

Tema 3

Regresión lineal simple: estadística y contraste de hipótesis

(actualizadas el 07-07-2023)

Tema 3. Regresión lineal simple: estadística y contraste de hipótesis

3.1 Supuestos del modelo lineal clásico

3.2 Propiedades probabilísticas del modelo

3.3 Distribución muestral de los estimadores MCO

3.4 Contrastes de hipótesis sobre un solo parámetro: el estadístico t

Bibliografía

- Ezequiel Uriel (2013): Capítulo 2 (epígrafe 2.5) y Capítulo 4 (4.1 y 4.2)
- Wooldridge (2015): Capítulo 2 (2.5) y Capítulo 4 (4.1 a 4.3)
- Stock y Watson (2012): Capítulo 4 (4.4 y 4.5) y Capítulo 5 (5.1 y 5.2)

3.1 Supuestos del modelo lineal clásico

... o hipótesis estadísticas básicas

en el tema 2 ...

- Recordamos nuestro planteamiento: partimos de que $y = f(x)$ y queremos cuantificar/estimar el efecto de x sobre y . Para ello planteamos un MRLS: $y = \beta_1 + \beta_2 x + u$
- OK, sabemos obtener estimaciones, PERO **¿me puedo fiar de las estimaciones?** Si, por ejemplo $\hat{\beta}_2 = 0.3$, ¿cuán seguro estoy de que el efecto de x sobre y , de que β_2 , es 0.3? ¿Estamos seguros de que no es 0.31 o 0.29 ... ?
- Es difícil que β_2 sea exactamente 0.3; como mucho, si MCO fuese un buen método de estimación podríamos pensar que efectivamente β_2 estará próximo a 0.3, pero **¿cuanto de próximo?** ¿Podemos dar un rango de valores entre los que con una alta probabilidad se encuentre β_2 ?
- Al final del tema podremos, pero antes, para ver la importancia de esto, imaginaros que la estimación puntual fuese $\hat{\beta}_2 = 0.3$ pero este resultado puede ser compatible con una **estimación por intervalos** de $[0.25; 0.35]$ o, si hay mucha incertidumbre, de $[-0.2, 0.8]$.

tenemos que cerrar (o ampliar) el MRLS

- Como veis entramos en el terreno de la probabilidad y la **inferencia**, pero con el modelo tal como lo tenemos ahora no podemos responder a este tipo de preguntas, no podemos avanzar más, no podemos contestar a las preguntas anteriores
- Para poder avanzar y hacer más útil el MRL, Tenemos que **incorporar al modelo una serie de supuestos o hipótesis**
- Comenzaremos añadiendo un conjunto de hipótesis "sencillas" que puede que no sean del todo realistas al analizar un determinado fenómeno económico, pero aún así comenzaremos con ellas y **las revisitaremos en los últimos temas.**
- Este conjunto de hipótesis sencillas se las conoce como **hipótesis estadísticas básicas** (h.e.b) o clásicas

MRL + h.e.b

- Al MRL + las h.e.b se le conoce como **modelo lineal** básico o **clásico**.
- Un resultado **muy importante** que obtendremos consiste en que **si se cumplen** las hipótesis estadísticas básicas (h.e.b) el método de MCO es un "buen" método para estimar los efectos de x en y
- En concreto, veremos que **si se cumplen las h.e.b, los estimadores MCO del MRL son ELIO**.
- Por contra, si alguna de las h.e.b no se cumple, los estimadores MCO **pueden** dejar de ser ELIO
- ELIO significa **estimador lineal, insesgado y óptimo**

Veamos cuales son las h.e.b

Hipótesis estadísticas básicas (h.e.b)

I) Hipótesis sobre la **forma funcional**, sobre el modelo

- 1) El modelo es: $y_i = \beta_1 + \beta_2 x_i + u_i$

Esta hipótesis parece una tautología. Todo nuestro análisis es condicional a que el modelo que estamos planteando es correcto. Hipótesis de **correcta especificación**. Esta hipótesis implica o supone que :

- el modelo (la relación entre x e y) es **lineal en parámetros**
- **no hay variables omitidas** ni variables irrelevantes

II) Hipótesis sobre la **perturbación** (u): es el bloque principal de hipótesis, al que más tiempo les dedicaremos. Contiene 5 hipótesis: de la segunda a la sexta

III) Hipótesis sobre los **regresores** (x): son 3 hipótesis: de la séptima a la novena. Diremos muy poco sobre ellas

IV) Hipótesis sobre los **parámetros** (β)

- 10) Los parámetros del modelo son fijos (!!)

Hipótesis sobre la perturbación (u)

- 2) Las u_i son v.a. no observables.
- 3) $E(u_i) = 0$
- 4) $Var(u_i) = \sigma^2$ para $i = 1, \dots, N$ **(HOMOCEDASTICIDAD)**
- 5) $Cov(u_i, u_j) = 0$ para $i \neq j$ **(NO AUTOCORRELACIÓN)**
- 6) $u_i \longrightarrow N$ **(NORMALIDAD)**

Todas ellas se pueden expresar conjuntamente como:

$$u_i \longrightarrow N(0, \sigma^2)$$

Hipótesis sobre los regresores (x)

- 7) Los regresores son no estocásticos, o sea, los **regresores son fijos** (!!).
 - 7*) Los regresores se distribuyen independientemente del término de perturbación: $E(x, u) = 0$
- 8) La matriz de datos de los regresores debe cumplir que:
 - 8.1) $N \geq k$ (Hay que tener al menos tantas observaciones como parámetros β)
 - 8.2) Los k regresores deben ser **linealmente independientes**; es decir, no pueden existir relaciones lineales exactas entre los regresores. **(NO COLINEALIDAD PERFECTA)**
- 9) Los regresores no tienen errores de medida.

3.2 Propiedades probabilísticas del modelo

No confundir con las propiedades descriptivas

¿por qué son tan importantes las h.e.b?

- En seguida veremos que **si se cumplen las h.e.b** se obtienen varios resultados (propiedades probabilísticas) que nos permitirán hacer inferencia sobre los β y predicciones sobre y
- Entre los resultados que obtendremos destaca uno de ellos: demostraremos que **si se cumplen las h.e.b**, entonces el método de MCO es un “buen” método para estimar los parámetros de un MRL; en concreto **los estimadores MCO serán insesgados y óptimos (ELIO)**.
- Este resultado fue obtenido en el teorema de Gauss-Markov. El teorema demuestra que en un MRL, el estimador MCO (MCO) de los β es el estimador lineal e insesgado **óptimo**; es decir, el estimador MCO es el estimador eficiente dentro de la clase de estimadores lineales e insesgados.
- Puntualización: para el teorema de Gauss-Markov no es necesaria la hipótesis de normalidad
- ELIO o BLUE: best linear unbiased estimator.

¿entendéis por qué es importante que sean ELIO? lo vemos enseguida

propiedades probabilísticas (I): distribución de y

- En el MRL, **si se cumplen TODAS las h.e.b**, tenemos que:

$$y_i \longrightarrow N(\beta_1 + \beta_2 x_i, \sigma^2)$$

- En palabras: el regresando (y) es una v.a que se distribuye como una Normal con valor esperado igual a $E(\beta_1 + \beta_2 x_i)$ y varianza igual a σ^2

propiedades probabilísticas (II): los $\hat{\beta}$ son ELIO

- esto ya lo sabíamos, es el teorema de Gauss-Markov, pero sí además se cumple la hipótesis de normalidad, entonces ...

propiedades probabilísticas (III): distribución de $\hat{\beta}$

- En el MRL, **si se cumplen TODAS las h.e.b**, tenemos que:

$$\hat{\beta}_j \longrightarrow N(\beta_j, \sigma_{\hat{\beta}_j}^2)$$

- En palabras:

¿entendéis por qué son importantes las propiedades probabilísticas?

- A partir de la propiedad probabilísticas (I) podremos efectuar predicciones sobre y
- A partir de la propiedad probabilísticas (III) podremos efectuar estimación por intervalos y contrastes de hipótesis sobre los parámetros (β)

¿por qué es importante que los $\hat{\beta}$ sean ELIO?

- para entenderlo hay que saber que significa insesgadez y optimalidad cuando hablamos de un estimador.

Insesgadez

- Un estimador es insesgado si $E(\hat{\beta}) = \beta$
- Parece fácil, pero ...

Insesgadez

- Un estimador es insesgado si $E(\hat{\beta}) = \beta$
- Los estimadores MCO son variables aleatorias (!!) ya que dependen de u . Según los valores concretos que tome la u , según la muestra concreta que utilicemos, obtendremos unas estimaciones concretas.
- La insesgadez **no nos garantiza** que las estimaciones acertarán, no. Lo que significa la insesgadez es que si dispusiésemos de muchas muestras de datos, entonces podríamos estimar muchas veces, y entonces, la media de esas estimaciones acertaría.
- Repito: con una estimación concreta, con una muestra concreta, no sabemos si acertaremos, de hecho, casi seguro que no acertaremos, pero si que sabemos que teóricamente, **si estimásemos muchas veces, acertaríamos en media**. La media de las estimaciones tendería a acertar.

Insesgadez

- La propiedad de insesgadez "no es la bomba", solamente nos garantiza que con muchas muestras tenderíamos a acertar.
 - Lo que si que no sería deseable es tener un estimador sesgado; es decir, que ni siquiera acertase en media, que ni siquiera tendiese a acertar si tuviésemos muchas muestras.

puntualizaciones sobre la insesgadez

- Una estimación no puede ser insesgada, ya que una vez se ha estimado es un valor concreto
- la insesgadez es una propiedad de los estimadores; del procedimiento por el cual obtenemos las estimaciones
- Con una sola muestra, que es lo habitual, no podemos garantizar que nuestra estimación esté cercana al parámetro que queremos estimar.

Insesgadez junto con optimalidad

- De acuerdo, la falta de sesgo es una propiedad deseable, pero tampoco es la bomba, **PERO** recordad que Gauss-Markov además implica que el estimador es óptimo: que la varianza del estimador es la mínima (dentro de la clase de estimadores ELI)
- Tener las dos propiedades juntas (insesgadez y mínima varianza) si es realmente interesante, ya que implica que los estimadores MCO son los que, a priori, maximizan la probabilidad de acertar con una sola estimación; por lo tanto hace que "nos fiemos" de ellos y los usemos.
- Los estimadores MCO se usan porque, **si se cumplen las h.e.b**, son ELIO, lo que hace que sean los estimadores que **maximizan la probabilidades de "acertar"**.

puntualizaciones sobre la optimalidad

- Los estimadores MCO, bajo las h.e.b, son óptimos, son los que tienen menor varianza, OK, pero aún así, nada nos garantiza que la varianza sea pequeña
- El tamaño de la varianza depende de $\sigma_{\hat{\beta}_j}^2 = \frac{\sigma^2}{N \text{Var}(x_j) (1-R_j^2)}$

3.3 Distribución muestral de los estimadores MCO

Ya la hemos visto como propiedad probabilística (III)

Distribución muestral de $\hat{\beta}$

- En el MRL, **si se cumplen TODAS las h.e.b**, tenemos que:

$$\hat{\beta}_j \longrightarrow N(\beta_j, \sigma_{\hat{\beta}_j}^2)$$

- En palabras: los estimadores, el estimador MCO de β_j , se distribuye de forma Normal, con valor esperado $E(\hat{\beta}_j) = \beta_j$; es decir, es insesgado, y con una varianza que denotamos por $\sigma_{\hat{\beta}_j}^2$
- Con este resultado **podremos** en breve hacer inferencia sobre los β , pero ... ¿que nos falta para ello?

falta aproximar/estimar $\sigma_{\hat{\beta}_j}^2$

- Puede demostrarse que $\sigma_{\hat{\beta}_j}^2 = \frac{\sigma^2}{N \text{Var}(x_j) (1-R_j^2)}$

estimando, obteniendo un estimador para $\sigma^2_{\hat{\beta}_j}$

- La varianza del estimador $\sigma^2_{\hat{\beta}_j} = \frac{\sigma^2}{N \text{Var}(x_j) (1-R_j^2)}$ depende de cuatro componentes:
 - El tamaño muestral N
 - La varianza del regresor cuyo efecto queremos calcular $\text{Var}(x_j)$
 - La colinealidad de x_j con el resto de regresores: $(1 - R_j^2)$
 - La varianza de las perturbaciones σ^2
- ¿Podemos calcular la $\text{Var}(\hat{\beta}_j)$? ¿Que nos falta para poder calcularla/estimarla?

para calcular/estimar $\text{Var}(\hat{\beta}_j)$ necesito primero estimar σ^2

Estimando σ^2

- Esto parece complicado, ya que las u son variables no observables
- Pero los residuos (\hat{u}) constituyen una aproximación adecuada a u
- Un estimador insesgado para σ^2 es : $\hat{\sigma}^2 = \frac{SCR}{N-k}$

¿recuerdas para que necesitamos $\hat{\sigma}^2$

- Si queremos hacer inferencia/contrastes sobre los β , necesitamos estimar su varianza, $Var(\hat{\beta}_j)$, y para ello necesitamos estimar la varianza de las perturbaciones

Finalmente, el estimador de la varianza de los estimadores (!!)

- Por lo tanto,

$$\hat{\sigma}_{\hat{\beta}_j}^2 = \frac{\hat{\sigma}^2}{N Var(x_j) (1 - R_j^2)} = \frac{SCR/(N - k)}{N Var(x_j) (1 - R_j^2)}$$

¿Cómo ha estimado Gretl la desviación típica de los estimadores?

Modelo 4: MCO, usando las observaciones 1-526

Variable dependiente: salario

	coeficiente	Desv. típica	Estadístico t	valor p	
const	-1,60221	0,731067	-2,192	0,0289	**
educacion	0,555099	0,0500552	11,09	8,57e-26	***
experiencia	0,0187467	0,0120402	1,557	0,1201	
antiguedad	0,138829	0,0211572	6,562	1,29e-10	***
mujer	-1,74241	0,266818	-6,530	1,57e-10	***
casada	0,556565	0,286742	1,941	0,0528	*
color	-0,0658096	0,426567	-0,1543	0,8775	
Media de la vble. dep.	5,896103	D.T. de la vble. dep.	3,693086		
Suma de cuad. residuos	4523,823	D.T. de la regresión	2,952359		
R-cuadrado	0,368218	R-cuadrado corregido	0,360914		
F(6, 519)	50,41424	Valor p (de F)	8,27e-49		
Log-verosimilitud	-1312,288	Criterio de Akaike	2638,576		
Criterio de Schwarz	2668,433	Crit. de Hannan-Quinn	2650,267		

- La desviación típica (o mejor, error estándar) del estimador nos informa de la **credibilidad de la estimación** y el margen de error que podemos esperar en las estimaciones.
- Para calcular las $\hat{\sigma}_{\hat{\beta}_j}^2$ hace falta tener $\hat{\sigma}^2$. Gretl ofrece $\hat{\sigma}^2$ en el epígrafe "D.T. de la regresión"
- Recuerda que podemos calcular $\hat{\sigma}^2$ como $\frac{SCR}{N-k}$

3.4 Contrastes de hipótesis sobre un solo parámetro: el estadístico t

es uno de los usos más habituales de un modelo de regresión

generalmente queremos hacer inferencia, queremos hacer preguntas a nuestro modelo

- Hasta ahora sabíamos estimar por MCO, y sabíamos que si se cumplen las h.e.b los estimadores MCO son los más fiables, pero ...

Modelo 4: MCO, usando las observaciones 1-526				
Variable dependiente: salario				
	coeficiente	Desv. típica	Estadístico t	valor p
const	-1,60221	0,731067	-2,192	0,0289 **
educacion	0,555099	0,0500552	11,09	8,57e-26 ***
experiencia	0,0187467	0,0120402	1,557	0,1201
antigüedad	0,138829	0,0211572	6,562	1,29e-10 ***
mujer	-1,74241	0,266818	-6,530	1,57e-10 ***
casada	0,556565	0,286742	1,941	0,0528 *
color	-0,0658096	0,426567	-0,1543	0,8775
Media de la vble. dep.	5,896103	D.T. de la vble. dep.	3,693086	
Suma de cuad. residuos	4523,823	D.T. de la regresión	2,952359	
R-cuadrado	0,368218	R-cuadrado corregido	0,360914	
F(6, 519)	50,41424	Valor p (de F)	8,27e-49	
Log-verosimilitud	-1312,288	Criterio de Akaike	2638,576	
Criterio de Schwarz	2668,433	Crit. de Hannan-Quinn	2650,267	

- ¿realmente el efecto de la educación sobre el salario es 0.55?
- ¿estamos relativamente seguros de que la educación tienen efecto positivo en el salario?
- La experiencia, ¿tiene efecto en el salario?
- ¿Hay brecha salarial por genero?

Todas son cuestiones sobre el fenómeno económico analizado, pero en términos del modelo son cuestiones acerca del valor de los parámetros (β)

Contrastes de hipótesis (recordando algunas ideas)

- Los contrastes serán siempre **sobre los parámetros poblacionales** (β): se toma una decisión sobre alguna característica de la población en función de los resultados de una muestra
- Las restricciones que se contrastan se recogen en la **hipótesis nula** (H_0)
- Se define también una **hipótesis alternativa** (H_1) que será la conclusión del contraste si la evidencia está lo suficientemente en contra de la H_0
- Para poder hacer un contraste se necesita algo, **un estadístico, con distribución conocida** si la H_0 fuese cierta (**bajo la H_0**)

Contrastes de de hipótesis (recordando algunas ideas)

- También se necesita una **regla de decisión** sobre si rechazar o no la H_0
- Generalmente se especifica un **nivel de significación** (α) que indica el margen tolerancia frente al error tipo I (rechazar la H_0 cuando está es cierta)
- El nivel de significación (α) junto con la hipótesis alternativa (H_1) definen la **región de rechazo**
- Si el estadístico de prueba toma un valor que pertenece a la región crítica, entonces rechazaremos la H_0 al $\alpha\%$
- Por contra, si el valor muestral del estadístico de prueba no pertenece a la región crítica, entonces no podremos rechazar la H_0 (al $\alpha\%$)

Contrastes de de hipótesis (3 etapas)

Podemos pensar que un contraste de hipótesis tienen o conlleva **tres etapas**:

- 1) **Establecer la H_0 y H_1** (hay que pasar la pregunta económica a formato estadístico)
- 2) Elegir un **estadístico de prueba** (que podamos calcular, que sea factible calcular y con distribución conocida bajo la H_0)
- 3) Definir la **regla de decisión** (que nos permitirá rechazar (o no) la H_0). Esto en la práctica consiste en fijar α , generalmente al 5%.
 - En la práctica, se divide el espacio en dos regiones (región de no rechazo y región crítica); si el valor del estadístico cae en la región crítica, entonces, rechazaré la hipótesis nula. (Si el valor del estadístico no cae en la región crítica no podré rechazar)

Contrastes de de hipótesis (rechazar o no rechazar)

Al final, cuando hagamos un contraste de hipótesis, rechazaremos o no la H_0 , PERO, un contraste no nos dará nunca una certeza 100%; solo nos permite tomar una decisión en función de si la H_0 es más o menos compatible o no con los datos que tenemos

- Si rechazamos la H_0 al $\alpha\%$ indica que con la datos, con la muestra que tenemos, solo hay un $\alpha\%$ de probabilidad de que H_0 sea cierta; es decir, la evidencia muestral está lo suficientemente en contra de la H_0 como para decir que los datos rechazan la H_0 .
- Si no rechazamos la H_0 no significa que estemos seguros 100% de que esta sea cierta; si no que la evidencia no está lo suficientemente en contra de la H_0 como para rechazarla.

Contrastes de de hipótesis (p-value)

- Los ordenadores, el software estadístico como Gretl, cuando les pedimos que nos hagan un contraste de hipótesis, nos ofrecen, además del valor que toma el estadístico, nos ofrecen el **p-value** (en Gretl lo llama "valor p")
- Al **p-value** se le conoce también como **nivel de significación crítico** (α')
- Si tenemos el **p-value asociado a un contraste**, entonces no tenemos necesidad de consultar las tablas estadísticas.
- Para un determinado nivel de significación (α):
 - Rechazaremos la H_0 si $\text{p-value} < \alpha$
 - No rechazaremos la H_0 si $\text{p-value} > \alpha$
- El **p-value** o nivel de significación crítico (α') es un indicador del nivel de admisibilidad de la H_0
- Cuanto mayor sea el **p-value** mayor confianza tenemos en que la H_0 es cierta y, por tanto, más complicado será rechazarla

Contrastes de de hipótesis (recordando)

- Supón que se ha estimado un MRL por MCO y $\hat{\beta}_2 = 0.5$ ¿Es $\beta_2 = 0.5$?
- Supón que se ha estimado un MRL por MCO y $\hat{\beta}_2 = 0.5$ ¿Rechazaremos la $H_0 : \beta_2 = 0.7$?
- Supón que se ha estimado un MRL por MCO y se ha rechazado que $H_0 : \beta_2 = 0.7$ ¿Estamos completamente seguros de que β_2 no es 0.7?
- Supón que se ha estimado un MRL por MCO y el t -ratio para contrastar $H_0 : \beta_2 = 0.7$ es 1.5. ¿Rechazamos la H_0 ?
- Supón que se ha estimado un MRL por MCO y el p-value asociado al t -ratio para contrastar $H_0 : \beta_2 = 0.7$ es 0.09. ¿Rechazamos la H_0 ?

obteniendo el t -ratio

- Bajo las h.e.b, $\hat{\beta}_j \longrightarrow N(\beta_j, \sigma_{\hat{\beta}_j}^2)$
- Por lo tanto, $\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma_{\hat{\beta}_j}^2}} \longrightarrow N(0, 1)$
- O, lo que es lo mismo: $\frac{\hat{\beta}_j - \beta_j}{\sigma_{\hat{\beta}_j}} \longrightarrow N(0, 1)$
- Si sustituimos $\sigma_{\hat{\beta}_j}$ por su estimador $\hat{\sigma}_{\hat{\beta}_j}$ tenemos el t -ratio:

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}_{\hat{\beta}_j}^2}} = \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}} \longrightarrow t_{N-k}$$

- El t -ratio es ...

Contrastes sobre un único parámetro con el t -ratio

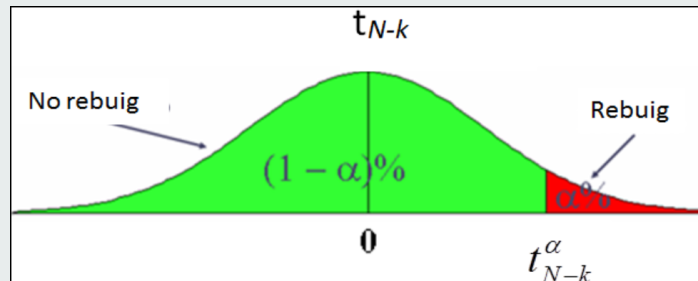
- Con el t -ratio podemos efectuar contrastes del tipo: $H_0 : \beta_j = \beta_j^0$
- Por ejemplo: $H_0 : \beta_3 = 7$
- Utilizaremos el estadístico t junto con la correspondiente regla de rechazo para determinar si rechazamos o no la hipótesis nula, H_0
- Además de la hipótesis nula (H_0) necesitamos una alternativa (H_1) y un nivel de significación (α)

El estadístico t : alternativas de una y dos colas

- La H_1 puede ser a una o dos colas
 - $H_1 : \beta_2 > 4$ y $H_1 : \beta_2 < 2$ son alternativas de **una cola**
 - $H_1 : \beta_2 \neq 5$ es una alternativa a **dos colas**

El estadístico t: alternativas de una cola (cola derecha)

- Por ejemplo: $H_0 : \beta_j = 0$ frente a $H_1 : \beta_j > 0$
- Rechazaremos si observamos un valor del estadístico “suficientemente” alejado de cero por la derecha. Valores negativos del estadístico no proveen evidencia a favor de H_1
- Con siempre hay que fijar el nivel de significación (α), o probabilidad de rechazar H_0 cuando en realidad es cierta. Habitualmente α se fija en el 5%
- Tras seleccionar un nivel de significación, α , buscamos el percentil $(1 - \alpha)$ -esimo en las tablas de la distribución apropiada (en este caso una t de student con $(N - k)$ grados de libertad, y le denominamos valor crítico .
- Rechazaremos la nula si el valor del estadístico t es mayor que el valor crítico. Si el estadístico t es menor que el valor crítico, no rechazamos la nula.



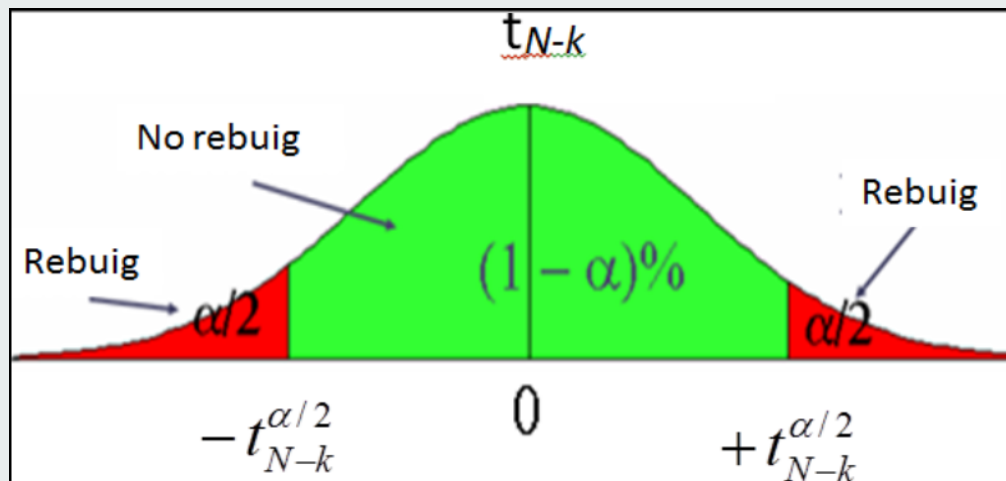
El estadístico t: alternativas de una cola (cola izquierda)

- Por ejemplo: $H_0 : \beta_j = 0$ frente a $H_1 : \beta_j < 0$
- Rechazaremos si observamos un valor del estadístico “suficientemente” alejado de cero por la izquierda. Valores positivos del estadístico no proveen evidencia a favor de H_1
- Evidentemente rechazaríamos la nula si el estadístico t en la muestra toma un valor menor que $-t_{N-k}^{\alpha}$
- No se rechazaría la H_0 si el valor del estadístico en mi muestra es mayor que $-t_{N-k}^{\alpha}$

Dibuje usted mismo la zona de rechazo y de no rechazo

Una cola vs. dos colas

- Cuando la alternativa es a una cola, la región de rechazo se concentra en una cola de la distribución. Además el signo del estadístico t es importante.
- Si H_1 se especifica a dos colas , aunque el contraste se haga al $\alpha\%$, el valor crítico (el de tablas) estará basado en $\alpha/2$.
- Rechazaremos $H_0 : \beta_j = 3$ frente a $H_1 : \beta_j \neq 3$ si el valor del estadístico **en valor absoluto** supera el valor critico; es decir si $|t - ratio| > t_{N-k}^{\alpha/2}$



Cálculo de p-valores para contrastes t

- Hemos repasado el "enfoque clásico" contrastes de hipótesis, que se basa en, tras especificar las hipótesis nula y alternativa, escoger un nivel de significación (α) que determina la región crítica, para luego comparar el valor muestral del estadístico con el valor crítico (de tablas) y concluir que la H_0 se rechaza o no al $\alpha\%$
- En cierto sentido, el enfoque clásico es arbitrario, pues se ha de fijar α
- Una vez fijado $\alpha\%$, la H_0 es rechazada o no, pero no sabemos si el rechazo o no rechazo es fuerte o débil.
- En lugar de fijar $\alpha\%$, consideremos la siguiente cuestión: dado el valor del estadístico, ¿cuál es el menor nivel de significación al que rechazaríamos la nula?
- Ese nivel se conoce como p-valor del contraste ("probabilidad de encontrar un valor que sea mayor al estadístico estimado")
- Una vez que el p-valor ha sido calculado, es sencillo realizar un contraste clásico. Para cualquier nivel de significación se rechazará la H_0 si $\text{p-valor} < \alpha\%$

contraste de significatividad individual

- El contraste de **significatividad individual** contrasta si un parámetro del modelo es **no nulo**. Por ejemplo:

$$\begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases}$$

- Es **uno de los contrastes más habituales**. Y como tal, **lo realiza por defecto Gretl** y cualquier programa estadístico que estime modelos de regresión.

Modelo 4: MCO, usando las observaciones 1-526				
Variable dependiente: salario				
	coeficiente	Desv. típica	Estadístico t	valor p
-----	-----	-----	-----	-----
const	-1,60221	0,731067	-2,192	0,0289 **
educacion	0,555099	0,0500552	11,09	8,57e-26 ***
experiencia	0,0187467	0,0120402	1,557	0,1201
antigüedad	0,138829	0,0211572	6,562	1,29e-10 ***
mujer	-1,74241	0,266818	-6,530	1,57e-10 ***
casada	0,556565	0,286742	1,941	0,0528 *
color	-0,0658096	0,426567	-0,1543	0,8775
Media de la vble. dep.	5,896103	D.T. de la vble. dep.	3,693086	
Suma de cuad. residuos	4523,823	D.T. de la regresión	2,952359	
R-cuadrado	0,368218	R-cuadrado corregido	0,360914	
F(6, 519)	50,41424	Valor p (de F)	8,27e-49	
Log-verosimilitud	-1312,288	Criterio de Akaike	2638,576	
Criterio de Schwarz	2668,433	Crit. de Hannan-Quinn	2650,267	

- Sin embargo, el contraste $H_0 : \beta_2 = 0.5$ frente a $H_1 : \beta_2 \neq 0.5$ ya no nos lo ofrece automáticamente Gretl, **tendremos que hacerlo nosotros**.

Intervalo de confianza para los β

- Dada una estimación, su desviación típica (o error estándar) y un nivel de significación, podemos obtener **intervalos de confianza (IC)** para, por ejemplo, β_j
- Concretamente, se construirá un IC **al** $(1 - \alpha)\%$ de la siguiente manera:

$$\hat{\beta}_j \pm t_{N-k}^{\alpha/2} \hat{\sigma}_{\hat{\beta}_j}$$

- Podemos pensar que el intervalo de confianza **recoge los valores verosímiles** para β_j
- Los **IC y contrastes están relacionados**; de hecho, dado un IC al $(1 - \alpha)\%$:
 - no se podrá rechazar mediante un contraste al $\alpha\%$ y a dos colas ningún valor para la hipótesis nula que esté dentro del IC al 95%.
 - Por contra, todo valor para β_j que no esté dentro del intervalo sería rechazado en un contraste al $\alpha\%$ a dos colas.

Significatividad económica vs significatividad estadística

- La significatividad estadística se determina por el valor del t-ratio mientras que la significatividad económica está relacionada con la magnitud (y signo) de las estimaciones
- Poner demasiado énfasis en la significatividad estadística puede llevar a concluir que una variable es “importante” para explicar el regresando, incluso aunque el efecto estimado sea muy modesto.
- Con tamaños de muestra grande los parámetros se suelen estimar de forma precisa: los errores estándar suelen ser pequeños lo que suele resultar en significatividad estadística, incluso aunque esa variable tenga un efecto parcial reducido.
- Aunque una variable sea estadísticamente significativa también hay que analizar el valor estimado del coeficiente para dar una idea de su importancia práctica o económica.