

UNIDAD 3



MÉTODOS Y TÉCNICAS DE INVESTIGACIÓN

TEMA 1:

Proceso de selección de muestra

ÍNDICE

1. Unidad 3: Contextualización del problema	3
<i>Tema 1: Proceso de selección de la muestra</i>	3
<i>Objetivo:.....</i>	3
<i>Introducción:</i>	3
2. Información de los subtemas	4
2.1 <i>Subtema 1: Muestreo.....</i>	4
2.2 <i>Subtema 2: Muestra Probabilística.....</i>	11
2.3 <i>Subtema 3: Muestra no Probabilística.....</i>	15
<i>Conclusión</i>	15
3. Bibliografía	16

1. Unidad 3: Metodología de la investigación

Tema 1: Proceso de selección de la muestra

Objetivo:

Conocer las diferentes técnicas de muestreo para obtener la muestra de una población.

Introducción:

Este apartado tiene como objetivo seleccionar la muestra que se utilizará para validar la investigación. La selección se realizará mediante la herramienta de muestreo.

Se conocerá la definición de la muestra, la cual es un conjunto de elementos extraídos de una población. La población puede ser finita o infinita y mediante el uso de fórmulas se obtendrá el conjunto de elementos que se va a analizar.

También revisaremos los tipos de muestreos probabilísticos, no probabilísticos y sus características.

En los muestreos probabilísticos veremos tres opciones: muestra aleatoria simple, estratificada y por racimos. De los no probabilísticos: por conveniencia, cuotas y bola de nieve.

Todos estos conceptos ayudarán a diferenciar y escoger el muestreo que más se apegue a nuestra necesidad y de esta forma seleccionar la muestra más representativa de la población.

2. Información de los subtemas

2.1 Subtema 1: Muestreo

A diario hacemos uso de la herramienta del muestreo, por ejemplo, cuando queremos saber de nuestro estado de salud general nos realizamos una serie de exámenes a partir de una muestra de sangre. Mediante la examinación de esta muestra conoceremos los índices de glóbulos rojos, blancos, plaquetas, concentración de glucosa, etc. Con estos resultados se puede realizar un análisis y detectar enfermedades o afectación médica.

Por lo tanto, para que nuestra investigación tenga validez utilizamos el muestreo. El muestreo es un instrumento o proceso mediante el cual se selecciona un conjunto de individuos, que serán los más representativos de la población a investigar.

Es importante indicar que este proceso de muestreo se lo realiza cuando la población es muy amplia y no es conveniente realizar un censo. Por lo tanto, el muestreo tiene como función determinar qué segmento de la población debe examinarse.

Por consiguiente, el muestreo es

Útil

- Obtenemos información confiable
- Incurre en menos gastos

Necesario

- Única forma de estudiar la población
- Posibilita profundizar en el análisis de las variables
- Permite que el estudio se realice en menor tiempo
- Permite tener mayor control de las variables a estudiar.

“Al momento de realizar el muestreo se debe tener en cuenta las siguientes decisiones técnicas:

- **Procedimiento de selección de elementos:** Se trata de elegir entre las opciones de muestreos probabilísticos y no probabilísticos.
- **Tamaño de la muestra:** Básicamente se trata de aplicar la fórmula correspondiente según el procedimiento de selección elegido.
- **Nivel de confianza y probabilidad de error:** Se trata de tomar decisiones respecto a la precisión de las estimaciones.

- **Probabilidad de selección de las unidades de muestreo:** Establecer si todas las unidades tendrán la misma probabilidad de ser elegidas o probabilidades distintas de selección.” (Vivanco, 2005)

El muestreo es útil en cualquier proceso porque está acompañado de un proceso inverso, al que llamamos generalización de resultados. Es decir, lo que se realiza para conocer el universo es:

- Extraer una muestra del mismo
- Medir un dato u opinión
- Proyectar en el universo el resultado observado en la muestra. (Ochoa, 2015)

Muestra

“La muestra es, en esencia, un subgrupo de la población. Digamos que es un subconjunto de elementos que pertenecen a ese conjunto definido en sus características al que llamamos población.” (Sampieri)

La muestra debe ser una representación apropiada de la población, en donde se represente los rasgos fundamentales de la población. Es representativa porque tiene “Nivel de confianza y Margen de error”.

Para lograr la representación adecuada de la población se debe considerar el siguiente “diseño de la muestra:

- Definir la población de estudio.
- Obtener el marco muestral pertinente.
- Escoger la técnica de muestreo más adecuada.
- Tomar decisiones sobre el tamaño y el margen de error de la muestra.” (Morone)

“El marco muestral está compuesto por un listado de los elementos que componen el universo y que permitirán su identificación, además es la base de los muestreos probabilísticos.” (Morone)

Las siguientes fórmulas ayudarán con el cálculo de la muestra, es decir, cuántos elementos serán seleccionados en la población.

Hay dos tipos de población:

Finita: Está formado por un número limitado de elementos. Ejemplo:

- Los habitantes de una región del Ecuador
- Número de estudiantes en una institución Universitaria
- Número de empleados en una empresa
- La flota de vehículos de una transportadora

Para este tipo de población se utiliza la siguiente fórmula:

$$n = \frac{N * Z^2 * p * q}{e^2 * (N-1) + Z^2 * p * q}$$

*Ilustración 1. Fórmula para población finita.
Elaborado a partir de Barojas, S. A. (2005)*

Infinita: Está formada por un número extremadamente grande de elementos, donde no se pueden contabilizar todos sus elementos, ya que existe un número ilimitado

Ejemplo:

- Población de grillos en el mundo
- Número de estrellas en el universo
- Cantidad de granos de área
- Número de gotas de agua en un lago

Para este tipo de población se utiliza la siguiente fórmula:

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

*Ilustración 2. Fórmula para población infinita.
Elaborado a partir de Barojas, S. A. (2005)*

A continuación, se detalla el significado de las variables

Variable	Significado
n	Tamaño de la muestra
Z	Nivel de confianza
p	Variable positiva
q	Variable negativa
N	Tamaño de la población
E	Precisión o error

*Tabla 1. Significado de las variables.
Elaborado a partir de Barojas, S. A. (2005)*

Como se indicó en el cuadro anterior, la variable n representa el tamaño de la muestra, es la cantidad de elementos que se seleccionarán para el estudio, los cuales serán representativos para la población.

La variable Z es el nivel de confianza esperado, es el grado de certeza o probabilidad expresada en porcentaje. El nivel de confianza indica la probabilidad de que los resultados de nuestra investigación sean ciertos y es colocado por el investigador. Los valores de Z se obtienen de la tabla de la distribución normal estándar $N(0,1)$.

Los valores más usados son: 90% - 95% - 99%. Al escoger un nivel de confianza por ejemplo 95% estaremos indicando que dicho porcentaje es la probabilidad de que los resultados de la investigación sean verdaderos. El porcentaje restante en este caso 5% representará el error.

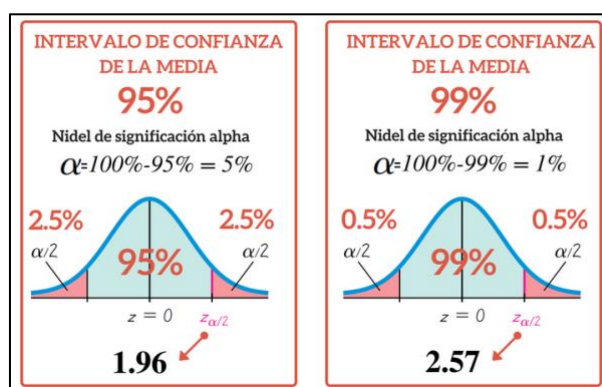


Ilustración 3. Intervalo de confianza.
Google images.

En la siguiente tabla se detalla los valores críticos por cada porcentaje de confianza.

% Error	Nivel de Confianza	Valor crítico (Z)
10	90%	1.645
5	95%	1.96
1	99%	2.58

Ilustración 4. Nivel de confianza y su Valor crítico.
Elaborado a partir de Barojas, S. A. (2005)

La variable “ p ” indica la proporción de elementos que poseen la característica de estudio, es la probabilidad de éxito o proporción esperada. En ciertos casos este valor es desconocido al no haber antecedentes, pero el comúnmente usado es 0.5.

La variable “ q ” a diferencia de la p indica la probabilidad de fracaso, es decir, es la proporción de elementos que no poseen la característica de estudio. Estas variables son complementarias y sumadas dan el valor de 1 ($p + q = 1$). Por ejemplo: Si la

variable p es igual a 0.5, la variable q sería igual a $1-p$, es decir $1 - 0.5$ dando como resultado 0.5.

Para conocer “ p ” es necesario tener investigaciones pasadas. En caso de no conocer “ p ” se le debe dar el mismo valor del que ocurra el evento que se está estudiando como que no ocurre, entonces $p = 50\%$ y $q = 50\%$.

La variable N es el total de la población o Universo. Son elementos en los cuales se pueden presentar características susceptibles a ser estudiadas, dependiendo si la población es finita o infinita se debe aplicar las fórmulas anteriormente mencionadas.

La variable “ e ” indica la precisión o error entre el valor observado (muestra) versus el valor verdadero (población). Representa el límite aceptable de error muestral. Siendo 5% (0.5) el valor estándar usado en las investigaciones. Este valor también es colocado por el investigador.

Por ejemplo: Si tenemos los resultados de una encuesta la cual indica que 100 individuos adquirirían productos de belleza de la marca “La Belleza” y hay un error muestral del 5%, se interpretaría que los productos los comprarán entre 95 y 105 individuos.

Existe una relación entre el error, nivel de confianza y tamaño de la muestra, al modificar cualquiera de los tres parámetros altera los restantes:

- Reducir el margen de error obliga a aumentar el tamaño de la muestra.
- Aumentar el nivel de confianza obliga a aumentar el tamaño de la muestra.
- Si aumenta el tamaño de mi muestra, puedo reducir el margen de error o incrementar el nivel de confianza. (Ochoa, 2015)

Ejemplo del cálculo de la muestra para población finita e infinita

En una fábrica de shampoo se producen diariamente 22350 envases de 250 ml. Para certificar que el peso del contenido sea correcto, se toma aleatoriamente algunos envases y se pesan. Se conoce que la variabilidad positiva es de $p=0.7$. Si se quiere garantizar un nivel de confianza de 90% y un porcentaje de error del 5%. ¿Cuántos envases se debe pesar?

$$n = \frac{N * Z^2 * p * q}{e^2 * (N - 1) + Z^2 * p * q}$$

Datos		
N=	22350	
Z=	90%	1,645
p=	0,7	0,7
q=	(1 - 0,7)	0,3
e=	5%	0,05

Nivel de Confianza	Valor crítico (Z)
90%	1,645
95%	1,96
99%	2,58

$$n = \frac{12700,73}{56,44077}$$

n= 225

Ilustración 5. Desarrollo de ejemplo - población finita.
Elaboración propia.

Se requiere hacer un estudio de estudiantes que reprueban a nivel superior. El estudio no tiene antecedentes, pero se desea garantizar un nivel de confianza del 95% y un porcentaje de error máximo de un 5% ¿Cuál es el tamaño de la muestra?

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

Datos		
N=	-	
Z=	95%	1,960
p=	0,5	0,5
q=	(1 - 0,5)	0,5
e=	5%	0,05

Nivel de Confianza	Valor crítico (Z)
90%	1,645
95%	1,960
99%	2,58

$$n = \frac{0,9604}{0,0025}$$

n= 384

Ilustración 6. Desarrollo de ejemplo - población infinita.
Elaboración propia.

A continuación, se detalla los tipos de muestreo:

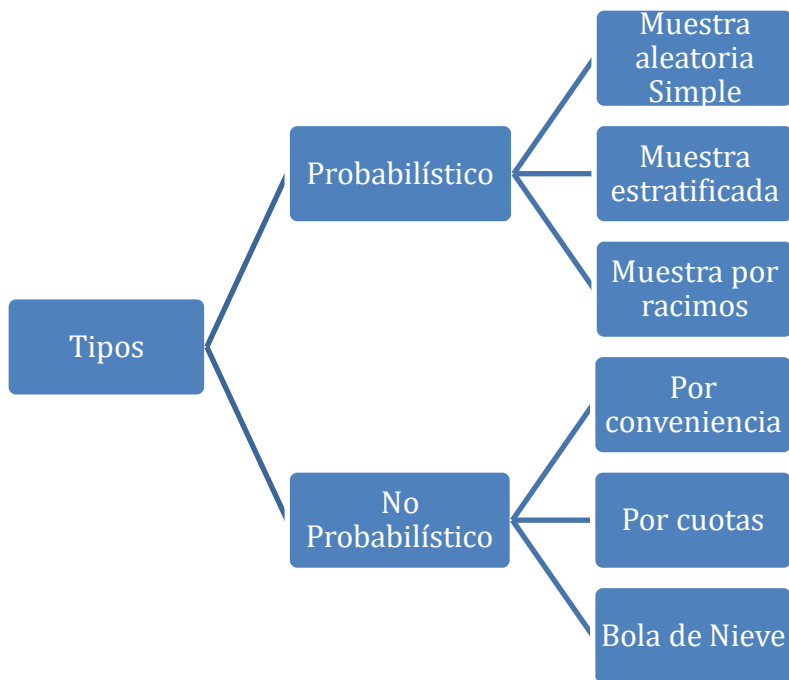


Ilustración 7. Tipos de muestreo.
Elaborado a partir Sampieri.

2.2 Subtema 2: Muestra Probabilística

“Es un subgrupo de la población en el que todos los elementos tienen la misma posibilidad de ser elegidos.” (Sampieri)

“Para que el muestreo sea probabilístico debe cumplir las siguientes condiciones:

- Todos los elementos o unidades de la población tengan la misma probabilidad de ser elegidos.
- Se apliquen procedimientos e instrumentos que garanticen que la selección sea aleatoria.” (Rojas, 2011)

Al cumplir con estas condiciones se certifica que cada elemento es escogido al azar y son representantes de la muestra seleccionada.

Muestra Aleatoria Simple

La muestra aleatoria simple se basa en que cada elemento de la población tiene la misma oportunidad de ser seleccionado para el estudio, por lo general la selección es al azar. Para este muestreo es necesario tener la lista completa de los elementos que conforman la población, es decir, la población debe ser finita con el fin de asignarle un número a cada elemento, de 1 a N.

Es importante mencionar que al ser un muestreo probabilístico no se considera los intereses que tenga el investigador.

Un ejemplo muy común de este muestreo es el método de la lotería, todos los números tienen la misma posibilidad de ser seleccionados. Otro ejemplo son los sorteos, puede haber una población pequeña de 100 personas que se les asignará un número. Dichos números son escritos en papelitos (en este caso 100 papelitos) que, a su vez, se almacenarán en una tómbola. Se calcula la muestra y supongamos que arroja el valor de 20. Entonces agita la tómbola y de ella sacamos 20 números al azar.

En base a lo anteriormente indicado, este tipo de muestra tiene como ventaja que es sencillo y al ser una selección al azar cada elemento tiene la misma probabilidad de ser seleccionado. Sin embargo, no convendría utilizarlo con poblaciones amplias pues se malversaría los recursos de costo y tiempo.

Pasos para seleccionar una muestra aleatoria simple:

- Definir la población
- Identificar un marco muestral
- Asignar un número único a cada elemento
- Obtener el tamaño de la muestra
- Seleccionar al azar

Muestra estratificada

“Es cuando la población se divide en segmentos y se selecciona una muestra para cada segmento.” (Sampieri)

Esta muestra se basa en hacer comparaciones entre segmentos de la misma muestra, es decir, una vez seleccionada la muestra se dividirá en sub-poblaciones o sub-grupos llamados estratos para luego comparar los resultados.

Su principal característica es asegurar que dichos estratos estén representados adecuadamente en la muestra y que tengan números iguales entre estratos.

Por ejemplo: Suponiendo que queremos evaluar el desempeño de los empleados de acuerdo al tiempo que llevan laborando en la empresa “Los Emprendedores”. La misma tiene 10000 colaboradores y cada uno ellos diferente tiempo laborando.

Para el estudio seleccionamos una población de 5000 empleados. Los años se segmentan de la siguiente forma, suponiendo que haya empleados que tengan como mínimo 1 año y máximo 15 años laborando.

- 1 año – 5 años = 2000
- 5 años – 10 años = 2500
- 11 años – 15 años = 500

Conocemos que nuestra población es de 5000 empleados y como muestra nos da un valor de 964:

$$n = \frac{N \cdot Z^2 \cdot p \cdot q}{e^2 \cdot (N-1) + Z^2 \cdot p \cdot q}$$

Z	1,96
p	50%
q	50%
N	10000
E	3%

9604
9,9595

n	964
---	-----

Ilustración 8. Cálculo de la muestra.
Elaboración propia.

Obtenida la muestra y definida la cantidad de estratos aplicaremos la siguiente fórmula que calcula la cantidad de elementos por cada estrato.

$$n_i = n \cdot \frac{N_i}{N}$$

siendo N el número de elementos de la población, n el de la muestra, N_i el del estrato i

Ilustración 9. Fórmula para cálculo de elementos por estrato.
Google images.

$$n_i = n \cdot \frac{N_i}{N}$$

n	964
Ni	2000
N	5000
Ni/N	0,4
n	964
ni	386

n	964
Ni	2500
N	5000
Ni/N	0,5
n	964
ni	482

n	964
Ni	500
N	5000
Ni/N	0,1
n	964
ni	96

Ilustración 10. Aplicación de Formula para el cálculo de estrato.
Elaboración propia.

Los resultados son los siguientes:

Estrato	Población	Muestra por Estrato
1 año - 5 años	2000	386
5 años - 10 años	2500	482
11 años - 15 años	500	96
TOTAL	5000	964

Tabla 2. Resultado de muestreo.
Elaboración propia.

Muestra por racimos

“En donde las unidades se encuentran encapsuladas en determinados lugares físicos.”
(Sampieri)

El muestreo por racimo implica una selección en una o más etapas, pero cada etapa es escogida al azar con la finalidad que tenga la misma probabilidad de ser seleccionados. Por cada racimo se escogerá la variable que se desea medir.

Este muestreo se lo utiliza en estudios donde cada grupo o racimo tiene variaciones, pero son similares entre sí.

Por ejemplo: Suponiendo que se desea determinar el porcentaje de docentes de primaria a nivel nacional que están dispuestos a utilizar las pizarras digitales como nuevo recurso en la impartición de clases.

Para el estudio, los racimos o grupos se deberían formar de la siguiente manera:

Primer grupo o racimo: Escoger cualquier región del país: Costa, Sierra, Oriente o Galápagos.

Segundo grupo o racimo: Supongamos que al realizar el muestreo aleatorio en el primer grupo la región Costa fue seleccionada, entonces para este segundo grupo elegimos 1 provincia de las 7 que conforman la región Costa. Así mismo se aplicará el muestreo aleatorio.

Tercer grupo o racimo: Si en el segundo grupo fue seleccionada la provincia Guayas, proseguimos a elegir el tipo de institución educativa ya sea pública o privada.

Cuarto grupo o racimo: Suponiendo que en el anterior grupo se escogió las instituciones educativas públicas, en este último grupo se seleccionará los docentes de primaria.

Se ratifica que para cada estudio la cantidad de grupo o racimos varía ya que esto depende de la variable que el investigador le interese estudiar.

2.3 Subtema 3: Muestra no Probabilística

En el muestreo No Probabilístico la elección de los elementos o individuos no tienen la misma posibilidad de ser elegidos, más bien depende de la decisión del investigador. Al ser selecciones informales no se puede realizar generalizaciones ya que la muestra obtenida no es representativa.

Por conveniencia:

En este tipo de muestreo el investigador selecciona intencionadamente los elementos o individuos que considera representativos de la población.

Por ejemplo: 20 Ingenieros en Sistemas se postularon para el cargo de Jefe de Sistema de la compañía XYZ, las 20 personas tienen experiencia en el cargo. Sin embargo, solo se selecciona 5 personas que tengan 3 o más años de experiencia.

Por cuotas

Este tipo de muestreo se maneja mediante cuotas, las cuotas son el número de individuos que reúnen ciertas características representativas en la investigación.

A continuación, se describe la siguiente cuota

- 20 estudiantes de 11 a 15 años, de sexo masculino y que vivan en Mariscal Sucre.

Determinada la cuota se seleccionan los primeros que cumplan con las características de la cuota.

Es parecido al muestreo estratificado, pero no se elige al azar.

Bola de Nieve

Este tipo de muestreo se lo utiliza en poblaciones muy pequeñas, debido a que al primer individuo seleccionado se le pide que identifique a otro que cumpla con las características de la investigación. Al igual que una bola de nieve esta va creciendo a medida que los individuos ya seleccionados recluten a otros. Por ejemplo, si el estudio es sobre el cáncer de mamas, se selecciona a una mujer y esta a su vez reclutará a familiares, amigas o conocidas que han pasado por lo mismo y esas mismas a otras.

Conclusión

El muestreo es de vital importancia en cualquier investigación ya que analizar la población completa no siempre es posible y mediante la selección de la muestra podemos analizar los elementos y dar resultados concluyentes.

Se recalca que la muestra debe ser lo suficientemente representativa para que los resultados obtenidos puedan generalizarse a la población.

3. Bibliografía

- » Barojas, S. A. (2005). Formulas para el cálculo de la muestra en investigaciones de salud. *Salud en Tabasco*.
- » Morone, G. (s.f.). Métodos y Técnicas de la Investigación Científica.
- » Ochoa, C. (Febrero de 2015). *Netquest*. Obtenido de <https://www.netquest.com/>
- » Rojas, V. M. (2011). *Metodología de la Investigación - Diseño y Ejecución*. Bogotá.
- » Sampieri, R. H. (s.f.). *Metodología de la Investigación. 6ta Edicion*.
- » Vivanco, M. (2005). Muestreo Estadístico. Diseño y Aplicaciones. Editorial Universitaria.