



Prueba de Hipótesis

Pamela E. Pairo

En la clase de hoy:

- Prueba de Hipótesis para una muestra
- Ejemplos

Recreo  

- Ejercitación

En la clase anterior...

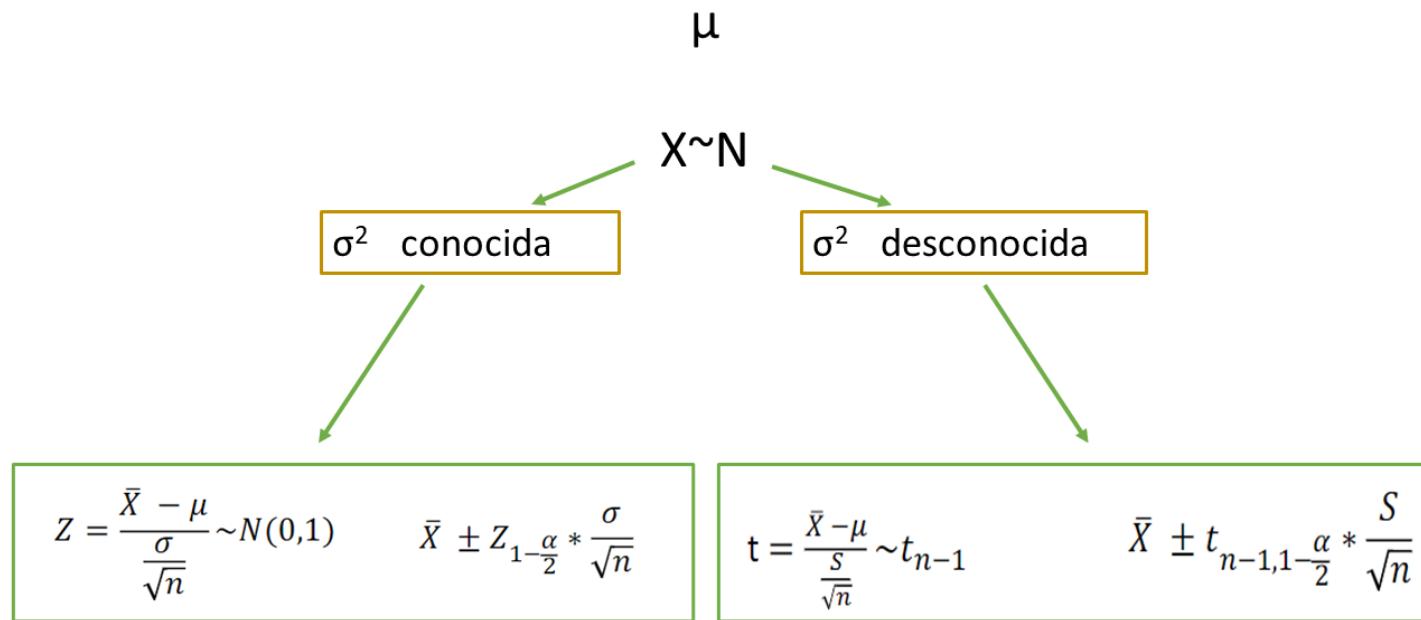
Estadística Inferencial

Permite estimar parámetros poblacionales

Probar hipótesis formuladas por sobre una población

Construir modelos estadísticos y efectuar predicciones

Intervalo de confianza para la media poblacional



¿Qué es una prueba de hipótesis?

- Es un proceso para determinar la validez de una aseveración hecha sobre la población basándose en evidencia muestral
- Es una afirmación sobre la población, a nivel de sus parámetros
- Debe plantearse antes de obtener la muestra

La prueba de hipótesis es un procedimiento de toma de decisión , relacionada principalmente con la elección entre dos conjuntos posibles de valores del parámetro

Definiciones importantes

Hipótesis de investigación **no es lo mismo que** la hipótesis estadísticas

Hipótesis de investigación: denotada por H_i expresa el objetivo de la investigación

Hipótesis estadística

- **La hipótesis nula**, denotada por H_0 , es el status quo o estado actual (lo que se cree hasta el momento) o la que asegura que no hay diferencias en la población. Es la hipótesis del no efecto
- **La hipótesis alternativa**, denotada por H_1 o H_a , es lo opuesto a la hipótesis nula, el cambio en la población que el investigador/a espera sea verdadero

Definamos las hipótesis: Ejemplo 1

Supongamos que se quiere determinar la efectividad de incorporar la vitamina E a la dieta de cerdos a fin de mejorar el aumento de peso, que actualmente es de 100g/día.

Definir la Variable de estudio

X= ?

Planteo de las hipótesis

H₀ ?

H_a ?

Ejemplo 2

La concentración de almidón en plantas de lechuga en la provincia de Buenos Aires, sigue una distribución normal con media 28 g y varianza $25 (\mu\text{g})^2$. Se quiere poner a prueba la hipótesis de que la concentración media de almidón disminuye con un tratamiento de abono.

Definir la Variable de estudio

$X = ?$

Planteo de las hipótesis

$H_0 ?$

$H_a ?$

Ejemplo 3

Se quiere conocer el el ajuste de un torno mecánico. Si esta bien ajustado , la media es de 2 cm., pero cuando la herramienta de corte adopta una posición incorrecta, dicho promedio se altera, aumentando o disminuyendo, sin llegar a modificarse el desvio estándar.

Definir la Variable de estudio

X= ?

Planteo de las hipótesis

H₀ ?

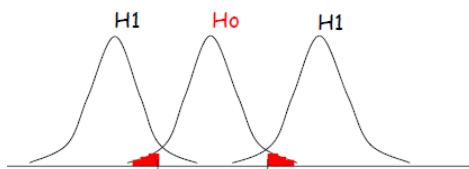
H_a ?

Planteo de hipótesis

$$H_0: \theta = \theta_o$$

$$H_1: \theta \neq \theta_o$$

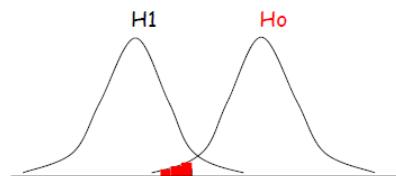
Prueba bilateral



$$H_0: \theta \geq \theta_o$$

$$H_1: \theta < \theta_o$$

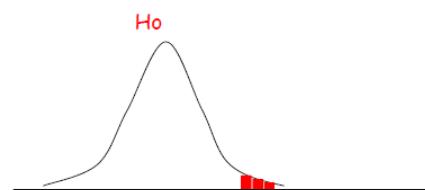
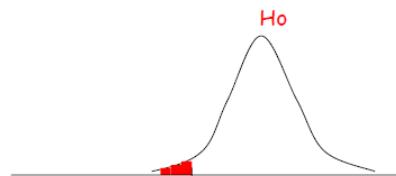
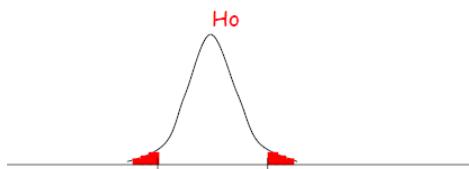
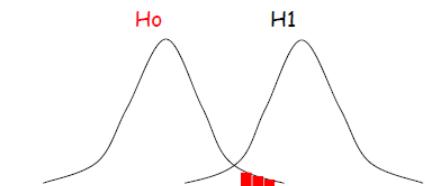
Prueba unilateral
izquierda



$$H_0: \theta \leq \theta_o$$

$$H_1: \theta > \theta_o$$

Prueba unilateral
derecha



Pasos a seguir para plantear una PH

- 1- Plantear las hipótesis estadísticas.
- 2- Seleccionar un estimador del parámetro y el estadístico de prueba (con distribución conocida).
- 3- Fijar el nivel de significación de la prueba.
- 4- Establecer la condición de rechazo (CR) y regla de decisión (RD).
 - CR: bajo qué probabilidades o valores del estadístico de prueba se debería rechazar H_0 .
 - RD: explicar en términos del problema cómo se procede.
- 5- Contrastar la muestra con la distribución teórica, calcular el p-valor y concluir.

Ejemplo

La duración media de una muestra de 100 tubos fluorescentes producidos por una compañía con un nuevo protocolo resulta ser de 1570 horas. Si la duración media de todos los tubos producidos por la compañía es 1600 horas con un desvío estándar de 120 horas, comprobar si la duración media ha cambiado con el nuevo protocolo. Asumir que duración en horas de los tubos es una variable con distribución normal.

Planteo de las hipótesis

2) Seleccionar un estimador del parámetro y el estadístico de prueba (con distribución conocida).

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

3) Fijar el nivel de significación de la prueba.

4) Establecer la región crítica y la regla de decisión.

Obteniendo los valores críticos

Como estamos trabajando en H_0 , trabajamos en la distribución Z.

```
qnorm(p=0.05/2) # Prueba Bilateral  
## [1] -1.959964
```

Si tuviésemos una Prueba unilateral:

```
qnorm(p=0.05, lower.tail = TRUE) #PU derecha  
## [1] -1.644854  
  
qnorm(p=0.05, lower.tail = FALSE) #PU Izquierda  
## [1] 1.644854
```

5) Contrastar la muestra con la distribución teórica, tomar una decisión, calcular el p-valor y concluir.

El P-valor

p-valor se define como la probabilidad de que un valor estadístico calculado sea posible dada una hipótesis nula cierta.

Si cumple con la condición de ser menor al nivel de significancia impuesto arbitrariamente, entonces la hipótesis nula será, eventualmente, rechazada.

Graficamente???

Concluyendo

Si se rechaza H_0 :

- La evidencia muestral contradice H_0
- Hay pruebas concluyentes contra H_0
- La prueba es significativa

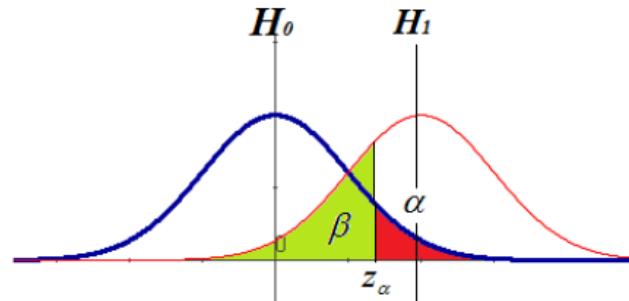
Si **no** se rechaza H_0 :

La evidencia muestral no contradice H_0

- No hay evidencias contra H_0
- La prueba no es concluyente

Riesgos al tomar una decisión

Potencia, Error de Tipo I y Tipo II



		Realidad	
		H_0 verdadera	H_0 falsa
Decisión basada en la muestra	No rechazo H_0	Decisión correcta Probabilidad $1-\alpha$	Error de tipo II Probabilidad β
	Rechazo H_0 Acepto H_1	Error de tipo I Probabilidad α (nivel de significación)	Decisión correcta Probabilidad $1-\beta$ (potencia)

Riesgos al tomar una decisión

Definiciones

$\alpha = P(\text{error tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ es verdadera})$

$\beta = P(\text{error tipo II}) = P(\text{no rechazar } H_0 / H_0 \text{ es falsa})$

$1 - \beta = \text{Potencia} = P(\text{rechazar } H_0 / H_0 \text{ es falsa})$ ó Poder o capacidad de la prueba estadística para detectar diferencias cuando estas realmente existen

Problema 1

Un cliente recibe habitualmente una partida de medidores eléctricos que, según las especificaciones del contrato, el promedio de las pérdidas debe ser menor o igual a 1 watt. Una muestra de 10 medidores, de una partida recién recibida, arroja una pérdida media de 1,06watts. Se sabe además, por experiencia anterior, que las pérdidas se distribuyen Normalmente con un desvío de 0,1watts.

- Asumiendo en un 10% la probabilidad de rechazarla indebidamente, ¿puede aceptarse la misma?

Problema 2

El espesor óptimo para obleas de silicio utilizadas en cierto tipo de circuito integrado debe ser superior a 245 nm. Un fabricante de productos electrónicos revisa todas las entregas eligiendo una muestra de 50 obleas del lote y estimando su espesor promedio. En la última entrega obtuvo una media de espesor de 246.18 nm y una desviación estándar de 3.6 nm.

“Si el espesor medio del lote fuese 248nm, que probabilidad hay de que el fabricante acepte el lote”. ¿A qué probabilidad se refiere?