DÓCIMA DE HIPÓTESIS

INFERENCIA

- Inferencia
  - Estimación
    - Puntual
    - Por intervalos
  - Prueba de hipótesis
    - Paramétricas
    - No paramétricas

Pruebas de Hipótesis o dócimas de hipótesis

Introducción

En la vida todos tomamos decisiones, para ello se analizan alternativas y luego se decide en función de convicciones o preferencias personales. Las pruebas de hipótesis siguen casi el mismo proceso, salvo que necesitamos información estadística.

La experiencia sobre el comportamiento de algún índice, o la exigencia del cumplimiento de alguna norma nos lleva a realizar proposiciones sobre el valor de algún parámetro estadístico.

Estas proposiciones se deben contrastar con la realidad (mediante el muestreo de datos) para tomar una decisión entre rechazar o no rechazar la proposición.

Estas proposiciones se denominan **Hipótesis** y el procedimiento para decidir si se rechazan o no se rechazan se denomina **Prueba de Hipótesis.**

Pruebas de Hipótesis

Introducción

Una prueba de hipótesis es una herramienta de análisis de datos que puede formar parte de un experimento comparativo más completo.

Una **hipótesis Estadística** es un proposición (una afirmación de que algo es verdadero). Las hipótesis se realizan, sobre los parámetros de una población o sobre la distribución de probabilidad de una variable aleatoria.

1ro se identifica algo de interés y luego se plantean 2 hipótesis, que se denominan hipótesis nula H_0 y hipótesis alternativa H_1.
La H_0 por lo general es una afirmación sobre un parámetro poblacional que tiene un valor específico. Se llama nula porque es el punto inicial de una investigación.
La H_1 es el complemento de lo planteado en la H_0 .

Pruebas de Hipótesis

Introducción

Una vez que se establecen las **hipótesis nula y alternativa**, se trabaja bajo el **supuesto de que la primera es una afirmación verdadera,** hasta que hay suficientes evidencias para rechazarla.
Para entender mejor el proceso, suele establecerse un paralelismo entre el contraste de hipótesis y el juicio a una persona.
En nuestro país se parte de la premisa de que el acusado es inocente. Planteamos las hipótesis:

$$\begin{cases} H_0 : \text{El acusado es inocente} \\ H_1 : \text{El acusado no es inocente} \end{cases}$$

El acusado es inocente hasta que se presentan suficientes hechos que demuestren lo contrario. Si se consiguen evidencias que demuestren que el acusado es culpable, se “rechaza la hipótesis nula” de que el acusado es inocente por lo tanto corresponde considerar que <u>el acusado no es inocente</u> (hipótesis alternativa), y se decide condenarlo.

En el caso de que “no se rechace H_0” se considera que el acusado es inocente y se decide no condenar o absolver al acusado.

La decisión se toma en función de la hipótesis nula
(si se rechaza o no se rechaza la hipótesis nula)

Hipótesis nula e hipótesis alternativa

- Resumen de formas para hipótesis nula y alterna (μ valor de interés)

$$\begin{cases} H_0: \mu \geq \mu_0 \\ H_1: \mu < \mu_0 \end{cases} \qquad \begin{cases} H_0: \mu \leq \mu_0 \\ H_1: \mu > \mu_0 \end{cases} \qquad \begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

- La igualdad siempre aparece vinculada al la hipótesis nula.

- Una forma de facilitar la selección adecuada de las hipótesis es asignando a la hipótesis nula H_0 la desigualdad que contiene en signo igual.

Error tipo I y Error de tipo II

- Son 4 los resultados posibles que pueden obtenerse por el hecho de que la hipótesis nula sea verdadera o falsa y que la decisión sea **rechazar H_0** o **no rechazar H_0.**
- Recordemos que las hipótesis nula y alternativa son aseveraciones sobre la población que se complementan.
- No siempre es posible que las conclusiones sean verdaderas o correctas, entonces puede suceder lo que se muestra en el gráfico siguiente:

Decisión	H_0 verdadera	H_0 es falsa
No se rechaza H_0	Decisión Correcta	Error tipo II
Rechazar H_0	Error tipo I	Decisión Correcta

Error tipo I y Error de tipo II

- No se puede eliminar la posibilidad de errores en la prueba de hipótesis, pero si es posible considerar su probabilidad
- Se define como:

  - α=probabilidad de cometer un error tipo I

    (Probabilidad de rechazar H_0 siendo verdadera)
  - β=probabilidad de cometer error tipo II

    (Probabilidad de No rechazar H_0 siendo falsa).

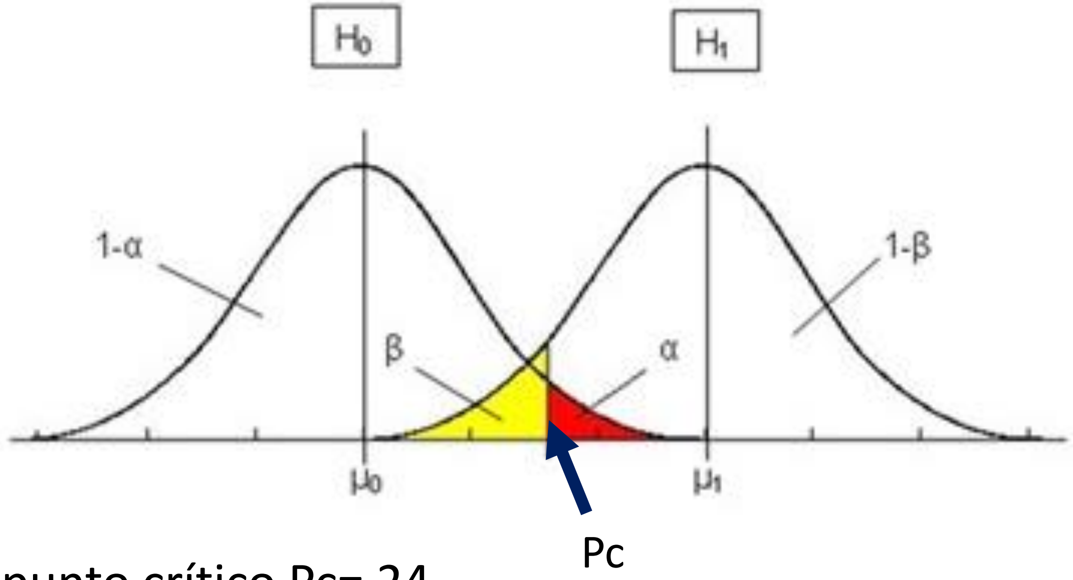
- La máxima probabilidad permisible se le llama nivel de significación para la prueba. Los valores acostumbrados son de 0,05 y 0,01.
- El valor de α **(nivel de significación)** debemos fijarlo previamente a la realización de la prueba estadística.

Error tipo I y Error de tipo II

- Para comprender como varían los errores, consideremos las siguientes hipótesis (POR ÚNICA VEZ):

$$\begin{cases} H_0: \mu = 20\ u \\ H_1: \mu = 30\ u \end{cases}$$

α=P(Error tipo I): abajo curva H_0
β=P(Error tipo II): abajo curva H1



Si consideramos que $\sigma^2 = (15)^2 = 225$, n=25 y el punto crítico Pc= 24

Calculemos α y β ¿qué relación hay entre ellos?

α=P(Error tipo I): abajo curva H_0 con μ_0 = 20

$$P(\bar{X} > 24) = P(z > \frac{24-\mu}{\frac{\sigma}{\sqrt{n}}}) = P(z > \frac{24-20}{\frac{15}{\sqrt{25}}}) =$$

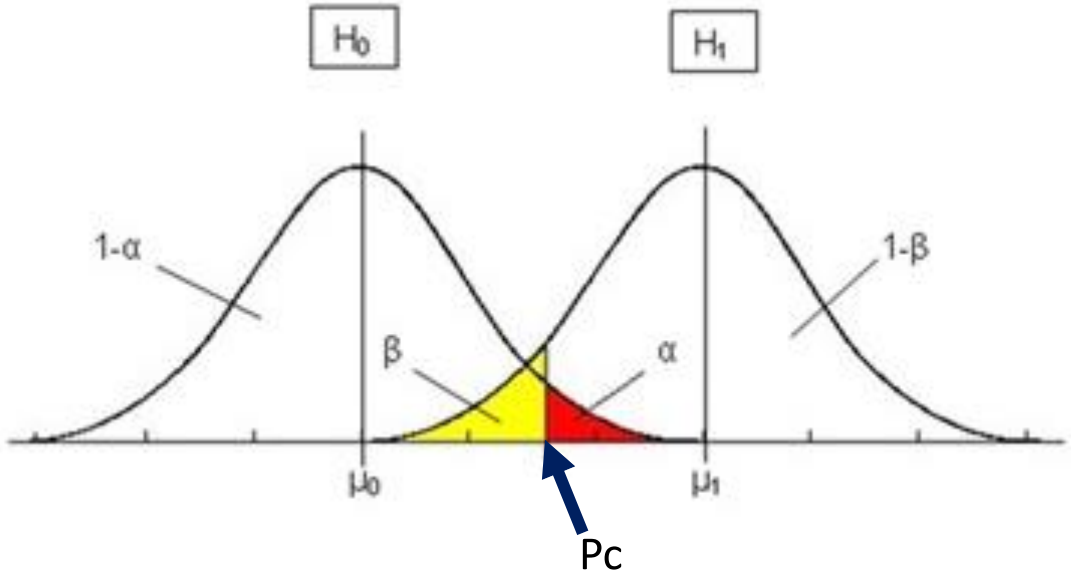
Como la zona está a la derecha del punto crítico

$$P(z > 4/3) = 1-0{,}9082 = 0{,}0918$$

β=P(Error tipo II): abajo curva H_1 con μ_1 = 30

$$P(\bar{X} < 24) = P(z < \frac{24-\mu}{\frac{\sigma}{\sqrt{n}}}) = P(z < \frac{24-30}{\frac{15}{\sqrt{25}}}) =$$

$$P(z < -2) = 0{,}0228$$



¿Qué pasa con α y β si se corre el punto crítico a la derecha?

Aumenta β y disminuye α

1.- Dadas las siguientes afirmaciones identifique el parámetro que corresponda a cada afirmación y exprese la hipótesis nula y la alternativa de forma simbólica.

frase	Completar las hipótesis nula y alternativa
1.- El gasto promedio es por lo menos $50 **El gasto promedio no es menor que $50**	$\begin{cases} H_0: \\ H_1: \end{cases}$
2.- A lo sumo el 25% de los estudiantes concurre a la universidad en auto.	
3.- La edad promedio es 13 años	
4.- La duración media de las baterías de celular es menor que 800 horas.	
5.- Para cumplir con la normativa, la varianza del nivel de impurezas en tanto por ciento en los envíos de un cierto producto químico no puede superar el valor (no más que) 4.	
6.- Cuanto mucho la inversión promedio fue de $1000 **La inversión promedio no supera los $1000**	
7.- La proporción de conductores que admite pasar en luz roja supera el 50%.	
8.- La estatura media de los jugadores de básquet es al menos de 1,75m	

2.- Dadas las frases anteriores, indique según el parámetro correspondiente, la variable pivotal o estadístico de la prueba a utilizar, dibújela y señale la región crítica use un nivel de significación del 5%.

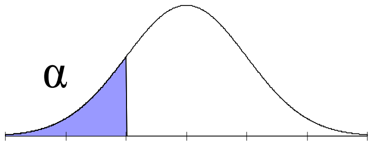
Parámetro **(μ, π, σ^2)**	**Variable pivotal** **(z, t, χ^2)**	**Gráfico de la variable y región crítica**

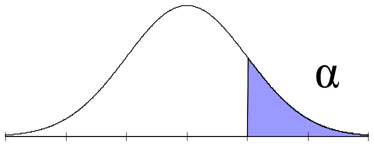
Hipótesis nula e hipótesis alternativa

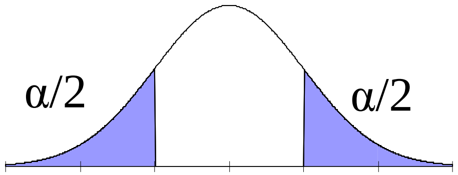
- Resumen de formas para hipótesis nula y alterna (μ valor de interés)

$$\begin{cases} H_0: \mu \geq \mu_0 \\ H_1: \mu < \mu_0 \end{cases} \qquad \begin{cases} H_0: \mu \leq \mu_0 \\ H_1: \mu > \mu_0 \end{cases} \qquad \begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

- Variable pivotal o estadístico de la prueba- (z ó t).
- Dibujar la variable y marcar la región crítica

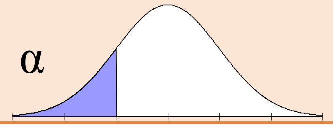
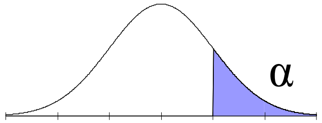
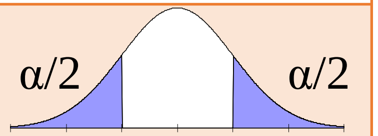
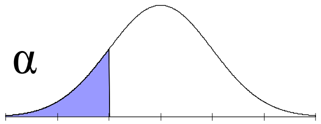
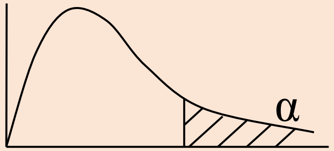
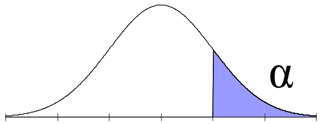
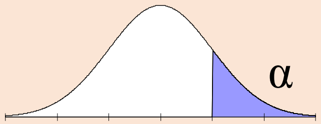
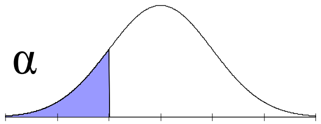






- La igualdad siempre aparece vinculada al la hipótesis nula.

- Una forma de facilitar la selección adecuada de las hipótesis es asignando a la hipótesis nula H_0 la desigualdad que contiene en signo igual.

frase	Completar las hipótesis nula y alternativa
1.- El gasto promedio es por lo menos \$50 El gasto promedio no es menor que \$50	$\boxed{Z\ (\sigma^2)\ o\ t\ (\cancel{\sigma^2})}$ $H_0: \mu \geq 50\$$ $H_1: \mu < 50\$$ 
2.- A lo sumo el 25% de los estudiantes concurre a la universidad en auto.	$H_0: \pi \leq 0{,}25$ $H_1: \pi > 0{,}25$  $\boxed{Z}$
3.- La edad promedio es 13 años	$\boxed{Z\ (\sigma^2)\ o\ t\ (\cancel{\sigma^2})}$ $H_0: \mu = 13$ años $H_1: \mu \neq 13$años 
4.- La duración media de las baterías de celular es menor que 800 horas.	$H_0: \mu \geq 800$hs $H_1: \mu < 800$hs  $\boxed{Z\ (\sigma^2)\ o\ t\ (\cancel{\sigma^2})}$
5.- Para cumplir con la normativa, la varianza del nivel de impurezas en tanto por ciento en los envíos de un cierto producto químico no puede superar el valor (no más que) 4.	$\boxed{\chi^2}$ $H_0: \sigma^2 \leq 0{,}04$ $H_1: \sigma^2 > 0{,}04$ 
6.- Cuanto mucho la inversión promedio fue de \$1000 La inversión promedio no supera los \$1000	$H_0: \mu \leq 1000\$$ $H_1: \mu > 1000\$$  $\boxed{Z\ (\sigma^2)\ o\ t\ (\cancel{\sigma^2})}$
7.- La proporción de conductores que admite pasar en luz roja supera el 50%.	$\boxed{Z}$ $H_0: \pi \leq 0{,}5$ $H_1: \pi > 0{,}5$ 
8.- La estatura media de los jugadores de básquet es al menos de 1,75m	$H_0: \mu \geq 1{,}75$ m $H_1: \mu < 1{,}75$ m  $\boxed{Z\ (\sigma^2)\ o\ t\ (\cancel{\sigma^2})}$

PASOS NECESARIOS PARA REALIZAR UNA DÓCIMA DE HIPÓTESIS

1.- Establecer la hipótesis nula y la hipótesis alterna (teniendo en cuenta el parámetro de interés).

2.- Elegir un nivel de significación α.

3.- Seleccionar el estadístico de prueba o variable pivotal.

4.- Dibujar el estadístico de la prueba o variable pivotal, marcar la región crítica y formular la regla de decisión.

5.- Reunir los datos muestrales y calcular el valor del estadístico de prueba o variable pivotal.

6.- Decisión y Conclusión en términos del problema.

Una muestra aleatoria de 100 muertes registradas en Estados Unidos el año pasado muestra una vida promedio de 71,8 años. Suponga que la desviación estándar poblacional es de 8,9 años. ¿Esto parece indicar que la vida media hoy en día es mayor a 70 años?. Utilice nivel de significación de 0,05.

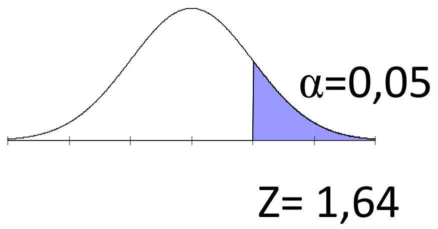
Variable: Años de vida de personas en Estados Unidos

1.- $\begin{cases} H_0: \mu \leq 70 \text{ años} \\ H_1: \mu > 70 \text{ años} \end{cases}$

2.- $\alpha = 0{,}05$

3.- $$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

4.-



<u>Regla de decisión</u>

Rechazo H_0 si Z cal $\geq 1{,}64$

No rechazo H_0 si Z cal $< 1{,}64$

5.- $$Zcal = \frac{71{,}8 - 70}{8{,}9 / \sqrt{100}} = 2{,}02$$

6.-Como Z cal > 1,64 **se rechaza H_0**

<u>Conclusión</u>: Con un nivel de significación de 0,05 tengo evidencias para rechazar la hipótesis nula y concluir que la vida media de hoy es mayor a 70 años.

Históricamente la proporción de clientes que compran con tarjeta de crédito en una determinada Farmacia es como mínimo del 25%, sin embargo la dueña de la farmacia piensa que esta cifra ha disminuido significativamente. De los últimos 1122 clientes 242 compraron con tarjeta de crédito, si α=1%. ¿Se está cumpliendo lo que piensa la dueña?.

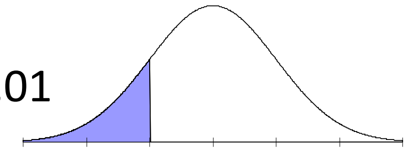
Variable: proporción de personas que compran con tarjeta de crédito en una Farmacia

1.- $\begin{cases} H_0: \pi \geq 0{,}25 \\ H_1: \pi < 0{,}25 \end{cases}$

2.- $\alpha = 0{,}01$

3.- $$z = \frac{p - \pi}{\sqrt{\frac{p(1-p)}{n}}}$$

4.- α=0,01



Z= -2,33

Regla de decisión

Rechazo H_0 si Z cal < -2,33
No rechazo H_0 si Z cal ≥ -2,33

5.- $p = \frac{242}{1122}$=0,215

$$z = \frac{0{,}215 - 0{,}25}{\sqrt{\frac{0{,}215(0{,}785)}{1122}}} = -\ 2{,}86$$

6.-Como Z cal < -2,33 **se rechaza H_0**

Conclusión: Con un nivel de significación de 0,01 o del 1%, tengo evidencias para rechazar la hipótesis nula y concluir que existen evidencias para pensar que se cumple lo que piensa la dueña (la proporción de clientes que compran con tarjeta de crédito en una determinada farmacia ha disminuido).

Una industria farmacéutica pone en marcha una nueva máquina de llenado de frascos de jarabe. La disposición técnica de control de calidad especifica que la varianza de llenado es a lo más de $0{,}02ml^2$.Se toma una muestra aleatoria de 20 frascos llenados por la nueva máquina y se encuentra que la varianza es $0{,}0225ml^2$. ¿Excede la nueva máquina la especificación técnica de control de calidad?

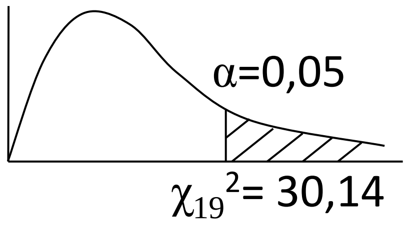
Variable: llenado de líquido de un frasco

1.- $\begin{cases} H_0: \sigma^2 \leq 0{,}02 \text{ ml}^2 \\ H_1: \sigma^2 > 0{,}02 \text{ ml}^2 \end{cases}$

2.- $\alpha = 0{,}05$

3.- $\chi^2_{(n-1)} = \dfrac{(n-1)\cdot S^2}{\sigma^2}$

4.-



Regla de decisión

Rechazo H_0 si $\chi^2 cal \geq 30{,}14$
No rechazo H_0 si $\chi^2 cal < 30{,}14$

5.- $\chi^2 cal = \dfrac{(19)\cdot 0{,}0225}{0{,}02} = 21{,}38$

6.-Como $\chi^2 cal < 30{,}14$ **no se rechaza H_0**

Conclusión: Con un nivel de significación de 0,05 tengo evidencias para no rechazar la hipótesis nula de que la nueva máquina no excede las especificaciones técnicas de control de calidad.

¡Muchas Gracias por su atención!