INSTITUTO DE ADMINISTRACIÓN PÚBLICA DEL ESTADO DE CHIAPAS, A.C.

ESTADISTICA ADMINISTRATIVA.

MATERIA:

DRA. ENRIQUE ANTONIO PANIAGUA MOLINA

CATEDRÁTICO:

3

ACTIVIDAD:

ESTADISTICA D-CONTROL L.

TEMA:

ERNESTO ROSS REYES

NOMBRE DEL ALUMNO:

TAPACHULA CHIS, A 20 DE SEPTIEMBRE DE 2015

SELECCIÓN DE LA MUESTRA

Las muestras se utilizan por economía de tiempo y recursos. Para seleccionar una muestra, lo primero que hay que hacer es definir la unidad de análisis (personas, organizaciones, periódico, comunidades, situaciones, eventos, etc.). El sobre que o quienes se van a recolectar datos depende del planteamiento del problema a investigar y de los alcances del estudio. Estas acciones nos llevarán al siguiente paso, que consiste en delimitar 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. Una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones (Selltiz et al., 1980). Es preferible establecer con claridad las características de la población, con la finalidad de delimitar cuales serán los parámetros muéstrales. La delimitación de las características de la población no solo depende de los objetivos del estudio, sino de otras razones prácticas.

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 no fueron seleccionados), 2) incluir casos que no deberían estar porque no forman parte de la población y 3) seleccionar casos que son verdaderamente inelegibles (Mertens, 2005).

¿Cómo seleccionar la 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. En realidad, pocas veces es posible medir a toda la población, por lo que obtenemos o seleccionamos una muestra y, desde luego, se pretende que este subconjunto sea un reflejo del conjunto de la población.

Básicamente categorizamos las muestras en dos grandes ramas: **las muestras no probabilísticas y las muestras probabilísticas.** En las últimas los elementos de la población tienen la 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.

¿Cómo se selecciona una muestra probabilística?

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, el principal objetivo en el diseño de una muestra probabilística es reducir al mínimo este error, al que se le llama error estándar (Kish, 1995).

Las muestras probabilísticas son esenciales en los diseños de investigación transeccionales, tanto descriptivo como correlaciónales-causales, donde se pretende hacer estimaciones de variables en la población. Estas variables se miden y se analizan con pruebas estadísticas en una muestra, donde se presupone que esta es probabilística y todos los elementos de la población tienen una misma probabilidad de ser elegidos.

El tamaño de la muestra.

Cuando se hace una muestra probabilística, uno debe preguntarse: dado que una población es de N, ¿cuál es el menor número de unidades muéstrales (personas, organizaciones, capítulos de telenovelas, etc.) que necesito para conformar una muestra (n) que me asegure un determinado nivel de error estándar, digamos menor de 0.01?

La respuesta a esta pregunta busca encontrar la probabilidad de ocurrencia de Y, así como que mi estimado de y se acerque a Y, el valor real de la población. Si establecemos el error estándar y lo fijamos en 0.01, sugerimos que esta fluctuación promedio de nuestro estimado y con respecto a los valores reales de la población Y no sea > 0.01, es decir, que de 100 casos, 99 veces mi predicción sea correcta y que el valor de y se sitúe en un intervalo de confianza que comprenda el valor de Y.

Resumiendo, para una determinada varianza [V) de Y, ¿qué tan grande debe ser mi muestra? Ello se determina en dos pasos:

1. = Tamaño provisional de la muestra1 = varianza de la muestra/varianza de la población

2 n = n’ /(1+n’IN)

Selección de la muestra

Si cambiamos el nivel de error tolerado y el nivel de confianza (0.01 o 1% de error y 99% de confianza, el tamaño de la muestra será mucho mayor, en este caso de 6 488.53 comerciales). El tamaño de la muestra es sensible al error y nivel de confianza que definamos. A menor error y mayor nivel de confianza, mayor tamaño de muestra requerido para representar a lapoblación o universo.

Procesa de la investigación cuantitativa

Quizá tengamos 300 católicos y dos o tres de otras religiones.Entonces es cuando preferimos obtener una muestra probabilística estratificada (el nombre nos dice que será probabilística y que se considerarán segmentos o grupos de la población, o lo que es igual: estratos).

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.

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. A continuación daremos un ejemplo que comprenda varios de los procedimientos descritos hasta ahora y que ilustra la manera como frecuentemente se hace
Una muestra probabilística en varias etapas.
.
¿**Cómo se lleva a cabo el procedimiento de selección de la muestra?**
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.

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 elegido.
El tamaño de una muestra también depende del número de subgrupos que nos interesan en una población, lo óptimo de una muestra depende de cuanto se aproxima su distribución a la distribución de las características de la población, esta aproximación mejora al incrementarse el tamaño de la muestra. La normalidad de las distribución de muestras grandes no obedece a la normalidad de la distribución de una población, la distribución de diversas variables a veces es normal y en ocasiones está lejos de serlo; sin embargo, la distribución de muestras de 100 o más elementos tiende a ser normal y esto sirve para el propósito de hacer estadística inferencial, sobre los valores de una población.

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

Muestreo al azar

Los investigadores la utilizan para seleccionar muestras telefónicas, involucra identificar áreas geográficas y sus correspondientes códigos telefónicos.

Muestra multietapas o polietápas

Para extraer la muestra hemos utilizado diversos procedimientos,

Una máxima del muestreo y el alcance del estudio

Los estudios exploratorios regularmente emplean muestras dirigidas o no probabilísticas. Las investigaciones experimentales, la mayoría de las veces utilizan muestras dirigidas. Los estudios no experimentales descriptivos o correlacionales – causales debe emplear muestras probabilísticas si quieren que sus resultados sean generalizados a una población.

En el enfoque cuantitativo las muestras probabilísticas son esenciales en diseños de investigación, se pretenden generalizar los resultados a una población, todos los elementos de la población al inicio tienen la misma probabilidad de ser elegidos, tal precisión depende del error de muestreo, llamado también error estándar.

Determinar el tamaño de la muestra y seleccionar los elementos muéstrales en forma aleatoria, el tamaño de la muestra se calcula mediante fórmulas, las muestras probabilísticas son: simples, estratificadas, sistemáticas y por racismo. La estratificada aumenta la precisión aumenta la precisión de la muestra e implica el uso deliberado de submuestras para cada estrato o categoría que sea relevante en la población, muestra por racimo o conglomerados implica diferencias entre la unidad de análisis y la unidad muestra, en este tipo de muestreo hay una selección en dos etapas. En la primera se seleccionan los racimos: escuelas, organizaciones, salones de clase; en la segunda y dentro de los racimos, a los participantes que van a ser medidos.

Existen tres procedimientos de selección: 1.- tómbola; 2.- cuadro de números aleatorios; 3.- selección automática. Todo procedimiento de selección depende de listado o base de datos, ya sea existente o construidas ad hoc. Los listados pueden ser guías telefónicas, listas de asociaciones, listas de escuelas oficiales, algunas de estos pueden ser archivos, hemerotecas y mapas así como internet.

Las muestras no probabilísticas también pueden llamarse muestras dirigidas, pues la elección de sujetos u objetos de estudios dependen del criterio del investigador. En el teorema del límite central se señala que una muestra de más de cien casos será una muestra con una distribución normal en sus características: sin embargo no debe confundirse con probabilidad, lo primero es necesario para efectuar pruebas estadísticas, lo segundo es requisito indispensable para hacer inferencias correctas sobre una población

EJERCICIOS ESTADISTICAS DESCRIPTIVAS

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'= U, V, W (una chimenea, cuesta más de cien mil, casa nueva)
U'=T, V, W (Casa de 3 o más baños, cuesta más de cien mil, casa nueva)
V'=T, U, W (Casa de 3 o más recamaras, una chimenea, casa nueva)
W= casa nueva

T ∩ U= (Casa tiene 3 o más recamaras; tiene una chimenea)
T ∩ V= (tres o más baños; cuesta más de cien mil pesos)
U' ∩ V= (Cuesta más de cien mil)
V U W= (Cuesta más de cien mil; casa nueva)
V' U W= (tres o más baños; una chimenea; casa nueva)
T U U= (tiene tres o más baños; tiene chimenea)

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 $S = \{1,2,3,4,5,6\}$
- Sub conjunto 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+ x + 2x + x + 2x + x = 1

S= (1, 2, 3, 4, 5,6)

S=2x+ x + 2x + x + 2x + x = 1

9X=1

X=1/9

- Esto se debe al postulado 2

  B= (4, 5,6)

  B= X + 2X + X

  SUSTITUYENDO VALORES

  B= 1/9 + 2 (1/9) + 1/9= 4/9

- La P(B) sería: 4/9

Ejercicio 3.- Calcula la muestra para una población desconocida con un 96% de confianza y 10% error. Para una prevalencia de .5 y .7

FORMULA

$$n = \frac{Z_{\propto}^{2}.\, p.\, q}{i^{2}}$$

Para prevalencia de .5

$Z^2\alpha$= 1.7506

P= .5

.q= 1-p = (1-.5) = .5

.i= .10

Sustituyendo en la formula

$$´n = \frac{(1.7506)^2(.5)(.5)}{(.10)^2} = \frac{(3.064)(.5)(.5)}{.01} = \frac{(3.064)(.25)}{.01} =$$

´n= 76.61 ≈ 77

Para una prevalencia de 0.7

$Z^2\alpha$= 1.7506

P= .7

.q= 1-p = (1-.7)

.i= .10

Sustituyendo en la fórmula

$$´n = \frac{(1.7506)^2(0.7)(1-0.7)}{(0.10)^2} = \mathbf{64.35 \approx 64}$$

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

FORMULA $$n = \frac{Z^2_{\propto}.N.p.q}{i^2(N-1)+Z^2_{\propto}.p.q}$$

Para una prevalencia de 0.5

$Z^2\alpha$= 1.7506

P= .5

.q= 1-p = (1-.5) = .5

.i= .05

Sustituyendo en la formula

$$´n = \frac{(2.3263)^2(350000)(.5)(.5)}{(.05)^2(350000-1)+(2.3263)^2(.5)(.5)} = \frac{(5.4116)(350000)(.25)}{(0.0025)(349999)+(5.4116)(.25)} =$$

$$´n = \frac{473515}{} = \frac{473515}{} =$$

(874.9975)+1.3529 876.3504

´n= 540.326 ≈ 540

Para una prevalencia de 0.7

$Z^2\alpha$= 1.7506

P= .5

.q= 1-p = (1-.7) = .3

.i= .05

Sustituyendo en la formula

$$´n= \frac{(2.3263)^2 (350000)(.7)(.3)}{(.05)^2(350000-1)+ (2.3263)^2(.7) (.3)} = \frac{(5.4116)(350000)(.21)}{(0.0025)(349999)+(5.4116)(.21)} =$$

$$´n= \frac{397752.60}{(874.9975)+1.364} = \frac{397752.60}{876.1339} =$$

´n= 453.98 ≈ 454

Ejercicio 5: De una Población de 1,176 padres de familia de la Cd. de Tuxtla Gutierrez. Se pretende conocer la aceptación de los programas educativos mediante caricaturas. se pretende obtener una muestra par saber el número de entrevistas y con ello obtener información estadísticamente confiable. Se asume un error estándar de 1.5% con un nivel de confiabilidad del 90%

Datos: FORMULA

N=1176

Z=1.2815

$$n = \frac{Z^2_{\propto.} N.p.q}{i^2(N-1) + Z^2_{\propto.}.p.q}$$

p=0.5

q= 0.5

i=0.015

$$´n = \frac{(1.2815)^2 (1176)(.5)(.5)}{(.015)^2(1176-1) + (1.2815)^2(.5)(.5)} = \frac{(1.6422)(1176)(.25)}{(0.000225)(1175)+(1.6422)(.25)} =$$

$$´n = \frac{482.8068}{(0.2643)+0.2643} = \frac{482.8068}{0.6746} =$$

´n=715.48 ≈ 715

Ejercicio 6.-

- Son los resultados de preguntarle 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 resultado

Estaturas (X)	f	Ni	xi * fi	xi^2 * fi
1.52	**1**	**1**	**1.52**	**2.3104**
1.54	**5**	**6**	**7.7**	**11.858**
1.55	**4**	**10**	**6.2**	**9.61**
1.58	**5**	**15**	**7.9**	**12.482**

1.6	2	17	3.2	5.12
1.62	4	21	6.48	10.4976
1.64	7	28	11.48	18.8272
1.66	3	31	4.98	8.2668
1.7	5	36	8.5	14.45
1.71	8	44	13.68	23.3928
1.73	6	50	10.38	17.9574
1.74	5	55	8.7	15.138
1.77	3	58	5.31	9.3987
1.8	1	59	1.8	3.24
1.83	1	60	1.83	3.3489
Total	60		99.66	165.8978

$$\bar{x} = \frac{\sum_{i=1}^{n} xi}{N}$$

Desviación estándar para datos agrupados:

$$\bar{x} = \frac{99.66}{60} = 1.661$$

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{x_i^2 f_i}{N} - \bar{x}^2}$$

$$\sigma = \sqrt[2]{\frac{165.8978}{60} - 1.661^2}$$

$$\sigma = \sqrt[2]{2.764963 - 2.75892}$$

$$\sigma = \sqrt[2]{.006043}$$

$$\sigma = 0.077736$$

Varianza para datos agrupados:

$$\sigma^2 = \sum_{i=1}^{n} \frac{x_i^2}{N} - \bar{x}^2$$

$$\sigma^2 = \frac{165.8978}{60} - 1.661^2$$

$$\sigma^2 = 2.764963 - 2.758921$$

$$\sigma^2 = .006042$$

$x = 1.7387$

$\sigma = 0.077736$

$\bar{x} = 1.6610$

$\sigma = -0.077736$

$\bar{x} = 1.5832$