

SELECCIÓN DE MUESTRA

Para seleccionar una muestra, lo primero que hay que hacer es definir la unidad de análisis (personas, organizaciones, periódicos, comunidades, situaciones, eventos, etc.). El sobre qué o quienes se van a recolectar datos depende del planteamiento del problema a investigar y de los alcances del estudio.

¿Sobre qué o quiénes se recolectarán datos?

Para el proceso cuantitativo la muestra es un subgrupo de la población de interés (sobre el cual se recolectarán datos, y que tiene que definirse o delimitarse de antemano con precisión), éste deberá ser representativo de la población.

¿Cómo se delimita una población?

Una vez que se ha definido cuál será la unidad de análisis, se procede a delimitar la población que va a ser estudiada y sobre la cual se pretende generalizar los resultados. Según Sellitz, una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones. Es preferible establecer con claridad las características de la población, con la finalidad de delimitar cuáles serán los parámetros muestrales.

La delimitación de las características de la población no sólo depende de los objetivos del estudio, sino de otras razones prácticas. Un estudio no será mejor por tener una población más grande; la calidad de un trabajo investigativo estriba en delimitar claramente la población con base en el planteamiento del problema.

Las poblaciones deben situarse claramente en torno a sus características de contenido, de lugar y en el tiempo.

¿Cómo seleccionar la muestra?

Básicamente categorizamos las muestras en dos grandes ramas: las "muestras no probabilísticas" y las "muestras probabilísticas". En estas últimas todos los elementos de la población tienen misma posibilidad de ser escogidos y se obtienen definiendo las características de la población y el tamaño de la muestra, y por medio de una selección aleatoria o mecánica de las unidades de análisis.

En las muestras no probabilísticas, la elección de los elementos no depende de la probabilidad, sino de causas relacionadas con las características de la investigación o de quien hace la muestra. Aquí el procedimiento no es mecánico, ni con base en fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de una persona o de un grupo de personas y, desde luego, las muestras seleccionadas obedecen a otros criterios de investigación.

Elegir entre una muestra probabilística o una no probabilística depende de los objetivos del estudio, del esquema de investigación y de la contribución que se piensa hacer con ella.

¿Cómo se selecciona una muestra probabilística?

El autor Kish, menciona que la elección entre la muestra probabilística y la no probabilística se determina con base en el planteamiento del problema, las hipótesis, el diseño de investigación y el alcance de sus contribuciones. Las muestras probabilísticas tienen muchas ventajas, quizá la principal sea que puede medirse el tamaño del error en nuestras predicciones. Se dice incluso que el principal objetivo en el diseño de una muestra probabilística es reducir al mínimo este error, al que se le llama error estándar.

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

Muestra probabilística estratificada

Cuando no basta que cada uno de los elementos muéstrales tengan la misma probabilidad de ser escogidos, sino que además es necesario estratificar la muestra en relación con estratos o categorías que se presentan en la población, y que además son relevantes para los objetivos del estudio, se diseña una muestra probabilística estratificada. Lo que aquí se hace es dividir a la población en subpoblaciones o estratos, y se selecciona una muestra para cada estrato.

Muestreo probabilístico por racimos

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

Muestra probabilística por racimos: subgrupo en el que las unidades de análisis se encuentran encapsuladas en determinados lugares físicos.

Listados y otros marcos muéstrales

Las muestras probabilísticas requieren la determinación del tamaño de la muestra y de un proceso de selección aleatoria que asegure que todos los elementos de la población tengan

la misma probabilidad de ser elegidos. Todo esto lo hemos visto, aunque nos falta exponer sobre algo esencial que precede a la selección de una muestra: el marco muestra! Éste constituye un marco de referencia que nos permite identificar físicamente los elementos de la población, la posibilidad de enumerarlos y, por ende, de proceder a la selección de los elementos muestrales.

Tamaño óptimo de una muestra

Las muestras probabilísticas requieren dos procedimientos básicos:

- 1) la determinación del tamaño de la muestra;
- 2) la selección aleatoria de los elementos muestrales.

Precisar adecuadamente el tamaño de la muestra puede tomarse muy complejo, esto depende del problema de investigación y la población a estudiar. Para el alumno y el lector en general, será muy útil comparar qué tamaño de muestra han empleado otros investigadores, a la luz de la revisión de la literatura.

Las muestras no probabilísticas

Las muestras no probabilísticas, también llamadas muestras dirigidas, suponen un procedimiento de selección informal. Se utilizan en muchas investigaciones cuantitativas y cualitativas.

Ejercicio 1

- Si la señora López compra una de las casas anunciadas para su venta en un diario de TGZ, T es el evento de que la casa tiene tres o más baños, U es el evento de que tiene una chimenea, V es el evento de que cuesta más de \$ 100 mil pesos y W es el evento de que es nueva.
- Describa (con palabras) cada uno de los siguientes eventos:

T' = Casas con > 3 baños

U' = No tienen chimenea
o 2 o más Chimeneas

V' = Casas de menos de
100 mil pesos

W' = Casas que no son
nuevas

$T \cap U$ = Casas de 3 o más
baños que tienen
una Chimenea

$T \cap V$ = Casas de 3 o más
baños que valen más de
100 mil pesos

$U' \cap V$ = Sin Chimenea y vale
menos de 100 mil pesos

$V \cup W$ = Unión de todas
las casas que valen más de
100 mil pesos y todas las
casas nuevas

Unión de casas de menos de
100 mil pesos y las casas
nuevas

$V' \cup W$ =

$T \cup U$ = Unión de las casas de
3 o más baños y las
que tienen una Chimenea

$T \cup V$ = Unión de todas las casas
que tienen 3 o más baños
y todas las que cuestan más de
100 mil pesos

$V \cap W$ = Casas que valen
más de 100 mil y son
nuevas

Ejercicio 2

- Un dado está arreglado de manera que cada número impar tiene el doble de probabilidad de ocurrir que un número par. Encuentra $P(B)$, donde B es el evento que un número mayor que 3 ocurra en un solo tiro del dado.

- Espacio muestral

Espacio muestral

$$S = \{1, 2, 3, 4, 5, 6\}$$

- Subconjunto B

Subconjunto B

$$B = \{4, 5, 6\}$$

- Probabilidad

- Si x es la probabilidad que ocurra un número par, $2x$ sería la probabilidad que ocurra un número impar.

- Entonces, encontramos que: $2x(1)$ + x + $2x(3)$ + x + $2x(5)$ + $x = 1$

- Esto se debe al postulado 2

- La $P(B)$ sería: $\frac{4}{9} = .4444 = 44.44\%$

1, 1, 2, 3, 3, 4, 5, 5, 6

$n=9$ lados

$$1 = \frac{2}{9} \quad 2 = \frac{1}{9} \quad 3 = \frac{2}{9} \quad 4 = \frac{1}{9} \quad 5 = \frac{2}{9} \quad 6 = \frac{1}{9}$$

$$\frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$$

Ejercicio 3

- Entre los ocho automóviles que un vendedor tiene en su sala de exhibición, el automóvil 1 es nuevo, tiene aire acondicionado, dirección hidráulica y asientos de cubo; el vehículo 2, tiene un año de uso, tiene aire acondicionado, pero no tiene ni dirección hidráulica ni asientos de cubo; el automóvil 3, tiene dos años de uso, tiene aire acondicionado y dirección hidráulica, pero no tiene asientos de cubo; la unidad 4 tiene tres años de uso, tiene aire acondicionado pero ni tiene ni dirección hidráulica ni asientos de cubo; el vehículo 5 es nuevo, no tiene aire acondicionado, ni dirección hidráulica ni asientos de cubo; el automóvil 6 tiene un año de uso, tiene dirección hidráulica, pero no tiene ni aire acondicionado ni asientos de cubo; el vehículo 7 tiene dos años de uso, no tiene aire acondicionado, ni dirección hidráulica ni asientos de cubo; y la unidad 8 tiene tres años de uso, no tiene aire acondicionado, pero tiene dirección hidráulica así como asientos de cubo.
- Si un cliente compra uno de estos automóviles y el evento de que compre un vehículo nuevo, por ejemplo, se representa con el conjunto (Automóvil 1, automóvil 5), indique en forma similar los conjuntos que representan los eventos de que:
 - A) Se decida por un automóvil sin aire acondicionado (Auto 5, Auto 6, Auto 7, Auto 8)
 - B) Escoja una unidad sin dirección hidráulica (Auto 2, Auto 4, Auto 5, Auto 7)
 - C) Escoja un vehículo con asientos de cubo (Auto 1, Auto 8)
 - D) Escoja un automóvil que tenga dos o tres años de uso (Auto 3, Auto 4, Auto 7, Auto 8)

a) AUTO 1, AUTO 5

b) AUTO 2, AUTO 4, AUTO 5, AUTO 7

Ejercicio 4

- Se lanza una moneda al aire una vez. Entonces si cae cara, se tira un dado una vez; si cae cruz, el dado se tira dos veces. Utilice la notación en la que $(H,2)$, por ejemplo, denota el evento de que la moneda cae cara y entonces el dado cae en 2, y $(T,2,1)$ denota el evento de que la moneda cae cruz y el dado se tira dos veces seguidas. Para enumerar:

- A) Los elementos del espacio muestral $S = \{x_1, x_2, x_3, x_4, x_5, x_6\}$
- B) Los elementos de S que corresponden al evento A de que caiga exactamente una cara $S = \{x_1, x_2, x_3, x_4\}$
- C) Los elementos de S que corresponden al evento B de que caiga un número mayor que 4 $S = \{x_5, x_6\}$

$x = \text{cara}$

$y = \text{cruz}$

A)

$$S = \{x_1, x_2, x_3, x_4, x_5, x_6, y_{11}, y_{12}, y_{13}, y_{14}, y_{15}, y_{16}, y_{21}, y_{22}, y_{23}, y_{24}, y_{25}, y_{26}, y_{31}, y_{32}, y_{33}, y_{34}, y_{35}, y_{36}, y_{41}, y_{42}, y_{43}, y_{44}, y_{45}, y_{46}, y_{51}, y_{52}, y_{53}, y_{54}, y_{55}, y_{56}, y_{61}, y_{62}, y_{63}, y_{64}, y_{65}, y_{66}\}$$