T3. El modelo lineal básico

Ana J. López y Rigoberto Pérez

Dpto Economía Aplicada. Universidad de Oviedo

Curso 2010-2011

Índice

El modelo lineal básico

Competencias

El análisis del modelo lineal básico sigue el mismo esquema descrito para el modelo lineal simple, si bien en este caso todos los desarrollos se presentan en notación matricial.

Este tema constituye el nucleo central de la asignatura y su objetivo es capacitar a los estudiantes para:

- Especificar un modelo lineal y sus hipótesis de trabajo en forma matricial.
- Deducir e interpretar los estimadores mínimo cuadráticos y máximo verosímiles.
- Analizar la capacidad explicativa del modelo estimado, comparándolo con otras alternativas.
- Realizar e interpretar los contrastes de significación individual.
- Enunciar y contrastar restricciones sobre los parámetros del modelo.
- Realizar predicciones ex post y llevar a cabo su evaluación.
- Manejar las opciones de Gretl relativas a estimación, contraste y predicción, interpretando adecuadamente las salidas correspondientes.

Modelo lineal básico

$$
\begin{aligned}
Y_1 &= \beta_1 + \beta_2 X_{21} + \cdots + \beta_k X_{k1} + u_1 \\
Y_2 &= \beta_1 + \beta_2 X_{22} + \cdots + \beta_k X_{k2} + u_2 \\
\vdots & \quad \vdots \quad \dots \quad \ddots \quad \ddots \quad \ddots \quad \dots \vdots \\
Y_n &= \beta_1 + \beta_2 X_{2n} + \cdots + \beta_k X_{kn} + u_n
\end{aligned}
$$

$$
\boxed{\underset{(n\times 1)}{\mathbf{y}} = \underset{(n\times k)}{\mathbf{X}} \underset{(k\times 1)}{\boldsymbol{\beta}} + \underset{(n\times 1)}{\mathbf{u}}}
$$

$$
\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}; \quad \mathbf{X} = \begin{pmatrix} 1 & X_{21} & \cdots & X_{k1} \\ 1 & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & \cdots & X_{kn} \end{pmatrix}; \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}; \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}
$$

Modelo lineal básico: hipótesis

$$\boxed{\mathbf{y} = \mathbf{X}\beta + \mathbf{u}}$$

Hipótesis sobre u

Vector de perturbaciones aleatorias Normal con esperanza nula y matriz de varianzas-covarianzas escalar

Hipótesis sobre X

Matriz de regresores no estocástica y de rango k

Hipótesis sobre β

Vector de parámetros estructurales

Hipótesis básicas sobre $\mathbf{u}$

$$E(\mathbf{u}) = \mathbf{0}$$
$$Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \sigma^2 \mathbf{I}_n$$

$$Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \begin{pmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2) \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}$$

$$\boxed{\begin{array}{lcr} Var(u_i) = E(u_i^2) = \sigma^2 & \Rightarrow & \text{Homocedasticidad} \\ Cov(u_i, u_j) = E(u_iu_j) = 0 & \Rightarrow & \text{No autocorrelación} \end{array}}$$

Hipótesis de normalidad

- $\mathbf{u}$ sigue una distribución normal multivariante: $\mathbf{u} \approx \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$
- $\mathbf{y}$ sigue una distribución normal multivariante: $\mathbf{y} \approx \mathcal{N}(\mathbf{X}\beta, \sigma^2 \mathbf{I}_n)$

Hipótesis básicas

Hipótesis sobre los regresores

- La matriz de regresores $\mathbf{X}$ es no estocástica
- La matriz de los regresores tiene rango k, $\rho(\mathbf{X}) = k$
- Las columnas de $\mathbf{X}$ son linealmente independientes
- $n > k$

Hpótesis sobre los parámetros

- $\boldsymbol{\beta}$ es un vector fijo

Estimación de los parámetros

Modelo teórico: $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$

Vector de estimadores de los parámetros: $\hat{\beta}$

Vector de valores estimados: $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$

Vector de residuos: $\hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\beta}$

Suma de cuadrados de residuos

$$\sum_{i=1}^{n} \hat{u}_i^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}} = \left(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n\right) \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{pmatrix}$$

Estimadores mínimo cuadráticos (MCO)

Función a minimizar:

$$\begin{aligned}\hat{\mathbf{u}}'\hat{\mathbf{u}} &= \left(\mathbf{y}-\mathbf{X}\hat{\beta}\right)'\left(\mathbf{y}-\mathbf{X}\hat{\beta}\right) = \mathbf{y}'\mathbf{y}-\hat{\beta}'\mathbf{X}'\mathbf{y}-\mathbf{y}'\mathbf{X}\hat{\beta}+\hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta} \\ &= \mathbf{y}'\mathbf{y}-2\hat{\beta}'\mathbf{X}'\mathbf{y}+\hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}\end{aligned}$$

Condición necesaria de extremo:

$$\begin{aligned}\frac{\partial\left(\hat{\mathbf{u}}'\hat{\mathbf{u}}\right)}{\partial\hat{\boldsymbol{\beta}}}=0 \quad &\Rightarrow \quad \frac{\partial\left(\mathbf{y}'\mathbf{y}-2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}+\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}\right)}{\partial\hat{\boldsymbol{\beta}}}=0 \\ &\Rightarrow \quad -2\mathbf{X}'\mathbf{y}+2\mathbf{X}'\mathbf{X}\hat{\beta}=0\end{aligned}$$

Estimadores de mínimos cuadrados ordinarios (MCO):

$$\boxed{\hat{\boldsymbol{\beta}}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}}$$

Estimadores MCO

$$\boxed{\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}}$$

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{pmatrix} \begin{pmatrix} 1 & X_{21} & \cdots & X_{k1} \\ 1 & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & \cdots & X_{kn} \end{pmatrix} =$$

$$= \begin{pmatrix} n & \sum_{i=1}^{n} X_{2i} & \cdots & \sum_{i=1}^{n} X_{ki} \\ \sum_{i=1}^{n} X_{2i} & \sum_{i=1}^{n} X_{2i}^2 & \cdots & \sum_{i=1}^{n} X_{2i}X_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} X_{ki} & \sum_{i=1}^{n} X_{ki}X_{2i} & \cdots & \sum_{i=1}^{n} X_{ki}^2 \end{pmatrix}$$

Estimadores máximo verosímiles

Hipótesis de normalidad

- $\mathbf{u} \approx \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, $\mathbf{y} \approx \mathcal{N}(\mathbf{X}\beta, \sigma^2\mathbf{I})$

$$L(y_1, y_2, \cdots, y_n, \sigma^2, \beta) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}[(\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}\beta)]}$$

$$\ln L(y_1, y_2, \cdots, y_n, \sigma^2, \beta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}\beta)$$

$$\frac{\partial \ln L}{\partial \beta} = -\frac{1}{2\sigma^2}(-2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\beta) = 0$$

$$\Rightarrow \mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y} \Rightarrow \boxed{\hat{\beta} = (X'X)^{-1}X'y}$$

Estimadores máximo verosímiles

$$\boxed{\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}} \quad \Rightarrow \quad \boxed{\text{EMV=EMC}}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{(\mathbf{y}-\mathbf{X}\hat{\beta})'(\mathbf{y}-\mathbf{X}\hat{\beta})}{2\sigma^4} = 0$$

$$\Rightarrow \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{2\sigma^4} = \frac{n}{2\sigma^2} \Rightarrow \boxed{\hat{\sigma}^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n}}$$

$\hat{\sigma}^2$ es un estimador sesgado de σ^2

Estimación de parámetros con Gretl

$$\boxed{demanda = \underbrace{\beta_1}_{const} + \beta_2 renta + \beta_3 precio + u}$$

Modelo 1: MCO, usando las observaciones 1--25\Variable dependiente: demanc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	5177.58	313.027	16.5404	0.0000
renta	177.506	83.7362	2.1198	0.0455
precio	−6.90275	0.765520	−9.0171	0.0000

En forma matricial

- matrix **y**=demanda
- matrix **X**=const renta precio
- matrix **b**=inv(**X'X**)***X'*****y**

Propiedades descriptivas de los estimadores MCO y MV

- La suma de los residuos es nula: $\sum_{i=1}^{n} \hat{u}_i = 0$
- Las medias de valores estimados y observados coinciden: $\bar{Y} = \bar{\hat{Y}}$
- El hiperplano de regresión pasa por el punto denominado "centro de gravedad": $\left(\bar{X}_2, \ldots, \bar{X}_k, \bar{Y}\right)$
- Los momentos de segundo orden entre regresores y residuos son nulos: $\mathbf{X}'\hat{\mathbf{u}} = 0$
- Los momentos de segundo orden entre valores estimados y residuos son nulos: $\hat{\mathbf{y}}'\hat{\mathbf{u}} = 0$

Características de los estimadores MCO y MV

Linealidad

$$\hat{\beta} = \mathbf{A}\mathbf{y} \quad ; \quad \mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \mathbf{u}) = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$$

Esperanza

$$E\left(\hat{\beta}\right) = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u}) = \beta$$

Matriz de Varianzas-Covarianzas

$$Cov\left(\hat{\beta}\right) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

Matriz de Varianzas-Covarianzas de los estimadores

$$
\begin{aligned}
Cov\left(\hat{\boldsymbol{\beta}}\right) &= E\left[\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\right] \\
&= E\left[\left((\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{u}\right)\left((\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{u}\right)'\right] \\
&= (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'E(\mathbf{u}\mathbf{u}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}} = \boxed{\sigma^2(\mathbf{X}'\mathbf{X})^{-1}}
\end{aligned}
$$

$$
Cov(\hat{\boldsymbol{\beta}}) = \begin{pmatrix}
\sigma^2_{\hat{\beta}_1} & \sigma_{\hat{\beta}_1\hat{\beta}_2} & \cdots & \sigma_{\hat{\beta}_1\hat{\beta}_k} \\
\sigma_{\hat{\beta}_2\hat{\beta}_1} & \sigma^2_{\hat{\beta}_2} & \cdots & \sigma_{\hat{\beta}_2\hat{\beta}_k} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{\hat{\beta}_k\hat{\beta}_1} & \sigma_{\hat{\beta}_k\hat{\beta}_2} & \cdots & \sigma^2_{\hat{\beta}_k}
\end{pmatrix}
$$

Matriz de Varianzas-Covarianzas de los estimadores

$$Cov\left(\hat{\boldsymbol{\beta}}\right) = E\left[\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\right]$$

$$= E\left[\begin{pmatrix} \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_2 - \beta_2 \\ \vdots \\ \hat{\beta}_k - \beta_k \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 - \beta_1 & \hat{\beta}_2 - \beta_2 & \cdots & \hat{\beta}_k - \beta_k \end{pmatrix}\right]$$

$$= E\left[\begin{pmatrix} \left(\hat{\beta}_1 - \beta_1\right)^2 & \left(\hat{\beta}_1 - \beta_1\right)\left(\hat{\beta}_2 - \beta_2\right) & \cdots & \left(\hat{\beta}_1 - \beta_1\right)\left(\hat{\beta}_k - \beta_k\right) \\ \left(\hat{\beta}_2 - \beta_2\right)\left(\hat{\beta}_1 - \beta_1\right) & \left(\hat{\beta}_2 - \beta_2\right)^2 & \cdots & \left(\hat{\beta}_2 - \beta_2\right)\left(\hat{\beta}_k - \beta_k\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\hat{\beta}_k - \beta_k\right)\left(\hat{\beta}_1 - \beta_1\right) & \left(\hat{\beta}_k - \beta_k\right)\left(\hat{\beta}_2 - \beta_2\right) & \cdots & \left(\hat{\beta}_k - \beta_k\right)^2 \end{pmatrix}\right]$$

Matriz de Varianzas-Covarianzas de los estimadores

$$Cov\left(\hat{\boldsymbol{\beta}}\right) = \begin{pmatrix} E\left[\hat{\beta}_1 - \beta_1\right]^2 & E\left[\left(\hat{\beta}_1 - \beta_1\right)\left(\hat{\beta}_2 - \beta_2\right)\right] & \cdots & E\left[\left(\hat{\beta}_1 - \beta_1\right)\left(\hat{\beta}_k - \beta_k\right)\right] \\ E\left[\left(\hat{\beta}_2 - \beta_2\right)\left(\hat{\beta}_1 - \beta_1\right)\right] & E\left[\hat{\beta}_2 - \beta_2\right]^2 & \cdots & E\left[\left(\hat{\beta}_2 - \beta_2\right)\left(\hat{\beta}_k - \beta_k\right)\right] \\ \vdots & \vdots & \ddots & \vdots \\ E\left[\left(\hat{\beta}_k - \beta_k\right)\left(\hat{\beta}_1 - \beta_1\right)\right] & E\left[\left(\hat{\beta}_k - \beta_k\right)\left(\hat{\beta}_2 - \beta_2\right)\right] & \cdots & E\left[\hat{\beta}_k - \beta_k\right]^2 \end{pmatrix}$$

$$Cov(\hat{\boldsymbol{\beta}}) = \begin{pmatrix} \sigma^2_{\hat{\beta}_1} & \sigma_{\hat{\beta}_1\hat{\beta}_2} & \cdots & \sigma_{\hat{\beta}_1\hat{\beta}_k} \\ \sigma_{\hat{\beta}_2\hat{\beta}_1} & \sigma^2_{\hat{\beta}_2} & \cdots & \sigma_{\hat{\beta}_2\hat{\beta}_k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\hat{\beta}_k\hat{\beta}_1} & \sigma_{\hat{\beta}_k\hat{\beta}_2} & \cdots & \sigma^2_{\hat{\beta}_k} \end{pmatrix}$$

Propiedades de los estimadores MCO y MV

Ausencia de sesgo

$$E\left(\hat{\boldsymbol{\beta}}\right) = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u}) = \boldsymbol{\beta}$$

Consistencia: $\hat{\beta} \underset{n\to\infty}{\longrightarrow} \beta$

$$\text{lím}_{n\to\infty} P\left(\left|\hat{\boldsymbol{\beta}}_{\mathbf{n}} - \boldsymbol{\beta}\right| \geq \epsilon\right) = 0$$
$$\text{lím}_{n\to\infty} Cov\left(\hat{\boldsymbol{\beta}}_n\right) = 0$$

Eficiencia: Los estimadores son óptimos

- $Var\left(\tilde{\boldsymbol{\beta}}\right) \geq Var\left(\hat{\boldsymbol{\beta}}\right)$, $\tilde{\boldsymbol{\beta}}$ lineal e insesgado [Teorema de Gauss-Markov]
- $Var\left(\tilde{\boldsymbol{\beta}}\right) \geq Var\left(\hat{\boldsymbol{\beta}}\right)$, $\tilde{\boldsymbol{\beta}}$ insesgado, $\mathbf{u} \approx \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ [Teorema de Rao]

Estimación de la matriz de Varianzas-Covarianzas

Estimador insesgado de la varianza

$$S^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k}$$

Estimador de la matriz de varianzas-covarianzas

$$S^2(\hat{\beta}) = S^2(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} S^2_{\hat{\beta}_1} & S_{\hat{\beta}_1\hat{\beta}_2} & \cdots & S_{\hat{\beta}_1\hat{\beta}_k} \\ S_{\hat{\beta}_2\hat{\beta}_1} & S^2_{\hat{\beta}_2} & \cdots & S_{\hat{\beta}_2\hat{\beta}_k} \\ \vdots & \vdots & \ddots & \vdots \\ S_{\hat{\beta}_k\hat{\beta}_1} & S_{\hat{\beta}_k\hat{\beta}_2} & \cdots & S^2_{\hat{\beta}_k} \end{pmatrix}$$

Inferencias sobre el modelo lineal

Distribuciones probabilísticas

Inferencias sobre β

$$\left.\begin{array}{l} \mathbf{u} \approx N(\mathbf{0}, \sigma^2 \mathbf{I}) \\ \mathbf{y} \approx N(\mathbf{X}\beta, \sigma^2 \mathbf{I}) \end{array}\right\} \quad \Rightarrow \hat{\boldsymbol{\beta}} \approx N\left(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}\right)$$

$$d_{\hat{\beta}_j} = \frac{\hat{\beta}_j - \beta_j}{S\left(\hat{\beta}_j\right)} \approx t_{n-k}$$

Inferencias sobre σ^2

$$\mathbf{u} \approx N(\mathbf{0}, \sigma^2 \mathbf{I}) \ \ ; \ \ S^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k}$$

$$\frac{(n-k)S^2}{\sigma^2} = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{\sigma^2} \approx \chi^2_{n-k}$$

Intervalos de confianza para los coeficientes en Gretl

Gretl. Salida del modelo: Análisis → Matriz de covarianzas de los coeficientes
Análisis → Intervalos de confianza para los coeficientes

Matriz de covarianzas de los coeficientes de regresión:

const	renta	precio	
97985.6	-12744.1	-137.877	const
	7011.75	-22.7114	renta
		0.586021	precio

Para cada coeficiente β_j, el intervalo de confianza se obtiene como:

$$\boxed{\left[\hat{\beta}_j - k_\alpha S_{\hat{\beta}_j}\ ,\ \hat{\beta}_j + k_\alpha S_{\hat{\beta}_j}\right]}$$

VARIABLE	COEFICIENTE	INTERVALO DE CONFIANZA 95 %	
const	5177.58	4528.40	5826.76
renta	177.506	3.84761	351.164
precio	-6.90275	-8.49035	-5.31516

Contrastes asociados a un modelo

$$\mathbf{Y} = \beta_1 + \beta_2\mathbf{X}_2 + \beta_3\mathbf{X}_3 + \beta_4\mathbf{X}_4 + \beta_5\mathbf{X}_5 + \mathbf{u}$$

Contraste global

¿Tiene sentido el modelo planteado para **Y**?

$$\boxed{H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0}$$

Contrastes individuales

¿Explica la variable **X** el comportamiento de **Y**? $\boxed{H_0 : \beta_2 = 0}$

Contrastes parciales

¿Son admisibles ciertas restricciones sobre el modelo?

$$\boxed{H_0 : \beta_1 + \beta_2 = 4,\ \beta_3 = 2\beta_5}$$

Contraste global de significación

- **Hipótesis**:

$$\boxed{\begin{array}{l} H_0 : \ \beta_2 = \beta_3 = \cdots = \beta_k = 0 \\ H_1 : \ \exists j = 2, \cdots, k \ , \ \beta_j \neq 0 \end{array}}$$

Bajo la hipótesis nula el modelo no tiene sentido

- **Supuesto de trabajo**: $\mathbf{u} \approx N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad \Rightarrow \quad \hat{\boldsymbol{\beta}} \approx N(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1})$
- **Discrepancia**:

$$d^* = \frac{\dfrac{\left(\hat{\boldsymbol{\beta}}' \mathbf{X}`\mathbf{y} - n\bar{Y}^2\right)}{(k-1)}}{\dfrac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{(n-k)}} \approx F^{k-1}_{n-k}$$

- **Nivel crítico**: $\boxed{p = P\left(F^{k-1}_{n-k} \geq F^*\right)}$
- **CONCLUSIÓN**: Para discrepancias muestrales elevadas y p bajos se rechaza la hipótesis y se concluye que hay alguna variable significativa en el modelo propuesto

Análisis de varianza

Varianza total (VT)

$$\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2=\sum_{i=1}^{n}Y_i^2-n\bar{Y}^2=\boxed{\mathbf{y}'\mathbf{y}-n\bar{Y}^2}$$

Varianza explicada (VE)

$$\begin{aligned}\sum_{i=1}^{n}\left(\hat{Y}_i-\bar{Y}\right)^2 &= \sum_{i=1}^{n}\hat{Y}_i^2-n\bar{Y}^2=\hat{\mathbf{y}}'\hat{\mathbf{y}}-n\bar{Y}^2=\hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}-n\bar{Y}^2\\ &= \hat{\beta}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}-n\bar{Y}^2=\boxed{\hat{\beta}'\mathbf{X}'\mathbf{y}-n\bar{Y}^2}\end{aligned}$$

Varianza no explicada (VNE)

$$\sum_{i=1}^{n}\left(\hat{u}_i-\bar{\hat{u}}\right)^2=\sum_{i=1}^{n}\hat{u}_i^2=\boxed{\hat{\mathbf{u}}'\hat{\mathbf{u}}}$$

Contraste F de significación global

Variación	Expresión	g.l.	Ratio
Explicada	$\hat{\beta}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2$	k-1	$\dfrac{\hat{\beta}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2}{k-1}$
No explicada	$\hat{\mathbf{u}}'\hat{\mathbf{u}}$	n-k	$\dfrac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k}$
Total	$\mathbf{y}'\mathbf{y} - n\bar{Y}^2$	n-1	$\dfrac{\mathbf{y}'\mathbf{y} - n\bar{Y}^2}{n-1}$

Contraste global F: $$\boxed{\dfrac{\dfrac{\hat{\beta}'\mathbf{X}'\mathbf{y} - n\bar{Y}^2}{k-1}}{\dfrac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k}} \approx F^{k-1}_{n-k}}$$

Análisis de la varianza con Gretl

Gretl. Salida del modelo: Análisis → ANOVA

Análisis de Varianza

	Suma de cuadrados	gl	Media de cuadrados
Regresión	2.51818e+07	2	1.25909e+07
Residuo	6.70464e+06	22	304757
Total	3.18864e+07	24	1.3286e+06

$$R^2 = \frac{2{,}51818e+07}{3{,}18864e+07} = 0{,}789733$$

$$F(2,22) = \frac{1{,}25909e+07}{304757} = 41{,}3145 \quad [Valor\ p = 3{,}55e-08]$$

Análisis de bondad de un modelo

- Coeficiente de determinación:

$$R^2 = 1 - \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2} = \frac{\hat{\beta}'\mathbf{X}'\hat{\mathbf{y}} - n\bar{Y}^2}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2}$$

Se cumple: $0 \leq R^2 \leq 1$; $F_{n-k}^{k-1} = \left(\frac{R^2}{1-R^2}\right)\left(\frac{n-k}{k-1}\right)$

- Coeficiente de determinación corregido (o ajustado):

$$\bar{R}^2 = 1 - \frac{\frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k}}{\frac{\mathbf{y}'\mathbf{y} - n\bar{Y}^2}{n-1}} = 1 - \left(\frac{n-1}{n-k}\right)\left(\frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{\mathbf{y}'\mathbf{y} - n\bar{Y}^2}\right)$$

$$= 1 - (1 - R^2)\left(\frac{n-1}{n-k}\right)$$

Se cumple: $\bar{R}^2 \leq R^2$

El coeficiente de determinación ajustado permite comparar varios modelos teniendo en cuenta el número de parámetros que contienen

Salida de modelo

Gretl. Modelo → Mínimos cuadrados ordinarios ...

```
Modelo 1: MCO, usando las observaciones 1984--2008 (T = 25)
Variable dependiente: demanda
```

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	410.176	63.4653	6.4630	0.0000
renta	33.0688	16.9773	1.9478	0.0643
precio	−0.567865	0.155207	−3.6588	0.0014

Media de la vble. dep.	309.2000	D.T. de la vble. dep.	136.8369
Suma de cuad. residuos	275604.4	D.T. de la regresión	111.9262
R^2	**0.386706**	**R^2 corregido**	**0.330952**
$F(2,22)$	**6.935942**	**Valor p** (de F)	**0.004617**
Log-verosimilitud	−151.8215	Criterio de Akaike	309.6431
Criterio de Schwarz	313.2997	Hannan--Quinn	310.6573
$\hat{\rho}$	0.158537	Durbin--Watson	1.639491

Análisis de bondad de un modelo

Maximizar Verosimilitud y minimizar las medidas de información

Logaritmo de Verosimilitud	$\ln L = -\frac{n}{2}\left[1 + \ln\left(2\pi\frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n}\right)\right]$
Criterio de Akaike	$AIC = n\ln\left(\frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n}\right) + 2k + n[1 + \ln(2\pi)]$
Cirterio de Schwart	$SIC = -2\ln L + k\ln(n)$
Criterio de Hannan-kinn	$HQC = -2\ln L + 2k\ln(\ln(n))$

Comparación de modelos

Gretl. Salida del modelo: Archivo → Guardar como icono
Vista de iconos: Modelo j → Añadir a la tabla de modelos
Vista de iconos: Tabla de modelos

Estimaciones de MCO
Variable dependiente: demanda

	(1)	(2)	(3)
const	410,2**	470,3**	276,6**
	(63.47)	(58.73)	(64.39)
renta	33,07*		11.06
	(16.98)		(19.69)
precio	−0,5679**	−0,4608**	
	(0.1552)	(0.1537)	
n	25	25	25
$\bar{R}^2$	0.3310	0.2497	-0.0294
ℓ	-151.8	-153.8	-157.8

Desviaciones típicas entre paréntesis
* indica significativo al nivel del 10 por ciento
** indica significativo al nivel del 5 por ciento

Contraste de significación individual

- **Hipótesis**:

$$\boxed{\begin{array}{l} H_0: \ \beta_j = 0 \\ H_1: \ \beta_j \neq 0 \end{array}}$$

Bajo la hipótesis nula la variable X_j no tiene sentido

- **Supuesto de trabajo**: $\mathbf{u} \approx N(\mathbf{0}, \sigma^2 \mathbf{I}) \quad \Rightarrow \quad \hat{\boldsymbol{\beta}} \approx N(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1})$
- **Discrepancia**:

$$d_{\hat{\beta}_j} = \frac{\hat{\beta}_j - \beta_j}{S_{\hat{\beta}_j}} \approx t_{n-k}$$

- **Nivel crítico**: $\boxed{p = P\left(|t_{n-k}| > \left|d_{\hat{\beta}_j}\right|\right)}$
- **CONCLUSIÓN**: Si p es bajo (discrepancias muestrales elevadas) rechazamos la hipótesis y concluimos que la variable asociada al parámetro es significativa

Contraste de significación individual con Gretl

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	410.176	63.4653	6.4630	0.0000	***
renta	33.0688	16.9773	1.9478	0.0643	*
precio	−0.567865	0.155207	−3.6588	0.0014	***

Los asteriscos representan el nivel de significatividad de los coeficientes:

- *** El coeficiente de la variable es significativo al 1 %
- ** El coeficiente de la variable es significativo al 5 %
- * El coeficiente de la variable es significativo al 10 %
- El coeficiente de la variable no es significativo al 10 %

$$\widehat{\text{demanda}} = \underset{(63{,}465)}{410{,}176} + \underset{(16{,}977)}{33{,}0688}\,\text{renta} - \underset{(0{,}15521)}{0{,}567865}\,\text{precio}$$

$$T = 25 \quad \bar{R}^2 = 0{,}3310 \quad F(2,22) = 6{,}9359 \quad \hat{\sigma} = 111{,}93$$

(Desviaciones típicas entre paréntesis)

Contraste sobre subconjuntos de parámetros

- **Hipótesis**:

$$\boxed{\begin{array}{l} H_0: \ \mathbf{R}\beta = \beta^* \\ H_1: \ \mathbf{R}\beta \neq \beta^* \end{array}}$$

  - Bajo la hipótesis nula se imponen ciertas restricciones
  - **R**: Matriz de r filas (restricciones) y k columnas (parámetros)

- **Discrepancia**:

$$\left(\frac{\hat{u}'_R\hat{u}_R - \hat{u}'\hat{u}}{\hat{u}'\hat{u}}\right)\left(\frac{n-k}{r}\right) \approx F^r_{n-k}$$

  - $\hat{\mathbf{u}}_\mathbf{R}$: Residuos modelo restringido
  - $\hat{\mathbf{u}}$: Residuos modelo sin restricciones

- **Nivel crítico**: $\boxed{p = P\left(F^r_{n-k} \geq F^*\right)}$

- **CONCLUSIÓN**: Si p es bajo se rechaza la hipótesis y se concluye que las restricciones no son admisibles

Contrastes sobre parámetros. Ilustración

Modelo:

$$\mathbf{Y} = \beta_1 + \beta_2 \mathbf{X}_2 + \beta_3 \mathbf{X}_3 + \beta_4 \mathbf{X}_4 + \beta_5 \mathbf{X}_5 + \mathbf{u}$$

- **Hipótesis**: $H_0 : \beta_2 + \beta_4 = 4, \quad \beta_3 = 2\beta_5$
- $H_0 : \mathbf{R}\boldsymbol{\beta} = \boldsymbol{\beta}^*$

$$\mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -2 \end{pmatrix} \quad \beta^* = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

- **Modelo libre**: $\hat{\mathbf{Y}} = \hat{\beta}_1 + \hat{\beta}_2 \mathbf{X}_2 + \hat{\beta}_3 \mathbf{X}_3 + \hat{\beta}_4 \mathbf{X}_4 + \hat{\beta}_5 \mathbf{X}_5$
- **Modelo restringido**: $\hat{\mathbf{Y}}^* = \hat{\beta}_1^* + \hat{\beta}_2^* \mathbf{X}_2 + 2\hat{\beta}_5^* \mathbf{X}_3 + (4 - \hat{\beta}_2^*)\mathbf{X}_4 + \hat{\beta}_5^* \mathbf{X}_5$
- **Discrepancia**:

$$\boxed{\left(\frac{\hat{\mathbf{u}}'_{\mathbf{R}}\hat{\mathbf{u}}_{\mathbf{R}} - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}}\right)\left(\frac{n-5}{2}\right) \approx F^2_{n-5}}$$

Test de restricciones. Ilustración

Modelo libre

$$\widehat{Demanda} = \hat{\beta}_1 + \hat{\beta}_2 Renta + \hat{\beta}_3 Precio$$

Hipótesis

- La demanda autónoma es de 400 unidades
- El efecto marginal de renta sobre la demanda es 60 veces mayor y de signo contrario al del precio

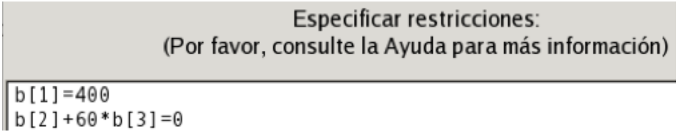
$$\boxed{\beta_1 = 400 \; ; \; \beta_2 + 60\beta_3 = 0}$$

Modelo restringido

$$\widehat{Demanda} = 400 - 60\hat{\beta}_3 Renta + \hat{\beta}_3 Precio$$

Test de restricciones. Ilustración Gretl

Gretl. Salida del modelo: Contrastes → Restricciones lineales



Conjunto de restricciones

1 : b[const] = 400

2 : b[renta] + 60*b[precio] = 0

Estadístico de contraste: F(2, 22) = 0.0218156, con valor **p = 978442**

Estimaciones restringidas:

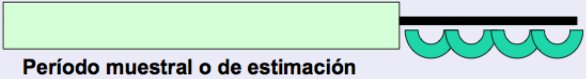
	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	400.000	7.57082e-07	5.283e+08	2.63e-194	***
renta	35.5710	5.67256	-5.742	6.45e-06	***
precio	-0.54285	0.0945427	-5.742	6.45e-06	***

Desviación típica de la regresión = 107.267

Predicción estática y dinámica

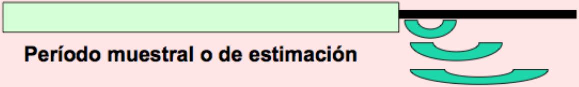
Horizonte h=1

Predicción estática con horizonte 1, basada en los datos registrados (one-step-ahead forecast)



Horizonte cambiante

Predicción dinámica con horizonte cambiante basada en valores previstos (multi-step forecast)



Elaboración de predicciones

Para elaborar predicciones a partir de un modelo $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$ será necesario asignar valores a las variables explicativas, mediante el vector:

$$\mathbf{x_0}' = (1, X_{20}, X_{30}, ..., X_{k0})$$

Predicción puntual

La predicción de **Y** vendrá entonces dada por:

$$\hat{\mathbf{Y}}_\mathbf{0} = \mathbf{x}_\mathbf{0}'\hat{\beta}$$

Error de predicción

$$e_{\hat{Y}_0} = Y_0 - \hat{Y}_0$$

Varianza del error de predicción

$$Var\left(e_{\hat{Y}_0}\right) = \sigma^2\left[1 + x_0{}'\left(X'X\right)^{-1}x_0\right]$$

Elaboración de predicciones

Intervalo de confianza para la predicción

Intervalo de confianza al nivel $1-\alpha$ para la predicción de Y condicionada al vector $\mathbf{x_0'}$

$$\left[\hat{Y}_0 - kS\sqrt{1+\mathbf{x_0'}(\mathbf{X'X})^{-1}\mathbf{x_0}}, \hat{Y}_0 + kS\sqrt{1+\mathbf{x_0'}(\mathbf{X'X})^{-1}\mathbf{x_0}}\right]$$

siendo k el valor tal que $P\left(|t_{n-k}| > k\right) = 1-\alpha$

Riesgo de predicción

El riesgo de la predicción dependerá de factores como:

- La bondad del modelo en el que se basa
- El tamaño muestral del modelo estimado (al aumentar n disminuye el riesgo, ya que afecta a S y $(\mathbf{X'X})^{-1}$)
- El horizonte de predicción h (las predicciones dinámicas se basan en predicciones anteriores, y por tanto el riesgo aumenta con el horizonte temporal)

Riesgo de predicción

Las predicciones pueden ir referidas al valor esperado de **Y** o al valor observado de esta variable.
En ambos casos la predicción puntual es la misma, pero en cambio el riesgo es mayor si la predicción va referida al valor observado. Más concretamente las bandas de confianza serían:

Para $\mathbf{Y}/\mathbf{x}_0'$

$$\left[\hat{Y}_0 - kS\sqrt{1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0},\ \hat{Y}_0 + kS\sqrt{1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0}\right]$$

Para $E(\mathbf{Y}/\mathbf{x}_0')$

$$\left[\hat{Y}_0 - kS\sqrt{\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0},\ \hat{Y}_0 + kS\sqrt{\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0}\right]$$