“PRUEBAS DE HIPÓTESIS CON MODELOS ECONOMÉTRICOS”

Dr. Roger Alejandro Banegas Rivero
Universidad Autónoma Gabriel René Moreno

Santa Cruz, Bolivia

Una introducción a estadísitica inferencial

Se requiere hacer inferencia con relación a la población mediante los parámetros de regresión lineal:

$$\hat{y}_t = 20.3 + 0.5091 x_t$$
$$(14.38)\ (0.2561)$$

Donde:

- $\hat{\beta} = 0.5091$ es una estimación puntual de una población desconocida, ¿Qué tan confiable fue la estimación de β?
- La confiabilidad del estimador puntual está medido por su desviación estándar

Prueba de hipótesis: algunos conceptos

- Se puede utilizar información a partir de una muestra para inferir conclusiones sobre una población.

- Siempre se tienen dos hipótesis que van juntas: la hipótesis nula (H_0) y la hipótesis alternativa (H_1).

- La hipótsis nula es la que se somete al soporte estadístico. La hipótesis alternativa representa el resultado de interés.

- Por ejemplo, supóngase que se desea verificar que el verdadero valor de β es 0.5. Se emplea la siguiente notación:

  $$H_0 : \beta = 0.5$$

  $$H_1 : \beta \neq 0.5$$

  Lo anterior se conoce como una prueba de hipótesis de dos colas.

Prueba de hipótesis de una cola

- En algunas ocasiones se tiene información apriori. Supóngase que se se conoce que $\beta > 0.5$ en lugar de $\beta < 0.5$. En este caso, se procede con una prueba de una cola:

  $H_0 : \beta = 0.5$

  $H_1 : \beta > 0.5$

  O también se tendría:

  $H_0 : \beta = 0.5$

  $H_1 : \beta < 0.5$

- Existen dos formas de verificar las hipótesis: 1) vía la significancia del parámetro (β); 2) a travéz de los intervalos de confianza del parámetro.

La probabilidad de distribución de los paramétros por Mínimos Cuadrados Ordinarios (MCO)

- Se asume que $u_t \sim N(0,\sigma^2)$

- Desde que los estimadores son una combinación de variables aleatorias :

$$\hat{\beta} = \sum w_t y_t$$

- Los pesos de las variables aleatorias están normalmente distribuidos:

$$\hat{\alpha} \sim N(\alpha, \mathrm{Var}(\alpha))$$
$$\hat{\beta} \sim N(\beta, \mathrm{Var}(\beta))$$

- ¿Qué sucede si los parámetros no están normalmente distribuidos? ¿Qué suscede si los parámetros se distribuyen normalmente?
- Si la respuesta es positiva, se cumple el supuesto del MCRL y el tamaño de la muestra para realizar la inferencia es adecuado.

La probabilidad de distribución de los paramétros por Mínimos Cuadrados Ordinarios (MCO)

- Se construyen variantes normales estándares a partir de $\hat{\alpha}$ y $\hat{\beta}$

$$\frac{\hat{\alpha}-\alpha}{\sqrt{\operatorname{var}(\alpha)}} \sim N(0,1) \quad \text{and} \quad \frac{\hat{\beta}-\beta}{\sqrt{\operatorname{var}(\beta)}} \sim N(0,1)$$

- Pero se desconocen las var(α) y var(β), así que:

$$\frac{\hat{\alpha}-\alpha}{se(\hat{\alpha})} \sim t_{T-k} \quad \text{and} \quad \frac{\hat{\beta}-\beta}{se(\hat{\beta})} \sim t_{T-K}$$

Prueba de hipótesis: Enfoque de la prueba de significancia

- Se asume que la regresión está dada por:

$$y_t = \alpha + \beta x_t + u_t \qquad \text{for } t=1,2,...,T$$

- Los pasos para realizar las pruebas de hipótesis son:

1. Estimar $\hat{\alpha}$, $\hat{\beta}$ y $SE(\hat{\alpha})$, $SE(\hat{\beta})$ en la manera usual

:

2. Calcular el estadístico "t"

$$test\ statistic = \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})}$$

Donde β^* es el valor de β bajo la hipótesis nula.

Enfoque de la prueba de signficancia (continuación)

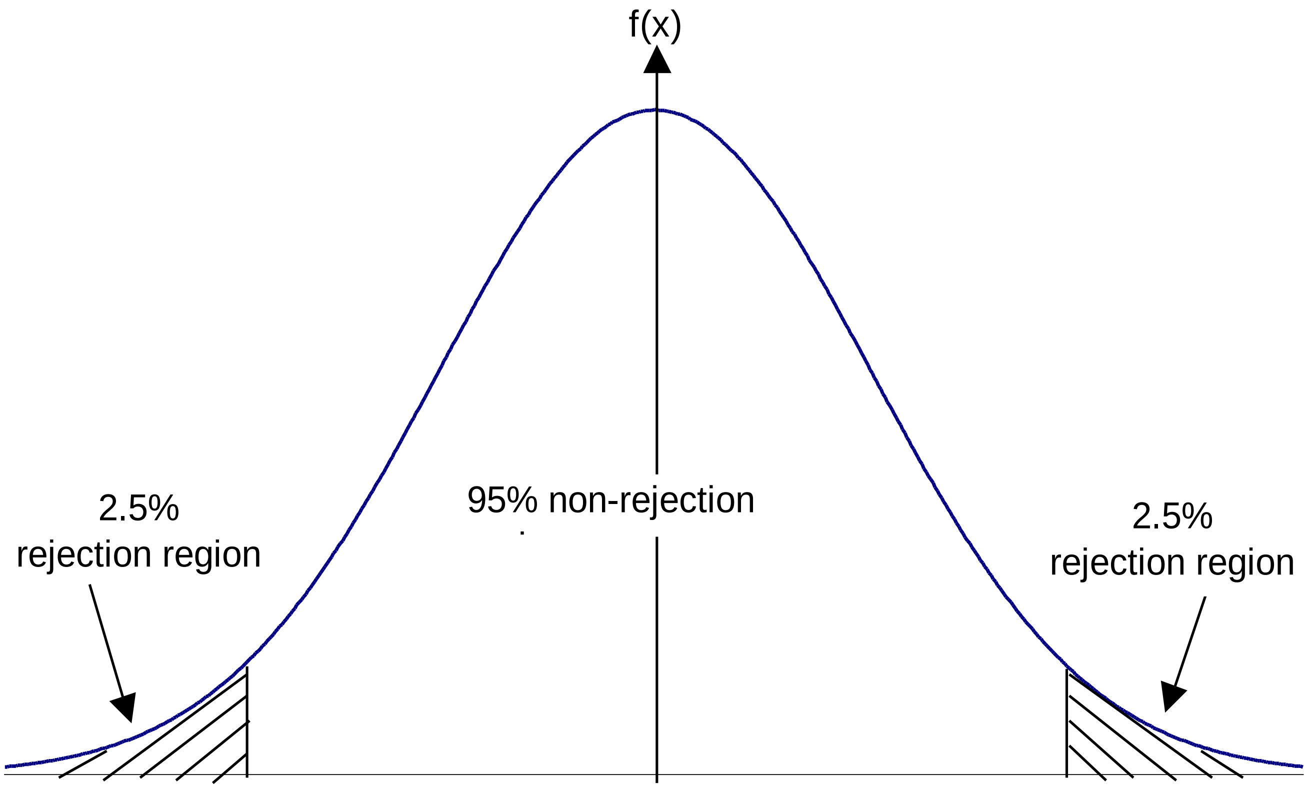
3. Se necesitan valores críticos para rechazar o no rechazar Ho. Los valores críticos pueden seguir una distribución *t-student* T-k grados de libertad.

4. Se elige un "nivel de signifciancia" que se denota por α. Constituye la zona de rechazo o no rechazo de Ho.

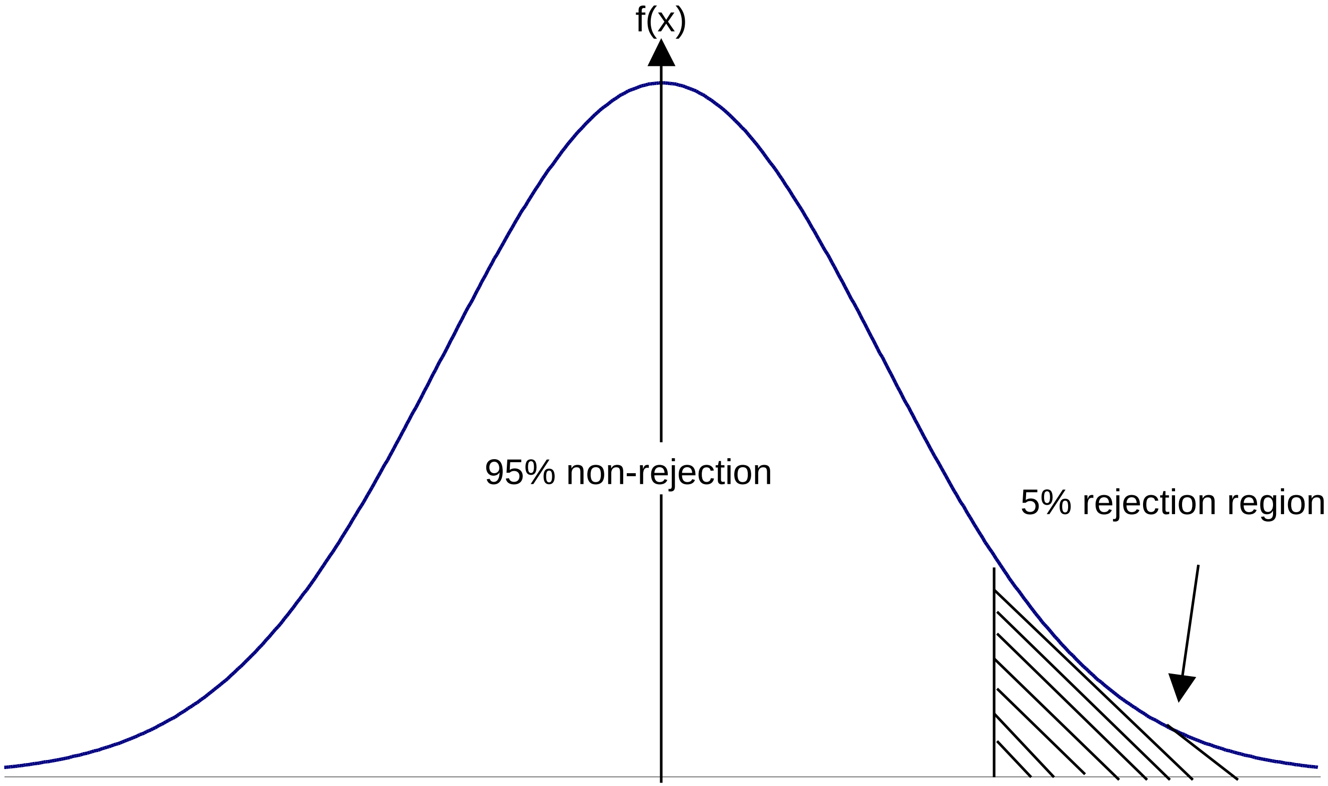
De forma convencional se emplea el 5%; sin embargo, se puede emplear el 10 y el 1% de forma respectiva.

Determinando la región de rechaza a un nivel de confianza

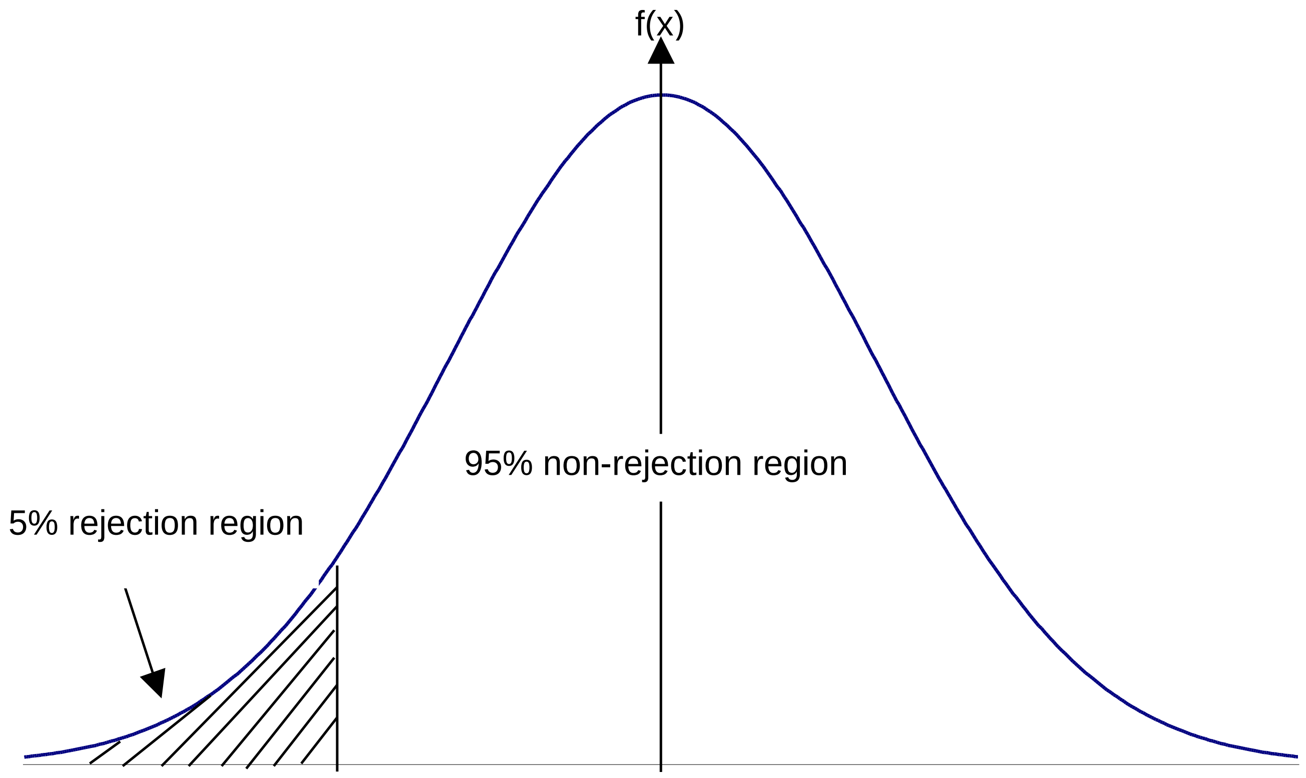
5. Se determina la región de rechazo a un nivel de significancia. Para una prueba de dos colas:



Zona de rechazo para 1-cola en prueba de hipótesis (Cola derecha)



Zona de rechazo para 1- cola en prueba de hipótesis (zona inferior)



Enfoque de significancia: criterios de decisión

6. Al emplear las tablas *t*-student se obtiene el valor crítico que se compara con el *t estimado o calculado*.

7. Si el *t estimado* > *t crítico*, cae región de rechazo, por tanto, se rechaza Ho.

Si el *t estimado* < *t crítico*, no cae en región de rechazo, por tanto, no se rechaza Ho.

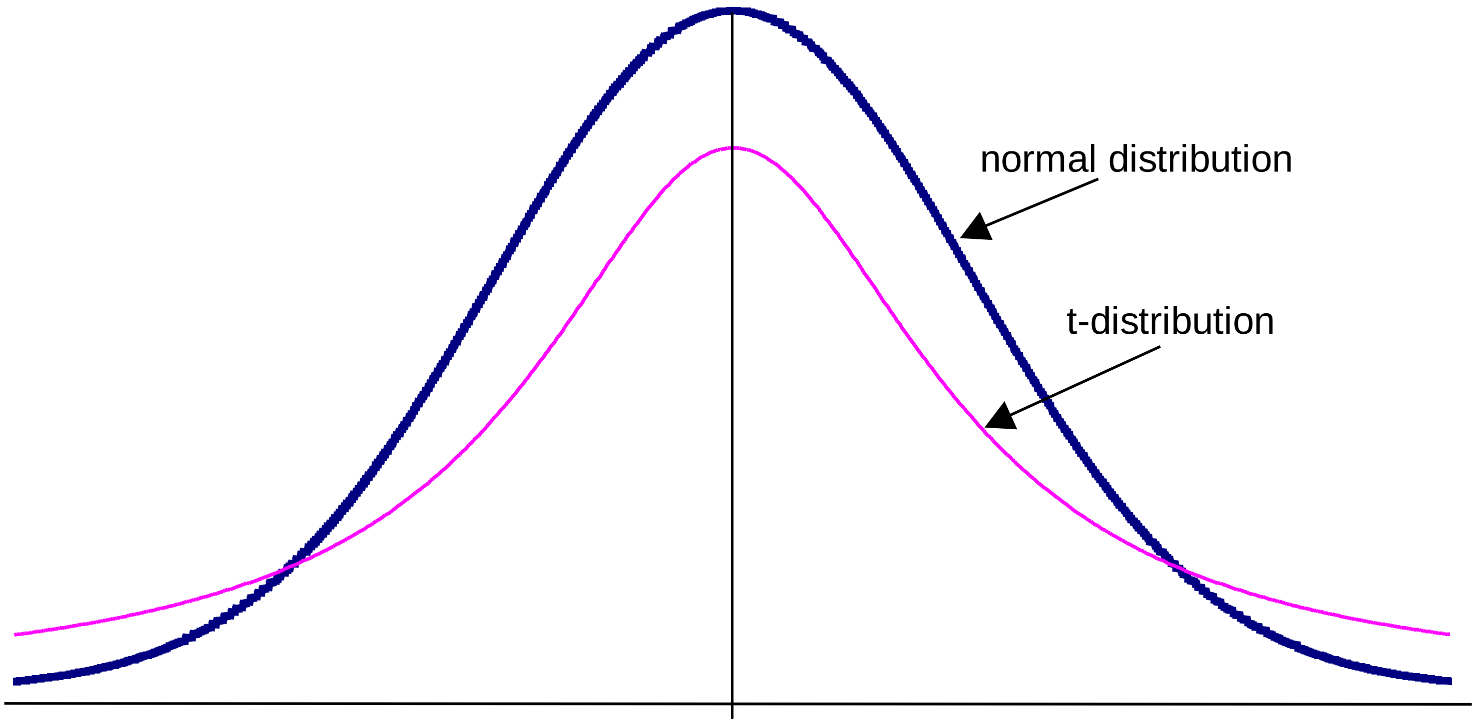
Una nota sobre la distribución *t* y la distribución normal

- Se deberá comprender la distribución normal y su característica de forma “campana”.

- Se puede normalizar una variable $\sim N(0,1)$ al realizar Z – Prom. Z / σz

- Existe una relación específica entre la distribución *t* y la distribución normal. Ambas son simetrícas y están centradas sobre cero. La distribución *t* tiene otro paramétro: los grados de libertad (T- k).

¿Cuál es la forma de la distribución *t*?



Comparación entre la distribución *t* y la distribución normal

- En el límite, una distribución *t* con un número infitinto de grados de liberto es equivalente a una distribución normal $t(\infty) = N(0{,}1)$

- Ejemplo de tablas estadísticas:

Nivel de significancia:	$N(0,1)$	$t(40)$	$t(4)$
50%	0	0	0
5%	1.64	1.68	2.13
2.5%	1.96	2.02	2.78
0.5%	2.57	2.70	4.60

- Las razón para estimar una distribución t en lugar de la distribución normal, se debe que se trabaja con σ^2, la varianza de las perturbaciones

Enfoque de intervalos de confianza para probar hipótesis

- Un ejemplo: se estima un parametro, digamos 0.93, y a un intervalo de confianza del 95%, el rango de este parámetro está entre 0.77 y 1.09.

- Esto significa que a una probabilidad del 95%, el efecto de X sobre Y persiste de forma positiva conteniendo el verdadero valor de β *(pero desconocido)*.

- Los intervalos de confianza son invariables con dos colas, aunque con una cola debe construirse por separado.

¿Cómo llevar a cabo una prueba de hipótesis mediante el uso de intervalos de confianza?

1. Calcule $\hat{\alpha}$, $\hat{\beta}$ y $SE(\hat{\alpha})$, $SE(\hat{\beta})$ de forma previa.

2. Elija un nivel de significancia, α, (ejemplo el 5%) que es equivalente a seleccionar el intervalo: $(1-\alpha)\times 100\%$.

3. Se emplea la tabla t para obtener los valores críticos, se necesitan T-k grados de libertad.

4. El intervalor de confianza está dado por:

$$(\hat{\beta} - t_{crit} \times SE(\hat{\beta}), \hat{\beta} + t_{crit} \times SE(\hat{\beta}))$$

5. Aplicar la prueba: Si el valor de β (β*) permanece dentro del intervalo de confianza, luego, se rechaza la hipótesis que $\beta = \beta$*, en otro caso, no se rechaza la hipótesis.

Intervalos de confianza versus prueba de significancia

- Debe notarse que ambos enfoques proporcionan la misma información del efecto de X sobre Y.

- Bajo el enfoque de significancia, no se rechaza H_0 si $\beta = \beta^*$ if cae dentro de la región de no rechazo de Ho:

$$-t_{crit} \leq \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})} \leq +t_{crit}$$

- Si despejan los términos:

$$-t_{crit} \times SE(\hat{\beta}) \leq \hat{\beta} - \beta^* \leq +t_{crit} \times SE(\hat{\beta})$$

$$\hat{\beta} - t_{crit} \times SE(\hat{\beta}) \leq \beta^* \leq \hat{\beta} + t_{crit} \times SE(\hat{\beta})$$

- Se llega al enfoque de intervalos de confianza.

Un ejemplo: construyendo pruebas de significancia e intervalos de confianza

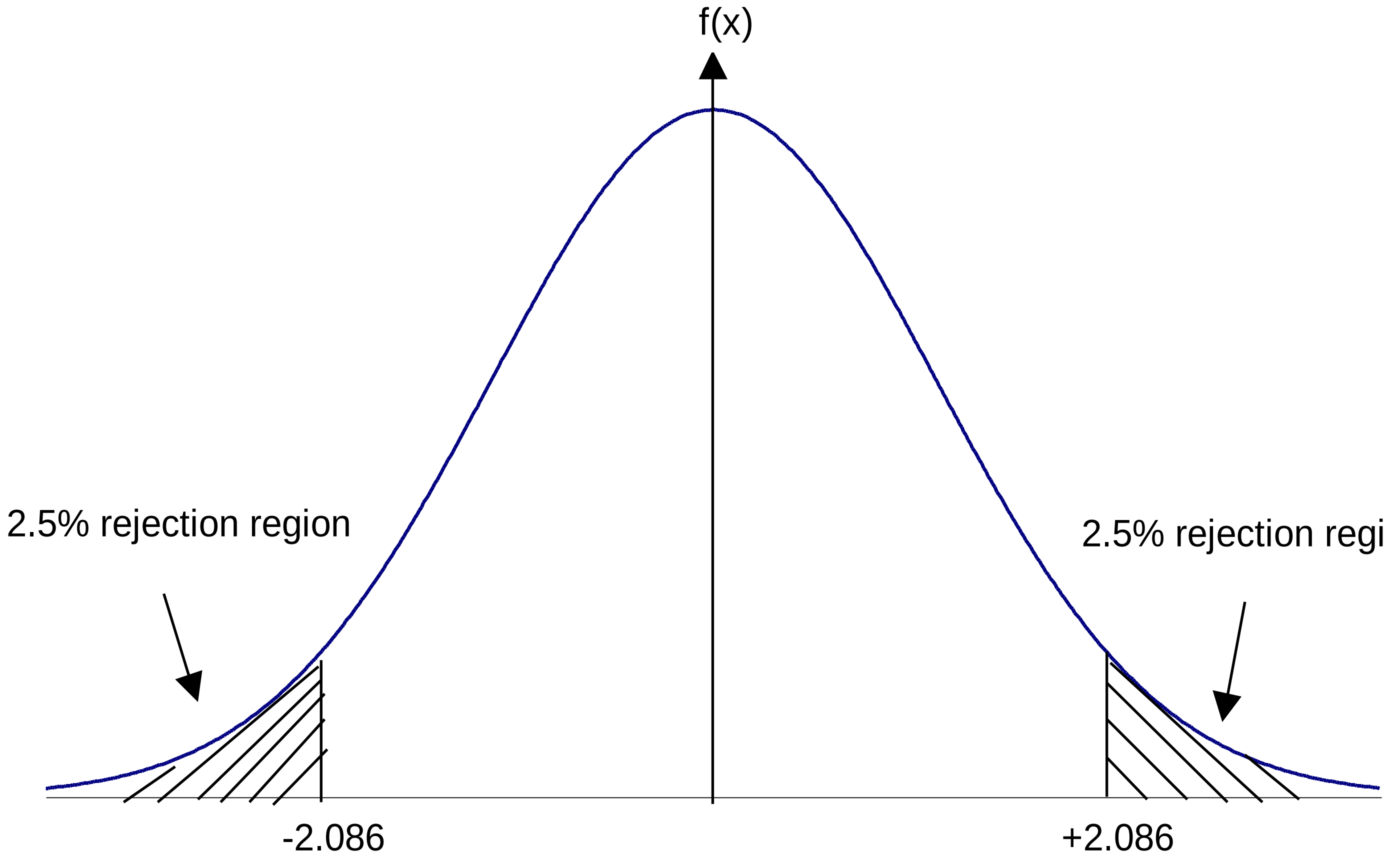
- Al emplear los resultados de la regresión,

$$\hat{y}_t = \underset{(14.38)}{20.3} + \underset{(0.2561)}{0.5091} x_t \quad , \quad T{=}22$$

- Se desea emplear los dos enfoques, para verificar la hipótesis nula que β =1 en contra de H1 para dos colas

- El primer paso consiste en estimar el *t crítico*. Se requiere $t_{crit} = t_{20;5\%}$

Determinando la región de rechazo



Desarrollando la prueba de hipótesis

- La hipótesis son:

$$H_0 : \beta = 1$$
$$H_1 : \beta \neq 1$$

La prueba de significancia:

$$test\ stat = \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})}$$
$$= \frac{0.5091 - 1}{0.2561} = -1.917$$

No se rechaza H_0

el *estadístico t* cae en la región de no rechazo.

Enfoque de intervalos de confianza:

$$\hat{\beta} \pm t_{crit} \times SE(\hat{\beta})$$
$$= 0.5091 \pm 2.086 \times 0.2561$$
$$= (-0.0251, 1.0433)$$

No existe consitencia el signo del parámetro, no se rechaza Ho.

El parámetro se encuentra en el rango, no se rechaza Ho.

Probando otras hipótesis

- ¿Qué sucede si se quiere probar que $H_0 : \beta = 0$ or $H_0 : \beta = 2$?
- Se modifica la notación:

$$\begin{array}{ll} & H_0 : \beta = 0 \\ \text{vs.} & H_1 : \beta \neq 0 \end{array}$$

$$\begin{array}{ll} & H_0 : \beta = 2 \\ \text{vs.} & H_1 : \beta \neq 2 \end{array}$$

Cambiando el nivel de significancia para la prueba

- De forma convencial se utiliza un nivel de significancia del 5%

- En algunos casos, (e.g. $H_0 : \beta = 1$), se puede cambiar los resultados para el intervalo de confianza cuando se cambia el nivel de significancia .

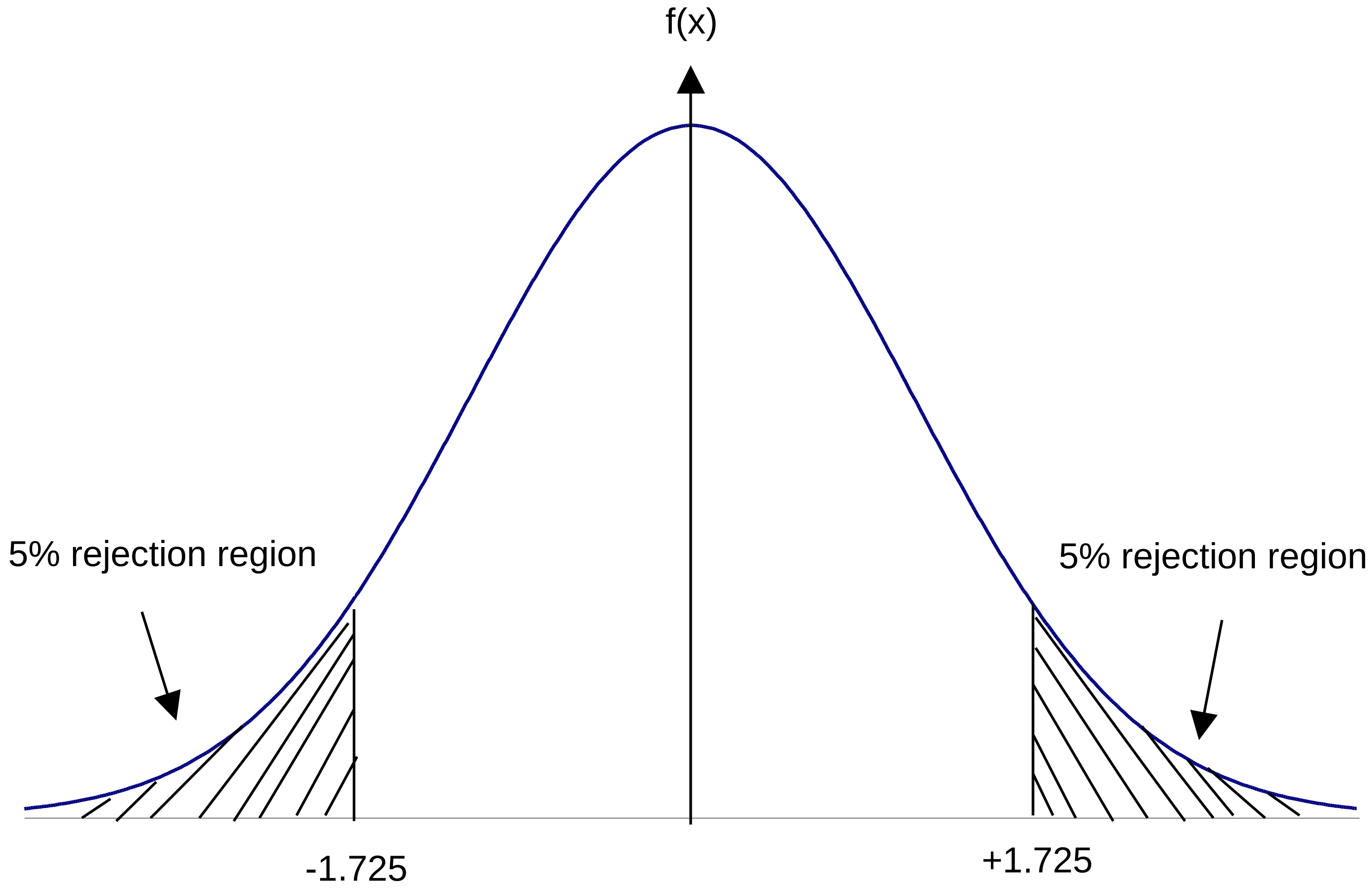
- La prueba de significiancia es más confiable que el enfoque de intervalos de confianza, es decir:

$$test\ stat = \frac{\hat{\beta} - \beta^*}{SE(\hat{\beta})}$$

$$= \frac{0.5091 - 1}{0.2561} = -1.917$$

- Supongáse, que se cambia el nivel de significancia al 10%,

Cambiando el nivel de significancia al 10%: Dos zonas de rechazo



Cambiando el nivel de significancia en la prueba de hipótesis: ¿cómo afectan a las conclusiones?

- $t_{20;10\%}$ = 1.725. Ahora, el *estadístico t* cae en la región de rechazo H_0.
- Hay que tener cuidado cuando las conclusiones varían de acuerdo a un cambiao en el nivel de significancia: ojo!!!!

¿ X influye de forma +/- sobre Y?

Ejemplo:

Los resultados indican que al 5% de significancia, X no influye sobre Y.

Los resultados indican que al 10% de significancia, X influye de forma positiva sobre Y (10% de probabilidad de cometerar ERROR DE TIPO I).

Alguna terminología adicional

- Si se rechaza la hipótesis nula (Ho) al 5%, se dice que los resultados son significativos de forma estadística (o estadísticamente significativos).

- Hagáse notar que la significancia estadística difiere de la significancia práctica.

Los errores que se pueden cometer al realizar pruebas de hipótesis

- De forma usual, se rechaza H_0 si el estadístico *t* es significativo a un nivel de significancia.

- Existen dos grandes errores en las pruebas de hipótesis.

1. Rechazar H_0 cuando en verdad no se puede rechazar: **error tipo I**.
2. No rechazar H_0 cuando de hecho se debe rechazar: **error tipo II**.

		Reality	
		H_0 is true	H_0 is false
Result of Test	Significant (reject H_0)	Type I error $= \alpha$	√
	Insignificant (do not reject H_0)	√	Type II error $= \beta$

La disyuntiva entre errores de Tipo I y Tipo II

- La probabilidad de eror del tipo I es igual a α, es decir el nivel de significancia.
- ¿Qué sucede si se convierte el nivel de significancia más exigente, por ejemplo del 5% al 1%?

  (Reducción del tamaño de la significancia)
- Se reduce la posibilidad de cometaer un error de tipo I.. Pero se incrementa la porbabilidad de cometer un error de tipo II :

Reduce size of test → more strict criterion for rejection → reject null hypothesis less often

↗ less likely to falsely reject

↘ more likely to incorrectly not reject

- Siempre hay una disyuntiva entre error de tipo I y error de tipo II. **¿ Qué tipo de error es más grave?**

La disyuntiva entre errores de Tipo I y Tipo II

- Biblográfía

Brooks, Chr. (2008) Chapter 2: A brief overview of the classical linear regression model in *Introductory Econometrics for Finance*, Cambrigdge University Press.