Unidad 8
Test de Hipótesis

Parte 3 : Test para una población

En esta Unidad estudiaremos:

01
Definiciones
Hipótesis estadística.
Hipótesis Nula y Alternativa.
Test de hipótesis.

02
Metodología
Paso a paso para realizar un test.

03
Test para una población
Test para la media.
Test para la proporción..

04
Test para dos poblaciones
Diferencia de medias.
Medias pareadas.
Diferencia de proporciones

05
Test e IC
Relación entre los dos métodos de estimación.

06
Errores
Que se pueden cometer al hacer un test.

03

Test para una población

Test para la media.
Test para la proporción..

Test para la media en población Normal

Caso 1: $X_1, X_2, \ldots X_n$ iid como $X \sim N(\mu, \sigma)$, σ desconocido ó

Caso 2: $X_1, X_2, \ldots X_n$ iid como $X \sim F_X$, con $EX = \mu$ y $Var(X) = \sigma^2$ (desconocido), y n grande para aplicar el T.C.L.

$$\frac{\bar{X}-\mu_0}{s/\sqrt{n}} \sim t_{(n-1)}$$

(distribución exacta o aproximada, según el caso)

Ejemplo 1

La empresa de Energía Eléctrica afirma que una aspiradora gasta 46 Kw/h. al año. Si una muestra aleatoria de 12 hogares indica que las aspiradoras gastan en promedio 42 Kw/h. al año con una desviación estándar dada por s = 11.9 Kw/h.
Pruebe con un nivel de significación de 0.05 que el gasto de energía eléctrica de las aspiradoras es distinto de 46 Kw/h. por año.

Información:
X="Gasto de energía de una aspiradora en Kw/h. al año."
$X_1, \ldots, X_{12}$ una muestra aleatoria de 12 hogares
$\bar{x} = 42$ Kw/h. al año
$s = 11.9$ Kw/h.

$\alpha = 0{,}05$
Probar que el gasto anual es distinto de 46 Kw/h.

Metodología para hacer un test

1. Definir $\mathbf{H_0}$ y $\mathbf{H_1}$.

2. Definir el **estadístico** a utilizar y **su distribución de probabilidad bajo** $\mathrm{H_0}$. La forma general de este estadístico será:

$$\text{Estadístico} = \frac{Estimador - Parámetro\ bajo\ H_0}{Desviación\ estándar\ del\ estimador}$$

3. Dado el nivel de significación α, determinar la región crítica y con los datos tomar la decisión. (O bien calcular **valor p** para tomar la decision)

Ejemplo 1

1. $H_0: \mu = 46$
   $H_1: \mu \neq 46$

2. $\dfrac{\bar{X}-\mu_0}{s/\sqrt{n}} \sim t_{(n-1)}$ Donde n = 12 y usamos distribución t aproximada porque suponemos X con distribución Normal y varianza desconocida.

3. $\alpha = 0.05$

Región crítica: $|t_{observado}| > t_{crítico}$

Donde $t_{observado} = \dfrac{42-46}{11.9/\sqrt{12}} = -1.16$ y $t_{crítico} = t_{11,\ 0.975} = 2.20$

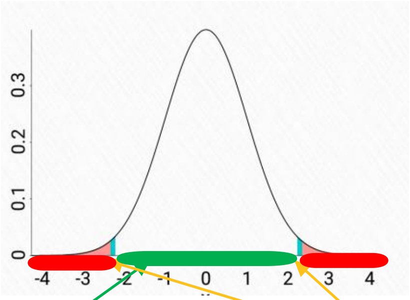
Decisión: no hay evidencia suficiente para rechazar H_0.

Concluimos que el gasto de energía eléctrica de las aspiradoras no es distinto de 46 Kw/h. por año.

Ejemplo 1

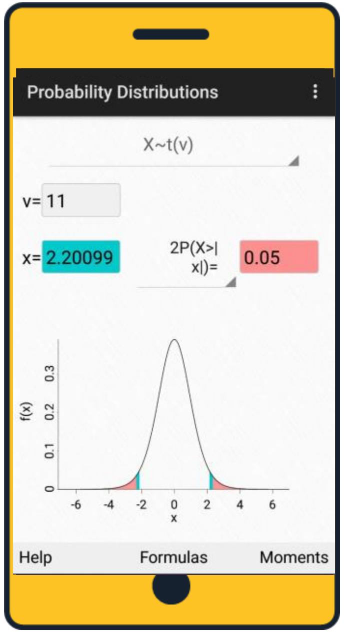
3. $\alpha = 0.05$

Región crítica: $|t_{observado}| > t_{crítico}$



$$t_{observado} = \frac{42-46}{11.9/\sqrt{12}} = -1.16$$

$$t_{crítico} = t_{11,\,0.975} = 2.20$$

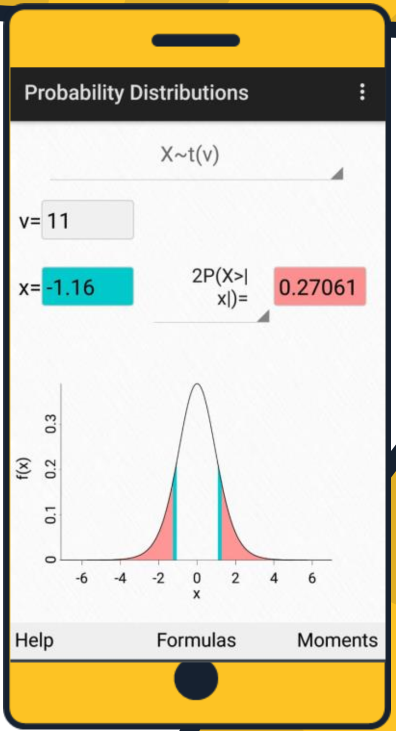


Ejemplo 1

3. Otra forma: utilizando el valor p:

Valor $p = P(|t_{11}| > |t_{observado}|) = 0.2706$

Como Valor $p > 0.05$, no hay evidencia para rechazar H_0.



Test para la proporción p en población Bernoulli

$X_1, X_2, \ldots X_n$ iid como $X \sim B(p)$, p desconocido y n grande para aplicar el T.C.L.

Como $E(X)=p$, usaremos nuevamente la media muestral como estimador.

$$\frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)/_n}} \approx N(0,1)$$

Ejemplo 2

El fabricante afirma que el 10% de los artículos son defectuosos. Los clientes afirman que $p > 0.1$. Se toma una muestra aleatoria de 100 artículos y se obtienen 12 defectuosos. ¿Qué se puede concluir?

Información:

X="Artículo es defectuoso." $X \sim B(p)$

$X_1, \ldots, X_{100}$ una muestra aleatoria de 100 artículos.

$$\bar{x} = 12/100$$

Probar que la proporción de defectuosos es distinta de 0.10.

Ejemplo 2

1. H_0: $p = 0.10$
   H_1: $p \neq 0.10$

2. $$\frac{\bar{X}-p_0}{\sqrt{p_0(1-p_0)/n}} \approx N(0,1)$$

   Donde n = 100 y usamos distribución Normal aproximada porque X tiene distribución Bernoulli y n es suficientemente grande.

3. Valor p = $P(|Z| > |Z_{observado}|)$

$$z_{observado} = \frac{0.12 - 0.10}{\sqrt{0.10x0.90/100}} = 0.67$$

Valor p = $P(|Z| > |Z_{observado}|) = 0.5028$

Decisión: no hay evidencia suficiente para rechazar H_0.
Concluimos que no hay evidencia de que la proporción de defectuosos difiere de 0.10.

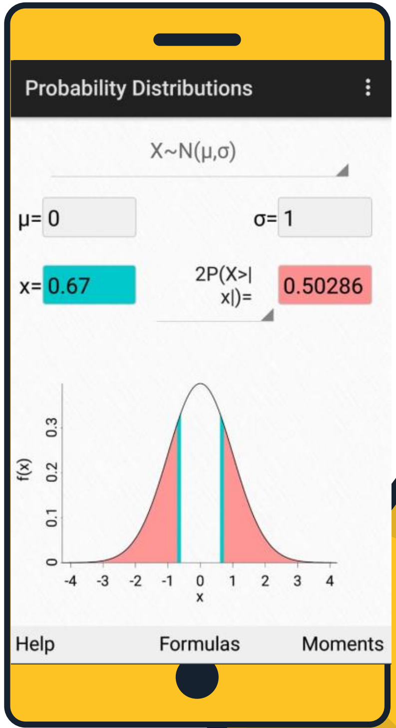
Ejemplo 2

3. Valor p = $P(|Z| > |Z_{observado}|)$

$$z_{observado} = \frac{0.12 - 0.10}{\sqrt{0.10x0.90/100}} = 0.67$$

Valor p = $P(|Z| > |Z_{observado}|)$ =0.5029

Como Valor p > 0.05 , no hay evidencia para rechazar H_0.



Observaciones

- Cuando los datos no dan suficiente evidencia para rechazar H_0, ésta no se puede rechazar. Los datos no dan suficiente evidencia cuando el valor observado del estadístico $t_{observado}$ es un valor de alta probabilidad (valor p grande) o equivalentemente cuando el valor observado del estadístico $t_{observado}$ no cae en la región de rechazo.

- Todos los procedimientos de test son similares, sólo cambian los estadísticos a utilizar.

Fin de la Parte 3