

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.

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



Algunos derechos reservados. Esta obra está bajo una licencia Reconocimiento-Compartirlgua 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

Índice	3
1. Introducción	6
1.1. El punto de vista estadístico: Regresión como descomposición ortogonal	6
1.2. El punto de vista del Análisis Económico: Regresión como modelo explicativo	8
2. Modelo Clásico de Regresión Lineal	10
2.1. Tres primeros supuestos en el Modelo Clásico de Regresión Lineal	12
3. Estimación MCO (Mínimos Cuadrados Ordinarios)	21
4. Propiedades algebraicas de la estimación MCO	34

5. Propiedades estadísticas del estimador MCO $\hat{\beta}$ x	36
6. Distribución de los estimadores MCO bajo Normalidad	42
7. Estimación por intervalos	47
8. Contrastación de hipótesis	50
9. Ejercicios	53
10. Trasparencias	62
11. Bibliografía	63

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 \(1997\)](#)¹:

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

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

1. Introducción

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



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



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

donde el conjunto de información es $\mathcal{D} : (\mathbf{X} = \mathbf{x})$



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

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



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

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

lectura estadística: de izquierda a derecha.

Siempre es cierta. No implica causalidad ni conclusiones teóricas



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

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

lectura estadística: de izquierda a derecha.

Siempre es cierta. No implica causalidad ni conclusiones teóricas

lectura teórica: de derecha a izquierda.

Interpretación puede ser falsa (regresiones espurias)

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



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

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

2. Modelo Clásico de Regresión Lineal



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



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)**



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!)

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



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



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

$$Y_1 = \beta_1 + \beta_2 X_1 + U_1$$



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

por lo tanto

$$Y_1 = \beta_1 + \beta_2 X_1 + U_1$$

$$Y_2 = \beta_1 + \beta_2 X_2 + U_2$$



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

por lo tanto

$$Y_1 = \beta_1 + \beta_2 X_1 + U_1$$

$$Y_2 = \beta_1 + \beta_2 X_2 + U_2$$

...

$$Y_N = \beta_1 + \beta_2 X_N + U_N$$

Ejemplo 1. [función de consumo:]

$$CON_n = \beta_1 + \beta_2 RD_n + U_n$$



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

es decir

$$\mathbb{E}_{U|\mathbf{x}}(\mathbf{U} | \mathbf{x}) = \begin{bmatrix} \mathbb{E}_{U_1|\mathbf{x}}(U_1 | \mathbf{x}) \\ \mathbb{E}_{U_2|\mathbf{x}}(U_2 | \mathbf{x}) \\ \vdots \\ \mathbb{E}_{U_n|\mathbf{x}}(U_n | \mathbf{x}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathbb{E}_{U_n|\mathbf{x}}(U_n | \mathbf{x}) \equiv \mathbb{E}_{U_n|\mathbf{x}}(U_n | x_1, \dots, x_N)$$

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

Ejemplo 2. [función de consumo:]

Estricta exogeneidad implica que para el individuo i -ésimo

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



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

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



■ **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

$$\text{Var}(\mathbf{Y}) = \text{Var} \left(\begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix} \right) \equiv E(\mathbf{Y}\mathbf{Y}') - E(\mathbf{Y})E(\mathbf{Y}')$$

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

$$\text{Var}(\mathbf{Y}) = \text{Var} \left(\begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix} \right) \equiv E(\mathbf{Y}\mathbf{Y}') - E(\mathbf{Y})E(\mathbf{Y}')$$

$$= \begin{bmatrix} E(Y_1^2) & E(Y_1 Y_2) & \cdots & E(Y_1 Y_N) \\ & E(Y_2^2) & \cdots & E(Y_2 Y_N) \\ & & \ddots & \vdots \\ & & & E(Y_N^2) \end{bmatrix} - \begin{bmatrix} [E(Y_1)]^2 & E(Y_1)E(Y_2) & \cdots & E(Y_1)E(Y_N) \\ & [E(Y_2)]^2 & \cdots & E(Y_2)E(Y_N) \\ & & \ddots & \vdots \\ & & & [E(Y_N)]^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{Y_1}^2 & \sigma_{Y_1 Y_2} & \cdots & \sigma_{Y_1 Y_N} \\ & \sigma_{Y_2}^2 & \cdots & \sigma_{Y_2 Y_N} \\ & & \ddots & \vdots \\ & & & \sigma_{Y_N}^2 \end{bmatrix}$$

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

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

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

$$\begin{aligned} \text{Var}(\mathbf{a}\mathbf{Y}) &= E(\mathbf{a}\mathbf{Y}\mathbf{Y}'\mathbf{a}') - E(\mathbf{a}\mathbf{Y})E(\mathbf{Y}'\mathbf{a}') \\ &= \mathbf{a} [E(\mathbf{Y}\mathbf{Y}') - E(\mathbf{Y})E(\mathbf{Y}')] \mathbf{a}' \\ &= \mathbf{a} \text{Var}(\mathbf{Y}) \mathbf{a}' \end{aligned}$$

sacando factores comunes



$$\begin{aligned}\text{Var}_{\mathbf{U}|\mathbf{X}}(\mathbf{U} | \mathbf{x}) &= \begin{pmatrix} \text{Var}_{U_1|\mathbf{X}}(U_1 | \mathbf{x}) & \dots & \text{Cov}_{U_1 U_n|\mathbf{X}}(U_1, U_n | \mathbf{x}) \\ \vdots & \ddots & \vdots \\ \text{Cov}_{U_n U_1|\mathbf{X}}(U_n, U_1 | \mathbf{x}) & \dots & \text{Var}_{U_n|\mathbf{X}}(U_n | \mathbf{x}) \end{pmatrix} \\ &= \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} = \sigma^2 \mathbf{I}_{[n \times n]}\end{aligned}$$

3. Estimación MCO (Mínimos Cuadrados Ordinarios)



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

- número de observaciones ≥ 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 $\text{Var}(\mathbf{X}_n) \neq 0$.



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



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

$$\hat{b} = \frac{s_{xy}}{s_x^2} \quad (3.1)$$

y

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

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



Las perturbaciones U_n no son observables



Las perturbaciones U_n no son observables

Pero las podemos estimar sustituyendo β 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” a 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	Coefficiente	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

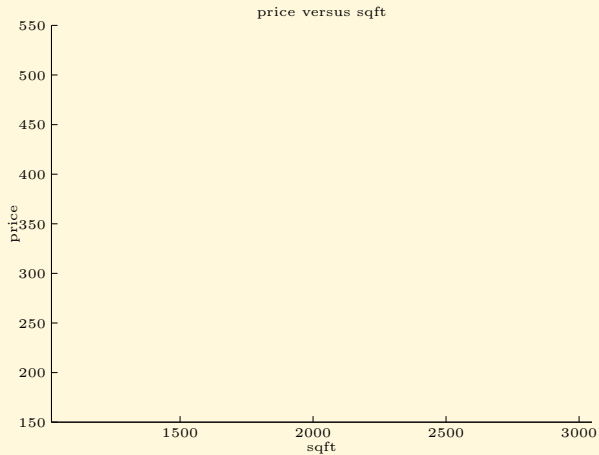
Por lo tanto, el precio de venta *esperado* de un piso con una superficie de 1800 pies cuadrados, $E_{Y|X}(Y|1800)$, será de

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

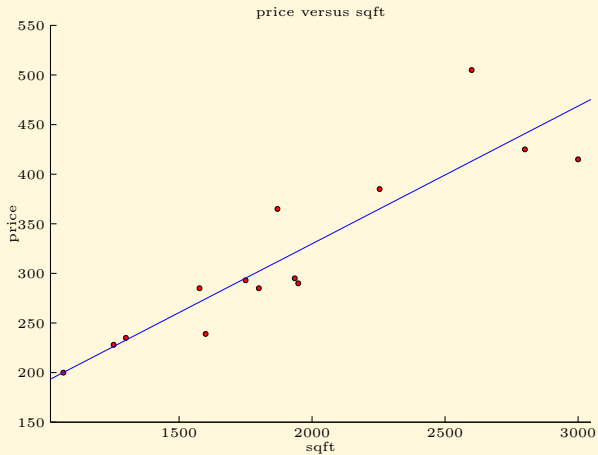
sin embargo en la muestra hay un piso con dicha superficie cuyo precio es $y_7 = 285$. Esta discrepancia es el error de estimación, que es una estimación de la perturbación (o factor desconocido) U_7 (el error \hat{e}_7 puede deberse por tanto a que dicho piso esta en una mala situación, dispone de pocos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

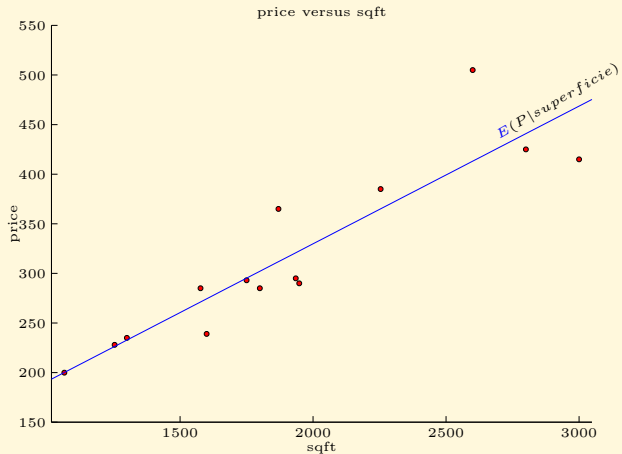
t	Precio	Superficie	Precio estimado $E(P \text{superficie})$	Error \hat{e}
1	199.9	1065	200.1200	-0.22000
2	228.0	1254	226.3438	1.65619
3	235.0	1300	232.7263	2.27368
4	285.0	1577	271.1602	13.83984
5	239.0	1600	274.3514	-35.35142
6	293.0	1750	295.1640	-2.16397
7	285.0	1800	302.1015	-17.10148
8	365.0	1870	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

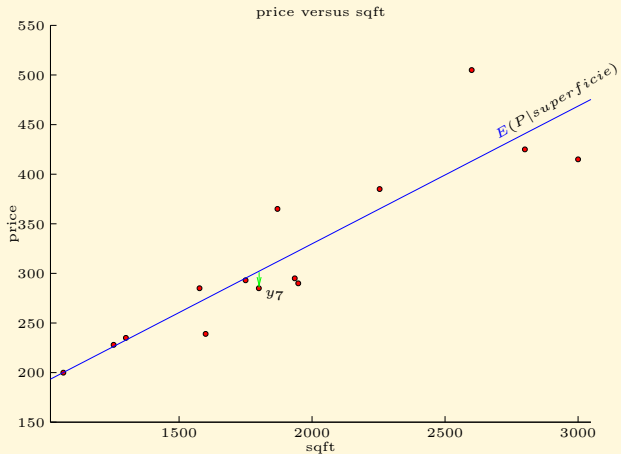
Cuadro 2: Superficie (en pies al cuadrado), precio de venta (en miles de dólares), precio estimado, y errores estimados.

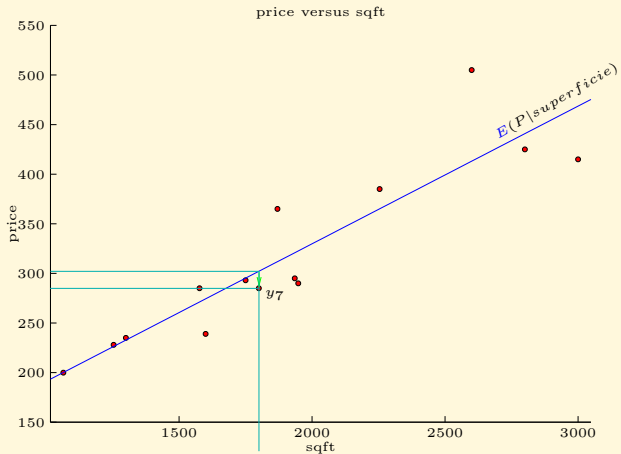


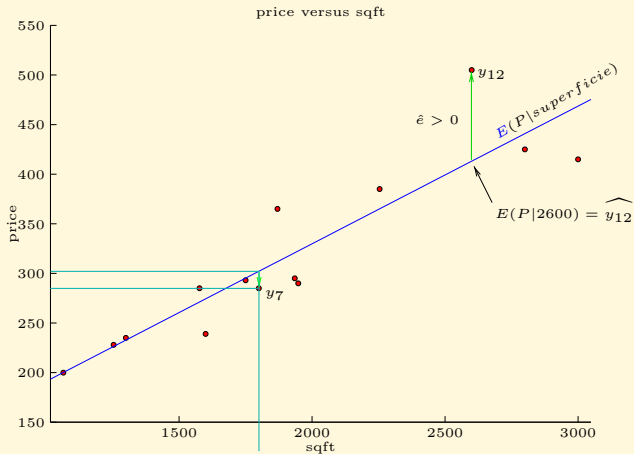


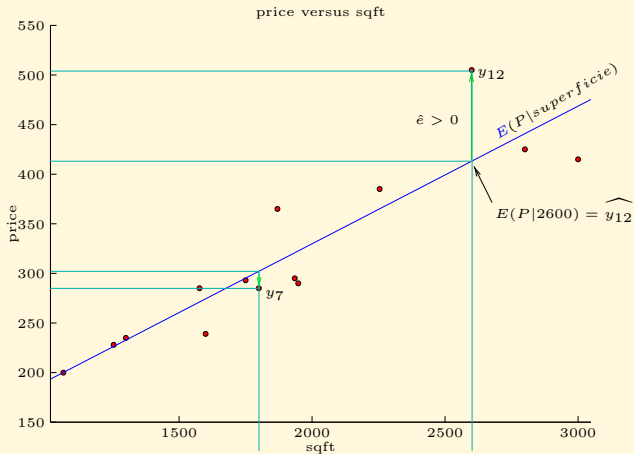












GNU Gretl (este ejemplo)

Continuación del ejemplo “precio de las viviendas” en la página~41

4. Propiedades algebraicas de la estimación MCO



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

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

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

Así pues, de 3.1 en la página~24 resulta

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

Así pues, de 3.1 en la página~24 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,

$$\hat{b} = \sum_n m_n y_n, \tag{5.1}$$

donde

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



Estimaciones:

$$\hat{b} = \sum m_n y_n,$$

Por tanto, \hat{b} es combinación lineal de los datos y_n ;

**Estimaciones:**

$$\hat{b} = \sum m_n y_n,$$

Por tanto, \hat{b} es combinación lineal de los datos y_n ; entonces \hat{a} también lo es (véase 3.2 en la página ~24).

**Estimaciones:**

$$\hat{b} = \sum m_n y_n,$$

Por tanto, \hat{b} es combinación lineal de los datos y_n ; entonces \hat{a} también lo es (véase 3.2 en la página ~24).

Estimadores:

$$\hat{b}_{|x} = \sum m_n Y_n.$$

y

$$\hat{a}_{|x} = \frac{\sum Y_n}{N} - (\hat{b}_{|x}) \cdot \bar{x}.$$

**Estimaciones:**

$$\hat{b} = \sum m_n y_n,$$

Por tanto, \hat{b} es combinación lineal de los datos y_n ; entonces \hat{a} también lo es (véase 3.2 en la página ~24).

Estimadores:

$$\hat{b}_{|x} = \sum m_n Y_n.$$

y

$$\hat{a}_{|x} = \frac{\sum Y_n}{N} - (\hat{b}_{|x}) \cdot \bar{x}.$$

Así que podemos calcular los momentos de los estimadores



Estimaciones:

$$\hat{b} = \sum m_n y_n,$$

Por tanto, \hat{b} es combinación lineal de los datos y_n ; entonces \hat{a} también lo es (véase 3.2 en la página ~24).

Estimadores:

$$\hat{b}_{|x} = \sum m_n Y_n.$$

y

$$\hat{a}_{|x} = \frac{\sum Y_n}{N} - (\hat{b}_{|x}) \cdot \bar{x}.$$

Así que podemos calcular los momentos de los estimadores

Análogamente $\widehat{y_{n|x}} = \widehat{a_{|x}} + (\widehat{b_{|x}}) \cdot x_n;$ y $\widehat{e_{n|x}} = Y_n - \widehat{y_{|x}}.$



Por 5.1 en la página ~37 sabemos que $\hat{b} \mid \mathbf{x} = \sum m_t Y_n$.



Por 5.1 en la página ~37 sabemos que $\hat{b} \mid \mathbf{x} = \sum m_t Y_n$.

Entonces,

$$\hat{b} \mid \mathbf{x} = \sum m_t (a + bx_t + U_n)$$



Por 5.1 en la página ~37 sabemos que $\hat{b} \mid \mathbf{x} = \sum m_t Y_n$.

Entonces,

$$\begin{aligned}\hat{b} \mid \mathbf{x} &= \sum m_t (a + bx_t + U_n) \\ &= a \sum m_t + b \sum m_t x_t + \sum m_t U_n\end{aligned}$$



Por 5.1 en la página ~37 sabemos que $\hat{b} \mid \mathbf{x} = \sum m_t Y_n$.

Entonces,

$$\begin{aligned}\hat{b} \mid \mathbf{x} &= \sum m_t (a + bx_t + U_n) \\ &= a \sum m_t + b \sum m_t x_t + \sum m_t U_n = b + \sum m_t U_n\end{aligned}\quad (5.2)$$

Por 5.1 en la página ~37 sabemos que $\hat{b} \mid \mathbf{x} = \sum m_t Y_n$.

Entonces,

$$\begin{aligned}\hat{b} \mid \mathbf{x} &= \sum m_t (a + bx_t + U_n) \\ &= a \sum m_t + b \sum m_t x_t + \sum m_t U_n = b + \sum m_t U_n\end{aligned}\quad (5.2)$$

y

$$E_{U_n \mid \mathbf{x}}(\hat{b} \mid \mathbf{x}) = b + \sum m_t E_{U_n \mid \mathbf{x}}(U_n \mid \mathbf{x}) = b.$$



Por 5.1 en la página ~37 sabemos que $\hat{b} \mid \mathbf{x} = \sum m_t Y_n$.

Entonces,

$$\begin{aligned} \hat{b} \mid \mathbf{x} &= \sum m_t (a + bx_t + U_n) \\ &= a \sum m_t + b \sum m_t x_t + \sum m_t U_n = b + \sum m_t U_n \end{aligned} \quad (5.2)$$

y

$$E_{U_n \mid \mathbf{x}} (\hat{b} \mid \mathbf{x}) = b + \sum m_t E_{U_n \mid \mathbf{x}} (U_n \mid \mathbf{x}) = b.$$

Por tanto el estimador es **insesgado** (lo mismo ocurre para \hat{a})



Además

$$\text{Var}(\hat{b} \mid \mathbf{x}) = \text{Var}\left(b + \sum m_t U_n\right) \quad \text{de (5.2)}$$



Además

$$\begin{aligned}\text{Var}(\hat{b} \mid \mathbf{x}) &= \text{Var}\left(b + \sum m_t U_n\right) && \text{de (5.2)} \\ &= \sum \text{Var}(m_t U_n) && \text{pues Cov}(U_i, U_j) = 0 \text{ si } i \neq j\end{aligned}$$

Además

$$\begin{aligned}\text{Var}(\hat{b} \mid \mathbf{x}) &= \text{Var}\left(b + \sum m_t U_n\right) && \text{de (5.2)} \\ &= \sum \text{Var}(m_t U_n) && \text{pues Cov}(U_i, U_j) = 0 \text{ si } i \neq j \\ &= \sum m_t^2 \text{Var}(U_n)\end{aligned}$$

Además

$$\begin{aligned}\text{Var}(\hat{b} \mid \mathbf{x}) &= \text{Var}\left(b + \sum m_t U_n\right) && \text{de (5.2)} \\ &= \sum \text{Var}(m_t U_n) && \text{pues Cov}(U_i, U_j) = 0 \text{ si } i \neq j \\ &= \sum m_t^2 \text{Var}(U_n) \\ &= \sigma^2 \sum m_t^2\end{aligned}$$



Además

$$\begin{aligned}\text{Var}(\hat{\mathbf{b}} \mid \mathbf{x}) &= \text{Var}\left(b + \sum m_t U_n\right) && \text{de (5.2)} \\ &= \sum \text{Var}(m_t U_n) && \text{pues Cov}(U_i, U_j) = 0 \text{ si } i \neq j \\ &= \sum m_t^2 \text{Var}(U_n) \\ &= \sigma^2 \sum m_t^2 \\ &= \frac{\sigma^2}{T s_{\mathbf{x}}^2} && (5.3)\end{aligned}$$

Además

$$\begin{aligned}
 \text{Var}(\hat{b} \mid \mathbf{x}) &= \text{Var}\left(b + \sum m_t U_n\right) && \text{de (5.2)} \\
 &= \sum \text{Var}(m_t U_n) && \text{pues Cov}(U_i, U_j) = 0 \text{ si } i \neq j \\
 &= \sum m_t^2 \text{Var}(U_n) \\
 &= \sigma^2 \sum m_t^2 \\
 &= \frac{\sigma^2}{T s_x^2} && (5.3)
 \end{aligned}$$

Y

$$\text{Var}(\hat{a} \mid \mathbf{x}) = \frac{\sigma^2 \sum x^2}{N s_x^2}; \quad \text{Cov}(\hat{a} \mid \mathbf{x}, \hat{b} \mid \mathbf{x}) = \frac{-\sigma^2 \cdot \bar{x}}{T \cdot s_x^2} \quad (5.4)$$

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

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

$$\text{Dt}(\hat{a}_{|x}) = \sqrt{\sigma^2 \cdot (9.1293e - 01)} = \sqrt{\frac{\sigma^2 \bar{x}^2}{T \cdot s_x^2}}$$
$$\text{Dt}(\hat{b}_{|x}) = \sqrt{\sigma^2 \cdot (2.3044e - 07)} = \sqrt{\frac{\sigma^2}{T \cdot s_x^2}}$$

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

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

$$\text{Dt}(\hat{a}_{|x}) = \sqrt{\sigma^2 \cdot (9.1293e - 01)} = \sqrt{\frac{\sigma^2 \bar{x}^2}{T \cdot s_x^2}}$$
$$\text{Dt}(\hat{b}_{|x}) = \sqrt{\sigma^2 \cdot (2.3044e - 07)} = \sqrt{\frac{\sigma^2}{T \cdot s_x^2}}.$$

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

Continuación del ejemplo “precio de las viviendas” en la página ~48

6. Distribución de los estimadores MCO bajo Normalidad



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

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



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.

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 Y_n , tienen distribución Normal.

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 Y_n , tienen distribución Normal.

Por tanto

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



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



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

Así pues,

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

y

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



El parámetro σ^2 es desconocido T0.7

El parámetro σ^2 es desconocido T0.7

La cuasivarianza de \hat{e}

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

es estimador insesgado de σ^2 .

(veáse [Novales, 1997](#), Apéndice 13.A)



sustituyendo σ^2 por su estimador, $\widehat{s_e^2}$, tenemos el estadístico \mathcal{T} del parámetro

$$\frac{\widehat{a} - a}{\sqrt{\frac{\widehat{s_e^2}}{\sum (x_i - \bar{x})^2}}} = \mathcal{T}^a \sim t_{N-2}; \quad \frac{\widehat{b} - b}{\sqrt{\frac{\widehat{s_e^2}}{N} \frac{\sum x_t^2}{\sum (x_t - \bar{x})^2}}} = \mathcal{T}^b \sim t_{N-2};$$

7. Estimación por intervalos

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

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

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

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

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

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

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

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

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

$$\begin{aligned} Dt(\widehat{a}_{|x}) &= \sqrt{\sigma^2 \cdot (9.1293e - 01)} = \sqrt{\frac{\sigma^2 \overline{x^2}}{N \cdot s_x^2}} \\ Dt(\widehat{b}_{|x}) &= \sqrt{\sigma^2 \cdot (2.3044e - 07)} = \sqrt{\frac{\sigma^2}{N \cdot s_x^2}}. \end{aligned}$$

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

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

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

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

Véase los resultados de estimación en el ejemplo del precio de las viviendas (página 29).

No conocemos $\sigma_{U_n}^2$, pero podemos sustituirla por la la cuasi-varianza de los errores estimados; puesto que $\widehat{s}_e^2 = \frac{\widehat{e}' \widehat{e}}{N-2} = \frac{18273.6}{14-2} = 1522.8$:

$$\widehat{Dt}(\widehat{a}_{|x}) = \sqrt{(1522.8) \cdot (9.1293e - 01)} = \sqrt{\frac{(1522.8) \cdot \overline{x^2}}{N \cdot s_x^2}} = 37.285;$$

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

Véase los resultados de estimación en el ejemplo del precio de las viviendas (página 29).

$$\text{Por otra parte, } \widehat{Cov}(\widehat{a}_{|x}, \widehat{b}_{|x}) = (1522.8) * (-4.4036e - 04) = \frac{-\widehat{s}_e^2 \cdot \overline{x}}{N \cdot s_x^2} = -0.671$$

No conocemos $\sigma_{U_n}^2$, pero podemos sustituirla por la la cuasi-varianza de los errores estimados; puesto que $\widehat{s_e^2} = \frac{\widehat{e}' \widehat{e}}{N-2} = \frac{18273.6}{14-2} = 1522.8$:

$$\widehat{Dt}(\widehat{a}_{|x}) = \sqrt{(1522.8) \cdot (9.1293e - 01)} = \sqrt{\frac{(1522.8) \cdot \overline{x^2}}{N \cdot s_x^2}} = 37.285;$$

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

Véase los resultados de estimación en el ejemplo del precio de las viviendas (página 29).

$$\text{Por otra parte, } \widehat{Cov}(\widehat{a}_{|x}, \widehat{b}_{|x}) = (1522.8) * (-4.4036e - 04) = \frac{-\widehat{s_e^2} \cdot \overline{x}}{N \cdot s_x^2} = -0.671$$

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 la página ~51

8. Contrastación 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 significativad 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

$$\frac{52.351}{37.285} = 1.4041$$

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

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 %.

$$H_0 : b = 0.15$$

$$H_1 : b < 0.15$$

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

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

sustituyendo lo que hemos calculado, tenemos que

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

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

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

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

EJERCICIO 1. 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.

Solución: Ya que

$$\begin{aligned} E_{U_n|\mathbf{x}}(Y_n | \mathbf{x}) &= E_{U_n|\mathbf{x}}(a + bx_t + U_n | \mathbf{x}) \\ &= a + bx_t + E_{U_n|\mathbf{x}}(U_n | \mathbf{x}) \\ &= a + bx_t \end{aligned}$$

pues a , b , y x_t son ctes

por el supuesto: $E_{U_n|\mathbf{x}}(U_n | \mathbf{x}) = 0$

Ejercicio 1

EJERCICIO 2. (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 σ^2 es la varianza de U_i , cuyo valor es desconocido.

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 σ^2 es la varianza de U_i , cuyo valor es desconocido.

(a) Estime por MCO los parámetros a y b del modelo.

Solución:

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 σ^2 es la varianza de U_i , cuyo valor es desconocido.

(a) Estime por MCO los parámetros a y b del modelo.

Solución:

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 2

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 2

(b) ¿Cuál es el beneficio esperado para una empresa que incurriera en unos gastos de formación de personal de 3?

Solución:

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 2

(b) ¿Cuál es el beneficio esperado para una empresa que incurriera en unos gastos de formación de personal de 3?

Solución: Según el modelo estimado, una empresa que incurra en unos gastos de 3 debería tener unos beneficios de

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

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 2

- (b) ¿Cuál es el beneficio esperado para una empresa que incurriera en unos gastos de formación de personal de 3?

Solución: 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 2

- (c) 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.

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 2

- (b) ¿Cuál es el beneficio esperado para una empresa que incurriera en unos gastos de formación de personal de 3?

Solución: 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 2

- (c) 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.

Solución: 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:

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

Puesto que

$$\widehat{y} = E_{U|X_{F\triangleright}}(Y | \mathbf{x}_{F\triangleright}),$$

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 2

- (d) 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.

Solución:

y los de la empresa F:

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

Puesto que

$$\widehat{y} = E_{U|X_{F\triangleright}}(Y | \mathbf{x}_{F\triangleright}),$$

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 2

- (d) 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.

Solución: 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

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

Ejercicio 2

- 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

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

Ejercicio 2

- (e) 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”?

Solución:

Solución: Las hipótesis son:

$$H_0 : b = 1$$

$$H_1 : b < 1$$

La región crítica de una sola cola es

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

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

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

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

Ejercicio 2

10. Transparencias

Lista de Transparencias

- 1 Descomposición ortogonal y causalidad
- 2 Modelo de regresión
- 3 Modelo Clásico de Regresión Lineal
- 4 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 $\hat{\beta} \mid x$
- 16 Esperanza del estimador MCO $\hat{b} \mid x$
- 17 Varianzas y covarianzas de los estimadores MCO
- 18 Supuesto 5: Distribución Normal de las perturbaciones

- 19 Distribución del estimador MCO $\widehat{\beta}_j \mid x$
- 20 Estimación de la varianza residual
- 21 Distribución cuando la varianza de U es desconocida
- 22 Partes del temario

11. Bibliografía

- Novales, A. (1997). *Estadística y Econometría*. McGraw-Hill, Madrid, primera ed. ISBN 84-481-0798-5. 5, 45
- Peña, D. (2001). *Fundamentos de Estadística*. Alianza Editorial, Madrid. ISBN 84-206-8696-4. 5
- Peña, D. (2002). *Regresión y diseño de experimentos*. Alianza Editorial, Madrid. ISBN 84-206-8695-6. 5
- Peña, D. y Romo, J. (1997). *Introducción a la Estadística para la Ciencias Sociales*. McGraw-Hill, Madrid. ISBN 84-481-1617-8. 5
- Ramanathan, R. (1998). *Introductory Econometrics with Applications*. Harcourt College Publisher, Orlando. 26

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