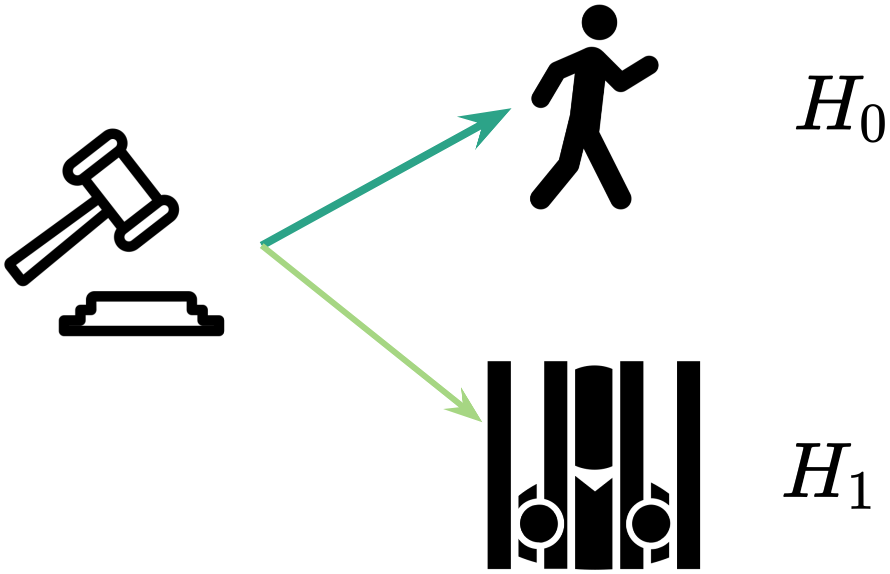
Probabilidad y estadística

Clase 7

Test de hipótesis

Motivación

Un test de hipótesis es una manera formal de elegir entre dos opciones, o hipótesis.



Inocente hasta que se demuestre lo contrario

Formalicemos esta idea

es una partición Θ

Sea una m.a. $\underline{X} = X_1, \ldots, X_n$ de una población con distribución perteneciente a una familia $F_\theta(x)$ con $\theta \in \Theta$. Sean Θ_0 y Θ_1 tales que $\Theta_0 \cap \Theta_1 = \emptyset$ y $\Theta_0 \cup \Theta_1 = \Theta$. Un test para este problema será una regla basada en $\underline{X}$ para decidir entre las dos hipótesis

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta_1$$

Inocente Culpable

Definición: Se llama test a una función $\delta(\underline{X})$ que puede tomar valores 0 o 1.

Formalicemos esta idea

En un test de hipótesis se plantean dos hipótesis:

H_1: es la hipótesis del investigador (lo que se desea probar). Se la conoce como hipótesis alternativa

$\boxed{H_0}$: se formula con el objetivo de ser rechazada. Se la conoce como hipótesis nula.

y adoptamos la H_1.

Diremos que rechazamos la hipótesis nula cuando $\delta(\underline{X}) = 1$, en caso contrario decimos que no hay evidencia suficiente para rechazar H_0.

Las hipótesis se formulan sobre parámetros de la **población**, y no de la muestra

Tipos de error

Error de tipo I: Es el error que se comete al rechazar una hipótesis nula que era verdadera. Buscamos que esto tenga muy baja probabilidad.

Encarcelar a una persona inocente

$$\mathbb{P}(EI) = \mathbb{P}_{\theta\in\Theta_0}(\delta(\underline{X}) \overset{H_1}{=} 1)$$

bajo H_0

Error de tipo II: Es el error que se comete al no rechazar una hipótesis nula que era falsa.

Dejar libre a una persona culpable

$$\mathbb{P}(EII) = \mathbb{P}_{\theta\in\Theta_1}(\overbrace{\delta(\underline{X}) = 0}^{H_0})$$

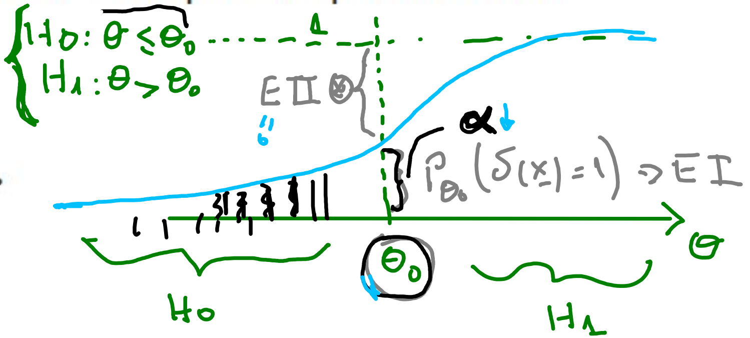
bajo H_1

	$\delta(\underline{x})=0$	$\delta(\underline{x})=1$
	Retain Null	Reject Null
sana H_0 true	$\surd$	type I error
enferma H_1 true	type II error	$\surd$

Potencia del test

$$1 - \pi_\delta(\theta_0) = 1 - \mathbb{P}_{\theta_0}(\delta(\underline{X}) = 1) = \mathbb{P}_{\theta_0}(\delta(\underline{X}) = 0) = EII$$

Def: Se llama potencia del test a la probabilidad de rechazar una hipótesis nula en función del parámetro desconocido sobre el que se plantea la hipótesis.



$$\boxed{\pi_\delta(\theta)} = \mathbb{P}_\theta(\delta(\underline{X}) = 1).$$

Luego podemos reescribir

$$\mathbb{P}(EI) = \pi_\delta(\theta),\ \theta \in \theta_0 \quad \text{y} \quad \mathbb{P}(EII) = 1 - \pi_\delta(\theta),\ \theta \in \Theta_1$$

Ejercicio 1

$$\begin{cases} H_0: \mu \leq 100 \\ H_1: \mu > 100 \end{cases}$$

Supongamos que en las especificaciones de procedimientos de una planta de energía nuclear se establece que la resistencia media de soldadura debe superar 100lb/plg. Supongamos que somos el director del equipo de inspección del ente regulador estatal que debe determinar si la planta cumple con las especificaciones. Se planea seleccionar una muestra al azar de soldaduras y realizar pruebas en cada una de ellas. $\delta(\underline{x})$

1. ¿Cuáles son las hipótesis a testear? $\boxed{\text{estafa // riesgo}}$ gasto // No riesgo
2. Explicar que significa en este contexto el error de tipo I y el de tipo II, y discutir cuáles son las consecuencias de cometer cada tipo de error

Nivel de significación

Def: Se llama nivel de significación del test a la máxima probabilidad de cometer un error de tipo I:

$$\alpha = \sup_{\theta \in \theta_0} \pi_\delta(\theta)$$

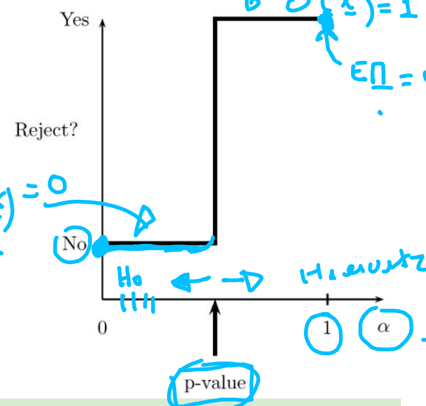
H_0

Nivel de significación y p-valor

EI: rechazar H_0 cuando H_0 es V.

Def: Se llama p-valor de un test al menor nivel de significación para el cual se rechaza H_0 $\boxed{\text{para una observación dada}}$

Otra forma: probabilidad de ver una muestra tanto o más extrema que la que se ha observado



$\delta(\underline{x})=1$ $E_{II}=0$

$\delta(\underline{x})=0$

cambiamos α

H_0 H_1 muestra

p-valor	evidencia
<.01	evidencia muy fuerte contra H0
.01-.05	evidencia fuerte contra H0
.05-.1	evidencia débil contra H0
>.1	poca o nula evidencia contra H0

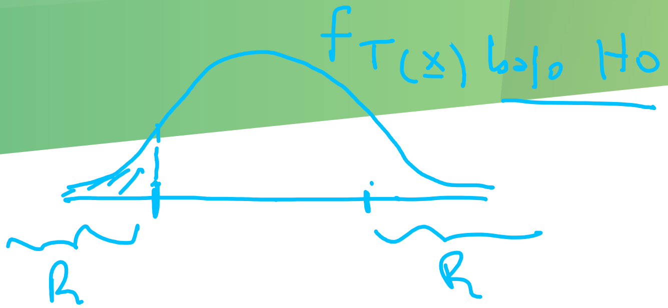
inocente

un p-valor grande no implica evidencia fuerte a favor de H0, y puede deberse a:

1. H0 es verdadera
2. H0 es falsa pero el test tiene poca potencia. (no podemos diferenciar para esa muestra.)

En este video hay una explicación simpática e intuitiva de lo que representa el p-valor!

Sobre la construcción de reglas de decisión



En la práctica, las reglas de decisión se construyen basándose en una estadística de la m.a. $\underline{X}$, es decir que son de la forma $\delta(\underline{X}) = \mathbb{1}\{T(\underline{X}) \in \mathcal{R}\}$. A $\mathcal{R}$ la conoce como región crítica o de rechazo .

(α)

depende de α

$EI = P(\text{Rechazar Ho cuando Ho era cierta}) \leq \alpha$

Podemos usar el método de pivote que vimos para intervalos de confianza o el EMV

Tipos de hipótesis

En general, tenemos 3 tipos de hipótesis que deseamos testear:

$$H_0 : \theta \geq \theta_0 \; vs. \; H_1 : \theta < \theta_0$$
$$H_0 : \theta \leq \theta_0 \; vs. \; H_1 : \theta > \theta_0$$

Hipótesis unilaterales

$$H_0 : \theta = \theta_0 \; vs. \; H_1 : \theta \neq \theta_0$$

Hipótesis bilaterales

$\frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0,1)$

Si el pivote es decreciente en θ, un test para

pivote

$$H_0 : \theta \leq \theta_0 \; vs. \; H_1 : \theta > \theta_0 \text{ será } \delta(\underline{X}) = \boldsymbol{I}\{U_{\theta_0} > k_\alpha\}$$

nivel de significación

$$H_0 : \theta \geq \theta_0 \; vs. \; H_1 : \theta < \theta_0 \text{ será } \delta(\underline{X}) = \boldsymbol{I}\{U_{\theta_0} < k_\alpha\}$$

$$H_0 : \theta = \theta_0 \; vs. \; H_1 : \theta \neq \theta_0 \text{ será } \boxed{\delta(\underline{X})} = \underbrace{\boldsymbol{I}\{U_{\theta_0} < k_{\alpha/2}\} + \boldsymbol{I}\{U_{\theta_0} > k_{1-\alpha/2}\}}$$

Ejercicio 2

Esto se lo conoce como t-test

La longitud de ciertas barras de acero sigue una distribución normal de varianza desconocida. El proveedor de las barras asegura que la media de las mismas es de 4 m.
Si en una muestra de 15 barras el promedio de las longitudes fue de 3.88 m y el desvío estándar muestral de 0.2 m, ¿qué puede decir respecto de la afirmación del proveedor?

X

μ

$$\begin{cases} H_0: \mu = 4 = \mu_0 \\ H_1: \mu \neq 4 \end{cases}$$

$$U = \frac{\bar{x} - \mu_0}{s}\sqrt{n} \sim t_{n-1}$$

Ho

Data

n = 15

$\bar{x} = 3{,}88$

$S = 0{,}2$

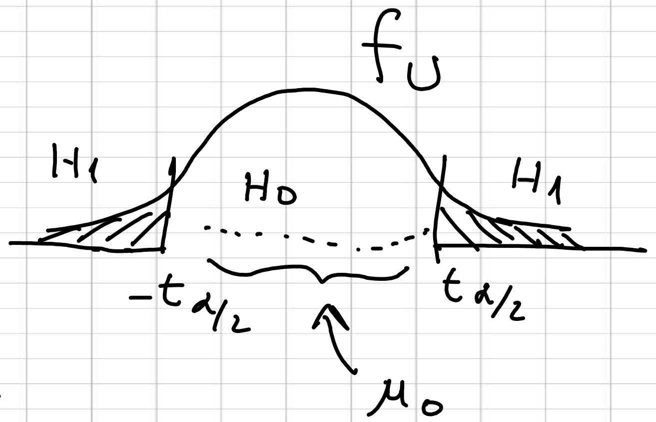
$$U_{0,5} = \frac{3{,}88 - 4}{0{,}2}\sqrt{15} = -2{,}32$$

Regla de decisión con nivel de signif. $\underline{\alpha = 0{,}05}$

$$P_{\mu_0}\left(\underline{\underline{\delta}}(\underline{x}) = 1\right) = \alpha$$

$$P_{\mu_0}\left(U < -t_{\alpha/2}\right) + P_{\mu_0}\left(U > t_{\alpha/2}\right) = \alpha$$

Prob de EI



IC

$$1 - P\left(-t_{\frac{\alpha}{2}} < U < t_{\alpha/2}\right) = \alpha$$

$$\boxed{P\left(-t_{\frac{\alpha}{2}} < U < t_{\alpha/2}\right) = 1 - \alpha} \leftarrow \text{intervalo de confianza}$$

$P(a < \mu < b)$ $\quad \frac{\bar{x}-\mu}{S}\sqrt{n}$

$P(\mu \in IC)$

$$IC_{\mu,\,95\%} = \left(\underbrace{\overset{3{,}88}{\bar{X}} - \overset{-2{,}145}{t_{\frac{\alpha}{2}}}}_{3{,}77} \cdot \frac{\overset{0{,}2}{S}}{\underset{15}{\sqrt{n}}} \; ; \; \underbrace{\bar{X} + \overset{2{,}145}{t_{\frac{\alpha}{2}}} \frac{S}{\sqrt{n}}}_{3{,}99}\right) = (3.77\,,\,3.99)$$

$$P\left(\mu \in I(\underline{x})\right) = 1 - \alpha$$

$$\delta(\underline{x}) = \mathbb{1}\left\{\underset{4}{\mu_0} \notin IC\right\} = 1$$

(el IC No contiene a μ_0)

$$\delta(\underline{x}) = \mathbb{1}\left\{ \overbrace{|U|}^{T(\underline{x})} > \underbrace{t_{\frac{\alpha}{2}}}_{= F_U^{-1}\left(1-\frac{\alpha}{2}\right)} \right\}$$

$$\delta(\underline{x}) = \mathbb{1}\left\{ |-2,32| > 2,145 \right\} = 1 \Rightarrow \text{Rechazo } H_0.$$

$$p\text{-valor} = P\left(|U| \geq U_{obs}\right)$$

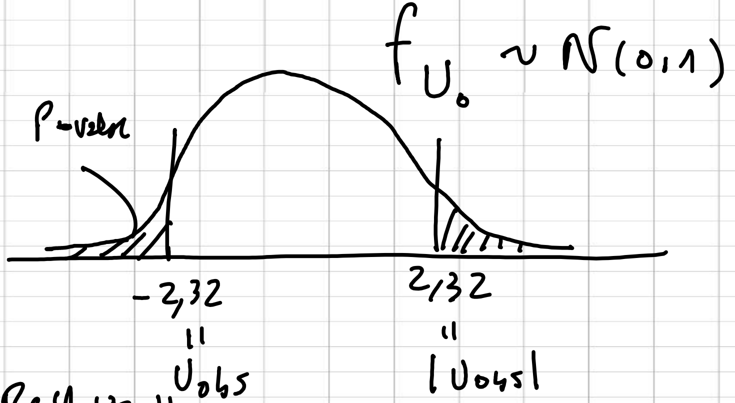
$$= P\left(|U| \geq 2,32\right)$$

simetría

$$= 2\,P\left(U < -2,32\right)$$

$$= 2\,F_U(-2,32) = 2 \times 0,01 = 0,02 < \alpha = 0,05 \Rightarrow \text{Rechazar } H_0$$

Rechazo H_0.



Test con nivel de significación asintótico

Def: Sea $\underline{X}_n = (X_1, \ldots, X_n)$ una m.a. de una población con distribución $F_\theta(x),\ \theta \in \Theta$. Se desea testear

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta_1$$

Se dirá que una sucesión de test tiene nivel de significación asintótico α si

$$\lim_{n\to\infty} \sup_{\theta\in\Theta_0} \pi_\delta(\theta) = \alpha$$

Nuevamente, podemos basar el test en la distribución asintótica de EMV, o bien usando TCL.

→ obtener la distri. asintótica del Pivote $\frac{\bar{X}-\mu}{\sigma}\sqrt{n} \sim N(0,1)$

Ejercicio 3

Región de rechazo.

p: prob. de estar de acuerdo con la construcción

m.a. $X_1, \dots, X_{100}$ con $X_1 \sim Ber(p)$

Se realiza un estudio para estimar la proporción de residentes en una ciudad que están a favor de la construcción de una planta de energía nuclear. En una muestra de $n=$ 100 personas se observó que la proporción de individuos que se encuentran a favor fue de 0.62 $= \hat{p} = \frac{1}{n} \sum_{i}^{100} x_i$

1. ¿Puede decirse con un nivel de significación de 0.01 que la mayor parte de la población está a favor de la construcción de la planta nuclear? Sí.
2. Hallar el p-valor. (estadístico)

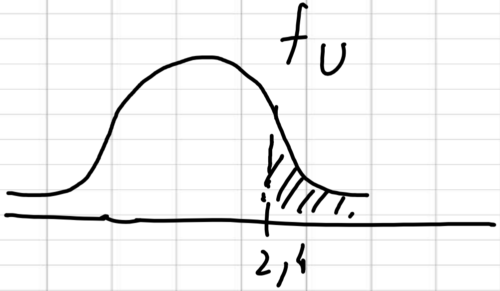
$$\begin{cases} H_0: p \leq 0{,}5 = p_0 \\ H_1: p > 0{,}5 \end{cases}$$

$$U = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)}} \sqrt{n} \xrightarrow[n \to \infty]{TCL} N(0,1)$$

$$\left[S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 \right]$$

$$\delta(\underline{x}) = \mathbb{1}\{U > z_\alpha\}, \quad z_\alpha = F_U^{-1}(1 - 0{,}01) = 2{,}33$$

$$\delta(\underline{x}) = \mathbb{1}\{2{,}4 > 2{,}33\} = 1 \Rightarrow \text{rechazo } H_0.$$

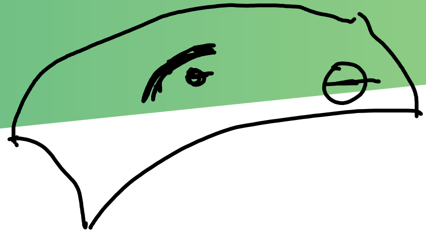


$$p_{valor} = P(U \geq 2{,}4 = U_{obs})$$

$$= 1 - P(U < 2{,}4) = 1 - F_U(2{,}4) = \underline{0{,}0082} < 0{,}01 = \alpha$$

$\Rightarrow$ Rechazo H_0.

EMV (distribución asintótica)



def: La matriz de información de Fisher $I(\theta)$ está dada por la varianza del score $s(X,\theta) = \frac{\partial}{\partial\theta}\log f(X;\theta)$ y resulta equivalente a $I(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\right]$.

EMV : $\hat{\mathbb{E}}\, s(x,\theta) = 0$

teo: el EMV es asintóticamente normal, y

$\lim_{n\to\infty} \frac{\hat{\theta}-\theta}{\text{se}} \overset{d}{=} N(0,1)$, con $\boxed{\text{se} \approx 1/\sqrt{nI(\theta)}}$. Incluso

$\lim_{n\to\infty} \frac{\hat{\theta}-\theta}{\hat{\text{se}}} \overset{d}{=} N(0,1)$ cuando consideramos $\hat{\text{se}} = 1/\sqrt{nI(\hat{\theta})}$.

poblacional

Es decir, $\hat{\theta}_n \approx N(\theta, \text{se}^2)$. $= N\left(\theta, \frac{1}{n} I(\theta)^{-1}\right)$

Wald test

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta \neq \theta_0$$

Asumiendo que $\hat{\theta}$ es asintóticamente normal: $\frac{\hat{\theta}-\theta}{\hat{\text{se}}} \overset{d}{=} N(0,1)$ (EMV)

El Wald test con nivel de significación asintótica α tiene como función de decisión

$$\delta(\underline{X}) = \mathbf{1}\left\{ \overbrace{\left|\frac{\hat{\theta}-\theta}{\hat{\mathbf{se}}}\right| > \mathbf{z}_{\alpha/2}}^{\text{Región de rechazo.}} \right\}$$

donde

$$z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$$

* El likelihood ratio test permite generalizar el test basado en el EMV a parámetros vectoriales.

Ejercicio 4

Analizar el Ejemplo 3 con el test de Wald.

$$\to I(p) = -E_p\left(\frac{\partial^2}{\partial p^2}\log f(x|p)\right)$$

$$= -E_p\left[\frac{\partial^2}{\partial p^2}\log\left(p^x(1-p)^{1-x}\right)\right]$$

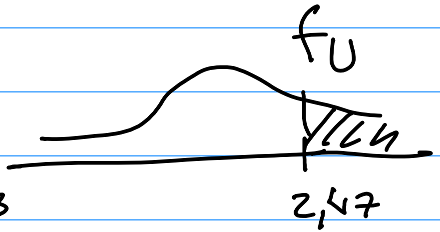
$$= -E_p\left[-\frac{x}{p^2}-\frac{1-x}{(1-p)^2}\right] = \frac{Ex}{p^2}+\frac{E(1-x)}{(1-p)^2} = \frac{p}{p^2}+\frac{1-p}{(1-p)^2} = \frac{1}{p}+\frac{1}{1-p}$$

$$= \frac{1-p+p}{p(1-p)} = \frac{1}{p(1-p)} \to \operatorname{var} X$$

$$\underline{\text{Wald}}:\ \delta_1(\underline{x}) = \mathbb{1}\left\{\frac{\hat{p}-p_0}{1/\sqrt{n I(\hat{p})}} > z_\alpha\right\} = \mathbb{1}\left\{\overbrace{\frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}}^{2,4} > \overbrace{z_\alpha}^{\sim 2,33}\right\}$$

p-valor de δ_1 es 0,0083

$$\delta_2(\underline{X}) = \mathbb{1}\left\{ \overbrace{\frac{\hat{p} - p_0}{\sqrt{\hat{p}(1-\hat{p})}}\sqrt{n}}^{2,47} > \overbrace{z_\alpha}^{2,33} \right\} = 1 \Rightarrow \text{Rechazo } H_0 \checkmark$$



p-valor de δ_2 es $F_U(2,47) = 0,0068$

es menor $\Rightarrow$ + evidencia de rechazar H_0.

Otra Por TCL :

$$\frac{\hat{p} - p}{S} \sqrt{n}$$