Test de Hipótesis

En ocasiones, el interés de la investigación se centra en dilucidar si un parámetro es compatible con un valor predeterminado. A partir de conocimientos previos o mediante un razonamiento lógico, se pueden elaborar hipótesis o conjeturas sobre el fenómeno o parámetro objeto de estudio (por ejemplo, establecer la hipótesis de que la media de una población toma un valor determinado).

Los test o pruebas de hipótesis permiten verificar la veracidad de alguna hipótesis establecida acerca de una población, a partir de la información disponible en la muestra, determinando si los valores difieren significativamente de los esperados por la hipótesis o si las diferencias observadas son debidas solo al azar.

Test de hipótesis

Es una **regla de decisión** para decidir si una hipótesis es verdadera o no, con su probabilidad de error.

Las hipótesis pueden clasificarse según si:

- Especifican un valor concreto o un intervalo de valores para los parámetros de una variable.
- Establecen la igualdad de las distribuciones de dos o más variables (poblaciones).
- Determinan la forma de la distribución de la variable.

Se decide entre dos hipótesis complementarias :

- la hipótesis nula (H_0)
- la hipótesis alternativa (H_1).

Al tomar una decisión, podemos cometer errores:

Error de tipo I Es el error que se comete al decidir rechazar la hipótesis nula (H_0) cuando en realidad es verdadera .
Error de tipo II Es el error que se comete al decidir no rechazar la hipótesis nula (H_0) cuando en realidad es falsa .

	Estado real	
Decisión tomada	**H_0 cierta**	**H_1 cierta**
Rechazar H_0	Error de tipo I	Decisión correcta
Dejar de rechazar H_0	Decisión correcta	Error de tipo II

α = P(Error tipo I) $\leftarrow$ nivel de significación

β = P(Error tipo II)

Potencia = $\pi(\theta) = 1 - \beta$ = P(rechazar la hipótesis nula (H_0) cuando en realidad es falsa)

Pasos a seguir en un problema de Test de Hipótesis

1. Definir las variables aleatorias.
2. Establecer las suposiciones necesarias.
3. Plantear las hipótesis nula y alternativa.
4. Definir el estadístico de prueba.
5. Fijar el nivel de significación.
6. Determinar la zona de rechazo, teniendo en cuenta el nivel de significación.
7. Con los valores de la muestra, calcular el valor del estadístico de prueba y tomar la decisión.
8. Interpretar la decisión en términos estadísticos y del problema.

Una vez definidas la variable y las suposiciones, aparece la pregunta

¿Quienes son H_0 y H_1?.

En todo test, la hipótesis nula está explicitada por una igualdad o una afirmación positiva: una cierta proporción es igual a 0.50, una cierta media es igual a 200, etc.

En los test de hipotésis el error α se fija de antemano. Como ese erro está asociado a las decisiones por H_1, la hipótesis alternativa debe ser aquella que, de afirmarse, se desea estar seguri (salvo error α) de que es correcta. Es decir, es la hipótesis a controlar.

Ejemplo 1:

Un laboratorio hace una propaganda en la que dice que el tiempo medio en que su aspirina calma el dolor de cabeza es menor que 15 minutos. El director de otro laboratorio sospecha que lo que dice la propaganda no es cierto. Quiere decidr si le hace juicio o no acusandolo de realizar propaganda desleal. ¿Qué hipótesis planteará?

X = tiempo (minutos) en que la aspirina de la propaganda calma el dolor de cabeza.

Podemos suponer $X \sim N(\mu, \sigma^2)$, donde

μ = tiempo medio en que la aspirina de la propaganda calma el dolor de cabeza

La propaganda afirma que $\mu < 15$. El director del otro laboratorio sospecha que lo que dice la propaganda no es cierto y quiere decidir si $\mu < 15$ ó $\mu > 15$

- Posibilidades a plantear

I) $H_0: \mu \geq 15$ $\quad H_1: \mu < 15$ $\longrightarrow$ No lo acusa de propaganda desleal.

II) $H_0: \mu \leq 15$ $\quad H_1: \mu > 15$ $\longrightarrow$ Lo Acusa de propaganda desleal.

- Posibles errores:
  - Decidir que $\mu < 15$ y en realidad $\mu > 15$: equivale a no acusar lo de propaganda desleal y realmente la propaganda es falsa
  - Decidir que $\mu > 15$ y en realidad $\mu < 15$: equivale a acusarlo de propaganda desleal cuando en realidad no lo es.

A uno de los errores le podemos fijar un valor de probabilidad pequeño de que ocurra (α), mientras que el otro queda indeterminado. Por esto

$\alpha = P$ (Hacer juicio acusandolo de propaganda desleal y en realidad la propaganda es cierta)

$$\alpha = P(\underbrace{\text{decidir que } \mu > 15}_{\text{Rechazar } H_0} \text{ y en realidad } \underbrace{\mu < 15}_{H_0 \text{ es verdadera}})$$

Entonces $\mu < 15$ debe estar en H_0 y $\mu > 15$ debe estar en H_1:

$$H_0)\ \mu \leq 15 \quad \text{vs} \quad H_1)\ \mu > 15$$

Regla practica para plantear hipótesis:

De las dos hipótesis posibles, lo que se quiere probar con más seguridad se debe colocar en H_1 (Siempre que se pueda !!!)

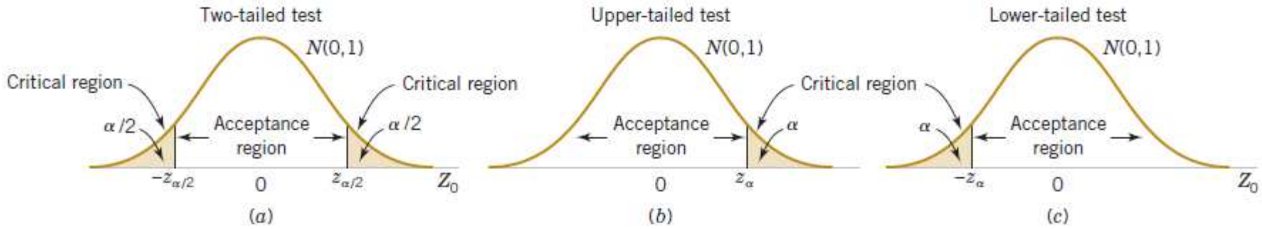
Test para una muestra

Para la media con varianza conocida

Suposición: $X \sim \mathcal{N}(\mu;\sigma)$ con σ conocida

Estadístico de prueba $\rightarrow$ $\boxed{\mathcal{Z} = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}}$

		Zona de rechazo (Z.R)
1 -$H_0 : \mu \leq \mu_0$	$H_1 : \mu > \mu_0$	$\mathcal{Z} \geq z_\alpha$
2- $H_0 : \mu \geq \mu_0$	$H_1 : \mu < \mu_0$	$\mathcal{Z} \leq -z_\alpha$
3- $H_0 : \mu = \mu_0$	$H_1 : \mu \neq \mu_0$	$\lvert\mathcal{Z}\rvert \geq z_{\alpha/2}$



Si $n=25$, $\overline{X}=17$, $\sigma=6$, $\alpha=0.05$:

- $H_0)\ \mu \leq 15$ vs $H_1)\ \mu > 15$, $X \sim N(\mu, 36)$
- $Z = \frac{\overline{X}-15}{6/\sqrt{25}} \longrightarrow Z_{Obs} = \frac{17-15}{6/\sqrt{25}} = 1{,}667$
- $\alpha = 0.05 \Rightarrow R.R = \{Z \geq Z_\alpha\} = \{Z \geq 1.645\}$

  $Z_{Obs} = 1{,}667 > 1{,}645$

Decisión: Rechazar H_0

En términos del problema: el tiempo medio en que la aspirina de la propaganda calma el dolor de cabeza es mayor que 15 minutos. Como la probabilidad de error de esta afirmación es menor que 5%, el director del otro laboratorio decide hacer juicio acusandolo de propaganda desleal.

Decisión en base a Intervalos de Confianza (solo para hipótesis bilateral)

Supongamos que con la muestra obtenida no puedo rechazar $H_0 = \mu = \mu_0$: Esto equivale a decir que

$$\left| \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\alpha/2} \right|$$

$$-z_{\alpha/2} \leq \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\alpha/2}$$

$$\overline{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \overline{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Decisión

- Para σ conocido:
  si $\mu_0 \notin (\overline{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \overline{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$, **rechazo** H_0.

- Para σ desconocido:
  si $\mu_0 \notin (\overline{X} - t_{n-1,\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \overline{X} + t_{n-1,\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$, **rechazo** H_0.