

INTRODUCCIÓN A LA ECONOMETRÍA

3er curso LE y LADE

Tema 3a

Dpto. de Econometría y Estadística (EA3)

UPV—EHU

3 El Modelo de Regresión Lineal (II). Inferencia y Predicción.

3.1a Distribución del Estimador de Mínimo Cuadrados bajo el supuesto de Normalidad.

Estimador MCO bajo Normalidad

- Si $Y = X\beta + u$, donde $u \sim \mathcal{N}(0, \sigma^2 I_T)$, entonces (recuérdese) el estimador MCO:

$$\begin{aligned}\hat{\beta}_{\text{MCO}} &= (X'X)^{-1}X'Y = \beta + (X'X)^{-1}X'u \\ &= \beta + \Gamma'u \quad \text{es lineal en perturbaciones.}\end{aligned}$$

- Por lo tanto, la misma distribución **Multivariante Normal**, con (recuérdese)

$$\begin{cases} E(\hat{\beta}) &= \beta, \\ \text{Var}(\hat{\beta}) &= \sigma^2(X'X)^{-1}. \end{cases}$$

- Es decir:

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})$$

Estimador MCO bajo Normalidad (casos)

Dado que $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})$:

- Para el k -ésimo coeficiente:

$$\hat{\beta}_k \sim \mathcal{N}(\beta_k, \sigma^2 a_{kk})$$

donde a_{kk} es el $(k+1)$ -ésimo elemento diagonal de $(X'X)^{-1}$

- por ejemplo: $\hat{\beta}_1 \sim \mathcal{N}(\beta_1, \sigma^2 a_{11})$,

$a_{11} = 2^\circ$ elemento diagonal.

- Para un conjunto de combinaciones lineales:

$$R\hat{\beta} \sim \mathcal{N}(R\beta, \sigma^2 R(X'X)^{-1}R')$$

- Para un subvector de $\hat{\beta}$: $R = [0_s \dots 0_s | I_s]$; entonces

$$\hat{\beta}^s \sim \mathcal{N}(\beta^s, \sigma^2 A_{ss})$$

donde $\beta^s =$ subvector de β , $A_{ss} =$ submatriz de $(X'X)^{-1}$.

Estimador MCO bajo Normalidad (casos)2

- En particular, si $R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow$

$$R \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \beta^* \text{ (sin intercepto):}$$

- y

$$(X'X)^{-1} = \begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{pmatrix};$$

- entonces

$$\hat{\beta}^* \sim \mathcal{N}(\beta^*, \sigma^2 \diamond)$$

Residuos MCO bajo Normalidad

- De forma similar, si $u \sim \mathcal{N}(0, \sigma^2 I_T)$,

Entonces,

$$\hat{u} \sim \mathcal{N}(0, \sigma^2 M)$$

- En particular, para el 4º residuo:

$$\hat{u}_t \sim \mathcal{N}(0, \sigma^2 m_{44})$$

donde m_{44} es el 4º elemento diagonal de la matriz M .

3.1b Contrastar Hipótesis: un Repaso.

Hipótesis y Contrastes (rep1)

- Punto de partida:

$$\left. \begin{array}{l} Y = X\beta + u \\ u \sim \mathcal{N}(0, \sigma^2 I_T) \end{array} \right\} \left\{ \begin{array}{l} \hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1}) \\ \hat{u} \sim \mathcal{N}(0, \sigma^2 M) \end{array} \right.$$

- Hipótesis:** "conjetura sobre la fn de dn de(l) parámetro(s)".

Por ejemplo:

- en MRLS: $\hat{\beta} \sim \mathcal{N}(\beta, v)$; sea $\beta = 2,5$.
- en MRLG: $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})$; sea $\beta_1 + \dots + \beta_K = 1$.
- en general: T^a Ec. \rightsquigarrow hipótesis
p.ej.: Fn Cobb-Couglas:

$$Y_t = e^{\beta_0} L_t^{\beta_1} K_t^{\beta_2} e^{u_t}$$

con rendimientos de escala constantes : $\beta_1 + \beta_2 = 1$

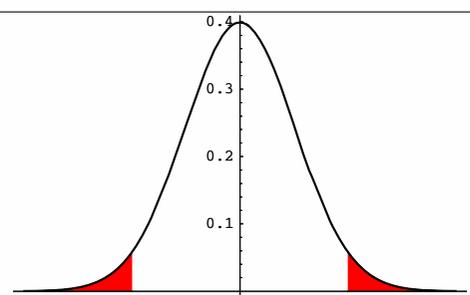
- Contraste:** "procedimiento para rechazar o aceptar la hipótesis"

Hipótesis y Contrastes (rep2)

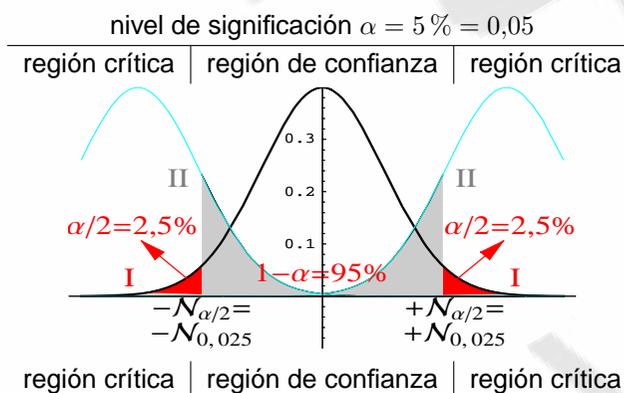
elementos	pasos						
a) hipótesis para contrastar (sobre estimador)	$H_0 : \dots$ frente a $H_a : \dots$ (disjunta)						
b) dn estimador	obtener estadístico de contraste con dn tabulada bajo H_0 :						
c) regla decisión	<p style="text-align: center;">estadístico calculado</p> <table style="width: 100%; border: none;"> <tr> <td style="border: none; text-align: center;">\in región crítica ("grande")</td> <td style="border: none; text-align: center;">\notin región crítica ("pequeño")</td> </tr> <tr> <td style="border: none; text-align: center;">↓</td> <td style="border: none; text-align: center;">↓</td> </tr> <tr> <td style="border: none; text-align: center;">Rechazar</td> <td style="border: none; text-align: center;">no Rechazar</td> </tr> </table>	\in región crítica ("grande")	\notin región crítica ("pequeño")	↓	↓	Rechazar	no Rechazar
\in región crítica ("grande")	\notin región crítica ("pequeño")						
↓	↓						
Rechazar	no Rechazar						

Hipótesis y Contrastes (rep2-cont)

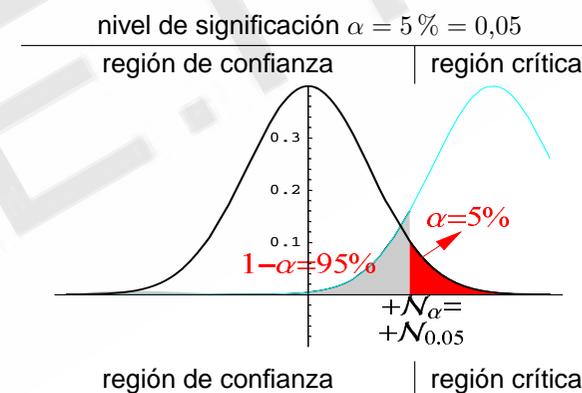
Ejemplo:

a)	$H_0 : \beta = 2,5$ frente a $H_a : \beta \neq 2,5$	$(\text{Var}(\beta)=4)$
b)	$\hat{\beta} \sim \mathcal{N}(\beta, 4) \rightsquigarrow z = \frac{\hat{\beta} - \beta}{2} \sim \mathcal{N}(0, 1)$	
c)	$z = \frac{\hat{\beta} - 2,5}{2} \in$ 	

Hipótesis y Contrastes: Región Crítica



Hipótesis y Contrastes: Región crítica (una cola)



Hipótesis y Contrastes: Distribuciones (repass)

1. Def de χ^2 (ji-cuadrado):

$$\left. \begin{array}{l} Z_i \sim \text{iid } \mathcal{N}(0, 1) \\ Z \sim \mathcal{N}(0, I_m) \end{array} \right\} Z'Z = \sum_{i=1}^m Z_i^2 \sim \chi^2(m) \quad \left\{ \begin{array}{l} E(\chi^2(m)) = m \\ \text{Var}(\chi^2(m)) = 2m \end{array} \right.$$

1b. $Z \sim \mathcal{N}(\mu, \Omega) \Rightarrow (Z - \mu)' \Omega^{-1} (Z - \mu) \sim \chi^2(m)$

2. Def de t (Student): $\left. \begin{array}{l} Z \sim \mathcal{N}(0, 1), \quad W \sim \chi^2(m) \\ Z, W \text{ independientes} \end{array} \right\} \frac{Z}{\sqrt{W/m}} \sim t(m)$

3. Def de \mathcal{F} (Snedecor): $\left. \begin{array}{l} V \sim \chi^2(n), \quad W \sim \chi^2(m) \\ V, W \text{ independientes} \end{array} \right\} \frac{V/n}{W/m} \sim \mathcal{F}_m^n$

4b. $n = 1 \Rightarrow \frac{Z^2}{W/m} \sim \mathcal{F}_m^1 \equiv t(m)^2$

Hipótesis y Contrastes: Resultado útil

De $\hat{u} \sim \mathcal{N}(0, \sigma^2 M)$:

■ $\frac{\text{SCR}}{\sigma^2} = \sum (\hat{u}_t^2 / \sigma^2) = \sum \mathcal{N}(0, 1)^2 \text{'s} \sim \chi^2(T-K-1)$

■ Entonces: $\frac{\hat{\sigma}^2}{\sigma^2} = \frac{\text{SCR}}{\sigma^2(T-K-1)} = \frac{\text{SCR}}{\sigma^2(T-K-1)} = \chi^2/\text{d.f.'s}$

◆ $\frac{\text{expr}}{\sigma} \sim \mathcal{N}(0, 1)$:

◆ $\frac{\text{expr}}{\hat{\sigma}} = \frac{\text{expr}/\sigma}{\hat{\sigma}/\sigma} = \frac{\text{expr}/\sigma}{\sqrt{\hat{\sigma}^2/\sigma^2}} = \frac{\mathcal{N}(0, 1)}{\sqrt{\chi^2/\text{d.f.'s}}} = t$

◆ $\frac{\text{expr}}{\sigma^2} \sim \chi^2(n)$:

◆ $\frac{\text{expr}}{\hat{\sigma}^2} = \frac{\text{expr}/\sigma^2}{\hat{\sigma}^2/\sigma^2} \Rightarrow \frac{\frac{\text{expr}}{\sigma^2}/n}{\hat{\sigma}^2/\sigma^2} = \frac{\chi^2(n)/n}{\chi^2/\text{d.f.'s}} \sim \mathcal{F}$

■ En resumen: $\sigma^2 \rightarrow \hat{\sigma}^2 \Rightarrow \begin{array}{l} !\mathcal{N}(0, 1) \rightarrow t !! \\ !\chi^2 \rightarrow \mathcal{F} !! \end{array}$

3.2a Contraste para Significación de un único parámetro. Intervalos de Confianza.

Contraste de Signif de un parámetro: dn

■ Estandarizar $\hat{\beta}_i \sim \mathcal{N}(\beta_i, \sigma^2 a_{ii})$

$$\frac{\hat{\beta}_i - \beta_i}{\sqrt{\text{Var}(\hat{\beta}_i)}} = \frac{\hat{\beta}_i - \beta_i}{\sigma \sqrt{a_{ii}}} = \frac{\hat{\beta}_i - \beta_i}{\sigma_{\hat{\beta}_i}} \sim \mathcal{N}(0, 1)$$

■ cambiar σ por $\hat{\sigma}$:

$$\frac{\hat{\beta}_i - \beta_i}{\hat{\sigma} \sqrt{a_{ii}}} = \frac{\hat{\beta}_i - \beta_i}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_i)}} = \frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}} \sim t(T-K-1)$$

■ Nota: $\sigma_{\hat{\beta}_i} \rightarrow S_{\hat{\beta}_i} \Rightarrow !\mathcal{N}(0, 1) \rightarrow t !!$

Contraste de Signif de un parámetro: regla

- $$\frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}} \sim t_{(T-K-1)}$$
- ¿Qué Contraste?**

$$\begin{cases} H_0 : \beta_i = c & (\text{contraste informativo}) \\ H_0 : \beta_i = 0 & (\text{contraste de significatividad}) \end{cases}$$
- Recuérdese:** Hipótesis \rightsquigarrow estadístico \rightsquigarrow regla...
- Contraste de Significatividad:**
 - Hipótesis:** $H_0 : \beta_i = 0$ frente a $H_a : \beta_i \neq 0$
 - Estadístico:** $t = \frac{\hat{\beta}_i}{S_{\hat{\beta}_i}} \sim t_{(T-K-1)}$ bajo H_0 :
 - Regla:** $|t| = \left| \frac{\hat{\beta}_i}{S_{\hat{\beta}_i}} \right| > t_{\alpha/2(T-K-1)} \Rightarrow$ rechazar H_0 :
 - $\Rightarrow \beta_i$ es (estadísticamente o significativamente) distinto de cero
 - $\Rightarrow X_i$ es una variable (estadísticamente) relevante o significativa.
- de forma similar para el contraste informativo $H_0 : \beta_i = c$ (Ejercicio: ¡Inténtalo!!)**

Contraste de Signif de un parámetro: regla (cont)

- Regla:** $|t| = \left| \frac{\hat{\beta}_i}{S_{\hat{\beta}_i}} \right| > t_{\alpha/2(T-K-1)} \Rightarrow$ rechazar H_0 :
- nivel de significación $\alpha = 5\% = 0,05$

región crítica | región de confianza | región crítica

$\alpha/2$ $1-\alpha=95\%$ $\alpha/2$

$-t_{\alpha/2}$ $+t_{\alpha/2}$

región crítica | región de confianza | región crítica

Intervalo de confianza para β_i

- Recuérdese que** $\frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}} \sim t_{(T-K-1)}$
- región crítica | región de confianza | región crítica

$\alpha/2$ $1-\alpha=95\%$ $\alpha/2$

$-t_{\alpha/2}$ $+t_{\alpha/2}$

región crítica | región de confianza | región crítica
- es decir:** $\Pr[-t_{\alpha/2} \leq \frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}} \leq +t_{\alpha/2}] = 1 - \alpha$
- $$\Pr[\hat{\beta}_i - t_{\alpha/2} S_{\hat{\beta}_i} \leq \beta_i \leq \hat{\beta}_i + t_{\alpha/2} S_{\hat{\beta}_i}] = 1 - \alpha$$

$IC_{1-\alpha}(\beta_i)$

Intervalo de confianza para β_i (cont)

- Es decir:**

$$IC_{1-\alpha}(\beta_i) = [\hat{\beta}_i \pm t_{\alpha/2} S_{\hat{\beta}_i}]$$
- p.ej. para $\alpha = 5\%$, $T-K-1 = 25$, $\hat{\beta}_i = 2,12$ y $S_{\hat{\beta}_i} = 0,08$:**

$$IC_{95\%}(\beta_i) = [\hat{\beta}_i \pm t_{2,5\%}(25) S_{\hat{\beta}_i}]$$

$$= [\hat{\beta}_i \pm 2,06 S_{\hat{\beta}_i}] = [2,12 \pm 2,06 \cdot 0,08] = [1,9552 ; 2,2848]$$
- contraste por medio del intervalo de confianza:**
- Hipótesis:** $H_0 : \beta_i = c$ frente a $H_a : \beta_i \neq c$
- Intervalo:** $IC_{95\%}(\beta_i)$
- Regla:** Rechazar H_0 : si $c \notin CI_{95\%}(\beta_i)$, con significación del 5%
- p.ej. $\beta_i = 0$?** \Rightarrow Rechazar $\Rightarrow \beta_i$ es significativo (al nivel del 5%).

Contraste de única Combinación Lineal

- Supongamos MRLG restringido con 1 restricción ($q = 1$):
 $R\beta = r$ pero ahora más sencillo...
 $R = d'$ (cualquier fila de $K+1$ valores d_0, d_1, \dots, d_K) y
 $r = c$ (cualquier valor escalar);
- Sea $H_0 : \nu = d'\beta = d_0\beta_0 + d_1\beta_1 + \dots + d_K\beta_K = c$
 es decir,
 un contraste informativo sobre el valor c que toma una única combinación lineal ν de los parámetros.

Contraste de única Combinación Lineal: Ejemplo

- Sea la fn de Cobb-Douglas linealizada

$$\log Y_t = \alpha + \beta_L \log L_t + \beta_K \log K_t + u_t$$

$$d' = [0 \quad 1 \quad 1] \text{ y } c = 1:$$

$$H_0 : \nu = d'\beta = [0 \quad 1 \quad 1] \begin{pmatrix} \alpha \\ \beta_L \\ \beta_K \end{pmatrix} = \beta_L + \beta_K = c = 1$$

- es decir, $H_0 : \beta_L + \beta_K = 1$;
 el contraste de la hipótesis de **rendimientos de escala constantes**.

Contraste de única Combinación Lineal: dn

- Dado que $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2(X'X)^{-1})$, tenemos que

$$d'\hat{\beta} \sim \mathcal{N}(d'\beta, \sigma^2 d'(X'X)^{-1}d)$$

$$\hat{\nu} \sim \mathcal{N}(\nu, \text{Var}(\hat{\nu}))$$

$$\text{donde } \text{Var}(\hat{\nu}) = \sigma^2 \sum_{i,j=0}^K d_i d_j a_{ij}$$

- Como antes, estandarizando $\hat{\nu}$

$$\frac{\hat{\nu} - \nu}{\sqrt{\text{Var}(\hat{\nu})}} \sim \mathcal{N}(0, 1)$$

- Por lo tanto (recuérdese que $\sigma \rightarrow \hat{\sigma}$):

$$\Rightarrow \frac{\hat{\nu} - \nu}{S_{\hat{\nu}}} \sim \mathbf{t}(T-K-1)$$

$$\text{donde } S_{\hat{\nu}} = \hat{\sigma} \sqrt{\sum_{i,j=0}^K d_i d_j a_{ij}}$$

Contraste de única Combinación Lineal: regla

- $\frac{\hat{\nu} - \nu}{S_{\hat{\nu}}} \sim \mathbf{t}(T-K-1)$
- ¿Qué Contraste? $\{H_0 : \nu (= d'\beta) = c \text{ (contraste informativo)}$

- Recuérdese:** Hipótesis \rightsquigarrow estadístico \rightsquigarrow regla...

- Contraste de una combinación lineal:

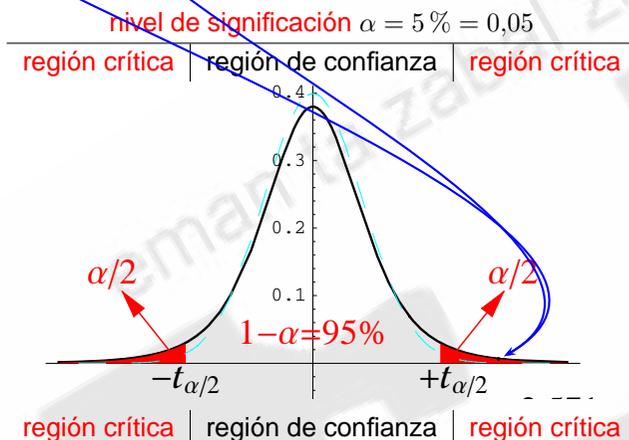
- ♦ **Hipótesis:** $H_0 : \nu = c$ frente a $H_a : \nu \neq c$
- ♦ **Estadístico:**

$$t = \frac{\hat{\nu} - c}{S_{\hat{\nu}}} \sim \mathbf{t}(T-K-1) \text{ bajo } H_0:$$

- ♦ **Regla:** $|t| > \mathbf{t}_{\alpha/2}(T-K-1) \Rightarrow \text{rechazar } H_0$;
 \Rightarrow valor de la combinación lineal no es correcta.
- ♦ **compárese** con el contraste de un único parámetro β_k , ¿algunas similitudes?.

Contraste de única Combinación Lineal: regla (c)

- Regla: $|t| = \left| \frac{\hat{v} - c}{S_{\hat{v}}} \right| > t_{\alpha/2}(T-K-1) \Rightarrow$ rechaza H_0 :



Contraste de única Combinación Lineal: Ejemplo

- En la fn de Cobb-Douglas linealizada:
 $\widehat{\log Y}_t = \hat{\alpha} + \hat{\beta}_L \log L_t + \hat{\beta}_K \log K_t, \quad T = 53;$
- $\widehat{\log Y}_t = 2,10 + 0,67 \log L_t + 0,32 \log K_t, \quad \hat{\sigma}^2 = 4;$

$$(X'X)^{-1} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & -1 \\ 0 & -1 & 7 \end{pmatrix}$$

- Contrastar la H_0 : rendimientos de escala constantes al nivel de significación del $\alpha = 5\%$:

Contraste de única Combinación Lineal: Ej (cont)

- Hipótesis: $H_0 : \beta_L + \beta_K = 1$ frente a $H_a : \beta_L + \beta_K \neq 1$
- Estadístico:

$$\begin{aligned} \hat{v} &= \hat{\beta}_L + \hat{\beta}_K \\ &= 0,67 + 0,27 = 0,89 \end{aligned}$$

$$\begin{aligned} S_{\hat{v}} &= \sqrt{\text{Var}(\hat{\beta}_L) + \text{Var}(\hat{\beta}_K) + 2\text{Cov}(\hat{\beta}_L, \hat{\beta}_K)} \\ &= \hat{\sigma} \sqrt{a_{11} + a_{22} + 2a_{12}} \\ &= 2\sqrt{4 + 7 + 2(-1)} = 2\sqrt{9} = 6 \end{aligned}$$

$$\begin{aligned} t &= \frac{\hat{v} - 1}{S_{\hat{v}}} \\ &= \frac{0,89 - 1}{6} = \frac{-0,11}{6} = -0,018. \end{aligned}$$

- Regla: $|t| = 0,018 < t_{0,025}(50) = 2,01 \Rightarrow$ no rechazar H_0 :
 \Rightarrow "rendimientos de escala constantes" es apoyado por datos.

3.2b Contraste de Significatividad Conjunta.

Contraste de Significatividad Conjunta: dn

- $H_0 : \beta_1 = \beta_2 = \dots = \beta_K = 0 \rightsquigarrow$
- $H_0 : \beta^* = \mathbf{0} \rightsquigarrow$

$$\hat{\beta}^* \sim \mathcal{N}(\mathbf{0}, \sigma^2 \begin{bmatrix} a_{00} & a_{01} & \dots & a_{0K} \\ a_{10} & a_{11} & \dots & a_{1K} \\ \vdots & \vdots & \dots & \vdots \\ a_{K0} & a_{K1} & \dots & a_{KK} \end{bmatrix}) \sim \mathcal{N}(\mathbf{0}, \sigma^2 (x'x)^{-1})$$

- Estandarizar y construir Suma Cuadrados:

$$\frac{\hat{\beta}^{*'} x' x \hat{\beta}^*}{\sigma^2} \sim \chi^2(K) \text{ bajo } H_0 :$$

- Entonces (recuérdese el cambio $\sigma^2 \rightarrow \hat{\sigma}^2$):

$$F = \frac{\hat{\beta}^{*'} x' x \hat{\beta}^* / K}{\hat{\sigma}^2} \sim \mathcal{F}_{TK-1}^K$$

Contraste de Significatividad Conjunta: regla

$$F = \frac{\hat{\beta}^{*'} x' x \hat{\beta}^* / K}{\hat{\sigma}^2} \sim \mathcal{F}_{TK-1}^K \text{ bajo } H_0 :$$

- **Contraste de significatividad conjunta:** $\{H_0 : \beta^* = 0\}$
- **Recuérdese:** Hipótesis \rightsquigarrow estadístico \rightsquigarrow regla...
 - ◆ **Hipótesis:** $H_0 : \beta^* = 0$ frente a $H_a : \beta^* \neq 0$ (es decir $\exists \beta_i \neq 0$)
 - ◆ **Estadístico:**

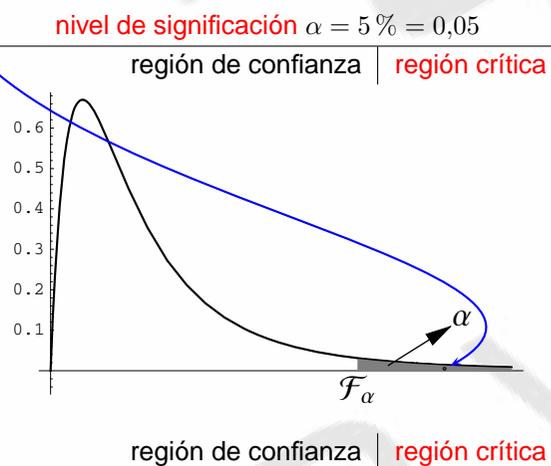
$$F = \frac{\hat{\beta}^{*'} x' x \hat{\beta}^* / K}{\hat{\sigma}^2} = \frac{\hat{y}' \hat{y} / K}{\hat{u}' \hat{u} / (T-K-1)} = \frac{\text{SCE} / K}{\text{SCR} / (T-K-1)}$$

$$= \frac{(\text{SCE}/\text{SCT}) / K}{(\text{SCR}/\text{SCT}) / (T-K-1)} = \frac{R^2 / K}{(1 - R^2) / (T-K-1)} \sim \mathcal{F}_{TK-1}^K \text{ bajo } H_0 :$$

- ◆ **Regla:** $F > \mathcal{F}_{\alpha}(K, T-K-1) \Rightarrow$ rechazar H_0 :
 - \Rightarrow todos los coefs son conjuntamente significativos (diferentes de cero)
 - \Rightarrow la regresión es (estadísticamente) relevante.

Contraste de Significatividad Conjunta: regla (co

- **Regla:** $F > \mathcal{F}_{\alpha}(K, T-K-1) \Rightarrow$ rechazar H_0 :



Contraste de Significatividad Conjunta: Ejemplo

- En el ejemplo anterior (fn de Cobb-Douglas linealizada):

$$\widehat{\log Y}_t = \hat{\alpha} + \hat{\beta}_L \log L_t + \hat{\beta}_K \log K_t, \quad T = 53;$$

$$\widehat{\log Y}_t = 2,10 + 0,67 \log L_t + 0,32 \log K_t, \quad \hat{\sigma}^2 = 4; R^2 = 0,88$$

- Contrastar significatividad conjunta

con el $\alpha = 5\%$ de significación:

$$F = \frac{R^2 / K}{(1 - R^2) / (T-K-1)}$$

$$= \frac{0,88 / 2}{(1 - 0,88) / (50)} = \frac{0,44}{0,024} = 183,33 > \mathcal{F}_{0,05}(2, 50) = 3,19$$

- $\Rightarrow \beta_K$ y β_L son conjuntamente significativos
- \Rightarrow regresión es relevante (estadísticamente).