Capítulo 10: Test de hipótesis

Una hipótesis es una suposición o conjetura sobre la naturaleza, cuyo valor de verdad o falsedad no se conoce. Una hipótesis estadística es una hipótesis sobre una o más poblaciones estadísticas o sobre un fenómeno aleatorio.

se refiere

Hipótesis estadística → **Población**

Análisis → **Muestra**

Ensayo de hipótesis: Procedimiento que se sigue para tratar de averiguar si una hipótesis es verdadera o falsa

Análisis para comprobar la verdad o falsedad de una hipótesis se realiza sobre una muestra

Ejemplo (ej. 1.17)

Partida de 25 piezas, k defectuosas.

Reviso 5 piezas seleccionadas al azar

X: "Cant. de piezas def. en la muestra de 5" → Hipergeométrica

$$P_k(x) = \frac{\binom{k}{x}\binom{25-k}{5-x}}{\binom{25}{5}}, \quad \max(0, k-20) \le x \le \min(5,k)$$

con $k \in \Theta = \{0, 1, \ldots, 25\}$

suponemos que hay un acuerdo con el vendedor

Acuerdo: como máximo k_0 def. en la partida...

La partida ¿cumple con lo acordado?

2 alternativas para K $\begin{cases} k \in \Theta_1 = \{0, 1, \dots, k_0\} & \text{ó} \\ k \in \Theta_2 = \{k_0+1, \dots 25\} \end{cases}$

Lo expresamos como: Decidir entre 2 hipótesis

$H_0: k \in \Theta_1$ vs $H_1: k \in \Theta_2$

y debo decidir sobre los valores obs de X

Observar que las hipótesis son sobre un parámetro desconocido que nunca voy a poder conocer (por lo tanto son hipótesis) y la decisión se toma a partir de los valores observados en mi muestra

Generalizando:

Predicción que se deriva de la teoría q. estoy analizando

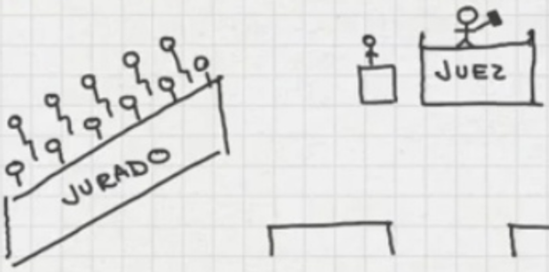
H_1: (Hipótesis del investigador) ó Hipótesis Alternativa

H_0: Hipótesis nula. Objeto del ensayo.
(Se formula con el propósito de ser rechazada)

La conclusión fuerte en un ensayo de hipótesis unicamente se tiene en el rechazo de H_0

Un ensayo puede rechazar o no la hipótesis nula. Cuando no, se dice que "no hay pruebas concluyentes en contra de la hipótesis" pero no quedaría probado que la hipótesis es verdadera. En cambio, si se rechaza H_0 quedaría probado que es falsa y se concluye que H_1 es verdadera

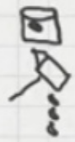
Comienzos...



¿Y si se equivoca?
Encarcela a un inocente
o libera a un culpable

H_0: Inocente vs. H_1: Culpable

¿Evidencia suficiente? Si $\longrightarrow$ cárcel



(Si no encuentra evidencia suficiente, no va a la cárcel, pero no significa que sea inocente...)

Definición Tipos de errores al tomar una decisión.

Error de tipo I: Se comete cuando se rechaza una hipótesis nula que era verdadera.
Debe tener MUY baja probabilidad

Error de tipo II: Se comete cuando no se rechaza una hipótesis nula que era falsa

[EI es más grave que EII]

Pensarlo en los ejemplos

Potencia del test: probabilidad de rechazar la hipotesis nula ⊛

Test de hipótesis es una Regla de decisión entre H_0 y H_1, y la expresamos como una función de la muestra aleatoria $\delta(\underline{X})$ que puede tomar los valores 0 o 1. Si $\delta(\underline{X}) = 1$ se rechaza H_0, y en caso contrario no se rechaza.

⊛ En función del parametro desconocido sobre el cual se plantea la hipotesis

Volviendo al lote, rechazaba el lote si encontraba 1 o más piezas def. en las 5 revisadas.

$$H_0: k \in \Theta_1 \quad vs \quad H_1: k \in \Theta_2$$

(Rechazar el lote es rechazar H_0)

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } X > 0 \longrightarrow \text{se rechaza el lote} \\ 0 & \text{si } X = 0 \longrightarrow \text{no se rechaza el lote} \end{cases}$$

, (no es único... ¿será óptimo?

Definición: Sea $\underline{X} = (X_1, \dots X_n)$ una m.a. de una población con distribución $F_\theta(x)$, $\theta \in \Theta$. Sean Θ_1 y Θ_2 tales que $\Theta_1 \cup \Theta_2 = \Theta$ y $\Theta_1 \cap \Theta_2 = \emptyset$. Un test para este problema es una regla de decisión basada en $\underline{X}$ para decidir entre 2 hipótesis:

$$H_0 : \theta \in \Theta_1 \quad \text{vs.} \quad H_1 : \theta \in \Theta_2$$

Entonces

$$P(\text{"Error I"}) = P_\theta(\text{"rechazar } H_0\text{"}), \quad \theta \in \Theta_1$$
$$P(\text{"Error II"}) = P_\theta(\text{"No rech. } H_0\text{"}), \quad \theta \in \Theta_2$$

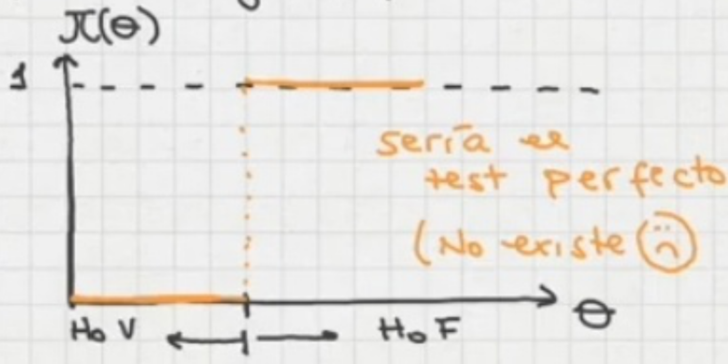
Y la potencia del test

$$\pi_\delta(\theta) = P_\theta(\text{"rech. } H_0\text{"}) = P_\theta(\delta(\underline{X}) = 1)$$
$$\overset{(*)}{=} E_\theta(\delta(\underline{X}))$$

(*) Como $\delta(\underline{X})$ es una Bernoulli, esa prob. coincide con la esperanza

Obs:

$$P(\text{"Error I"}) = \pi_\delta(\theta), \quad \theta \in \Theta_1$$
$$P(\text{"Error II"}) = 1 - \pi_\delta(\theta), \quad \theta \in \Theta_2$$



Voy a buscar un test óptimo de manera que la función de potencia se parezca lo más posible a ese mundo ideal. No se pueden disminuir ambos errores en simultaneo para un determinado tamaño de muestra (aumentar uno nos significa disminuir el otro y viceversa).

Para que ambos errores sean pequeños se debe aumentar la cantidad de observaciones

Nivel de Significación del test: es la máxima probabilidad de cometer un error de tipo I   maximo de la funcion potencia

$$\alpha = \sup_{\theta \in \Theta_1} \pi_\delta(\theta)$$

Se llama p-valor de un test al menor nivel de significación para el cual se rechaza H_0, para una observación dada.

(Es la prob. de encontrar un valor tan o más extremo que el que se encontró con la muestra observada)

Llamaremos

$$\beta(\theta) = P_\theta(\text{"Error II"}) = 1 - \pi_\delta(\theta), \theta \in \Theta_2$$

(curva característica operativa)

Entrada en calor

La producción diaria en cierta empresa es aleatoria con distribución $\mathcal{N}(10, 1.5^2)$. Se propone comprar una máquina que promete una distribución $\mathcal{N}(11, 1.5^2)$. Plantear las hipótesis correspondientes al problema y el test que permita decidir
un

La nueva maquina es igual a lo que tengo o produce con una media mejor. Solo compro la maquina si encuentro evidencia de que es mejor

La maq nueva es tan buena como lo que tengo

La maquina nueva es mejor

$$H_0: \mu = 10 \quad \text{vs} \quad H_1: \mu = 11$$

Ahora, voy a tomar una muestra, observarla, y decidir en base a ella. Supongamos $n=16$

Criterio de decisión razonable

Rechazaría H_0 si la probabilidad de observar mi muestra cuando $\mu = 11$ es mucho mayor que cuando $\mu = 10$, es decir:

verosimilitud cuando $\mu=11$ evaluada en la m.a

$$T = \frac{f_{\mu=11}(\underline{X})}{f_{\mu=10}(\underline{X})} > k_\alpha$$

k_α: cte que depende de α

max prob que soporto de cometer un error del tipo I

Error del tipo I: Comprar la maquina cuando es tan buena como lo que tengo

Voy a rechazar H_0 solo si encuentro evidencia suficiente de que μ no es 10 sino que es 11

Ⓗ solo tiene dos valores $\{10, 11\}$

Obs: las hipótesis se plantean sobre un parámetro desconocido μ, del cual nunca voy a saber su verdadero valor

Fijado α, proponemos

$$\delta(\underline{x}) = \begin{cases} 1 & \text{si} \quad T > k_\alpha \\ 0 & \text{si} \quad T \leq k_\alpha \end{cases}$$

De manera que cumpla que

$$\alpha = P_{\mu=10}\left(\delta(\underline{x}) = 1\right)$$

(ver que cumple con la definición)

Ahora necesito que me digan quién es $f_\theta(x)$ y ya puedo resolver el ejercicio.

Test del cociente de verosimilitu

Test para hipótesis simple vs Hip. simple

↑ (H): tiene un solo elemento

$$H_0: \theta = \theta_1 \quad \text{Vs.} \quad H_1: \theta = \theta_2$$

Regla de decisión (Test)

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } \dfrac{f_{\theta_2}(\underline{X})}{f_{\theta_1}(\underline{X})} > k_\alpha \\ 0 & \text{si no} \end{cases}$$

Dado un α fijo, debemos hallar k_α que cumpla

$$\alpha = P_{\theta_1}(\delta(\underline{X}) = 1)$$

Supongamos que $X_1, \dots X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ nuestra m.a., y que por alguna razón conocemos a σ^2.

$$H_0: \mu = \mu_1 \quad \text{Vs} \quad H_1: \mu = \mu_2$$

Desarrollemos el cociente de verosimilitud (sí, es una cuenta larga)

$$f_\mu(\underline{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2} = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2}$$

$$\rightarrow \frac{f_{\mu_2}(\underline{x})}{f_{\mu_1}(\underline{x})} = \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_2)^2}}{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_1)^2}} > k_\alpha$$

$$-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_2)^2-(x_i-\mu_1)^2 > \ln k_\alpha$$

$$\sum_{i=1}^{n}\cancel{x_i^2}-2\mu_2 x_i+\mu_2^2-\cancel{x_i^2}+2x_i\mu_1-\mu_1^2 < -2\sigma^2\ln k_\alpha$$

$$2(\mu_1-\mu_2)\sum_{i=1}^{n}x_i+n(\mu_2^2-\mu_1^2) < -2\sigma^2\ln k_\alpha$$

si $\mu_1 > \mu_2$

$$\sum_{i=1}^{n}x_i < k'_\alpha$$

↳ depende de muchas constantes y de α

Resultando

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } \sum_{i=1}^{n} X_i < k'_\alpha \\ 0 & \text{si no} \end{cases}$$

Para hallar el valor k'_α usamos la 2ª condición:

$$\alpha = P_{\mu_1}(\delta(\underline{X})=1) = P_{\mu_1}\left(\sum_{i=1}^{n}X_i < k'_\alpha\right)$$

$X_i \sim N(\mu_1,\sigma^2)$

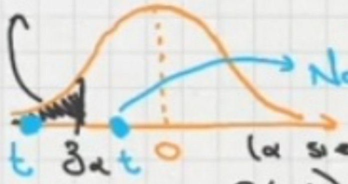
$$= P_{\mu_1}\left(\underbrace{\frac{\sum_{i=1}^{n}X_i - n\cdot\mu_1}{\sqrt{n\sigma^2}}}_{\text{Esta será la regla}} < z_\alpha\right)$$

$\sum_{i=1}^{n}X_i \sim N(n\mu_1, n\sigma^2)$ (cap. 8)

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } \dfrac{\sum_{i=1}^{n} X_i - n\mu_1}{\sqrt{n\sigma^2}} < z_\alpha \\ 0 & \text{si no} \end{cases}$$

↳ Estadístico del test

Zona de rechazo



es un test con nivel de significación α

(α siempre es pequeño)

Obs: $\delta(\underline{X})$ depende de α, n, la m.a. y el valor del parámetro Bajo Ho

La regla será la misma si

$$H_1: \mu < \mu_1$$

Muestra → $\underline{x}$. $\quad t = \dfrac{\sum_{i=1}^{n} x_i - n\mu_1}{\sqrt{n\sigma^2}}$

nota que $\sum X_i$ es el estadístico suficiente para μ

Estudiemos la función potencia

$$\pi_\delta(\mu) = P_\mu(\delta(\underline{X}) = 1)$$

$$= P_\mu\left(\frac{\sum_{i=1}^{n} X_i - n\mu_1}{\sqrt{n\sigma^2}} < z_\alpha\right)$$

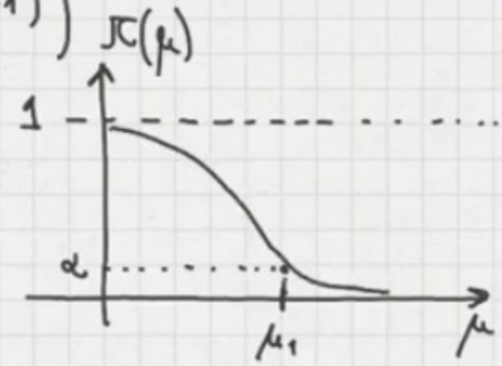
$$= P_\mu\left(\sum_{i=1}^{n} X_i < z_\alpha \cdot \sqrt{n\sigma^2} + n\mu_1\right)$$

estandarizo

$$= \Phi\left(z_\alpha - \frac{\sqrt{n}}{\sigma}(\mu - \mu_1)\right)$$

Para el caso de $\mu_1 > \mu_2$

- Si μ aumenta, $\pi(\mu)$ dism.
- $\pi_\delta(\mu)$ es decreciente.



Estudiemos la función potencia

$$\pi_\delta(\mu) = P_\mu(\delta(\underline{x})=1)$$

$$= P_\mu\left(\frac{\sum_{i=1}^n X_i - n\mu_1}{\sqrt{n\sigma^2}} < z_\alpha\right)$$

$$= P_\mu\left(\sum_{i=1}^n X_i < z_\alpha \cdot \sqrt{n\sigma^2} + n\mu_1\right)$$

estandarizo

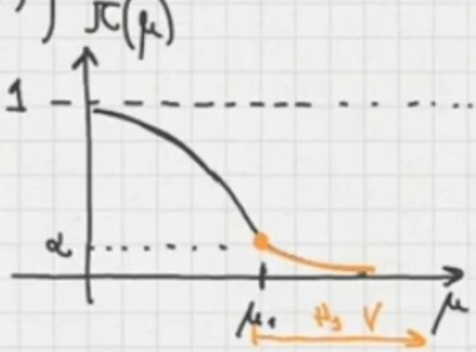
$$= \Phi\left(z_\alpha - \frac{\sqrt{n}}{\sigma}(\mu-\mu_1)\right)$$

Para el caso de $\mu_1 > \mu_2$

- Si μ aumenta, $\pi(\mu)$ dism.

Si $H_0: \mu \geq \mu_1$ vs $H_1: \mu < \mu_1$

Vale el mismo test



Ejemplo: Se mide el grado de impurezas de un producto químico. El método de medición está afectado por un error que se supone $\mathcal{N}(0, 0.01)$. Además, los errores correspondientes a distintas mediciones son independientes entre sí. Se sabe que el producto es aceptable si el grado de impurezas es menor a 0.7. Se hacen 64 mediciones y se quiere encontrar un test de modo que la prob. de aceptar el producto cuando este no satisfaga las condiciones sea menor que 0.05

¿Qué sabemos? θ: grado de impurezas (no lo conozco)

X_i = "i-ésima medición del grado de impureza"

$$X = \theta + N, \quad N \sim \mathcal{N}(0, 0.01)$$

$$\longrightarrow X \sim \mathcal{N}(\theta, 0.01) \qquad n = 64$$

Tengo que elegir mis hipótesis tal que

$$\underbrace{P(\text{"Aceptar el prod. cuando no satisface"})}_{\text{Error I}} < \underbrace{0.05}_{\alpha}$$

$$H_0: \theta \geq 0.7 \quad \text{vs.} \quad H_1: \theta < 0.7$$

Si Rechazo H_0, acepto el producto.

Es equivalente

$$H_0: \theta = 0.7 \quad \text{vs} \quad H_1: \theta < 0.7$$

Vimos recien que el test de CV para este caso llegaba a

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } \sum_{i=1}^{64} X_i < k'_\alpha \\ 0 & \text{si no} \end{cases}$$

Ahora,

$$\alpha = 0.05 = P_{\theta=0.7}(\delta(\underline{X}) = 1)$$

$$= P_{\theta=0.7}\left(\sum_{i=1}^{64} X_i < k'_\alpha\right)$$

$$= P_{\theta=0.7}\left(\frac{\sum_{i=1}^{64} X_i - 64 \cdot 0.7}{\sqrt{64 \cdot 0.01}} < z_{0.05}\right)$$

Es equivalente

$$H_0: \theta = 0.7 \quad \text{vs} \quad H_1: \theta < 0.7$$

Vimos recien que el test de CV para este caso llegaba a

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } \sum_{i=1}^{64} X_i < k'_\alpha \\ 0 & \text{si no} \end{cases}$$

Entendamos: Quiero decidir que θ es chiquito, y lo voy a decidir cuando la suma de mis obs. sean muy chiquitas ¿Cuan chiquitas?

Ahora,

$$\alpha = 0.05 = P_{\theta=0.7}\left(\delta(\underline{X}) = 1\right)$$

$$= P_{\theta=0.7}\left(\sum_{i=1}^{64} X_i < k'_\alpha\right)$$

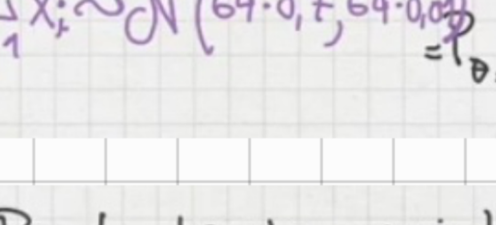
¿Por que? Porque si $\theta = 0{,}7$ entonces $\sum_{i=1}^{64} X_i \sim \mathcal{N}(64 \cdot 0{,}7,\ 64 \cdot 0{,}01)$

$$= P_{\theta=0.7}\left(\frac{\sum_{i=1} X_i - 64 \cdot 0{,}7}{\sqrt{64 \cdot 0{,}01}} < z_{0.05}\right)$$

Por lo tanto, mi test para este problema será

$$\underbrace{\delta(\underline{X})}_{\text{Una V.A.}} = \begin{cases} 1 & \text{si } \dfrac{\sum_{i=1}^{64} X_i - 64 \cdot 0{,}7}{\sqrt{64 \cdot 0{,}01}} < -1{,}6449 \\ 0 & \text{si no} \end{cases}$$

T. Estadistico del test. SOLO depende de la m. a.

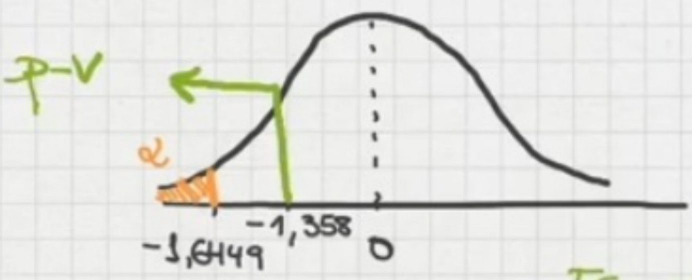


Muestra: $\sum_{i=1}^{64} x_i = 47{,}71 \rightarrow t = \dfrac{47{,}71 - 64 \cdot 0{,}7}{\sqrt{64 \cdot 0{,}01}} = -1{,}35$

No Rech Ho → No tengo evidencia Suficiente para aceptar el producto.

Ahora a entender el p-valor!!!

¡Que pena! yo quería aceptarlo...
¿Que info me da mi muestra?



En este caso
$P-V = \Phi(-1,358) = 0{,}08$

Es el menor nivel de significación para el cual se rechaza Ho para una obs dada

Es la prob. de encontrar otra muestra tanto o mas extrema, si Ho fuera V.
(es bastante poco probable), si $PV < 0.05$ Rech Ho siempre