
El presente resumen debe complementarse con la Bibliografía Obligatoria.

A. Estadística Inferencial

1. Parámetros: valores de la población → DESCONOCIDO
2. Estadísticos: Valores que arroja la muestra → CONOCIDO

A través de la **Estadística Inferencial** los parámetros pueden Estimarse (Inferirse) a través de los estadísticos.

Entonces, la estadística inferencial es una rama de la estadística que comprende los métodos y procedimientos para deducir propiedades (hacer inferencias) de una población, a partir de una pequeña parte de la misma (muestra).

B. Pruebas de Hipótesis

Las pruebas de Hipótesis permiten comparar Estadísticos de una o más muestras con respecto a los parámetros de la población.

Las pruebas de hipótesis implican las siguientes operaciones:

1. Formular una H_a y una H_0
2. Obtener la distribución de la Muestra (análisis de los datos obtenidos)
3. Establecer una norma de decisión (para rechazar o aceptar la H_0)
4. Calcular el estadístico adecuado
5. Tomar una decisión

C. Tipos de prueba de Hipótesis

A) Pruebas paramétricas → Variables intervalares y de razón (continuas)

B) Pruebas no paramétricas → Variables nominales y ordinales (categóricas)

C.1. Prueba Z (paramétrica)

- a) Comparación de un porcentaje observado sobre otro hipotético
- b) Comparación de dos porcentajes observados

C.2. Prueba de Ji Cuadrado (no paramétrica)

- a) Comparación de un tabla observado con otra teórica
- b) Comparación de dos tablas observadas
- c) Prueba de independencia de dos variables categóricas

D. Prueba Z

Para aplicar la Prueba Z los datos deben cumplir las siguientes condiciones:

- 1) El tamaño de la muestra debe ser mayor o igual a 30 unidades. De ser menor se utiliza la prueba *t de student*
- 2) La prueba Z basa en la distribución Normal Estándar

D.1. Prueba Z. Comparación de un porcentaje Hipotético con otro observado

Veamos un ejemplo:

Un supermercado local posee -según sus propietarios- una participación en el mercado cercana al 44%. A pesar del relativo éxito, sus dueños desean desprenderse del negocio ya que para seguir creciendo deben expandirse a otras zonas, esfuerzo que no están dispuestos a realizar ya que les supondría arriesgar en inversiones gran parte de patrimonio logrado.

En consecuencia han decidió vender la empresa. Una multinacional se mostró interesada en comprar. La firma de capitales franceses realizó un estudio con el objeto de verificar la participación del supermercado.

Luego de realizar un estudio sobre una muestra probabilística de 900 elementos, la multinacional observó que la participación del supermercado era, en realidad, del 39%.

El jefe del estudio, con criterio, se preguntó lo siguiente:

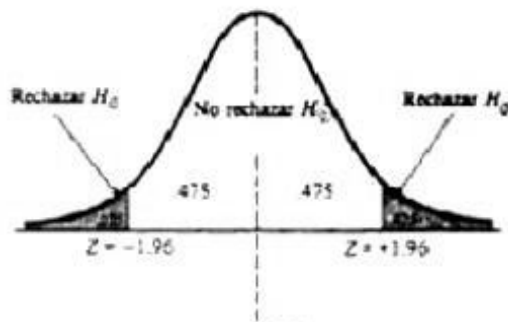
¿Cuál es la probabilidad de obtener una muestra que arroje 39% si la probabilidad real es del 44%? ¿Las diferencias observadas son productos del azar o son realmente significativas?

En consecuencia realizó la prueba correspondiente:

1ro. Elaboramos la Ho:

$H_0 = P = P$ (las proporciones son iguales (PRUEBA DE DOS COLAS))

2do. Determinamos un nivel de significancia. Es un valor de certeza que fija el investigador. (de certeza a no equivocarse) . Un valor frecuentemente utilizado es $\alpha = 0,05$



Nivel de significancia: 0,05
Nivel de confianza: 95% (0,95)

3ro. Calculamos el Error Estándar del Porcentaje Hipotético (S_p)

$$S_p = \sqrt{\frac{p \cdot q}{n}}$$

p = porcentaje hipotético del universo que tiene la característica estudiada (44%)

$q = 1 - p$ (100% - 44%)

n = tamaño de la muestra (900)

$$S_p = 1,65$$

4to. Luego se convierte la diferencia entre el porcentaje hipotético y el porcentaje observado en unidades de desviación estándar o unidades Z

$$Z = \frac{\% \text{ hipot.}(44) - \% \text{ observ}(39)}{S_p(1,65)}$$

$$Z = 3,02$$

5to. Ahora corresponde determinar si la diferencia expresada en unidades de desviación estándar ocurrió por azar o no.

Recordemos que habíamos decidido trabajar con un 95% de nivel de confianza, es decir con un 0,05 de significancia.

95% de nivel de confianza equivale a $\mp 1,96$ unidades de desviación estándar. Entonces si el valor estadístico (calculado) supera al valor crítico de Z \rightarrow RECHAZO la H₀.

Existe una probabilidad inferior a 0,05 de obtener un 39% (en realidad un valor distinto ya que se trata de una prueba de dos colas) si el valor real es del 44%.

Digámoslo de otra forma: la probabilidad de obtener un porcentaje de participación del 39% (en realidad un valor distinto ya que se trata de una prueba de dos colas) cuando la participación es del 44% es de 1 en 749 ($\alpha = 0,00125623$)

Las diferencias son estadísticamente significativas. No fueron producto del azar.

D.2. Prueba Z. Comparación de dos porcentajes observados

Medimos las intención de compra de un producto X en dos ciudades: Rafaela y Santa Fe

Intención de compra	Rafaela → n(raf)	Santa Fe → n(sfe)
Compraría	160 (40%)	250 (50%)
No compraría	240 (60%)	250 (50%)
Total	400	500

1. Principio general de la Ho es que ambas proporciones son iguales:

Ha= Las proporciones son distintas

Ho= las proporciones son iguales

2. Establecemos una regla de decisión. Optaremos por trabajar con un 95% de nivel de confianza, es decir con una probabilidad máxima de 5 en 100 ($\rho = 0,05$)

3. Calculamos el Error Estándar del porcentaje

$$Sdif = \sqrt{\frac{P(raf) \cdot Q(raf)}{n(raf)} + \frac{P(Sfe) \cdot Q(Sfe)}{n(Sfe)}}$$

$$Sdif = \sqrt{\frac{40\% \cdot 60\%}{400} + \frac{50\% \cdot 50\%}{500}}$$

$$Sdif = 3,3\%$$

4. Luego se debe convertir la diferencia en errores de desviación estándar o unidades Z

$$Z = \frac{P(raf) - P(Sfe)}{Sdif}$$

$$Z = \frac{40\% - 50\%}{3,3\%} = \pm 3,03$$

5. Se compara el valor crítico ($Z = \pm 1,96$) con el calculado ($Z = \pm 3,03$). Si el último excede al primero se rechaza la H_0 . Hay una baja probabilidad (menor a 0,05) de obtener esa diferencia a $\pm 3,03$ unidades Z .

Las muestras son estadísticamente significativas. Las diferencias no se deben al azar.

D. Chi cuadrado (No paramétrica)

Para su aplicación requiere:

- 1) Los datos estén compilados en una tabla
- 2) Las variables hayan sido medidas en una escala no métrica o bien sean categóricas
- 3) Los datos estén expresados en frecuencias absolutas
- 4) Cada celda de la tabla contenga un valor mayor o igual a 5

La prueba de Chi^2 permite:

- 1) Comparar una Tabla Hipotética con otra observada
- 2) Comparar dos tablas observadas
- 3) Probar la independencia de dos variables nominales u ordinales

D.1. Chi cuadrado. Comparación de un tabla de frecuencias observadas con otra teórica

El objetivo de la prueba es comparar dos conjuntos de datos para verificar si las diferencias son significativas o se deben al azar.

1. Formulamos la H_0

$H_a =$ Las proporciones son distintas

$H_0 =$ Las proporciones de la población son iguales

2. Construcción de una tabla de frecuencias Observadas (f_o) y (f_t)

Para calcular la frecuencia teórica de cada celda se realiza la siguiente ecuación:

$$f_t = \frac{\text{Total de cada columna} \cdot \text{Total de cada fila}}{\text{Gran total}}$$

3. Una vez construida la tabla de frecuencias teóricas, se calcula el estadístico de X^2

$$X^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

4. Se calculan los Grados de Libertad (G.L.) de la tabla.

$$G.L. = (N^{\circ} \text{ de renglones} - 1) \cdot (N^{\circ} \text{ de columnas} - 1)$$

5. Luego se buscan en la Tabla de distribución de probabilidad de Chi² para un nivel de significancia determinado y una cantidad X de grados de libertad el estadístico de Chi² (Valor crítico)

6. Se compara el valor calculado con el valor crítico de Chi². Si el primero excede al segundo se rechaza la Hipótesis nula.

Se puede concluir que existe suficiente evidencia estadística para concluir que las diferencias son significativas y no se debieron al azar en un nivel de 0,05.

D.1. Chi cuadrado: Comparación de dos tablas observadas

El objetivo de la prueba es comparar dos conjuntos de datos para verificar si las diferencias son significativas o se deben al azar.

1. Formulamos la *H₀*

H_a = Las proporciones son distintas

H₀ = Las proporciones de la población son iguales

2. Se reduce la tabla de mayor tamaño a fin de equiparar las poblaciones. Por lo tanto se construye una tabla teórica de ésta.

Para calcular la frecuencia teórica de cada celda de la tabla teórica que se construirá, se realiza la siguiente ecuación:

$$\text{Frecuencia Teórica} = \frac{n \text{ muestra más pequeña}}{n \text{ muestra más grande}} \cdot (\text{Valor de la celda})$$

3. Se calcula el estadístico de X^2

$$X^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

4. Se calculan los Grados de Libertad (G.L.) de la tabla.

$$G.L. = (N^{\circ} \text{ de renglones} - 1) \cdot (N^{\circ} \text{ de columnas} - 1)$$

5. Luego se busca en la **tabla de distribución de probabilidad de Chi^2** para un nivel de significancia determinado y una cantidad X de grados de libertad el estadístico de Chi^2 (Valor crítico)

6. Se compara el valor calculado con el valor crítico de Chi^2 . Si el primero excede al segundo se rechaza la hipótesis nula.

Bibliografía

- ❑ Guaragna, Beatriz y Fridman, Alejandro. (2003. Cap. 10 y 11) Investigación de Mercado en el Siglo XXI. Un enfoque desde el Cono Sur. Buenos Aires: de las Ciencias
- ❑ Weiers, Ronald. (1986. Cap. 12) *Investigación de mercados*. México: Prentice Hall Hispanoamericana.
- ❑ Boyd H., Westfall R., Stasch S. (1992. Cap. 14 y 15) *Investigación de Mercados - Texto y casos*. México: Noriega.