ESTADÍSTICA INFERENCIAL

Población $\xrightarrow{\text{M.A.S}}$ Muestra (n)

Población: μ, σ, σ^2, $p = \pi$

Muestra (n): $\bar{x}$, $S = \hat{\sigma}$, $S^2 = \hat{\sigma}^2$, $\hat{p}$

?

→ **PARÁMETROS POBLACIONALES**: valor constante que caracteriza la población.

→ **ESTIMADORES**: funciones matemáticas que dependen de la muestra

ECONOMETRÍA

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i$$

M.A.S

Población: β_0, β_1, β_2, $\vdots$, β_k, σ^2

Muestra (n): $\hat{\beta}_0$, $\hat{\beta}_1$, $\vdots$, $\hat{\beta}_k$, $\hat{\sigma}^2$

conclusiones

→ **ESTIMADORES**

→ **PARÁMETROS** desconocidos

INFERENCE

INFERENCIA
- Estimación puntual $\hat{\beta}_2 = 0'2$
- Intervalos de Confianza IC
  $\beta_2 \in [0'1, 0'3]$ $1-\alpha = 0'95$
- Contrastes de hipótesis

- **ESTIMACIÓN PUNTUAL**

$$\hat{\beta} = (X'X)^{-1}X'Y = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

- **INTERVALOS DE CONFIANZA** $\beta_j \in [LI, LS]$ $1-\alpha \equiv$ CONFIANZA

$\alpha \equiv$ SIGNIFICACIÓN o ERROR

$$\beta_j \in \left[\hat{\beta}_j \mp \overbrace{t_{n-k-1} \cdot ee(\hat{\beta}_j)}^{e_{max}} \right] \quad 1-\alpha$$

$s.e(\hat{\beta}_j)$

$S\hat{\beta}_j$

$p(LI < \beta_j < LS) \simeq 1-\alpha$

$1-\alpha \equiv$ CONFINZA: la prob aproximada de que el verdadero valor de β_j esté dentro del intervalo.

$\alpha \equiv$ Significación o Error tipo I

α es la prob. aproximada de que el verdadero valor del parámetro NO esté dentro del intervalo.

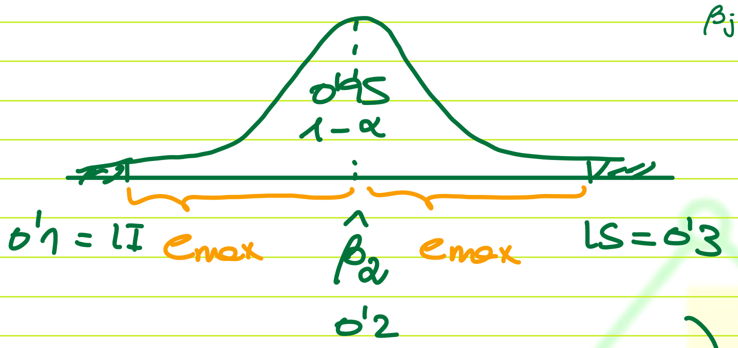
Ejemplo: $\hat{\beta}_2 = 0'2$ $\qquad \beta_2 \in [0'1, 0'3]$ $\quad 1-\alpha = 0'95$

$p(0'1 < \beta_2 < 0'3) = 0'95$

$$\beta_j \in \Big[\hat{\beta}_j \mp \overbrace{t_{n-k-1} \cdot ee(\hat{\beta}_j)}^{e_{max}}\Big] \quad 1-\alpha$$

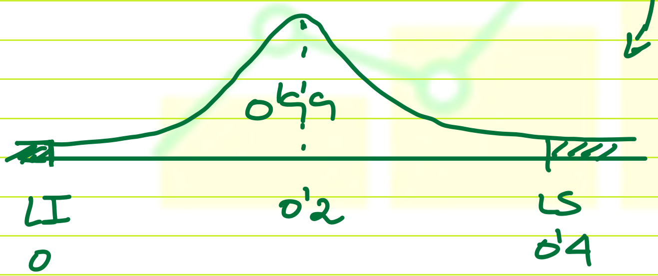
s.e $(\hat{\beta}_j)$

$S_{\hat{\beta}_j}$



$$\hat{\beta}_2 = \frac{0'3 + 0'1}{2} = 0'2$$

$(1-\alpha)\uparrow \longrightarrow AI \uparrow$



- **CONTRASTES DE HIPÓTESIS**

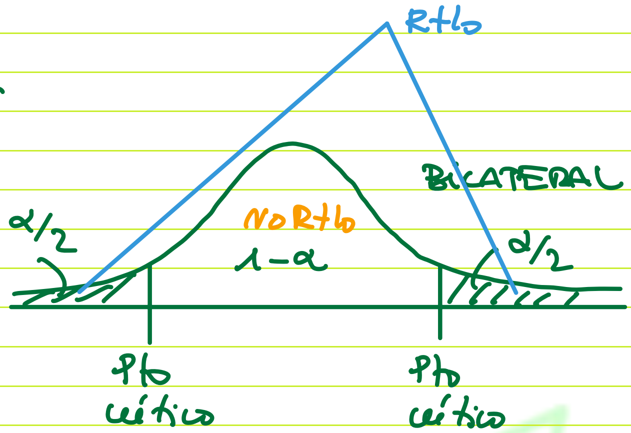
H_0 : hipótesis nula

H_1 : hipótesis alternativa (mutuamente excluyente de H_0)

TIPOS DE CONTRASTES

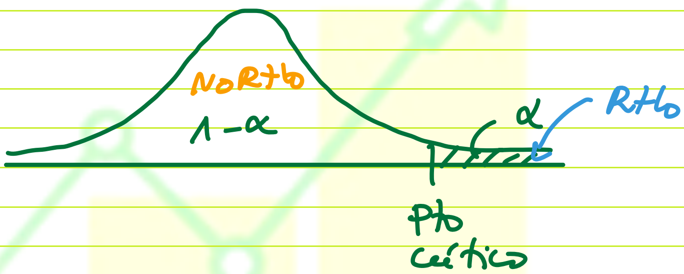
1) $H_0 : \beta_2 = 0$

$H_1 : \beta_2 \neq 0$



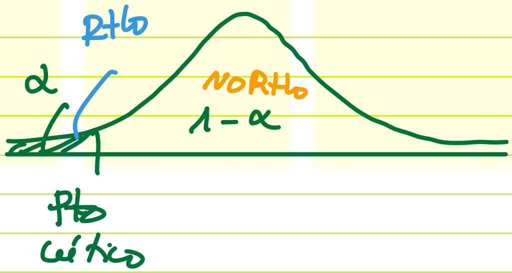
2) $H_0 : \beta_2 = 0'2$

$H_1 : \beta_2 > 0'2$



3) $H_0 : \beta_2 = 0'2$

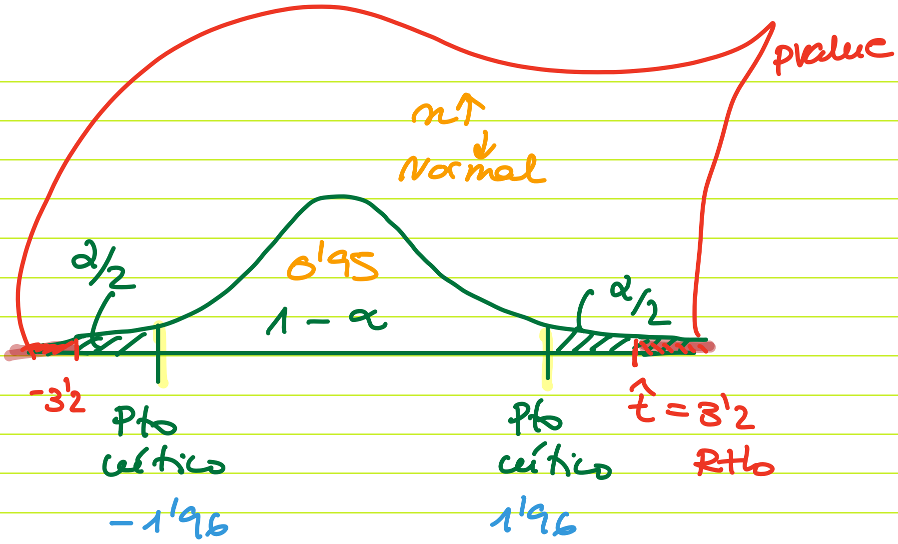
$H_1 : \beta_2 < 0'2$



Ejemplo

$H_0: \beta_2 = 2$

$H_1: \beta_2 \neq 2$



$\hat{t} \simeq \hat{z} = 3'2 > 1'96 \longrightarrow$ RHo

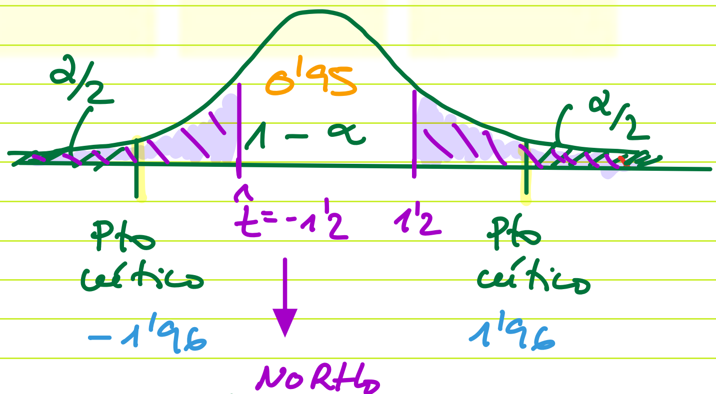
pvalue = área acumulada por el estadístico →

→ compara con α.

pvalue $> \alpha \longrightarrow$ No Rechazamos Ho
pvalue $< \alpha \longrightarrow$ Rechazamos Ho

$H_0: \beta_2 = 2$

$H_1: \beta_2 \neq 2$



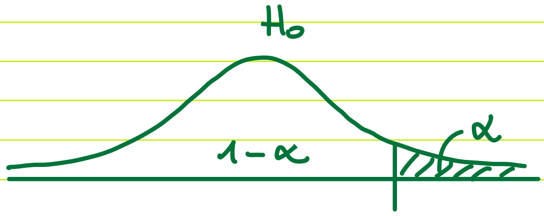
$\hat{t} \simeq \hat{z} = -1'2 \qquad -1'96 < \hat{t} < 1'96 \longrightarrow$ No RHo

pvalue $> \alpha = 0'05 \longrightarrow$ No RHo

ERRORES EN UN CONTRASTE DE HIPÓTESIS

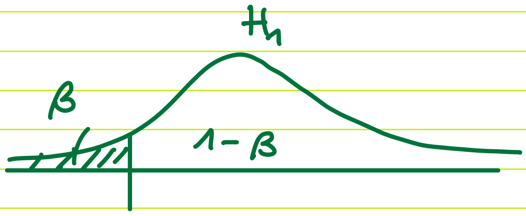
$1-\alpha \equiv$ CONFIANZA

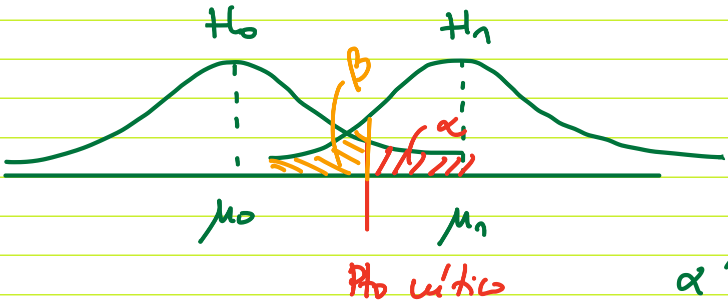
$\alpha \equiv$ ERROR TIPO I



$1-\beta \equiv$ POTENCIA DEL CONTRASTE

$\beta \equiv$ ERROR TIPO II





$\cancel{\alpha+\beta=1}$

$\cancel{\alpha-\beta=1}$

$\alpha\uparrow \iff \beta\downarrow$

$n\uparrow \Rightarrow (\alpha \text{ y } \beta)\downarrow$

	H_0 cierta	H_0 falsa
No Rechazo	$1-\alpha$ CONFIANZA	β ERROR TIPO II
Rechazo H_0	α ERROR TIPO I	$1-\beta$ POTENCIA DEL CONTRASTE

$p(\text{No R}H_0 / H_0 \text{ cierta}) = 1-\alpha$

$p(\text{R}H_0 / H_0 \text{ cierta}) = \alpha$

$p(\text{No R}H_0 / H_0 \text{ falsa}) = \beta$

$p(\text{R}H_0 / H_0 \text{ falsa}) = 1-\beta$

* Relación entre IC y CH

Ejemplo:

$H_0: \beta_2 = 0$

$H_1: \beta_2 \neq 0$

$1-\alpha = 0'95$

Información: $\beta_2 \in [0'2, 0'6]$ $1-\alpha = 0'95$

Como el cero (lo que contrastamos en H_0) no está dentro entonces rechazamos H_0.

$H_0: \beta_2 = 0'4$

$H_1: \beta_2 \neq 0'4$

Información: $\beta_2 \in [0'2, 0'6]$ $1-\alpha = 0'95$

- Como 0'4 está dentro del intervalo entonces no rechazamos H_0 con una confianza del 95%
- Al 99% de confianza no rechazamos H_0

* Al 90% de confianza no se puede concluir nada.

MODELO DE REGRESIÓN LINEAL SIMPLE

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

- y_i – variable explicada (endógena)
- $x_i \rightarrow$ variable explicativa (exógena)
- $\left.\begin{matrix}\beta_0 \\ \beta_1\end{matrix}\right\}$ Parámetros del modelo
- $\varepsilon_i \rightarrow$ término de perturbación $E(\varepsilon_i) = 0$
- Modelo de Regresión Lineal Simple poblacional

  Función de Regresión poblacional estocástica

$$y_i = \underbrace{\beta_0 + \beta_1 x_i}_{\substack{\text{parte determinista} \\ \text{(constante)}}} + \underbrace{\varepsilon_i}_{\text{parte aleatoria}}$$

- Función de Regresión poblacional FRP

$$E(Y_i / x_i) = \beta_0 + \beta_1 x_i$$

- Función de Regresión muestral FRM

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

- $\hat{\varepsilon}_i = y_i - \hat{y}_i \longrightarrow y_i = \hat{y}_i + \hat{\varepsilon}_i$

- Función de Regresión muestral estocástica

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\varepsilon}_i$$

- Modelo de Regresión Lineal Múltiple Poblacional o Función de Regresión Poblacional Estocástica:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i$$

- Función de Regresión Poblacional FRP :

$$E\left(Y_i | X_i\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}$$

- Función de Regresión Muestral FRM :

$$\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_{1i} + \widehat{\beta}_2 X_{2i} + \ldots + \widehat{\beta}_k X_{ki}$$

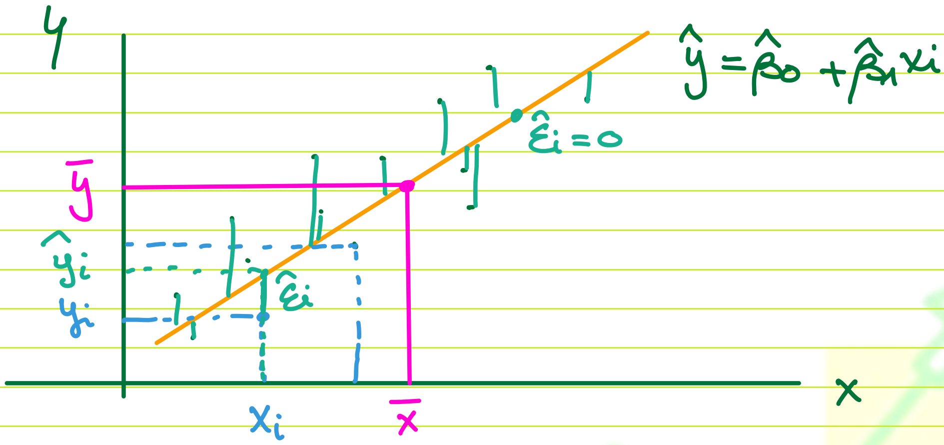
- Función de Regresión Muestral Estocástica :

$$Y_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_{1i} + \widehat{\beta}_2 X_{2i} + \ldots + \widehat{\beta}_k X_{ki} + \widehat{\varepsilon}_i$$

MRLS

Pendiente: $\hat{\beta}_1 = \frac{S_{xy}}{S_x^2} = \frac{Cov(x,y)}{Var(x)}$

Ordena: $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$



* COEFICIENTE DE BONDAD DE AJUSTE R^2

* COEFICIENTE DE CORRELACIÓN LINEAL $r_{xy} = \rho_{xy}$

- $r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$ $[-1, 1]$ ⟶ adimensional

S_{xy} tiene unidades $\left\{ \begin{array}{ll} S_{xy} = 0 & \text{No existe R.L} \\ S_{xy} > 0 & \text{Existe R.L positiva} \\ S_{xy} < 0 & \text{Exist R.L negativa} \end{array} \right.$

r_{xy} → mide la intensidad de Relación lineal

$S_{xy} = 0 \longleftrightarrow r_{xy} = 0$ $\hat{\beta}_1 = 0$

$S_{xy} > 0 \longleftrightarrow r_{xy} > 0$

$S_{xy} < 0 \longleftrightarrow r_{xy} < 0$

COEFICIENTE DE BONDAD DE AJUSTE

$R^2 = (r_{xy})^2 \quad [0, 1] \quad$ % Explicación

* Covarianza

$$S_{xy} = Cov(x,y) = \frac{\sum (x_i - \bar{x})(y_j - \bar{y})}{n-1}$$

$S_{xy} = 0 \longrightarrow$ No exist relación lineal

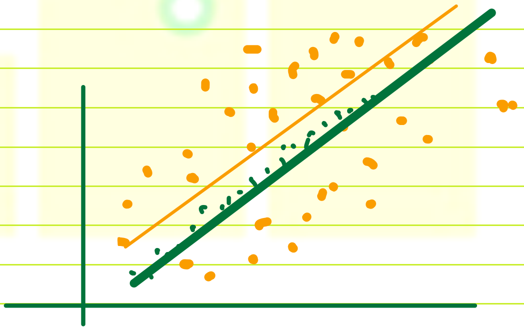
$S_{xy} = 0 \; \cancel{\longrightarrow} \; X$ e Y son independientes

X e Y son independiente $\Longrightarrow S_{xy} = 0$

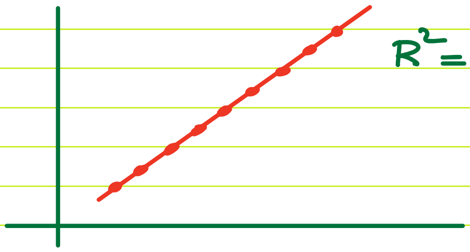
$r_{xy} \simeq 0'5$

$r_{xy} \longrightarrow 1$

$S_{xy} > 0 \longrightarrow r_{xy} > 0 \longrightarrow \hat{\beta}_1 > 0$



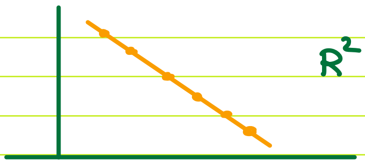
$r_{xy} = 1$



$R^2 = 1 \quad 100\%$

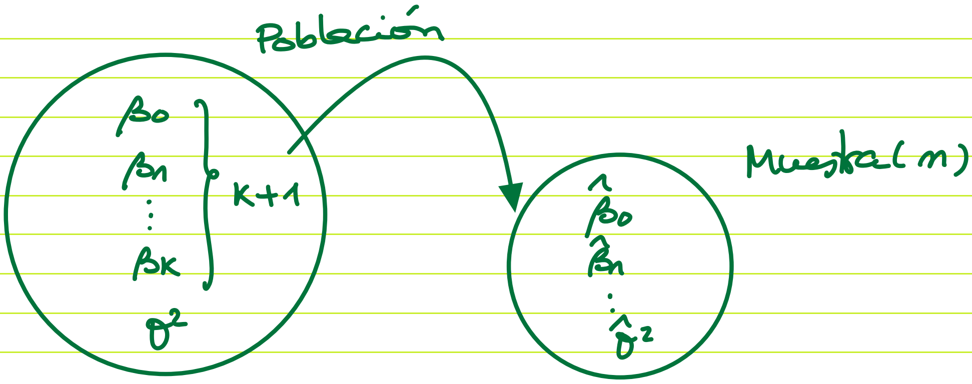
$S_{xy} < 0 \longrightarrow r_{xy} < 0 \longrightarrow \hat{\beta}_1 < 0$

$r_{xy} = -1$



$R^2 = 1 \quad 100\%$

PROPIEDADES DE LOS ESTIMADORES



MRLS :

Modelo de Regresión Poblacional :

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Función de Regresión Muestral (FRM)

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

$E(a) = a$
$Var(a) = 0$

<u>Propiedades de los estimadores</u>

* ERROR DE ESTIMACIÓN : $e(\hat{\beta}) = \hat{\beta} - \beta$

* <u>SESGO</u> : $E(e(\hat{\beta})) = E(\hat{\beta} - \beta) = E(\hat{\beta}) - E(\beta) =$

$$= E(\hat{\beta}) - \beta$$

$$\boxed{\text{Sesgo}(\hat{\beta}) = E(\hat{\beta}) - \beta}$$

- Sesgo$(\hat{\beta}) = 0 \longrightarrow E(\hat{\beta}) = \beta \longrightarrow$ INSESGADO
- Sesgo$(\hat{\beta}) \neq 0 \longrightarrow E(\hat{\beta}) \neq \beta \longrightarrow$ SESGADO

* Varianza :

Varianza mínima $\longrightarrow$ EFICIENTE

Estimadores insesgado $\rightarrow$ comparamos varianzas

$$Var(\hat{\beta}_1) < Var(\hat{\beta}_2)$$

$\hat{\beta}_1$ es más eficiente que $\hat{\beta}_2$

* ERROR CUADRÁTICO MEDIO

$$ECM(\hat{\beta}) = \left[Sesgo(\hat{\beta})\right]^2 + Var(\hat{\beta})$$

insesgado $\longrightarrow ECM(\hat{\beta}) = Var(\hat{\beta})$

sesgado $\longrightarrow ECM(\hat{\beta}) > Var(\hat{\beta})$

sesgado

$$\boxed{ECM(\hat{\beta}) \geq Var(\hat{\beta})}$$

insesgado

* CONSISTENCIA

$$\lim_{n \to \infty} ECM(\hat{\beta}) = 0 \iff \text{consistente}$$