Dpto. de Economía Cuantitativa
Universidad Complutense de Madrid

Econometría I

Tema 2 Contrastes de Hipótesis y Previsión

Marcos Bujosa

Material de apoyo para el curso *Econometría I*

Actualizado el: 26 de septiembre de 2007 Versión 2.01

Puede encontrar la última versión de este material en:

http://www.ucm.es/info/ecocuan/mbb/index.html#ectr1

Índice

Este es un material de apoyo a las clases. En ningún caso sustituye a los libros de texto que figuran en el programa de la asignatura; textos que el alumno debe estudiar para afrontar el examen final con ciertas garantías de éxito.

Referencias recomendadas para la asignatura: Novales (1993), Wooldridge (2006), Verbeek (2004)

Otra referencia seguida en la elaboración de este material es el capítulo 1 de Hayashi (2000), que se puede descargar desde:
http://www.pupress.princeton.edu/chapters/s6946.pdf

Contrastes de Hipótesis y Previsión

1. Introducción

⇑ **Contrastes de hipótesis paramétricas** **1**

Hipótesis afirmación sobre uno o varios parámetros poblacionales

- H_0: hipótesis *nula*
- H_1: hipótesis complementaria (*alternativa*)

Contraste de hipótesis es una regla que establece

- para que valores muestrales $\mathbf{x}$ se rechaza H_0 (*región crítica, RC*)
- para que valores muestrales $\mathbf{x}$ no se rechaza H_0 (*región de aceptación, RA*)

Toma de decisión rechazo o no rechazo de H_0

⇑ **Contrastes de hipótesis paramétricas** **2**

Caracterizamos RC mediante un estadístico $g(X)$.
Ejemplo

- Tren sale cada hora (tardo 10' en llegar al andén)
- H_0: me da tiempo
  H_1: NO me da tiempo
- $g(X)$: media de los relojes de los presentes
- $RC = \{\mathbf{x} \text{ tales que: } g(\mathbf{x}) \geq 45'\}$ (*nivel de significación*)
- Pregunto la hora, y decido si voy al andén

Pero el estadístico podría ser

- $g^*(X)$: media de los relojes de más de 130 euros.
- $RC^* = \{\mathbf{x} \text{ tales que: } g^*(\mathbf{x}) \geq 49'\}$ (*nivel de significación*)

⇑ **Etapas de un contraste** **3**

1. Planteamiento de hipótesis nula

$$H_0 : \quad X \sim f(x; \boldsymbol{\theta}) ; \qquad \boldsymbol{\theta} \in \Theta_0$$

   y alternativa

$$H_1 : \quad X \sim f(x; \boldsymbol{\theta}) ; \qquad \boldsymbol{\theta} \in \Theta_1$$

   donde $\Theta = \Theta_0 \cup \Theta_1$, y $\Theta_0 \cap \Theta_1 = \emptyset$

2. Elección del estadístico $g(X)$
3. División del espacio muestral en dos regiones: RC y RA
   (dado un nivel de significación α)
   $RC \cap RA = \emptyset$; $\quad RC \cup RA =$ espacio muestral
   - ¿Donde está mi muestra?
   - Cálculo del estadístico sobre la muestra: $g(\mathbf{x})$
   - Rechazo o no rechazo H_0 (y toma de decisión)

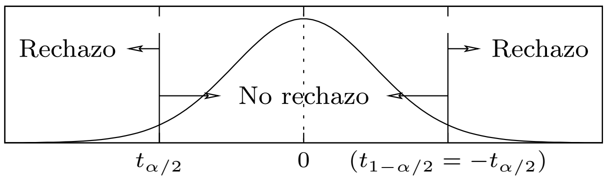
2. Contraste de hipótesis sobre coeficientes individuales de la regresión

2.1. Contrastes de dos colas

Contraste de la t: de dos colas **4**

1. $H_0 : \beta_j = b_j; \qquad H_1 : \beta_j \neq b_j$
2. (Ecuación (6.3), página 39 Tema 1) $\dfrac{\widehat{\beta_j} - b_j}{\widehat{\mathrm{Dt}}\left(\widehat{\beta_j}\right)} \equiv \mathcal{T}^j \underset{H_0}{\sim} t_{N-k}$

   Distribución t con $T-k$ grados de libertad

3. 

Donde
$$\widehat{\mathrm{Dt}}\left(\widehat{\beta_j}\right) \equiv \sqrt{\widehat{\mathrm{Var}}\left(\widehat{\beta_j}\right)}.$$

Ejemplo 1. [**continuación de "precio de las viviendas":**] Podemos contrastar la significación individual de la estimación de a.

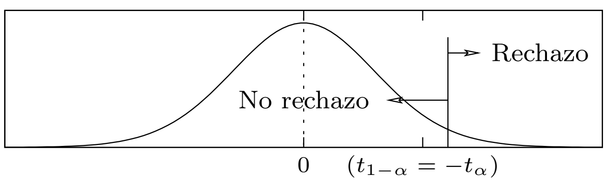
Puesto que no tiene mucho sentido que un piso con una superficie igual a cero tenga precio distinto de cero; podemos contrastar:
$$H_0: \ a = 0; \qquad H_1: \ a \neq 0$$
En este caso la región crítica debe ser
$$RC = \left\{ \mathbf{x} \;\middle|\; \frac{\widehat{a} - 0}{\sqrt{\widehat{\mathfrak{s}^2}\left((\mathbf{x}'\mathbf{x})^{-1}\right)_{11}}} > k_1 \text{ o } \frac{\widehat{a} - 0}{\sqrt{\widehat{\mathfrak{s}^2}\left((\mathbf{x}'\mathbf{x})^{-1}\right)_{11}}} < k_2 \right\},$$
donde $\frac{\widehat{a}-0}{\sqrt{\widehat{\mathfrak{s}^2}\left((\mathbf{x}'\mathbf{x})^{-1}\right)_{11}}} \equiv \mathcal{T}^a \underset{H_0}{\sim} t_{12}$. (o bien, para este caso particular, de (5.2))
$$RC = \left\{ \mathbf{x} \;\middle|\; \frac{\widehat{a} - 0}{\sqrt{\widehat{\mathfrak{s}^2}\frac{\sum x_t^2}{N\sum(x_t-\overline{x})^2}}} > k_1 \text{ o } \frac{\widehat{a} - 0}{\sqrt{\widehat{\mathfrak{s}^2}\frac{\sum x_t^2}{N\sum(x_t-\overline{x})^2}}} < k_2 \right\},$$
Para $\alpha = .05$ resulta $k_1 = t_{12,\ \alpha/2} = 2.16 = -k_2$, es decir
$$\widehat{\mathcal{T}}^a = \frac{52.351}{37.285} = 1.4041 \qquad \text{no rechazamos } H_0 \text{ con nivel de significación del } 5\,\%.$$
Véase los resultados de estimación en el ejemplo del precio de las viviendas (página 17 del Tema 1).

2.2. Contrastes de una sola cola

Contraste de la t: de una sóla cola (derecha) **5**

1. $$H_0 : \beta_j = b_j; \qquad H_1 : \beta_j > b_j$$
2. (Ecuación (6.3), página 39 Tema 1) $\dfrac{\widehat{\beta_j} - b_j}{\widehat{\mathrm{Dt}}\left(\widehat{\beta_j}\right)} \equiv \mathcal{T}^j \underset{H_0}{\sim} t_{N-k}$

   Distribución t con $T-k$ grados de libertad

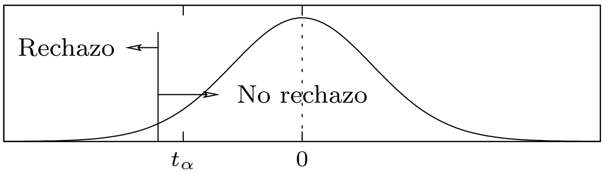
3. 

⇑ **Contraste de la t: de una sóla cola (izquierda)** 6

1.
$$H_0 : \beta_j = b_j; \qquad H_1 : \beta_j < b_j$$

2. (Ecuación (6.3), página 39 Tema 1) $\dfrac{\widehat{\beta_j} - b_j}{\widehat{\mathrm{Dt}}\left(\widehat{\beta_j}\right)} \equiv \mathcal{T}^j \underset{H_0}{\sim} t_{N-k}$

Distribución t con $T-k$ grados de libertad



3.

Ejemplo 2. [**continuación de "precio de las viviendas":**]
Un experto en el mercado de la vivienda afirma que un aumento de un pie cuadrado en la superficie supone un incremento de (como poco) 150 dolares, pero nunca menos. A la luz de los datos, ¿podemos creer al experto a un nivel de significación del 2.5 %?

$$H_0 : \; b = 0.15; \quad H_1 : \; b < 0.15$$

La región critica de una sóla cola es

$$RC = \left\{ \mathbf{x} \;\middle|\; \frac{\widehat{b} - 0.15}{\sqrt{\widehat{\mathfrak{s}^2}\left((\mathbf{x'}\,\mathbf{x})^{-1}\right)_{22}}} < k \right\}$$

(o bien, para este caso particular (5.2))

$$RC = \left\{ \mathbf{x} \;\middle|\; \frac{\widehat{b} - 0.15}{\sqrt{\frac{\widehat{\mathfrak{s}^2}}{\sum (x_t - \overline{x})^2}}} < k \right\}$$

sustituyendo lo que hemos calculado, tenemos que

$$\widehat{\mathcal{T}}^b = \frac{0.139 - 0.15}{0.01873} = -0.58729 > t_{12,\;0.025} = -2.16$$

no podemos rechazar la opinión del experto. Véase los resultados de estimación en el ejemplo del precio de las viviendas (página 17 del Tema 1).

2.3. Reglas de decisión para contrastes de una cola y de dos colas

⇑ **Contraste t de cola derecha: regla de decisión** 7

Dado un valor hipotético b_j para β_j;
un $\widehat{\mathcal{T}}^j$ *"grande y positivo"* es signo en contra de H_0.

Paso 1. Decidir nivel de significación α y buscar en tablas **valor crítico** $t_{N-k,\;1-\alpha}$ tal que, si H_0 es cierta

$$P\left(\mathcal{T}^j > t_{N-k,\;1-\alpha}\right) = \alpha$$

Paso 2. Rechazar H_0 si: $\quad \widehat{\mathcal{T}}^j > t_{N-k,\;1-\alpha}$.
No rechazar H_0 en caso contrario.

Donde $\widehat{\mathcal{T}}^j$ es la realización del estadístico (i.e., la variable aleatoria) $\mathcal{T}^j$ para una muestra concreta $\mathbf{x}$.

Contraste t de cola derecha: regla de decisión con el p-valor 8

Dado un valore hipotético b_j para β_j;
un $\widehat{\mathcal{T}}^j$ *"grande y positivo"* es signo en contra de H_0.
Paso 1. Calcular

$$p = P\left(\mathcal{T}^j > \widehat{\mathcal{T}}^j\right).$$

Paso 2. No rechazar H_0 si: $p > \alpha$. Rechazar H_0 en caso contrario.

Contraste t de cola izquierda: regla de decisión 9

Dado un valor hipotético b_j para β_j;
$\widehat{\mathcal{T}}^j$ *"negativo y grande en valor absoluto"*: signo en contra de H_0.
Paso 1. Decidir nivel de significación α y buscar en tablas **valor crítico** $t_{N-k,\ \alpha}$ tal que, si H_0 es cierta

$$P\left(\mathcal{T}^j < t_{N-k,\ \alpha}\right) = \alpha$$

Paso 2. Rechazar H_0 si: $\widehat{\mathcal{T}}^j < t_{N-k,\ \alpha}$
No rechazar H_0 en caso contrario.

Contraste t de cola izquierda: regla de decisión con el p-valor 10

Dado un valor hipotético b_j para β_j;
$\widehat{\mathcal{T}}^j$ *"negativo y grande en valor absoluto"*: signo en contra de H_0.
Paso 1. Calcular

$$p = P\left(\mathcal{T}^j < \widehat{\mathcal{T}}^j\right).$$

Paso 2. No rechazar H_0 si: $p > \alpha$. Rechazar H_0 en caso contrario.

Contraste de la t de dos colas: regla de decisión 11

Dado un valor hipotético b_j para β_j;
una desviación *"grande"* de $\widehat{\mathcal{T}}^j$ de cero es signo en contra de H_0.
Paso 1. Decidir nivel de significación α y buscar en tablas **valor crítico** $t_{N-k,\ \alpha/2}$ tal que, si H_0 es cierta

$$P\left(t_{N-k,\ \alpha/2} < \mathcal{T}^j < t_{N-k,\ 1-\alpha/2}\right) = 1-\alpha$$

Paso 2. No rechazar H_0 si:

$$t_{N-k,\ \alpha/2} < \mathcal{T}^j < t_{N-k,\ 1-\alpha/2}$$

rechazar H_0 en caso contrario.

Contraste t de dos colas: regla de decisión con el p-valor 12

Dado un valor hipotético b_j para β_j;
una desviación *"grande"* de $\widehat{\mathcal{T}}^j$ de cero es signo en contra de H_0.
Paso 1. Calcular

$$p = P\left(\mathcal{T}^j > \left|\widehat{\mathcal{T}}^j\right|\right) \times 2;$$

Paso 2. No rechazar H_0 si: $p > \alpha$. Rechazar H_0 en caso contrario.

"recuadro resultados de estimación en precio de las viviendas" (pag. 17 del Tema 1):

Ejercicio 3. Discuta la veracidad o falsedad de la siguiente afirmación:
Si la hipótesis nula no es rechazada, entonces es cierta.

EJERCICIO 4. Discuta la veracidad o falsedad de la siguiente afirmación:

Si se rechaza H_0 con un nivel de significación $\alpha = 0.05$ también se rechaza H_0 con un nivel de significación $\alpha = 0.10$

EJERCICIO 5. En el siguiente modelo de regresión simple, $Y = \alpha + \beta X + \varepsilon$, Y mide la tasa de crecimiento del PIB real y X la tasa de crecimiento de la masa monetaria. Tenemos una muestra de tamaño 35. Queremos contrastar la hipótesis de neutralidad monetaria en esta muestra (es decir, la masa monetaria no tiene efecto real sobre el nivel del PIB). Proponga la hipótesis nula y alternativa que, según la t$^{\text{a}}$ económica, crea más conveniente para realizar el contraste, así como la región crítica. ¿Cual es la distribución del estadístico de contraste?

3. Contraste de hipótesis sobre combinaciones lineales de coeficientes de la regresión

⇑ **Hipótesis lineales** 13

$$H_0 : \mathbf{R}\boldsymbol{\beta} = \boldsymbol{r},$$

$\underset{[r\times k]}{\mathbf{R}}$ es matriz con rango $(\mathbf{R}) = r$, $\quad (r \leq k)$; $\quad$ y $\underset{[r\times 1]}{\boldsymbol{r}}$ es vector.
Las r ecuaciones son restricciones sobre valores de los coeficientes.
Condición rango $(\mathbf{R}) = r$, garantiza:

- no restricciones redundantes
- no restricciones incompatibles

Ejemplo 6. [**ecuación de salarios:** (continuación Ejemplo 2 en la página 6 del tema 1)]

Supongamos que queremos contrastar si educación y antigüedad tienen el mismo efecto en el incremento del salario, y que además, la experiencia no tiene ningún efecto (por tanto $r = 2$)

$$\beta_2 = \beta_3 \qquad \text{y} \qquad \beta_4 = 0.$$

En forma matricial, $H_0 : \mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$, donde

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{r} = \begin{bmatrix} 0 \\ 0 \end{bmatrix};$$

donde $\mathbf{R}$ cumple la condición de rango completo.

Añadiendo restricciones que no cumplen la condición de rango:

- Supongamos que adicionalmente imponemos que
$$\beta_2 - \beta_3 = \beta_4.$$
Esta es una restricción redundante, pues ya se cumple con las dos primeras restricciones; en forma matricial
$$\mathbf{R} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix}, \quad \boldsymbol{r} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$
- Supongamos que imponemos una condición incompatible con las dos primeras:
$$\beta_4 = 0.5,$$
que evidentemente es incompatible con $\beta_4 = 0$. Matricialmente
$$\mathbf{R} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{r} = \begin{bmatrix} 0 \\ 0 \\ 0.5 \end{bmatrix}.$$
De nuevo la condición de rango se incumple.

3.1. El test F

⇑ ratio-F 14

Bajo supuestos 1 a 5; y si H_0: $\mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$ cierta,
donde rango $\left(\underset{[r\times k]}{\mathbf{R}}\right) = r$, definimos el **ratio-*F***:

$$\mathcal{F} = \frac{\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)' \left[\mathbf{R}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)/r}{\widehat{s}^2} \tag{3.1}$$

$$= \left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)' \left[\mathbf{R}\,\widehat{\mathrm{Var}}\left(\widehat{\boldsymbol{\beta}}\right)\mathbf{R}'\right]^{-1} \left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)/r \tag{3.2}$$

por (6.2) del Tema 1 (pág. 39)

$$\mathcal{F} \underset{H_0}{\sim} F_{r,N-k}$$

Nota 1. Sean $W \sim \chi^2_{(r)}$ y $\frac{\widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}}}{\sigma^2} \sim \chi^2_{(N-k)}$ dos variables aleatorias independientes, entonces (Demostración en (Mittelhammer, 1996, Teorema 6.21, página 345))

$$\frac{W/r}{\frac{\widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}}}{\sigma^2}/(N-k)} \underset{H_0}{\sim} F_{r,N-k}$$

Proposición 3.1 (Distribución del ratio *F*:). *El ratio-F se distribuye como una F con r y $N-k$ grados de libertad, es decir $\mathcal{F} \underset{H_0}{\sim} F_{r,N-k}$*

Demostración. La demostración es similar a la de la Proposición 6.7 (pág. 40 del Tema 1) [1]. Primero escribimos el estadístico en la forma

$$\mathcal{F} = \frac{W/r}{\frac{\widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}}}{\sigma^2}/(N-k)}$$

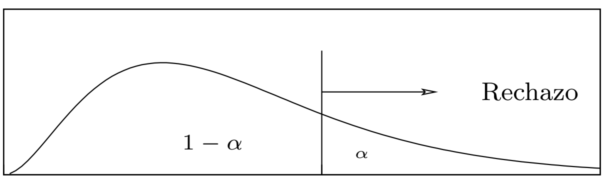
donde $W = \left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)' \left[\sigma^2\,\mathbf{R}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)$.

Los pasos a demostrar son:

1. $W \underset{H_0}{\sim} \chi^2_{(r)}$
   (Bajo H_0 se verifica que $\mathrm{Var}\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r} \,\middle|\, \mathbf{x}\right) = \sigma^2\,\mathbf{R}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{R}'$ y puesto que $\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right) \underset{H_0}{\sim} \mathrm{N}$ aplicando la Nota 6.4, pag. 39 del Tema 1 queda demostrado.)
2. $\frac{\widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}}}{\sigma^2} \underset{H_0}{\sim} \chi^2_{(N-k)}$.
   (es la Proposición 6.5 del Tema 1).
3. W es independiente de $\frac{\widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}}}{\sigma^2}$
   (la demostración de este punto es idéntica al razonamiento empleado en la Proposición 6.7, pág. 40 del Tema 1).

La demostración se completa con la Nota 1. □

⇑ Contraste de la F 15

1. $H_0 : \mathbf{R}\,\boldsymbol{\beta} \;=\; \boldsymbol{r}$; $\qquad H_1 : \mathbf{R}\,\boldsymbol{\beta} \;\neq\; \boldsymbol{r}$
2. $\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)' \left[\mathbf{R}\,\widehat{\mathrm{Var}}\left(\widehat{\boldsymbol{\beta}}\right)\mathbf{R}'\right]^{-1} \left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)/r \underset{H_0}{\sim} F_{r,N-k}$
3. 

   0 $\quad F_{1-\alpha}$

[1] y la puede encontrar completa en Hayashi (2000, pp. 41)

3.2. Reglas de decisión para el contraste de la F

⇑ **Contraste F: regla de decisión** 16

Paso 1. Calcular ratio-F mediante (3.2), un valor *"grande"* de $\widehat{\mathcal{F}}$ es signo en contra de $H_0: \; \mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$.

Paso 2. Decidir nivel de significación α y buscar en tablas **valor crítico** $F_{r,N-k,1-\alpha}$ tal que, si H_0 es cierta

$$P\left(\mathcal{F} > F_{r,N-k,1-\alpha}\right) = \alpha$$

Paso 3. Rechazar H_0 si: $\quad \widehat{\mathcal{F}} > F_{r,N-k,1-\alpha}$
No rechazar H_0 en caso contrario.

Donde $\widehat{\mathcal{F}}$ es la realización del estadístico (i.e., la variable aleatoria) $\mathcal{F}$ para una muestra concreta $\mathbf{x}$.

⇑ **Contraste F: regla de decisión con el p-valor** 17

Paso 1. Calcular el ratio-F mediante (3.2), un valor *"grande"* de $\widehat{\mathcal{F}}$ es signo en contra de $H_0: \; \mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$.

Paso 2. Calcular

$$p = P\left(\mathcal{F} > \widehat{\mathcal{F}}\right).$$

Paso 3. No rechazar H_0 si: $p > \alpha$. Rechazar H_0 en caso contrario.

3.3. t versus F

⇑ **t versus F** 18

Contrastación de hipótesis individual es caso particular, donde $r = 1$ y

$$\underset{[1\times k]}{\mathbf{R}} = \begin{bmatrix} 0 & \cdots & 0 & \underset{(j)}{1} & 0 & \cdots & 0 \end{bmatrix}, \quad \boldsymbol{r} = b_j$$

(3.2) se reduce a

$$\begin{aligned} \mathcal{F} &= \left(\widehat{\beta_j} - b_j\right)' \left[\mathbf{R}\,\widehat{\mathrm{Var}}\left(\widehat{\boldsymbol{\beta}}\right)\mathbf{R}'\right]^{-1} \left(\widehat{\beta_j} - b_j\right)/1 \\ &= \left(\widehat{\beta_j} - b_j\right)' \left[\widehat{\mathrm{Var}}\left(\widehat{\beta_j}\right)\right]^{-1} \left(\widehat{\beta_j} - b_j\right) \underset{H_0}{\sim} F_{1,N-k} \end{aligned} \tag{3.3}$$

que es cuadrado del estadístico $\mathcal{T}$ de (6.3), página 39 del Tema 1 [a].

[a] $F_{1,N-k}$ es el cuadrado de una t_{N-k}.

Nota 2. Nótese que

$$\begin{aligned} \mathcal{F} &= \left(\widehat{\beta_j} - b_j\right)' \left[\mathbf{R}\,\widehat{\mathrm{Var}}\left(\widehat{\boldsymbol{\beta}}\right)\mathbf{R}'\right]^{-1} \left(\widehat{\beta_j} - b_j\right)/1 \\ &= \left(\widehat{\beta_j} - b_j\right)' \left[\widehat{\mathrm{Var}}\left(\widehat{\beta_j}\right)\right]^{-1} \left(\widehat{\beta_j} - b_j\right) \underset{H_0}{\sim} F_{1,N-k} \quad \text{pues se selecciona el elemento } j \text{ de la diagonal} \\ &= \frac{\left(\widehat{\beta_j} - b_j\right)^2}{\widehat{\mathrm{Var}}\left(\widehat{\beta_j}\right)} = \left(\frac{\widehat{\beta_j} - b_j}{\widehat{\mathrm{Dt}}\left(\widehat{\beta_j}\right)}\right)^2 = (\mathcal{T})^2 \qquad \text{por ser un escalar} \end{aligned}$$

donde

$$\mathcal{T} \underset{H_0}{\sim} t_{N-k}$$

Si el contrate individual $\mathcal{T}$ es de una sóla cola, no se puede realizar mediante $\mathcal{F}$. El estadístico $\mathcal{F}$ sólo tiene en cuenta las desviaciones al cuadrado, por lo tanto, penaliza de igual manera una diferencia positiva entre H_0 y la estimación obtenida, que si la diferencia es negativa.

Nota 3. No sólo el contraste de siginificación individual tiene una distribución $\mathcal{T}^2$. Siempre que $\underset{[1\times k]}{\mathbf{R}} =$

$\begin{bmatrix} r_1, & r_2, & \cdots & r_k \end{bmatrix}$ y, consecuentemente, $\boldsymbol{r}$ sea un escalar (es decir, siempre que haya una única restricción lineal), el estadístico resultante es $\mathcal{F} = (\mathcal{T})^2$;veámoslo:

$$\mathcal{F} = \left(r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k} - b\right)' \left[\mathbf{R}\,\widehat{\text{Var}}\left(\widehat{\boldsymbol{\beta}}\right)\mathbf{R}'\right]^{-1} \left(r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k} - b\right)/1$$

operando dentro del corchete tenemos:

$$= \left(r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k} - b\right)' \left[\widehat{\text{Var}}\left(\ r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k}\right)\right]^{-1} \left(r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k} - b\right)$$

y por ser una expresión escalar:

$$= \frac{\left(r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k} - b\right)^2}{\widehat{\text{Var}}\left(\ r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k}\right)} = \left(\frac{r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k} - b}{\widehat{\text{Dt}}\left(\ r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k}\right)}\right)^2 = (\mathcal{T})^2,$$

ya que $r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k}$ es una combinación lineal de Normales, es una variable aleatoria escalar con distribución Normal[2]; así pues, la expresión dentro del paréntesis cumple las mismas propiedades descritas en la Proposición 6.7 en la página˜40 del Tema 1; y por tanto

$$\frac{\left(r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k} - b\right)}{\widehat{\text{Dt}}\left(\ r_1\widehat{\beta_1} + \cdots + r_k\widehat{\beta_k}\right)} = \frac{\mathbf{R}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)}{\widehat{\text{Dt}}\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}}\right)} = \mathcal{T} \underset{H_0}{\sim} t_{N-k} \qquad \text{donde } \mathbf{R}\boldsymbol{\beta} = b.$$

Significación conjunta del modelo Con el contraste $\mathcal{F}$ se puede contrastar la significación conjunta de varios (o todos) los coeficientes de la regresión.

Por ejemplo, para el contraste de significación conjunta basta con:

$$\underset{[(k-1)\times k]}{\mathbf{R}} = \left[\ \underset{[(k-1)\times 1]}{\mathbf{0}} \qquad \underset{[(k-1)\times(k-1)]}{\mathbf{I}}\right]; \qquad \boldsymbol{r} = \underset{[(k-1)\times 1]}{\mathbf{0}}.$$

En este caso:

H_0 : todos los coeficientes (excepto el de la constante) son nulos;

H_1 : al menos uno es distinto de cero.

Este contraste no es equivalente a realizar k contrastes individuales *por separado.*

Este estadístico F para la contrastación de significación conjunta es el que aparece en las regresiones de los programas informáticos.

GNU Gretl (precio vivienda simulado)
"recuadro resultados de estimación en precio de las viviendas" (pag. 17 del Tema 1):
GNU Gretl MLG peso bbs

A resolver en clase

EJERCICIO 7. [**Novales (1993, Ejercicio 4.7; pag 153)**]

(a) Contrastar $H_0 : 3\beta_2 - \beta_3 = 7$ frente a $H_1 : 3\beta_2 - \beta_3 \neq 7$ en el modelo

$$Y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + U_t;$$

si se han obtenido los siguientes resultados: $\widehat{\beta_2}$ =4.5; $\widehat{\beta_3}$ =5.0; SRC =40; N =25;

$$(\mathbf{x}'\,\mathbf{x})^{-1} = \begin{pmatrix} 50 & 13 & 21 & -120 \\ & 4 & 2 & -20 \\ & & 6 & -10 \\ & & & 10 \end{pmatrix}$$

---

[2]nótese que si los estimadores de las covarianzas de los estimadores $\widehat{\text{Cov}}\left(\widehat{\beta_i}, \widehat{\beta_j}\right)$ fuesen cero, entonces

$$\left[\mathbf{R}\,\widehat{\text{Var}}\left(\widehat{\boldsymbol{\beta}}\right)\mathbf{R}'\right] = \left[r_1^2\widehat{\text{Var}}\left(\widehat{\beta_1}\right) + \cdots + r_k^2\widehat{\text{Var}}\left(\widehat{\beta_k}\right)\right]$$

es decir, la suma de las varianzas multiplicadas por los factores r_j^2; pero como en general dichas covarianzas nunca son cero, la expresión es más complicada pues se deben añadir los términos $2r_ir_j\widehat{\text{Cov}}\left(\widehat{\beta_i}, \widehat{\beta_j}\right)$, para todos los pares $i \neq j$.

(b) ¿Cómo se contrastaría la hipótesis anterior junto con la hipótesis $\beta_4 = 0$, si se ha estimado $\widehat{\beta_4} = -1.5$?

4. Regiones e intervalos de confianza

Sea $\underset{[1\times k]}{\mathbf{R}}$, de la Nota 3 en la página 9 (o bien de (3.2) y (3.3)) se deduce la primera expresión de la siguiente trasparencia:

⇑ **Intervalo de confianza de una combinación lineal de betas** **19**

De la Nota 3 en la página 9 se deduce que si $\underset{[1\times k]}{\mathbf{R}}$

$$\frac{\mathbf{R}\left(\widehat{\boldsymbol\beta}-\boldsymbol\beta\right)}{\widehat{\mathrm{Dt}}\left(\mathbf{R}\widehat{\boldsymbol\beta}\right)} \underset{H_0}{\sim} t_{N-k}$$

Entonces

$$P\left(t_{N-k;\;\alpha/2} < \frac{\mathbf{R}\left(\widehat{\boldsymbol\beta}-\boldsymbol\beta\right)}{\widehat{\mathrm{Dt}}\left(\mathbf{R}\widehat{\boldsymbol\beta}\right)} < t_{N-k;\;1-\alpha/2}\right) = 1-\alpha \tag{4.1}$$

Un intervalo de confianza en una función aleatoria que cumple cierta condición. En el caso de una combinación lineal de $\boldsymbol\beta$ la condición es que el invervalo verifique que

$$P\left(\quad \mathbf{R}\boldsymbol\beta \quad\in\quad \mathrm{IC}^{\mathbf{R}\widehat{\boldsymbol\beta}}_{1-\alpha}(\mathbf{W}) \quad\right) = 1-\alpha, \qquad \text{para cualquier } \boldsymbol\beta\in\Theta,$$

donde $\mathbf{W} = \left[\boldsymbol{Y}, \mathbf{X}\right]$ y donde, para la muestra concreta $\mathbf{w}$, el conjunto de valores $\boldsymbol{b}$ es el intervalo:

$$\begin{aligned}\mathrm{IC}^{\mathbf{R}\widehat{\boldsymbol\beta}}_{1-\alpha}(\mathbf{w}) &= \left\{\boldsymbol{b} \text{ tales que } t_{N-k;\;\alpha/2} < \frac{\mathbf{R}\left(\widehat{\boldsymbol\beta}-\boldsymbol{b}\right)}{\widehat{\mathrm{Dt}}\left(\mathbf{R}\widehat{\boldsymbol\beta}\right)} < t_{N-k;\;1-\alpha/2}\right\}\\ &= \left\{\boldsymbol{b} \text{ tales que } \mathbf{R}\boldsymbol{b} \in \left[\ \mathbf{R}\widehat{\boldsymbol\beta} \ \pm\ t_{N-k;\;\alpha/2}\cdot\widehat{\mathrm{Dt}}\left(\mathbf{R}\widehat{\boldsymbol\beta}\right)\right]\right\}\end{aligned}$$

⇑ **Intervalo de confianza de una combinación lineal de betas** **20**

$$P\left(\quad \mathbf{R}\boldsymbol\beta \quad\in\quad \mathrm{IC}^{\mathbf{R}\widehat{\boldsymbol\beta}}_{1-\alpha}(\mathbf{W}) \quad\right) = 1-\alpha \quad \text{donde}$$

$$\mathrm{IC}^{\mathbf{R}\widehat{\boldsymbol\beta}}_{1-\alpha}(\mathbf{W}) = \left\{\boldsymbol{b} \text{ tales que } \mathbf{R}\boldsymbol{b} \in \left[\ \mathbf{R}\widehat{\boldsymbol\beta} \ \pm\ t_{N-k;\;\alpha/2}\cdot\widehat{\mathrm{Dt}}\left(\mathbf{R}\widehat{\boldsymbol\beta}\right)\right]\right\}$$

y donde $\mathbf{W} = \left[\boldsymbol{Y}, \mathbf{X}\right]$; y donde para una muestra concreta $\mathbf{w}$

$$\mathrm{IC}^{\mathbf{R}\widehat{\boldsymbol\beta}}_{1-\alpha}(\mathbf{w}) = \left\{\boldsymbol{b} \text{ tales que } \mathbf{R}\boldsymbol{b} \in \left[\ \mathbf{R}\widehat{\boldsymbol\beta} \ \pm\ t_{N-k;\;\alpha/2}\cdot\widehat{\mathrm{Dt}}\left(\mathbf{R}\widehat{\boldsymbol\beta}\right)\right]\right\}$$

En el caso del intervalo de un parámetro individual (3.3) $\underset{[1\times k]}{\mathbf{R}} = \begin{bmatrix}0 & \cdots & 0 & \underset{(j)}{1} & 0 & \cdots & 0\end{bmatrix}$,

$$\mathrm{IC}^{\widehat{\beta_j}}_{1-\alpha}(\mathbf{w}) = \left[\ \widehat{\beta_j} \ \pm\ t_{N-k;\;\alpha/2}\cdot\widehat{\mathrm{Dt}}\left(\widehat{\beta_j}\right)\right]$$

La amplitud de los intervalos de confianza es una medición de la precisión de los estimadores.

Ejemplo 8. [**continuación de "precio de las viviendas":**] Los intervalos de confianza de los parámetros a y b son de la forma

$$\mathrm{IC}^{\widehat{\beta_j}}_{1-\alpha}(\mathbf{w}) = \left[\ \widehat{\beta_j} \ \pm\ t_{N-k;\;\alpha/2}\cdot\widehat{\mathrm{Dt}}\left(\widehat{\beta_j}\right)\right]$$

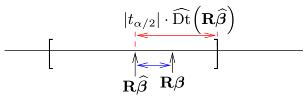
por tanto, En el caso del efecto marginal de la superficie sobre el precio y de la constante son respectivamente

$$b \in \left[0.139 \pm (t_{(12,\alpha/2)}) \cdot 0.01873\right]; \qquad a \in \left[52.3509 \pm (t_{(12,\alpha/2)}) \cdot 37.285\right];$$

Intervalos de confianza y contrastes bilaterales **21**

$\mathbf{R}\boldsymbol{\beta}$ pertenece al intervalo de confianza $\text{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{w})$, si $\mathbf{R}\boldsymbol{\beta} \in \left[\mathbf{R}\widehat{\boldsymbol{\beta}} \pm t_{\alpha/2} \cdot \widehat{\text{Dt}}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}\right)\right]$; Por tanto $\mathbf{R}\boldsymbol{\beta}$ está en el intervalo de confianza si

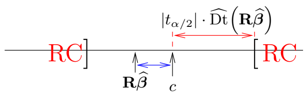
$$\left|\mathbf{R}\boldsymbol{\beta} - \mathbf{R}\widehat{\boldsymbol{\beta}}\right| \leq |t_{\alpha/2}| \cdot \widehat{\text{Dt}}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}\right) \tag{4.2}$$



Por otra parte **no rechazamos** H_0: $\mathbf{R}\boldsymbol{\beta} = c$ $[H_1$: $\mathbf{R}\boldsymbol{\beta} \neq c]$ **si:** $\mathbf{R}\widehat{\boldsymbol{\beta}} \in \left[c \pm |t_{\alpha/2}| \cdot \widehat{\text{Dt}}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}\right)\right]$; es decir, si

$$\left|\mathbf{R}\widehat{\boldsymbol{\beta}} - c\right| \leq |t_{\alpha/2}| \cdot \widehat{\text{Dt}}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}\right) \tag{4.3}$$

donde por H_0: $\mathbf{R}\boldsymbol{\beta} = c$



Comparando (4.2) y (4.3) es sencillo comprobar que todas las hipótesis $H_0: \ \mathbf{R}\boldsymbol{\beta} = c^*$ que cumplen (4.3) (i.e., que no se rechazarían), son puntos que pertenecen al intervalo de confianza de $\mathbf{R}\boldsymbol{\beta}$ (por (4.2)). Por tanto el intervalo con nivel de confianza $(1-\alpha)$ contiene todas las hipótesis nulas $H_0: \ \mathbf{R}\boldsymbol{\beta} = c^*$ que no se rechazarían en un contraste **bilateral** con una significación α.

Intervalos y contrastes **22**

$$\boxed{\text{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{w}) \quad = \quad \text{Hipótesis no rechazables con significación } \alpha}$$

Intervalos de confianza frente a contrastes de hipótesis **23**

Contrastes bilaterales al $\alpha\,\%$ e intervalos de confianza al $1-\alpha\,\%$; donde $\underset{[1\times k]}{\mathbf{R}}$

- Rechazar $H_0: \mathbf{R}\boldsymbol{\beta} = c$; frente $H_1: \mathbf{R}\boldsymbol{\beta} \neq c$;

  $$\text{si y sólo si} \quad c \notin \text{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{w})$$

- Rechazar $H_0: \ \beta_j = b_j$; frente $H_1: \ \beta_j \neq b_j$;

  $$\text{si y sólo si} \quad b_j \notin \text{IC}_{1-\alpha}^{\widehat{\beta_j}}(\mathbf{w})$$

Ejemplo 9. [**continuación de "precio de las viviendas"**] Valores de $t_{\alpha/2}$ con GRETL

⇑ **Regiones de confianza** 24

Si $\mathbf{R}$ es de tipo:

$$\underset{[r\times k]}{\mathbf{R}} = \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \end{pmatrix};$$

de 3.1 en la página˜8 se deduce que la región

$$P\left(\frac{\left(\widehat{\boldsymbol{\beta}_*}-\boldsymbol{\beta}_*\right)'\mathbf{x_*}'\mathbf{x_*}\left(\widehat{\boldsymbol{\beta}_*}-\boldsymbol{\beta}_*\right)/r}{\widehat{\mathfrak{s}}^2} < F_{r,\ N-k,\ \alpha}\right) =$$

$$P\left(\left(\widehat{\boldsymbol{\beta}_*}-\boldsymbol{\beta}_*\right)'\mathbf{x_*}'\mathbf{x_*}\left(\widehat{\boldsymbol{\beta}_*}-\boldsymbol{\beta}_*\right) < r\widehat{\mathfrak{s}}^2\cdot F_{r,\ N-k,\ \alpha}\right) = 1-\alpha$$

es una elipsoide en $\mathbb{R}^r$, donde $\boldsymbol{\beta}_* = \begin{bmatrix}\beta_1 & \beta_2 & \beta_4 & \beta_5\end{bmatrix}'$ y $\mathbf{x_*} = \begin{bmatrix}\boldsymbol{x}_{\blacktriangledown 1} & \boldsymbol{x}_{\blacktriangledown 2} & \boldsymbol{x}_{\blacktriangledown 4} & \boldsymbol{x}_{\blacktriangledown 5}\end{bmatrix}$

Y por tanto la región de confianza debe verificar

$$P\Big(\quad \mathbf{R}\boldsymbol{\beta} \quad \in \quad \mathrm{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{W}) \quad\Big) = 1-\alpha$$

donde $\mathrm{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{w})$ es el conjunto de puntos $\boldsymbol{b}$ pertenecientes $\mathbb{R}^r$ que verifican lo siguiente:

$$\mathrm{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{w}) = \left\{\boldsymbol{b} \text{ tales que } \left(\widehat{\boldsymbol{\beta}}-\boldsymbol{b}\right)'\mathbf{x}'\mathbf{x}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{b}\right) < r\widehat{\mathfrak{s}}^2\cdot F_{r,\ N-k,\ \alpha}\right\}$$

que es una elipsoide si $r > 1$

Ejemplo 10. $k=2$; $\quad H_0:\ \beta_1 = a,\quad$ y $\quad \beta_2 = b$
$\mathbf{R}\boldsymbol{\beta}{=}\boldsymbol{r}$; $\quad \mathbf{R} = \left(\begin{smallmatrix}1 & 0\\ 1 & 0\end{smallmatrix}\right)$; $\quad \boldsymbol{y} = \left(\begin{smallmatrix}a\\ b\end{smallmatrix}\right)$.

Solución:

solución tentativa pero incorrecta No rechazar si

$$(a,b) \in \text{región tal que } \begin{cases}\widehat{\beta_1} - \left|t_{\alpha/2}\right|\cdot\widehat{\mathrm{Dt}}\left(\widehat{\beta_1}\right) < a < \widehat{\beta_1} + \left|t_{\alpha/2}\right|\cdot\widehat{\mathrm{Dt}}\left(\widehat{\beta_1}\right)\\ \widehat{\beta_2} - \left|t_{\alpha/2}\right|\cdot\widehat{\mathrm{Dt}}\left(\widehat{\beta_2}\right) < b < \widehat{\beta_2} + \left|t_{\alpha/2}\right|\cdot\widehat{\mathrm{Dt}}\left(\widehat{\beta_2}\right)\end{cases}$$

que es un rectángulo (formado por el producto cartesiano de los intervalos de confianza individuales).

solución correcta No rechazar si

$$(a,b) \in \left\{\text{región tal que } \left(\widehat{\beta_1}-a,\quad \widehat{\beta_2}-b\right)\left(\widehat{\mathrm{Var}}\left(\widehat{\beta}\right)\right)^{-1}\begin{pmatrix}\widehat{\beta_1}-a\\ \widehat{\beta_2}-b\end{pmatrix} < 2\cdot F_{r,N-k}(\alpha)\right\}$$

que es una elipse.

□

⇑ **Intervalo de confianza para la varianza** 25

Por Proposición 6.5 en la página˜39 del Tema 1

$$\frac{N-k}{\sigma^2}\widehat{\mathfrak{s}}^2_{\widehat{e}} = \frac{\widehat{e}'\,\widehat{e}}{\sigma^2} \underset{H_0}{\sim} \chi^2_{(N-k)}.$$

Por tanto

$$P\left(\chi^2_{(N-k;\ \alpha/2)} < \frac{\widehat{e}'\,\widehat{e}}{\sigma^2} < \chi^2_{(N-k;\ 1-\alpha/2)}\right) = 1-\alpha$$

o bien

$$P\left(\frac{\widehat{e}'\,\widehat{e}}{\chi^2_{(N-k;\ 1-\alpha/2)}} < \sigma^2 < \frac{\widehat{e}'\,\widehat{e}}{\chi^2_{(N-k;\ \alpha/2)}}\right) = 1-\alpha \tag{4.4}$$

⇑ **Intervalo de confianza para la varianza** 26

$$P\Big(\quad \sigma^2 \quad \in \quad \mathrm{IC}_{1-\alpha}^{\widehat{\sigma^2}}(\mathbf{W}) \quad \Big) = 1-\alpha$$

donde

$$\mathrm{IC}_{1-\alpha}^{\widehat{\sigma^2}}(\mathbf{W}) = \left[\frac{\widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}}}{\chi^2_{(N-k;\;1-\alpha/2)}}, \quad \frac{\widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}}}{\chi^2_{(N-k;\;\alpha/2)}}\right]$$

5. Estimación bajo restricciones lineales generales

⇑ **Estimación restringida** 27

Motivos:

- análisis previo → restricciones plausibles
  (estimaciones más precisas, si información correcta)
- comparación entre estimación restringida y no restringida permite contrastar la validez de las restricciones

Ejecución:

- por sustitución
- método de mínimos cuadrados restringidos linealmente (MCRL)

Ejemplo 11. [**Estimación restringida mediante sustitución (Gujarati, 2003, pag. 266–270):**]
Supongamos que tenemos el siguiente modelo en logaritmos proveniente de una función de Cobb-Douglas:

$$\ln Y_t = \beta_1 + \beta_2 \ln K_t + \beta_3 \ln L_t + U$$

Consideremos la siguiente restricción: $\beta_2 + \beta_3 = 1$

La estimación MCO imponiendo la restricción de *rendimientos constantes a escala* se puede obtener sencillamente re-escribiendo el modelo de siguiente modo:

$$\begin{aligned} \ln Y_t &= \beta_1 + \beta_2 \ln K_t + (1-\beta_2)\ln L_t + U \\ \ln Y_t - \ln L_t &= \beta_1 + \beta_2(\ln K_t - \ln L_t) + U \\ \ln \frac{Y_t}{L_t} &= \beta_1 + \beta_2 \ln \frac{K_t}{L_t} + U \end{aligned}$$

y estimando por MCO el modelo con las nuevas variables redefinidas.

⇑ **Mínimos cuadrados restringidos** 28

Bajo supuestos 1 a 5, buscamos estimación $\widehat{\boldsymbol{\beta}^*}$ de $\boldsymbol{\beta}$ que cumpla el conjunto de restricciones lineales:

$$\mathbf{R}\widehat{\boldsymbol{\beta}^*} = \boldsymbol{r},$$

$\underset{[r\times k]}{\mathbf{R}}$ es de rango completo; rango $(\mathbf{R}) = r \quad (r \leq k)$; y $\underset{[r\times 1]}{\boldsymbol{r}}$ un vector.
La Estimación de Mínimos Cuadrados con Restricciones Lineales (MCRL)

$$\widehat{\boldsymbol{\beta}^*} = \widehat{\boldsymbol{\beta}} + (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\boldsymbol{r} - \mathbf{R}\widehat{\boldsymbol{\beta}}\right) \tag{5.1}$$

logra $SRC(\widehat{\boldsymbol{\beta}^*})$ mínima, donde $\widehat{\boldsymbol{e}}^* = \boldsymbol{y} - \mathbf{x}\widehat{\boldsymbol{\beta}^*}$.
El Estimador es

$$\widehat{\boldsymbol{\beta}^*} = \widehat{\boldsymbol{\beta}} + (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\boldsymbol{r} - \mathbf{R}\,\widehat{\boldsymbol{\beta}}\right) \tag{5.2}$$

Proposición 5.1. *De todas las estimaciones $\widetilde{\boldsymbol{\beta}}$ que verifican la restricción lineal $\mathbf{R}\widetilde{\boldsymbol{\beta}} = \boldsymbol{r}$, la estimación de Mínimos Cuadrados con Restricciones Lineales, $\widehat{\boldsymbol{\beta}^*}$, alcanza la mínima suma de residuos al cuadrado $SRC(\widetilde{\boldsymbol{\beta}})$.*

Demostración. Primero demostramos que $\widehat{\boldsymbol{\beta}^*}$ cumple la restricción:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}^*} &= \widehat{\boldsymbol{\beta}} + (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\boldsymbol{r}-\mathbf{R}\widehat{\boldsymbol{\beta}}\right) \\
\mathbf{R}\widehat{\boldsymbol{\beta}^*} &= \mathbf{R}\widehat{\boldsymbol{\beta}} + \mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\boldsymbol{r}-\mathbf{R}\widehat{\boldsymbol{\beta}}\right) && \text{premultiplicando por } \mathbf{R} \\
\mathbf{R}\widehat{\boldsymbol{\beta}^*} &= \mathbf{R}\widehat{\boldsymbol{\beta}} + \left(\boldsymbol{r}-\mathbf{R}\widehat{\boldsymbol{\beta}}\right) && \text{pues } \mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}']^{-1}=\mathbf{I} \\
\mathbf{R}\widehat{\boldsymbol{\beta}^*} &= \boldsymbol{r}
\end{aligned}$$

La demostración de la segunda parte también es sencilla. Puesto que

$$\begin{aligned}
\widehat{\boldsymbol{e}}^* &= \boldsymbol{y} - \mathbf{x}\widehat{\boldsymbol{\beta}^*} \\
&= \boldsymbol{y} - \mathbf{x}\widehat{\boldsymbol{\beta}} - (\mathbf{x}\widehat{\boldsymbol{\beta}^*} - \mathbf{x}\widehat{\boldsymbol{\beta}}) && \text{sumando y restando } \mathbf{x}\widehat{\boldsymbol{\beta}} \\
&= \widehat{\boldsymbol{e}} - \mathbf{x}\left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right); && (5.3)
\end{aligned}$$

tenemos

$$\begin{aligned}
\widehat{\boldsymbol{e}}^{*\prime}\,\widehat{\boldsymbol{e}}^* &= \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}} + \left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right)'\mathbf{x}'\,\mathbf{x}\left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right) - \left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right)'\underbrace{\mathbf{x}'\,\widehat{\boldsymbol{e}}}_{\mathbf{0}} - \underbrace{\widehat{\boldsymbol{e}}'\,\mathbf{x}}_{\mathbf{0}'}\left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right) \\
&= \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}} + \left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right)'\mathbf{x}'\,\mathbf{x}\left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right), && (5.4)
\end{aligned}$$

ya que $\mathbf{x}'\,\widehat{\boldsymbol{e}} = \mathbf{0}$ por construcción del estimador MCO. Y debido a que $(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})'\,\mathbf{x}'\,\mathbf{x}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})$ es una suma de cuadrados (y por tanto semi-definido positivo), se deduce que

$$SRC(\widetilde{\boldsymbol{\beta}}) = \widehat{\boldsymbol{e}}^{*\prime}\,\widehat{\boldsymbol{e}}^* \geq \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}} = SRC(\widehat{\beta}).$$

□

En los libros de texto encontrará una demostración alternativa derivando el Lagrangiano de un problema de minimización con restricciones lineales (cf., Novales (1993, pag. 132)).

5.1. Propiedades estadísticas del estimador MCRL $\widehat{\boldsymbol{\beta}^*}$

⇑ **Estimador MCRL $\widehat{\boldsymbol{\beta}^*}$** 29

El estimador **siempre verifica** la condición: $\mathbf{R}\,\widehat{\boldsymbol{\beta}^*} = \boldsymbol{r}$

Si la restricción impuesta $\mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$ se cumple en realidad (si es cierta), de (5.2)

$$\mathrm{E}\left(\widehat{\boldsymbol{\beta}^*}\,\middle|\,\mathbf{x}\right) = \boldsymbol{\beta} \qquad \text{¡Ojo! sólo si se cumple restricción. ¿probable?}$$

Sea o no sea cierta la restricción $\mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$,

$$\mathrm{Var}\left(\widehat{\boldsymbol{\beta}^*}\,\middle|\,\mathbf{x}\right) = \mathrm{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right) - \sigma^2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}$$

es decir

$$\mathrm{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right) \geq \mathrm{Var}\left(\widehat{\boldsymbol{\beta}^*}\,\middle|\,\mathbf{x}\right)$$

Ejercicio 12. Demuestre las propiedades estadísticas de la trasparencia anterior.

Ejercicio 13. Si los valores $\boldsymbol{\beta}$ son desconocidos ¿cuál es la probabilidad de que una restricción impuesta al modelo se cumpla realmente? (Piense en la relación entre casos favorables y casos posibles cuando los parámetros son números reales.)

La matriz $\sigma^2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}$ es definida positiva, y por tanto el estimador restringido es más eficiente que el estimador MCO sin restringir. Esto no contradice el Tª de Gauss-Markov (en la página˜35 del Tema 1) por dos razones:

- el conjunto de información empleado en una y otra estimación es distinto (más información en el caso de estimación restringida lleva a una mayor eficiencia, es decir mayor precisión (o menor incertidumbre)).

- el Tª de Gauss-Markov se refiere a estimadores insesgados; y, en la práctica, la probabilidad de que $\widehat{\boldsymbol{\beta}^*}$ sea un estimador insesgado es cero (véase el ejercicio propuesto más arriba), salvo que se conozcan los verdaderos valores $\boldsymbol{\beta}$ y se imponga una restricción que realmente cumplan dichos parámetros; algo que jamás ocurre en los estudios o análisis de problemas reales.

5.2. Contraste de la F mediante sumas residuales

Hay una expresión alternativa del estadístico $\mathcal{F}$ empleado para la contrastación de la hipótesis nula $H_0 : \mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$. De (5.4) deducimos que

$$\widehat{\boldsymbol{e}}^{*\prime}\,\widehat{\boldsymbol{e}}^* - \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}} = \left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right)' \mathbf{x}'\,\mathbf{x}\left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right),$$

y por (5.1) sabemos que $\left(\widehat{\boldsymbol{\beta}^*} - \widehat{\boldsymbol{\beta}}\right) = (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\boldsymbol{r} - \mathbf{R}\widehat{\boldsymbol{\beta}}\right)$; así pues, tenemos que

$$\widehat{\boldsymbol{e}}^{*\prime}\,\widehat{\boldsymbol{e}}^* - \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}} = \left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)$$

Sustituyendo $\widehat{\boldsymbol{e}}^{*\prime}\,\widehat{\boldsymbol{e}}^* - \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}}$ por $\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)$ en el numerador del estadístico $\mathcal{F}$ de la Ecuación 3.1 en la página~8; y teniendo en cuenta que $\mathfrak{s}^2_{\widehat{e}} = \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}}\,/(N-k)$, podemos re-escribir el estadístico tal como aparece en la Ecuación (5.5), más abajo:

⇑ **Contraste de la F mediante sumas residuales** **30**

Teniendo en cuenta que $\mathbf{x}'\,\widehat{\boldsymbol{e}} = \mathbf{0}$

$$\widehat{\boldsymbol{e}}^{*\prime}\,\widehat{\boldsymbol{e}}^* - \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}} = \left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)$$

Así pues, de (3.1)

$$\mathcal{F} = \frac{(\widehat{\boldsymbol{e}}^{*\prime}\,\widehat{\boldsymbol{e}}^* - \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}})/r}{\widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}}\,/(N-k)} = \frac{N-k}{r}\cdot\frac{SRC^* - SRC}{SRC} \underset{H_0}{\sim} F_{r,N-k} \tag{5.5}$$

donde $H_0 : \mathbf{R}\boldsymbol{\beta} = \boldsymbol{r}$

Nótese que bajo la hipótesis nula, el numerador debe ser cero, y que cuanto mayor es el numerador (cuanto más difiere de cero), mayor evidencia hay en contra de que la H_0 sea cierta; así pues, este contraste es de una sóla cola, la cola de la derecha.

Ejercicio 14. Sabiendo que $R^2 = 1 - \frac{SRC}{STC}$ (i.e., el caso general) demuestre que el estadístico (5.5) se puede expresar como

$$\mathcal{F} = \frac{N-k}{r}\cdot\frac{R^2 - R^{2*}}{1 - R^2}$$

Pista. En la estimación restringida $SRC^* = (1 - R^{2*})\cdot STC$ y en la estimación sin restringir $SRC = (1 - R^2)\cdot STC$, donde STC es común a ambas (¿por qué?).

Contraste de significación global de modelos con cte. En este contraste, las hipótesis son:

$$H_0 : \quad \beta_2 = \cdots = \beta_k = 0 \quad \text{frente a} \quad H_1 : \quad \text{al menos un } \beta_j \neq 0,$$

es decir, que bajo H_0 el modelo es $\boldsymbol{Y} = \beta_1 + \boldsymbol{U}$

Recuérdese que en el Modelo con sólo una cte. como variable explicativa (**Modelo 1**, pág. 14 del Tema 1) resulta que $\widehat{\boldsymbol{y}} = \overline{\boldsymbol{y}}$, y por tanto, $SEC = 0$. Además, por ser un **modelo con término cte.** (pág. 23 del Tema 1) resulta que $SRC = STC - SEC$;. y por tanto, $R^2 = \frac{SEC}{STC}$.

En el contraste de significación global el modelo restringido resulta ser

$$\boldsymbol{Y} = \beta_1 + \boldsymbol{U},$$

ya que se contrasta que todos los coeficientes de los regresores no constantes son simultáneamente nulos; y el modelo sin restringir es

$$\boldsymbol{Y} = \beta_1 + \beta_2 \boldsymbol{X}_{\blacktriangledown 2} + \cdots + \beta_k \boldsymbol{X}_{\blacktriangledown k} \boldsymbol{U}$$

Como en este caso especial el modelo tiene término constante, podemos re-escribir (5.5) como:

$$\mathcal{F} = \frac{N-k}{r} \cdot \frac{(STC - SEC^*) - (STC - SEC)}{STC - SEC} \qquad \text{por ser modelo con término cte.}$$

$$= \frac{N-k}{r} \cdot \frac{SEC - SEC^*}{STC - SEC}$$

$$= \frac{N-k}{r} \cdot \frac{SEC}{STC - SEC} \qquad \text{pues en el modelo restringido } SEC^* = 0$$

$$= \frac{N-k}{r} \cdot \frac{R^2}{1-R^2} \underset{H_0}{\sim} F_{k-1,N-k} \qquad \text{dividiendo y multiplicando por } STC$$

⇑ **Contraste de la F mediante coef. de determinación** 31

$$\mathcal{F} = \frac{N-k}{r} \cdot \frac{R^2 - R^{2^*}}{1-R^2} \underset{H_0}{\sim} F_{r,N-k}$$

Contraste de significación global de modelos con cte.

$$\mathcal{F} = \frac{N-k}{k-1} \cdot \frac{R^2}{1-R^2} \underset{H_0}{\sim} F_{k-1,N-k} \qquad \text{(caso especial)}$$

A resolver en clase

Ejercicio 15. [(**Ejemplo 4.6 de Novales, 1993, pp. 137**)] Consideremos el modelo

$$Y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 x_{t3} + U_t.$$

donde las variables se hayan en diferencias respecto a la media, y las sumas de los productos cruzados calculados a partir de 103 observaciones muestrales son:

	y	x_1	x_2	x_3
y	10	3	−3	−2
x_1	3	10	4	−2
x_2	−3	4	8	0
x_3	−2	−2	0	6

Contrastar la hipótesis lineal H_0: $\beta_1 + \beta_2 = 0$ frente a H_1: $\beta_1 + \beta_2 \neq 0$.

5.2.1. Test de Chow

- Novales (1993, Sección 4.10, pps. 139–140)
- Novales (1997, Sección 15.3.4, pps. 572)
- Wooldridge (2006, Sección 7.4 y Sección 13.1)

⇑ **Contrastes de cambio estructural: *Test de Chow*** 32

H_0: parámetros no varian en la muestra (No cambio estructural)
H_1: σ^2 cte., pero betas toman dos conjuntos de valores.
Modelo sin restringir

$$\begin{aligned} Y_n &= \mathbf{x}\boldsymbol{\beta}_A + U_n \quad n = 1,\ldots,p \\ Y_n &= \mathbf{x}\boldsymbol{\beta}_B + U_n \quad n = p+1,\ldots,N \end{aligned},$$

Modelo restringido H_0 : $\boldsymbol{\beta}_A = \boldsymbol{\beta}_B$, es decir,

$$Y_n = \mathbf{x}\boldsymbol{\beta} + U_n \quad n = 1,\ldots,N \;,$$

$U_n \sim N(0,\sigma^2)$ para $n = 1,\ldots,N$ en ambos modelos.

⇑ Contrastes de cambio estructural: *Test de Chow* 33

Modelo sin restringir $2k$ parámetros estimados ($\boldsymbol{\beta}_A$, $\boldsymbol{\beta}_B$);
y además $\text{SRC} = \text{SRC}_A + \text{SRC}_B$.
Modelo restringido k restricciones lineales ($\beta_{A_j} = \beta_{B_j}$; $j = 1 \dots k$)
Por lo tanto,

$$\begin{aligned} \mathcal{F} &= \tfrac{N-2k}{k} \tfrac{\text{SRC}^* - \text{SRC}}{\text{SRC}} \\ &= \tfrac{N-2k}{k} \tfrac{\text{SRC}^* - (\text{SRC}_A + \text{SRC}_B)}{(\text{SRC}_A + \text{SRC}_B)} \end{aligned}$$

6. Previsión

- Novales (1993, Sección 4.12, pps. 146–153)
- Wooldridge (2006, Sección 6.4)

⇑ Previsión 34

Sea $\underset{[N\times 1]}{\boldsymbol{Y}} = \underset{[N\times k]}{\mathbf{X}} \underset{[k\times 1]}{\boldsymbol{\beta}} + \underset{[N\times 1]}{\boldsymbol{U}}$ bajo los supuestos 1 a 4; y $\widehat{\boldsymbol{\beta}}$; $\widehat{\mathfrak{s}^2}$ (Con N obs.)

Nueva obs. $\{\boldsymbol{x}_{p\triangleright}\}$;
(donde $p > N$ y donde $\{U_p, \boldsymbol{X}_{p\triangleright}\}$ ortogonales a las anteriores)

Como predictor de $Y_p \longrightarrow$ fun. lineal de los nuevos datos

$$\widehat{y_p} = \boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}} = x_{p1}\widehat{\beta_1} + \cdots + x_{pk}\widehat{\beta_k}.$$

cuyo valor esperado verifica

$$\mathrm{E}\left(\boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}} \,\middle|\, \mathbf{x}\right) = \boldsymbol{x}_{p\triangleright}\mathrm{E}\left(\widehat{\boldsymbol{\beta}} \,\middle|\, \mathbf{x}\right) = \boldsymbol{x}_{p\triangleright}\boldsymbol{\beta} = \mathrm{E}(Y_p \mid \boldsymbol{x}_{p\triangleright})$$

Sea el MLG

$$\underset{[N\times 1]}{\boldsymbol{Y}} = \underset{[N\times k]}{\mathbf{X}} \underset{[k\times 1]}{\boldsymbol{\beta}} + \underset{[N\times 1]}{\boldsymbol{U}}$$

bajo los supuestos 1 a 4. Sea $\widehat{\boldsymbol{\beta}}$ el estimador MCO con la muestra $\{Y_n,\ \boldsymbol{X}_{n\triangleright}\}_{n=1}^{N}$. (es decir, sólo hasta N)

Supongamos que añadimos un nuevo dato $\{U_p, \boldsymbol{X}_{p\triangleright}\}$; donde $p > N$ (previsión fuera de la muestra). Supongamos además que las variables $\{U_p, \boldsymbol{X}_{p\triangleright}\}$ son ortogonales a $\mathbf{U}$ y $\mathbf{X}$, y que $\mathrm{E}(U_p \mid \boldsymbol{x}_{p\triangleright}) = 0$ y $\mathrm{Var}(U_p \mid \boldsymbol{x}_{p\triangleright}) = \sigma^2$.

Queremos prever Y_p conocido el vector de regresores $\boldsymbol{x}_{p\triangleright}$. Además queremos que el predictor sea una función lineal del vector de nuevos datos $\boldsymbol{x}_{p\triangleright}$. Como predictor de Y_p condicionado a $\boldsymbol{x}_{p\triangleright}$ (en realidad también condicionado a toda la matriz $\mathbf{x}$) empleamos

$$\widehat{y_p} = \boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}} = x_{p1}\widehat{\beta_1} + x_{p2}\widehat{\beta_2} + \cdots + x_{pk}\widehat{\beta_k}.$$

es decir, la combinación lineal de los datos $\boldsymbol{x}_{p\triangleright}$ que viene determinada por los estimadores de los betas (con las N primeras obs. [3]). Y cuya esperanza es $\mathrm{E}\left(\boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}} \,\middle|\, \mathbf{x}\right) = \boldsymbol{x}_{p\triangleright}\mathrm{E}\left(\widehat{\boldsymbol{\beta}} \,\middle|\, \mathbf{x}\right) = \boldsymbol{x}_{p\triangleright}\boldsymbol{\beta}$ que coincide con el valor esperado de Y_p, ya que

$$\mathrm{E}(Y_p \mid \boldsymbol{x}_{p\triangleright}) = \mathrm{E}(\boldsymbol{X}_{p\triangleright}\boldsymbol{\beta} + U_p \mid \boldsymbol{x}_{p\triangleright}) = \mathrm{E}(\boldsymbol{X}_{p\triangleright}\boldsymbol{\beta} \mid \boldsymbol{x}_{p\triangleright}) + \mathrm{E}(U_p \mid \boldsymbol{x}_{p\triangleright}) = \boldsymbol{x}_{p\triangleright}\boldsymbol{\beta}$$

Así, $\widehat{y_p}\big|\boldsymbol{x}_{p\triangleright}$ es un predictor condicionado e insesgado de la esperanza de Y_p condicionada a $\boldsymbol{x}_{p\triangleright}$.
Por tanto, dada la nueva información $\boldsymbol{x}_{p\triangleright}$, como predicción de Y_p, emplearemos la estimación puntual de la esperanza de Y_p condicionada a dicha información.

[3] Nótese que sólo podemos estimar por MCO con las N primeras observaciones, ya que, aunque disponemos de $N+1$ observaciones de los regresores, sólo tenemos N observaciones del regresando Y

Previsión
35

Prever ≡ Inferir la distribución de
$$Y_p = \boldsymbol{x}_{p\triangleright}\boldsymbol{\beta} + U_p,$$
para algún $\boldsymbol{x}_{p\triangleright}$ dado; y donde $U_p \sim \mathrm{N}\left(0\,,\,\sigma^2\right)$.

- $Y_p \sim \mathrm{N}\left(\boldsymbol{x}_{p\triangleright}\boldsymbol{\beta}\,,\,\sigma^2\right)$, donde $\boldsymbol{\beta}$ y σ^2 son desconocidos.
- Para estimar $\mathrm{E}(Y_p)$ se utiliza el predictor $\widehat{y_p} \equiv \boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}}$
- Para estimar $\mathrm{Var}(Y_p)$ se utiliza el estimador $\widehat{\mathfrak{s}^2}$
- $\widehat{y_p}$ es un *predictor puntual* de Y_p
  (y también un estimador puntual de $\mathrm{E}(Y_p \mid \boldsymbol{x}_{p\triangleright})$).

Error de previsión: momentos
36

$$\widehat{e_p} = Y_p - \widehat{y_p}. = \boldsymbol{x}_{p\triangleright}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right) + U_p$$

Tomando esperanzas $\quad \mathrm{E}\left(\widehat{e_p} \mid \mathbf{x}^{\diamond}\right) = 0$

donde $\boldsymbol{U}^{\diamond} = \begin{bmatrix}\boldsymbol{U}\\ U_p\end{bmatrix}$ y $\mathbf{X}^{\diamond} = \begin{bmatrix}\mathbf{X}\\ \boldsymbol{X}_{p\triangleright}\end{bmatrix}$

Y la varianza
$$\begin{aligned}\mathrm{Var}\left(\widehat{e_p} \mid \mathbf{x}^{\diamond}\right) =& \mathrm{Var}(U_p \mid \boldsymbol{x}_{p\triangleright}) + \boldsymbol{x}_{p\triangleright}\mathrm{Var}\left(\widehat{\boldsymbol{\beta}} \,\middle|\, \mathbf{x}\right)\boldsymbol{x'}_{p\triangleright}\\ =& \left[\sigma^2 + \sigma^2\boldsymbol{x}_{p\triangleright}(\mathbf{x'}\,\mathbf{x})^{-1}\boldsymbol{x'}_{p\triangleright}\right]\end{aligned}$$

Eficiencia: ¿y si empleamos un estimador lineal e insesgado $\widetilde{\boldsymbol{\beta}}$?

Tomando esperanzas
$$\mathrm{E}\left(\widehat{e_p} \mid \mathbf{x}^{\diamond}\right) = \boldsymbol{x}_{p\triangleright}\left(\boldsymbol{\beta} - \mathrm{E}\left(\widehat{\boldsymbol{\beta}} \,\middle|\, \mathbf{x}\right)\right) + \mathrm{E}(U_p \mid \boldsymbol{x}_{p\triangleright}) = 0$$
donde $\boldsymbol{U}^{\diamond} = \begin{bmatrix}\boldsymbol{U}\\ U_p\end{bmatrix}$ y $\mathbf{X}^{\diamond} = \begin{bmatrix}\mathbf{X}\\ \boldsymbol{X}_{p\triangleright}\end{bmatrix}$

Y la varianza
$$\begin{aligned}\mathrm{Var}\left(\widehat{e_p} \mid \mathbf{x}^{\diamond}\right) =& \mathrm{E}\left(\left[\boldsymbol{x}_{p\triangleright}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right) + U_p\right]\left[\boldsymbol{x}_{p\triangleright}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right) + U_p\right]' \,\middle|\, \mathbf{x}^{\diamond}\right)\\ =& \boldsymbol{x}_{p\triangleright}\mathrm{E}\left(\left[\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right]\left[\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right]' \,\middle|\, \mathbf{x}\right)\boldsymbol{x'}_{p\triangleright} + \mathrm{E}\left({U_p}^2 \mid \boldsymbol{x}_{p\triangleright}\right) + 2\mathrm{E}\left(\boldsymbol{x}_{p\triangleright}\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right)U_p \,\middle|\, \mathbf{x}^{\diamond}\right)\\ =& \boldsymbol{x}_{p\triangleright}\mathrm{Var}\left(\widehat{\boldsymbol{\beta}} \,\middle|\, \mathbf{x}\right)\boldsymbol{x'}_{p\triangleright} + \mathrm{Var}(U_p \mid \boldsymbol{x}_{p\triangleright})\\ =& \left[\sigma^2\boldsymbol{x}_{p\triangleright}(\mathbf{x'}\,\mathbf{x})^{-1}\boldsymbol{x'}_{p\triangleright} + \sigma^2\right]\end{aligned}$$
puesto que $\mathrm{E}(U_p \mid \boldsymbol{x}_{p\triangleright}) = 0$; y que $\mathrm{E}\left(\left[\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right] \,\middle|\, \mathbf{x}\right) = \mathbf{0}$; además de que U_p es ortogonal a $\left[\boldsymbol{U}, \mathbf{X}\right]$, y por lo tanto también es ortogonal a $\widehat{\boldsymbol{\beta}}$, por ser función lineal de $\left[\boldsymbol{U}, \mathbf{X}\right]$; y entonces
$$2\boldsymbol{x}_{p\triangleright}\cdot\mathrm{E}\left(\left(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}\right)U_p \,\middle|\, \mathbf{x}^{\diamond}\right) = 0.$$

Error de previsión: distribución
37

Bajo normalidad $\widehat{e_p} \sim \mathrm{N}\left(0\,,\,\sigma^2\left[1 + \boldsymbol{x}_{p\triangleright}(\mathbf{x'}\,\mathbf{x})^{-1}\boldsymbol{x'}_{p\triangleright}\right]\right)$ y
$$\frac{\widehat{e_p}}{\mathrm{Dt}\left(\widehat{e_p}\right)} = \frac{Y_p - \widehat{y_p}}{\sqrt{\sigma^2 + \sigma^2\boldsymbol{x}_{p\triangleright}(\mathbf{x'}\,\mathbf{x})^{-1}\boldsymbol{x'}_{p\triangleright}}} \sim \mathrm{N}\left(0\,,\,1\right)$$
Así
$$\frac{Y_p - \widehat{y_p}}{\sqrt{\widehat{\mathfrak{s}^2} + \widehat{\mathfrak{s}^2}\boldsymbol{x}_{p\triangleright}(\mathbf{x'}\,\mathbf{x})^{-1}\boldsymbol{x'}_{p\triangleright}}} \sim t_{N-k}$$

Previsión: Estimación intervalos de confianza y otras probab. 38

$$\mathrm{IC}^{\widehat{y_p}}_{1-\alpha}(\mathbf{w}) = \left[\boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}} \pm t_{N-k;\,\alpha/2} \cdot \widehat{\mathrm{Dt}}\left(\widehat{e_p}\right)\right]$$

$$P\left(Y_p \geq h\right) = P\left(\mathcal{T} \geq \frac{h-\widehat{y_p}}{\widehat{\mathrm{Dt}}\left(\widehat{e_p}\right)}\right); \qquad \text{donde } \mathcal{T} \sim t_{N-k}$$

Bajo normalidad

$$\left(\mathrm{E}(Y_p \mid \boldsymbol{x}_{p\triangleright}) - \widehat{\mathrm{E}(y_p)}\right) = \left(\boldsymbol{x}_{p\triangleright}\boldsymbol{\beta} - \boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}}\right) \sim \mathrm{N}\left(0\,,\,\sigma^2\boldsymbol{x}_{p\triangleright}(\mathbf{x}'\,\mathbf{x})^{-1}\boldsymbol{x}'_{p\triangleright}\right)$$

$$\mathrm{IC}^{\widehat{\mathrm{E}(y_p)}}_{1-\alpha}(\mathbf{w}) = \left[\boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}} \pm t_{N-k;\,\alpha/2} \cdot \widehat{\mathrm{Dt}}\left(\boldsymbol{x}_{p\triangleright}\widehat{\boldsymbol{\beta}}\right)\right]$$

EJERCICIO 16. Resuelva el ejercicio propuesto nº 4 del profesor José Alberto Mauricio. http://www.ucm.es/info/ecocuan/ectr1/index.html#Material.

7. Ejercicios

EJERCICIO 17. Sea Y la creación de empleo trimestral de las 17 Comunidades Autónomas (CCAA) españolas (excluidas Ceuta y Melilla) y X un indicador que mide el crecimiento en el nivel de actividad de la comunidad a lo largo del último año. Queremos estimar un modelo de regresión lineal simple que relacione Y en funcion de X. La media muestral de Y es 0.04, la de X es 0.03; sus desviaciones típicas son 0.2 y 0.15, respectivamente, y la covarianza es 0.01.

(a) Presente el modelo de regresión lineal y estime por MCO sus parámetros.

(b) La siguiente tabla muestra los residuos MCO de la estimación anterior correspondientes a cada una de las 17 comunidades autónomas (CCAA)

CCAA Nº	residuo	CCAA Nº	residuo	CCAA Nº	residuo
1	0.79	7	0.24	13	-0.68
2	1.00	8	0.67	14	1.10
3	0.84	9	0.45	15	0.61
4	-0.51	10	-0.81	16	-0.60
5	-0.33	11	1.54	17	-2.40
6	-1.33	12	-0.60		

¿Qué interpretación tienen los residuos en el marco de este modelo? En función de dicha interpretación, comente los aspectos más destacables de esta tabla de residuos con relación a cómo ha sido la creación de empleo en las diferentes CCAA en el trimestre.

(c) Suponga que desea llevar a cabo el contraste de significación individual de la pendiente del modelo de regresión. Indique qué hipótesis alternativa es la más apropiada para este contraste. ¿Qué estadístico de contraste y qué región crítica son más apropiados?.

EJERCICIO 18. Sean los siguientes datos:

Familia	y_n	x_n	$x_n y_n$	x_n^2	$\widehat{Y_n}$	$\widehat{U_n}$	$x\widehat{U_n}$	$\widehat{U_n}^2$
A	4	2	8	4	4.50	-0.50	-1	0.25
B	7	3	21	9	6.25	0.75	2.25	0.5625
C	3	1	3	1	2.75	0.25	0.25	0.0625
D	9	5	45	25	9.75	-0.75	-3.75	0.5625
E	17	9	153	81	16.75	0.25	2.25	0.0625
sumas	40	20	230	120	40	0	0	1.5

Cuadro 1:

donde Y es el gasto semanal de las familias, y X es su ingreso semanal.

Además se sabe que las varianzas y covarianzas muestrales son:

$$N \cdot s_y^2 = 124, \quad N \cdot s_x^2 = 40, \quad N \cdot s_{xy} = 70,$$

donde N es el tamaño muestral.

Suponga que plantea el siguiente modelo

$$Y_n = a + bx_n + U_n,$$

donde U_n son otros factores que afectan al consumo familiar distintos de sus ingresos. Se sabe que la distribución conjunta de dichos factores es:

$$\mathbf{U} \sim N(\mathbf{0},\ \sigma^2\,\mathbf{I}),$$

donde $\mathbf{I}$ es una matriz identidad de orden 5, y σ^2 es la varianza de U_n, cuyo valor es desconocido.

(a) Estime por MCO los parámetros del modelo.

(b) Estime por MCO los parámetros del modelo pero dando un intervalo de confianza del 95 %.

(c) Contraste la hipótesis de que "la propensión marginal es igual a uno" frente a "es menor que uno" con un nivel de significación del 10 %. ¿Cuál es el p-valor de la estimación del "de la propensión marginal estimada"?

(d) ¿Cuál es la familia que más ha gastado respecto al nivel esperado, dados su ingresos? ¿y la que menos?

(e) ¿Cuál es el gasto esperado para una familia que tuviera unos ingresos semanales iguales a 4?

8. Bibliografía

Gujarati, D. N. (2003). *Basic Econometrics*. McGraw-Hill, cuarta ed. ISBN 0-07-112342-3. International edition. 14

Hayashi, F. (2000). *Econometrics*. Princeton University Press, Princeton, New Jersey. ISBN 0-691-01018-8. 2, 8

Mittelhammer, R. C. (1996). *Mathematical Statistics for Economics and Business*. Springer-Verlag, New York, primera ed. ISBN 0-387-94587-3. 8

Novales, A. (1993). *Econometría*. McGraw-Hill, segunda ed. 2, 10, 15, 17, 18

Novales, A. (1997). *Estadística y Econometría*. McGraw-Hill, Madrid, primera ed. ISBN 84-481-0798-5. 17

Verbeek, M. (2004). *A Guide to Modern Econometrics*. John Wiley & Sons, Inc., segunda ed. 2

Wooldridge, J. M. (2006). *Introducción a la econometría. Un enfoque moderno*. Thomson Learning, Inc., segunda ed. 2, 17, 18

9. Trasparencias

Lista de Trasparencias

A. Desarrollo casos generales

Regiones de confianza para conjuntos para varias combinaciones lineales de parámetros: El caso general, en el que $\underset{[r\times k]}{\mathbf{R}}$ es una matriz cualquiera de rango completo k; de 3.1 en la página˜8 se deduce que

$$P\left(\frac{\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\boldsymbol{r}\right)'\left[\mathbf{R}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\boldsymbol{r}\right)/r}{\widehat{\mathfrak{s}^2}}<F_{r,\;N-k,\;\alpha}\right)=$$

$$P\left(\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\boldsymbol{r}\right)'\left[\mathbf{R}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\boldsymbol{r}\right)<r\widehat{\mathfrak{s}^2}\,F_{r,\;N-k,\;\alpha}\right)=1-\alpha.$$

Y por tanto la región de confianza debe verificar

$$P\Big(\quad\mathbf{R}\boldsymbol{\beta}\quad\in\quad\mathrm{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{W})\quad\Big)=1-\alpha$$

donde $\mathrm{IC}_{1-\alpha}^{\mathbf{R}\boldsymbol{\beta}}(\mathbf{w})$ es el conjunto de puntos $\boldsymbol{b}$ pertenecientes $\mathbb{R}^r$ que verifican lo siguiente:

$$\mathrm{IC}_{1-\alpha}^{\mathbf{R}\widehat{\boldsymbol{\beta}}}(\mathbf{w})=\left\{\boldsymbol{b}\;\text{ tales que }\left(\mathbf{R}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{b}\right)\right)'\mathbf{x}'\mathbf{x}\left(\mathbf{R}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{b}\right)\right)<r\widehat{\mathfrak{s}^2}\cdot F_{r,\;N-k,\;\alpha}\right\}$$

que es una elipsoide si $r>1$ (y un intervalo si $r=1$).

Previsión m periodos: Añadir previsión m periodos.

Soluciones a los Ejercicios

Ejercicio 3. Falso: El test de hipótesis sólo puede indicar que con los datos observados no rechazamos la hipótesis nula H_0 con una probabilidad φ, es decir, asumiendo el riesgo de que, con probabilidad φ, no rechacemos H_0 aún siendo falsa (error tipo II).

Ejercicio 3

Ejercicio 4. Verdadero: la región crítica asociada al nivel de significación del 10 % contiene a la del 5 %; por lo tanto si se ha "caído" en la región del 5 %, entonces también se ha "caído" en la del 10 %.

Ejercicio 4

Ejercicio 5. $H_0 : \beta = 0$, frente a $H_1 : \beta > 0$. Nótese que, por t^{a} económica, lo más razonable es proponer un contraste de cola superior. La región crítica más conveniente es por tanto: $RC = \left\{\hat{\beta} > k\right\} = \left\{\frac{\hat{\beta}-\beta}{dt(\hat{\beta})} > t_{1-\alpha}\right\}$, donde $\hat{\beta}$ es el estimador MCO y $t_{33,\ 1-\alpha}$ es el valor tabulado de una t-Student con 33 grados de libertad que deja a la izquierda un área igual a $1-\alpha$.

Ejercicio 5

Ejercicio 7(a) $\mathbf{R}\boldsymbol{\beta} = [0,\ 3,\ -1,\ 0]\begin{bmatrix}\beta_1 & \beta_2 & \beta_3 & \beta_4\end{bmatrix}' = 7;\quad r=1;\quad$ k=4; y $t_{21,\ 0.975}$ =2.08 (dos colas $\alpha = 0.05$)

Por una parte $\widehat{s^2_{e}} = \frac{\text{SRC}}{N-k} = 40/21$; por otra parte

$$\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}' = \begin{bmatrix}0, & 3, & -1, & 0\end{bmatrix}\begin{bmatrix}v_{11} & v_{12} & v_{13} & v_{14}\\ v_{12} & v_{22} & v_{23} & v_{24}\\ v_{13} & v_{23} & v_{33} & v_{34}\\ v_{14} & v_{24} & v_{34} & v_{44}\end{bmatrix}\begin{bmatrix}0\\3\\-1\\0\end{bmatrix}$$
$$=9v_{22} - 6v_{23} + v_{33} \;=\; 9\cdot 4 - 6\cdot 2 + 6 = 30$$

Así pues, puesto que $\frac{\mathbf{R}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})}{\widehat{\text{Dt}}(\mathbf{R}\,\widehat{\boldsymbol{\beta}})} \underset{H_0}{\sim} t_{N-k}$ (de la Nota 3 en la página˜9)

$$\widehat{\mathcal{T}} = \frac{\mathbf{R}\,\widehat{\boldsymbol{\beta}} - c}{\widehat{\text{Dt}}\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}}\right)} = \frac{3\widehat{\beta_2} - \widehat{\beta_3} - 7}{\sqrt{\frac{40}{21}\cdot 30}} = 0.198 < t_{N-k,\ 0.975} = 2.08$$

El nivel crítico p (*p-value*): t(21): área a la derecha de 0.198 = 0.422475
(valor a dos colas = 0.844949; complemento = 0.155051).

□

Ejercicio 7(b) Ahora, $r = 2$, $F_{2,\ 21,\ 0.95} = 3.47$ ($\alpha = 0.05$) y

$$\mathbf{R} = \begin{bmatrix}0 & 3 & -1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \qquad \boldsymbol{r} = \begin{bmatrix}7\\0\end{bmatrix}$$

El estadístico a utilizar es el ratio-F

$$\mathcal{F} = \frac{\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)/r}{\widehat{s^2}};$$

Los resultados parciales son

$$\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}' = \begin{bmatrix}30 & -50\\ -50 & 10\end{bmatrix}$$
$$\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1} = \begin{bmatrix}-0.0045455 & -0.0227273\\ -0.0227273 & -0.0136364\end{bmatrix}$$
$$\mathbf{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r} = \begin{bmatrix}1.5\\ -1.5\end{bmatrix}$$

Por tanto

$$\widehat{\mathcal{F}} = \frac{\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\,\widehat{\boldsymbol{\beta}} - \boldsymbol{r}\right)/r}{\widehat{s^2}}$$
$$= \begin{bmatrix}1.5 & -1.5\end{bmatrix}\begin{bmatrix}-0.0045455 & -0.0227273\\ -0.0227273 & -0.0136364\end{bmatrix}\begin{bmatrix}1.5\\ -1.5\end{bmatrix}\cdot\frac{1}{2}\cdot\frac{21}{40} = 0.016108$$

□

Ejercicio 12. La insesgadez del estimador es inmediata, puesto que el último sumando de la Ecuación 5.1 en la página˜14 es cero cuando la restricción es cierta.

Veamos la segunda. En la Ecuación (5.2) podemos ver como el estimador restringido es la suma del estimador MCO, menos una función lineal de dicho estimador y mas un término constante; por ello la varianza será la varianza del estimador MCO, mas la varianza de la función del estimador menos dos veces la covarianza del estimador con la transformación lineal de dicho estimador (por tanto una cte. por la varianza del estimador MCO). Así pues: $\operatorname{Var}\left(\widehat{\boldsymbol{\beta}^*}\,\middle|\,\mathbf{x}\right)$

$$\begin{aligned}
\operatorname{Var}\left(\widehat{\boldsymbol{\beta}^*}\,\middle|\,\mathbf{x}\right) =& \operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right) + \operatorname{Var}\left((\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}\,\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right)\\
&- 2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right)\\
=& \operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right) + (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right)\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\\
&- 2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right)
\end{aligned}$$

y sustituyendo $\operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right)$ por su valor $\sigma^2(\mathbf{x}'\,\mathbf{x})^{-1}$

$$\begin{aligned}
=& \operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right) + (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\sigma^2(\mathbf{x}'\,\mathbf{x})^{-1})\,\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\\
&- 2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\sigma^2(\mathbf{x}'\,\mathbf{x})^{-1})
\end{aligned}$$

y puesto que $\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}' = \mathbf{I}$

$$\begin{aligned}
=& \operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right) + \sigma^2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\\
&- 2\sigma^2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\\
=& \operatorname{Var}\left(\widehat{\boldsymbol{\beta}}\,\middle|\,\mathbf{x}\right) - \sigma^2(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}
\end{aligned}$$

Ejercicio 12

Ejercicio 13. La probabilidad es cero si no conocemos los verdaderos valores de $\boldsymbol{\beta}$.

Ejercicio 13

Ejercicio 14. STC es comun, puesto que la variable dependiente $\boldsymbol{y}$ es la misma sea el modelo restringido o no. Por tanto

$$\frac{SRC^* - SRC}{SRC} = \frac{(1-R^{2^*})STC - (1-R^2)STC}{(1-R^2)STC} = \frac{R^2 - R^{2^*}}{1-R^2}$$

Ejercicio 14

Ejercicio 15. Vamos a emplear el estadístico $\mathcal{F}$ expresado mediante sumas residuales (5.5) para contrastar la hipótesis.

Para ello estimamos primero por MCO:

Por una parte tenemos $(\mathbf{x}'\,\mathbf{x})^{-1} = \begin{bmatrix} 10 & 4 & -2 \\ 4 & 8 & 0 \\ -2 & 0 & 6 \end{bmatrix}^{-1}$; y por otra $\mathbf{x}'\,\boldsymbol{y} = \begin{bmatrix} 3 \\ -3 \\ -2 \end{bmatrix}$.

Así pues,

$$\widehat{\boldsymbol{\beta}} = (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{x}'\,\boldsymbol{y} = \begin{bmatrix} \frac{23}{44} \\ \\ \frac{-28}{44} \\ \\ \frac{-7}{44} \end{bmatrix};$$

y puesto que SRC $= \widehat{\boldsymbol{e}}'\,\widehat{\boldsymbol{e}} = \boldsymbol{y}'\,\boldsymbol{y} - \widehat{\boldsymbol{\beta}'}\,\mathbf{x}'\,\boldsymbol{y}$ (de T$^{\text{a}}$ de Pitágoras y $\widehat{\boldsymbol{y}}'\,\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{y}}'\,\boldsymbol{y}$), entonces

$$\text{SRC} = 10 - \begin{bmatrix} \frac{23}{44} & \frac{-28}{44} & \frac{-7}{44} \end{bmatrix} \begin{bmatrix} 3 \\ -3 \\ -2 \end{bmatrix} = 10 - \left(\tfrac{23}{44}3 + \tfrac{28}{44}3 + \tfrac{7}{44}2\right) = 6.20$$

Ahora estimamos por MCR:

$$\widehat{\boldsymbol{\beta^*}} = \widehat{\boldsymbol{\beta}} + (\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{x}'\,\mathbf{x})^{-1}\mathbf{R}'\right]^{-1}\left(\boldsymbol{r} - \mathbf{R}\widehat{\boldsymbol{\beta}}\right) = \begin{bmatrix} \frac{4}{7} \\ \frac{-4}{7} \\ \frac{-1}{7} \end{bmatrix}$$

donde $\mathbf{R} = (1, 1, 0)$ y $\boldsymbol{r} = 0$.

De nuevo $\text{SRC}^* = \boldsymbol{y}'\,\boldsymbol{y} - \widehat{\boldsymbol{\beta^*}}'\,\mathbf{x}'\,\boldsymbol{y}$, y por tanto

$$\text{SRC}^* = 10 - \left(\tfrac{4}{7}3 + \tfrac{4}{7}3 + \tfrac{1}{7}2\right) = 6.28$$

Empleando (5.5) obtenemos:

$$\widehat{\mathcal{F}} = \frac{N-k}{r}\frac{\text{SRC}^* - \text{SRC}}{\text{SRC}}$$

$$= \frac{100}{1}\frac{6.28 - 6.20}{6.20} = 1.29 < F_{1,100,\ 0.95} = 3.94$$

Hay una segunda forma de obtener SRC^* estimando por MCO el modelo con las restricciones incorporadas: Si sustituimos la restricción $(\beta_1 + \beta_2 = 0)$ en el modelo tenemos:

$$Y_t = \beta_1(x_{t1} - x_{t2}) + \beta_3 x_{t3} + U_t = \beta_1 x^*_{t1} + \beta_3 x_{t3} + U_t,$$

donde $x^*_{t1} = x_{t1} - x_{t2}$.

Ahora

$$\mathbf{x^*}'\,\mathbf{x^*} = \begin{pmatrix} 10 & -2 \\ -2 & 6 \end{pmatrix} \quad \text{y} \quad \mathbf{x^*}'\,\boldsymbol{y} = \begin{pmatrix} 6 \\ -2 \end{pmatrix}$$

puesto que

$$\begin{aligned} \sum (x^*_{t1})^2 &= \sum (x_{t1})^2 + \sum (x_{t2})^2 - 2\sum (x_{t1}x_{t2}) && = 10 + 8 - 2\cdot 4 = 10 \\ \sum (x^*_{t1}x_{t3}) &= \sum (x_{t1}x_{t3}) - \sum (x_{t2}x_{t3}) && = -2 - 0 = -2 \\ \sum (x^*_{t1}y_t) &= \sum (x_{t1}y_t) - \sum (x_{t2}y_t) && = 3 - (-3) = 6 \end{aligned}$$

Así pues, $\widehat{\boldsymbol{\beta^*}} = (\mathbf{x^*}'\,\mathbf{x^*})^{-1}\mathbf{x^*}'\,\boldsymbol{y} = \begin{pmatrix} 4/7 \\ -1/7 \end{pmatrix}$.

De nuevo $\text{SRC}^* = \boldsymbol{y}'\,\boldsymbol{y} - \widehat{\boldsymbol{\beta^*}}'\,\mathbf{x}'\,\boldsymbol{y}$, y

$$\text{SRC}^* = 10 - \left(\tfrac{4}{7}6 + \tfrac{1}{7}2\right) = 6.28;$$

que es el mismo resultado que obtuvimos empleando $\widehat{\boldsymbol{\beta^*}}$.

Ejercicio 15

Ejercicio 17(a) El modelo es $Y_n = a + x_n \cdot b + U_n$

Las estimaciones MCO de a y b son:

$$\widehat{b} = \frac{s_{xy}}{s_x^2} = \frac{0.01}{0.15^2} = 0.44; \qquad \widehat{a} = \overline{y} - \overline{x}\cdot\frac{s_{xy}}{s_x^2} = 0.04 - 0.03\cdot 0.44 = 0.027;$$

Las estimación MCO de cada y_n es: $\widehat{y_n} = 0.027 + 0.44 \cdot x_n$

□

Ejercicio 17(b) El residuo n-ésimo es: $\widehat{e_n} = y_n - \widehat{y_n} = y_n - \text{E}(Y_n \mid x_n)$ es decir, es la diferencia entre la creación de empleo observada en la comunidad n-ésima y la creación de empleo *esperada* dado el dato de crecimiento de la actividad de dicha comunidad.

Por tanto, residuos positivos indican mayor creación de empleo de lo esperado (dado el crecimiento de la actividad), y residuos negativos indican menor creación de empleo de lo esperado (dado el crecimiento de la actividad).

Por tanto, la comunidad que parece haber tenido un comportamiento mejor (dada la información disponible) es la comunidad número 11, y la que peor lo ha tenido es la número 17.

□

Ejercicio 17(c) No cabe esperar que crecimientos positivos del nivel de actividad perjudiquen la creación de empleo; por lo tanto el contraste debe ser de una sóla cola. Así pues:

$$H_0 : b = 0; \quad H_1 : b > 0; \quad \text{estadístico: } \frac{\left(\widehat{b} - b\right)}{\sqrt{\frac{\widehat{\mathfrak{s}^2}}{\sum(x_i - \overline{x})^2}}} \underset{H_0}{\sim} t_{15}; \quad RC = \left\{ \mathbf{x} \left| \frac{\widehat{b}}{\sqrt{\frac{\widehat{\mathfrak{s}^2}}{\sum(x_i - \overline{x})^2}}} > k \right. \right\}$$

□

Ejercicio 18(a)

1. Por una parte:
$$\widehat{b} = \frac{\widehat{\sigma_{xy}}}{\widehat{\sigma_x^2}} = \frac{s_{xy}}{s_x^2} = \frac{70}{40} = 1.75$$
por otra, las medias muestrales son
$$\overline{x} = \frac{\sum x_n}{5} = \frac{20}{5} = 4; \qquad \overline{y} = \frac{\sum y_n}{5} = \frac{40}{5} = 8;$$
por lo que
$$\widehat{a} = \overline{y} - \widehat{b} \cdot \overline{x} = 8 - 1.75 \cdot 4 = 1.$$

□

Ejercicio 18(b) El estimador MCO se distribuye Normal con esperanza igual al verdadero valor de los parámetros estimados, y varianza desconocida.

- **(Parámetro a)** Buscamos los valores A y B tales que
$$P\left(A \leq \frac{\left(\widehat{a} - a\right)}{\sqrt{\frac{\widehat{\mathfrak{s}^2} \sum x_n^2}{N \sum(x_n - \overline{x})^2}}} \leq B\right) = (1 - \alpha)$$
Donde $\frac{\left(\widehat{a} - a\right)}{\sqrt{\frac{\widehat{\mathfrak{s}^2} \sum x_n^2}{N \sum(x_n - \overline{x})^2}}}$ se distribuye como una t de Student con N-2 grados de libertad, y $\widehat{\mathfrak{s}^2}$ es el estimador de la cuasi-varianza muestral. Por tanto A y B son los valores que aparecen en las tablas, y que determinan un intervalo centrado en cero con una probabilidad asociada del 95 %; es decir, $A = -3.182$, y $B = 3.182$, y $\widehat{\mathfrak{s}^2} = 1.5/(N-2) = 0.5$. Así pues,
$$P\left(-3.182 \leq \frac{\left(1 - a\right)}{\sqrt{\frac{0.5 \cdot 120}{5 \cdot 40}}} \leq 3.182\right) = 0.95$$
$$P\left(-3.182 \cdot \sqrt{0.3} \leq \left(1 - a\right) \leq 3.182 \cdot \sqrt{0.3}\right) = 0.95$$
$$P\left(1 + 3.182 \cdot \sqrt{0.3} \geq a \geq 1 - 3.182 \cdot \sqrt{0.3}\right) = 0.95,$$
es decir,
$$a \in [-0.7429, \ 2.7429]$$
con probabilidad 0.95
- **(Parámetro b)** Del mismo modo, buscamos los valores A y B tales que
$$P\left(A \leq \frac{\left(\widehat{b} - b\right)}{\sqrt{\frac{\widehat{\mathfrak{s}^2}}{\sum(x_n - \overline{x})^2}}} \leq B\right) = (1 - \alpha)$$
Donde $\frac{\left(\widehat{b} - b\right)}{\sqrt{\frac{\widehat{\mathfrak{s}^2}}{\sum(x_n - \overline{x})^2}}}$ se distribuye como una t de Student con N-2 grados de libertad; por tanto
$$P\left(-3.182 \leq \frac{\left(1.75 - b\right)}{\sqrt{\frac{0.5}{40}}} \leq 3.182\right) = 0.95$$
$$P\left(-3.182 \cdot \sqrt{0.0125} \leq \left(1.75 - b\right) \leq 3.182 \cdot \sqrt{0.0125}\right) = 0.95$$
$$P\left(1.75 + 3.182 \cdot \sqrt{0.0125} \geq b \geq 1.75 - 3.182 \cdot \sqrt{0.0125}\right) = 0.95,$$
es decir,
$$b \in [1.3842, \ 2.0958]$$
con probabilidad 0.95

$\square$

Ejercicio 18(c) Las hipótesis son:

$$H_0: \ b = 1$$
$$H_1: \ b < 1$$

La región critica de una sola cola es

$$RC = \left\{ \mathbf{x} \;\middle|\; \frac{\widehat{b} - 1}{\sqrt{\frac{\widehat{s^2}}{\sum (x_n - \overline{x})^2}}} < k \right\},$$

donde k es el valor de la tablas para una t de Student de tres grados de libertad, ya que el estadístico de la parte izquierda de la desigualdad tiene dicha distribución. Para $\alpha = 0.1$, tenemos que $k = t\,3,\ 0.1 = -1.638$ sustituyendo tenemos que

$$\frac{1.75 - 1}{\sqrt{0.0125}} = 6.7082 > k = t\,3,\ 0.1 = -1.638$$

por lo que no rechazamos H_0.

El p-valor es la probabilidad de

$$\begin{aligned} P\left(\widehat{b} \le 1.75 \;\middle|\; H_0\right) =& P\left(\frac{\widehat{b} - b}{\sqrt{\frac{\widehat{s^2}}{\sum (x_n - \overline{x})^2}}} \le \frac{1.75 - b}{\sqrt{\frac{\widehat{s^2}}{\sum (x_n - \overline{x})^2}}} \;\middle|\; H_0 \right) \\ =& P\left(W \le \frac{1.75 - 1}{\sqrt{0.0125}} = 6.7082 \right) \simeq 1, \end{aligned}$$

donde W se distribuye como una t de Student con tres grados de libertad.

$\square$

Ejercicio 18(d) La familia B es la que más ha gastado (7) respecto al nivel esperado (6.25)

La familia D es la que menos ha gastado (9) respecto al nivel esperado (9.75)

$\square$

Ejercicio 18(e) Según el modelo estimado, una familia con ingresos de 4 debería gastar

$$\widehat{y} = 1 + 1.75 \cdot 4 = 8$$

$\square$