

**ESTADISTICA ADMINISTRATIVA**

**MATERIA:**

**DR. ENRIQUE ANTONIO PANIAGUA MOLINA**

**CATEDRÁTICO:**

**3**

**ACTIVIDAD:**

**EJERCICIOS DE ESTADISTICA DESCRIPTIVA Y CONTROL DE LECTURA (CAPITULO 8).**

**TEMA:**

**LUIS GARCIA SOLIS**

**NOMBRE DEL ALUMNO:**

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

**CONTROL DE LECTURA**

**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.

Para el proceso cuantitativo la muestra es un subgrupo de la población de interés (sobre el cual se recolectarán datos, y que tiene que definirse o delimitarse de antemano con precisión), este deberá ser representativo de la población. El investigador pretende que los resultados encontrados en la muestra logren generalizarse o extrapolarse a la población.

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. La delimitación de las características de la población no solo depende de los objetivos del estudio, sino de otras razones prácticas. Un estudio no será mejor por tener una población mas grande; la calidad de un trabajo investigativo estriba en delimitar claramente la población con base en el planteamiento del problema. Las poblaciones deben situarse claramente en torno a sus características de contenido, de lugar y en el tiempo.

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). El primer paso para evitar tales errores es una adecuada delimitación del universo o población. Los criterios que cada investigador cumpla depende de sus objetivos de estudio, lo importante es establecerlos de manera mas especifica. Toda investigación debe ser transparente, asi como estar sujeta a crítica y réplica.

¿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 os 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. Elegir entre una muestra probabilística o una no probabilística depende de los objetivos del estudio, del esquema de investigación y de la contribución que se piensa hacer con ella.

¿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 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.

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 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.

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 mas 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 ESTADÍSTICAS 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 (tiene chimenea, cuesta más de cien mil, casa nueva)*

***U’****=T, V, W (tiene 3 o más baños, cuesta más de cien mil, casa nueva)*

***V’****=T, U, W (tiene 3 o más baños, tiene chimenea, casa nueva)*

***W´****= T, U, V (tiene 3 o más baños, cuesta más de cien mil, cuesta más de cien mil)*

***T ∩ U****= (el evento es nulo)*

***T ∩ V****= (el evento es nulo)*

***U’ ∩ V****= (Cuesta más de cien mil)*

***V U W****= (Cuesta más de cien mil; casa nueva)*

***V’ U W****= (tiene 3 o más baños; una chimenea; casa nueva)*

***T U U****= (tiene 3 o más baños; tiene chimenea)*

***T U V****= (tiene 3 o más baños; cuesta más de cien mil)*

***T ∩ W****=  (el evento es nulo)*

**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)

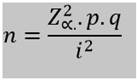
**B**= X + 2X + X

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

FÓRMULA

**Z²α**= 1.7506

**P**= .5

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

**.i**= .10

Sustituyendo en  la fórmula

**´n**=  https://lh4.googleusercontent.com/fLzcRan-p17U05H-fFV_7I8-711WtQXxpD3WoRGlpCQjo8xBBx5G6nsw4whbANo-cQMn-uktn2gkQJ4iOHEkIRmEnOArToE0ZkLo-VSgfqcQvOfbJmDexbHYq34SCHyE9A=s1600     **=** **76.61 ≈ 77**

**Para una prevalencia de 0.7**

**Z²α**= 1.7506

**P**= .7

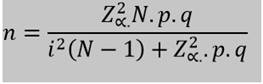
**.q**= 1-p = (1-.7)

**.i**= .10

Sustituyendo en  la fórmula

**´n**=  https://lh6.googleusercontent.com/AqaSH05Mlq0MKk_cYkHeanDppTlYW-7_Nmpq7_j_W4Haz23KSkbHJWoPX7fawtkfq61RGGEfRV7nDvpj7T6WnB5xsoTrbD7C2h-zlRAesFvY4cfXyOpP_fPp9NGAhVMDyA=s1600     **=** **64.35 ≈ 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

FÓRMULA****

**Para una prevalencia de 0.5**

**Z²α**= 2.3263

**P**= .5

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

**.i**= .05

Sustituyendo en  la fórmula

**´n=**    https://lh3.googleusercontent.com/uaBQwa70Hu3kusygWa2PK78aLMrtQAKUIVomKKuryrEGUZVETGUDpkWGX3xpWsiHiBAukkoLSSvg84KkGGexuz0W3rCqm-maTHC4GQfen-SA9gyWFpM4U0-Ddi82JgDPBns=s1600    **n=**      https://lh5.googleusercontent.com/HX9dGxEGq3lNqpylX-ZM4vGr2Or-uCfulp3YrmcOStAgn5qgI0RLauOzDk5xLgdQ2cPW4DudCc-Er6gP63X3zQ85ktZZFvnTsYWtLfo4N2iiWPjIiLk7AyxxY-cCmH5RSQ=s1600

**´n=**    **540.33** **≈ 540**

**Para una prevalencia de 0.7**

**Z²α**= 2.3263

**P**= .5

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

**.i**= .05

Sustituyendo en  la fórmula

**´n=**    https://lh5.googleusercontent.com/Ob7f843KEkBeJRunJ6U7cw_u0Zrx5OQe_OtT4d5gT3S3RrgAZg4z1byG9C8esb95EILdmywg4gfy652AQ_rm3_puQ6RSwvI1wsIlKW0h85xlq2qlINX5OstiljHFyLYt=s1600    ´**n=**      https://lh3.googleusercontent.com/cEdBGxjZPw4C9dGdNoLR3rnA6V1--d-5k79Au1k-lj8iIZcFdd7cWDPdrdp9Y1RcP74z3No-jDQgsKisRybZC15NUQUIzTi9biw2CU_KEIi6B-TSkXk9CWDFq4SJvngz=s1600

**´n=**    **453.98** **≈ 454**

**Ejercicio 5:** De una Población de 1,176 padres de familia de la Cd. 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%

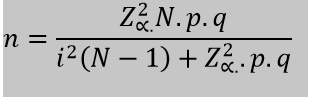
Datos:

N=1176

Z=1.2815

p=0.5

q= 0.5

i=0.015         Fórmula:       

´n= (1.2815)² (1176)(.5)(.5)             =       (1.6422)(1176)(.25)             =

  (.015)²(1176-1)+ (1.2815)²(.5) (.5)      (0.000225)(1175)+(1.6422)(.25)

´n=  482.8068           =      482.8068     =

    (0.2643)+0.2643             0.6746

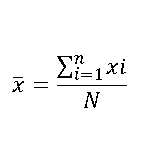
**´n=715.48 ≈ 715**

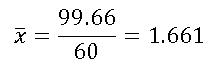
**Ejercicio 6:**

* 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 resultados.

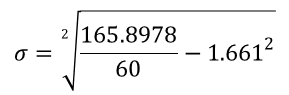
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Estaturas (X)** | **f** | **Ni** | **xi \* fi** | **xi 2 \* fi** |
| **1.52** | **1** | **1** | **1.52** | **2.3104** |
| **1.54** | **5** | **6** | **7.7** | **11.858** |
| **1.55** | **4** | **10** | **6.2** | **9.61** |
| **1.58** | **5** | **15** | **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:**

****

****

**https://lh5.googleusercontent.com/665Ws8HEgM80HwlinLX8r9uf6EZ4p0vjsrr8otFPN1RbtDjqAImueq-3ErA7CxHMlgZDg2BIkINUlzFekwXmr6VGF9_YzgqHVrVkoWzhKPUHktCsWzEm_EfjoJWabS7nRg=s1600**

****

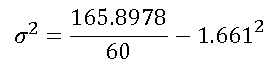
**https://lh5.googleusercontent.com/Kk8O359fxgYQ1350wGwr6Jc90caIBcKCQN0hz8f22jZ3di_aMO5wLyyGjxtFDxCyKDmaASnmM2YEN1Ng6_MQ1oaX-UCD8wtk5wj7jQs4D__Gp86fHmm-gLj5elv3f49eFQ=s1600**

**https://lh6.googleusercontent.com/vaG5BE6qRpGcaGdR22N2zlR3wdounK72hEaPPLMzJK6JjGimNHK2qQ4MV3VzQLN3rOW-UEI9LUyPxh3I7lZvrNgi5tDbVfiBEMmLgmnGVoOS30GP_UVLT_uwPhczJN4y4w=s1600**

**https://lh5.googleusercontent.com/o7woCxi5cEK_x__W0p0qmPlg3zhzUcwEPynO9lqri4yveqRK_lI7N96eieqWx8wG4kRr-raFiHmK_41M2eiee1WKN8KVQFy-TDCF_-_UsytNVm9lHIPSlLCyLSWDzwexACc=s1600**

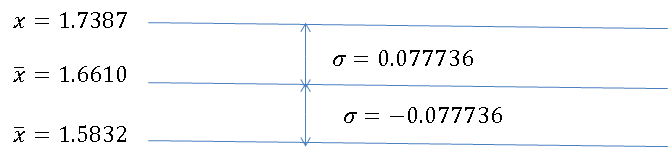
**Varianza para datos agrupados:**

**https://lh3.googleusercontent.com/DbW06MfLpUnGZS9cL_plMLSQ-0znsK1sL4iXtZw1Lyg_VV64aE7ftC5zfhxwznoaXPz2f3jiAVO49t2lo8UGOhYzOGxGZLcrDo_M35e-RjTi8xsIoAK7KwS1evjD1oFw2A=s1600**

****

**https://lh3.googleusercontent.com/Hdk7SJ5tjbWxdLr38a0AcBFow_-z2umphelaywV13cZAVXX5g1Kv4YcNkiqtNqoPIdrgIlspJNFpKAb1VtUK185OwP5Fmj2BMVzgGXlbBveO0w1gWZEYwYlPMDbmvBIsIw=s1600**

**https://lh5.googleusercontent.com/6leq0oFCBk7a6-yW-S4GfNaQXc70zu_06OSgoP-PBvYe-pOSoSB9Bx_-4U5Hf9iE4IopokURqneIj4drvMPfLAkR6KrslZ91dvjXMFV2sE41wuhsUrf_v-obFXcjsZTI5jg=s1600**

****