Lic. Belinda García Isasi

MAESTRIA EN ADMINISTRACION Y POLITICAS PUBLICAS. ESTADISTICA ADMINISTRATIVA. DR. ENRIQUE ANTONIO PANIAGUA MOLINA.

CAP.8 selección de la muestra.

EJERCICIOS DE ESTADISTICA DESCRIPTIVA.

ACT.3

CAPITULO. 8 SELECCIÓN DE LA MUESTRA.

En una investigación la mayoría de las veces necesitamos realizar el estudio de una muestra aunque hay algunas investigaciones que no requieran una muestra. Las muestras son utilizadas para economizar recurso y tiempo.

Dependiendo del planteamiento de nuestra investigación, sabremos a quien o a quienes debemos estudiar, interrogar, hacer entrevista, cuestionario etc.

Para seleccionar la muestra debemos definir la unidad de análisis, en el caso de mi investigación la muestra se hará a un subgrupo de mujeres que hayan destacado en la política en la ciudad de Tapachula Chiapas. En años anteriores la mujeres ocupaban un segundo lugar a nivel política los mejores puestos, como son las diputaciones locales, las presidencias municipales, las secretarias de ayuntamiento, por lo general eran ocupados por hombres, a la mujer no la tomaban en cuenta aunque esta tuviera conocimientos y estudios profesionales.

Una vez definido cuál será la unidad de análisis, debemos delimitar la población que se habrá de estudiar, y sobre la cual generaremos resultados. La calidad de un trabajo de investigación, estriba en delimitar claramente la población con base en el planteamiento del problema.

Al seleccionar una muestra debemos evitar tres errores:

1. No elegir a casos que deberían ser parte de la muestra (participantes que deberían estar y no están seleccionados)
2. Incluir a casos que no deberían estar porque forman parte de la población.
3. Seleccionar casos que son verdaderamente inelegibles.

La adecuada delimitación del universo o población nos ayuda a evitar errores, toda investigación debe ser transparente, así como estar sujeta a crítica y replica.

Todas las muestras bajo el enfoque cuantitativo deben de ser representativas.

Existen dos tipos de muestras:

1.- las muestras no probabilísticas.- subgrupo de la población en el que todos los elementos de esta tienen la misma posibilidad de ser elegidos.

2.- las muestras probabilísticas.- puede medirse el tamaño del error en nuestras predicciones, el principal objetivo de una muestra probabilística es reducir el mínimo de este error estándar.

Para hacer una muestra probabilística es necesario entender los siguientes términos y sus definiciones:

La población, a la que se le suele denominar como ii, es un conjunto de elementos.

La muestra a la que se simboliza como n, es un subconjunto de la población N.

En una población N (previamente delimitada por objetivos dela investigación), nos interesa establecer valores de las características de los elementos N.

Nos interesa conocer valores promedio en la población, lo cual se expresa como:

Y=al valor de una variable determinada (Y) que nos interesa conocer, digamos un promedio.

También nos interesa conocer.

V=la varianza de la población con respecto a determinadas variables (la varianza indica la variabilidad).

Para seleccionar el tamaño de la muestra probabilística, nos debemos preguntar: dado que una población es de N, ¿Cuál es el menor número de unidades muéstrales que necesito para conformar una muestra(n) que me asegure un determinado nivel de error estándar, digamos menor de 0.01?

Por ejemplo en el caso de mi investigación acerca del papel dela mujer en la política en el municipio de Tapachula Chiapas.

Lo primero será determinar o conocer N (población o universo)

Lo segundo es establecer el error aceptable y el nivel de confianza que suele ser de 1 o 5%en el caso de error. Debemos tener la seguridad de que la muestra sea representativa de la población con 95% o 99% en cuanto a nivel de confianza.

¿Cómo se lleva a cabo el procedimiento de selección de la muestra?

Se utilizan tres procedimientos de selección:

-Tómbola.- los números elegidos al azar conformaran la muestra.

- Números random o números aleatorios.- se utiliza un atabla de números que implica un mecanismo de probabilidad muy bien diseñado.

- Listados y otros marcos muéstrales.- este constituye un marco de referencia que nos permite identificar físicamente los elementos dela población, la posibilidad de enumerarlos y, por ende, de proceder a la selección de los elementos muéstrales. (Archivos, mapas, volúmenes, periodos registrados.)

Tamaño óptimo de la muestra:

Las muestras probabilísticas requieren de dos procedimientos básicos:

1. La determinación del tamaño de la muestra.
2. La selección aleatoria de los elementos muéstrales.

Muestras no probabilísticas.

También se les llama dirigidas, se utilizan en investigaciones cualitativas y cuantitativas.

Desde la visión cuantitativa la ventaja de una muestra no probabilística es su utilidad para determinado diseño de estudio se requiere no tanto una representatividad de elementos de una población, sino una cuidadosa y controlada selección de sujetos con ciertas características especificadas previamente en el planteamiento del problema.

En cuanto al enfoque cualitativo, al no interesar tanto la posibilidad de generalizar los resultados, las muestras no probabilísticas o dirigidas son de gran valor, pues logran obtener los casos que interesan al investigados y que llegan a ofrecer una gran riqueza para la recolección y análisis de los datos.

**EJERCICIOS DE ESTADISTICA DESCRIPTIVA.**

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:**



Respuesta

Consideramos que los eventos no son excluyentes, entonces

***T´*** **=** es el evento de que la casa tiene una chimenea, y/o cuesta más de $100 mil pesos y/o es nueva

***U´*** **=** es el evento de que la casa tiene tres o más baños, y/o cuesta más de $100 mil pesos y/o es nueva

***V´* =** es el evento de que la casa tiene tres o más baños, y/o tiene una chimenea y/o es nueva

***W´* =** es el evento de que la casa tiene tres o más baños, y/o tiene una chimenea y/o cuesta más de $100 mil pesos

***T*** ***U* =** es el evento de que la casa tiene tres o más baños y tiene una chimenea

***T*** ***V* =** es el evento de que la casa tiene tres o más baños y cuesta más de $100 mil pesos

***U´*** ***V*** = es el evento de que la casa cuesta más de $100 mil pesos

***V W =*** es el evento de que la casa cuesta más de $100 mil pesos o la casa es nueva

***V´ W =*** es el evento de que la casa tiene tres o más baños, y/o tiene una chimenea y/o la casa es nueva

***T U=*** es el evento de que la casa tiene tres o más baños o la casa tiene una chimenea

***T V=*** es el evento de que la casa tiene tres o más baños o la casa cuesta más de $100 mil pesos

***V***  ***W* =** es el evento de que la casa cuesta más de $100 mil pesos y es nueva

1. **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, *2x* sería la probabilidad que ocurra un número impar.**
  + **Entonces, encontramos que: *2x* + *x* + *2x* + *x* +*2x* + *x* = 1 *lo que implica que* x= 1**

**9**

* + - **Esto se debe al postulado 2**
  + **La P(B) sería: 1 + 2 + 1 = 4**

**9 9 9 9**

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

n = Z 2 x p x q

i2

Respuesta:

El tamaño de la muestra con una prevalencia de 0.5 es

p = 0.5 entonces

q = 1-p

= 0.5

i, el error permitido es del 10%, entonces

i = 0.1

el nivel de confianza es 96%, es decir (1- )= .96, entonces =.04, es decir = 4%, al buscar en las tablas de la distribución Normal (0,1) encontramos que Z.04= 1.75, por lo tanto

n = (1.75)2 (0.5)(0.5)

(.1)2

n = 0.7656

.01

n = 76.56

Por lo tanto, el tamaño de muestra es de 77.

El tamaño de la muestra con una prevalencia de 0.7 es

p = 0.7 entonces

q = 1-p

= 0.3

i, el error permitido es del 10%, entonces

i = 0.1

el nivel de confianza es 96%, es decir (1- )= .96, entonces =.04, es decir = 4%, al buscar en las tablas de la distribución Normal (0,1) encontramos que Z.04= 1.75, por lo tanto

n = (1.75)2 (0.7)(0.3)

(.1)2

n = 0.3675

.01

n = 36.75

Por lo tanto, el tamaño de muestra es de 37.

1. **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**

n = Z 2 x N x p x q \_

i2(N-1)+Z 2 x p x q

Respuesta:

N = 350,000 familias

El tamaño de la muestra con una prevalencia de 0.5 es

p = 0.5 entonces

q = 1-p

= 0.5

i, el error permitido es del 5%, entonces

i = 0.05

el nivel de confianza es 99%, es decir (1- )= .99, entonces =.01, es decir = 1%, al buscar en las tablas de la distribución Normal (0,1) encontramos que Z.01= 2.33, por lo tanto

n = \_(2.33)2 (350,000)(0.5)(0.5)­ ­ ­

(.05)2 (350,000-1)+2.33 (0.5)(0.5)

n = 475,028.75

873.0825

n = 544.08

Por lo tanto, el tamaño de muestra es de 545.

Respuesta:

N = 350,000 familias

El tamaño de la muestra con una prevalencia de 0.7 es

p = 0.7 entonces

q = 1-p

= 0.3

i, el error permitido es del 5%, entonces

i = 0.05

el nivel de confianza es 99%, es decir (1- )= .99, entonces =.01, es decir = 1%, al buscar en las tablas de la distribución Normal (0,1) encontramos que Z.01= 2.33, por lo tanto

n = \_\_(2.33)2 (350,000)(0.7)(0.3)­ ­ ­ \_

(.05)2 (350,000-1)+2.33 (0.7)(0.3)

n = 399,024.15

872.9893

n = 457.07

Por lo tanto, el tamaño de muestra es de 458.

1. **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\Toshiba\Downloads\Error%20estándar.docx) **de 1.5% con un nivel de confiabilidad del 90%**

n = n´ \_ n´= ­\_S2­­ ­

1+ n´/N 2

Donde S2 = p(1-p) y 2 = (se)2

n = tamaño de la muestra

N = tamaño de la población

S2 = varianza muestral

2 = varianza poblacional

se = error estándar

p = % de confiabilidad

Entonces tenemos que

N = 1,176

se = 1.5%, es decir se = 0.015

p = 90%, es decir p = 0.9

S2 = .9(1-.9), entonces S2 = 0.09

2 = (0.015)2 entonces 2 = 0.0002

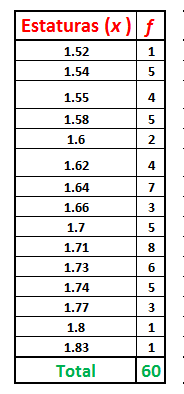
n´ = \_0.09 \_ , n´ = 450

0.0002

n = 450 \_ , por lo tanto n = 325.46

1 + 450/1,176

Entonces el tamaño de la muestra es 326.



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

**Obtén la media aritmética (para datos agrupados)**

ESTATURA (x) frecuencia (f) x . f

1.52 1 1.52

1.54 5 7.70

1.55 4 6.20

1.58 5 7.90

1.60 2 3.20

1.62 4 6.48

1.64 7 11.48

1.66 3 4.98

1.70 5 8.50

1.71 8 13.68

1.73 6 10.38

1.74 5 8.70

1.77 3 5.31

1.80 1 1.80

1.83 \_1\_ 1.83

60 99.66

Entonces la media aritmética es

***x*** = 99.66

60

Por lo tanto ***x*** =1.66, lo cual quiere decir que la estatura promedio de los 60 trabajadores del departamento de limpia municipal de SCLC es de un metro y sesenta y seis centímetros.

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

ESTATURA (x) (x – ***x***) (x – ***x***)2 f . (x – ***x***)2

1.52 -0.14 0.0196 0.0196

1.54 -0.12 0.0144 0.0720

1.55 -0.11 0.0121 0.0484

1.58 -0.08 0.0064 0.0320

1.60 -0.06 0.0036 0.0072

1.62 -0.04 0.0016 0.0064

1.64 -0.02 0.0004 0.0028

1.66 0.00 0.0000 0.0000

1.70 0.04 0.0016 0.0080

1.71 0.05 0.0025 0.0200

1.73 0.07 0.0049 0.0294

1.74 0.08 0.0064 0.0320

1.77 0.11 0.0121 0.0363

1.80 0.14 0.0196 0.0196

1.83 0.17 0.0289 0.0289

0.3626

Entonces la desviación estándar es:

= 0.3626/60

Por lo tanto = 0.0777, quiere decir que “las estaturas de los 60 trabajadores se alejan” 7.77 centímetros de la estatura media que es 1.66 mts. En la tabla de estaturas y frecuencias, se observa que hay 25 personas cuyas estaturas son superiores en 7.77 cms. a la estatura media.

Entonces el 41.67% (25/60) de los empleados sobrepasan en más de 7.77 cms. la estatura media.

La varianza es entonces

*v* = 2

*v* = (0.0777)2

*v* = 0.0060, hay que recordar que como la varianza es la desviación estándar elevada al cuadrado, el número que arroja la varianza ya no está dado en las mismas unidades que la media y la desviación estándar. Entonces nos da una medida de dispersión de los datos observados en el estudio en cuestión con respecto a la media de los mismos.