

Dpto. de Economía Cuantitativa  
Universidad Complutense de Madrid  
**Introducción a la Econometría**

Tema 5 — Modelo clásico de regresión lineal

Marcos Bujosa y Gustavo A. Marrero

Material de apoyo para el curso *Introducción a la Econometría* de la licenciatura en Economía de la Universidad Complutense de Madrid.

© 2003–2007 Marcos Bujosa y Gustavo A. Marrero  
Actualizado el: 10 de abril de 2007

Versión 4.0

Copyright © 2003–2007 Marcos Bujosa y Gustavo A. Marrero



Algunos derechos reservados. Esta obra está bajo una licencia Reconocimiento-CompartirIgual de Creative Commons. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-sa/2.5/es/deed.es> o envíe una carta a Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Puede encontrar la última versión de este material en:

<http://www.ucm.es/info/ecocuan/mbb/index.html#ietria>

## Índice

<b>Índice</b>	<b>1</b>
<b>1. Introducción</b>	<b>2</b>
1.1. El punto de vista estadístico: Regresión como descomposición ortogonal . . . . .	2
1.2. El punto de vista del Análisis Económico: Regresión como modelo explicativo . . . . .	3
<b>2. Modelo Clásico de Regresión Lineal</b>	<b>4</b>
2.1. Tres primeros supuestos en el Modelo Clásico de Regresión Lineal . . . . .	4
2.2. Modelo Clásico de Regresión Lineal para Muestras Aleatorias . . . . .	9
2.3. Regresores No Estocásticos . . . . .	9
<b>3. Estimación MCO (Mínimos Cuadrados Ordinarios)</b>	<b>9</b>
<b>4. Propiedades algebraicas de la estimación MCO</b>	<b>12</b>
<b>5. Propiedades estadísticas del estimador MCO <math>\hat{\beta} \mid x</math></b>	<b>12</b>
<b>6. Distribución de los estimadores MCO bajo Normalidad</b>	<b>15</b>
<b>7. Estimación por intervalos</b>	<b>15</b>
<b>8. Contrastación de hipótesis</b>	<b>16</b>
<b>9. Ejercicios</b>	<b>17</b>
<b>10. Bibliografía</b>	<b>18</b>
<b>. Soluciones a los Ejercicios</b>	<b>19</b>

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.

El programa se cubre con los siguientes capítulos de libro de texto [Novales \(1997\)](#)<sup>1</sup>:

**Capítulos 1 a 3:** Estos temas han sido cubiertos en asignaturas anteriores, y debido a su bajo nivel de complejidad no se verán en clase (aunque forman parte del programa).

**Capítulos 4 a 6:** Estos temas han sido cubiertos en las asignaturas [Estadística I](#) y [II](#). Se realizará un breve repaso en clase (una semana o semana y media como máximo), asumiendo que el alumno es capaz de preparar por su cuenta esta parte.

**Capítulos 7 y 8:** completos

**Capítulo 9:** secciones 9.4 a 9.6

**Capítulos 10 y 12:** completos

**Tema 5. Análisis de relaciones. Modelo Lineal Simple** (Tema 13 de EyE) (1 semana y media o menos)

*Capítulo 13 del manual de Alfonso Novales “Estadística y Econometría”*

1. Introducción al análisis de relación.
2. Análisis gráfico. Gráfico x-y. Importancia de las escalas en series temporales.
3. Análisis estadístico de relaciones. La covarianza y el coeficiente de correlación. Ventajas e inconvenientes.
4. Relación matemática Vs. relación Estadística. El modelo econométrico.
5. El modelo de regresión simple.
6. Introducción. ¿Que es un estadístico? ¿Que es un estimador ? ¿Que es una estimación?
7. Propiedades de un estimador.
8. El método de mínimos cuadrados ordinarios. Relación con la aproximación lineal de la esperanza condicionada.
9. Distribución del estimador de  $\beta$ . Estimación puntual e Intervalos de Confianza
10. Interpretación de los residuos de un modelo. La varianza residual.

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>

## 1. Introducción

### 1.1. El punto de vista estadístico: Regresión como descomposición ortogonal

↑	<u>Descomposición ortogonal y causalidad</u>	1
$Y = E(Y \mathcal{D}) + U$		
donde el conjunto de información es $\mathcal{D} : (\mathbf{X} = \mathbf{x})$ ; por tanto		
$Y = E(Y \mathbf{X}) + U$		
donde $E_{\mathbf{y} \mathbf{x}}(Y \mathbf{x})$ es una función arbitraria		
<b>lectura estadística:</b> de izquierda a derecha.		
Siempre es cierta. No implica causalidad ni conclusiones teóricas		
<b>lectura teórica:</b> de derecha a izquierda.		
Interpretación puede ser falsa (regresiones espurias)		

De [Spanos \(1999\)](#), Capítulo 7, en particular la Sección 7.5.3)

Sea  $Y$  una variable aleatoria con segundo momento finito, es decir,  $E(|Y|^2) < \infty$ , y un conjunto de información  $\mathcal{D}$ ; entonces siempre podemos encontrar una descomposición de  $Y$  como la siguiente:

$$Y = E(Y|\mathcal{D}) + U \tag{1.1}$$

donde

$E(Y|\mathcal{D})$ : es el componente sistemático

<sup>1</sup>Otros excelentes manuales en castellano son [Peña \(2001\)](#), [Peña \(2002\)](#) y [Peña y Romo \(1997\)](#).

$U$ : es el componente NO-sistemático

La existencia de dicha descomposición<sup>2</sup> está garantizada siempre que  $E(|Y|^2) < \infty$ .

Ambos componentes de  $Y$  satisfacen las siguientes propiedades

1.  $E_{U|D}(U | \mathcal{D}) = 0$
2.  $E_{U|D}(U^2 | \mathcal{D}) = \text{Var}_{U|D}(Y | \mathcal{D}) < \infty$
3.  $E\left(U \cdot \left[E(Y | \mathcal{D})\right]\right) = 0$  por tanto **ambos componentes son ortogonales**.

Supondremos que disponemos de una sucesión de variables aleatorias  $Y_n$  (para  $n = 1, \dots, N$ ) y de una matriz de variables aleatorias  $\mathbf{X}$ ; y que nuestro conjunto de información  $\mathcal{D}$  es

$$\mathcal{D} : (\mathbf{X} = \mathbf{x})$$

es decir, *el conjunto de variables aleatorias  $\mathbf{X}$  (en total  $N$  variables) ha tomado conjuntamente la matriz de valores  $\mathbf{x}$ .*

Siendo así, la descomposición ortogonal para cada  $Y_n$  queda como sigue:

$$Y_n = E(Y_n | \mathbf{X}) + U_n$$

Nótese que esta es una descomposición *puramente estadística*. Únicamente nos dice que si disponemos de cierta información acerca de las variables  $\mathbf{X}$ , podemos descomponer la variable  $Y_n$  en dos partes. Pero no hay una *teoría económica* detrás; por tanto no dice si hay relaciones de causalidad entre las variables. Podría ocurrir que:

1. bien las variables  $\mathbf{X}$  generaran parcialmente a  $Y$  (y por tanto, al conocer  $\mathcal{D} : (\mathbf{X} = \mathbf{x})$  sabemos qué parte de  $Y$  es debida a  $\mathbf{X}$  y qué parte no)
2. o bien que  $Y$  causa (o genera) las variables  $\mathbf{X}$  (y por tanto, al observar  $\mathcal{D} : (\mathbf{X} = \mathbf{x})$  sabemos qué cabe esperar que ha ocurrido con la variable causante  $Y$ ; como cuando vemos llover por la ventana, y entonces sabemos que hay nubes en el cielo)
3. o bien, que hay alguna otra causa común (y quizá desconocida) que genera conjuntamente tanto a  $Y$  como a  $\mathbf{X}$  (y observar lo que ha ocurrido con  $\mathbf{X}$  (la información  $\mathcal{D}$ ) nos indica que cabe esperar que ha ocurrido con  $Y$  (puesto que tienen un causante común).

La descomposición ortogonal

$$Y_n = E(Y_n | \mathbf{X}) + U_n$$

se lee de **izquierda a derecha** (es decir, “*puedo descomponer  $Y_n$  en las dos partes descritas a la derecha*”), y no hay una *teoría* detrás.

## 1.2. El punto de vista del Análisis Económico: Regresión como modelo explicativo

Como economistas deseamos que la descomposición estadística de más arriba sea reflejo de las relaciones teóricas entre  $\mathbf{X}$  y  $Y$ . En este sentido queremos leer la relación de **derecha a izquierda**, es decir  $Y$  (por ejemplo el consumo) está generado por una función de las variables  $\mathbf{X}$  (por ejemplo una función de la renta) junto a otras causas distintas de la renta ( $U$ ).

Esta visión sugiere algunos de los nombres dados tanto para  $Y$  como para  $\mathbf{X}$ . No obstante (y a pesar de los nombres), no debemos nunca perder de vista que la descomposición ortogonal es una relación estadística que siempre<sup>3</sup> podemos encontrar; pero que en general no permite sacar conclusiones teóricas de ella (regresiones espurias). Sólo en aquellos casos en que las variables situadas a derecha e izquierda provienen de un modelo teórico bien establecido, que nos sugiere qué variables son causantes (y por ello las situamos a derecha) y cuáles son causadas (izquierda) quizá podamos sacar conclusiones. La palabra “quizá”, se debe a que con frecuencia los datos disponibles no miden aquellos conceptos empleados en los modelos teóricos (consumo permanente, preferencias, nivel de precios, utilidades, aversión al riesgo, etc.), o bien a que los modelos no están correctamente especificados (temas que se verán en otros cursos de econometría).

<sup>2</sup>Si interpretamos las variables aleatorias con varianza finita como elementos de un espacio vectorial, entonces  $E(Y | \mathcal{D})$  representa una proyección ortogonal, y la descomposición (1.1) es análoga al teorema de proyección ortogonal (Luenberger, 1968), con  $E(Y | \mathcal{D})$  como el mejor predictor en el sentido de la propiedad ECSV4 en la página 36 del Tema 2.

<sup>3</sup>siempre y cuando  $E(|Y_n|^2) < \infty$

↑
Modelo de regresión
2

$$Y_n = h(\mathbf{X}) + U_n$$

donde :

- $Y_n$ : Vble endógena, objetivo, explicada (o regresando)
- $\mathbf{X} = [X_1, \dots, X_n]'$  : Vble exógena, de control, explicativa (o regresor)
- $U_n$ : factor desconocido o perturbación

Llamamos a  $Y$  *vble. endógena* (porque consideramos que se determina su valor o características a través del modelo), *vble. objetivo* (porque es una magnitud que deseamos controlar, por ejemplo la inflación si somos la autoridad monetaria) o simplemente **regresando**.

El vector columna  $\mathbf{X} = [X_1, \dots, X_n]'$  : formada por  $N$  variables que llamamos *exógenas* (porque consideramos que vienen dadas de manera externa al modelo), o *vbles. de control* (porque tenemos capacidad de alterar su valor para, a través del modelo, *controlar*  $Y$ ; por ejemplo fijar la oferta monetaria o los tipos de interés para controlar la inflación), o simplemente **regresor**.

$U_n$  es el efecto conjunto de otras variables o circunstancias que influyen en la observación de  $Y_n$ , y que decidimos no contemplar en el modelo por alguna razón (dificultad o imposibilidad de observarlas) o sencillamente porque las desconocemos. También puede ser sencillamente un error cometido al medir  $Y_n$ . Llamamos a  $U_n$  **perturbación**.

## 2. Modelo Clásico de Regresión Lineal

↑
Modelo Clásico de Regresión Lineal
3

Modelo especial en el que la descomposición ortogonal

$$Y_n = E(Y_n | \mathbf{X}) + U_n$$

es tal que:

$E_{U_n|\mathbf{X}}(Y_n | \mathbf{x})$  es **función lineal** de  $x_n$

$\text{Var}_{U_n|\mathbf{X}}(Y_n | \mathbf{x})$  es **constante** (**homocedasticidad**)

¿Qué debo suponer para que esto se cumpla?  
(¡al menos como lectura estadística!)

En el análisis de regresión estamos interesados en estimar los dos primeros momentos de  $Y_n$  condicionados a  $\mathbf{X} = \mathbf{x}$ , es decir,  $E_{U_n|\mathbf{X}}(Y_n | \mathbf{x})$  y  $\text{Var}_{U_n|\mathbf{X}}(Y_n | \mathbf{x})$ .

El modelo *Modelo Clásico de Regresión Lineal* es un caso particular en el que  $E_{U_n|\mathbf{X}}(Y_n | \mathbf{x})$  es función lineal de  $x_n$  (del dato con subíndice  $n$ , es decir, del instante  $n$ , o de la empresa  $n$ , o del país  $n$ , o del individuo  $n$ , ...) y  $\text{Var}_{U_n|\mathbf{X}}(Y_n | \mathbf{x})$  es una función constante (por tanto  $Y_n | \mathbf{x}$  es *homocedástica*).

A continuación, vamos a describir los tres supuestos de un modelo econométrico que garantizan la existencia de una descomposición ortogonal como la del modelo clásico de regresión lineal. El cuarto supuesto, que garantiza que la estimación de la relación lineal es única, lo veremos en la sección siguiente.

### 2.1. Tres primeros supuestos en el Modelo Clásico de Regresión Lineal

↑
Supuesto 1: linealidad
4

$h(\cdot)$  es lineal:  $Y_n = h(X_n) + U_n = \beta_1 + \beta_2 X_n + U_n$   
por lo tanto

$$\begin{aligned} Y_1 &= \beta_1 + \beta_2 X_1 + U_1 \\ Y_2 &= \beta_1 + \beta_2 X_2 + U_2 \\ \dots & \quad \dots \\ Y_N &= \beta_1 + \beta_2 X_N + U_N \end{aligned}$$

por tanto

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} = \beta_1 \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \beta_2 \cdot \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}$$

o

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{U}$$

$[\mathbf{Y}]$      $[\mathbf{X}]$      $[\boldsymbol{\beta}]$      $[\mathbf{U}]$   
 $[N \times 1]$      $[N \times 2]$      $[2 \times 1]$      $[N \times 1]$

donde

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}; \quad \mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_N \end{bmatrix}; \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}; \quad \mathbf{U} = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}$$

o bien  $\mathbf{X} = [\mathbf{1}, \mathbf{X}]$ ;    donde     $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$

**Ejemplo 1. [función de consumo:]**

$$CON_i = \beta_1 + \beta_2 RD_i + U_i$$

donde  $CON_i$  y  $RD_i$  son el consumo y la renta disponible del individuo  $i$ -ésimo respectivamente, y  $U_n$  son otros factores que afectan al consumo del individuo  $i$ -ésimo distintos a su renta disponible (activos financieros, estado de ánimo, etc.).

Aquí la variable exógena  $Y$  es el consumo ( $CON$ ), y los regresores son  $X_1 = 1$  (una constante) y  $X_2$  la renta disponible ( $RD$ ).

Supuesto 2: Esperanza condicional de  $\mathbf{U}$  – Estricta exogeneidad 5

$$E_{\mathbf{U}|\mathbf{X}}(\mathbf{U} | \mathbf{x}) = \mathbf{0}_{[n \times 1]}$$

es decir

$$E_{\mathbf{U}|\mathbf{X}}(\mathbf{U} | \mathbf{x}) = \begin{bmatrix} E_{U_1|\mathbf{X}}(U_1 | \mathbf{x}) \\ E_{U_2|\mathbf{X}}(U_2 | \mathbf{x}) \\ \vdots \\ E_{U_n|\mathbf{X}}(U_n | \mathbf{x}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$E_{U_n|\mathbf{X}}(U_n | \mathbf{x}) \equiv E_{U_n|\mathbf{X}}(U_n | x_1, \dots, x_N)$$

para  $n = 1, \dots, N$ .

$$E_{U_n|\mathbf{X}}(U_n | \mathbf{x}) \equiv E_{U_n|\mathbf{X}}\left(U_n \mid \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}\right)$$

para  $n = 1, \dots, N$ .

**Ejemplo 2. [función de consumo: (continuación del Ejemplo 1)]**

Estricta exogeneidad implica que para el individuo  $i$ -ésimo

$$E_{U_i|\mathbf{X}}(U_i | \mathbf{rd}) = E_{U_i|\mathbf{X}}(U_i | (rd_1, rd_2, \dots, rd_k)) = 0,$$

es decir, la esperanza de la perturbación  $i$ -ésima, condicionada a todas y cada una de las rentas disponibles, es cero.

↑

Supuesto 2: Esperanza condicional de  $U$  – Estricta exogeneidad

6

$$E_{U_n | \mathbf{X}}(U_n | \mathbf{x}) = \mathbf{0}_{[n \times 1]} \Rightarrow \begin{cases} E(U_n \mathbf{X}) = \mathbf{0} & \text{ortogonalidad } U_n \perp \mathbf{X} \\ E(U_n) = 0 \\ \text{Cov}(U_n, \mathbf{X}) = 0 \end{cases}$$

(ortogonalidad entre lo que conozco  $\mathbf{X}$  y lo que desconozco  $U_n$ )

**Comentario.** En el caso de regresión con datos temporales, la exogeneidad estricta implica que los regresores son ortogonales a las perturbaciones *pasadas, presentes y futuras*. Esta es una restricción muy fuerte, que no se cumple en general con datos temporales (se discutirá en *Econometría II*).

A continuación aparecen las demostraciones de la transparencia anterior T0.6:

**Proposición 2.1.** Si  $E_{U_n | \mathbf{X}}(U_n | \mathbf{x}) = 0$ , entonces  $E(U_n \mathbf{X}) = \mathbf{0}_{[n \times 1]}$

*Demostración.*

$$\begin{aligned} E(U_n \mathbf{X}) &= \int \cdots \int u_n \mathbf{x} f_{U_n | \mathbf{X}}(u_n, \mathbf{x}) du_n dx_N \cdots dx_1 \\ &= \int \cdots \int u_n \mathbf{x} f_{U_n | \mathbf{X}}(u_n | \mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) du_n dx_N \cdots dx_1 \\ &= \int u_n \left[ \int \cdots \int \mathbf{x} f_{\mathbf{X}}(\mathbf{x}) dx_N \cdots dx_1 \right] f_{U_n | \mathbf{X}}(u_n | \mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) du_n \\ &= \int u_n [E(\mathbf{X})] f_{U_n | \mathbf{X}}(u_n | \mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) du_n \\ &= [E(\mathbf{X})] \int u_n f_{U_n | \mathbf{X}}(u_n | \mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) du_n \\ &= E(\mathbf{X}) \cdot E_{U_n | \mathbf{X}}(U_n | \mathbf{x}) \\ &= E(\mathbf{X}) \cdot \mathbf{0} = \mathbf{0}_{[n \times 1]} \end{aligned} \quad \text{por hipótesis}$$

□

Una importante implicación de  $E_{U_n | \mathbf{X}}(U_n | \mathbf{x}) = 0$ , es que entonces  $E(U_n) = 0$  ya que

$$\begin{aligned} E(U_n) &= E_{\mathbf{X}}(E(U_n | \mathbf{X})) && \text{por el T}^a \text{ de las esperanzas iteradas.} \\ &= E_{\mathbf{X}}(0) = 0 && \text{por ser } E_{U_n | \mathbf{X}}(U_n | \mathbf{x}) \text{ las realizaciones de } E(U_n | \mathbf{X}) \end{aligned}$$

Y de los dos resultados anteriores se deriva que

$$\text{Cov}(U_n, \mathbf{X}) = E(U_n \mathbf{X}) - E(U_n) \cdot E(\mathbf{X}) = \mathbf{0}_{[n \times 1]} - 0 \cdot E(\mathbf{X}) = \mathbf{0}_{[n \times 1]}$$

**EJERCICIO 1. [Relación si y sólo si entre la función de regresión lineal y los supuestos 1 y 2]**  
 Demuestre que los supuestos 1 y 2 implican la primera condición del *Modelo Clásico de Regresión Lineal*, esto es, que la función de regresión de  $Y_n$  sobre los regresores es lineal

$$E_{U_n | \mathbf{X}}(Y_n | \mathbf{x}) = \beta_1 + x_n \beta_2.$$

Recíprocamente, demuestre que si dicha condición se verifica para todo  $t = 1, \dots, T$ , entonces necesariamente se satisfacen los supuestos 1 y 2.

*Solución:*

$$\begin{aligned} E_{U_n | \mathbf{X}}(Y_n | \mathbf{x}) &= E_{U_n | \mathbf{X}}(\beta_1 + X_n \beta_2 + U_n | \mathbf{x}) && \text{por el Supuesto 1} \\ &= \beta_1 + x_n \beta_2 + E_{U_n | \mathbf{X}}(U_n | \mathbf{x}) && \text{puesto que } X_n = x_n \\ &= \beta_1 + x_n \beta_2 && \text{por el Supuesto 2.} \end{aligned}$$

Recíprocamente, suponga que  $E_{U_n|\mathbf{x}}(Y_n | \mathbf{x}) = \beta_1 + x_n\beta_2$  para todo  $n = 1, \dots, N$ . Definamos  $U_n = Y_n - E_{U_n|\mathbf{x}}(Y_n | \mathbf{x})$ . Entonces, por construcción el Supuesto 1 se satisface ya que  $U_n = Y_n - [\beta_1 + \beta_2 X_n]$ . Por otra parte

$$E_{U_n|\mathbf{x}}(U_n | \mathbf{x}) = E_{U_n|\mathbf{x}}(Y_n | \mathbf{x}) - E_{U_n|\mathbf{x}}(E(Y_n | \mathbf{X}) | \mathbf{x}) \quad \text{por la definición que aquí damos a } U_n \\ = 0;$$

pues  $E_{U_n|\mathbf{x}}(E(Y_n | \mathbf{X}) | \mathbf{x}) = E_{U_n|\mathbf{x}}(Y_n | \mathbf{x})$ , ya que:

$$E_{U_n|\mathbf{x}}(E(Y_n | \mathbf{X}) | \mathbf{x}) = \int \left[ \int y_t f_{U_n|\mathbf{x}}(U_n | \mathbf{x}) du_n \right] f_{U_n|\mathbf{x}}(U_n | \mathbf{x}) du_n \\ = \int \left[ \int (U_n + \beta_1 + x_n\beta_2) f_{U_n|\mathbf{x}}(U_n | \mathbf{x}) du_n \right] f_{U_n|\mathbf{x}}(U_n | \mathbf{x}) du_n \\ = \beta_1 + x_n\beta_2 + \int \left[ \int U_n f_{U_n|\mathbf{x}}(U_n | \mathbf{x}) du_n \right] f_{U_n|\mathbf{x}}(U_n | \mathbf{x}) du_n \\ = \beta_1 + x_n\beta_2 + E_{U_n|\mathbf{x}}(E(U_n | \mathbf{X}) | \mathbf{x}) \\ = \beta_1 + x_n\beta_2 + E_{U_n|\mathbf{x}}(U_n | \mathbf{x}) = E_{U_n|\mathbf{x}}(\beta_1 + X_n\beta_2 + U_n | \mathbf{x})$$

Ejercicio 1

⬆
Supuesto 3: Perturbaciones esféricas
7

- homocedasticidad**

$$E_{U_n|\mathbf{x}}(U_n^2 | \mathbf{x}) = \sigma^2 \quad \text{para } n = 1, 2, \dots, N$$
  
- no autocorrelación**

$$E_{U_i U_j | \mathbf{x}}(U_i U_j | \mathbf{x}) = 0 \quad \text{si } i \neq j \quad \text{para } i, j = 1, 2, \dots, N$$

**Nota 1.** Definimos la matriz de varianzas y covarianzas de un vector columna  $\mathbf{Y}$  como

$$\text{Var}(\mathbf{Y}) = \text{Var} \left( \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix} \right) \equiv E(\mathbf{Y}\mathbf{Y}') - E(\mathbf{Y})E(\mathbf{Y}') \\ = \begin{bmatrix} E(Y_1^2) & E(Y_1 Y_2) & \dots & E(Y_1 Y_N) \\ & E(Y_2^2) & \dots & E(Y_2 Y_N) \\ & & \ddots & \vdots \\ & & & E(Y_N^2) \end{bmatrix} - \begin{bmatrix} [E(Y_1)]^2 & E(Y_1)E(Y_2) & \dots & E(Y_1)E(Y_N) \\ & [E(Y_2)]^2 & \dots & E(Y_2)E(Y_N) \\ & & \ddots & \vdots \\ & & & [E(Y_N)]^2 \end{bmatrix} \\ = \begin{bmatrix} \sigma_{Y_1}^2 & \sigma_{Y_1 Y_2} & \dots & \sigma_{Y_1 Y_N} \\ & \sigma_{Y_2}^2 & \dots & \sigma_{Y_2 Y_N} \\ & & \ddots & \vdots \\ & & & \sigma_{Y_N}^2 \end{bmatrix}$$

Sea  $\mathbf{a}_{[m \times N]}$ , entonces,

$$\text{Var}(\mathbf{aY}) = E(\mathbf{aY}\mathbf{Y}'\mathbf{a}') - E(\mathbf{aY})E(\mathbf{Y}'\mathbf{a}') \\ = \mathbf{a} [E(\mathbf{Y}\mathbf{Y}') - E(\mathbf{Y})E(\mathbf{Y}')] \mathbf{a}' \quad \text{sacando factores comunes} \\ = \mathbf{a} \text{Var}(\mathbf{Y}) \mathbf{a}'$$

↑
Supuestos 2 y 3: Implicación conjunta
8

$$\text{Var}_{\mathbf{U}|\mathbf{x}}(\mathbf{U} | \mathbf{x}) = \begin{pmatrix} \text{Var}_{U_1|\mathbf{x}}(U_1 | \mathbf{x}) & \text{Cov}_{U_1 U_2|\mathbf{x}}(U_1, U_2 | \mathbf{x}) & \dots & \text{Cov}_{U_1 U_N|\mathbf{x}}(U_1, U_N | \mathbf{x}) \\ \text{Cov}_{U_2 U_1|\mathbf{x}}(U_2, U_1 | \mathbf{x}) & \text{Var}_{U_2|\mathbf{x}}(U_2 | \mathbf{x}) & \dots & \text{Cov}_{U_2 U_N|\mathbf{x}}(U_2, U_N | \mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}_{U_N U_1|\mathbf{x}}(U_N, U_1 | \mathbf{x}) & \text{Cov}_{U_N U_2|\mathbf{x}}(U_N, U_2 | \mathbf{x}) & \dots & \text{Var}_{U_N|\mathbf{x}}(U_N | \mathbf{x}) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} = \sigma^2 \mathbf{I}_{[N \times N]}$$

$$\begin{aligned} \text{Var}_{\mathbf{U}|\mathbf{x}}(\mathbf{U} | \mathbf{x}) &= \text{E}_{\mathbf{U}|\mathbf{x}}(\mathbf{U}\mathbf{U}' | \mathbf{x}) - \text{E}_{\mathbf{U}|\mathbf{x}}(\mathbf{U} | \mathbf{x}) \text{E}_{\mathbf{U}|\mathbf{x}}(\mathbf{U}' | \mathbf{x}) \\ &= \text{E}_{\mathbf{U}|\mathbf{x}} \left( \begin{bmatrix} U_1 \\ \vdots \\ U_N \end{bmatrix} \begin{bmatrix} U_1 & \dots & U_N \end{bmatrix} \middle| \mathbf{x} \right) - \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \dots & 0 \end{bmatrix} && \text{por el Supuesto 2} \\ &= \begin{bmatrix} \text{E}_{U_1|\mathbf{x}}(U_1^2 | \mathbf{x}) & \text{E}_{U_1 U_2|\mathbf{x}}(U_1 U_2 | \mathbf{x}) & \dots & \text{E}_{U_1 U_N|\mathbf{x}}(U_1 U_N | \mathbf{x}) \\ \text{E}_{U_2 U_1|\mathbf{x}}(U_2 U_1 | \mathbf{x}) & \text{E}_{U_2|\mathbf{x}}(U_2^2 | \mathbf{x}) & \dots & \text{E}_{U_2 U_N|\mathbf{x}}(U_2 U_N | \mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{E}_{U_N U_1|\mathbf{x}}(U_N U_1 | \mathbf{x}) & \text{E}_{U_N U_2|\mathbf{x}}(U_N U_2 | \mathbf{x}) & \dots & \text{E}_{U_N|\mathbf{x}}(U_N^2 | \mathbf{x}) \end{bmatrix} - \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \\ &= \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} && \text{por el Supuesto 3} \end{aligned}$$

El supuesto de que la matriz de varianzas y covarianzas de la perturbaciones (condicionada a  $\mathbf{x}$ ) es  $\sigma^2$  veces la matriz identidad (*perturbaciones esféricas*)

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

es una restricción muy fuerte, ya que implica:

1. que la dispersión (la varianza) del efecto de término perturbación asociada a cada observación (o a cada instante, o a cada individuo, etc) es idéntica a la de las demás (no sabemos exactamente a que se debe la perturbación que afecta a cada  $Y_n$  pero la dispersión (incertidumbre) de ese efecto es idéntica para todos).

Dicho de otra forma: las perturbaciones  $U_n$  son *homedásticas*, ya que

$$\text{Var}_{U_n|\mathbf{x}}(U_n | \mathbf{x}) = \sigma^2 \quad \text{para todo } n = 1 : N.$$

2. que la covarianza entre las perturbaciones de observaciones distintas (o de instantes ,o individuos diferentes) es cero. Dicho de otra forma: las perturbaciones no tienen *correlación serial*, ya que

$$\text{Cov}_{U_i U_j|\mathbf{x}}(U_i, U_j | \mathbf{x}) = 0 \quad \text{para } i \neq j.$$

Esto añadido al supuesto de distribución conjunta Normal (ver Supuesto 5 más adelante T0.18) significará que las perturbaciones son *independientes* para las distintas observaciones.

**Nota 2 (Relación entre la función cedástica contante y los supuestos 1, 2 y 3).** Nótese que con los tres supuestos también se cumple la segunda condición del modelo clásico de regresión lineal ya que

$$\text{Var}_{U_n|\mathbf{x}}(Y_n | \mathbf{x}) = \text{Var}_{U_n|\mathbf{x}}(\beta_1 + \beta_2 X_n + U_n | \mathbf{x}) = \text{Var}_{U_n|\mathbf{x}}(U_n | \mathbf{x}) = \sigma^2$$

## Variación de los supuestos 2 y 3 en algunos casos especiales:

### 2.2. Modelo Clásico de Regresión Lineal para Muestras Aleatorias

Si  $(\mathbf{Y}, \mathbf{X})$  es una muestra aleatoria simple, i.e.  $\{Y_n, X_n\}$  es i.i.d. para  $n = 1, \dots, N$ ; entonces,

$$\begin{aligned} E_{U_i|\mathbf{X}}(U_i | \mathbf{x}) &= E_{U_i|X_i}(U_i | x_i) \\ E_{U_i|\mathbf{X}}(U_i^2 | \mathbf{x}) &= E_{U_i|X_i}(U_i^2 | x_i) \\ \text{y también } E_{U_i U_j|\mathbf{X}}(U_i U_j | \mathbf{x}) &= E_{U_i|X_i}(U_i | x_i) E_{U_j|X_j}(U_j | x_j) \quad \text{para } i \neq j \end{aligned}$$

Con lo que los los supuestos 2 [T0.5](#) y 3 [T0.7](#) quedan reducidos a

**supuesto 2'**:  $E_{U_i|X_i}(U_i | x_i) = 0$

**supuesto 3'**:  $E_{U_i|X_i}(U_i^2 | x_i) = \sigma^2 > 0$

para todo  $n = 1, \dots, N$

(Nótese que los regresores están referidos exclusivamente a la observación  $n$ -ésima)

En general *este supuesto no es adecuado para modelos con datos de series temporales* ya que las muestras no son i.i.d. (no son muestras aleatorias simples puesto que suele haber correlación entre los datos).

**EJERCICIO 2.** Demuestre que

$$E_{U_i U_j|\mathbf{X}}(U_i U_j | \mathbf{x}) = E_{U_i|X_i}(U_i | x_i) E_{U_j|X_j}(U_j | x_j) \quad \text{para } i \neq j$$

para el caso de muestras aleatorias simples (*m.a.s.*)

*Pista.*

$$E_{U_i U_j|\mathbf{X}}(U_i U_j | \mathbf{x}) = E_{U_i U_j|\mathbf{X}}(E(U_i | \mathbf{X} U_j) \cdot U_j | \mathbf{x})$$

debido a que  $\{U_i, X_i\}$  es independiente de  $\{U_j, X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_N\}$  para  $i \neq j$ , junto con el teorema de las esperanzas iteradas.

### 2.3. Regresores No Estocásticos

Si los regresores son no estocásticos, es decir son la matriz determinista  $\mathbf{x}$ , entonces no es necesario distinguir entre funciones de densidad condicionales,  $f_{U_n|\mathbf{X}}(U_n | \mathbf{x})$ , e incondicionales,  $f_{U_N}(U_n)$ ; por tanto los supuestos 2 [T0.5](#) y 3 [T0.7](#) quedan reducidos a

**supuesto 2''**:  $E(U_n) = 0$

**supuesto 3''**:  $E(U_n^2) = \sigma^2 > 0$  y  $E(U_i U_j) = 0$  para  $i \neq j$

para todo  $n, i, j = 1, \dots, N$

(Estos son los supuestos empleados en la mayoría de libros de texto, como por ejemplo en [Novales \(1993\)](#))

*Este caso no puede suponerse con modelos autorregresivos o de ecuaciones simultáneas.*

Quedan un cuarto supuesto acerca del rango de la matriz de regresores (que garantiza la unicidad en la estimación) y un quinto supuesto acerca de la distribución conjunta de  $\mathbf{U}$  que enunciaremos más adelante (véase Supuesto 4 [T0.9](#) y Supuesto 5 [T0.18](#))

## 3. Estimación MCO (Mínimos Cuadrados Ordinarios)

Necesitamos un cuarto supuesto que nos permita obtener estimaciones de  $\beta_1$  y  $\beta_2$

⬆
Supuesto 4: Independencia lineal de los regresores
9

El rango de  $\mathbf{X}_{[N \times 2]}$  es 2 con probabilidad 1.

- número de observaciones  $\geq 2$
- Vectores columna  $\mathbf{1}$ ,  $\mathbf{X}$  linealmente indep.

Por tanto  $\mathbf{X}$  no es un vector de constantes  $a$ , es decir,  $P(\mathbf{X} \neq a \cdot \mathbf{1}) = 1$ ;  
O lo que es lo mismo  $\text{Var}(X_n) \neq 0$ .

Se dice que existe *multicolinealidad perfecta* cuando el Supuesto 4 **NO** se satisface.

↑
Modelo Lineal Simple
10

Por el **Supuesto 1**  $Y_n = a + bX_n + U_n$ ; entonces

$$\begin{aligned} E_{U_n|X_n}(Y_n | x_n) &= E_{U_n|X_n}(a + bX_n + U_n | x_n) = \\ &= a + bx_n + E_{U_n|X_n}(U_n | x_n) \\ &= a + bx_n. \end{aligned}$$

Por lo tanto, sabemos que  $E_{U_n|X_n}(Y_n | x_n)$  es

$$E_{U_n|X_n}(Y_n | x_n) = \underbrace{E(Y) - E(X)}_a \frac{\text{Cov}(Y, X)}{\text{Var}(X)} + \underbrace{\frac{\text{Cov}(X, Y)}{\text{Var}(X)}}_b \cdot x_n;$$

para todo  $x_n \in \mathbb{R}_X$ ,

Véanse las ecuaciones (10.3) y (10.4) Sección 10.4 (**Esperanza condicional cuando es una función lineal**) del Tema 2, página 37.

Véanse las ecuaciones (10.3) y (10.4) Sección 10.4 (**Esperanza condicional cuando es una función lineal**) del Tema 2, página 37.

↑
Modelo Lineal Simple: Metodo de los momentos
11

Estimaremos los parámetros  $a$  y  $b$  sustituyendo los momentos teóricos por los muestrales; es decir

$$\hat{b} = \frac{s_{xy}}{s_x^2} \tag{3.1}$$

y

$$\hat{a} = \bar{y} - \frac{s_{xy}}{s_x^2} \bar{x} = \bar{y} - \hat{b} \bar{x} \tag{3.2}$$

Supuesto 4 (independencia lineal de regresores)  $\Rightarrow$  solución única.

↑
Término de error
12

Las perturbaciones  $U_n$  no son observables  
 Pero las podemos estimar sustituyendo  $\beta$  por  $\hat{\beta}$  para una muestra concreta  $\{y_n, x_n\}_{n=1}^N$  de  $\{Y_n, X_n\}_{n=1}^N$ .

$$\hat{e}_n = y_n - (\hat{a} + \hat{b}x_n) = y_n - \hat{y}_n; \quad \text{donde } \hat{y}_n = \hat{a} + \hat{b}x_n$$

De manera que:  $y_n = \hat{y}_n + \hat{e}_n$

**Ejemplo 3.** [precio de las viviendas:]

$t$	Precio	Superficie
1	199.9	1065
2	228.0	1254
3	235.0	1300
4	285.0	1577
5	239.0	1600
6	293.0	1750
7	285.0	1800
8	365.0	1870
9	295.0	1935
10	290.0	1948
11	385.0	2254
12	505.0	2600
13	425.0	2800
14	415.0	3000

**Cuadro 1:** Superficie (en pies al cuadrado) y precio de venta de los pisos (en miles de dólares) (Ramanathan, 1998, pp. 78)

Planteamos el modelo  $Y_n = a + bX_n + U_n$ , donde  $Y_n$  es el precio del piso  $t$ -ésimo,  $X_n$  es su superficie, y  $U_n$

son otros factores que influyen en el precio del piso, pero “ortogonales” al la superficie del mismo (situación, estado de mantenimiento, servicios, etc.). Deseamos saber cual es el efecto *marginal* del incremento de la superficie de un piso en su precio, y disponemos de datos de 14 pisos (Cuadro 1 en la página anterior). Por lo tanto necesitamos estimar el valor del parámetro  $b$ .

Calculando los momentos muestrales de  $x$  e  $y$  obtenemos

$$\hat{a} = \bar{y} - \bar{x} \frac{s_{xy}}{s_x^2} = 52.3509; \quad \hat{b} = \frac{s_{xy}}{s_x^2} = 0.13875$$

$$\widehat{\text{price}} = 52,3509 + 0,138750 \text{ sqft}$$

(1,404)            (7,407)

$$T = 14 \quad \bar{R}^2 = 0,8056 \quad F(1, 12) = 54,861 \quad \hat{\sigma} = 39,023$$

(entre paréntesis, los estadísticos  $t$ )

Estimaciones MCO utilizando las 14 observaciones 1–14  
Variable dependiente: price

Variable	Coeficiente	Desv. típica	Estadístico $t$	valor p
const	52,3509	37,2855	1,4041	0,1857
sqft	0,138750	0,0187329	7,4068	0,0000
Media de la var. dependiente		317,493		
D.T. de la variable dependiente		88,4982		
Suma de cuadrados de los residuos		18273,6		
Desviación típica de los residuos ( $\hat{\sigma}$ )		39,0230		
$R^2$		0,820522		
$\bar{R}^2$ corregido		0,805565		
Grados de libertad		12		
Criterio de información de Akaike		144,168		
Criterio de información Bayesiano de Schwarz		145,447		

Salida del programa “libre” **Gretl** (Gnu Regression, Econometrics and Time-series Library)

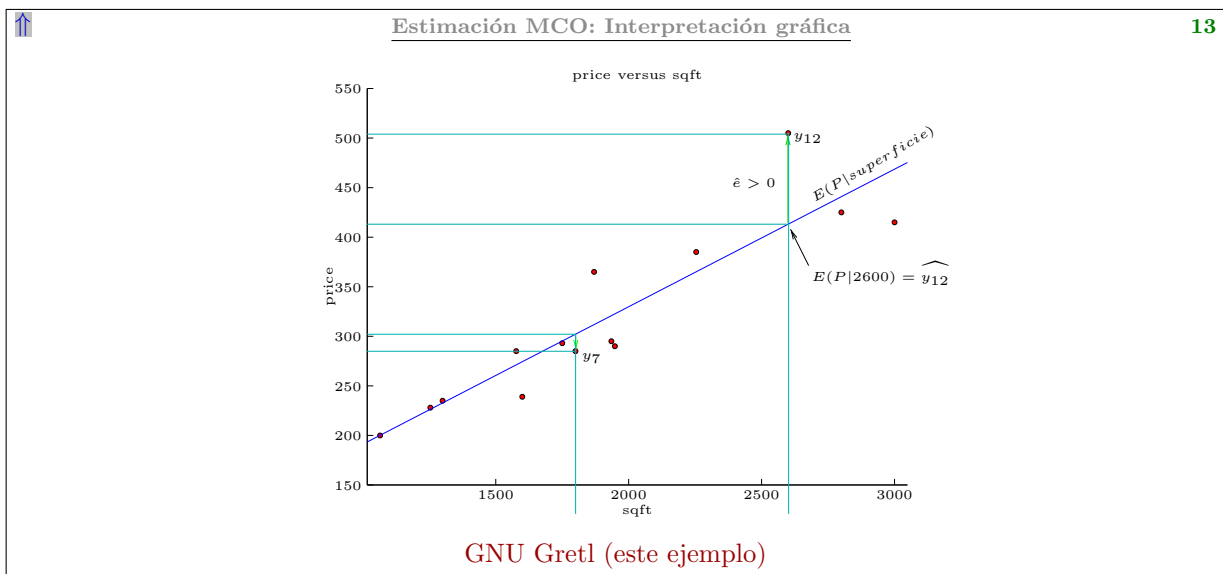
Por lo tanto, el precio de venta *esperado* de un piso con una superficie de 1800 pies cuadrados,  $E_{yx}(Y | 1800)$ , será de

$$\hat{y}_7 = 52.3509 + 0.139 \cdot 1800 = 302.551$$

sin embargo en la muestra hay un piso con dicha superficie cuyo precio es  $y_7 = 285$ . Esta discrepancia es el error de estimación, que es una estimación de la perturbación (o factor desconocido)  $U_7$  (el error  $\hat{e}_7$  puede deberse por tanto a que dicho piso esta en una mala situación, dispone de pocos servicios, etc.; es decir, características distintas a su superficie)

$t$	Precio	Superficie	Precio estimado $E(P   \text{superficie})$	Error $\hat{e}$
1	199.9	1065	200.1200	-0.22000
2	228.0	1254	226.3438	1.65619
3	235.0	1300	232.7263	2.27368
4	285.0	1577	271.1602	13.83984
5	239.0	1600	274.3514	-35.35142
6	293.0	1750	295.1640	-2.16397
7	285.0	1800	302.1015	-17.10148
8	365.0	1870	311.8140	53.18600
9	295.0	1935	320.8328	-25.83278
10	290.0	1948	322.6365	-32.63653
11	385.0	2254	365.0941	19.90587
12	505.0	2600	413.1017	91.89826
13	425.0	2800	440.8518	-15.85180
14	415.0	3000	468.6019	-53.60187

**Cuadro 2:** Superficie (en pies al cuadrado), precio de venta (en miles de dólares), precio estimado, y errores estimados.



Continuación del ejemplo “precio de las viviendas” en la página ~14

#### 4. Propiedades algebraicas de la estimación MCO

Mínimos cuadrados ordinarios: Propiedades algebraicas 14

$$\sum \hat{e}_n = 0$$

$$\sum x_n \hat{e}_n = 0$$

**EJERCICIO 3.** Demuestre las dos igualdades de la transparencia anterior.

#### 5. Propiedades estadísticas del estimador MCO $\hat{\beta} | x$

**Nota 3.** Sean  $x$  e  $y$  vectores de orden  $N$ , entonces

$$\sum_n (x_n - \bar{x})(y_n - \bar{y}) = \sum_n y_n(x_n - \bar{x}) \quad \text{para } n = 1, \dots, N.$$

*Demostración.*

$$\begin{aligned} \sum_n (x_n - \bar{x})(y_n - \bar{y}) &= \sum_n y_n(x_n - \bar{x}) - \bar{y} \sum_n (x_n - \bar{x}) \\ &= \sum_n y_n(x_n - \bar{x}) - \bar{y} \cdot 0 = \sum_n y_n(x_n - \bar{x}) \quad \text{para } n = 1, \dots, N. \end{aligned}$$

□

Así pues, de 3.1 en la página ~10 resulta

$$\begin{aligned} \hat{b} &= \frac{\sum_n (x_n - \bar{x})(y_n - \bar{y})/N}{\sum_n (x_n - \bar{x})^2/N} = \frac{\sum_n y_n(x_n - \bar{x})}{\sum_n (x_n - \bar{x})^2} \\ &= \sum_n y_n \left( \frac{x_n - \bar{x}}{\sum_n (x_n - \bar{x})^2} \right). \end{aligned}$$

es decir,

$$\hat{b} = \sum_n m_n y_n, \quad (5.1)$$

donde

$$m_n = \frac{x_n - \bar{x}}{\sum_n (x_n - \bar{x})^2}.$$

Por tanto,  $\hat{b}$  es una combinación lineal de los datos  $y_n$  (donde  $m_n$  son los coeficientes de dicha combinación); y entonces  $\hat{a}$  también es combinación lineal de los datos  $y_n$  (véase 3.2 en la página ~10).

<u>Estimador MCO <math>\hat{\beta} \mid \mathbf{x}</math></u>	15
<p><b>Estimaciones:</b></p> $\hat{b} = \sum_n m_n y_n,$ <p>Por tanto, <math>\hat{b}</math> es combinación lineal de los datos <math>y_n</math>; entonces <math>\hat{a}</math> también lo es (véase 3.2 en la página ~10).</p> <p><b>Estimadores:</b></p> $\hat{b}_{\mid \mathbf{x}} = \sum_n m_n Y_n.$ <p>y</p> $\hat{a}_{\mid \mathbf{x}} = \frac{\sum Y_n}{N} - (\hat{b}_{\mid \mathbf{x}}) \cdot \bar{x}.$ <p>Así que podemos calcular los momentos de los estimadores</p> <p>Análogamente <math>\widehat{y_n}_{\mid \mathbf{x}} = \hat{a}_{\mid \mathbf{x}} + (\hat{b}_{\mid \mathbf{x}}) \cdot x_n;</math>    y    <math>\widehat{e_n}_{\mid \mathbf{x}} = Y_n - \widehat{y}_{\mid \mathbf{x}}.</math></p>	

Se puede verificar que

$$\sum m_t = 0 \quad (5.2)$$

$$\sum m_t^2 = \frac{1}{\sum x_t^2} = \frac{1}{\sum (x_t - \bar{x})^2} = \frac{1}{T s_x^2} \quad (5.3)$$

$$\sum m_t x_t = \sum m_t (x_t - \bar{x}) = 1 \quad (5.4)$$

(Novales, 1997; Gujarati, 2003, pag. 488–491 y pag. 100 respectivamente).

**EJERCICIO 4.** Demuestre que las ecuaciones 5.2 a 5.4 son ciertas.

	Esperanza del estimador MCO $\hat{b} \mid \mathbf{x}$	16
<p>Por 5.1 en la página anterior sabemos que <math>\hat{b} \mid \mathbf{x} = \sum m_t Y_n</math>. Entonces,</p> $\hat{b} \mid \mathbf{x} = \sum m_t (a + bx_t + U_n)$ $= a \sum m_t + b \sum m_t x_t + \sum m_t U_n = b + \sum m_t U_n \quad (5.5)$ <p>y</p> $E_{U_n \mid \mathbf{x}}(\hat{b} \mid \mathbf{x}) = b + \sum m_t E_{U_n \mid \mathbf{x}}(U_n \mid \mathbf{x}) = b.$ <p>Por tanto el estimador es <b>insesgado</b> (lo mismo ocurre para <math>\hat{a}</math>)</p>		

**EJERCICIO 5.** Verifique que el estimador MCO del parámetro  $a$  es insesgado.

	Varianzas y covarianzas de los estimadores MCO	17
<p>Además</p> $\begin{aligned} \text{Var}(\hat{b} \mid \mathbf{x}) &= \text{Var}\left(b + \sum m_t U_n\right) && \text{de (5.5)} \\ &= \sum \text{Var}(m_t U_n) && \text{pues Cov}(U_i, U_j) = 0 \text{ si } i \neq j \\ &= \sum m_t^2 \text{Var}(U_n) \\ &= \sigma^2 \sum m_t^2 \\ &= \frac{\sigma^2}{T s_x^2} && \text{de (5.3)} \end{aligned} \quad (5.6)$ <p>Y</p> $\text{Var}(\hat{a} \mid \mathbf{x}) = \frac{\sigma^2 \sum x^2}{N s_x^2}; \quad \text{Cov}(\hat{a} \mid \mathbf{x}, \hat{b} \mid \mathbf{x}) = \frac{-\sigma^2 \cdot \bar{x}}{T \cdot s_x^2} \quad (5.7)$		

**EJERCICIO 6.** Demuestre que las igualdades de (5.7) son ciertas.

**Ejemplo 5.** [continuación de “precio de las viviendas”:]

Las desviaciones típicas de  $\hat{a} \mid \mathbf{x}$  y  $\hat{b} \mid \mathbf{x}$  son:

$$\text{Dt}(\hat{a} \mid \mathbf{x}) = \sqrt{\sigma^2 \cdot (9.1293e - 01)} = \sqrt{\frac{\sigma^2 \bar{x}^2}{T \cdot s_x^2}}$$

$$\text{Dt}(\hat{b} \mid \mathbf{x}) = \sqrt{\sigma^2 \cdot (2.3044e - 07)} = \sqrt{\frac{\sigma^2}{T \cdot s_x^2}}$$

Pero no conocemos  $\sigma_{U_n}^2$ .

Continuación del ejemplo “precio de las viviendas” en la página siguiente

## 6. Distribución de los estimadores MCO bajo Normalidad

↑ Supuesto 5: Distribución Normal de las perturbaciones 18

Para conocer la distribución completa necesitamos un supuesto más sobre la distribución conjunta de  $\mathbf{U}$ :

$$\mathbf{U} | \mathbf{x} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad \Rightarrow \quad \mathbf{Y} | \mathbf{x} \sim N(a + b\mathbf{x}, \sigma^2 \mathbf{I})$$

donde  $\mathbf{I}$  es la matriz identidad.

Ya que  $\hat{a}_{|\mathbf{x}}$  y  $\hat{b}_{|\mathbf{x}}$  combinaciones lineales de  $\mathbf{Y}_n$ , tienen distribución Normal.

Por tanto

$$\hat{a}_{|\mathbf{x}} \sim N\left(a, \frac{\sigma^2 \sum x^2}{N s_{\mathbf{x}}^2}\right); \quad \hat{b}_{|\mathbf{x}} \sim N\left(b, \frac{\sigma^2}{N s_{\mathbf{x}}^2}\right)$$

↑ Distribución del estimador MCO  $\hat{\beta}_j | \mathbf{x}$  19

(a partir de ahora también denotaremos un estimador  $\hat{\theta} | \mathbf{x}$  sencillamente con  $\hat{\theta}$ )

Así pues,

$$\hat{\beta}_j \sim N\left(\beta_j, \text{Var}\left(\hat{\beta}_j\right)\right)$$

y

$$\frac{\hat{\beta}_j - \beta_j}{\text{Dt}\left(\hat{\beta}_j\right)} \sim N(0, 1)$$

↑ Estimación de la varianza residual 20

El parámetro  $\sigma^2$  es desconocido T0.7

La cuasivarianza de  $\hat{\mathbf{e}}$

$$\hat{s}_{\mathbf{e}}^2 \equiv \frac{\sum \hat{e}_n^2}{N-2}$$

es estimador insesgado de  $\sigma^2$ .  
(veáse [Novales, 1997](#), Apéndice 13.A)

↑ Distribución cuando la varianza de  $\mathbf{U}$  es desconocida 21

sustituyendo  $\sigma^2$  por su estimador,  $\hat{s}_{\mathbf{e}}^2$ , tenemos el estadístico  $\mathcal{T}$  del parámetro

$$\frac{\hat{a} - a}{\sqrt{\frac{\hat{s}_{\mathbf{e}}^2}{\sum (x_i - \bar{x})^2}}} = \mathcal{T}^a \sim t_{N-2}; \quad \frac{\hat{b} - b}{\sqrt{\frac{\hat{s}_{\mathbf{e}}^2}{N \sum (x_t - \bar{x})^2}}} = \mathcal{T}^b \sim t_{N-2};$$

## 7. Estimación por intervalos

Vimos que para dar una estimación por intervalos, fijado  $(1 - \alpha)$ , debíamos buscar los números *min* y *max* tales que

$$P(\min \leq \mathcal{T} \leq \max) = (1 - \alpha)$$

donde *min* y *max* son  $\pm t_{(N-2, \alpha/2)}$ ; entonces

$$\beta_k \in \left[ \hat{\beta}_k \pm t_{(N-n, \alpha/2)} \cdot \widehat{\text{Dt}}\left(\hat{\beta}_k\right) \right] \quad \text{con Prob.} = (1 - \alpha).$$

### Ejemplo 5. [continuación de “precio de las viviendas”:]

Las desviaciones típicas de  $\hat{a}_{|\mathbf{x}}$  y  $\hat{b}_{|\mathbf{x}}$  son:

$$\text{Dt}\left(\hat{a}_{|\mathbf{x}}\right) = \sqrt{\sigma^2 \cdot (9.1293e - 01)} = \sqrt{\frac{\sigma^2 \mathbf{x}^2}{N \cdot s_{\mathbf{x}}^2}}$$

$$\text{Dt}\left(\hat{b}_{|\mathbf{x}}\right) = \sqrt{\sigma^2 \cdot (2.3044e - 07)} = \sqrt{\frac{\sigma^2}{N \cdot s_{\mathbf{x}}^2}}.$$

No conocemos  $\sigma_{\hat{U}_n}^2$ , pero podemos sustituirla por la cuasi-varianza de los errores estimados; puesto que  $\hat{s}_{\hat{\epsilon}}^2 = \frac{\hat{\epsilon}'\hat{\epsilon}}{N-2} = \frac{18273.6}{14-2} = 1522.8$ :

$$\widehat{\text{Dt}}(\hat{a}_{|\mathbf{x}}) = \sqrt{(1522.8) \cdot (9.1293e - 01)} = \sqrt{\frac{(1522.8) \cdot \bar{\mathbf{x}}^2}{N \cdot s_{\mathbf{x}}^2}} = 37.285;$$

$$\widehat{\text{Dt}}(\hat{b}_{|\mathbf{x}}) = \sqrt{(1522.8) \cdot (2.3044e - 07)} = \sqrt{\frac{1522.8}{N \cdot s_{\mathbf{x}}^2}} = 0.01873$$

Véase los resultados de estimación en el ejemplo del precio de las viviendas (página 11).

Por otra parte,  $\widehat{\text{Cov}}(\hat{a}_{|\mathbf{x}}, \hat{b}_{|\mathbf{x}}) = (1522.8) \cdot (-4.4036e - 04) = \frac{-\hat{s}_{\hat{\epsilon}}^2 \cdot \bar{\mathbf{x}}}{N \cdot s_{\mathbf{x}}^2} = -0.671$

Por tanto, la estimación del efecto marginal de la superficie sobre el precio es,

$$b \in [0.139 \pm (t_{(12, \alpha/2)}) \cdot 0.01873];$$

y para la constante

$$a \in [52.351 \pm (t_{(12, \alpha/2)}) \cdot 37.285];$$

Continuación del ejemplo “precio de las viviendas” en esta página

## 8. Contratación de hipótesis

Esta parte de las notas se verá después de cubrir el Tema 6 del programa.

**Ejemplo 5.** [continuación de “precio de las viviendas”:] Podemos contrastar la significatividad 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:

$$H_0 : a = 0$$

$$H_1 : a \neq 0$$

En este caso la región crítica debe ser

$$RC = \left\{ \mathbf{x} \left| \frac{\hat{a} - 0}{\sqrt{\hat{s}^2 \frac{\sum x_t^2}{N \sum (x_t - \bar{x})^2}}} > k_1 \text{ o } \frac{\hat{a} - 0}{\sqrt{\hat{s}^2 \frac{\sum x_t^2}{N \sum (x_t - \bar{x})^2}}} < k_2 \right. \right\},$$

Para  $\alpha = .05$  resulta  $k_1 = t_{12, \alpha/2} = 2.16 = -k_2$ , es decir

$$\frac{52.351}{37.285} = 1.4041 \quad \text{no rechazamos } H_0 \text{ con nivel de significación del 5\%}.$$

Un experto en el mercado de la vivienda afirma que un aumento de un pie cuadrado en la superficie supone un incremento de 150 dolares (¡si no más!). A la luz de los datos, podemos creer al experto a un nivel de significación del 2.5 %.

$$H_0 : b = 0.15$$

$$H_1 : b < 0.15$$

La región crítica de una sólo cola es

$$RC = \left\{ \mathbf{x} \left| \frac{\hat{b} - 0.15}{\sqrt{\frac{\hat{s}^2}{\sum (x_t - \bar{x})^2}}} < k \right. \right\}$$

sustituyendo lo que hemos calculado, tenemos que

$$\frac{0.139 - 0.15}{0.01873} = -0.58729 > t_{12, 0.025} = -2.16$$

no podemos rechazar la opinión del experto.

## 9. Ejercicios

A resolver en clase

**EJERCICIO 7.** Demuestre que en el modelo de regresión simple  $Y_n = a + bx_t + U_n$  el supuesto  $E_{U_n|\mathbf{x}}(U_n | \mathbf{x}) = 0$  implica  $E_{U_n|\mathbf{x}}(Y_n | \mathbf{x}) = a + bx_t$ ; y  $U$  es la perturbación aleatoria del modelo.

A resolver en clase

**EJERCICIO 8.** (Consta de 5 apartados)

Sean los siguientes datos:

Empresa	$y_i$	$x_i$	$x_i y_i$	$x_i^2$
A	1	1	1	1
B	3	2	6	4
C	4	4	16	16
D	6	4	24	16
E	8	5	40	25
F	9	7	63	49
G	11	8	88	64
H	14	9	126	81
sumas	56	40	364	256

**Cuadro 3:**

donde  $y$  son beneficios, y  $x$  son gastos en formación de personal de una empresa.

Además se sabe que las varianzas y covarianzas muestrales son tales que:

$$N \cdot s_y^2 = \sum (y_i - \bar{y})^2 = 132,$$

$$N \cdot s_x^2 = \sum (x_i - \bar{x})^2 = 56,$$

$$N \cdot s_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = 84,$$

donde  $N$  es el tamaño muestral.

Suponga que se plantea el siguiente modelo

$$Y_i = a + bx_i + U_i,$$

donde  $U_i$  son otros factores que afectan a los beneficios distintos de sus gastos en formación (el término de error). 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 8, y  $\sigma^2$  es la varianza de  $U_i$ , cuyo valor es desconocido.

- Estime por MCO los parámetros  $a$  y  $b$  del modelo.
- ¿Cuál es el beneficio esperado para una empresa que incurriera en unos gastos de formación de personal de 3?
- Calcule los residuos de la empresa E y F. ¿Que indica en este caso el signo de los residuos? La comparación de los residuos para estas empresas ¿contradice el hecho de que F tiene mayores beneficios que E? Justifique su respuesta.
- Estime por MCO un intervalo de confianza del 95% para el parámetro  $b$  del modelo, sabiendo que la suma de los residuos al cuadrado es 6.
- Contraste la hipótesis de que “la pendiente del modelo es uno” frente a que “es menor que uno” con un nivel de significación del 10%. ¿Cuál es el p-valor de la estimación de “dicha pendiente”?

Mas ejercicios completos de regresión en el Tema 6 del programa.

### Lista de Transparencias

- Descomposición ortogonal y causalidad
- Modelo de regresión
- Modelo Clásico de Regresión Lineal
- Supuesto 1: linealidad

- 5 Supuesto 2: Esperanza condicional de  $U$ – Estricta exogeneidad
- 6 Supuesto 2: Esperanza condicional de  $U$ – Estricta exogeneidad
- 7 Supuesto 3: Perturbaciones esféricas
- 8 Supuestos 2 y 3: Implicación conjunta
- 9 Supuesto 4: Independencia lineal de los regresores
- 10 Modelo Lineal Simple
- 11 Modelo Lineal Simple: Metodo de los momentos
- 12 Término de error
- 13 Estimación MCO: Interpretación gráfica
- 14 Mínimos cuadrados ordinarios: Propiedades algebraicas
- 15 Estimador MCO  $\hat{\beta} \mid \mathbf{x}$
- 16 Esperanza del estimador MCO  $\hat{b} \mid \mathbf{x}$
- 17 Varianzas y covarianzas de los estimadores MCO
- 18 Supuesto 5: Distribución Normal de las perturbaciones
- 19 Distribución del estimador MCO  $\hat{\beta}_j \mid \mathbf{x}$
- 20 Estimación de la varianza residual
- 21 Distribución cuando la varianza de  $U$  es desconocida
- 22 Partes del temario

## 10. Bibliografía

- Gujarati, D. N. (2003). *Basic Econometrics*. McGraw-Hill, cuarta ed. ISBN 0-07-112342-3. International edition. 13
- Hayashi, F. (2000). *Econometrics*. Princeton University Press, Princeton, New Jersey. ISBN 0-691-01018-8. 2
- Luenberger, D. G. (1968). *Optimization by vector space methods*. Series in decision and control. John Wiley & Sons, Inc., New York. 3
- Novalés, A. (1993). *Econometría*. McGraw-Hill, segunda ed. 9
- Novalés, A. (1997). *Estadística y Econometría*. McGraw-Hill, Madrid, primera ed. ISBN 84-481-0798-5. 2, 13, 15, 19
- Peña, D. (2001). *Fundamentos de Estadística*. Alianza Editorial, Madrid. ISBN 84-206-8696-4. 2
- Peña, D. (2002). *Regresión y diseño de experimentos*. Alianza Editorial, Madrid. ISBN 84-206-8695-6. 2
- Peña, D. y Romo, J. (1997). *Introducción a la Estadística para la Ciencias Sociales*. McGraw-Hill, Madrid. ISBN 84-481-1617-8. 2
- Ramanathan, R. (1998). *Introductory Econometrics with Applications*. Harcourt College Publisher, Orlando. 10
- Spanos, A. (1999). *Probability Theory and Statistical Inference. Econometric Modeling with Observational Data*. Cambridge University Press, Cambridge, UK. ISBN 0-521-42408-9. 2

### Partes del temario

- Tema 1 IntEctr-T01
- Tema 2 IntEctr-T02
- Tema 3 IntEctr-T03
- Tema 4 IntEctr-T04
- Tema 5 IntEctr-T05
- Tema 6 IntEctr-T06
- Tema 7 IntEctr-T07

### Soluciones a los Ejercicios

**Ejercicio 2.** Puesto que  $\{U_i, X_i\}$  es independiente de  $\{U_j, X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_N\}$  tenemos que

$$E_{U_j, X_j}(U_i | \mathbf{x}, U_j) = E_{U_i, X_i}(U_i | x_i).$$

Así

$$\begin{aligned} E_{U_i, U_j, \mathbf{X}}(U_i U_j | \mathbf{x}) &= E_{U_i, U_j, \mathbf{X}}(E(U_i U_j | \mathbf{X}, U_j) \cdot U_j | \mathbf{x}) && \text{por Teorema esperanzas iteradas} \\ &= E_{U_i, U_j, \mathbf{X}}(E(U_i | \mathbf{X}, U_j) \cdot U_j | \mathbf{x}) && \text{por linealidad de la esperanza condicional} \\ &= E_{U_i, U_j, \mathbf{X}}(E(U_i | X_i) \cdot U_j | \mathbf{x}) && \text{por ser m.a.s.} \\ &= E_{U_i, X_i}(U_i | x_i) E_{U_j, X_j}(U_j | x_j) && \text{por ser m.a.s.} \end{aligned}$$

Ejercicio 2

**Ejercicio 3.**

1.

$$\sum \widehat{e}_n = \sum (y_n - \widehat{y}_n) = \sum (y_n - \widehat{a} - \widehat{b}x_n) = \sum y_n - N \underbrace{\left[ \frac{\sum y_n}{N} - \widehat{b} \frac{\sum x_n}{N} \right]}_{\widehat{a}} - \widehat{b} \sum x_n = 0$$

2.

$$\begin{aligned} \sum x_n \widehat{e}_n &= \sum [x_n(y_n - \widehat{a} - \widehat{b}x_n)] \\ &= \sum x_n y_n - \sum x_n \widehat{a} - \sum \widehat{b} x_n^2 \\ &= \sum x_n y_n - \sum \left( \frac{\sum x_n y_n}{N} - \widehat{b} \frac{\sum x_n^2}{N} \right) - \sum \widehat{b} x_n^2 \\ &= \sum x_n y_n - \sum \left( \frac{\sum x_n y_n}{N} \right) - \sum \left( \widehat{b} \frac{\sum x_n^2}{N} \right) - \sum \widehat{b} x_n^2 \\ &= \sum x_n y_n - N \left( \frac{\sum x_n y_n}{N} \right) - N \left( \widehat{b} \frac{\sum x_n^2}{N} \right) - \sum \widehat{b} x_n^2 \\ &= 0 \end{aligned}$$

Ejercicio 3

**Ejercicio 5.** De 3.2 en la página 10 sabemos que

$$\widehat{a} = \bar{y} - \widehat{b} \bar{x} = \frac{1}{n} \sum y_t - \widehat{b} \frac{1}{n} \sum x_t.$$

Por lo tanto el estimador condicionado es

$$\widehat{a} | \mathbf{x} = \frac{1}{n} \sum Y_n - \left( \widehat{b} | \mathbf{x} \right) \frac{1}{n} \sum x_t$$

cuya esperanza es

$$\begin{aligned} E_{U, \mathbf{X}}(\widehat{a} | \mathbf{x}) &= \frac{1}{n} \sum E_{U, \mathbf{X}}(Y_n | \mathbf{x}) - E_{U, \mathbf{X}}(\widehat{b} | \mathbf{x}) \frac{1}{n} \sum x_t \\ &= \frac{1}{n} \sum E_{U, \mathbf{X}}(Y_n | \mathbf{x}) - b \frac{1}{n} \sum x_t && \text{puesto que } \widehat{b} \text{ es insesgado} \\ &= \frac{1}{n} \sum E_{U, \mathbf{X}}(a + bx_t + U_n | \mathbf{x}) - b \frac{1}{n} \sum x_t && \text{sustituyendo } Y_n \\ &= \frac{1}{n} \sum a + b \frac{1}{n} \sum x_t + \frac{1}{n} E_{U, \mathbf{X}}(U_n | \mathbf{x}) - b \frac{1}{n} \sum x_t && \text{cancelando términos} \\ &= a + \frac{1}{n} E_{U, \mathbf{X}}(U_n | \mathbf{x}) = a. \end{aligned}$$

Ejercicio 5

**Ejercicio 6.** Novales (véase 1997, capítulo 13)

Ejercicio 6

**Ejercicio 7.** Ya que

$$\begin{aligned} E_{U, \mathbf{X}}(Y_n | \mathbf{x}) &= E_{U, \mathbf{X}}(a + bx_t + U_n | \mathbf{x}) \\ &= a + bx_t + E_{U, \mathbf{X}}(U_n | \mathbf{x}) && \text{pues } a, b, \text{ y } x_t \text{ son ctes} \\ &= a + bx_t && \text{por el supuesto: } E_{U, \mathbf{X}}(U_n | \mathbf{x}) = 0 \end{aligned}$$

## Ejercicio 7

**Ejercicio 8(a)**

1. Por una parte:

$$\hat{b} = \frac{s_{xy}}{s_x^2} = \frac{84}{56} = 1.5$$

por otra, las medias muestrales son

$$\bar{x} = \frac{\sum x_i}{8} = \frac{40}{8} = 5; \quad \bar{y} = \frac{\sum y_i}{8} = \frac{56}{8} = 7;$$

por lo que

$$\hat{a} = \bar{y} - \hat{b} \cdot \bar{x} = 7 - 1.5 \cdot 5 = -0.5.$$

□

**Ejercicio 8(b)** Según el modelo estimado, una empresa que incurra en unos gastos de 3 debería tener unos beneficios de

$$\hat{y} = \hat{a} + \hat{b}x = -0.5 + 1.5 \cdot 3 = 4$$

□

**Ejercicio 8(c)** Los residuos de la empresa E serán:

$$y_E - \hat{y}_E = y_E - (\hat{a} + \hat{b}x_E) = 8 - (-0.5 + 1.5 \cdot 5) = 8 - 7 = 1$$

y los de la empresa F:

$$y_F - \hat{y}_F = y_F - (\hat{a} + \hat{b}x_F) = 9 - (-0.5 + 1.5 \cdot 7) = 9 - 10 = -1.$$

Puesto que

$$\hat{y} = E_{U|\mathbf{x}_{F>}}(\mathbf{Y} | \mathbf{x}_{F>}),$$

un signo positivo para el residuo de cierta empresa significa que ésta ha logrado unos beneficios mayores que los esperados (dado su nivel de gasto en formación de personal,  $x$ ). Por el contrario, un residuo negativo significa que la empresa ha obtenido unos beneficios menores de los esperados por el modelo (dado su gasto en formación).

La comparación entre empresas con distinta inversión en formación no es apropiada para valorar los datos sobre beneficios (sólo lo es entre empresas con mismo nivel de gasto en formación). La empresa F tiene mayores beneficios que los de E, pero, dado su nivel de gasto en formación (7), estos beneficios deberían haber sido aún mayores (el valor esperado es 10).

□

**Ejercicio 8(d)** El estimador MCO se distribuye Normal con esperanza igual al verdadero valor de los parámetros estimados, y varianza desconocida.

- Buscamos los valores  $A$  y  $B$  tales que

$$P\left(A \leq \frac{\hat{b}_1 \mathbf{x} - b}{\sqrt{\frac{\hat{s}_e^2}{\sum (x_i - \bar{x})^2}}} \leq B\right) = (1 - \alpha)$$

Donde  $\frac{\hat{b}_1 \mathbf{x} - b}{\sqrt{\frac{\hat{s}_e^2}{\sum (x_i - \bar{x})^2}}}$  se distribuye como una  $t$  de Student con  $N - 2$  grados de libertad; 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,  $A = -2.447$ , y  $B = 2.447$ , y  $\hat{s}_e^2 = 6/(N - 2) = 1$ . Así pues, la estimación del intervalo de confianza de parámetro desconocido  $b$  es

$$IC_{0.95}^b(\mathbf{w}) = [1.5 \pm 2.447 \cdot \sqrt{1/56}]$$

□

**Ejercicio 8(e)** Las hipótesis son:

$$H_0 : b = 1$$

$$H_1 : b < 1$$

La región crítica de una sola cola es

$$RC = \left\{ \mathbf{x} \mid \frac{\widehat{b}_{|\mathbf{x}} - 1}{\sqrt{\frac{\widehat{s}_e^2}{\sum(x_i - \bar{x})^2}}} < k \right\},$$

donde  $k$  es el valor de la tabla para una  $t$  de Student de seis grados de libertad, ya que el estadístico de la parte izquierda de la desigualdad tiene dicha distribución. Para  $\alpha = 0.1$ , tenemos que  $k = t_{6, 0.1} = -1.44$ . Sustituyendo tenemos que

$$\frac{1.5 - 1}{\sqrt{1/56}} = 3.74 > k = t_{6, 0.1} = -1.44$$

por lo que no rechazamos  $H_0$ .

El  $p$ -valor es la probabilidad de

$$\begin{aligned} P(\widehat{b}_{|\mathbf{x}} \leq 1.5 \mid H_0) &= P\left( \frac{\widehat{b}_{|\mathbf{x}} - b}{\sqrt{\frac{\widehat{s}_e^2}{\sum(x_i - \bar{x})^2}}} \leq \frac{1.5 - b}{\sqrt{\frac{\widehat{s}_e^2}{\sum(x_i - \bar{x})^2}}} \mid H_0 \right) \\ &= P\left( W \leq \frac{1.5 - 1}{\sqrt{1/56}} = 3.74 \right) \simeq 0.999 \end{aligned}$$

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

□