 1  $N(0, 1)$ . En cada desviación típica se ha marcado el área que engloba, es decir, la probabilidad de encontrar casos dentro. Por ejemplo, entre la media 0 y una desviación típica positiva se encuentra el 34,13 % de los casos. Entonces, entre más una desviación típica y menos una desviación típica se encuentra el 68,26 % de los casos. Entre  $-2$  desviaciones típicas y  $+2$  desviaciones típicas se encuentra el 95,44 % de los casos. El valor exacto para un área del 95 % es el rango entre  $-1,96$  y  $+1,96$ . En resumen, todas las probabilidades para cada valor son conocidas en una curva normal.

Figura 7. Curva normal tipificada



Todas las distribuciones pueden ser transformadas y tipificadas restando a sus valores la media y dividiéndolos por la desviación típica:

$$Z = \frac{x - \mu_x}{\sigma_x}$$

Cuando restamos a la media los valores por encima de la media, tenemos valores positivos, mientras que los que están por debajo de la media los tienen negativos. Ahora la media vale 0. Cuando dividimos estos valores por la desviación típica, tenemos los valores de la variable medidos en desviaciones típicas. Si el valor es 1, quiere decir que el valor está a una desviación típica de la media, etc.

La distribución muestral puede ser tipificada:

$$Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

Cada observación de la distribución muestral es un estadístico muestral, en este caso la media de la muestra:  $\bar{x}$ .

Por el teorema central del límite, sabemos que la media de la distribución muestral es la media poblacional:

$$\mu_{\bar{x}} = \mu_x$$

Mientras que el error típico (la desviación típica de la distribución muestral) es una proporción de la desviación típica poblacional relativa a la raíz cuadrada del tamaño de la muestra:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

Entonces, para saber el rango de valores que debe tener nuestro intervalo de confianza, solo hay que elegir los valores de la  $Z$  de nuestra distribución muestral que contengan la proporción de casos suficientes para el nivel de confianza con el cual queremos trabajar. Es decir, si elegimos un valor de la  $Z = \pm 1,96$ , sabemos (por la distribución de la figura 7) que en esta área estarán el 95 % de los casos.

A diferentes valores de  $Z$  corresponden diferentes niveles de confianza:

$$Z_{0,99} = 2,57$$

$$Z_{0,98} = 2,33$$

$$Z_{0,95} = 1,96$$

Recordemos que queremos calcular entre qué valores de nuestra variable (la altura, la tolerancia o la media de la opinión sobre el acoso escolar) se encuentra el parámetro poblacional a partir de la muestra que hemos obtenido de esta población.

Entonces escribimos el intervalo en el que se encuentra la  $Z$  con el 95 % de probabilidad:

$$\Pr\{-1,96 < Z < 1,96\} = 0,95$$

La  $Z$  es el resultado de tipificar la distribución muestral (sustituimos  $Z$  por la fórmula para tipificar una variable):

$$\Pr\left\{-1,96 < \frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}} < 1,96\right\} = 0,95$$

Aislamos el parámetro poblacional que queremos conocer, la media poblacional:

$$\Pr\left\{-1,96\left(\frac{\sigma_x}{\sqrt{n}}\right) < \bar{x} - \mu_x < 1,96\left(\frac{\sigma_x}{\sqrt{n}}\right)\right\} = 0,95$$

$$\Pr\left\{\bar{x} - 1,96\left(\frac{\sigma_x}{\sqrt{n}}\right) < \mu_x < \bar{x} + 1,96\left(\frac{\sigma_x}{\sqrt{n}}\right)\right\} = 0,95$$

Ahora ya podemos ver que los límites del intervalo, dentro del cual está la media poblacional con un 95 % de probabilidades, se calculan restando y sumando de la media muestral el producto de la puntuación  $Z_{0,95}$ , que es  $\pm 1,96$ , y el error típico (o desviación típica de la distribución muestral).

## 2.5. ¿Qué factores inciden en la deducción?

$$\Pr\left\{\bar{x} - Z_{n.c.}\left(\frac{\sigma_x}{\sqrt{n}}\right) < \mu_x < \bar{x} + Z_{n.c.}\left(\frac{\sigma_x}{\sqrt{n}}\right)\right\} = n.c.$$

¿De qué depende la amplitud del **intervalo de confianza** o el **margen de error**  $\epsilon$ ? En la fórmula anterior se han destacado los elementos que inciden en el tamaño del intervalo:

$$\epsilon = \pm Z_{n.c.} \left( \frac{\sigma_x}{\sqrt{n}} \right)$$

En primer lugar, tenemos el **nivel de confianza** elegido: cuanto más confianza necesitamos tener en los resultados (con el mismo tamaño de muestra), mayor será el margen de error (mayor es el valor absoluto de  $Z$ ).

En segundo lugar, tenemos la **variabilidad** (desviación típica) de la población: cuanto mayor sea la dispersión de los datos, mayor será el margen de error con un nivel de confianza y tamaño de la muestra dados.

Finalmente, tenemos el **tamaño de la muestra**. El tamaño de la muestra  $n$  divide la expresión, por lo tanto, cuanto mayor sea  $n$ , más pequeño será el margen de error.

Esta es la fórmula para calcular el tamaño de la muestra teniendo en cuenta el nivel de confianza deseado en la estimación de los resultados  $Z_{n.c.}$ , el margen de error que se quiere cometer  $\epsilon$  y la dispersión poblacional. Normalmente, en los cálculos de las muestras por encuestas suponen que hay una variación máxima en las variables que se quieren estudiar para curarse en salud.

$$n = \frac{Z^2 \cdot \sigma_x}{\epsilon^2}$$

Es notable que el tamaño de la población  $N$  de la que extraemos la muestra  $n$  no aparece en este resultado. La razón es que, cuando la población es grande (a partir de 100.000 individuos se considera una «población infinita»), su tamaño es irrelevante en la precisión y confianza de los resultados obtenidos. Si la población es más pequeña que 100.000 individuos, a medida que la  $N$  disminuye, la precisión y el nivel de confianza crecen si el tamaño de la muestra se mantiene igual.

Esto quiere decir que una muestra de 1.500 individuos será igualmente precisa y tendrá el mismo nivel de confianza tanto si trata de problemas referidos a una ciudad como Barcelona, como a otra de España.

En los Estados Unidos de América, por ejemplo, las encuestas preelectorales tienen la misma magnitud de muestra que tienen las típicas encuestas preelectorales realizadas para las elecciones generales españolas. Naturalmente, aquí aparece el problema de la heterogeneidad de la población. El tamaño de la población no afecta si esta se mantiene igualmente diversa, pero este supuesto es difícil de mantener a medida que la población se hace grande. En los Estados Unidos de América las muestras pueden ser pequeñas porque la relativa homogeneidad cultural del país lo permite, pero en el Estado español la diversidad cultural y política es mucho mayor y, rutinariamente, en las encuestas de opinión política se tienen que sobredimensionar las regiones donde hay un subsistema de partidos claramente diferenciado, como por ejemplo Euskadi o Cataluña.

Los investigadores prefieren el muestreo probabilístico porque la elección aleatoria permite utilizar la teoría de la probabilidad para estimar la precisión de los hallazgos obtenidos a partir de una muestra. En los resultados de cualquier estudio obtenidos por medio de una muestra, siempre se tienen que indicar los niveles de confianza y los intervalos de confianza.

### **Criminología**

En criminología, a pesar de que muchas veces las condiciones para tener muestras probabilísticas son difíciles y los supuestos en los que se basa la teoría de la probabilidad, por la poca calidad de los datos, casi nunca se cumplen estrictamente, es bueno estar familiarizado con la lógica que hay detrás de estos cálculos para evaluar correctamente los datos de vuestras investigaciones o los que aparecen en informes e investigaciones.

### 3. El muestreo probabilístico

A continuación, se da cuenta de diferentes diseños de muestreos probabilísticos básicos que pueden ser utilizados solos o en combinación, en función de los objetivos y las características de la investigación que se tenga que llevar a cabo. Es importante que, como investigadores y como consumidores de investigación, conozcamos cuáles son los fundamentos teóricos del muestreo y seamos capaces de elegir los métodos de muestreo más acordes con los objetivos de nuestra investigación y poder reconocer si las investigaciones realizadas por otros investigadores tienen problemas que puedan hacer dudar de las conclusiones a las que llegan en sus estudios o informes.

#### 3.1. El marco de muestreo

La operación previa necesaria para hacer un muestreo probabilístico es disponer o construir un marco de muestreo.

El **marco de muestreo** es la lista de todos los elementos de los cuales se tendrá que elegir la muestra probabilística, es decir, es una lista de la población que se quiere estudiar.

No siempre es posible conseguir una lista completa y por eso también se habla de **casi lista**. Hay que tener presente que el marco de muestreo es el que determina la población que efectivamente será muestreada y sobre la cual se pueden hacer legítimamente extrapolaciones con la muestra que se haya obtenido.

Las muestras que se hacen en organizaciones son las más simples porque normalmente hay un listado de sus miembros. Por ejemplo, en las escuelas o las universidades está el listado del profesorado, el personal de gestión y los estudiantes matriculados. Pero incluso en este caso, se tiene que pensar que es difícil obtener un listado completo de todos los que trabajan o estudian en la escuela. Es posible que haya personal temporal, estudiantes en prácticas u otros tipos no estándares que no aparezcan en el listado y, por lo tanto, en el marco de muestreo. El problema puede ser más o menos irrelevante, dependiendo del tipo de investigación que estemos llevando a cabo, pero tendríamos que ser conscientes y estar prevenidos a la hora de interpretar los resultados de la muestra. Las listas censales de los municipios o de los organismos oficiales son muy buenas cuando las usamos como marcos de muestreo, pero no son fáciles de obtener. Las encuestas de opinión suelen basarse en las guías telefónicas, pero, como hemos comentado, el avance de la telefonía móvil hace que cada vez sean más problemáticas a la hora de basar un marco de muestreo de la

población residente en un territorio. En general, los listados que se pueden obtener tienden a ser incompletos y, por lo tanto, pueden esconder sesgos de los que los investigadores no son conscientes.

Los marcos de muestreo se tienen que pensar como una **definición operativa de la población** que deseamos estudiar. Por ejemplo, si queremos estudiar la población de los magistrados que trabajan en las audiencias provinciales de España para saber su opinión sobre el punitivismo penal, tendríamos que encontrar formas que nos permitieran llegar a listados amplios de cada una de las audiencias provinciales. La información sobre los magistrados es relevante para los abogados y es probable que los listados de las asociaciones de abogados provinciales recojan esta información y la hagan pública. Naturalmente, basar el marco de muestreo en estas listas no oficiales puede ser problemático porque es posible que no se recoja esta información del mismo modo en todas las provincias o los datos estén igualmente actualizados en todas partes. Como investigadores, siempre tenemos que considerar los casos que pueden quedar fuera del marco de muestreo.

#### Encuestas electorales

En las encuestas electorales, el marco de muestreo se basa en los ciudadanos con derecho a voto de un territorio. Por esta razón, el marco de muestreo tiene que excluir a los menores de edad, los ciudadanos extranjeros sin derecho de voto, etc.

### 3.2. El muestreo simple aleatorio

El **muestreo simple aleatorio** consiste en un procedimiento que garantiza que todas las unidades de la población tienen la misma probabilidad de ser elegidas.

En primer lugar, se tiene que disponer del marco de muestreo y atribuir a cada elemento de la lista un único número. Después se tiene que utilizar un procedimiento de tipo lotería para elegir a tantos individuos o casos como haga falta para la muestra.

El procedimiento podría ser poner los números en un trozo de papel en una urna e ir extrayendo tantos números como haga falta. El procedimiento preferido hasta la presencia generalizada de ordenadores era utilizar las tablas de números aleatorios que aparecían en los apéndices de los libros de estadística. Actualmente, los programas estadísticos sirven para generar números aleatorios o para, directamente, elegir elementos del marco de muestreo.

### 3.3. El muestreo aleatorio sistemático

El **muestreo aleatorio sistemático** es una variación del muestreo simple aleatorio. Consiste en elegir los casos del listado que contenga la población con unos intervalos fijos.

Tenemos el tamaño de la población  $N$  y sabemos el tamaño de la muestra  $n$  que queremos extraer. Entonces solo se tiene que calcular la **fracción o intervalo de muestreo**  $f$  ( $f = N/n$ ) que nos indique cada cuántos casos elegiremos uno.

Supongamos que tenemos el listado con las 3.000 incidencias tratadas por una comisaría de policía en un periodo determinado, y queremos elegir 500 para analizar. El tamaño de la población es  $N = 3.000$ , y el tamaño de la muestra es  $n = 500$ . Entonces,  $3.000/500 = 6$  es la fracción de muestreo. Para evitar sesgos, se elige aleatoriamente un número del 1 al 6 para ser el primero incluido en la muestra y, a partir de este, se elige el sexto número que sigue en la lista.

El muestreo aleatorio sistemático producirá resultados aleatorios si no hay ningún elemento sistemático en el orden de casos en el marco de muestreo. Si los casos se ordenan en algún tipo de patrón cíclico, entonces se puede producir un sesgo en la muestra.

Si, por ejemplo, la lista de las incidencias tratadas en la comisaría sigue el orden temporal, es posible que sigan algún tipo de patrón y, por ejemplo, los casos que se han tratado los fines de semana, cuando hay gente que sale de fiesta, sean diferentes de los del resto de días de la semana. Si nuestra fracción de muestreo sigue esta periodicidad, es posible que la muestra esté sesgada.

Como el muestreo sistemático es una forma intuitivamente atractiva para elegir nuestros casos, siempre tenemos que examinar nuestras listas para detectar si hay algún tipo de patrón que pudiera producir un sesgo en la muestra que la invalidara para extraer conclusiones aplicables a la población.

### 3.4. El muestreo estratificado

El muestreo estratificado es una modificación del muestreo aleatorio simple y del muestreo sistemático, que hemos visto que sirve para evitar que nuestra muestra sea invalidada por el hecho de que, eligiendo muestras de forma puramente aleatoria, algunos subgrupos pequeños de la población pero interesantes en el análisis queden fuera de la muestra.

El **muestreo estratificado** es un método para mejorar la representatividad del aleatorio simple, que consiste en identificar los subgrupos dentro de la población que son relevantes en la investigación que se quiere llevar a cabo, garantizando que estos subgrupos estén presentes suficientemente en nuestra muestra eligiendo casos aleatoriamente dentro de cada subgrupo o estrato.

#### Primeros estudios de encuesta de opinión

Este tipo de problema se produjo realmente en algunos de los primeros estudios de encuesta de opinión que se hicieron durante la Segunda Guerra Mundial entre los soldados americanos. Se disponía de un listado en el que las unidades estaban ordenadas de forma que se empezaba por los mandos (sargentos, cabos) y después seguían los soldados rascos. Como las unidades eran aproximadamente del mismo tamaño, había un patrón sistemático. Como la fracción de muestreo se aproximó a este ciclo, el resultado fue una muestra sesgada que contenía muchos más mandos de los que correspondía si se hubiera elegido una muestra realmente aleatoria.

Desde el punto de vista de la teoría de la probabilidad, mejora la precisión y el nivel de confianza de la muestra porque, eligiendo casos en grupos más homogéneos y, por lo tanto, con menos dispersión respecto del fenómeno que queremos estudiar, disminuye el margen de error de la muestra.

En el ejemplo de la muestra de centros escolares de secundaria para hacer una investigación sobre el acoso escolar, que hemos utilizado para ilustrar una muestra sesgada, decíamos que teníamos motivos (basados en las investigaciones previas o en la intuición) para pensar que la forma como era tratado el acoso escolar podía depender de dos características de los centros: primero, si estaban situados en una zona rural o en una zona metropolitana y, segundo, si eran centros privados o públicos. Tener en cuenta los cuatro subgrupos de centros que aparecen al cruzar estas dos características y que supuestamente son más homogéneos respecto del acoso escolar, hará que nuestra muestra gane en representatividad porque eliminamos el componente de aleatoriedad en la pertenencia de cualquiera de estos grupos. Además, si alguno de estos subgrupos es muy pequeño (por ejemplo, el número de centros privados rurales), pero es necesario sacar conclusiones en nuestro estudio de él, solo por el error de muestreo, sería posible que los casos en la muestra fueran insuficientes.

En otro estudio que también se pregunte sobre el acoso dentro de las escuelas, pero que se quiera centrar en medir cuáles son las percepciones sobre estos comportamientos de los estudiantes en general y, en especial, de los que están directamente implicados, la cuestión de la muestra se tendrá que enfocar de forma diferente, pero también se puede tener que recurrir al muestreo estratificado. Por ejemplo, la población que quiere ser estudiada puede ser definida como los estudiantes matriculados en un centro escolar. En este estudio, los casos son los estudiantes y no los centros escolares de Cataluña; no estamos preocupados por averiguar las estrategias sobre el acoso en los centros escolares de secundaria, sino por medir los comportamientos y las actitudes de los estudiantes ante el acoso. Por lo tanto, el marco de muestreo será el listado de todos los estudiantes matriculados en aquel curso. Entonces, los investigadores pueden estratificar la muestra por todas aquellas características relevantes para la investigación según las pistas que nos proporciona la investigación anterior y la disponibilidad en el marco de muestreo. Seguramente se tendrá que estratificar por curso, puesto que es probable que el acoso cambie de los primeros cursos a los finales. Además de la cuestión de la edad y la madurez que están relacionadas con el curso, es interesante saber cuáles son los elementos que desencadenan el acoso y si hay un patrón de desarrollo en el tiempo. Otra característica para estratificar es el género de los estudiantes. Una vez efectuados los estratos, se tendría que elegir aleatoriamente (con alguno de los dos sistemas comentados) dentro de los estratos una submuestra proporcional al tamaño de cada estrato.

Las muestras se pueden estratificar tanto como se quiera (o se pueda) puesto que en general, cuando se hace, estamos ganando en representatividad y precisión de la muestra. El objetivo de la estratificación es conseguir que la población se agrupe homogéneamente en relación con las características relevantes de la investigación que se lleva a cabo.

Los estratos buscan la homogeneidad dentro de los estratos y la heterogeneidad entre los estratos.

Naturalmente, cuando el investigador estratifica una muestra se está imponiendo un trabajo muy complejo –es más difícil que hacer un muestreo aleatorio simple– para garantizar que la muestra obtenida no esté sesgada. Por eso solo se estratificará si resulta beneficioso para garantizar que las preguntas que se plantea la investigación podrán ser respondidas o si los costes de hacerlo son relativamente bajos. Por ejemplo, en el listado de los centros escolares está disponible la situación en el territorio y si son de propiedad privada o pública. Del mismo modo, en el listado de los estudiantes de un centro escolar tenemos información sobre el curso en el que están matriculados y su género. En los dos casos es viable hacer la estratificación.

Imaginemos que en el estudio del acoso en un centro escolar una de las preguntas importantes es si la etnia de los estudiantes influye en la emergencia de este comportamiento. Entonces seguramente será imprescindible estratificar nuestra muestra para garantizar que los estudiantes de etnias minoritarias también estén en la muestra. Sin la presencia de estos estudiantes en la muestra no será posible responder a una pregunta clave de la investigación. Hacer una muestra aleatoria ya es complicado, pero si resulta que por el azar finalmente no hemos capturado los casos que nos hacían más falta, será un derroche injustificable. Por lo tanto, tendríamos que hacer el esfuerzo de estratificar por la etnia de los estudiantes, a pesar de que este tipo de información sobre los estudiantes seguramente no será tan fácil de obtener como el género o el curso.

### 3.5. El muestreo estratificado no proporcional

Cuando nos encontramos con un caso como el que hemos descrito en el párrafo anterior, es posible que el grupo que nos interesa estudiar sea tan pequeño que, si la muestra aleatoria se hace proporcional en el peso de los grupos en la población, no queden suficientes casos para poder responder a las preguntas planteadas.

En el ejemplo anterior, si en el centro escolar hay un 10 % de niños de etnia gitana y hay recursos para hacer una muestra de tamaño  $n = 150$ , entonces solo tendremos quince niños de etnia gitana en la muestra. Si también tenemos en cuenta que, además de la etnia, tenemos otras preguntas (entre otras, el efecto del género o del curso en el que están), fácilmente esta cantidad no nos permita comprobar el efecto de la etnia que queremos averiguar.

El **muestreo estratificado no proporcional** resuelve este problema sobredimensionando la submuestra de los grupos o estratos más pequeños o complejos. Aparentemente esta operación se contradice con lo que hemos afirmado antes: la muestra representativa, no sesgada, tiene que tener la misma proporción en los grupos relevantes para la pregunta de investigación que la que hay en la población. Pero hemos de tener en cuenta que, si hemos aplicado alguna desproporción a la hora de hacer la muestra –como sucede cuando lo hemos hecho conscientemente, por diseño–, al analizar los resultados podremos «rectificar» la desproporción para obtener los resultados correctos en el ámbito poblacional. Cada estrato tendrá un **coeficiente de ponderación  $W$**  diferente, que permite analizar los resultados de la muestra correctamente en el ámbito poblacional.

Por ejemplo, si queremos estimar en la escuela la proporción de estudiantes que han sufrido acoso en la población, deberemos tener en cuenta que los estudiantes de los grupos grandes, que han sido infrarrepresentados en la muestra, tienen que contar más; mientras que los grupos pequeños en la población (por ejemplo, los niños de etnia gitana), que han sido sobrerrepresentados en la muestra, tendrán que contar menos. Hay tres tipos de formas de establecer la proporcionalidad de los grupos:

1) La **afijación simple** es la que solo tiene en cuenta la existencia de estratos y elige una submuestra del mismo tamaño para todos los estratos que hemos distinguido. Si se utilizara en el ejemplo de la escuela, la muestra tendría la misma cantidad de niños que de niñas, la misma cantidad de niños de cada curso, la misma cantidad de niños de la etnia mayoritaria que de las etnias minoritarias.

2) La **afijación proporcional** es la que hemos visto en el apartado anterior, en el que la proporción de los estratos en la muestra es la misma que la de la población.

3) La **afijación óptima** no solamente tiene en cuenta la proporción de cada estrato en la población, sino que también tiene en cuenta la dispersión que se observa en cada estrato respecto del problema que se quiere estudiar. Cuanto mayor sea la dispersión en un estrato, mayor será la proporción de población que se elegirá.

Naturalmente, para poder realizar este tipo de afijación se requiere conocimiento de la variable de interés de nuestra investigación, que normalmente tendremos que encontrar en las investigaciones previas sobre el caso que queremos estudiar. Si esta dispersión nos es desconocida y es importante para el diseño de nuestra muestra, es aconsejable realizar un estudio piloto (si nuestra investigación se basa en encuestas, una encuesta piloto en los diferentes estratos para estimar la magnitud de la dispersión en cada uno de estos estratos).

Este tipo de muestreo no proporcional ya lo hemos comentado en cuanto a las encuestas electorales españolas. En las típicas encuestas de opinión política del Centro de Investigaciones Sociológicas, la afijación de la muestra suele no ser proporcional en la población que vive en las diferentes regiones españolas. Las regiones con un sistema de partidos más complejo –hay más dispersión de voto porque más partidos políticos obtienen una cantidad sustantiva de votos–, como por ejemplo Cataluña o Euskadi, tienen una sobre-representación. Esto quiere decir que a la hora de hacer los análisis estadísticos para el conjunto del electorado español, se tendrán que tener en cuenta los pesos que corresponden a cada individuo de acuerdo con su estrato. Si no se ponderara la muestra de este modo, por ejemplo, la media de la autoubicación en la escala ideológica izquierda-derecha, la media mostraría un valor más hacia la izquierda del que hay en realidad, puesto que los electores de Euskadi y de Cataluña tienden a ubicarse en posiciones más hacia la izquierda que el resto de comunidades autónomas españolas.

### **Encuestas de victimización**

Las muestras de las encuestas de victimización en los Estados Unidos de América son un ejemplo de afijación no proporcional. Como su objetivo es tener una muestra de víctimas, determinados tipos de delitos, como por ejemplo los robos o las violaciones, son mucho más frecuentes en las grandes áreas metropolitanas que en las áreas rurales (es decir, la variabilidad de los delitos es mayor en las áreas urbanas que en las áreas rurales, que son más homogéneas) y, por lo tanto, la muestra de la encuesta de victimización sobrerrepresenta las áreas urbanas.

En las encuestas de victimización del Reino Unido, se hace una muestra desproporcionadamente mayor de las áreas rurales con fuerzas de policía más pequeñas, con objeto de obtener suficientes casos y poder comparar el funcionamiento de las diferentes policías que operan en todo el territorio.

Como se puede comprobar en estos ejemplos, el muestreo, que consiste en decidir quién o qué, dónde y cuándo se recogerán las evidencias que servirán para elaborar nuestros argumentos, no es una cuestión estándar para cualquier tipo de estudio. Conocer bien la pregunta de investigación que se intenta responder y cómo se ha intentado responderla en la investigación previa es importante para poder diseñar correctamente nuestro sistema de muestreo. Al mismo tiempo, toda investigación, para lograr los conocimientos teóricos de la realidad, tiene que considerar las restricciones que imponen los recursos de los que se dispone: esencialmente, tiempo y dinero. El reto es hacer diseños de muestreo que permitan responder eficazmente a las preguntas planteadas en la investigación y que tengan en cuenta las inevitables restricciones para que sean viables y eficientes.

El sistema de muestreo aleatorio siguiente es una muestra de cómo se pueden encontrar soluciones imaginativas a los problemas prácticos que plantea el diseño de muestras.

### **3.6. El muestreo de conglomerados (agrupamientos o clústeres)**

Tanto el muestreo aleatorio simple como el muestreo estratificado tienen un par de características que tienden a hacerlos muy costosos. Por un lado, requieren la existencia previa de un listado completo y detallado de todos los elementos de la población que se quiere estudiar. En estudios como las grandes encuestas de opinión no se suele disponer de esta información tan precisa (por eso se tiende a aproximarla con las listas de las guías telefónicas). Por otro lado, si los elementos de la población están muy dispersos en el territorio, re-

sulta muy costoso ir a todos los puntos en los que, por azar, se ha elegido un caso de la muestra. Esto es especialmente costoso para hacer estudios de encuestas presenciales. Si un individuo elegido para la muestra vive en un pueblecito de los Pirineos, un encuestador se tiene que desplazar solo para hacerle la entrevista a él. Y esto se tendría que repetir para cada individuo en todas las localidades donde la población estuviera distribuida. Este tipo de muestreo resulta inviable.

La idea que hay detrás del muestreo por conglomerados o clústeres es que, en muchos casos, la población está agrupada en subpoblaciones, dentro de las cuales es posible encontrar toda la variedad de individuos presentes en la población en relación con el problema que se quiere estudiar.

Un **conglomerado** o **clúster** es una unidad natural que contiene individuos de la población que se quiere estudiar.

#### Conglomerados

Por ejemplo, una escuela, una comisaría, un juzgado, una prisión, una ciudad, un bloque de viviendas, un vagón de tren de viajeros, todos pueden ser utilizados como conglomerados en una investigación.

La existencia de estas unidades naturales o conglomerados reduce los problemas de las muestras aleatorias a dimensiones tratables. En primer lugar, se pueden compilar con más facilidad los listados completos y detallados de estas unidades muestrales, sin tener que tener el listado completo y detallado de todos los elementos de la población. Una vez que se tiene el listado de los conglomerados, se puede realizar una muestra aleatoria simple. En segundo lugar, una vez que se han elegido los conglomerados que forman parte de la muestra, en una segunda fase, se analizan todos los miembros de cada conglomerado o una muestra aleatoria de estos hasta completar la muestra necesaria para el estudio. En el caso de estudios de encuestas presenciales, esta característica sirve para disminuir los costes de los desplazamientos.

Volvemos al estudio sobre el acoso escolar en secundaria. Los centros de secundaria son conglomerados, unidades naturales en las que podemos encontrar a los estudiantes, los elementos de la población de estudiantes de secundaria de Cataluña. Podemos listar todos los centros escolares de Cataluña y extraer una muestra aleatoria. Después, un equipo de entrevistadores se podría desplazar a los centros escolares elegidos y entrevistar a todos los estudiantes de la escuela o, entonces, realizar un muestreo aleatorio simple de los miembros de la escuela, utilizando el listado de estudiantes matriculados, como hemos visto anteriormente. De este modo obtendremos una muestra que se acerque bastante a una muestra aleatoria de la población de estudiantes en Cataluña y, además, podremos realizar las encuestas personales con un coste razonable.

Esta conclusión optimista se basa en un supuesto que no siempre es asumible: los agrupamientos o clústeres tienen que ser en su composición interna suficientemente **heterogéneos** respecto de las características que queremos estudiar en la población, y tienen que ser muy parecidos entre ellos, **homogéneos**, en esta heterogeneidad.

Esto quiere decir que en cada centro escolar tendríamos que poder encontrar todos los tipos de comportamientos y de actitudes respecto del acoso escolar que encontramos en el conjunto de la población. Si hay agrupamientos

que tienen mayoritariamente un determinado tipo de actitudes y otros agrupamientos que tienen otros, la elección aleatoria de los agrupamientos no nos permitirá garantizar razonablemente que podamos lograr un muestreo cercano al aleatorio.

Por lo que ya hemos dicho de la investigación sobre el acoso escolar en secundaria, es posible que en este caso no sea razonable pensar que los centros escolares (los agrupamientos) son muy parecidos entre ellos en los comportamientos y las actitudes ante el acoso. Si algunos de los agrupamientos tienen una proporción mayor de estudiantes que aceptan el acoso que otros, por ejemplo, supongamos que los centros privados aceptan con más facilidad o hacen menos caso de los casos de acoso que los centros público (esto solo es un supuesto), entonces, cuando estamos haciendo una elección aleatoria de centros escolares podemos estar incorporando inadvertidamente un sesgo a la muestra.

La idea de que hay detrás del agrupamiento natural de la población en unidades puede ser aprovechada de muchas formas en la investigación en ciencias sociales. Los estudios de comportamiento electoral desde el desarrollo de las encuestas científicas buscan, como el Santo Grial, la población ideal que se comporta electoralmente exactamente como el conjunto de la población, de modo que solo hay que saber cómo se comportará esta población para acertar quién será el próximo presidente de los Estados Unidos de América o cuáles serán los resultados electorales de Cataluña. En Cataluña se ha dicho en algún momento que esta población podría ser Granollers.

En algunos estudios especialmente complejos se puede utilizar imaginativamente la idea de conglomerado. Por ejemplo, la Autoridad Metropolitana de Barcelona del Transporte que agrupa diferentes compañías de transporte público en diferentes modos (esencialmente metro, Renfe, FGC y líneas de autobuses) está interesada en establecer cuál es el grado de cumplimiento de los usuarios de las diferentes líneas gestionadas por cada uno de los operadores con objeto de poder distribuir los beneficios de los sistemas de pago integrados, que son válidos para todos los transportes públicos. Para conocer el grado de cumplimiento, se tienen que hacer inspecciones aleatorias de los viajeros de cada línea, pero ¿cómo se puede hacer si no tenemos listados completos y detallados de los pasajeros? La solución es aprovechar que los viajeros están agrupados naturalmente en los vehículos (vagones de metro o tren o autobuses), elegir aleatoriamente algunas de estas unidades muestrales e inspeccionar a todos los pasajeros para saber si han comprado billete o no.

Hemos visto que el procedimiento del muestreo por conglomerados está pensado para dar respuesta a los problemas prácticos del muestreo aleatorio simple, pero también hemos visto que su punto débil es suponer que las unidades muestrales son iguales o muy parecidas entre ellas, de modo que contienen en sí mismas toda la variedad que hay en la población (el lema de los conglomerados es: heterogeneidad interna y homogeneidad externa). Pero sabemos que en muchos casos este supuesto no se puede mantener. De modo que se tiene que suponer que siempre que se utiliza el muestreo por conglomerados disminuye la precisión y el nivel de confianza de los resultados obtenidos.

El muestreo por conglomerados siempre se lleva a cabo, como mínimo, en dos fases: primero hay una selección aleatoria de los conglomerados (las unidades muestrales) y después, dentro de cada conglomerado, se seleccionan los

### Magic Town

La idea del clúster también es el tema que hay detrás de una película americana del año 1947. El protagonista de *Magic Town*, Rip Smith (James Stewart), dirige una empresa de sondeos de opinión que está a punto de quebrar por los costes de hacer los estudios. Por casualidad, descubre una ciudad, Grandview, que es el espejo perfecto de la opinión de todos los EE. UU. y esto le salva la empresa... de momento, puesto que es capaz de hacer buenos estudios rápidamente y casi sin costes. Pero la crisis no tarda en llegar porque, cuando los ciudadanos de Grandview descubren que son observados y utilizados como modelo de todo el país, cambian su comportamiento, se informan y se forman sobre las cuestiones cívicas y políticas que les preguntan y, como resultado, dejan de ser representativos de la opinión dominante de todos los Estados Unidos de América.

La película se inspira en los estudios sobre Middletown que hicieron la pareja de sociólogos Robert y Helen Lynd en los años treinta del siglo pasado.

individuos finales de la muestra. Por esta razón, a veces también se denomina *muestreo multinivel de conglomerados*. De hecho, el muestreo por conglomerados podría tener tantas etapas o fases como poblaciones naturales podemos distinguir encajadas las unas dentro de las otras. Por ejemplo, podemos tener todos los centros escolares de Cataluña y hacer una muestra aleatoria, y a continuación, dentro de estas escuelas se pueden distinguir los grupos clase y elegir aleatoriamente los que sean convenientes para completar la muestra. Finalmente, podríamos seleccionar a todos los estudiantes del grupo clase que se estudiarán. Pero como hemos dicho, cada vez que utilizamos una etapa de muestreo por conglomerados estaremos perdiendo precisión en la muestra final que obtengamos.

El procedimiento probabilístico siguiente intenta resolver esta dificultad haciendo algo más complejo el muestreo por conglomerados teniendo en cuenta los estratos, de forma que se compense la potencial pérdida de precisión producto de la utilización de conglomerados.

### **3.7. El muestreo multinivel (o de conglomerados con estratificación)**

El **muestreo multinivel** combina el muestreo por conglomerados con el muestreo estratificado. El problema de la heterogeneidad entre los conglomerados del procedimiento anterior se resuelve a fuerza de agruparlos en estratos homogéneos. Entonces, el problema potencial de la pérdida de precisión quedará compensado o mitigado.

En el ejemplo del acoso escolar en la educación secundaria, la estratificación por zona urbana y rural de la escuela o por propiedad privada o pública ya se ha comentado y puede ser aplicada a la hora de elegir las diferentes escuelas (**agrupamientos** o conglomerados). Se pueden añadir tantos estratos como se considere que puede ser útil para mejorar la homogeneidad interna de las unidades. Por ejemplo, podemos sospechar que el tamaño del centro escolar puede ser un factor que afecte a la forma como se puede producir y al acoso. Por eso se podría añadir en este nivel una estratificación en tres niveles que tenga en cuenta el número de estudiantes matriculados en los centros. Hemos visto que en el interior de los centros se pueden distinguir los grupos clase como unidades naturales y utilizarlos para hacer una selección aleatoria. Pero en este nivel también se puede estratificar por factores que pueden afectar al fenómeno que queremos estudiar. En este caso una alternativa obvia es entre los grupos por la mañana y los grupos por la tarde. Teniéndolo en cuenta, podremos tener una muestra más precisa y representativa de la población.

En general, en los muestreos multinivel a cualquier nivel se pueden crear conglomerados y estratos. Todo depende de cómo se puede aprovechar más eficientemente la información que tenemos de la forma como se dispone la población, y las diferencias que puede haber entre los diferentes conglomerados. Si sabemos que los conglomerados son suficientemente homogéneos, no habrá que estratificarlos.

También se recomienda maximizar el número de agrupamientos de modo que estos contengan pocos individuos en su interior para que haya menos pérdida de precisión en la muestra total. Por ejemplo, si queremos hacer una encuesta sobre la sensación de inseguridad en una ciudad, podemos distinguir en una

primera etapa el listado completo de las calles que hay en la ciudad. Si hay diferencias importantes entre los diferentes barrios, tendríamos que estratificar las calles según la pertenencia a estos barrios. A partir del listado que identifica las calles, se tendría que hacer un muestreo aleatorio simple entre estas calles. En otra etapa, se tendrían que listar todas las viviendas que encontramos en cada calle y seleccionar aleatoriamente el número de casas necesario para completar la muestra. Finalmente, entre las personas que viven en cada vivienda se puede volver a seleccionar aleatoriamente una.

### **Encuestas nacionales de victimización delictiva**

Las encuestas nacionales de victimización delictiva que se hacen en los Estados Unidos de América siguen un procedimiento multinivel con agrupamientos. El objetivo de la encuesta es representar a la población americana de más de 12 años que vive en casas. Esta especificación es relevante, porque no incluirá la población que vive en viviendas institucionales, como por ejemplo los militares. Por este mismo motivo, tampoco podrán proporcionar estimaciones de los delitos que sufren los establecimientos comerciales o los negocios. Como en los EE. UU. no hay una lista de las viviendas, se tiene que hacer una muestra con agrupamientos con diferentes fases. En una primera fase, se definen como unidades de muestreo primarias las diferentes áreas administrativas en las que se divide el país. Estas áreas son estratificadas según sean grandes áreas metropolitanas, condados no metropolitanos o agrupaciones de condados del alrededor de las áreas metropolitanas. Las áreas más grandes son incluidas de forma automática en la muestra, mientras que las otras áreas son estratificadas por tamaño, densidad de población y número de delitos reportados, entre otras variables, para hacer un muestreo estratificado proporcional a la población de cada estrato. En una segunda fase, se listan todas las viviendas del área dedicando mucho esfuerzo para que la lista sea completa y esté al día. En la última encuesta del año 2018, estos procedimientos han servido para encuestar a 160.000 personas en 95.000 viviendas.

Fuente: M. G. Maxfield y E. R. Babbie (2015). *Research methods for criminal justice and criminology* (7.ª ed., p. 220-221). Stamford, CT: Cengage Learning.

A pesar de que los científicos prefieren los muestreos probabilísticos, porque evitan sesgos conscientes o inconscientes en la selección de nuestros casos y, además, permiten hacer estimaciones de los errores muestrales, muchas veces no es posible llevarlos a cabo. En estas situaciones tenemos que utilizar los procedimientos de muestreo no probabilísticos.

## 4. Los muestreos no probabilísticos

El rasgo característico de los muestreos no probabilísticos es el desconocimiento de la probabilidad de que un individuo forme parte de la muestra, puesto que no disponemos de un listado completo y detallado de todos los individuos de la población. Esta situación afecta a muchas investigaciones criminológicas que normalmente se interesan por comportamientos delictivos o socialmente desviados y que, por lo tanto, se esconden. A pesar de esta limitación, los procedimientos no probabilísticos se esfuerzan por evitar al máximo los sesgos en su captura de casos para la muestra. Y a veces, atendido al tipo de conductas que se quieren estudiar, estos procedimientos pueden producir resultados superiores a los que se habrían obtenido aunque fuera posible hacer una muestra aleatoria.

### 4.1. El muestreo intencional

El **muestreo intencional** consiste en seleccionar los casos de la muestra de acuerdo con el conocimiento de la población y de los elementos que la forman.

Por ejemplo, en el estudio sobre el tratamiento que reciben las víctimas de violencia de pareja en los juzgados, es importante que se tenga en cuenta el punto de vista de muchos colectivos: las víctimas, los jueces, los abogados, las personas de los servicios sociales que las atienden, etc. No es posible tener un listado completo de cada uno de estos colectivos a la hora de elegir aleatoriamente una muestra. Por lo tanto, el investigador solo necesitará encontrar a unos cuantos miembros de estos colectivos con objeto de tener una idea de cómo valoran el tratamiento que reciben las víctimas, qué problemas aparecen y cómo creen que los pueden mejorar. Para llevar a cabo una evaluación de la implementación de una política o programa de tratamiento de las víctimas, lo importante es tener una variedad de puntos de vista suficiente sobre la cuestión y esto se consigue de forma adecuada con la muestra intencional. En este caso, la información obtenida puede ser mejor que tener una muestra representativa de cada uno de estos grupos a partir de la cual se pueda extrapolar información a la población.

A veces las investigaciones que se están llevando a cabo son tan complejas que los criterios cruciales a la hora de elegir los casos provienen de la facilidad de disponer de acceso a las instituciones públicas.

Por ejemplo, en un estudio sobre los flujos de los casos de delitos sexuales a menores de la Administración de justicia española, el equipo liderado por el profesor Josep M. Tamarit (2017) se centró en la circunscripción de Lleida porque sus dimensiones reducidas y el conocimiento del entorno permitió reconstruir la evolución de los casos dentro del sistema a pesar de todas las dificultades. Se pudo argumentar convincentemente que los resultados de la investigación, a pesar de que limitados a la provincia de Lleida, podían ser extrapolados al conjunto del sistema judicial español por la tendencia a la uniformidad de este sistema.

En otros casos, el interés del investigador es conseguir casos que contengan la máxima variabilidad respecto de algunos factores. Entonces, puede ser conveniente centrarse en casos que presenten situaciones muy diferentes para poder comparar cuál es el efecto de los factores que se quieren estudiar.

Todos los estudios que utilicen este tipo de procedimiento de muestreo tendrán que argumentar sus decisiones a la hora de elegir unos casos determinados. Cuando lo hagan, intentarán justificar la decisión y hacer ver cómo los criterios a la hora de elegir los casos no introducen un sesgo en la muestra o minan la credibilidad de los resultados a los que se llega con el estudio.

## 4.2. El muestreo por cuotas

El **muestreo por cuotas** es un intento de hacer una muestra representativa de la población, tal como lo hacen los muestreos probabilísticos, pero sin realmente utilizar procedimientos aleatorios para seleccionar los casos.

Los dos tipos de procedimientos utilizan información parecida de la población, y esto puede confundir a quien no los conoce lo suficiente. Pero, como veremos, a pesar del parecido en los objetivos y la utilización de la información, los dos procedimientos tienen aproximaciones a la cuestión de la obtención de la muestra bastante diferentes.

El **muestreo por cuotas** parte de la distribución de las características que se conocen de la población objetivo. Por ejemplo, para conseguir las muestras que se utilizan a la hora de hacer encuestas telefónicas, se elabora una matriz con las proporciones de la población que tiene combinaciones de diferentes variables sociodemográficas: el género, la edad, el nivel educativo, el lugar de residencia, etc. Calculan qué porcentaje de la población son mujeres, de edad mediana con un nivel educativo universitario, que viven en la provincia de Barcelona, etc. Una vez que tienen esta matriz, pueden saber las **cuotas**, la cantidad de población que habrá que entrevistar con la combinación de características de cada celda, para conseguir que la muestra sea idéntica a la población con las proporciones de las combinaciones de las características sociodemográficas. Finalmente, a los entrevistadores se les pide que busquen a los individuos que cumplan las cuotas que les son encomendadas.

El resultado que obtenemos con este tipo de muestreo se asemeja al que se obtiene con el muestreo probabilístico, pero tiene el problema de que, a la hora de llenar las celdas, puede haber sesgo en los individuos que acaban entrando en la muestra: es probable que haya individuos que tengan más probabilidad que otros de formar parte de la muestra por muchas razones que no controlamos del todo (más disponibles, más extravertidos, etc.). Este problema es especialmente importante en las encuestas preelectorales, en las que no todos

los electores están igualmente movilizados y dispuestos a revelar sus intenciones de voto, lo que finalmente puede estropear las predicciones a pesar de las correcciones que se hagan.

En las encuestas presenciales, una estrategia para evitar la discrecionalidad de los entrevistadores para llevar a cabo una muestra de cuota es darles unas instrucciones muy precisas de las viviendas que pueden contactar para obtener una entrevista. Este sistema es conocido como **rutas aleatorias** (aunque realmente no sea un sistema propiamente aleatorio, sino que quiera acercarse a él). En las instrucciones se dice el punto de la población en el que tiene que empezar la ruta, cuántas viviendas tiene que pasar después de girar hacia la derecha o izquierda, antes de llamar para pedir hacer la entrevista.

Las muestras por cuota no son muestras aleatorias pero se combinan con las rutas aleatorias, que eliminan una buena parte de la discrecionalidad a la hora hacer la encuesta, producen unos resultados que se han calificado de semi-aleatorios. De hecho, algunos de los estudios más rigurosos llevados a cabo por el Centro de Investigaciones Sociológicas se llevan a cabo utilizando procedimientos de muestreo multinivel, combinando agrupamientos y estratificaciones y acaban seleccionando las unidades últimas (los individuos) por rutas aleatorias y cuotas de sexo y edad. A pesar de que la muestra no es estrictamente aleatoria, en la ficha técnica se dice el error muestral en el que se incurre.

### Ficha técnica del barómetro de cultura política del CIS, Estudio 2.387

**Ámbito:**

Nacional. Se incluyen las provincias insulares y se excluyen Ceuta y Melilla.

**Universo:**

Población española de ambos sexos de 18 años y más.

**Tamaño de la muestra:**

Diseñada: 2.500 entrevistas.

Realizada: 2.484 entrevistas.

**Afijación:**

Proporcional.

**Ponderación:**

No procede.

**Puntos de muestreo:**

168 municipios y 47 provincias.

**Procedimiento de muestreo:**

Polietápico, estratificado por conglomerados, con selección de las unidades primarias de muestreo (municipios) y de las unidades secundarias (secciones) de forma aleatoria proporcional, y de las unidades últimas (individuos) por rutas aleatorias y cuotas de sexo y edad.

Los estratos se han formado por el cruce de las diecisiete comunidades autónomas como el tamaño de hábitat, dividido en siete categorías: menor o igual a 2.000 habitantes; de 2.001 a 10.000; de 10.001 a 50.000; de 50.001 a 100.000; de 100.001 a 400.000; de 400.001 a 1.000.000, y más de 1.000.000 de habitantes.

Los cuestionarios se han aplicado mediante entrevista personal en los domicilios.

**Error muestral:**

Para un nivel de confianza del 95,5 % (dos sigmas), y  $P = Q$ , el error es de  $\pm 2$  % para el conjunto de la muestra y en el supuesto de muestreo aleatorio simple.

**Fecha de realización:**

Del 24 al 28 de marzo de 2000.

### 4.3. Las muestras de conveniencia

Las **muestras de conveniencia** son las que se fían de los sujetos que están disponibles en un lugar o un espacio determinado para formar parte de la muestra.

#### Muestras de conveniencia

Por ejemplo, en los estudios de percepción de seguridad en las calles, es lógico utilizar los mismos peatones de la calle.

Los procedimientos para evaluar las protestas y manifestaciones tienen que usar a la gente disponible en la calle para pedirles su participación. Se establecen protocolos a la hora de elegir a la gente de manera sistemática (por ejemplo, contactar con cada diez manifestantes que caminan por el lado derecho de la vía), de manera que se evite que los encuestadores tiendan a elegir inconscientemente a los manifestantes que tengan elementos que les llamen la atención. Los manifestantes contactados que se han manifestado receptivos no son entrevistados en aquel momento, sino que se les pide que dejen el teléfono o la dirección de correo electrónico con objeto de poder hacer la entrevista en otro momento con más calma.

Del mismo modo, en estudios de criminalidad en barrios conflictivos o zonas de ocio, es posible establecer procedimientos de muestreo sistemáticos de población disponible para recoger las opiniones y experiencias relevantes con objeto de orientar las políticas de mejora de la zona.

#### 4.4. Las muestras de bola de nieve

El **muestreo de bola de nieve** consiste en identificar en un primer momento a un individuo o un pequeño grupo de individuos que forma parte de la población que queremos estudiar y pedirles que nos ayuden a identificar a otros individuos como ellos que quieran participar en el estudio.

Este tipo de muestreo se utiliza habitualmente en los estudios de delincuentes o de personas con comportamientos no aceptados socialmente. Normalmente, este tipo de muestreo se utiliza para llevar a cabo estudios cualitativos en profundidad de los sujetos estudiados. No se busca la representatividad estadística, sino entender las interpretaciones que los sujetos hacen de sus actividades.

En el muestreo de bola de nieve muchas veces es crucial el individuo que hace de primer contacto. Lo más sencillo a la hora de identificar a los delincuentes es fiarse de aquellos que están fichados por la policía, pero precisamente estos pueden ser una fuente de sesgo, o desviarse de la población que tenemos como objetivo.

En general, los muestreos no aleatorios, como por ejemplo las muestras intencionales, de conveniencia y de bola de nieve, son útiles cuando es imposible determinar la probabilidad de los elementos que acaban formando parte de la muestra. Son estrategias de muestreo especialmente útiles cuando la población objetivo de la investigación es difícil de situar o identificar, como es el caso de muchos estudios criminológicos.



## Bibliografía

- Brady, H. E. y Collier, D. (2010). *Rethinking social inquiry: Diverse tools, shared standards*. Lanham: Rowman & Littlefield.
- Chatterjee, S. K. (2003). *Statistical Thought: A Perspective and History*. Oxford: Oxford University Press.
- Copleston, F. (1981). *Historia de la filosofía* (vol. 3). Barcelona: Ariel.
- Glaser, B. G. y Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: De Gruyter.
- Kahneman, D., Slovic, P., y Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Lynd, R. S. y Lynd, H. M. (1935). *Middletown in transition: A study in cultural conflicts*. California: Harcourt.
- Manheim, J. B. y Rich, R. C. (1988). *Análisis político empírico. Métodos de investigación en ciencia política*. Madrid: Alianza.
- Maxfield, M. G. y Babbie, E. R. (2015). *Research methods for criminal justice and criminology* (Seventh edition). Stamford, CT: Cengage Learning.
- Nisbett, R. y Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Tamarit Sumalla, J. M., Guardiola Lago, M. J., Padró-Solanet, A., y Hernández-Hidalgo, P. (2017). The judicial pursuit of the sexual victimization of children: How the criminal justice system processes cases. *International Review of Victimology*, 23(2), 123-144. doi:10.1177/0269758016680867

