

**MAESTRIA EN ADMINISTRACIÓN Y POLÍTICAS PÚBLICAS**

**Materia: Estadística Administrativa**

**Actividad: 3 Tema: «Ejerc. Estadísticas Descriptivas-Control de Lectura»**

**ASESOR:**

**Dr. Enrique Antonio Paniagua Molina**

**MAESTRANTE:**

**Emperatriz González Alfaro**

*Tapachula de Córdova y Ordoñez, Chiapas; 21 de Septiembre de 2015.*

**“Selección de la Muestra”**

Cuando realizamos una investigación, sabemos que lo importante es conocer los sujetos, objetos, sucesos o comunidades de estudio (las unidades de análisis); es por ello, que durante la selección de una muestra lo primero que hay que formar es la definición de la unidad de análisis que nos llevará a la delimitación de la población. Cabe señalar que no siempre podemos obtener una muestra, pero en la mayoría de los entornos sí realizamos un estudio en una muestra, un ejemplo claro es cuando se realiza un censo.

Al seleccionar la muestra debemos evitar tres errores que pueden presentarse: 1) no elegir a casos que deberían ser parte de la muestra (participantes que deberían estar y no fueron seleccionados), 2) incluir a casos que no deberían estar porque no forman parte de la población y 3) seleccionar casos que son verdaderamente inelegibles (Mertens, 2005).

La muestra es un subgrupo de la población, son elementos a ese conjunto definido, pertenecen a ese conjunto definido en sus características al que llamamos población. Los tipos de muestras son las siguientes:

Muestras no probabilísticas: Es la elección de los elementos donde no depende de la probabilidad, sino de causas relacionadas con las características de la investigación o de quien hace la muestra.

Las muestras probabilísticas: Son esenciales en los diseños de investigación transeccionales, tanto descriptivos como correlacionales-causales, donde se pretende hacer estimaciones de variables en la población; por tal motivo, es primordial determinar el tamaño de la muestra y seleccionar los elementos muéstrales, de manera que todos tengan la misma posibilidad de ser elegidos.

Es importante saber que 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.

Para el procedimiento de selección de muestra las unidades de análisis o los elementos muéstrales se eligen siempre aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido:

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.* Aquí 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. Para dar algunos ejemplos tenemos la tabla 8.3.

En la primera columna se encuentran unidades de análisis que frecuentemente vamos a estudiar. En la segunda, sugerimos posibles racimos donde se encuentran dichos elementos.

Tabla 8.3 Ejemplo de racimos o *clusters*

Unidad de análisis Posibles racimos

Adolescentes Preparatorias

Obreros Industrias

Amas de casa Mercados

Niños Colegios

Muestrear por racimos implica diferenciar entre la unidad de análisis y la unidad muestra! La unidad de análisis indica quiénes van a ser medidos, o sea, los participantes o casos a quienes en última instancia vamos a aplicar el instrumento de medición. La unidad muestral (en este tipo de muestra) se refiere al racimo por medio del cual se logra el acceso a la unidad de análisis. El muestreo por racimos supone una selección en dos etapas, ambas con procedimientos probabilísticos. En la primera, se seleccionan los racimos, siguiendo los pasos ya señalados de una muestra probabilística simple o estratificada. En la segunda, y dentro de estos racimos, se selecciona a los sujetos u objetos que van a medirse.

Para ello se hace una selección que asegure que todos los elementos del racimo tienen la misma probabilidad de ser elegidos.

El procedimiento se lleva acabo cuando iniciamos nuestra exposición sobre la muestra probabilística, señalamos que los tipos de muestra dependen de dos cosas: del tamaño de la muestra y del procedimiento de selección. De lo primero hemos hablado con todo detalle, de lo segundo trataremos ahora. Se determina el tamaño de la muestra *n,* pero ¿cómo seleccionar los elementos muéstrales? Se precisa el número de racimos necesario y ¿cómo se seleccionan los sujetos dentro de cada racimo? Hasta el momento sólo hemos dicho que los elementos se eligen de manera aleatoria, pero ¿cómo se hace esto? Las unidades de análisis o los elementos muéstrales se eligen siempre aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido. Se utilizan tres procedimientos de selección:

**Tómbola:** Consiste en enumerar todos los elementos muéstrales del uno al número n. Hacer fichas o papeles, e ir sacando números al azar.

**Números Ramdon o números aleatorios:** Es la utilización de números que implica un mecanismo de probabilidad muy bien diseñado, fueron generados con una especie de ruleta electrónica.

**Selección Sistemática de elementos muestrales:** La selección es muy útil e implica elegir dentro de una determinada población determinados elementos a partir de un intervalo que se determina por el tamaño de la población y el tamaño de la muestra.

¿De que requieren Las muestras probabilísticas? Requieren de 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. Es por ello, que los listados existentes sobre una población son variados: guías telefónicas listas de miembros de asociaciones, directorios especializados, listas oficiales de escuelas en la zona, bases de datos de los alumnos de una universidad o de los clientes de una empresa, nóminas de una organización, entre otros.

**Listados y otros marcos muéstrales:** Las muestras probabilísticasrequieren 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 permita identificar físicamente los elementos de la población, la posibilidad de enumerarlos y, por ende, de proceder a la selección de los elementos muéstrales (los casos de la muestra). Normalmente se trata de un listado existente o una lista que es necesario confeccionar *ad hoc,* con los casos de la población.

**Los listados existentes sobre una población son variados:** Pueden ser guías telefónicas, listas de miembros de las asociaciones, directorios especializados, listas oficiales de escuelas de la zona, bases de datos de los alumnos de una universidad o de los clientes de una empresa, registros médicos, catastros, nóminas de una organización, etc. En todo caso hay que tener en cuenta lo completo de una lista, su exactitud, su veracidad, su calidad y su nivel de cobertura en relación con el problema a investigar y la población que va a medirse, ya que todos estos aspectos influyen en la selección de la muestra.

**Archivos:** Un gerente de reclutamiento y selección de una empresa quiere precisar si algunos datos que se dan en una solicitud de trabajo están correlacionados con el ausentismo del empleado. Es decir, si a partir de datos como edad, género, estado civil, nivel educativo y duración en otro trabajo, es factible predecir la conducta de ausentismo. Para establecer correlaciones se considerará como población a todas las personas contratadas durante 10 años. Se relacionan sus datos en la solicitud de empleo con los registros de faltas.

**Mapas:** Los mapas son muy útiles como marco de referencia en muestras de racimos. Por ejemplo, un investigador quiere saber qué motiva a los compradores de las tiendas de autoservicio. A partir de una lista de tiendas de cada cadena competidora, marca sobre un mapa de la ciudad, todas las tiendas de autoservicios, las cuales constituyen una población de racimos, pues en cada tienda seleccionada entrevistará a un número de clientes. El mapa le permite ver la población (tiendas de autoservicio) y su situación geográfica, de manera que elige zonas donde coexistan diferentes tiendas competidoras, para asegurarse de que el consumidor de la zona tenga todas las posibles alternativas.

Para el tamaño óptimo de una muestra las muestras probabilísticas requieren de dos procedimientos básicos: la determinación del tamaño de la muestra, el cual fue descrito en su modalidad más simple en la sección sobre el tamaño de la muestra y la selección aleatoria de los elementos muéstrales.

Existen varios tipos de medición como el BRIE que es el barómetro del Real Instituto Elcano; el eurobarómetro es otra encuesta que abarca a diversos países de la Unión Europea y su muestra es de aproximadamente 1000 personas por país, exceptuando Alemania y el Reino Unido.

Para finalizar, es importante saber que una buena definición de la unidad de análisis, puede delimitar claramente la población en base a los objetivos de estudio y saber las características del contenido y lugar.

**“Ejercicios de Estadística Descriptiva”**

**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:****

**Respuestas:**

**T´=** La casa tiene una chimenea, cuesta más de $ 100 mil pesos y es nueva.

**U´=** Tiene tres o más baños, cuesta más de $ 100 mil pesos y es nueva.

**V´=** Tiene tres o más baños, tiene chimenea y es nueva.

**W´=** Tiene tres o más baños, tiene chimenea y cuesta más de $ 100 mil pesos

**T intersección U=** Vacío.

**T intersección V=** Vacío.

**U´ intersección V=** Cuesta más de $ 100 mil pesos.

**V u W=** Cuesta más de $ 100 mil pesos y es nueva.

**V´u W=** Tiene tres o más baños, tiene chimenea y es nueva.

**T u U=** Tiene tres o más baños y tiene chimenea.

**T u V=** Tiene tres o más baños y cuesta más de $ 100 mil pesos

**V intersección W=** Vacío.

**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



Sub conjunto B



Probabilidad

Si *x* es la probabilidad que ocurra un número par, \_\_y\_\_ sería la probabilidad que ocurra un número impar. Considerando que y es doblemente probable

Entonces, encontramos que: \_2y\_+ *x* + \_2y\_ + *x* + \_2y\_+ *x* = 1

Esto se debe al postulado 2;

La P(B) sería: x + 2y + x = 2x + 2y; si x= 1/6, y= 1/6; entonces 2(1/6) + 2(1/6) = 4/6

**Ejercicio 3**

Calcula la muestra para una población desconocida con un 96% de confianza y 10% error. Para una prevalencia de .5 y .7

Para prevalencia de .5:

Z= 1.7506

p= 0.5

q= 1-p = 0.5

i = 0.1

n= (1.7506)2 (0.5) (0.5) = (3.0646) (0.5) (0.5) = **76.61**

(0.1)2 (0.01)

n= **77**

Para prevalencia de .7:

Z= 1.7506

p= 0.7

q= 1-p = 0.3

i = 0.1

n= (1.7506)2 (0.7) (0.3) = (3.0646) (0.7) (0.3) = **64.35**

(0.1)2 (0.01)

n= **64**

**Ejercicio 4.**

Calcula la muestra para una población de 350,000 familias, con un 99% de confianza y 5% error. Para una prevalencia de .5 y .7

Para prevalencia de .5:

Z= 2.3263 (2.575)

p= 0.5

q= 1-p = 0.5

i = 0.05

N= 350,000

n= (2.3263)2 (350,000) (0.5) (0.5) .

(0.05)2 (350,000-1) + (2.3263)2 (0.5) (0.5)

n= (5.4116) (350,000) (0.5) (0.5) = 473,521.27 = **540.33**

(0.0025) (349,999) + (1.3529) 876.35

n= **540**

Para prevalencia de .7:

Z= 2.3263

p= 0.7

q= 1-p = 0.3

i = 0.05

N= 350,000

n= (2.3263)2 (350,000) (0.7) (0.3) .

(0.05)2 (350,000-1) + (2.3263)2 (0.7) (0.3)

n= (5.4116) (350,000) (0.7) (0.3) = 397,752.60 = **453.98**

(0.0025) (349,999) + (1.1364) 876.13

n= **454**

**Ejercicio 5.**

De una población de 1,176 padres de familia de la ciudad de Tuxtla Gutiérrez, se pretende conocer la aceptación de los programas educativos mediante caricaturas. Se pretende obtener una muestra para saber el número de entrevistas y con ello obtener información estadísticamente confiable. Se asume un [error standard](file:///C:\Users\oem\Downloads\Error%20estándar.docx) de 1.5% con un nivel de confiabilidad del 90%

n = n´ n´ = s2

1 + (n´/ N) V2

Si:

N = 1,176

s2 = p (1-p) = 0.9 (1-0.9) = 0.09

V2= (se)2 = (0.015)2 = 0.000225

p = 0.9

Entonces:

n´ = 0.09 = 400

0.000225

n = 400 = 400 = 298.48

1 + (400/1176) 1.3401

**n = 298**

**Ejercicio 6.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Estatura (x)** | **f** | **f . x** | **.**  **f . (xi – x)2** |
| 1.52 | 1 | 1.52 | 0.0196 |
| 1.54 | 5 | 7.70 | 0.0720 |
| 1.55 | 4 | 6.20 | 0.0484 |
| 1.58 | 5 | 7.90 | 0.0320 |
| 1.60 | 2 | 3.20 | 0.0072 |
| 1.62 | 4 | 6.48 | 0.0064 |
| 1.64 | 7 | 11.48 | 0.0028 |
| 1.66 | 3 | 4.98 | 0 |
| 1.70 | 5 | 8.50 | 0.0080 |
| 1.71 | 8 | 13.68 | 0.0200 |
| 1.73 | 6 | 10.38 | 0.0294 |
| 1.74 | 5 | 8.70 | 0.0320 |
| 1.77 | 3 | 5.31 | 0.0363 |
| 1.80 | 1 | 1.80 | 0.0196 |
| 1.83 | 1 | 1.83 | 0.0289 |
| Total | 60 | 99.66 | 0.3626 |

Son los resultados de preguntarles la estatura a 60 trabajadores del departamento de limpia municipal de SCLC.

Obtén la media aritmética (para datos agrupados)

Obtén la desviación estándar y la varianza (para datos agrupados)

Interpreta los resultados

**RESPUESTAS:**

.

Media aritmética= x = Sumatoria (f . x) = 99.66 = 1.66

n 60

Desviación estándar = Raíz cuadrada de 0.3626 = 0.0777

60

Varianza = V = el cuadrado de la desviación estándar = (0.0777)2 = 0.0060

Con base en los resultados obtenidos podemos concluir que la media en la estatura de los 60 trabajadores es de 1.66, correspondiente a 3 trabajadores de los 60; lo que nos indica que 57 trabajadores; es decir, el 94.68%; ostenta diferentes estaturas respecto a la media. Así mismo, la estatura más baja corresponde a un trabajador con 1.52 y la más alta a otro con 1.83; lo que nos indica que la dispersión de la población, con respecto a la media, oscila entre estas dos estaturas. También podemos concluir que 28 trabajadores tienen una estatura por debajo de la media y 29 trabajadores, arriba de la media.

**BIBLIOGRAFIA**

[**METODOLOGÍA DE LA INVESTIGACIÓN**](http://www.iapchiapasenlinea.mx/download.php?file=resources/recurso_187.pdf), Hernández Sampierí Roberto, Fernández Collado, Carlos, Pilar Baptísta Lucio, 4ta edición, editorial Mc Graw Hill, Pags 267-305.