CAPÍTULO 6

La selección de la muestra en el proceso de investigación

Vanesa Martello y Agustín Cleve

> No podemos estudiar todos y cada uno de los casos de aquello que nos interesa, ni tampoco tendríamos que desear hacerlo. Todo emprendimiento científico intenta descubrir algo pasible de ser aplicado a todas las cosas de una cierta clase estudiando unos pocos ejemplos, cuyos resultados, como se dice en la jerga, sean "generalizables" a la totalidad de los miembros de esa clase (Becker, 1998, p.95).

El presente capítulo aborda las nociones de población, muestra y muestreo dentro del proceso de investigación, así como los tipos de muestreo probabilístico y no probabilístico para la selección de los casos de estudio. Pero antes de adentrarnos en estas definiciones, es necesario que nos ubiquemos dentro del proceso de investigación para comprender en qué momento en particular se lleva adelante la construcción de la muestra.

Piovani (2018) señala que todo diseño de investigación implica decisiones en torno a:

- La construcción del objeto
- La selección de las unidades
- La recolección de la información
- El análisis de la información

En este capítulo, desarrollamos esta segunda gran decisión de todo diseño de investigación que se vincula a la pregunta: ¿cómo seleccionaremos las unidades sobre las cuales vamos a relevar información? Y esta selección de unidades nos llevará a la conformación de la muestra de estudio.

En un conocido manual de metodología, Hernández Sampieri (2010) se pregunta “¿En una investigación siempre tenemos una muestra?” a lo cual responde:

> No siempre, pero en la mayoría de las situaciones sí realizamos el estudio con una muestra. Sólo cuando queremos realizar un censo debemos incluir en el estudio a todos los casos (personas, animales, plantas, objetos) del universo o la población. Por ejemplo, los estudios motivacionales en empresas suelen abarcar a todos sus empleados para evitar que los excluidos piensen que su opinión no se toma en cuenta. Las muestras se utilizan por economía de tiempo y recursos. (p. 172)

En toda indagación científica es necesario acotar espacio-temporalmente las poblaciones que constituyen el objeto de estudio. Si bien la definición y delimitación de la población responde a criterios teóricos, debe adecuarse a las posibilidades empíricas y prácticas (Dimarco y Martello, 2016). En este capítulo, se busca proveer elementos conceptuales y prácticos para orientar a les estudiantes en este proceso decisivo.

Definiciones de población, muestra y muestreo

Cuando se formula el problema y los objetivos de investigación, se realiza un recorte empírico que nos posibilitará saber en dónde buscaremos la información que necesitamos. Dicha información específica debe obtenerse de lo que se denomina unidad de análisis, es decir, de los casos o entidades que inevitablemente se deben observar para responder a los objetivos propuestos. Por ejemplo, si interesa estudiar la carrera de Trabajo Social, se deberán recolectar datos organizacionales, siendo la carrera en sí la unidad de análisis. Distinta situación tendríamos si el objetivo fuera conocer las expectativas y satisfacción de les estudiantes sobre la carrera, en cuyo caso elles mismes, en tanto individuos, constituirían la unidad de análisis. Asimismo, si se pretende analizar medios de información sobre protestas sociales, las noticias como productos sociales serían la unidad de análisis.

Evidentemente, no es factible entrevistar a la totalidad de estudiantes o examinar todas las noticias que componen la unidad de análisis. En este sentido, es pertinente realizar la pregunta: ¿cuántos casos quiero estudiar de la unidad de análisis? ¿Qué características deben tener? ¿Cómo los voy a elegir? Ello conduce rápidamente a las nociones de muestra y muestreo, referidas a cuántas unidades considerar, sus características y representatividad. Tales interrogantes resultan fundamentales para avanzar en la investigación.

Definir con precisión la población de interés es un paso esencial en el proceso investigativo. La **población** refiere al conjunto total de casos que cumplen determinados criterios establecidos por les investigadores para delimitar la problemática a examinar. Debe quedar expresada sin ambigüedades en términos de contenido, lugar y período temporal.

Una vez delimitada la población, se debe seleccionar una **muestra,** esto es, un subconjunto de la población sobre el cual se recolectará la información. El procedimiento para seleccionar la muestra se denomina **muestreo**.

Existen dos tipos de muestreo: probabilístico y no probabilístico. En el muestreo probabilístico todos los elementos de la población tienen una probabilidad conocida y distinta de cero de ser escogidos para formar parte de la muestra. Ello permite hacer inferencias estadísticas y generalizar resultados. En contraste, en el muestreo no probabilístico la selección se basa en otros criterios deliberados por les investigadores. Su representatividad es limitada.

Recapitulando, la selección de casos involucra diferentes momentos: primero, definir la población o universo a partir de criterios teórico-metodológicos derivados de los objetivos del estudio. Segundo, establecer conceptualmente los criterios de selección de casos/unidades y su

aplicación concreta sobre el terreno. Tercero, optar por las técnicas e instrumentos que permitirán acceder a la muestra deseada. En metodologías cuantitativas, el muestreo aleatorio o no aleatorio se sitúa en esta etapa. Cabe destacar que estos pasos se encuentran interrelacionados, cada uno influye sobre el otro. Por ejemplo, la forma de acceso a los casos puede obligar a redefinir los propósitos teóricos.

Tipos de muestreo y representatividad

La muestra en investigación científica debe reflejar los rasgos característicos o comportamientos regulares y necesarios de aquello que se investiga. Su función es identificar, a través de ella, esos aspectos distintivos del objeto de estudio. Esta capacidad de representar aquello que es regular o característico en el objeto investigado se denomina representatividad de la muestra. Si consideramos a ese objeto como el universo de referencia, la muestra representativa puede definirse como aquella que reproduce fielmente las propiedades esenciales de ese universo. Esto implica que la muestra contiene de modo apropiado los rasgos principales del conjunto mayor al que pertenece.

Decir que una muestra es representativa es lo mismo que decir que tenemos fundamento suficiente para pensar que la estructura de la muestra es análoga a la estructura del universo (Samaja, 1993).

Pongamos como ejemplo que queremos realizar un estudio para conocer las opiniones sobre la legalización del aborto de les residentes adultes de la ciudad de La Plata. Nuestra hipótesis es que la edad y la religión de les encuestades van a incidir en su posicionamiento al respecto.

En este caso, la población objetivo serían todos los habitantes adultos de La Plata. Para que la muestra sea representativa de esta población, necesitamos incluir participantes de los diferentes segmentos relevantes:

- Diferentes grupos etarios (jóvenes de 18 a 35 años, adultes de 36 a 55, mayores de 56) en una proporción similar a la que exista en la ciudad.
- Creyentes de distintas religiones (católicos, evangélicos, otras), también respetando su peso poblacional.
- Personas de distintos género y nivel educativo.

Si construyéramos una muestra de 500 casos, esta debería respetar la distribución real de la población de La Plata en esas variables clave, entonces podremos considerarla representativa y extrapolar los resultados sobre la opinión del aborto al total de los habitantes adultos de la ciudad. En cambio, si nuestra muestra estuviera compuesta, por ejemplo, sólo por mujeres católicas menores a 35 años, no podría constituirse como una muestra representativa de la totalidad de la población de la ciudad de La Plata.

En síntesis, la muestra en investigación científica debe ser representativa, en el sentido de exhibir adecuadamente las propiedades distintivas y patrones fundamentales del universo

investigado. En los apartados siguientes, se desarrollan con mayor detenimiento los tipos muestreo probabilísticos y no probabilísticos.

Muestreos probabilísticos

En el muestreo probabilístico, todos los elementos de la población tienen una probabilidad conocida y distinta de cero de ser seleccionados para la muestra y el criterio de selección es aleatorio. De esta manera, no será la persona encargada de realizar la investigación quien decidirá qué casos sí y qué casos no constituirán parte de la muestra, sino que la selección de la misma será producto del azar. Esto permite estimar la representatividad de la muestra y el error muestral. Es el más recomendable para estudios cuantitativos que buscan resultados generalizables. No obstante, su realización es costosa y dificultosa puesto que requiere amplios conocimientos en estadística y además contar con un listado completo de la población de estudio, documento que se denomina *marco muestral*. Para ello, existen personas especializadas en el diseño de muestras probabilísticas.

Se pueden encontrar cuatro tipos de muestreo probabilístico:

En el *muestreo aleatorio simple*, los casos se seleccionan de modo que todos tengan la misma probabilidad de ser elegidos. Esto se logra asignando un número o código a cada unidad poblacional del marco muestral y extrayendo aleatoriamente la cantidad requerida para la muestra. Aunque conceptualmente es sencillo, presenta limitaciones prácticas en poblaciones muy grandes.

Este muestreo permite estimar la probabilidad de que la muestra sea representativa del universo. Es decir, adjudicar un valor a esa probabilidad con cierto margen de error y nivel de confianza.

Para realizar un muestreo aleatorio simple se deben seguir los siguientes pasos:

1) Asignar un número a cada caso de la población.

2) Utilizar algún método mecánico o aleatorio (como bolas en una bolsa, tablas de números aleatorios, generación de números aleatorios con una calculadora o computadora, softwares especializados, etc.) para seleccionar los casos necesarios y completar el tamaño de muestra requerido. Supongamos que tenemos una población de 1000 unidades y queremos una muestra de 100 casos. Se colocan 1000 bolillas en un bolillero y las primeras 100 que salgan formarán parte de la muestra.

Sin embargo, este procedimiento tiene limitaciones cuando la población es muy heterogénea. Además, es necesario contar con un listado completo de la población para poder aplicarlo correctamente.

Es importante distinguir entre estimar la probabilidad de ocurrencia de algo y asegurar que ocurrirá. Si un evento tiene 99 % de probabilidad, es muy probable, pero existe 1 % de chances

de que no suceda. La probabilidad señala cuán posible es que ocurra, no garantiza que efectivamente ocurrirá.

El concepto de "error muestral" suele generar confusión. Cuando un estudio reporta por ejemplo que el margen de error es de 5 % con 95 % de confianza, eso no significa que la muestra tenga un 5 % de probabilidad de estar errada. El margen indica la amplitud máxima que la estimación de la muestra podría diferir del verdadero valor poblacional. Es decir, establece un rango en el cual esperamos que se encuentre el parámetro real, considerando el tamaño de muestra y nivel de confianza.

A modo de ejemplo, una encuesta política muestra que el 60 % de los votantes prefieren al Candidato A, con margen de error de +/- 5 %. Esto significa que el porcentaje real en la población total podría estar entre 55 % y 65 % (60% ± 5%). Pero no quiere decir que haya 5 % de probabilidades de que la muestra esté equivocada.

De hecho, la muestra tiene sólo dos opciones: o representa adecuadamente a la población o no lo hace. No hay una probabilidad intermedia de "estar errada", como sugiere ese 5 %. Una vez extraída la muestra, o acierta el valor poblacional dentro del margen de error o no lo hace.

En resumen, el margen de error no se debe interpretar como la probabilidad de equivocación de la estimación muestral, sino como la amplitud máxima esperada entre la estimación y el verdadero valor poblacional. La muestra es representativa o no lo es.

Por lo tanto, el muestreo aleatorio simple posibilita estimar la probabilidad de obtener una muestra representativa, lo cual es distinto a asegurar que lo será. Permite calcular el nivel de confianza sobre la representatividad de la muestra extraída.

El *muestreo aleatorio sistemático* es una técnica de muestreo probabilístico en la cual se selecciona a les participantes del estudio de una lista ordenada de acuerdo a un intervalo fijo de selección. Por ejemplo, si queremos obtener una muestra de 100 estudiantes de una población de 2000 utilizando un muestreo sistemático, primero enumeraríamos a toda la población de 1 a 2000. Luego, calculamos el intervalo de muestreo dividiendo el tamaño de la población (2000) entre el tamaño deseado de la muestra (100), dando un intervalo de 20. Después se selecciona aleatoriamente un número entre 1 y 20 como punto de partida, supongamos que sale el número 12. A partir de ahí, el 12 será el primer caso de nuestra muestra y luego cada vigésimo estudiante en la lista sería elegido para formar parte de la misma (es decir, el estudiante 32, 52, 72... hasta llegar a 100 estudiantes).

La ventaja de este tipo de muestreo es que es más fácil y rápido de implementar que otros métodos probabilísticos. Solo requiere definir un intervalo y aplicarlo sistemáticamente. Si la lista de la población está ordenada aleatoriamente, la muestra será muy representativa y es libre de sesgos de les investigadores en la selección. Su desventaja es que si la lista de la población tiene algún orden o periodicidad puede introducir un sesgo en la muestra. Por ejemplo, imaginemos que tenemos una lista de 500 estudiantes universitaries. La lista está ordenada por aquelles que trabajan tiempo completo primero, luego las personas que trabajan medio tiempo, y finalmente aquellas que no trabajan.

Queremos una muestra de 100 estudiantes, aplicando un intervalo de muestreo sistemático de 5 (500/100=5). Al seleccionar cada 5° estudiante en la lista, es muy probable que la mayoría de les seleccionados sean estudiantes que trabajan tiempo completo, ya que elles están al inicio de la lista. Habría muy poques o ningune estudiante que no trabaja, porque elles se encuentran al final.

En este caso, el muestreo sistemático terminaría por sesgar la muestra, sobrerrepresentando a les estudiantes con empleos de tiempo completo y subrepresentando o excluyendo a quienes no trabajan. La periodicidad de la lista (ordenada por situación laboral) genera el sesgo. Por ello, para estudiar estudiantes con diferente situación laboral sería mejor utilizar otro tipo de muestreo probabilístico como el aleatorio simple, que no refleje el orden original.

Otra desventaja es que no permite estratificar la muestra: No podemos dividir o clasificar a la población en subgrupos (por género, edad, etc.) y garantizar cierto número de cada uno en la muestra. Simplemente toma los elementos según el intervalo sin importar a qué subgrupo pertenezcan. Otro problema es que es sensible a pérdidas de participantes: Si en la lista original la persona en la posición 12 rechaza participar, esto rompe el intervalo sistemático de 20. Entonces la muestra ya no sería equidistante y representativa. Lo mismo si las personas 52 y 72 también rechazan.

Por lo tanto, el muestreo sistemático es una técnica eficiente cuando se cuenta con un marco muestral completo y ordenado aleatoriamente. Hay que tener cuidado con posibles sesgos derivados de patrones en la lista original. Se recomienda considerar otros métodos si se necesita estratificar o si hay riesgo de no respuesta.

En el *muestreo aleatorio estratificado* se procede dividiendo a la población en subgrupos y así se obtienen muestras aleatorias de cada uno, logrando representatividad de los grupos. En el muestreo aleatorio estratificado, se divide a la población en diferentes grupos homogéneos llamados estratos que se seleccionan en función de una característica relevante para el estudio, como la carrera que estudia, el lugar de residencia, el género o el estado civil, entre otros.

El objetivo de este tipo de muestreo es asegurarse de que todos los estratos de interés estén representados adecuadamente en la muestra. Esto es importante porque permite obtener una visión más completa y precisa de la población en estudio, ya que se consideran las características específicas de cada estrato.

Dentro de cada estrato, se puede aplicar el muestreo aleatorio simple o el muestreo sistemático para seleccionar los elementos concretos que formarán parte de la muestra. Esto implica seleccionar aleatoriamente casos de cada estrato de manera independiente.

Supongamos que la población está compuesta por 500 estudiantes de Trabajo Social y queremos obtener una muestra de 100 estudiantes.

Primero dividimos a la población en estratos según dos características:

1. Situación laboral:
   - Trabajan (300 estudiantes)
   - No trabajan (200 estudiantes)

2. Turno que cursan:

- Mañana (250 estudiantes)
- Tarde (250 estudiantes)

De este modo tenemos 4 estratos:

- Trabajan y cursan mañana (150 estudiantes)
- Trabajan y cursan tarde (150 estudiantes)
- No trabajan y cursan mañana (100 estudiantes)
- No trabajan y cursan tarde (100 estudiantes)

Como el 60 % de la población trabaja, el 60 % de la muestra (60 estudiantes) se selecciona al azar entre los dos estratos de estudiantes que trabajan (30 de cada uno).

El 40 % restante (40 estudiantes) se selecciona entre los dos estratos de estudiantes que no trabajan (20 de cada uno).

De este modo todos los subgrupos están representados en la muestra final de 100 en su proporción real en la población de 500 estudiantes.

El muestreo aleatorio estratificado tiene la ventaja de reducir el error muestral para un tamaño dado de la muestra, ya que se considera la variabilidad dentro de cada estrato. Al seleccionar una muestra representativa de cada estrato, se obtiene una estimación más precisa de las características de interés en la población total.

Sin embargo, el muestreo aleatorio estratificado puede presentar dificultades, especialmente en la etapa de identificación y selección de los estratos. Requiere un conocimiento detallado de la población y de las características relevantes para la estratificación. Además, puede resultar más complejo y requerir más recursos en comparación con otros métodos de muestreo.

El muestreo aleatorio estratificado busca superar las dificultades de otros métodos de muestreo, al considerar diferentes estratos homogéneos dentro de la población. Esto permite obtener una muestra representativa de cada estrato y obtener estimaciones más precisas de las características de interés en la población total.

El *muestreo aleatorio por conglomerados* consiste en dividir a la población en grupos o conglomerados (por ejemplo, por región geográfica[10], escuelas, barrios, etc.), y luego seleccionar aleatoriamente algunos conglomerados de los cuales se recolectarán todos los datos.

Por ejemplo, queremos hacer un estudio sobre métodos de enseñanza en cursos de metodología de la investigación en todo el país.

La población objetivo son todos los cursos de metodología en todas las carreras de grado. Dado que son muchísimos no es viable estudiarlos a todos.

Usaremos entonces un muestreo por conglomerados:

1) Primero se seleccionan aleatoriamente 15 universidades públicas y privadas del país.

2) Dentro de cada universidad se eligen 3 facultades al azar (por ejemplo, Cs. Económicas, Humanidades, Ingeniería).

3) Por último, dentro de cada facultad se selecciona 1 curso de metodología de la carrera.

[10] El muestreo por conglomerados es ampliamente utilizado en nuestro país por la Encuesta Permanente de Hogares.

Así obtenemos una muestra de 15 universidades x 3 facultades x 1 curso = 45 cursos de metodología.

Los conglomerados están compuestos por: las universidades, las facultades y, finalmente, los cursos específicos. Todos fueron elegidos al azar, pero garantizando la representación nacional. Una vez seleccionada la muestra, se recopila información de todos los cursos mediante encuestas a docentes y estudiantes. Los resultados encontrados se generalizarían a nivel poblacional dentro de ciertos márgenes de error.

Sus ventajas se asocian con que son más económicos y fáciles que hacer una muestra aleatoria de toda la población y se obtienen datos de subgrupos específicos (los cursos de metodología). No obstante, entre sus desventajas se encuentran que son menos precisos que el muestreo aleatorio simple, pudiendo excluir segmentos enteros de la población. Además, las unidades dentro de un conglomerado suelen ser muy parecidas entre sí. Por ejemplo, si sólo se eligen grandes ciudades, se excluye a quienes viven en áreas rurales. O si se eligen escuelas privadas, no representa a estudiantes de escuelas públicas.

En conclusión, el muestreo por conglomerados es útil para estudios exploratorios o cuando no existe un marco muestral completo, pero puede sesgar la muestra si los conglomerados no representan la diversidad de la población.

Cálculo del tamaño muestral en muestreos probabilísticos

El tamaño muestral influye sobre el error aleatorio. Un tamaño muy pequeño genera imprecisión, mientras uno muy grande implica costos y recursos innecesarios.

Para muestreos probabilísticos existen fórmulas que consideran: el nivel de confianza deseado, el error muestral aceptable, la variabilidad de la población y el tipo de muestreo. El investigador debe determinar estos parámetros.

El tamaño de muestra requerido en un muestreo aleatorio depende principalmente de tres factores:

1. La variabilidad del fenómeno estudiado: mientras más heterogénea o dispersa sea la población en la característica de interés, se necesita una muestra más grande para obtener la misma precisión.
2. El margen de error deseado: cuanto menor margen de error se busque, mayor debe ser el tamaño de muestra. Un margen de error menor implica mayor precisión.
3. El nivel de confianza: este indica la probabilidad de que el rango de error de la muestra capture el verdadero valor poblacional. Un nivel de confianza mayor requiere un tamaño de muestra mayor.

El tamaño del universo o población total no influye directamente en la determinación del tamaño muestral, sino estos tres factores.

Por ejemplo, para estimar con un 5 % de margen de error y 95 % de confianza la proporción de votantes a favor de un candidato, necesitaríamos una muestra distinta si la elección es muy

reñida (50 % votaría por él) que si su apoyo es muy bajo (5 % votaría por él). En el primer caso la población es más heterogénea respecto a esa preferencia.

En síntesis, el tamaño de una muestra probabilística depende de la variabilidad del fenómeno, el nivel de precisión y confianza deseados, no del tamaño total de la población que queremos inferir. Esto se aplica a todas las variantes de muestreo probabilístico.

Muestreos no probabilísticos

En los muestreos no probabilísticos la elección de los casos no depende de la probabilidad sino de otros criterios deliberados del investigador. No permite un cálculo preciso del error muestral y son útiles cuando no se dispone de marco muestral.

Tal como plantea Galtung (1966), aunque la exploración de un fenómeno mediante muestreos no probabilísticos es siempre tentativa, esto no significa que estas muestras no puedan ser válidas y representativas bajo determinadas condiciones. Según este autor, las muestras no probabilísticas pueden constituir "buenas muestras" cuando:

- El fenómeno bajo estudio tiene muy poca variabilidad, es decir, es bastante homogéneo y regular en la población. En este caso una pequeña muestra intencional podría describirlo adecuadamente.
- Cuando se conoce bien la variabilidad del fenómeno por estudios previos o el investigador tiene criterios sólidos para identificar las principales variaciones.
- Cuando el objetivo es explorar en profundidad los aspectos variables de un fenómeno. Aquí una muestra pequeña pero estratégica puede ser útil, especialmente en etapas iniciales.

En síntesis, las muestras no probabilísticas pueden resultar válidas, representativas y aportar conocimiento relevante sobre un fenómeno cuando se dan estas condiciones contextuales específicas.

Entre los principales tipos de muestreo no probabilístico, podemos mencionar:

Muestreo por conveniencia: se seleccionan casos accesibles que acepten participar de la investigación. Supongamos que un investigador quiere estudiar los hábitos de lectura en estudiantes universitarios de su ciudad. Como muestreo por conveniencia, el investigador podría seleccionar para su estudio a los 100 estudiantes de la carrera en donde da clases. Los selecciona por ser accesibles y fáciles de contactar al ser alumnos de la misma institución.

La ventaja es que es muy rápido y práctico obtener la muestra, al ser estudiantes a los que tiene fácil acceso. La desventaja es que sólo representan a una pequeña porción (su facultad) y carreras afines, no a todos los estudiantes universitarios. Por lo tanto, aunque los resultados pueden explorar los hábitos de ese grupo específico de 100 alumnos, no se podrían generalizar a toda la población universitaria de la ciudad ni hacer inferencias por el sesgo en la selección de la muestra.

El muestreo por conveniencia selecciona los casos más fáciles de obtener para el investigador, pero no representan necesariamente a toda la población, limitando la posibilidad de generalizar los resultados.

Muestreo por cuotas: es una técnica de muestreo no probabilístico en la que se divide a la población en segmentos o cuotas según sus características relevantes (género, edad, locación, etc.) y luego se seleccionan participantes de cada cuota para conformar una muestra con la misma distribución de esas características que la población total. El principio de división por grupos lo relaciona con el muestreo estratificado, pero la principal diferencia entre ambos es que el muestreo por cuotas no es aleatorio.

Si tenemos una población de 1000 estudiantes universitarios, queremos una muestra de 100 estudiantes mediante cuotas por:

- Carrera: 50 % de Sociales, 30 % de Económicas, 20 % de Ingeniería.
- Género: 50 % mujeres, 50 % varones
- Edad: 40 % de 18-22 años, 40 % de 23-25 años, 20 % de 26-30 años.

Entonces la muestra se conformaría con:

- 50 estudiantes de Sociales (25 mujeres, 25 varones según las edades)
- 30 de Económicas (15 mujeres, 15 varones según edades)
- 20 de Ingeniería (10 mujeres, 10 varones según edades)

La muestra de 100 tiene una distribución proporcional en cuotas de carrera, género y edad igual a la población total. Pero al no ser aleatoria, no podemos inferir ni generalizar resultados. Sólo válido para ese grupo específico.

Bola de nieve: es un procedimiento utilizado en general en estudios cualitativos para obtener información de personas o grupos que comparten ciertas características o experiencias específicas. En esta técnica, el proceso de selección de participantes se inicia con un pequeño número de sujetos que cumplen con los criterios de estudio. Luego, estos participantes iniciales ayudan a identificar y reclutar a otros participantes que también puedan ser relevantes para el estudio, creando así una "bola de nieve" de participantes reclutados a medida que el estudio avanza.

Este muestreo es especialmente útil cuando se investigan grupos o poblaciones que son difíciles de alcanzar o identificar a través de métodos de muestreo más tradicionales. Es eficaz para estudiar comunidades cerradas, subculturas, grupos minoritarios y situaciones donde no existe una lista preexistente de miembros para seleccionar.

Si en una investigación el objetivo es explorar las experiencias de acoso callejero que viven las mujeres en la ciudad de La Plata, al no haber un registro oficial, se decide utilizar un muestreo bola de nieve:

1) Se contacta a 5 mujeres que han relatado situaciones de acoso en redes sociales.

2) A cada una se le pide que recomiende a 2 mujeres de su círculo que también hayan vivido episodios de acoso callejero en la ciudad.

3) Esas 10 mujeres referidas (2da ronda) son contactadas y entrevistadas sobre sus experiencias. También se les pide 2 referidas cada una.

4) El proceso se repite hasta llegar a 50 mujeres participantes en la muestra final.

Si bien este procedimiento permite encontrar casos y obtener testimonios de un grupo de mujeres afectadas, la muestra se va homogenizando y autocorrelacionando. No representa necesariamente a todas las mujeres de la ciudad.

Los resultados servirán para explorar en profundidad las experiencias de acoso en ese grupo, pero no para hacer generalizaciones sobre la población. Se requeriría un muestreo probabilístico aleatorio para tal fin.

Si bien la representatividad es limitada, los muestreos no probabilísticos son valiosos para estudios cualitativos y exploratorios. Es necesario señalar, antes de terminar este apartado, que existen otros tipos de muestreo no probabilísticos, como el muestreo teórico que no desarrollaremos en este capítulo.

Conclusiones

En este capítulo, hemos trabajado las nociones de población, muestra y muestreo, así como los diferentes tipos de muestreo que encontramos en investigación. Es importante remarcar que en investigación cuantitativa raras veces se estudia a la población total, sino que se procede a relevar información sobre un subconjunto, la muestra. Esta debe ser representativa de la población para lograr resultados válidos y generalizables. El muestreo probabilístico es el más recomendable para estudios cuantitativos y permite la inferencia estadística.

El muestreo no probabilístico es pertinente para investigaciones cualitativas y también en algunas cuantitativas, sobre todo cuando no existe un marco muestral claro. Complementar métodos probabilísticos y no probabilísticos equilibra las limitaciones de cada uno. No existe un método universalmente mejor. La elección dependerá del planteo del problema y las circunstancias concretas.

Les investigadores deben evaluar cuidadosamente las opciones de muestreo para seleccionar aquel diseño que maximice la representatividad, viabilidad y economía del estudio. Las decisiones tomadas tendrán implicaciones significativas sobre la calidad de los datos recolectados y la validez de los resultados obtenidos.

Referencias

Archenti, N., Marradi, A. y Piovani, J. I. (2007). *Metodología de las Ciencias Sociales.* Buenos Aires: EMECE.

Becker, H. (1998). *Trucos del oficio: Cómo conducir su investigación en ciencias sociales*. Buenos Aires: Siglo XXI editores.

Baranger, D. (1992). *Construcción y análisis de datos. Introducción al uso de técnicas cuantitativas en la investigación socia*l. Posadas: Editorial Universitaria.

Dimarco, M. y Martello, V. (2016). Cap.5 “La selección de los casos. Sobre quiénes/qué vamos a relevar la información” en: Cuenca, A y Lozano, S. *La enseñanza de la investigación. Diálogo entre la teoría y el oficio del investigador en Trabajo Social*. La Plata: Edulp.

Flick, U. (2004). *Introducción a la investigación cualitativa.* Madrid: Morata.

Galtung, J. (1966). *Teoría y métodos de la investigación social*. Volumen 1. Buenos Aires: Editorial Eudeba.

Hernández Sampieri, R. (2010). Cap.8 “Selección de la muestra” en: *Metodología de Investigación*. Sexta edición. México: McGrawHill Education.

Polit, D y Hungler, B. (1997). *Investigación científica en ciencias de la salud*. México: McGraw Hill Interamericana.

Samaja, J. (1993). *Epistemología y Metodología: Elementos para una teoría de la investigación científica*. Buenos Aires: Eudeba.