

La Muestra

La manera más común de iniciar una investigación de campo es con la estimación de una muestra representativa de la población. Aunque no es un requisito imprescindible, es una realidad que la mayoría de las investigaciones no evalúan a todos los miembros de una población (i.e., censo). Realizar un censo es demandante de recursos económicos y de tiempo. Solo en situaciones especiales se hacen censos, como por ejemplo el censo nacional de población y vivienda que se realiza cada 5 años en nuestro país, o cuando la población es pequeña.

La muestra es una parte representativa de la población bajo estudio y su importancia radica en que es indicada en las siguientes situaciones:

1. Es una opción viable cuando no se puede realizar un censo por limitaciones en los recursos y el tiempo.
2. Se utiliza cuando se tiene una población infinita, de tal forma que es imposible revisar a cada integrante de la población. Por ejemplo, un proceso de producción continua.
3. Es la opción indicada cuando se desconoce la población total o no se tiene una ubicación exacta de sus integrantes.
4. Cuando las pruebas que se realizan implican la destrucción de las unidades muestrales. Por ejemplo una degustación de un alimento.
5. Cuando se requieren estimaciones precisas y confiables. En una investigación con muestreo se estima y controla el margen de error.

Utilizar una muestra en lugar de investigar a la población total también tiene sus inconvenientes, como por ejemplo el desconocimiento de la ubicación física de los

La Muestra

elementos de la población o la obtención de resultados erróneos si no se sigue un procedimiento riguroso en la selección de la muestra.

Algunos conceptos con los que debemos relacionarnos cuando hacemos investigación basada en una muestra son los siguientes:

Población: se refiere al conjunto de elementos con características similares que queremos estudiar. Por ejemplo, los habitantes de la ciudad de Saltillo, los empleados de una organización, los productos pertenecientes a un lote determinado de producción. También se le conoce como universo. Las poblaciones deben delimitarse en función de su contenido, lugar y tiempo. Por ejemplo, los niños de primaria que se hayan inscrito de hace 6 años a la fecha en la ciudad de Saltillo.

Muestra: es el subconjunto de la población que vamos a evaluar. Este subconjunto tiene las mismas características que el resto de los elementos de la población, es decir, es representativa de ella.

Marco muestral: es un listado que contiene a todos los elementos que integran a una población y que están disponibles para su selección. Por ejemplo, un listado de los restaurantes de comida rápida de la ciudad.

Muestreo: es el procedimiento estadístico que se sigue para estimar y seleccionar la muestra.

La Muestra

Unidad de análisis: denominada también como caso o elemento, identifica a aquellos elementos sobre los que vamos a obtener la información, y pueden ser sujetos, objetos, sucesos o fenómenos y organizaciones.

Las ventajas del muestreo sobre el censo son bastantes:

- Es más económico y rápido.
- Permite obtener datos con más detalle.
- Requiere menos personal.
- Es más controlable.
- Sus resultados se pueden extrapolar a toda la población.

Existen dos tipos de muestras que se diferencian por el procedimiento que siguen para estimarlas:

1. Muestras no probabilísticas: son aquellas en que los elementos se seleccionan atendiendo a criterios subjetivos y personales. Es decir, la elección de los elementos no depende de la probabilidad sino de los requerimientos específicos o características de la población. Aunque los resultados de este tipo de muestras no se pueden generalizar a toda la población, sí tienen las siguientes ventajas:
 - a. El investigador selecciona los casos de interés.

La Muestra

- b. Es de gran utilidad en estudios exploratorios.
- c. Tiene un costo más bajo que las muestras probabilísticas.
- d. Los elementos son fáciles de seleccionar ya que no es obligatorio que se escojan de un marco muestral.

Las principales técnicas para las muestras no probabilísticas son:

- a. Muestreo por conveniencia: la elección de los elementos es de acuerdo a la intención o conveniencia del investigador.
 - b. Muestreo de juicio o de experto: los elementos son seleccionados por una persona experta en el tema.
 - c. Muestreo por cuota: en esta técnica se elige una cantidad determinada de elementos para la muestra. Se basa en el conocimiento de ciertos estratos en la población pero los elementos en cada estrato no se eligen de manera aleatoria.
 - d. Bola de nieve: los elementos se seleccionan en base a información suministrada por un elemento ya identificado en la población.
2. Muestras probabilísticas: las unidades se seleccionan atendiendo a criterios científicos, y todos los elementos de la población tienen una probabilidad determinada de ser parte de la muestra y no por el criterio de los investigadores. Aunque requiere de un marco muestral de dónde seleccionar los elementos y es más costoso que el no probabilístico, tiene las siguientes ventajas:

La Muestra

- a. Se obtienen muestras representativas.
- b. Los resultados son generalizables a toda la población.
- c. Cada elemento tiene una posibilidad mayor a cero de ser seleccionado para la muestra.

Las técnicas de muestreo probabilístico son las siguientes:

- a. Muestreo aleatorio simple: cada elemento de la población tiene la misma probabilidad de ser seleccionado. Las técnicas más utilizadas son la tómbola y la tabla de números aleatorios.
- b. Muestreo aleatorio sistemático: las unidades se seleccionan sistemáticamente a partir de un número aleatorio. Se estima un número para saber cada cuantos elementos de una lista se va a seleccionar a un participante ($\text{Brinco} = \text{Población} / \text{tamaño de muestra}$). Por ejemplo $100/10=10$. Cada décimo elemento de la lista será parte de la muestra.
- c. Muestreo estratificado: la población se divide en grupos que tienen características similares entre sus miembros, pero diferentes a los miembros de otros grupos. Las dos técnicas más utilizadas son el muestreo estratificado simple y el proporcional. En el simple, a cada estrato le corresponde el mismo número de elementos. En el proporcional, la distribución se hace en base al tamaño de la población de cada estrato. Por ejemplo, si un estrato de la población tiene 80 personas y otro 20 personas, y si queremos una muestra del 10% con el muestreo estratificado simple elegiríamos 5 personas de cada grupo y con el proporcional 8 y 2.

La Muestra

- d. Muestreo por conglomerado: similar al anterior, pero en esta técnica se hace una selección aleatoria de los grupos que van a participar. Y los grupos son diferentes al interior pero iguales a los otros grupos. Por ejemplo, grupos de estudiantes del turno matutino y vespertino de una escuela.

Por último, veremos cómo se determina el tamaño que va a tener la muestra probabilística (n). Existen varias fórmulas que dependen del tipo de muestreo que estamos utilizando.

En este curso solo veremos la fórmula usando proporciones para poblaciones infinitas y finitas:

El primer paso es determinar el nivel de confianza con que se quiere trabajar (z). El nivel de confianza es la probabilidad de que el parámetro de la población que estamos estimando se encuentre dentro del intervalo de confianza. Se recomienda revisar las tablas para el nivel de confianza en un libro de estadística o de investigación de mercados. Pero los dos valores más comunes son $z=1.96$ para un 95% de confianza y $z= 1.65$ para el 90% de confianza.

El segundo paso es estimar las características de que ocurra un fenómeno determinado (p) o de que no ocurra (q). La suma debe dar uno. Un estándar es utilizar .5 y .5.

El tercer paso es determinar el grado de error máximo aceptable en los resultados. Generalmente se establece en un 5% pero puede llegar hasta un 10%.

La Muestra

El cuarto paso es aplicar una de las siguientes fórmulas:

Para la población infinita

$$n = \frac{p*q}{e^2}$$

Donde:

n: Tamaño de la muestra.

p: Proporción de la población que sí cumple con un criterio determinado.

q: Proporción de la población que no cumple con un criterio determinado.

e: Error.

Para la población finita

$$n = \frac{Z^2 p^* q^* N}{N e^2 + Z^2 p^* q}$$

Donde:

N: Universo o población.

La Muestra

Veamos un ejemplo. Determina el tamaño de muestra considerando una proporción de hombres del 45% y que observe un 95% de confianza y permita un error máximo del 5%. El tamaño de la población es de 3,000 individuos.

$$\frac{(1.96)^2 (.45) (.55) 3,000}{3,000 (.05)^2 + (1.96)^2 (.45) (.55)}$$

El tamaño de muestra es de 337 individuos.