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

## MAESTRIA EN ADMINISTRACIÓN Y POLÍTICAS PÚBLICAS

MATERIA: **ESTADÍSTICA ADMINISTRATIVA**

ACTIVIDAD: **03.- EJERCICIOS DE ESTADÍSTICA DESCRIPTIVA Y CONTROL DE LECTURA**

ALUMNO: **ING. JAIRO ALEXANDER LÓPEZ HERNÁNDEZ**

DOCENTE: **DR. ENRIQUE ANTONIO PANIAGUA MOLINA**

**TAPACHULA, CHIAPAS A 21 DE SEPTIEMBRE DE 2015**

## EJERCICIOS DE ESTADÍSTICA 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:

***T'*** = U, V, W (una chimenea, cuesta más de $ 100 mil pesos, casa nueva)

***U'*** = T, V, W (Casa de 3 o más baños, cuesta más de $ 100 mil pesos, casa nueva)

***V'*** = T, U, W (Casa de 3 o más baños, una chimenea, casa nueva)

***W'*** = T, U, W (Casa de 3 o más baños, una chimenea, cuesta más de $ 100 mil pesos)

***T*** ∩ ***U*** = Vacío (No tienen nada en común)

***T*** ∩ ***V*** = Vacío (No tienen nada en común)

***U'*** ∩ ***V*** = (Cuesta más de $ 100 mil pesos)

***V U W*** = (Cuesta más de $ 100 mil pesos, casa nueva)

***V' U W*** = (Casa de 3 o más baños, una chimenea, casa nueva)

***T U U*** = (Casa de 3 o más baños, una chimenea)

***T U V*** = (Casa de 3 o más baños, cuesta más de $ 100 mil pesos)

***V*** ∩ ***W*** = Vacío (No tienen nada en común)

**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

2/9 + 1/9 + 2/9 + 1/9 + 2/9 + 1/9 = 1

9/9 = 1 1 = 1

- Esto se debe al postulado 2

  - La P(B) sería: 4/9

B = (4, 5,6)

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

***P (B) = 0.444***

**3.-** Calcula la muestra para una población desconocida con un 96% de confianza y 10% error. Para una prevalencia de 0.5 y 0.7

FÓRMULA

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

Para una prevalencia de 0.5

Donde:

$Z^2\alpha$ = 1.7506

$p$ = 0.5

$q$ = (1- p) = (1 - 0.5) = 0.5

$i$ = 0.10

Sustituyendo en la fórmula

$$n = \frac{(1.7506)^2 \; x \; (0.5) \; x \; (0.5)}{(0.10)^2}$$

n = 76.615 ***n = 77***

Para una prevalencia de 0.7

Donde:

$Z^2\alpha$ = 1.7506

$p$ = 0.7

$q$ = (1 - p) = (1 – 0.7) = 0.3

$i$ = 0.10

Sustituyendo en la fórmula

$$n = \frac{(1.7506)^2 \; x \; (0.7) \; x \; (0.3)}{(0.10)^2}$$

n = 64.357 ***n = 64***

**4.-** Calcula la muestra para una población de 350,000 familias, con un 99% de confianza y 5% error. Para una prevalencia de 0.5 y 0.7

FÓRMULA

$$n = \frac{Z_{\alpha.}^{2} N.p.q}{i^{2}(N-1) + Z_{\alpha.}^{2}.p.q}$$

Para una prevalencia de 0.5

Donde:

$Z^2\alpha$ = 2.3263

$N$ = *350,000*

$p$ = 0.5

$q$ = (1 – p) = (1 - 0.5) = 0.5

$i$ = 0.05 Sustituyendo en la fórmula

$$n = \frac{(2.3263)^2 \ x \ (350000) \ x \ (0.5) \ x \ (0.5)}{(0.05)^2 \ (350000 - 1) + (2.3263)^2 (0.5)(0.5)}$$

n = 540.333 ***n = 540***

Para una prevalencia de 0.7

Donde:

$Z^2\alpha$ = 2.3263

$N$ = *350,000*

$p$ = 0.7

$q$ = (1 – p) = (1 - 0.7) = 0.3

$i$ = 0.05 Sustituyendo en la fórmula

$$n = \frac{(2.3263)^2 \ x \ (350000) \ x \ (0.7) \ x \ (0.3)}{(0.05)^2 \ (350000 - 1) + (2.3263)^2 (0.7)(0.3)}$$

n = 453.99 ***n = 454***

**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 estándar de 1.5% con un nivel de confiabilidad del 90%.

Datos:

$N$ = 1176

$Z^2\alpha$ = 1.2815

$p$ = 0.5

$q$ = (1- p) = (1 – 0.5) = 0.5

$i$ = 0.015

i=0.015

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

Sustituyendo en la fórmula

$$n = \frac{(1.2815)^2\ x\ (1176)\ x\ (0.5)\ x\ (0.5)}{(0.015)^2\ (1176-1) + (1.2815)^2(0.5)(0.5)}$$

n = 715.39 ***n = 715***

**6.-** Con los siguientes datos

| Estaturas (x) | f |
|---|---|
| 1.52 | 1 |
| 1.54 | 5 |
| 1.55 | 4 |
| 1.58 | 5 |
| 1.6 | 2 |
| 1.62 | 4 |
| 1.64 | 7 |
| 1.66 | 3 |
| 1.7 | 5 |
| 1.71 | 8 |
| 1.73 | 6 |
| 1.74 | 5 |
| 1.77 | 3 |
| 1.8 | 1 |
| 1.83 | 1 |
| Total | 60 |

- 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 resultados

| 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 |

Desviación estándar para datos agrupados:

$$\bar{x} = \frac{\sum_{i=1}^{n} xi}{N} \qquad \bar{x} = \frac{99.66}{60} = 1.661$$

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{x_i^2 f_i}{N} - \bar{x}^2} \qquad \sigma = \sqrt[2]{\frac{165.8978}{60} - 1.661^2}$$

$$\sigma = \sqrt[2]{2.764963 - 2.75892} \qquad \sigma = \sqrt[2]{.006043}$$

$$\sigma = 0.077736$$

Varianza para datos agrupados:

$$\sigma^2 = \sum_{i=1}^{n} \frac{x_i^2}{N} - \bar{x}^2 \qquad \sigma^2 = \frac{165.8978}{60} - 1.661^2$$

$$\sigma^2 = 2.764963 - 2.758921 \qquad \sigma^2 = .006042$$

$$x = 1.7387$$

$$\sigma = 0.077736$$

$$\bar{x} = 1.6610$$

$$\sigma = -0.077736$$

$$\bar{x} = 1.5832$$

## SELECCIÓN DE LA MUESTRA

Para poder recolectar los datos que nos van a servir en nuestra investigación debemos seleccionar una muestra para hacer el estudio. La muestra no es más que un subgrupo de la población de interés la cual debe ser ser representativo de la población ya que el investigador pretende que los resultados encontrados en la muestra logren generalizarse o extrapolarse a la población, al seleccionar la muestra hay que definir la unidad de análisis (personas, organizaciones, periódico, comunidades, situaciones, eventos, etc.); lo cual depende del planteamiento del problema a investigar y de los alcances del estudio, lo que a su vez nos permite delimitar la población de estudio.

Una vez que se ha definido cual 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.

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).

La muestra que podemos utilizar en nuestra investigación puede ser: no probabilísticas o probabilísticas. En la última 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.

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 correlacionales-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. Las unidades o elementos muéstrales tendrán valores muy parecidos a los de la población, de manera que las mediciones en el subconjunto nos darán estimados precisos del conjunto mayor.

En el caso de las muestras probabilísticas puede ser: muestra aleatoria simple, muestra estratificada, por racimos o clusters. La estratificada aumenta la precisión de la muestra e implica el uso deliberado de submuestras para cada estrato o categoría que sea relevante en la población; mientras que la muestra por racismo 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.

La muestra en nuestra investigación ya sea de tipo probabilística o no probabilísticas, la utilizamos por economía en tiempo y recursos ya que se agiliza el estudio de la población y se utiliza menor recurso humano, material, financiero que si se hiciera en toda la población objetos de estudio.

*BIBLIOGRAFÍA*

*Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, P. (2006). Selección de la muestra, Metodología de la investigación (pp. 235-270) México: McGraw-Hill.*