Econometría

Diplomado Banco Central de Honduras

Instituto de Economía

Pontificia Universidad Católica de Chile

Juan Ignacio Urquiza – Junio 2022

Propiedades estadísticas

- Supuestos:
  - Linealidad en parámetros (RLM.1).
  - Muestreo aleatorio (RLM.2).
  - Colinealidad imperfecta (RLM.3).
  - Media condicional cero (RLM.4).
  - Homocedasticidad (RLM.5).
- Se puede demostrar que:
  - Bajo los supuestos RLM.1 a RLM.4, los estimadores de MCO son insesgados.
  - Bajo los supuestos RLM.1 a RLM.5, son los de mínima varianza entre los estimadores lineales insesgados.

Supuestos

- Linealidad en parámetros (RLM.1):
  - La relación poblacional entre las variables sigue un modelo lineal:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u = \boldsymbol{X}\beta + u$$

  - Si el modelo analítico no es lineal, hay que linealizarlo.
  - Por ejemplo:

$$y = AK^{\beta_1}L^{\beta_2}$$

  - Entonces:

$$\rightarrow \ln y = \beta_0 + \beta_1 \ln K + \beta_2 \ln L$$

Supuestos

- Muestreo aleatorio (RLM.2):
  - Las observaciones provienen de una muestra aleatoria de la población:

$$\{(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i) : i = 1, 2, \ldots, n\}$$

  - Esto significa que las observaciones son independientes:

$$cov\left(u_i, u_j \middle| \boldsymbol{X}\right) = 0 \quad \forall i \neq j$$

  - En un contexto de series de tiempo, esto implica que los errores no tienen autocorrelación.

Supuestos

- Colinealidad imperfecta (RLM.3):
  - No hay relaciones lineales exactas entre las variables independientes.
  - Los regresores pueden estar correlacionados entre sí, sólo que no pueden hacerlo en forma perfecta.
  - Esto permite la identificación de los parámetros.
  - Sin embargo, tal como lo veremos más adelante, una colinealidad alta dificulta la identificación.

Supuestos

- Media condicional cero – $E(\boldsymbol{u}|\boldsymbol{X}) = \mathbf{0}$ (RLM.4):
  - También se puede escribir como $E(u|x_1, \dots, x_k) = 0$.
  - Significa que ninguno de los factores en el término del error correlaciona con las variables explicativas.
  - Implica que:

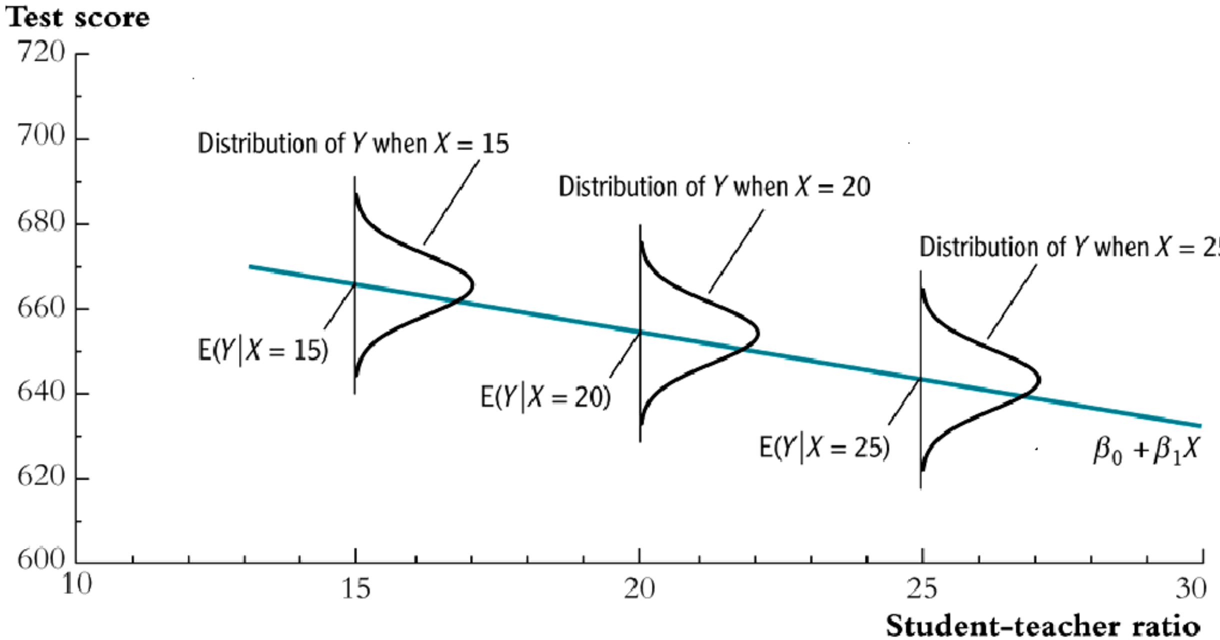
$$E(\boldsymbol{y}|\boldsymbol{X}) = E(\boldsymbol{X\beta} + \boldsymbol{u}|\boldsymbol{X}) = \boldsymbol{X\beta}$$

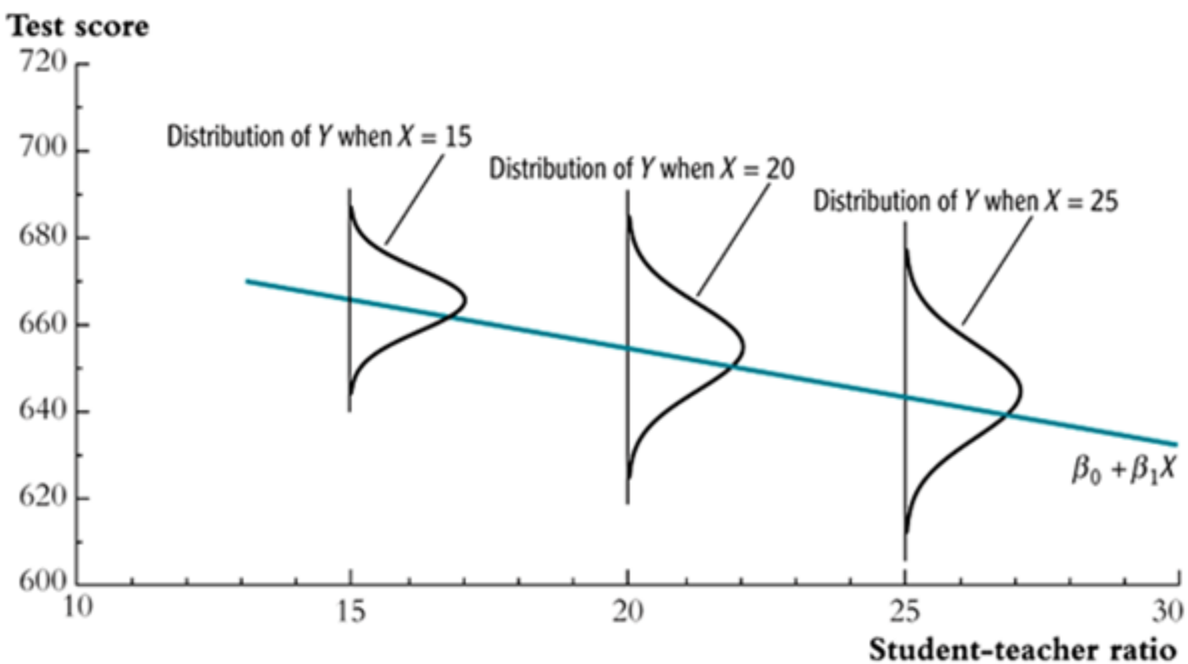
$$E[E(\boldsymbol{u}|\boldsymbol{X})] = E(\boldsymbol{u}) = \mathbf{0}$$

  - Este supuesto puede no cumplirse por varias razones:
    - Especificación incorrecta de la forma funcional, omisión de variables relevantes, errores de medición, etc.

Supuestos

- Homocedasticidad – $V(\boldsymbol{u}|\boldsymbol{X}) = \sigma^2\boldsymbol{I}$ (RLM.5):
  - También se puede escribir como $V(u|x_1, \ldots, x_k) = \sigma^2$.
  - Es decir, la varianza del error no depende de haber observado una realización particular de X.
  - Cuando este supuesto no se cumple, se dice que el modelo presenta heterocedasticidad.
  - ¿Gráficamente?


Test score
720
700
680
660
640
620
600
Distribution of Y when $X = 15$
Distribution of Y when $X = 20$
Distribution of Y when $X = 25$
$E(Y|X = 15)$
$E(Y|X = 20)$
$E(Y|X = 25)$
$\beta_0 + \beta_1 X$
10
15
20
25
30
Student–teacher ratio


Test score
720
700
680
660
640
620
600
Distribution of Y when $X = 15$
Distribution of Y when $X = 20$
Distribution of Y when $X = 25$
$\beta_0 + \beta_1 X$
10
15
20
25
30
Student–teacher ratio

Propiedades estadísticas

- Insesgamiento:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$
$$= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u})$$
$$= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}$$

  - Por lo tanto:

$$E\left(\widehat{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right) = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E(\boldsymbol{u}|\boldsymbol{X}) = \boldsymbol{\beta}$$
$$\rightarrow E\left(\widehat{\boldsymbol{\beta}}\right) = E\left[E\left(\widehat{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right)\right] = \boldsymbol{\beta}$$

  - Esto implica que el promedio de los estimadores $\widehat{\boldsymbol{\beta}}$ obtenidos a partir de todas las muestras aleatorias posibles es igual a $\boldsymbol{\beta}$.

Propiedades estadísticas

- <u>Varianza</u>:

$$\boxed{\boldsymbol{V}\left(\widehat{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}}$$

$$\begin{aligned}
\boldsymbol{V}\left(\widehat{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right) &= \boldsymbol{E}\left[\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\middle|\boldsymbol{X}\right] \\
&= \boldsymbol{E}\left[(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{u}\boldsymbol{u}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\middle|\boldsymbol{X}\right] \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{E}[\boldsymbol{u}\boldsymbol{u}'|\boldsymbol{X}]\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}} \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'(\sigma^2\boldsymbol{I})\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}} \\
&= \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}} \\
&= \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}
\end{aligned}$$

Teorema de Gauss-Markov

- Cualquier estimador lineal de $\boldsymbol{\beta}$ se puede escribir como:

$$\widetilde{\boldsymbol{\beta}} = \boldsymbol{A}'\boldsymbol{y} = \boldsymbol{A}'(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}) = \boldsymbol{A}'\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{A}'\boldsymbol{u}$$

- Para que $\widetilde{\boldsymbol{\beta}}$ sea insesgado:

$$\boldsymbol{E}\left(\widetilde{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right) = \boldsymbol{A}'\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{A}'\boldsymbol{E}(\boldsymbol{u}|\boldsymbol{X}) = \boldsymbol{A}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta} \leftrightarrow \boldsymbol{A}'\boldsymbol{X} = \boldsymbol{I}$$

- ¿Varianza?

$$\boldsymbol{V}\left(\widetilde{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right) = \boldsymbol{A}'\boldsymbol{V}(\boldsymbol{u}|\boldsymbol{X})\boldsymbol{A} = \sigma^2\boldsymbol{A}'\boldsymbol{A}$$

- Por lo tanto:

$$\boldsymbol{V}\left(\widetilde{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right) - \boldsymbol{V}\left(\widehat{\boldsymbol{\beta}}\middle|\boldsymbol{X}\right) = \sigma^2\left[\boldsymbol{A}'\boldsymbol{A} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\right]$$

$$= \sigma^2\left[\boldsymbol{A}'\boldsymbol{A} - \boldsymbol{A}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{A}\right]$$

Teorema de Gauss-Markov

- Por lo tanto:

$$\boldsymbol{V}(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}) - \boldsymbol{V}(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}) = \sigma^2 \boldsymbol{A}'[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\boldsymbol{A}$$
$$= \sigma^2 \boldsymbol{A}'\boldsymbol{M}\boldsymbol{A}$$

- Dado que $\boldsymbol{M}$ es simétrica e idempotente, entonces $\boldsymbol{A}'\boldsymbol{M}\boldsymbol{A}$ es semi-definida positiva.

- Esto implica que para cualquier vector $\boldsymbol{c}$:

$$\boldsymbol{V}(\boldsymbol{c}'\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}) - \boldsymbol{V}(\boldsymbol{c}'\widehat{\boldsymbol{\beta}}|\boldsymbol{X}) = \boldsymbol{c}'[\boldsymbol{V}(\widetilde{\boldsymbol{\beta}}|\boldsymbol{X}) - \boldsymbol{V}(\widehat{\boldsymbol{\beta}}|\boldsymbol{X})]\boldsymbol{c} \geq 0$$

- Por lo tanto, esto establece que el estimador de MCO es el mejor estimador lineal insesgado (MELI).

Varianza del estimador de MCO

- Bajo los supuestos RLM.1 a RLM.5, se cumple que:

$$V\left(\hat{\beta}_j \middle| \boldsymbol{X}\right) = \frac{\sigma^2}{SCT_j \times \left(1 - R_j^2\right)} \qquad \forall j = 1, 2, \ldots, k.$$

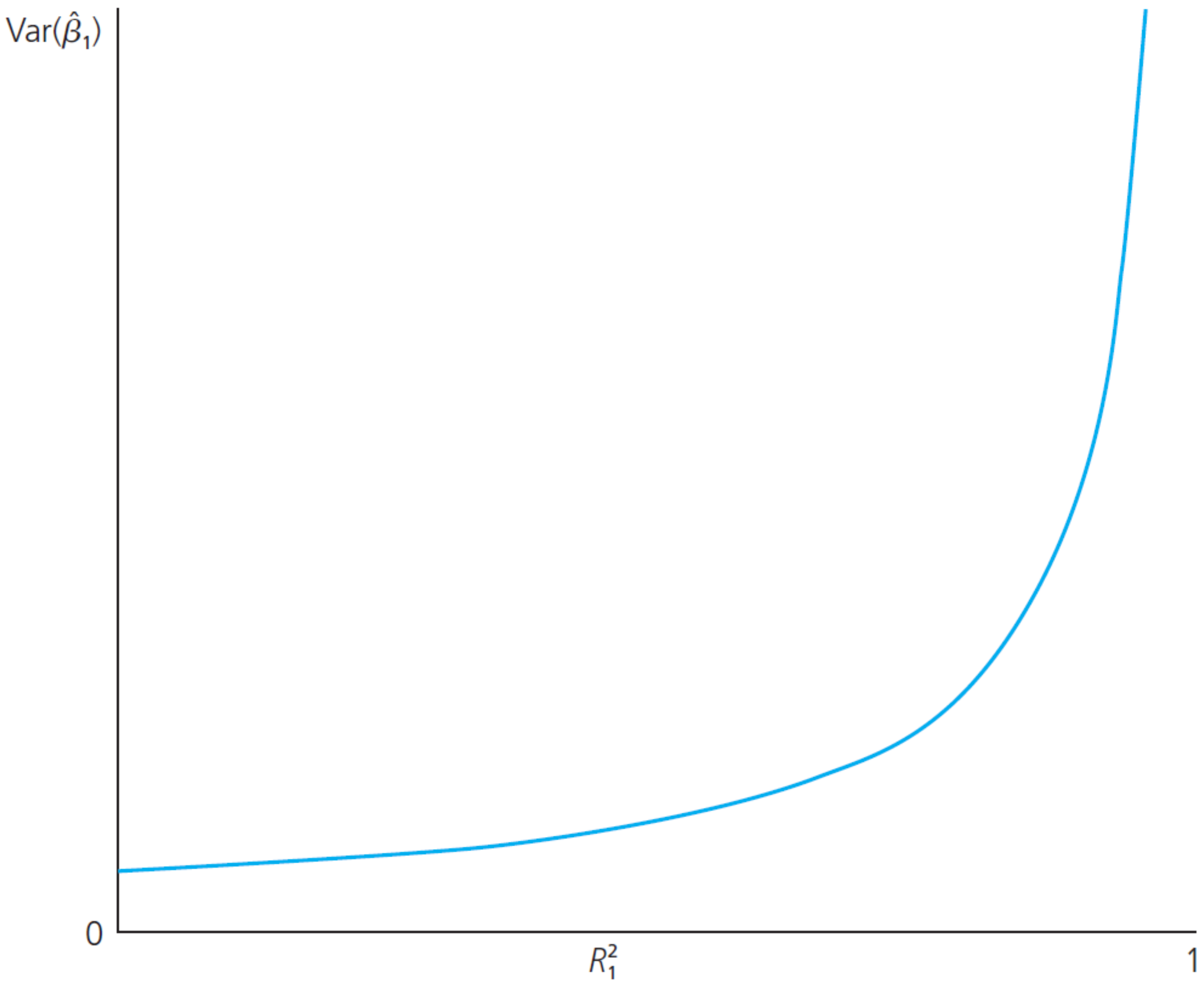
donde:

  - SCT_j es la suma de cuadrados totales de x_j.
  - R_j^2 es el R^2 resultante de la regresión de x_j (como variable dependiente) sobre todas las demás variables independientes. En otras palabras, el R_j^2 indica cuánto de la variable x_j es explicado por los otros regresores.

Varianza del estimador de MCO

- $V\left(\widehat{\beta}_j \middle| \boldsymbol{X}\right)$ es creciente en σ^2:
  - Mientras más ruido haya, más difícil será estimar el efecto parcial de x_j sobre y.
- $V\left(\widehat{\beta}_j \middle| \boldsymbol{X}\right)$ es decreciente en SCT_j:
  - Es decir, se prefiere una mayor variación muestral en x_j. Por ejemplo, aumentando el tamaño de la muestra.
- $V\left(\widehat{\beta}_j \middle| \boldsymbol{X}\right)$ es creciente en R_j^2:
  - El supuesto RLM.3 asegura que el $R_j^2 < 1$ pero nótese que cuando $R_j^2 \to 1$, entonces $V\left(\widehat{\beta}_j \middle| \boldsymbol{X}\right) \to \infty$. Esto se conoce como multicolinealidad.

Gráficamente



Multicolinealidad

- Se define como una correlación elevada pero no perfecta entre 2 o más variables explicativas.
- En muchos casos, la correlación elevada entre algunas variables resulta irrelevante. Por ejemplo, cuando x_2 y x_3 están muy correlacionadas entre sí pero éstas no se correlacionan con la variable de interés, x_1.
- Un intento por reducir la varianza del estimador podría ser eliminar algún regresor y con ello el disminuir el R_j^2. Sin embargo, como se verá más adelante, la omisión de variables relevantes produce estimadores sesgados.
- En definitiva, no implica violación alguna de los supuestos anteriores y, por lo tanto, MCO sigue siendo MELI.

Supuestos

- Normalidad – $u \sim N(0, \sigma^2)$ (RLM.6):
  - El error poblacional es independiente de las variables explicativas y está distribuido normalmente, con media cero y varianza σ^2.
  - Esto implica que:

$$y|\boldsymbol{X} \sim N(\boldsymbol{X}\beta, \sigma^2)$$

  - Este supuesto es clave para hacer inferencia acerca de los valores de parámetros.

Distribución del estimador de MCO

- La normalidad de los errores se traduce en distribuciones muestrales normales de los estimadores de MCO:

$$\widehat{\beta}_j|\boldsymbol{X} \sim N\left(\beta_j, V\left(\widehat{\beta}_j|\boldsymbol{X}\right)\right)$$

- Esto implica que, condicional en los regresores:

$$\frac{\widehat{\beta}_j - \beta_j}{\sqrt{V\left(\widehat{\beta}_j|\boldsymbol{X}\right)}} \sim N(0,1)$$

- Sin embargo, $V\left(\widehat{\beta}_j|\boldsymbol{X}\right)$ depende de σ^2 que en la práctica es desconocido. Por lo tanto, debemos estimarlo. ¿Cómo?

Estimador MCO de σ^2

- Dado que los errores son inobservables, se utilizan los residuos de MCO:

$$\hat{\sigma}^2 = \left(\frac{1}{n-k-1}\right) \times \sum_{i=1}^{n} \hat{u}_i^{\,2}$$

- Bajo los supuestos RLM.1 a RLM.5, $E(\hat{\sigma}^2 | \boldsymbol{X}) = \sigma^2$.
- Por lo tanto, bajo los supuestos RLM.1 a RLM.6, se cumple que, condicional en los regresores:

$$\frac{\hat{\beta}_j - \beta_j}{s.e.(\hat{\beta}_j)} \sim t_{n-k-1}$$

donde: $s.e.(\hat{\beta}_j) = \sqrt{\hat{V}(\hat{\beta}_j | \boldsymbol{X})}$.

```
. reg testscr str el_pct expn_stu
```

Source	SS	df	MS
Model	66409.8837	3	22136.6279
Residual	85699.7099	416	206.008918
Total	152109.594	419	363.030056

Number of obs	=	420
F(3, 416)	=	107.45
Prob > F	=	0.0000
R-squared	=	0.4366
Adj R-squared	=	0.4325
Root MSE	=	14.353

testscr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
str	-.2863992	.4805232	-0.60	0.551	-1.230955	.658157
el_pct	-.6560227	.0391059	-16.78	0.000	-.7328924	-.5791529
expn_stu	.0038679	.0014121	2.74	0.006	.0010921	.0066437
_cons	649.5779	15.20572	42.72	0.000	619.6883	679.4676

Errores estándar

```
. estat vce
```

Covariance matrix of coefficients of **regress** model

e(V)	str	el_pct	expn_stu	_cons
str	.23090254			
el_pct	-.00344264	.00152927		
expn_stu	.00042012	-3.220e-06	1.994e-06	
_cons	-6.7125805	.06060743	-.01879395	231.21388

Test sobre un parámetro individual

- Suponga que desea evaluar si el coeficiente que acompaña a X_j es estadísticamente distinto de a_j. Es decir:

$$H_0: \beta_j = a_j$$

$$H_1: \beta_j \neq a_j$$

- En el caso particular en que $a_j = 0$, se busca determinar si X_j es una variable "significativa o relevante". Si se rechaza H_0, decimos entonces que X_j es significativa.
- Para esto usamos el estadístico t construido bajo H_0:

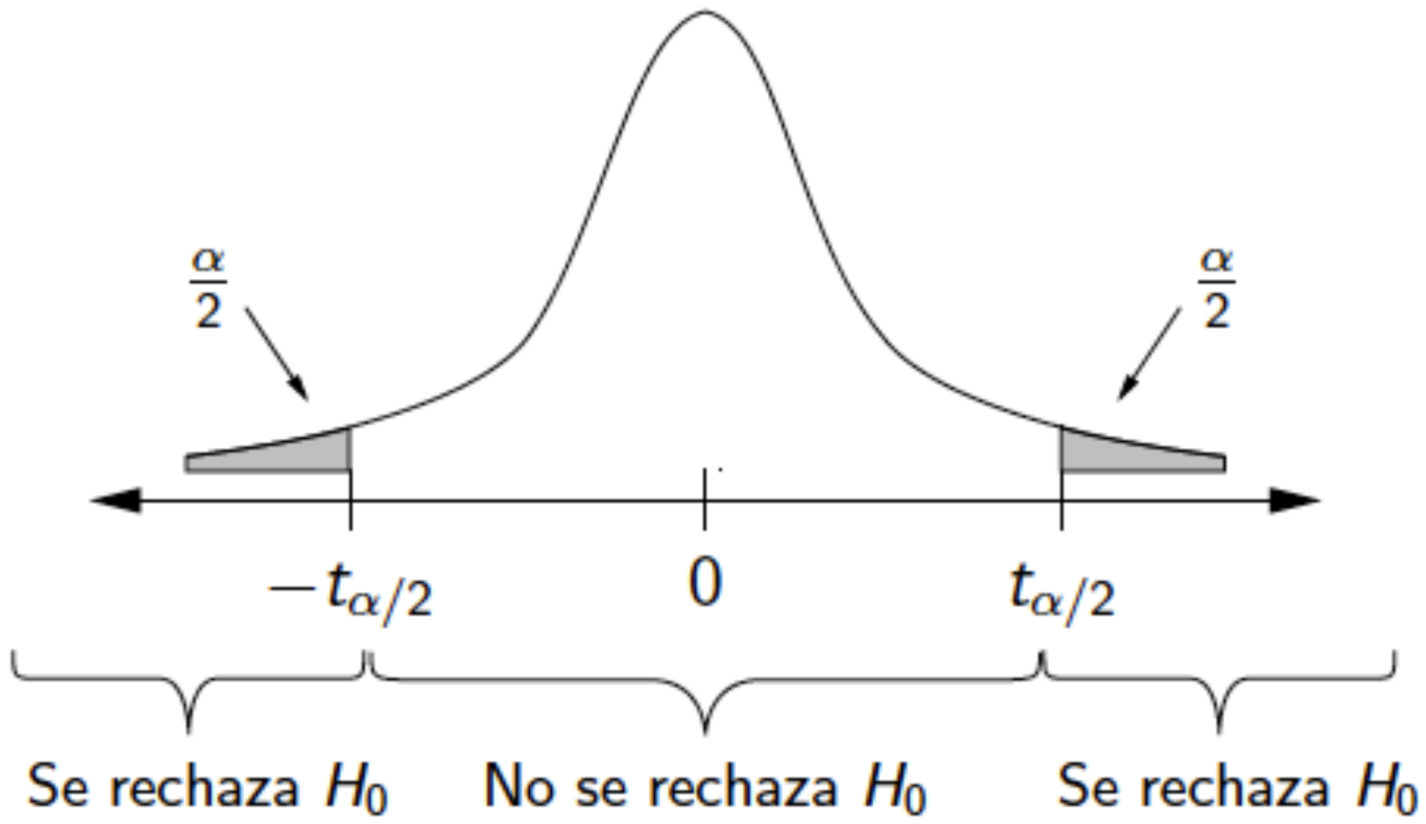
$$t_{\widehat{\beta}_j} = \frac{\widehat{\beta}_j - a_j}{s.e.\left(\widehat{\beta}_j\right)} \sim t_{n-k-1}$$

Test sobre un parámetro individual

- $t_{\widehat{\beta}_j}$ mide en cuántas desviaciones estándar estimadas de $\widehat{\beta}_j$ se aleja $\widehat{\beta}_j$ de a_j.
- Valores de $t_{\widehat{\beta}_j}$ alejados de 0 son evidencia contraria a H_0.
- Suponga que se fija en α la probabilidad de cometer error del tipo I.
- Sea $t_{\alpha/2}$ el percentil $(1-\alpha/2)$ de la distribución t-Student con $n-k-1$ grados libertad.
- Entonces, se rechaza H_0 cuando:

$$\left|\frac{\widehat{\beta}_j - a_j}{s.e.\left(\widehat{\beta}_j\right)}\right| \geq t_{\alpha/2}$$

Gráficamente



Intervalo de confianza

- Además, sabemos que:

$$\left|\frac{\widehat{\beta}_j - \beta_j}{s.e.\left(\widehat{\beta}_j\right)}\right| \geq t_{\alpha/2}$$

es equivalente a decir que:

$$Pr\left(-t_{\alpha/2} \leq \frac{\widehat{\beta}_j - \beta_j}{s.e.\left(\widehat{\beta}_j\right)} \leq t_{\alpha/2}\right) = (1 - \alpha)$$

- Esto implica que el intervalo de confianza para β_j es:

$$\widehat{\beta}_j - t_{\alpha/2} \times s.e.\left(\widehat{\beta}_j\right) \leq \beta_j \leq \widehat{\beta}_j + t_{\alpha/2} \times s.e.\left(\widehat{\beta}_j\right)$$

- Si β_j no está en el intervalo, entonces se rechaza H_0.

Valor-*p*

- Se define al valor-*p* de una prueba como:

$$Pr\left(|T| > \left|t_{\widehat{\beta_j}}\right|\right)$$

donde T es una variable t-Student con $n-k-1$ grados libertad y $t_{\widehat{\beta_j}}$ es el valor del estadístico de contraste obtenido.

- En otras palabras, el valor-*p* es el mínimo nivel de significancia a partir del cual se rechaza H_0.

- Por ejemplo, si se evalúa $H_0\colon \beta_j = 0$ y se obtiene un valor-*p* igual a 6.4%, se concluye que β_j es significativo a un nivel de significancia del 10% pero no al 5%.

- Entonces, valores-*p* pequeños son evidencia contraria a H_0.

Test sobre un parámetro individual

- Considere ahora una hipótesis alternativa del tipo:

$$H_1\!: \beta_j > a_j \quad o \quad H_1\!: \beta_j < a_j$$

- Sin pérdida de generalidad, suponga un test cuya hipótesis alternativa es:

$$H_1\!: \beta_j < a_j$$

- Esto significa que valores positivos del estadístico $t_{\widehat{\beta_j}}$ no representan evidencia contraria a H_0.
- Entonces, se rechaza H_0 en favor de H_1 cuando $t_{\widehat{\beta_j}} < -t_\alpha$.
- En estos test de 1 cola, el valor-*p* correspondiente es la mitad del valor-*p* del test de 2 colas.

Combinación lineal de parámetros

- Suponga ahora que queremos evaluar si $\beta_1 = \beta_2$. Esto se puede reescribir como:

$$H_0\colon \beta_1 - \beta_2 = 0$$

- Dado los estimadores de MCO siguen una distribución normal, su diferencia sigue también una distribución normal.
- Entonces, podemos utilizar el siguiente estadístico t:

$$\frac{\widehat{\beta_1} - \widehat{\beta_2}}{s.e.\left(\widehat{\beta_1} - \widehat{\beta_2}\right)} = \frac{\widehat{\beta_1} - \widehat{\beta_2}}{\sqrt{\hat{V}\left(\widehat{\beta_1}\right) + \hat{V}\left(\widehat{\beta_2}\right) - 2\times\hat{C}\left(\widehat{\beta_1}, \widehat{\beta_2}\right)}} \sim t_{n-k-1}$$

- En definitiva, podemos aplicar este mismo enfoque para evaluar cualquier combinación lineal de parámetros.

Pruebas de exclusión

- Considere ahora el siguiente MRL múltiple:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

- Suponga que se desea probar que cuando controlamos por X_1, las variables X_2 y X_3 no tienen efecto parcial sobre Y.
- Es decir:

$$H_0\colon\ \beta_2 = 0, \beta_3 = 0$$

- Hablamos entonces de 2 restricciones de exclusión bajo H_0.
- Si rechazamos H_0, decimos que X_2 y X_3 son conjuntamente significativas. El test no indica qué variable tiene un efecto significativo; puede ser una variable, la otra o ambas.

Pruebas de exclusión

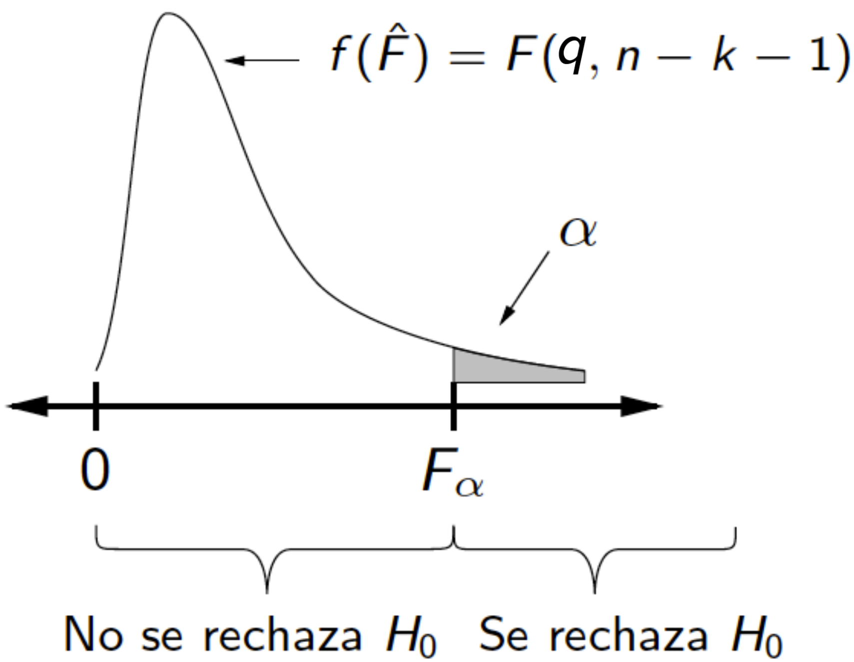
- Este test puede pensarse como una comparación entre un modelo libre y un modelo sobre el cual se imponen las restricciones.
- En particular, se puede demostrar que:

$$F = \frac{(SCR_r - SCR_{sr})}{SCR_{sr}} \times \frac{(n-k-1)}{q} \sim F(q, n-k-1)$$

  donde SCR_r es la suma de cuadrados residuales del modelo restringido, SCR_{sr} la del modelo libre o no restringido, $(n-k-1)$ son los grados de libertad del modelo libre, y q es el número de restricciones bajo H_0.
- ¿Intuición? Evaluar si aumento en la SCR al pasar del modelo libre al modelo restringido es lo suficientemente grande.

Gráficamente

- Valores de $\hat{F}$ alejados de 0 son evidencia contraria a H_0.
- Suponga que se fija en α la probabilidad de cometer error del tipo I.
- Entonces, se rechaza H_0 cuando $\hat{F} > F_\alpha(q, n-k-1)$.



```
. reg testscr str

      Source |       SS       df       MS              Number of obs =     420
-------------+------------------------------           F(  1,   418) =   22.58
       Model |  7794.11004     1  7794.11004           Prob > F      =  0.0000
    Residual |  144315.484   418  345.252353           R-squared     =  0.0512
-------------+------------------------------           Adj R-squared =  0.0490
       Total |  152109.594   419  363.030056           Root MSE      =  18.581

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -2.279808   .4798256    -4.75   0.000     -3.22298   -1.336637
       _cons |    698.933   9.467491    73.82   0.000     680.3231    717.5428
------------------------------------------------------------------------------

. reg testscr str el_pct expn_stu

      Source |       SS       df       MS              Number of obs =     420
-------------+------------------------------           F(  3,   416) =  107.45
       Model |  66409.8837     3  22136.6279           Prob > F      =  0.0000
    Residual |  85699.7099   416  206.008918           R-squared     =  0.4366
-------------+------------------------------           Adj R-squared =  0.4325
       Total |  152109.594   419  363.030056           Root MSE      =  14.353

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4805232    -0.60   0.551    -1.230955     .658157
      el_pct |  -.6560227   .0391059   -16.78   0.000    -.7328924   -.5791529
    expn_stu |   .0038679   .0014121     2.74   0.006     .0010921    .0066437
       _cons |   649.5779   15.20572    42.72   0.000     619.6883    679.4676
------------------------------------------------------------------------------
```

```
. reg testscr str el_pct expn_stu

      Source |       SS       df       MS              Number of obs =     420
-------------+------------------------------           F(  3,   416) =  107.45
       Model |  66409.8837     3  22136.6279           Prob > F      =  0.0000
    Residual |  85699.7099   416  206.008918           R-squared     =  0.4366
-------------+------------------------------           Adj R-squared =  0.4325
       Total |  152109.594   419  363.030056           Root MSE      =  14.353

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4805232    -0.60   0.551    -1.230955     .658157
      el_pct |  -.6560227   .0391059   -16.78   0.000    -.7328924   -.5791529
    expn_stu |   .0038679   .0014121     2.74   0.006     .0010921    .0066437
       _cons |   649.5779   15.20572    42.72   0.000     619.6883    679.4676
------------------------------------------------------------------------------

. test el_pct expn_stu

 ( 1)  el_pct = 0
 ( 2)  expn_stu = 0

       F(  2,   416) =  142.27
            Prob > F =    0.0000
```

$$\rightarrow F = \frac{(144315.48 - 85699.71)}{85699.71} \times \frac{(420 - 3 - 1)}{2} = 142.27$$

Contraste de regresión

- También conocido como test de significancia general o global.
- La hipótesis nula establece que ninguna variable explicativa tiene efecto sobre Y. Es decir:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

- Entonces, el modelo restringido es:

$$Y = \beta_0 + u,$$

  cuyo R^2 es igual a 0.
- Por lo tanto, el estadístico F se puede escribir a partir del R^2 del modelo no restringido:

$$F = \frac{R^2}{(1-R^2)} \times \frac{(n-k-1)}{k}$$

```
. reg testscr str el_pct expn_stu

      Source |       SS       df       MS              Number of obs =     420
-------------+------------------------------           F(  3,   416) =  107.45
       Model |  66409.8837     3  22136.6279           Prob > F      =  0.0000
    Residual |  85699.7099   416  206.008918           R-squared     =  0.4366
-------------+------------------------------           Adj R-squared =  0.4325
       Total |  152109.594   419  363.030056           Root MSE      =  14.353

------------------------------------------------------------------------------
     testscr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         str |  -.2863992   .4805232    -0.60   0.551    -1.230955     .658157
      el_pct |  -.6560227   .0391059   -16.78   0.000    -.7328924   -.5791529
    expn_stu |   .0038679   .0014121     2.74   0.006     .0010921    .0066437
       _cons |   649.5779   15.20572    42.72   0.000     619.6883    679.4676
------------------------------------------------------------------------------
```

```
. test str el_pct expn_stu

 ( 1)  str = 0
 ( 2)  el_pct = 0
 ( 3)  expn_stu = 0

       F(  3,   416) =  107.45
            Prob > F =    0.0000
```

$$\rightarrow F = \frac{0.4366}{(1-0.4366)} \times \frac{(420-3-1)}{3} = 107.45$$