## 1. INTRODUCCIÓN AL MUESTREO

### DEFINICIONES BÁSICAS

Se llama **UNIVERSO** al conjunto de Unidades Experimentales que poseen características comunes observables, y que se utilizan para obtener información sobre un hecho particular

El **Universo** queda determinado cuando se definen los objetivos del trabajo que se llevará a cabo y a quiénes se les realizarán las observaciones.

Los **Universos** pueden ser infinitos o finitos, dependiendo de la cantidad de elementos que los conforman y, consecuentemente las **Poblaciones** serán infinitas o finitas.

Se llama **POBLACIÓN** a cualquier variable particular que se estudia a un Universo

Cada Universo puede generar varias Poblaciones. Al establecer cuáles serán las variables que se le observarán a cada una de las Unidades Experimentales que conforman el Universo, quedan determinadas las distintas Poblaciones.

**N**: tamaño del Universo o de la Población.

Si fuese necesario tener información precisa sobre la totalidad de los elementos que forman el Universo, entonces se debería realizar un censo, pero si el tamaño del Universo es demasiado grande, la tarea censual resulta muy onerosa. Por otro lado la ejecución de un censo es imposible en aquellos casos donde hay que enfrentar a un Universo infinito, o cuando el proceso de medida o investigación de las características de cada elemento es destructivo. Por estos motivos, hay que utilizar otras técnicas estadísticas a los efectos de obtener datos útiles y confiables, necesarios para tener información acerca de la Población.

Una de esas técnicas es observar y medir a algunos elementos del Universo y con ellos poder conseguir la información necesaria para cumplir con los objetivos fijados para concretar el trabajo.

Se llama **MUESTRA** a un subconjunto o parte de una Población tomada de forma tal que con ella se pueda hacer un juicio acerca de esa Población completa

La Muestra también será un subconjunto del Universo.

**n**. tamaño de la Muestra.

Se llama **FRACCIÓN DE MUESTREO** al cociente entre el tamaño de la Muestra y el tamaño de la Población.

$$Fm = \frac{n}{N}$$

**INFERENCIA ESTADÍSTICA**

Se llama **INFERENCIA ESTADÍSTICA** a cualquier afirmación que se realiza sobre una determinada población basándose en los datos obtenidos con una muestra, pudiéndose tener, a partir del cálculo de probabilidad, una determinada medida de la incertidumbre que se genera.

Se **infiere** la población a partir de la muestra.

**MUESTREO**

Se llama **MUESTREO** al procedimiento mediante el cual se obtienen una o más muestras

Se llama **UNIDAD DE MUESTREO** a una unidad experimental o a un grupo de unidades experimentales que son tomadas para obtener una muestra

- MUESTREO PROBABILÍSTICO:

Cuando las Unidades Experimentales que componen la muestra son tomadas al azar, pudiéndose calcular a priori la probabilidad que tiene cada muestra y como consecuencia cada Unidad Experimental, de ser la obtenida.

El proceso de obtención de cada uno de los elementos que integrarán la muestra debe ser considerado como un experimento **aleatorio**.

Es considerado **objetivo** este muestreo porque, como la obtención de las Unidades Experimentales es realizada al azar, la inclusión de cada una de ellas en la muestra, no depende del sujeto que se encarga de realizar la muestra.

· MUESTREO INTENCIONAL:

**Cuando las Unidades Experimentales que componen la muestra son obtenidas siguiendo una regla o norma preestablecida.**

Cada elemento que integrará la muestra será "elegido" por el sujeto que realiza el trabajo, de acuerdo a un criterio que se fija previamente. Es un muestreo **subjetivo**.

La representatividad de la muestra depende de la intención de quién la realiza.

· MUESTREO SIN NORMA:

Cuando por razones de comodidad, costo o cualquier otra circunstancia, la obtención de las Unidades Experimentales que componen la muestra se realiza **sin una norma**, o regla o criterio definido. A cada elemento de la muestra los elige un sujeto pero sin un criterio fijo, por esto se considera que la obtención es **cuasi aleatoria**, luego el muestreo es **cuasi-objetivo**.

Este se puede utilizar cuando hay elementos de juicio suficientes como para suponer que la población es homogénea.

MÉTODOS DE OBTENCIÓN DE MUESTRA

Se estudiará la aplicación de los Métodos Estadísticos en el Muestreo PROBABILÍSTICO, por lo tanto, se supone que cada elemento que interviene en la muestra es tomado al azar. Para garantizar que la obtención de cada uno de ellos es realmente aleatoria, es conveniente la utilización de un bolillero o de una Tabla de Dígitos de Azar.

·MUESTREO SIMPLE AL AZAR (SIN REEMPLAZO):

Consiste en obtener al azar una muestra de **n** elementos, de entre los **N** que constituyen el Universo, de modo tal que todas las muestras posibles de tamaño **n** tengan la misma probabilidad de ser tomadas, como así también, que todos los elementos que integrarán la muestra tengan, en el momento de cada extracción, la misma probabilidad de ser obtenidas.

Ejemplo: un Universo finito de tamaño **N.**

La probabilidad de que una muestra particular sea la obtenida, es de acuerdo a la definición clásica:

$$\frac{1}{N}$$

n.

Sobre la base de esto, por ser equiprobables todas las muestras del tamaño **n**, la probabilidad de que un elemento cualquiera de la población forme parte de la muestra se puede calcular haciendo:

por lo tanto todos los elementos tienen la misma probabilidad de ser extraídos.

Si la población es infinita y las variables muestrales son independientes, las probabilidades de que cada de una de ellas asuma un determinado valor, es la misma que para cualquier valor que asuman las otras. Caso contrario, si hay dependencia estadística, dicha probabilidad dependerá del valor de las que ya fueron obtenidas.

Dado que las Unidades Experimentales son tomadas al azar, cualquiera sea la posición que ocupe en el Universo, para que se cumpla la condición de igualdad de probabilidad para todas ellas, el Muestreo Simple al Azar (MSA) debería ser utilizado solamente cuando se tiene real evidencia de la homogeneidad de la población de la cual se tomará la muestra.

·MUESTREO ESTRATIFICADO AL AZAR:

En aquellos casos en que la población es heterogénea, la gran variabilidad que presentan los valores de la variable en estudio hace que la Utilización de Muestreo Simple al Azar pueda proporcionar muestras no representativas y las conclusiones que surjan de ellas no serán del todo confiables. En estos casos el método de muestreo mas adecuado a utilizar es el **MUESTREO ESTRATIFICADO AL AZAR**, que consiste en particionar al Universo en estratos (o clases o subpoblaciones), dentro de los cuales sí la variable debe presentar homogeneidad, y de cada uno ellos se obtiene una muestra Simple al Azar.

Los símbolos a utilizar en este trabajo son:

**N** : tamaño del Universo

**h**.: cantidad de estratos

**Nj**: tamaño del j-ésimo estrato

La asignación del tamaño de la muestra a cada uno de los distintos estratos se llama AFIJACIÓN, y puede realizarse de alguna de las siguientes maneras:

a) Afijación igual o uniforme: el tamaño de muestra que le corresponde a cada estrato es igual para todos. Este tamaño se calcula haciendo el cociente entre el tamaño de la muestra, **n**, y la cantidad de estratos **h**.

b) Afijación proporcional: el tamaño de la muestra que le corresponde a cada estrato es proporcional al tamaño del estrato. Se calcula haciendo el producto entre la fracción de muestreo y el tamaño de cada estrato.

c) Afijación óptima: el tamaño de la muestra para cada estrato es proporcional al tamaño del estrato y al desvío estándar correspondiente. De esta manera se tiene en cuenta la falta de homogeneidad entre las subpoblaciones.

–

–

CUANDO SE TOMAN MUESTRAS UTILIZANDO EL MÉTODO DE **MUESTREO SIMPLE AL AZAR** O EL MÉTODO **MUESTREO ESTRATIFICADO AL AZAR**, LAS UNIDADES DE MUESTREO COINCIDEN CON LAS UNIDADES EXPERIMENTALES.

EL MUESTREO POR **CONGLOMERADOS POLIETÁPICO** TIENE COMO PARTICULARIDAD QUE LA UNIDAD DE MUESTREO ESTÁ FORMADA POR UN GRUPO DE UNIDADES EXPERIMENTALES.

·MUESTREO POR CONGLOMERADOS POLIETÁPICO:

Consiste en agrupar los elementos que conforman el Universo en conglomerados, de modo tal que entre ellos haya la suficiente homogeneidad como para representar una población. El proceso de obtención de la muestra se hace sobre la base de estos conglomerados, que constituyen las unidades de muestreo, y no sobre la base de unidades físicas, con excepción de la ultima etapa.

Es adecuado cuando las unidades experimentales están agrupadas en conglomerados y cada una de ellas, puede ser considerada como unidad de muestreo.

Si es de una etapa: se toma una muestra simple al azar de conglomerados y en cada uno de ellos se miden todos los elementos que constituyen el conglomerado.

Si es de dos etapas de muestreo o es bietapico: se toma una muestra simple al azar de conglomerados (primera etapa de muestreo) y de cada uno de estos conglomerados, se toma una muestra de unidades experimentales (segunda etapa de muestreo).

Si es de tres etapas o trietápico: primero se constituyen los conglomerados, y dentro de ellos se forman los subconglomerados, luego se toma una muestra simple al azar de conglomerados (primera etapa de muestreo), de cada uno de ellos una muestra simple al azar de subconglomerados (segunda etapa de muestreo), y por último de cada subconglomerado, una muestra simple al azar de unidades físicas (tercera etapa de muestreo).

·MUESTREO SISTEMÁTICO AL AZAR:

Consiste en ordenar a las **N** Unidades Experimentales que conforman el Universo, de acuerdo a como se fueron presentando, y obtener la muestra eligiendo, sistemáticamente ( de aquí el nombre del método), un elemento cada **c** unidades, tomado el primero de ello en forma aleatoria.

Si el universo es finito el número **c** es la parte entera del cociente entre el tamaño del universo y el tamaño de la muestra.

Si el universo es infinito el número **c** se elige arbitrariamente sobre la base de un buen saber y entender del estadístico que realiza el trabajo.

Este método es adecuado en el caso en las unidades experimentales que forman el universo se presentan con una determinada periodicidad. En este caso hay que evitar que el número **c** sea igual al período con que se presentan las unidades experimentales en el Universo, porque, si ello ocurriese perdería representatividad en la muestra.

**CARACTERÍSTICAS POBLACIONALES Y MUESTRALES**

Los conceptos que se desarrollan a continuación están referidos únicamente a las muestras que se obtienen mediante el método de Muestreo Simple al Azar.

PARÁMETRO ESTADÍSTICO

UNIVERSO FINITO Y PEQUEÑO

Si se tiene un Universo finito y lo suficientemente pequeño como para que el comportamiento probabilistico de las poblaciones no esté explicado por un modelo teórico especifico, se define:

Se llama **PARÁMETRO ESTADÍSTICO** a toda medida que resume información calculada con las variables poblacionales

Si no hay posibilidad de una equivocada interpretación, a los **PARÁMETROS ESTADÍSTICOS** se los denomina simplemente **PARÁMETROS**, y serán simbolizados con letras griegas.

Ejemplos de Parámetros: el total de elementos que presenta un determinado atributo, la proporción de elementos que presenta un determinado atributo, la media aritmética, la varianza, el desvío estándar, etc.

·Cuando se quiere representar un parámetro indefinido se utilizará la letra griega, Theta **q**.

·Para la media de población o media poblacional se utilizará la letra griega Mu **m**.

· Para la varianza de la población o varianza poblacional, se utilizará la letra griega sigma **s**, o con el supra índice 2 (Sigma cuadrado).

· Para el total de elementos con un determinado atributo en el Universo se utilizará la letra **Â.**

· Para la proporción de elementos que tienen un determinado atributo en el Universo, se utilizará la letra griega Pi, **p.**

UNIVERSO FINITO Y GRANDE O UNIVERSO INFINITO

Si la población es finita y lo suficientemente grande o infinita, como para que su comportamiento probabilístico esté explicado por una función de probabilidad, si la variable es discreta, o por una función de densidad de probabilidad, si se trata de una variable continua, entonces se define:

Se llama **PARÁMETRO ESTADÍSTICO** a todo parámetro matemático, o medida capaz de individualizar a una Población, dentro de una familia de Poblaciones con la misma función de densidad de probabilidad,

en caso de variables continuas, o la misma función de probabilidad en caso de variables discretas

Los PARÁMETROS no son constantes matemáticas, sino que pueden tomar cualquier valor que se encuentre dentro de un conjunto específico llamado: **ESPACIO PARAMÉTRICO** y simbolizado con la letra griega omega W.

**PROPOSICIÓN 1: LAS OPERACIONES ALGEBRAICAS REALIZADAS CON PARÁMETROS SON PARÁMETROS.**

ESTADÍGRAFO Y ESTIMADOR

–

Se llama **ESTADÍGRAFO** a toda función escalar, **h (x1, x2, x3...)**, generada con las variables muestrales

Dado que los estadígrafos son funciones generadas por variables aleatorias, también son variables aleatorias, luego, existirá una función de densidad de probabilidad o una función de probabilidad, lo que corresponda, que describa su comportamiento PROBABILÍSTICO, como así también es posible que tengan una Esperanza Matemática finita y una Varianza finita.

Cumplen distintos roles dentro del análisis inferencial.

Se llama **ESTIMADOR DE UN PARÁMETRO q** a todo **ESTADÍGRAFO** que proporcione información acerca de dicho parámetro

Se llama **ESTADÍGRAFO DE TRANSFORMACIÓN** a aquel estadígrafo que permite transformar al estimador en una variable que tenga una determinada distribución de probabilidad

ESTIMACIÓN

Dado que en la mayoría de los trabajos en donde se aplica el análisis estadístico los parámetros de las poblaciones son desconocidos, hay que llevar a cano el proceso de inferir o sacar conclusiones acerca de éstos a través de las variables muestrales.

Los métodos que se utilizan para ello son dos, que generalmente se complementan, a saber:

a. ESTIMACIÓN PUNTUAL

b. ESTIMACIÓN POR INTERVALO

Se llama **ESTIMACIÓN PUNTUAL** del parámetro q a un método de estimación que consiste en calcular el valor numérico único que asume el estimador luego de tomar la muestra y realizar las mediciones correspondientes. Este valor numérico se llama **PUNTO DE ESTIMACIÓN**.

Se llama **ESTIMACIÓN POR INTERVALO** del parámetro q a un método de estimación que consiste en calcular, con los datos de la muestra, los límites de un conjunto acotado de números reales, llamado **INTERVALO DE ESTIMACIÓN**.

DISTRIBUCIÓN DE LOS ESTIMADORES

Se llama **DISTRIBUCIÓN DE PROBABILIDAD DEL ESTIMADOR q^** a aquella función de densidad de probabilidad, o función de probabilidad, según corresponda, que describe su comportamiento probabilístico

Si se considera al muestreo probabilístico como un experimento aleatorio, y, si para cada muestra posible se calculase el valor numérico del Estimador, estos valores, acompañados del correspondiente valor de probabilidad, o densidad de probabilidad, constituyen la **DISTRIBUCIÓN DE PROBABILIDAD DEL ESTIMADOR.**

SESGO

Se llama **SESGO** a la diferencia entre, la esperanza matemática del estimador y el parámetro a estimar

$$\textbf{SESGO} = E(q) - q$$

ERROR MEDIO CUADRÁTICO

Se llama **ERROR MEDIO CUADRÁTICO**, **(EMC),** a la esperanza matemática del cuadrado de la diferencia entre el estimador y el parámetro.

**EMC ( q ) = E (q - q)**

El **ERROR MEDIO CUADRÁTICO** mide la variabilidad del estimador con respecto al parámetro que está estimando.

## PROPIEDADES DE LOS BUENOS ESTIMADORES

· Insesgamiento.

· Eficiencia.

· Consistencia.

· Suficiencia.

ESTIMADOR INSESGADO

El estimador q es un ESTIMADOR INSESGADO del parámetro q sí, y sólo si la esperanza matemática del estimador q es igual al parámetro q.

**E (q) = q**

Si se pudiesen calcular todos los valores numéricos de un determinado estimador y se calculase el promedio aritmético de todos ellos, y dicho promedio resultase igual al parámetro, entonces el estimador sería INSESGADO, en caso contrario el estimador sería SESGADO.

En algunas ocasiones esta propiedad puede no cumplirse directamente, pero sí cuando el tamaño de la muestra crece indefinidamente (tiende a infinito). Cuando se presenta esta situación, para definir la PROPIEDAD DE INSESGAMIENTO, se utiliza un concepto limite:

El estimador q es un estimador Asintóticamente Insesgado del parámetro q, sí, y sólo si el límite de la esperanza matemática del estimador q, cuando el tamaño de la muestra **n** tiende a infinito, es igual al parámetro q.

SI EL ESTIMADOR ES INSESAGADO, O ASINTÓTICAMENTE INSESGADO, EL ERROR MEDIO CUADRÁTICO ES IGUAL A LA VARIANZA DEL ESTIMADOR, DIRECTAMENTE O EN EL LÍMITE, CUANDO **n** TIENDE A INFINITO.

CONSISTENCIA

El estimador q es un ESTIMADOR CONSISTENTE del parámetro q sí y sólo, si se cumple que: el estimador q converge en probabilidad al parámetro q, cuando el tamaño de la muestra **n** tiende a infinito.

Un Estimador es Consistente si, a medida que el tamaño de la muestra crece indefinidamente (tiende a infinito), la probabilidad de que la diferencia entre el estimador y el valor del parámetro pueda hacerse tan pequeña como se quiera, tiende a la unidad.

EFICIENCIA

El estimador q es un ESTIMADOR EFICIENTE del parámetro q sí, y sólo si se cumple que el estimador q tiene la menor varianza que puede tener un estimador del parámetro q.

SUFICIENCIA

El estimador q es un ESTIMADOR SUFICIENTE sí, y sólo si se cumple que: el estimador q utiliza toda la información relevante, contenida en la muestra aleatoria, con respecto al parámetro q.

La importancia de esta propiedad reside en que, si un estimador insesgado del parámetro q es un estimador suficiente, entonces tendrá la varianza menor que aquel estimador insesgado no suficiente de dicho parámetro.

**GRADOS DE LIBERTAD**

Se llama **GRADOS DE LIBERTAD** a la cantidad de variables libres, o estadísticamente independientes, que intervienen en un problema o en una distribución asociada a un problema

**PROPOSICIÓN 2: LA CANTIDAD DE GRADOS DE LIBERTAD QUE TIENE LA SUMA DEL CUADRADO DE LAS DESVIACIONES CON RESPECTO A LA MEDIA ARITMÉTICA MUESTRAL ES IGUAL AL TAMAÑO DE LA MUESTRA MENOS UNO**

**ALGUNOS ESTIMADORES IMPORTANTES**

MEDIA ARITMÉTICA MUESTRAL

Se llama **MEDIA ARITMÉTICA MUESTRAL** (media muestral) a un estimador de la media poblacional, que se genera haciendo el cociente entre la suma de las variables muestrales y el tamaño de la muestra.

VARIANZA MUESTRAL

Se llama **VARIANZA MUESTRAL** a un estimador de la varianza poblacional que se genera haciendo el cociente entre la suma del cuadrado de las desviaciones con respecto a la media aritmética muestral y los grados de libertad

PROPORCIÓN MUESTRAL

Se llama **PROPORCIÓN MUESTRAL O PROPORCIÓN DE ELEMENTOS CON UN DETERMINADO ATRIBUTO EN LA MUESTRA** a un estimador de la proporción poblacional que se genera mediante el cociente entre la cantidad de elementos que si poseen un determinado atributo en la muestra y el tamaño de la muestra.

ESPERANZA Y VARIANZA DE LOS ESTIMADORES MEDIA MUESTRAL, PROPORCIÓN MUESTRAL Y VARIANZA MUESTRAL

MEDIA MUESTRAL

· La esperanza matemática de la MEDIA MUESTRAL siempre es la MEDIA POBLACIONAL

Por lo tanto la MEDIA MUESTRAL es un estimador INSESGADO de la MEDIA POBLACIONAL.

· La varianza de la MEDIA MUESTRAL de universos infinitos es

Por lo tanto si el universo es infinito, la MEDIA MUESTRAL es un estimador CONSISTENTE de la MEDIA POBLACIONAL.

· La varianza de la MEDIA MUESTRAL de universos finitos es

Por lo tanto, si el universo es finito, la MEDIA MUESTRAL es un estimador CONSISTENTE de la MEDIA

POBLACIONAL.

PROPORCIÓN MUESTRAL

· La esperanza matemática de la PROPORCIÓN MUESTRAL siempre es la PROPORCIÓN POBLACIONAL

Por lo tanto la PROPORCIÓN MUESTRAL es un estimador INSESGADO de la PROPORCIÓN POBLACIONAL

· La varianza de la PROPORCIÓN MUESTRAL de universos infinitos es

Por lo tanto si el universo es infinito, la PROPORCIÓN MUESTRAL es un estimador CONSISTENTE de la PROPORCIÓN POBLACIONAL.

· La varianza de la PROPORCIÓN MUESTRAL de universos finitos es

Por lo tanto, si el universo es finito, la PROPORCIÓN MUESTRAL es un estimador CONSISTENTE de la PROPORCIÓN POBLACIONAL.

VARIANZA MUESTRAL

· La esperanza matemática de la VARIANZA MUESTRAL en universos infinitos, siempre es la VARIANZA POBLACIONAL.

Por lo tanto, la VARIANZA MUESTRAL es un estimador INSESGADO de la VARIANZA POBLACIONAL.

· La VARIANZA MUESTRAL de universos finitos es

**DISTRIBUCIÓN DE ALGUNOS ESTIMADORES USUALES**

DISTRIBUCIÓN DE LA MEDIA MUESTRAL DE POBLACIONES NORMALES

**PROPOSICIÓN 3: EL ESTIMADOR MEDIA ARITMÉTICA MUESTRAL, POR ESTAR ORIGINADO EN UNA SUMA DE VARIABLES NORMALES, ES UNA COMBINACIÓN LINEAL DE VARIABLES ALEATORIAS NORMALES, POR LO TANTO TAMBIÉN TIENE DISTRIBUCIÓN NORMAL**

La esperanza matemática de la media aritmética muestral es:

Y la **VARIANZA** de la media aritmética muestral, si la población es **FINITA** es:

Y la **VARIANZA** de la media aritmética muestral, si la población es **INFINITA** es:

Por lo tanto, la **VARIABLE ESTANDARIZADA**, o **ESTADÍGRAFO DE TRANSFORMACIÓN** de la media aritmética muestral, para **POBLACIONES FINITAS** es:

Y para **POBLACIONES INFINITAS** es:

A este último estadígrafo también se lo puede escribir:

Esto significa que para el cálculo de probabilidad, y para toda la tarea inferencial donde se utilice la media aritmética muestral, de muestras obtenidas de poblaciones normales cuya varianza sea conocida, hay que utilizar la **distribución normal**.

PAG 48/49/50/51 LIBRO.

**FRACCIÓN DE MUESTREO**

Se llama **FRACCIÓN DE MUESTREO** al cociente entre el tamaño de la muestra y el tamaño de la población

La **FRACCIÓN DE MUESTREO** mide la proporción del tamaño de la muestra con respecto al tamaño de la población.

Si la fracción de muestreo es inferior a 0,10 la población se considera infinita, y al factor de la varianza y de la media muestral y de la proporción muestral $\frac{N-n}{N-1}$ se lo puede considerar igual a 1.

$$\text{Si } \mathbf{Fm} \leq 0{,}10 \Rightarrow \frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-1} = 1$$

**2.INTERVALOS DE CONFIANZA**

**DEFINICIÓN**

Se llama **INTERVALO DE CONFIANZA PARA EL PARÁMETRO q** a un conjunto cerrado y acotado de posibles valores del parámetro, cuyos límites, inferior y superior, son funciones del estimador, y la correspondiente probabilidad de que dicho intervalo cubra al verdadero valor del parámetro

La probabilidad recibe le nombre de **NIVEL DE CONFIANZA** o **PROBABILIDAD FIDUCIAL**.

Los límites del intervalo son variables aleatorias, por ser funciones de variables aleatorias, por lo tanto el intervalo es aleatorio, se desplaza por el eje de números reales. Luego, cuando sus límites asumen un determinado valor, el intervalo puede estar cubriendo o no, al verdadero del parámetro que se quiere estimar. El nivel de confianza, es la probabilidad de que el intervalo cubra a dicho valor

**1- e** : **Nivel de confianza**: probabilidad de que el intervalo **cubra** al verdadero valor del parámetro.

**e**: **Nivel de riesgo**: probabilidad de que el intervalo **no cubra** al verdadero valor del parámetro.

**INTERVALOS DE CONFIANZA PARA LA MEDIA POBLACIONAL Y PARA LA PROPORCIÓN POBLACIONAL**

Los intervalos de confianza para la media poblacional y la proporción poblacional son de tipo aditivo, esto

significa que al valor del estimador se le suma y resta una cantidad que depende del nivel de confianza elegido para realizar la estimación.

A este tipo de intervalos se los puede escribir, genéricamente,

O, expresado como intervalo

Donde **h** se llama **MULTIPLICADOR o FACTOR DE CONFIANZA** y es el valor del fractil 1 - e de la distribución

2

de probabilidad asociada al estimador q.

La cantidad que se suma y se resta al estimador se llama **Error de Muestreo**.

En los intervalos aditivos, el estimador es igual a la semisuma de los limites del intervalo y el error de muestreo es igual a la semidiferencia de los límites del intervalo.

INTERVALOS DE CONFIANZA PARA LA MEDIA POBLACIONAL DE POBLACIONES NORMALES

VARIANZA POBLACIONAL CONOCIDA. POBLACIONES INFINITAS.

Donde el **MULTIPLICADOR o FACTOR DE CONFIANZA Z**0, es el valor de abscisa de distribución normal estandarizada hasta donde se acumula una probabilidad igual a 1 - e

2

VARIANZA POBLACIONAL DESCONOCIDA. POBLACIONES INFINITAS

Donde el **MULTIPLICADOR o FACTOR DE CONFIANZA t**0, es el valor de abscisa de la distribución "t" de Student con (**n** – 1) g.1 hasta donde se acumula una probabilidad igual a 1 - e

2

VARIANZA POBLACIONAL DESCONOCIDA. POBLACIONES FINITAS.

Donde el **MULTIPLICADOR O FACTOR DE CONFIANZA t**0, es el valor de abscisa de la distribución "t" de **Student** con (**n** – 1) g.1 hasta donde se acumula una probabilidad igual a 1 - e

2

INTERVALO DE CONFIANZA PARA LA MEDIA POBLACIONAL DE POBLACIONES CUYA DISTRIBUCIÓN DE PROBABILIDAD ES DESCONOCIDA.

–

MUESTRAS GRANDES

Si la distribución de probabilidad de la población no es conocida, y el tamaño de la muestra es lo suficientemente grande como para que se cumpla el Teorema Central del Límite, entonces para poblaciones infinitas o finitas, respectivamente

O si la varianza poblacional es desconocida, para poblaciones infinitas o finitas respectivamente

Donde el **MULTIPLICADOR O FACTOR DE CONFIANZA Z**0, es el valor de abscisa de la ditribucion normal estandarizada hasta donde se acumula una probabilidad igual a 1 - e

2

MUESTRAS CHICAS

Si la distribución de probabilidad de la población no es conocida, y el tamaño de la muestra no es lo suficientemente grande como para que se cumpla el Teorema Central del Límite, entonces el nivel de confianza del intervalo de estimación es la cota inferior de probabilidad que surge de la aplicación del teorema de Tchebycheff.

Entonces para poblaciones con varianza poblacional conocida, infinitas o finitas, los intervalos de confianza son respectivamente

O si la varianza poblacional es desconocida para poblaciones infinitas o finitas son respectivamente

Donde el **MULTIPLICADOR o FACTOR DE CONFIANZA k**o es el valor que surge de la aplicación del mencionado Teorema

INTERVALO DE CONFIANZA PARA LA PROPORCIÓN DE ELEMENTOS CON UN DETERMINADO ATRIBUTO EN EL UNIVERSO

–

PROPORCIÓN POBLACIONAL DESCONOCIDA. UNIVERSOS INFINITOS

Donde el **MULTIPLICADOR O FACTOR DE CONFIANZA z**o es el valor de abscisa de la distribución normal estandarizada hasta donde se acumula una probabilidad igual a 1 - e

2

PROPORCIÓN POBLACIONAL DESCONOCIDA. UNIVERSOS FINITOS

Donde el **MULTIPLICADOR o FACTOR DE CONFIANZA Z**o es el valor de abscisa de la distribución normal estandarizada hasta donde se acumula una probabilidad igual a 1 - e

2

INTERVALO DE CONFIANZA PARA LA VARIANZA POBLACIONAL

POBLACIÓN INFINITA

–

Donde:

A: valor de abscisa de la distribución **ji** – cuadrado con (**n** – 1) grados de libertad, hasta donde se acumula una probabilidad igual a e

2

B: valor de abscisa de la distribución **ji** – cuadrado con (**n** – 1) grados de libertad, hasta donde se acumula una probabilidad igual a 1 - e

2

TAMAÑO DE MUESTRA

TAMAÑO DE MUESTRA PARA ESTIMAR LA MEDIA POBLACIONAL

Si la población es infinita el error de muestreo es

Por lo tanto,

Si la población es finita, el error de muestreo es

Por lo tanto,

**TAMAÑO DE MUESTRA PARA ESTIMAR LA PROPORCIÓN POBLACIONAL**

–

Si la población es infinita, el error de muestreo es

Por lo tanto,

Si la población es finita, el error de muestreo es

Por lo tanto,

## 3.PRUEBA DE HIPÓTESIS

DEFINICIONES BÁSICAS

Se llama **HIPÓTESIS ESTADÍSTICA** a cualquier afirmación o aseveración que se formula acerca de cualquier característica poblacional (el valor numérico de un parámetro, la forma funcional de una población, etcétera)

Se llama **HIPÓTESIS PARAMÉTRICA** a aquella **HIPÓTESIS ESTADÍSTICA** planteada para controlar o verificar el valor numérico de un parámetro

Se consideran sólo tres posibles situaciones del valor numérico del parámetro, a saber:

· El valor numérico del parámetro **q** es exactamente igual a un determinado valor postulado qo.

· El valor numérico del parámetro **q** es menor a un determinado valor postulado qo.

· El valor numérico del parámetro **q** es mayor a un determinado valor postulado qo.

De acuerdo con la situación que se trate, o sea, de acuerdo al o a los posibles valores del parámetro, la persona que ha encomendado una tarea estadística realizará una determinada acción.

Se llama **CURSO DE ACCIÓN** a la acción que se llevaría a cabo, si se conociese el verdadero valor del parámetro **q**

Se llama **DESIGUALDAD EQUIVALENTE A LA IGUALDAD** a aquella desigualdad entre el parámetro q y el valor postulado qo, que provoca el mismo **CURSO DE ACCIÓN** que se llevaría a cabo con la igualdad entre el valor del parámetro q y el valor del postulado qo.

Se llama **DESIGUALDAD NO EQUIVALENTE A LA IGUALDAD** a aquella desigualdad entre el parámetro q y el valor del postulado qo, que provoca un **CURSO DE ACCIÓN** distinto al que se llevaría a cabo con la igualdad entre el valor numérico del parámetro q y el valor del postulado qo.

Se llama **HIPÓTESIS NULA** a aquella hipótesis que establece que la diferencia entre el verdadero valor del parámetro q, y el valor que se postula qo, es cero

La hipótesis nula se simboliza **Ho**.

**La hipótesis nula debe plantearse como la igualdad entre el valor del parámetro y el valor postulado**.

$$Ho : q = qo$$

A los fines prácticos, esta igualdad puede estar acompañada o no por alguna de las dos desigualdades, según sea el **CURSO DE ACCIÓN** a seguir y la existencia o no de alguna **DESIGUALDAD EQUIVALENTE**.

Sobre la base de esto se puede distinguir dos tipos de hipótesis nulas:

· **HIPÓTESIS NULA ÚNICA**: una **HIPÓTESIS NULA** es **ÚNICA** cuando no hay **DESIGUALDAD EQUIVALENTE**.

En este caso se postula un único valor posible para el parámetro q. Si el único valor postulado del parámetro q se simboliza qo, entonces el planteo de la **HIPÓTESIS NULA ÚNICA** es:

**Ho: q = qo**

EL VALOR DEL PARÁMETRO q ES IGUAL AL VALOR POSTULADO q

Tambien puede expresarse:

**Ho: q - qo = 0**

LA DIFERENCIA ENTRE EL VERDADERO VALOR DEL PARÁMETRO q Y EL VALOR POSTULADO qO, ES CERO.

· **HIPÓTESIS NULA MÚLTIPLE**: una **HIPÓTESIS NULA** es **MÚLTIPLE** cuando hay una **DESIGUALDAD EQUIVALENTE**.

En este caso se postula un conjunto semicerrado de posibles valores del parámetro q. Si hay una desigualdad equivalente, ésta debe acompañar a la igualdad porque ambas provocan el mismo curso de acción; en este caso, se pueden diferenciar dos planteos de **HIPÓTESIS NULA MÚLTIPLE**, a saber:

* Si la **DESIGUALDAD EQUIVALENTE** es la desigualdad menor, entonces la **HIPÓTESIS NULA MÚLTIPLE** es:

**Ho : q £ qo**

Y se lee EL VALOR DEL PARÁMETRO q ES IGUAL O MENOR AL VALOR POSTULADO qO.

Tambien puede expresarse

**Ho : q - qo £ qo**

Y se lee: LA DIFERENCIA ENTRE EL VERDADERO VALOR DEL PARÁMETRO q Y EL VALOR POSTULADO qO, ES MENOR O IGUAL A CERO.

* Si la **DESIGUALDAD EQUIVALENTE** es la desigualdad mayor, entonces la **HIPÓTESIS NULA MÚLTIPLE** es

**Ho : q ³ qo**

Y se lee EL VALOR DEL PARÁMETRO q ES IGUAL O MAYOR AL VALOR POSTULADO qO.

Tambien puede expresarse

**Ho : q - qo ³ qo**

Y se lee: LA DIFERENCIA ENTRE EL VERDADERO VALOR DEL PARÁMETRO q Y EL VALOR POSTULADO qO, ES MAYOR O IGUAL A CERO.

**HIPÓTESIS ALTERNATIVA**: se llama **HIPÓTESIS ALTERNATIVA** a aquella hipótesis que debería cumplirse si la **HIPÓTESIS NULA** no es cierta.

La **HIPÓTESIS ALTERNATIVA** se simboliza **H1**

Según los objetivos de trabajo, la HIPÓTESIS ALTERNATIVA puede ser de dos tipos:

* **HIPÓTESIS ALTERNATIVA ÚNICA**: una **HIPÓTESIS ALTERNATIVA** es **ÚNICA** cuando hay un solo valor alternativo del parámetro q, o el q1, que debería ser en el caso de que la **HIPÓTESIS NULA** no sea cierta.

La expresión: **H1 : q = q1**

Se lee: SI LA HIPÓTESIS NULA NO ES CIERTA, ENTONCES, EL VALOR DEL PARÁMETRO $\theta$ DEBERÍA SER IGUAL A $\theta_1$.

* **HIPÓTESIS ALTERNATIVA MÚLTIPLE**: una **HIPÓTESIS ALTERNATIVA** es **MÚLTIPLE** cuando hay un conjunto abierto de posibles valores alternativos del parámetro $\theta$ , en caso de que se rechace la **HIPÓTESIS NULA.**

De acuerdo con el tipo de **HIPÓTESIS NULA** que se plantea, se distinguen tres formas mutuamente excluyentes de plantear la **HIPÓTESIS ALTERNATIVA MÚLTIPLE**

1. Si la HIPÓTESIS NULA es ÚNICA, o sea, si no hay DESIGUALDAD EQUIVALENTE

$$H_0 : \theta = \theta_0$$

Entonces se plantea la siguiente HIPÓTESIS ALTERNATIVA MÚLTIPLE

$$H_1 : \theta \neq \theta_0$$

A esta forma de plantear la HIPÓTESIS ALTERNATIVA ÚNICA se la interpreta:

SI LA HIPÓTESIS NULA NO ES CIERTA ENTONCES, EL VALOR DE PARÁMETRO $\theta$ ES DISTINTO (MAYOR O MENOR) A $\theta_0$.

Es un planteo "por distinto".

2. Si la HIPÓTESIS NULA no es ÚNICA, y si la DESIGUALDAD EQUIVALENTE es la desigualdad **menor**,

$$H_0: \theta \leq \theta_0$$

Entonces se plantea la siguiente HIPÓTESIS ALTERNATIVA MÚLTIPLE

$$H_1: \theta > \theta_0$$

A esta forma de plantear la HIPÓTESIS ALTERNATIVA MÚLTIPLE se la interpreta: SI LA HIPÓTESIS NULA NO ES CIERTA, ENTONCES, EL VALOR DEL PARÁMETRO $\theta$ ES MAYOR A $\theta_0$.

Es un planteo "por mayor".

3. Si la HIPÓTESIS NULA no es ÚNICA, y si la DESIGUALDAD EQUIVALENTE es la desigualdad **mayor**,

$$H_0 : \theta \geq \theta_0$$

Entonces se plantea la siguiente HIPÓTESIS ALTERNATIVA MÚLTIPLE

$$H_1: q < q_0$$

A esta forma de plantear la HIPÓTESIS ALTERNATIVA MÚLTIPLE se la interpreta: SI LA HIPÓTESIS NULA NO ES CIERTA, ENTONCES, EL VALOR DEL PARÁMETRO $q$ ES MENOR A $q_0$.

Es un planteo “por menor”.

**PRUEBA DE LA HIPÓTESIS NULA:** se llama **PRUEBA DE LA HIPÓTESIS NULA** a un método estadístico con el cual, a partir de los datos de una muestra aleatoria, se decide acerca de la veracidad o falsedad de la **HIPÓTESIS NULA** formulada, pudiéndose calcular la probabilidad de cometer un error en la decisión tomada.

La hipótesis que se prueba para decidir si debe ser rechazada o no, siempre es la **HIPÓTESIS NULA**, por lo tanto, en el resto de este trabajo a la **PRUEBA DE HIPÓTESIS NULA** se la llamará **PRUEBA DE HIPÓTESIS**.

**ESTADÍGRAFO DE PRUEBA:** se llama **ESTADÍGRAFO DE PRUEBA**, para pruebas paramétricas, a un estadígrafo apropiado, ep, con el que se realiza la **PRUEBA DE HIPÓTESIS**, que mida la discrepancia, **d**, entre el parámetro a probar y el estimador correspondiente y, además, tiene una distribución de probabilidad conocida.

Para poder establecer su discrepancia, en el ESTADÍGRAFO DE PRUEBA deben estar presentes tanto el parámetro a estimar como su correspondiente estimador. Genéricamente al ESTADÍGRAFO DE PRUEBA se lo simboliza

Y a su valor numérico, después de reemplazar las variables muestrales con el resultado de la muestra, se lo simboliza **ep.**

El ESTADÍGRAFO DE PRUEBA es una variable aleatoria que se genera transformando al estimador qÙ, por lo tanto su dominio **D**, es una transformación del espacio muestral **U.**

**REGIÓN CRITICA Y REGIÓN DE NO RECHAZO**

El método para realizar la PRUEBA DE HIPÓTESIS consiste en particionar al dominio, D, del **ESTADÍGRAFO DE PRUEBA** , **ep**, en dos subconjuntos o regiones mutuamente excluyentes. Según a cuál de las dos regiones pertenezca el valor numérico del **ESTADÍGRAFO DE PRUEBA**, se rechaza o no a la **HIPÓTESIS NULA**

Se llama **REGIÓN CRÍTICA (Rc)** al subconjunto del dominio D con el que se rechaza la **HIPÓTESIS NULA.**

Se llama **REGIÓN DE NO RECHAZO Ra = ( D – Rc)**, al subconjunto del dominio D con el que no se rechaza la **HIPÓTESIS NULA.**

Si hay una **DESIGUALDAD EQUIVALENTE** la **REGIÓN CRÍTICA** está formada por un subconjunto semicerrado. En este caso se dice que la prueba es **unilateral**.

Si no hay una **DESIGUALDAD EQUIVALENTE** la **REGIÓN CRÍTICA** está formada por dos subconjuntos semicerrados mutuamente excluyentes de igual tamaño. En este caso se dice que la prueba es **bilateral**.

**PUNTO CRÍTICO**: se llama **PUNTO CRÍTICO, pc,** a la frontera de la **REGIÓN CRITICA**.

Cuando la prueba es **unilateral** la **REGIÓN CRITICA** está formada por un conjunto semicerrado, por lo tanto, hay un solo **PUNTO CRÍTICO**. Si la prueba es **bilateral**, la **REGIÓN CRITICA** está formada por dos conjuntos semicerrados, luego hay dos **PUNTOS CRÍTICOS**. En cualquiera de las dos situaciones, los **PUNTOS CRÍTICOS** pertenecen a la **REGIÓN CRITICA**, por lo tanto, son puntos de rechazo de la **HIPÓTESIS NULA**.

**REGLA DE DECISIÓN**: se llama **REGLA DE DECISIÓN** a aquella regla que establece las pautas para rechazar la **HIPÓTESIS NULA** y se enuncia, "si el valor numérico del **ESTADÍGRAFO DE PRUEBA** pertenece a la **REGIÓN CRITICA**, entonces se rechaza la **HIPÓTESIS NULA**, en caso contrario, si el valor numérico del **ESTADÍGRAFO DE PRUEBA** no pertenece a la **REGIÓN CRITICA**, entonces no se rechaza la **HIPÓTESIS NULA**

Simbólicamente

Dada la hipótesis nula

**Ho : q = q o**

La regla de decisión establece que **hay que rechazarla** si, luego de obtener la muestra, hacer las mediciones correspondientes y calcular el valor numérico del **ESTADÍGRAFO DE PRUEBA ep**, éste **pertenece** a la **REGIÓN CRITICA**, y **no hay que rechazarla** si **no pertenece** a la **REGIÓN CRITICA**.

**ERROR DE TIPO 1**: se llama **ERROR DE TIPO 1** al hecho de rechazar la **HIPÓTESIS NULA** cuando la **HIPÓTESIS NULA** es cierta.

El **ERROR DE TIPO 1** se simboliza $\alpha$ **TI**.

Dada la hipótesis nula

**Ho : q = q o**

La regla de decisión establece que **hay que rechazarla** si, luego de obtener la muestra, hacer las mediciones correspondientes y calcular el valor numérico del **ESTADÍGRAFO DE PRUEBA**, **ep**, éste pertenece a la **REGIÓN CRÍTICA**.

Puede suceder que a pesar de que **ep** pertenezca a la **REGIÓN CRITICA**, la **HIPÓTESIS NULA** sea cierta. En este caso se estaría

cometiendo el **ERROR DE TIPO 1**.

$$\mathbf{ep}\ \hat{I}\ \hat{A}c\ /\ q = qo\ \mathbf{\dot{P}}\ \ (\mathbf{e_{T\,1)}}$$

**ERROR DE TIPO II**: se llama **ERROR DE TIPO II** al hecho de **no** rechazar la **HIPÓTESIS NULA** cuando la **HIPÓTESIS NULA** es falsa.

El **ERROR DE TIPO II** se simboliza **($e_{TII}$)**

Dada la hipótesis nula

**Ho : q = q o**

La regla de decisión establece que **no hay que rechazarla** si, luego de obtener la muestra, hacer las mediciones correspondientes y calcular el valor numérico del **ESTADÍGRAFO DE PRUEBA**, ep, éste no pertenece a la **REGIÓN CRÍTICA**.

Pero puede suceder que a pesar de que **ep** no pertenezca a la **REGIÓN CRITICA**, o sea, pertenezca a la **REGIÓN DE NO RECHAZO**, la **HIPÓTESIS NULA** no sea cierta. En este caso se estaría cometiendo el **ERROR DE TIPO II**.

$$\mathbf{ep}\ \hat{I}\ \hat{A}c\ /\ q\ ^{1}\ qo\ \mathbf{\dot{P}}\ \ (\mathbf{e_{T\,II)}}$$

**NIVEL DE SIGNIFICACIÓN**: se llama **NIVEL DE SIGNIFICACIÓN** a la probabilidad de cometer el **ERROR DE TIPO I**, o sea, a la probabilidad de rechazar la **HIPÓTESIS NULA**, cuando es cierta. El **NIVEL DE SIGNIFICACIÓN** se simboliza con la letra a, y mide el tamaño de la **REGIÓN CRITICA**

$$\mathbf{a = P\ (eTI)}$$

Dada la hipótesis nula

**Ho : q = q o**

Y utilizando la distribución de probabilidad del estadígrafo de prueba, es posible calcular la probabilidad de que su valor numérico pertenezca a la **REGIÓN CRITICA**, cuando la **HIPÓTESIS NULA** sea cierta

$$\mathbf{a = P\ (eTI) = P\ (}\mathbf{ep}\ \hat{I}\ \hat{A}c\ /\ q = qo)$$

**POTENCIA DE LA PRUEBA**: se llama **POTENCIA DE LA PRUEBA** a la probabilidad de no cometer el **ERROR DE TIPO II**, o sea, la probabilidad de rechazar la hipótesis nula cuando es falsa. La **POTENCIA DE LA PRUEBA** se simboliza con la letra

Dada la hipótesis nula

**Ho : q = q o**

Y utilizando la distribución de probabilidad del estadígrafo de prueba, es posible calcular la probabilidad de que el valor numérico de dicho estadígrafo pertenezca a la **REGIÓN CRITICA** cuando la **HIPÓTESIS NULA** no sea cierta, esto es, calcular la probabilidad de rechazar correctamente la hipótesis nula

Para poder calcular la potencia de la prueba, es necesario plantear una hipótesis alternativa única, o sea, establecer un solo valor alternativo para el parámetro

Si se calculan las potencias correspondientes a todos los valores posibles del parámetro y se las representan gráficamente se obtiene la gráfica de una función llamada **CURVA DE POTENCIA**, simbolizada **CP (qi)**

La **POTENCIA DE LA PRUEBA** es el complemento de la probabilidad de cometer el **ERROR DE TIPO II**

**ACCIÓN DERIVADA**: el rechazo de la **HIPÓTESIS NULA** inducirá al investigador a realizar una determinada acción con respecto al objeto de su investigación. Si por el contrario, no se rechaza la **HIPÓTESIS NULA**, entonces el investigador estará también inducido a realizar una acción, pero distinta. Cualesquiera de estas dos **ACCIONES** se **DERIVAN** del resultado de la prueba de hipótesis,

Se llama **ACCIÓN DERIVADA** a la acción que se lleva a cabo según el resultado de la decisión estadística que se tome, rechazar o no rechazar la **HIPÓTESIS NULA**.

**PASOS A SEGUIR PARA REALIZAR UNA PRUEBA DE HIPÓTESIS PARAMÉTRICA**

1. Establecer el *Parámetro* a probar.
2. Indicar los *Cursos de Acción*.
3. Verificar si hay una *Desigualdad Equivalente*.
4. Plantear la *Hipótesis Nula* y la *Hipótesis Alternativa*.
5. Indicar el *Estadígrafo de Prueba* a utilizar y su correspondiente *distribución de probabilidad*.
6. Establecer la *Región Crítica* y el o los *Puntos Críticos* teniendo en cuenta lo siguiente:

    a. Si la desigualdad no equivalente es la desigualdad menor, toda la región crítica está a la izquierda y el punto crítico es el fractil a.

    b. Si la desigualdad no equivalente es la desigualdad mayor, toda la región crítica está a la derecha y el punto crítico es el fractil (1- a).

    c. Si no hay desigualdad equivalente la región crítica se particiona en dos. Una parte a la izquierda cuyo punto crítico es el fractil a/2 y la otra parte a la derecha cuyo punto crítico es el fractil 1- a/2.

7. Plantear la regla de decisión estadística para rechazar o no la hipótesis nula.
8. Calcular el valor numérico del estadígrafo de prueba y verificar a qué región pertenece.
9. Tomar la decisión estadística.
10. Llevar a cabo la *Acción Derivada*.

**ESTADÍGRAFOS DE PRUEBA PARA LA PRUEBA DE HIPÓTESIS DE PARÁMETROS ESPECÍFICOS DE UNA POBLACIÓN**

La herramienta estadística fundamental para realizar una prueba de hipótesis es el **ESTADÍGRAFO DE PRUEBA**. Los estadígrafos que cumplen con las características requeridas para ser un estadígrafo de prueba son los estadígrafos de transformación de los estimadores.

ESTADÍGRAFO DE PRUEBA PARA LA PRUEBA DE HIPÓTESIS DE LA MEDIA POBLACIONAL DE POBLACIONES NORMALES

La distribución de probabilidad para la **media muestral** de poblaciones normales es la **distribución normal** en aquellos casos donde la varianza poblacional sea conocida. Pero, si la varianza poblacional es desconocida y se utiliza la **varianza muestral**, entonces, la distribución de probabilidad para la **media muestral** de poblaciones normales es la distribución **"t" de Student** con (n-1) grados de libertad.

Si el universo es finito, su tamaño forma parte de los estadígrafos de transformación, ya sea para la **distribución normal** como para la **"t" de Student** con (n-1) grados de libertad. Por lo tanto, cuando el parámetro a probar sea la media de la población m , para establecer el adecuado estadígrafo de prueba, hay que distinguir si el universo es infinito o finito, como así también si la varianza poblacional es conocida o desconocida.

POBLACIONES INFINITAS CON VARIANZA POBLACIONAL CONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones Normales Infinitas cuando **sí** se conoce la Varianza Poblacional es

Y su distribución de probabilidad es la **distribución normal**, ya que para esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

POBLACIONES FINITAS CON VARIANZA POBLACIONAL CONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones Normales Finitas cuando **sí** se conoce la Varianza Poblacional es

Y su distribución de probabilidad es la **distribución normal** ya que para esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

POBLACIONES INFINITAS CON VARIANZA POBLACIONAL DESCONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones Normales Infinitas cuando **no** se conoce la Varianza Poblacional es

Y su distribución de probabilidad es la **DISTRIBUCIÓN "t" de STUDENT** con **(n – 1)** grados de libertad, ya que para esas condiciones de población, éste es el estadígrafo de transformación del estimador.

POBLACIONES FINITAS CON VARIANZA POBLACIONAL DESCONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones Normales Finitas cuando **no** se conoce la Varianza Poblacional es

Y su distribución de probabilidad es la **DISTRIBUCIÓN "t" de STUDENT** con **(n – 1)** grados de libertad, ya que para esas condiciones de población, éste es el estadígrafo de transformación del estimador.

ESTADÍGRAFO DE PRUEBA PARA LA PRUEBA DE HIPÓTESIS DE LA MEDIA POBLACIONAL DE POBLACIONES NO NORMALES

POBLACIONES INFINITAS CON VARIANZA POBLACIONAL CONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones no Normales Infinitas cuando **sí** se conoce la Varianza Poblacional y el tamaño de la muestra es suficientemente grande (n>30) es

Ya que, por el Teorema Central del Límite, su distribución de probabilidad es, asintóticamente, la **Distribución Normal** y, en esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

## POBLACIONES FINITAS CON VARIANZA POBLACIONAL CONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones no Normales Finitas cuando **sí** se conoce la Varianza Poblacional es

Ya que, por el Teorema Central del Límite, su distribución de probabilidad es asintóticamente, la **Distribución Normal** y, en esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

## POBLACIONES INFINITAS CON VARIANZA POBLACIONAL CONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones no Normales Infinitas, si el tamaño de la muestra es suficientemente grande (n>30), cuando **no** se conoce la Varianza Poblacional es

Ya que, por el Teorema Central del Limite, su distribución de probabilidad es, asintóticamente, la **Distribución Normal** y, en esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

## POBLACIONES FINITAS CON VARIANZA POBLACIONAL DESCONOCIDA

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Media Poblacional de Poblaciones Normales Finitas cuando **no** se conoce la Varianza Poblacional es

Ya que, por el Teorema Central del Límite, su distribución de probabilidad es, asintóticamente, la **Distribución Normal** y, en esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

## ESTADÍGRAFO DE PRUEBA PARA LA PRUEBA DE HIPÓTESIS DE LA PROPORCIÓN DE ELEMENTOS CON UN DETERMINADO ATRIBUTO

## UNIVERSOS INFINITOS

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Proporción Poblacional de elementos que tienen un determinado atributo en Universos Infinitos, en el caso de que el tamaño de la muestra sea suficientemente grande (n>50), es

Ya que por el Teorema Central del Límite, su distribución de probabilidad es, asintóticamente, la **Distribución Normal**, y en esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

## UNIVERSOS FINITOS

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Proporción Poblacional de elementos que tienen un determinado atributo en Universos Finitos, en el caso de que el tamaño de la muestra sea suficientemente grande (n>50), es

Ya que por el Teorema Central del Límite, su distribución de probabilidad es, asintóticamente, la **Distribución Normal**, y en esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

ESTADÍGRAFO DE PRUEBA PARA LA PRUEBA DE HIPÓTESIS DE LA VARIANZA POBLACIONAL DE POBLACIONES NORMALES

POBLACIONES INFINITAS (CASO ÚNICO)

El **estadígrafo de prueba** que se debe utilizar para probar el valor numérico de la Varianza Poblacional de Poblaciones Normales Infinitas es,

Y su distribución de probabilidad es la **Distribución "Ji" Cuadrado** con (**n** –1) grados de libertad ya que para esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

**ESTADÍGRAFOS DE PRUEBA PARA LA PRUEBA DE HIPÓTESIS REFERIDA A LA COMPARACIÓN DE PARÁMETROS ESPECÍFICOS DE DOS POBLACIONES**

Se tratará el tema de comparar objetivamente dos poblaciones, utilizando el valor de los parámetros.

En algunas ocasiones se puede necesitar establecer si los promedios correspondientes a dos poblaciones pueden ser considerados iguales o distintos. En otras palabras, si la diferencia entre los promedios poblacionales es realmente distinta de cero o significativamente distinta de cero. El parámetro a probar, en este caso es el parámetro **DIFERENCIA DE MEDIAS POBLACIONALES**

Por ejemplo, es una empresa que tiene dos plantas de produccion se puede querer verificar si los costos medios por unidad producida en cada una de ellas son iguales.

No sólo se puede comparar los promedios, también se pueden comparar las varianzas, esto es, tratar de verificar si dos poblaciones son homoscedásticas, esto quiere decir si tienen o no la misma variabilidad. En otras palabras, si el cociente entre las varianzas poblacionales es igual a uno o no. El parámetro a probar es el parámetro **COCIENTE ENTRE LAS VARIANZAS**

Por ej, en el caso de los costos, se desea verificar si las dos plantas tienen igual sus varianzas. También se puede querer verificar si la proporción de elementos que tienen un determinado atributo es la misma para dos universos distintos. En otras palabras, si la diferencia entre las proporciones poblacionales es cero o no. El parámetro a probar es en este caso el parámetro **DIFERENCIA DE PROPORCIONES.**

Por ej, una empresa que fabrica jabones para lavar ropa está interesada en probar si la proporción de los consumidores que usan su marca es la misma en dos ciudades.

Los estadígrafos que cumplen con las características requeridas para ser un estimador de prueba son los estadígrafos de transformación de los estimadores.

ESTADÍGRAFO DE PRUEBA PARA COMPARAR LAS VARIANZAS POBLACIONALES DE DOS POBLACIONES NORMALES

El **estadígrafo de prueba** para establecer la homoscedasticidad de dos poblaciones normales, comparando las dos Varianzas Poblacionales de Poblaciones Normales y es

Y su distribución de probabilidad es la **DISTRIBUCIÓN F de SNEDECOR con n**1 – 1 grados de libertad en el numerador y **n**2 – 1 grados de libertad en el denominador, ya que para esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

ESTADÍGRAFO DE PRUEBA PARA COMPARAR LAS MEDIAS POBLACIONALES DE DOS POBLACIONES NORMALES

## ESTADÍGRAFO DE PRUEBA PARA COMPARAR LAS MEDIAS POBLACIONALES DE DOS POBLACIONES NORMALES SI LAS VARIANZAS POBLACIONALES SON CONOCIDAS

El **estadígrafo de prueba** para comparar dos Medias Poblacionales de Poblaciones Normales cuando las Varianzas Poblacionales son conocidas es

Y su distribución de probabilidad es la **distribución normal**, ya que para esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

## ESTADÍGRAFO DE PRUEBA PARA COMPARAR LAS MEDIAS POBLACIONALES DE DOS POBLACIONES NORMALES SI LAS VARIANZAS POBLACIONALES SON DESCONOCIDAS

### Varianzas poblacionales desconocidas pero iguales

El **estadígrafo de prueba** para comparar dos Medias Poblacionales de Poblaciones Normales cuando las varianzas poblacionales con desconocidas, pero se puede probar que son iguales es,

Y su distribución de probabilidad es la **distribución t de student** con (n1 + n2 – 2) grados de libertad, ya que para esas condiciones de la población, éste es el estadígrafo de transformación del estimador.

**Sa** se llama Varianza Amalgamada y es el promedio ponderado de las varianzas muestrales. Ponderadas por los respectivos grados de libertad

### Varianzas poblacionales desconocidas pero distintas

El **estadígrafo de prueba** para comparar dos Medias Poblacionales de Poblaciones Normales cuando las Varianzas Poblacionales son desconocidas pero no se puede probar que son iguales es

Y su distribución de probabilidad es, aproximadamente, la **distribución t de student** con **v** grados de libertad, donde

Ya que para esas condiciones de la población, éste es el estadígrafo de transformación del estimador,

ESTADÍGRAFO DE PRUEBA PARA COMPARAR LAS PROPORCIONES POBLACIONALES DE DOS POBLACIONES

El **estadígrafo de prueba** para comprar dos Proporciones Poblacionales es

Y su distribución de probabilidad es, asintóticamente, la **distribución normal**, ya que para esas condiciones, éste es el estadígrafo de transformación del estimador:

**Red de docentes de la UBA** - Clases particulares y cursos. Todas las materias
Primario - Secundario - CBC - Universitario -Idiomas - Computacion - Programación. Arte: Danzas y teatro.
info@clasesatodahora.com.ar / 4788-8542 y 4780-5977