Capítulo 16.

INFERENCIAESTADÍSTICA EN EL MODELO LINEAL GENERAL

En el proceso de investigación econométrica, la asociación entre variables se estudia en la etapa de análisis de información. Una manera de hacerlo es especificar un modelo lineal que relacione a las variables como se mostró en el capítulo precedente. Luego de especificado, el modelo debe ser estimado y evaluado para verificar, si se cumple o no, la relación establecida entre la variable dependiente y las exógenas. En este capítulo se evalúa dicha relación. Para ello se estudia si las variables exógenas son o no significativas, tanto en forma individual como conjunta. Se presentan una serie de pruebas para verificar la bondad del ajuste, basándose en contrastes de hipótesis estadísticas y en coeficientes que estudian si la relación lineal especificada es adecuada.

16.1. El coeficiente de determinación

Cuando se específica un modelo lineal general, se está evaluando cuánto de las variaciones que se producen en la variable dependiente o endógena, en el espacio y tiempo establecido, es determinada por las variaciones de las variables exógenas. Una de las maneras de hacerlo es descomponiendo la varianza de la variable endógena en dos componentes; la que proviene de las variaciones producidas por las variables exógenas y la que proviene de la parte aleatoria del modelo. Para ello, se puede realizar la siguiente descomposición de la varianza de la variable dependiente:

$$\frac{\sum_t (Y_t - \bar{Y})^2}{T-1} = \frac{\sum_t \left(\hat{Y}_t - \bar{Y}\right)^2}{k-1} + \frac{\sum_t \left(Y_t - \hat{Y}_t\right)^2}{T-k}$$

Donde es evidente que: $T-1 = k-1+T-k$; siendo T la cantidad de observaciones, T – K los grados de libertad del modelo y k – 1 la cantidad de variables exógenas. El primer sumando expresa las variaciones producidas por las variables exógenas y el segundo sumando las variaciones producidas por la parte aleatoria, representada por los residuos del modelo.

Trabajando con los numeradores:

$$\sum_t \left(Y_t - \overline{Y}\right)^2 = \sum_t \left(\hat{Y}_t - \overline{Y}\right)^2 + \sum_t \left(Y_t - \hat{Y}_t\right)^2 \qquad [1]$$

Esto equivale a decir: $SCT = SCE + SCR$

Donde,

SCT, es la suma de cuadrados de la variable endógena o lo que es lo mismo, es el numerador de la varianza de Y_t y se denominará *Suma de Cuadrados Totales*;

SCE, es la *Suma de Cuadrados Explicada* por las variables exógenas;

SCR, es la *Suma de Cuadrados de los Residuos* de la regresión.

En base a esta descomposición de la variabilidad de Y, se define el *coeficiente de determinación* como una medida de la capacidad explicativa del modelo, es decir, de la bondad de ajuste:

$$R^2 = \frac{\sum_t \left(\hat{Y}_t - \overline{Y}\right)^2}{\sum_t \left(Y_t - \overline{Y}\right)^2} = 1 - \frac{\sum_t e_t^2}{\sum_t \left(Y_t - \overline{Y}\right)^2} = 1 - \frac{SCR}{SCT} \qquad [2]$$

O lo que es lo mismo,

$$R^2 = 1 - \frac{\mathbf{e}'\mathbf{e}}{\mathbf{y}'\mathbf{y} - T\overline{Y}^2} = 1 - \frac{SCR}{SCT} \qquad [3]$$

Para demostrar esta igualdad, se descompone la suma del cuadrado de los residuos; teniendo en cuenta que, según definiciones estudiadas en el capítulo anterior, el error ($\mathbf{e}$) es la diferencia entre el vector de valores observados ($\mathbf{y}$) y el vector de valores estimados $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ se tiene:

$$\mathbf{e}'\mathbf{e} = \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)' \left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)$$

Realizando los productos correspondientes

$$= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

Se sabe, por demostraciones realizadas en el capítulo anterior, que $\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$; reemplazando esta expresión en el segundo y tercer término se tiene

$$= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'(\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}) - (\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e})'\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

Aplicando la propiedad de traspuesta y eliminando paréntesis

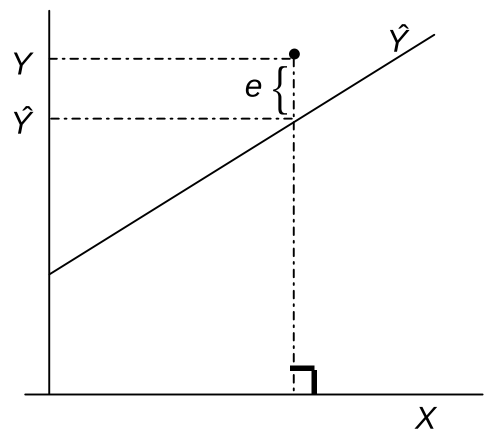
$$= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{e} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}'\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

Recordando que $\mathbf{X}'\mathbf{e} = \mathbf{0}$, por lo que $e'\mathbf{X} = \mathbf{0}$, y simplificando la expresión

$$\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

Pero $\mathbf{X}\hat{\boldsymbol{\beta}} = \hat{\mathbf{y}}$, por lo que

$$\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{y}}'\hat{\mathbf{y}} \qquad [4]$$



Restando en ambos miembros $T\overline{Y}^2$, y reordenando términos, se obtiene:

$$\begin{aligned} \mathbf{y'y} - T\overline{Y}^2 &= \mathbf{e'e} + \mathbf{\hat{y}'\hat{y}} - T\overline{Y}^2 \\ SCT &= SCR + SCE \end{aligned} \qquad [5]$$

Por lo tanto,

$$R^2 = \frac{SCE}{SCT} = \frac{SCT - SCR}{SCT} = 1 - \frac{SCR}{SCT} = 1 - \frac{\mathbf{e'e}}{\mathbf{y'y} - T\overline{Y}^2} \qquad [6]$$

Este coeficiente mide el *porcentaje de la variación* de la variable endógena, Y, que queda explicada en la regresión por la variación conjunta de las variables exógenas, X.

Ejemplo 16.1 Con los datos del Ejemplo 15.5, se calcula el coeficiente de determinación:

$$R^2 = \frac{SCE}{SCT} = \frac{133.0998}{133.2} = 0.999 = \left(1 - \frac{0.0858}{133.2}\right)$$

Esto significa que, la proporción de la variación de Y explicada por la regresión lineal es de 0.999 ó el 99,9%.

El coeficiente R^2 está comprendido entre 0 y 1.

Si $\sum_t e_t^2 = 0$, entonces $R^2 = 1$ y el modelo estimado se ajustaría perfectamente a los datos.

Por el contrario, si $\sum_t \left(Y_t - \overline{Y}\right)^2 = \sum_t e_t^2$, el modelo de regresión no explicaría algo de la variabilidad de $Y_t \Rightarrow R^2 = 0$.

Como conclusión, el modelo se ajusta mejor a los datos cuando más próximo está el coeficiente de determinación a la unidad.

El coeficiente de determinación, R^2, va a aumentar al añadir más regresores al modelo, sin que esto signifique que la nueva variable incluida sea relevante para explicar el comportamiento de Y. Una medida que posibilita analizar el número óptimo de variables a incorporar es el *coeficiente de determinación corregido*, que se define como:

$$\overline{R^2} = 1 - \frac{\sum_t e_t^2 / T-k}{\sum_t \left(Y_t - \overline{Y}\right)^2 / T-1} = 1 - \frac{\mathbf{e'e}/T-k}{\left(\mathbf{y'y} - T\overline{Y}^2\right)/T-1} \qquad [7]$$

Donde las sumas cuadráticas (residual y total) se corrigen por los grados de libertad; éstos son los términos que penalizan la inclusión desmedida de regresores en el modelo.

La incorporación de variables en el modelo aporta el beneficio de incrementar el valor de R^2 pero tiene el costo de disminuir los grados de libertad. Mientras el beneficio supere al costo, será conveniente incorporar nuevas variables; si el costo, en términos de grados de libertad, supera el beneficio de incorporar variables significa que ya no resulta conveniente la inclusión.

De esta forma, conforme se aumente el número de regresores, no está claro cual va a ser la variación que experimente el coeficiente de determinación corregido.

Ejemplo 16.1.a Para el ejemplo

$$\bar{R}^2 = 1 - \frac{0.0858/2}{133.2/4} = 0.9987$$

Para comparar el ajuste de varias especificaciones de acuerdo con el número de regresores utilizados, además del $\bar{R}^2$, existen otros criterios:

Criterio de Schwarz: $CS = \ln \frac{\mathbf{e'e}}{T} + \frac{k}{T} \ln T$ [8]

Criterio de Akaike: $CA = \ln \frac{\mathbf{e'e}}{T} + \frac{2k}{T}$ [9]

Habitualmente se buscan especificaciones que reduzcan la suma cuadrática de los residuos; sin embargo, todos los criterios llevan implícita una penalización que aumenta con el número de regresores.

16.2. Inferencia

Para contrastar determinadas hipótesis sobre los parámetros del modelo de regresión, se utilizan estadísticos cuya distribución exacta, bajo la hipótesis nula, depende de la distribución de los estimadores de $\boldsymbol{\beta}$ y σ^2.

De acuerdo a los supuestos establecidos sobre el modelo, se puede demostrar que para cualquier tamaño de muestra dado, las perturbaciones ε_t siguen distribuciones normales, independientemente distribuidas de media cero y varianza constante σ_ε^2. Por tanto, los elementos del vector $\boldsymbol{\varepsilon}$, se distribuyen independiente y conjuntamente según una ley normal multivariante, con vector de medias $\mathbf{0}$ y matriz de covarianzas $\sigma_{\boldsymbol{\varepsilon}}^2\mathbf{I}_T$. Como se estudió en el capítulo anterior, esto se puede especificar de la siguiente forma:

$$\mathrm{f}(\boldsymbol{\varepsilon}) = \mathrm{p}\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_\mathrm{T} \end{bmatrix} = \frac{1}{(2\pi)^{\mathrm{T}/2}(\sigma_{\boldsymbol{\varepsilon}}^2)^{\mathrm{T}/2}}\exp[-\frac{1}{2\sigma_{\boldsymbol{\varepsilon}}^2}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}] \qquad [10]$$

En [10], $\frac{1}{\sigma_{\boldsymbol{\varepsilon}}^2}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$ puede expresarse como $\boldsymbol{\varepsilon}'\left(\sigma_{\boldsymbol{\varepsilon}}^2\mathbf{I}_T\right)^{-1}\boldsymbol{\varepsilon}$, el cual define una forma cuadrática cuyo resultado es el siguiente escalar:

$$\frac{\varepsilon_1^2}{\sigma^2} + \frac{\varepsilon_2^2}{\sigma^2} + \cdots + \frac{\varepsilon_T^2}{\sigma^2}$$

Es decir, una suma de variables normales tipificadas, porque $E(\varepsilon) = 0$. De esta forma cada variable aleatoria normal ε_T se divide por σ^2. Así pues,

$$\frac{\varepsilon_1^2}{\sigma^2} + \frac{\varepsilon_2^2}{\sigma^2} + \cdots + \frac{\varepsilon_T^2}{\sigma^2} \sim \chi_T^2$$

es decir, la forma cuadrática $\boldsymbol{\varepsilon}'\left(\sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}_T\right)^{-1} \boldsymbol{\varepsilon} \sim \chi_T^2$

La ecuación anterior muestra explícitamente que la matriz de la forma cuadrática es la inversa de la matriz de covarianzas.

Dado que las perturbaciones son no observables, puede interesar conocer la distribución de la suma de cuadrados de los residuos vista en el capítulo anterior:

$$\mathbf{e}'\mathbf{e} = \boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}$$

Y para ello se utilizan las propiedades de las matrices simétricas e idempotentes y, en particular, que la forma cuadrática con variables aleatorias ε_T tipificadas se distribuye como

$$\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon} \sim \chi_{T-k}^2$$

Siendo $T-k$ el rango y la traza de la matriz $\mathbf{M}$, matriz simétrica e idempotente. Por tanto, y en este caso con $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}_T)$ y rango igual a la traza $T-k$

$$\frac{\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}}{\sigma_{\boldsymbol{\varepsilon}}^2} = \frac{\mathbf{e}'\mathbf{e}}{\sigma_{\boldsymbol{\varepsilon}}^2} \sim \chi_{T-k}^2 \qquad [11]$$

Resultado que será utilizado en los contrastes de validez del modelo.

Observación. Dado un vector $\mathbf{x}$ con distribución normal multivariante

$$\mathbf{x} \sim \mathbf{N}(\mathbf{0}, \mathbf{I})$$

considere la forma cuadrática $\mathbf{x}'\mathbf{A}\mathbf{x}$ donde $\mathbf{A}$ es idempotente con rango $r \le n$.

Representando la matriz de vectores característicos de $\mathbf{A}$ por $\mathbf{Q}$, entonces

$$\mathbf{Q}'\mathbf{A}\mathbf{Q} = \mathbf{D} \qquad [12]$$

donde $\mathbf{D}$ tendrá r unos y $n-r$ ceros en la diagonal principal.

$$= \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\ \ddots \\1\end{matrix}}\right\} \text{r términos = rango(A)} \\ \left.\vphantom{\begin{matrix}0\\ \ddots \\0\end{matrix}}\right\} \text{n-r términos} \end{matrix}$$

Definiendo: $\mathbf{y} = \mathbf{Q}'\mathbf{x}$, entonces, $\mathbf{x} = \mathbf{Q}\mathbf{y}$

Por lo tanto,

$$E(\mathbf{Y}) = E(\mathbf{Q}'\mathbf{X}) = \mathbf{Q}'E(\mathbf{X}) = 0$$
$$V(\mathbf{Y}) = E[\mathbf{Q}'\mathbf{X}(\mathbf{Q}'\mathbf{X})'] = E[\mathbf{Q}'\mathbf{X}\mathbf{X}'\mathbf{Q}] = \mathbf{Q}'E(\mathbf{X}\mathbf{X}')\mathbf{Q} = \mathbf{Q}'\mathbf{I}\mathbf{Q} = \mathbf{Q}'\mathbf{Q}$$

Al ser $\mathbf{Q}$ ortogonal, es decir vectores linealmente independientes,

$$\mathbf{Q}' = \mathbf{Q}^{-1} \Rightarrow \mathbf{Q}'\mathbf{Q} = \mathbf{Q}^{-1}\mathbf{Q} = \mathbf{I}$$

Por esto

$$V(\mathbf{y}) = \mathbf{I}$$

De esta forma las variables Y son normales tipificadas e independientes.

La forma cuadrática se puede expresar ahora usando [12] como:

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{y}'\mathbf{Q}'\mathbf{A}\mathbf{Q}\mathbf{y}$$

donde

$$\mathbf{Q}\mathbf{y} = \mathbf{x}$$
$$\mathbf{Q}'\mathbf{A}\mathbf{Q} = \mathbf{D}$$

De modo que

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{y}'\mathbf{Q}'\mathbf{A}\mathbf{Q}\mathbf{y} = Y_1^2 + Y_2^2 + \cdots + Y_T^2$$

Por lo tanto,

$$\mathbf{x}'\mathbf{A}\mathbf{x} \sim \chi^2(r) \qquad [13]$$

Generalizando, si $\mathbf{x} \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I})$ y $\mathbf{A}$ es idempotente con rango $r \leq n$, entonces $\frac{1}{\sigma^2}\mathbf{x}'\mathbf{A}\mathbf{x} \sim \chi^2(r)$

Proposición 16.1. Bajo el supuesto de normalidad de las perturbaciones aleatorias ε_t y conociendo, además, que $\mathbf{X}$ es una matriz no estocástica de rango completo por columnas, $\rho(\mathbf{X}) = k$, se tiene que $\hat{\boldsymbol{\beta}}$ es un vector aleatorio con distribución normal $k-$variante, ya que es función del vector aleatorio normal $\boldsymbol{\varepsilon}$. Por lo tanto, $\hat{\boldsymbol{\beta}}$ se distribuye normal *k-variante* con media $\boldsymbol{\beta}$ y varianza $\sigma^2(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}$

$$\hat{\boldsymbol{\beta}} \sim N_k\left[\boldsymbol{\beta},\ \sigma^2(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\right] \qquad [14]$$

La primera parte de la proposición queda demostrada considerando el resultado conocido de que

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

Se sabe también que

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$$
$$\mathbf{V}(\hat{\boldsymbol{\beta}}) = \sigma_\varepsilon^2(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}$$

Con lo cual $\hat{\boldsymbol{\beta}}$ es estimador insesgado y óptimo.

Por lo tanto, $\hat{\boldsymbol{\beta}} \sim N\left[\boldsymbol{\beta},\ \sigma^2(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\right]$

Tipificando,

$$\frac{\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}}{\left[\sigma^2(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\right]^{1/2}} \sim N\ [\mathbf{0}, \mathbf{I}]$$

Es decir, los k estimadores tienen una función de probabilidad normal multivariante

$$p(\hat{\boldsymbol{\beta}}) = p(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)$$
$$= \frac{1}{(2\pi)^{k/2} \left| \sigma_\varepsilon^2 (\mathbf{X'X})^{-1} \right|^{1/2}} \exp\left[-\frac{1}{2} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \left[\sigma_\varepsilon^2 (\mathbf{X'X})^{-1} \right]^{-1} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \right]$$

Proposición 16.2. Sea el vector aleatorio $\hat{\boldsymbol{\beta}}$ con distribución normal multivariante y sea $\mathbf{R}$ una matriz de orden $\mathbf{qxk}$, con rango $\rho(\mathbf{R}) = q$, entonces el vector $\mathbf{R}\hat{\boldsymbol{\beta}}$ tiene una distribución normal $q-$variante con media igual a $\mathbf{R}\boldsymbol{\beta}$ y varianza igual a $\sigma_\varepsilon^2 \mathbf{R}(\mathbf{X}'\,\mathbf{X})^{-1}\mathbf{R}'$

Corolario: $\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim N_q \left[\mathbf{0},\ \sigma_\varepsilon^2 \mathbf{R}(\mathbf{X}'\,\mathbf{X})^{-1}\mathbf{R}' \right]$ [15]

Si se tiene $\hat{\boldsymbol{\beta}} \sim N \left[\boldsymbol{\beta},\ \sigma_\varepsilon^2 (\mathbf{X}'\,\mathbf{X})^{-1} \right]$ y una matriz $\underset{qxk}{\mathbf{R}}$ donde $rango(\mathbf{R}) = q$

el producto de

$$\mathbf{R}\hat{\boldsymbol{\beta}} \sim N \left[\mathbf{R}\boldsymbol{\beta},\ \sigma_\varepsilon^2 \mathbf{R}(\mathbf{X}'\,\mathbf{X})^{-1}\mathbf{R}' \right]$$

de modo que

$$\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim N_q \left[0,\ \sigma_\varepsilon^2 \mathbf{R}(\mathbf{X}'\,\mathbf{X})^{-1}\mathbf{R}' \right]$$

Proposición 16.3. Si se combina al vector de perturbaciones $\boldsymbol{\varepsilon}$ del modelo, con distribución $N_t\ \left[\mathbf{0},\ \sigma_\varepsilon^2\right]$, con una matriz simétrica e idempotente $\mathbf{M}$, entonces,

(a) $\dfrac{\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}}{\sigma_\varepsilon^2} \sim \chi^2_{gl=\rho(M)}$

(b) Sea $\mathbf{N}$ otra matriz idempotente y simétrica, las distribuciones $\chi^2\ \dfrac{\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}}{\sigma_\varepsilon^2}$ y $\dfrac{\boldsymbol{\varepsilon}'\mathbf{N}\boldsymbol{\varepsilon}}{\sigma_\varepsilon^2}$ son independientes si y solo si $\mathbf{M}\cdot\mathbf{N}=\mathbf{0}$

(c) Dada una matriz $\mathbf{R}$ de orden $\mathbf{qxk}$, con rango $\rho(\mathbf{R})=q$; el vector aleatorio $\mathbf{R}\hat{\boldsymbol{\beta}}$, con distribución normal $q-$variante con media igual a $\mathbf{R}\boldsymbol{\beta}$ y varianza igual a $\sigma_\varepsilon^2\mathbf{R}\left(\mathbf{X}'\,\mathbf{X}\right)^{-1}\mathbf{R}'$, es independiente de $\dfrac{\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}}{\sigma_\varepsilon^2} \sim \chi^2_{gl=\rho(M)}$, si y solamente si $\mathbf{R}\cdot\mathbf{M}=\mathbf{0}$

Corolarios: Así, de la parte (a) se puede decir que, sobre la base de demostraciones anteriores, $\mathbf{e}'\mathbf{e}=\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}$, entonces,

$$\frac{e_1^2}{\sigma_\varepsilon^2}+\frac{e_2^2}{\sigma_\varepsilon^2}+\cdots+\frac{e_t^2}{\sigma_\varepsilon^2} \sim \chi^2_{T-k}\text{, o bien, } \frac{\mathbf{e}'\mathbf{e}}{\sigma_\varepsilon^2} \sim \chi^2_{T-k}$$

De esta última expresión -y conociendo que el estimador de σ_ε^2 es $S^2=\dfrac{\mathbf{e}'\mathbf{e}}{T-k}$- se puede establecer la importante conclusión:

$$\frac{(T-k)S^2}{\sigma_\varepsilon^2} \sim \chi^2_{T-k} \qquad [16]$$

Donde los grados de libertad se obtienen por el hecho de que $\rho(\mathbf{M}) = tr\mathbf{M}$, siendo $\mathbf{M}$ una matriz simétrica e idempotente con $tr\mathbf{M} = T-k$.

Pero también -y teniendo en cuenta las tres partes de esta proposición- se tiene que

$\hat{\boldsymbol{\beta}}$ y $\frac{\mathbf{e'e}}{\sigma_\varepsilon^2}$ se distribuyen independientemente .

Estas proposiciones son suficientes para establecer los procedimientos de inferencia para cualquier elemento del vector $\hat{\boldsymbol{\beta}}$ o sobre alguna combinación lineal de los mismos. El objetivo es contrastar hipótesis sobre los coeficientes de regresión del modelo.

Prueba de hipótesis de normalidad

Lo primero a realizar, antes de hacer cualquier inferencia en el modelo lineal general, es comprobar que -para el espacio y tiempo que se ha utilizado- se cumple con la hipótesis de normalidad de las perturbaciones aleatorias.

El contraste de Normalidad se realiza a través del estadístico de Jarque Bera que se define como

$$JB = \frac{T-k}{6}\left(A^2 + \frac{(C-3)^2}{4}\right) \sim \chi^2_2$$

Donde A representa la simetría de la distribución de los errores mínimo cuadráticos y C la curtosis de dicha distribución. Siendo k el número de coeficientes y T el número de unidades de observación. Bajo la hipótesis nula de normalidad, este estadístico se distribuye como chi cuadrado con 2 grados de libertad.

Observación. Dada una serie de T observaciones sobre una variable cualquiera, por ejemplo Z, la medida de asimetría de la distribución de la serie se calcula haciendo

$$A = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{Z_t - \bar{Z}}{\hat{\sigma}_z}\right)^3$$

Donde $\hat{\sigma}_Z$ es el estimador del desvío estándar. Si la distribución es simétrica, A=0

La curtosis mide la puntiagudez de la distribución de la serie. Se calcula haciendo

$$C = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{Z_t - \bar{Z}}{\hat{\sigma}_z}\right)^4$$

Si C=3, la distribución es mesocúrtica o normal; C<3 indica una distribución platicúrtica (forma chata, aplanada); C>3, distribución leptocúrtica (forma puntiaguda, alta frecuencia de observaciones alrededor de un mismo valor).

Pruebas de significatividad individual y conjunta de los coeficientes de regresión

Las pruebas que se desarrollan aquí, se restringen a aquellas hipótesis que pueden expresarse como combinaciones lineales de los coeficientes de regresión $\boldsymbol{\beta}$.

Las hipótesis a contrastar se pueden escribir, de forma general:

$$\begin{aligned} H_o &: \mathbf{R}\boldsymbol{\beta} = \mathbf{r} \\ H_1 &: \mathbf{R}\boldsymbol{\beta} \neq \mathbf{r} \end{aligned} \qquad [17]$$

Donde:

$\mathbf{R}$ es una matriz de orden $\mathbf{qxk}$, sus filas representan la cantidad de restricciones (q) y sus columnas la cantidad de parámetros (k) del modelo, siendo sus elementos los coeficientes que acompañan a los parámetros a contrastar.

El rango de la matriz **R** viene dado por el número de restricciones sobre los parámetros que se están contrastando

$\rho(\mathbf{R}) = \mathbf{q}$

$\mathbf{r}$ es un vector de tamaño $\mathbf{qx1}$, con $\mathbf{q} \geq \mathbf{1}$.

El estadístico de contraste se determina de acuerdo a las proposiciones anteriores. Así, por [15] se sabe que

$$\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim N_q \left[\mathbf{0},\ \sigma_\varepsilon^2 \mathbf{R}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{R}' \right]$$

Planteando la hipótesis nula $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ y reemplazando en $\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, se tiene que

$$\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{R}\boldsymbol{\beta} = \mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}$$

con lo cual

$$\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r} \sim N_q \left[\mathbf{0},\ \sigma_\varepsilon^2 \mathbf{R}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{R}' \right]$$

Asimismo, por la proposición 16.2 y tipificando la variable aleatoria $\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}$ se tiene que

$$(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})' \left[\sigma_\varepsilon^2 \mathbf{R}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{R}' \right]^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}) \sim \chi_q^2$$

Es decir, se distribuye como una chi cuadrado con grados de libertad igual al número de restricciones (e igual al rango de la matriz $\mathbf{R}$).

El problema habitual es que σ_ε^2 es desconocido y se estima a partir de $s^2 = \frac{\mathbf{e'e}}{T-k}$; utilizando la propiedad de que $\frac{\mathbf{e'e}}{\sigma_\varepsilon^2} \sim \chi_{T-k}^2$ se puede obtener, finalmente, el siguiente estadístico

$$F = \frac{\left[\left(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right)' \left(\mathbf{R} \left(\mathbf{X}' \mathbf{X} \right)^{-1} \mathbf{R}' \right)^{-1} \left(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right) \right] / q}{\mathbf{e'e} / T-k} \overset{H_0}{\sim} F_{q,\, T-k} \qquad [18]$$

Que, bajo la hipótesis nula, sigue una distribución F de Snedecor. Los grados de libertad vienen dados por el número de restricciones q y $T-k$, donde T es el número de observaciones y k el número de coeficientes estimados.

A este procedimiento se lo conoce como test de restricciones lineales y permite contrastar cualquier conjunto de restricciones lineales sobre los parámetros.

La Regla de decisión consiste en rechazar la hipótesis nula $H_o : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ con un nivel de significación α, cuando el valor muestral del estadístico F sea mayor que la ordenada de la distribución $F_{q,T-k}$ que deja a la derecha una probabilidad α; es decir, si:

$F \geq F_{\alpha;q,T-k}$ la hipótesis nula se rechaza

Ejemplo 16.1.b. Con los datos del ejemplo 15.5 se va a contrastar el conjunto de las hipótesis

$$H_0: \begin{cases} \beta_2 = 1{,}5 \\ \beta_3 = -0{,}1 \end{cases}$$

El modelo definido en el ejemplo 15.5 es

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i; \quad \forall i = 1,\dots,n$$

La matriz $\mathbf{R}$ y el vector $\mathbf{r}$ se construyen de la siguiente manera:

$$\mathbf{R}_{qxk=2x3} = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\underbrace{\beta_1 \quad \beta_2 \quad \beta_3}_{\text{coeficientes de } \beta_i \text{ en las restricciones}}} \qquad \mathbf{r}_{qx1=2x1} = \begin{bmatrix} 1{,}5 \\ -0{,}1 \end{bmatrix}$$

De acuerdo a [17], la hipótesis nula se define

$$\mathrm{H}_0 : \mathbf{R}\beta = \mathbf{r}$$

esto es

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1{,}5 \\ -0{,}1 \end{bmatrix}$$

El contraste de esta hipótesis se realiza con la expresión [18], para lo cual hay que resolver $\mathbf{R}\hat{\beta}-\mathbf{r}$.

En el ejemplo 15.5 $\widehat{\boldsymbol{\beta}} = \begin{bmatrix} 2.418523 \\ 1.033557 \\ 0.34 \end{bmatrix}$

por lo que

$$\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{r} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2.418523 \\ 1.033557 \\ 0.34 \end{bmatrix} - \begin{bmatrix} 1.5 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 1.033557 \\ 0.34 \end{bmatrix} - \begin{bmatrix} 1.5 \\ -0.1 \end{bmatrix}$$
$$= \begin{bmatrix} -0.4664 \\ 0.44 \end{bmatrix}$$

En el ejemplo 15.5.d se tiene el resultado de $(\mathbf{X}'\mathbf{X})^{-1}$

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{29800} \begin{bmatrix} 26396 & -1650 & -2384 \\ -1650 & 250 & 0 \\ -2384 & 0 & 596 \end{bmatrix}$$

De modo que $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$ será

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{R}' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \frac{1}{29800} \begin{bmatrix} 26396 & -1650 & -2384 \\ -1650 & 250 & 0 \\ -2384 & 0 & 596 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{R}' = \begin{bmatrix} 0.008389 & 0 \\ 0 & 0.02 \end{bmatrix}$$

Ahora se debe obtener la inversa de $\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$, para lo cual se debe calcular el determinante y la matriz adjunta; esto es:

$$\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} = \frac{1}{\left|\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right|} Adj\left(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right)$$

resolviendo, $\left|\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right| = \begin{bmatrix} 0.008389 & 0 \\ 0 & 0.02 \end{bmatrix} = 0.00016779$

$$\boldsymbol{Adj}(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}') = Adj\left(\begin{bmatrix} 0.008389 & 0 \\ 0 & 0.02 \end{bmatrix}\right) = \begin{bmatrix} 0.02 & 0 \\ 0 & 0.008389 \end{bmatrix}$$

por lo que

$$[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1} = \frac{1}{0.00016779}\begin{bmatrix} 0.02 & 0 \\ 0 & 0.008389 \end{bmatrix} = \begin{bmatrix} 119.2 & 0 \\ 0 & 50 \end{bmatrix}$$

Entonces

$$\begin{aligned}&\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{r}\right)'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{r}\right)\\ &\qquad = [-0.4664 \quad 0.44]\begin{bmatrix} 119.2 & 0 \\ 0 & 50 \end{bmatrix}\begin{bmatrix} -0.4664 \\ 0.44 \end{bmatrix}\end{aligned}$$

Resolviendo $\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{r}\right)'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{r}\right) = 35.609452$

De acuerdo al Ejemplo 15.5.f, $\mathbf{e}'\mathbf{e} = 0.085786$

En síntesis:

$\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\mathbf{r}\right)'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\mathbf{r}\right)=35.609452$

$\mathbf{e}'\mathbf{e}=0.085786$

q=2

$T-k=n-k=5-3=2$

El estadístico F expresado en [18] es

$$F=\frac{\left[\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\mathbf{r}\right)'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}\left(\mathbf{R}\widehat{\boldsymbol{\beta}}-\mathbf{r}\right)/\mathbf{q}\right]}{(\mathbf{e}'\mathbf{e})/(\mathbf{T}-\mathbf{k})}=\frac{35.609452/2}{0.085786/2}=415.09630$$

El valor teórico de, $F^{*}_{2,2;0.95}=19$

por lo que $F>F^{*}_{2,2;0.95}$

entonces se rechaza la hipótesis nula: $\beta_2=1{,}5$ y $\beta_3=-0{,}1$.

Casos particulares

1) Si la hipótesis a contrastar se refiere a los valores de todos los coeficientes, se tiene un contraste de la *significación conjunta* del modelo de regresión, el que informa si cambios en las variables explicativas X explican, en conjunto, las variaciones de la variable endógena y.

La hipótesis nula es:

$$H_0 : \beta_2 = \beta_3 = \cdots = \beta_k = 0 \qquad [19]$$
H_1 : alguna de las igualdades no se cumple

Y la matriz **R** de orden **(k − 1)xk** y el vector **r** de orden **(k − 1)** del estadístico, son en este caso:

$$R = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ . & . & . & . & \cdots & . \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \qquad r = \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ 0 \end{bmatrix}$$

Observación. En este caso el número de restricciones (q) es k-1, que representa las variables explicativas del modelo, dejando a un lado el término constante.

Es interesante señalar que esta hipótesis nula no incluye el coeficiente relacionado con el término independiente de la regresión β_1. De hecho, aunque todos los coeficientes $\beta_2, \beta_3, \cdots, \beta_k$ no fueran estadísticamente distintos de cero, el término independiente β_1 recogería aproximadamente la media de la variable endógena y podría ser distinto de cero.

El conjunto de hipótesis de [19] puede contrastarse con el coeficiente de determinación como:

$$F = \frac{R^2 / k-1}{\left(1-R^2\right)/ T-k} \overset{H_0}{\sim} F_{k-1,\, T-k} \qquad [20]$$

Si el valor del estadístico F es superior a la ordenada $F_{\alpha; k-1,\, T-k}$, se concluye que las variables $X_2,\ldots,X_k$, conjuntamente, tienen un efecto significativamente distinto de cero a la hora de explicar la variable dependiente.

Ejemplo 16.1.c. Para el ejemplo que se viene desarrollando

K=3

T=5

$R^2 = 0{,}999$

Reemplazando estos valores en [20]

$$F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} = \frac{0.999/2}{(1-0.999)/2} = 1551.958$$

El valor teórico para F con 2 grados de libertad en el numerador y denominador, con un nivel de confianza de 0,95 es

$$F^*_{2,2;0.95} = 19$$

La prueba indica que $F > F^*_{2,2;0.95}$

Por lo que se rechaza la hipótesis nula, lo cual quiere decir que el conjunto de variables explican el comportamiento de la dependiente.

2) Si la hipótesis que se desea contrastar se refiere al valor de un *solo coeficiente,* la hipótesis nula es del tipo:

$$\begin{aligned} &H_0 : \beta_i = \beta_{i0} \\ &H_1 : \beta_i \neq \beta_{i0} \end{aligned} \qquad [21]$$

La matriz **R** del estadístico sería un vector **1xk** de la forma $(0,\ldots,1,\ldots,0)$, con el **1** ocupando la posición $i-ésima$, mientras que el vector **r**, vendría dado por el escalar β_{i0}.

$$\underset{qxk}{\mathbf{R}} = \underset{1xk}{\mathbf{R}} = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} \quad \mathbf{r} = \beta_{io}$$

El estadístico de contraste, teniendo en cuenta a [18], toma la forma:

$$F = \frac{\dfrac{\left[\left(\begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}\begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_i \\ \vdots \\ \hat{\beta}_k \end{bmatrix} - \beta_{i0}\right)'\left(\begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}\left(\mathbf{X'X}\right)^{-1}\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}\right)^{-1}\left(\begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}\begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_i \\ \vdots \\ \hat{\beta}_k \end{bmatrix} - \beta_{i0}\right)\right]}{q}}{\dfrac{\mathbf{e'e}}{T-k}}$$

$$F = \frac{\frac{\left[\left(\hat{\beta}_i - \beta_{i0}\right)\left(a_{ii}\right)^{-1}\left(\hat{\beta}_i - \beta_{i0}\right)\right]}{1}}{\frac{\mathbf{e'e}}{T-k}}$$

$$F = \frac{\left(\hat{\beta}_i - \beta_{i0}\right)^2}{\frac{\mathbf{e'e}}{T-k} \; a_{ii}} \overset{H_0}{\sim} F_{1,\,T-k}$$

Se conoce que la distribución F, con *1* grado de libertad en el numerador y *v* grados de libertad en el denominador, equivale al cuadrado de la distribución t con *v* grados de libertad. Por esto es posible reexpresar lo anterior en:

$$t = \frac{\hat{\beta}_i - \beta_{i0}}{\left(\frac{\mathbf{e'e}}{T-k} \; a_{ii}\right)^{1/2}} \overset{H_0}{\sim} t_{T-k} \qquad [22]$$

Donde

a_{ii} es el elemento $i-ésimo$ de la diagonal principal de la matriz $(X'X)^{-1}$,

$\hat{\sigma}_{\varepsilon}^2 a_{ii} = S_{\hat{\beta}_i}^2$ es la varianza estimada del estimador $\hat{\beta}_i$; y

t_{T-k} indica la distribución t de student de $T-k$ grados de libertad.

En el caso particular de que $\beta_{i0} = 0$, se está contrastando la *significación individual* de la variable explicativa X_t.

La hipótesis nula del contraste es que, el efecto marginal de un cambio en X_t sobre el *valor medio* de Y_t es nulo:

$$H_0 = \beta_i = 0$$
$$H_1 = \beta_i \neq 0$$

El estadístico de contraste, toma la forma:

$$t = \frac{\hat{\beta}_i}{S\ a_{ii}^{1/2}} \overset{H_0}{\sim} t_{T-k}$$

Regla de decisión: Rechazar la hipótesis nula a un nivel de significación α, si:

$$\left| \frac{\hat{\beta}_i}{S_{\hat{\beta}_i}} \right| > t_{\alpha/2;T-k}$$

Donde $t_{\alpha/2;T-k}$, es la ordenada de la distribución t de student de $T-k$ grados de libertad, que deja a la derecha una probabilidad de $\alpha/2$.

Ejemplo 16.1.d comprobar la significatividad individual de las variables

Las hipótesis, para el Ejemplo 15.5, son

$H_0: \beta_2 = 0$	$H_0: \beta_3 = 0$
$H_1: \beta_2 \neq 0$	$H_1: \beta_3 \neq 0$

Teniendo en cuenta los resultados alcanzados en 15.5.d y 15.5.g, los estadísticos respectivos serán

$$t_{\beta_2} = \frac{1.03356 - 0}{\sqrt{0.00036}} = 54.487$$

$$t_{\beta_3} = \frac{0.34 - 0}{\sqrt{0.00086}} = 11.6093$$

El valor teórico de t para 2 grados de libertad y un nivel de confianza de 0,95 es

$t_{2;0,95} = 4{,}303$

β_2 y β_3 son significativamente distintos de cero, por lo que las variables $\mathbf{X_2}$ y $\mathbf{X}_3$ son significativas.

En general, y en este ejemplo en particular, se puede armar la tabla de análisis de la varianza para definir el estadístico F. Se plantea inicialmente la Tabla teórica y luego se completa con los datos del ejemplo:

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de suma de cuadrados
X	*SCE*	k-1	SCE/2
Residuos	*SCR*	T-k	SCR/2
Total	*SCT*	T-1	

Tabla 16.1. Análisis de la varianza para la regresión

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de suma de cuadrados
X	133.11	2	66.5571
Residuos	0,08577	2	0,04289
Total	133.1	4	

Tabla 16.2. Análisis de la varianza para la regresión (datos del ejemplo)

Ahora se calcula el estadístico F:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} = \frac{0.999/2}{(1-0.999)/2} = 1551.958$$

$$F = \frac{SCE/(k-1)}{SCR/(T-k)} = \frac{133.11/2}{(1-0.08577)/2} = 1551.958$$

Para ilustrar aún más la relación entre las distribuciones cabe recordar, de acuerdo con la proposición 16.1, que el estimador de cualquier parámetro de la relación lineal sigue una distribución normal univariante dada por

$$\hat{\beta}_i \sim N(\beta_i, \sigma_\varepsilon^2 a_{ii})$$

Donde, como antes, a_{ii} es el elemento i−ésimo de la diagonal principal de la matriz $(X'X)^{-1}$. Así,

$$\frac{\hat{\beta}_i - \beta_i}{\sigma_\varepsilon \sqrt{a_{ii}}} \sim N(0,1)$$

Y según el corolario de la proposición 16.3.

$$\frac{(T-k)S^2}{\sigma_\varepsilon^2} \sim \chi^2_{T-k}$$

Por lo que si se define el estadístico t,

$$t = \frac{\dfrac{\hat{\beta}_i - \beta_i}{\sigma_\varepsilon \sqrt{a_{ii}}}}{\dfrac{S\sqrt{T-k}}{\sigma_\varepsilon \sqrt{T-k}}}$$

queda

$$t = \frac{\hat{\beta}_i - \beta_i}{S\sqrt{a_{ii}}} \sim t_{T-k}; \quad \forall i = 1,\dots,k$$

Ejemplo 16.2 Para estimar el modelo

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \varepsilon_t; \forall t = 1,\dots,90$$

Se dispone de las matrices

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} 5 & -3 & 2 & 0 \\ -3 & 6 & -2 & -4 \\ 2 & -2 & 4 & 3 \\ 0 & -4 & 3 & 4 \end{bmatrix}; \quad \mathbf{X'y} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 2 \end{bmatrix}$$

Se conoce además que: $Y'Y = \sum_{t=1}^{90} Y_t^2 = 80$

y que la estimación MCO del modelo de regresión lineal es

$$\hat{Y}_t = \underbrace{11}_{(1.6)} - \underbrace{7}_{(1.7)} X_{2t} + \underbrace{12}_{(1.4)} X_{3t} + \underbrace{3}_{(1.4)} X_{4t}$$

Donde entre paréntesis se informa del desvío estándar de cada estimador.

Con esta información se solicita:

a) Calcular la SCR (tenga en cuenta que la SCE se puede escribir como $\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}$

b) Contrastar la significatividad individual de cada una de las variables del modelo.

c) Contrastar la significatividad conjunta de las variables del modelo

d) Contrastar la restricción lineal sobre el modelo, la que está expresada en la siguiente hipótesis, a un nivel de confianza de 0.95,

$H_0 = \beta_3 + 2\beta_2 = 3$

Los apartados a), b) y c) se dejan como ejercicio al lector. Para contrastar la hipótesis formulada en el apartado d) se deben definir las siguientes matrices, realizar y comprobar los cálculos correspondientes

$$\mathbf{R} = [0 \quad 2 \quad 1 \quad 0] \qquad \mathbf{r} = [3]$$

$$F = \frac{\left(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right)'\left(\mathbf{R}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{R}'\right)^{-1}\left(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right)/q}{\mathbf{e}'\mathbf{e}/T-k} = \frac{(-5)\frac{1}{20}(-5)}{\frac{1}{2}} = 2.5$$

Para el nivel de confianza establecido ¿los coeficientes β_3, β_2 satisfacen la restricción lineal planteada?

16.3. El modelo en forma de desviaciones

Es oportuno comenzar este tema con una observación. Algunos autores informan que el coeficiente de determinación varía en el intervalo $[0,1]$ *siempre y cuando* el modelo lineal tenga coeficiente o término independiente. Esta afirmación es errónea, habida cuenta de que el modelo de regresión lineal siempre tiene término independiente excepto cuando se especifica en forma de desviaciones (o variables desvíos), pero aún en este caso el coeficiente R^2 está comprendido entre 0 y 1.

Observación: Se dice lo anterior ya que el estimador mínimo cuadrado ordinario del modelo de regresión lineal es, como se ha demostrado, el mejor estimador lineal, insesgado y óptimo. La última propiedad dice que es el de menor varianza de entre todos los estimadores lineales e insesgados posibles. Esto, como se verá, se mantiene si el modelo se especifica en forma de desviaciones, pero es falso cuando el modelo se formula sin término independiente.

Se demostrará esta afirmación utilizando la regla del absurdo. Si se define el modelo sin término independiente,

$$Y_t = \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + \beta_k X_{kt} + \varepsilon_t; \quad \forall t = 1, \ldots, T$$

La diferencia entre este modelo y el formulado habitualmente es el término β_1, por lo que se podría expresar el modelo sin término independiente de la siguiente manera:

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + \beta_k X_{kt} + \varepsilon_t - \beta_1 X_{1t}; \quad \forall t = 1, \ldots, T$$

Donde X_{1t} es un vector de unos, **T x 1**, por lo que

$$\boldsymbol{\beta}_1\mathbf{X}_1 = \begin{bmatrix} \beta_1 \\ \beta_1 \\ \vdots \\ \beta_1 \end{bmatrix} = \boldsymbol{\beta}_1$$

Se puede demostrar que el estimador del modelo reformulado es

$$\hat{\hat{\boldsymbol{\beta}}} = (\mathbf{X}`\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} + \hat{\boldsymbol{\beta}}_1)$$

Este estimador es insesgado solo si $\mathrm{E}(\hat{\boldsymbol{\beta}}_1) = \boldsymbol{\beta}_1$

Partiendo de este supuesto, la varianza de este estimador es

$$\mathbf{V}(\hat{\hat{\boldsymbol{\beta}}}) = \sigma_\varepsilon^2(\mathbf{X}`\mathbf{X})^{-1} + V(\hat{\boldsymbol{\beta}}_1) > \mathbf{V}(\hat{\boldsymbol{\beta}})$$

Es decir, no es un estimador óptimo. Dentro de la teoría econométrica, esta observación debe tenerse presente a la hora de estimar un modelo.

Por lo tanto, la única forma de especificar un modelo econométrico sin término independiente, para su posterior estimación, es la presentación en forma de desviaciones de las variables con respecto a sus medias aritméticas, cuestión que se analiza seguidamente.

Ahora se desea realizar una estimación sin término independiente, para ello se especifica el modelo en forma de desviaciones de la siguiente manera

$$y_t = \beta_2 x_{2t} + \beta_3 x_{3t} + \cdots + \beta_k x_{kt} + \varepsilon_t; \quad \forall t = 1,\ldots,T$$

Donde las letras minúsculas representan variables desvíos respecto a su media, es decir:

$$y_t = Y_t - \overline{Y}$$
$$x_{kt} = X_{kt} - \overline{X}; \forall k$$
$$\varepsilon_t = \varepsilon_t; \text{ ya que } E(\varepsilon_t) = 0$$

Se utiliza, generalmente, para estimar el modelo en dos etapas. En la primera se estiman los coeficientes de regresión –que coinciden con los estimados en la regresión habitual- y en la segunda etapa la ordenada al origen, término o coeficiente independiente.

Se puede escribir matricialmente como

$$\mathbf{Ay} = \mathbf{AX_2\beta_2} + \boldsymbol{\varepsilon}$$

Donde

$\mathbf{A} = \mathbf{I} - \left(\frac{1}{\mathrm{T}}\right)\mathbf{ii'}$; siendo $\mathbf{A}$ una matriz de transformación, simétrica e idempotente, e **i** un vector de *T* unos

$\mathbf{Ay}$, es el vector endógeno representado en forma de desviaciones

$\mathbf{AX}$, es la matriz de variables explicativas en forma de desviaciones

$\boldsymbol{\beta}_2$, es el vector de los coeficientes del modelo (sin ordenada al origen)

$\mathbf{A\varepsilon} = \boldsymbol{\varepsilon}$

$\mathbf{Ai} = \mathbf{0}$, en general premultiplicando por $\mathbf{A}$ cualquier vector cuyos elementos sean idénticos, da como resultado el vector nulo

Por lo tanto,

$$\mathbf{A} = \mathbf{I} - \left(\frac{1}{T}\right)\mathbf{ii'} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} - \frac{1}{T}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

Las variables del modelo sufren una transformación; por ejemplo, para el caso de la variable endógena se tiene:

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} \Rightarrow \frac{1}{T}\mathbf{ii'}\,\mathbf{y} = \frac{1}{T}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} \overline{Y} \\ \overline{Y} \\ \vdots \\ \overline{Y} \end{bmatrix}$$

Entonces, $\mathbf{Ay} = \left\{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} - \frac{1}{T}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}\right\}\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} - \frac{1}{T} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix}$$

$$= \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} - \begin{bmatrix} \overline{Y} \\ \overline{Y} \\ \vdots \\ \overline{Y} \end{bmatrix} = \begin{bmatrix} Y_1 - \overline{Y} \\ Y_2 - \overline{Y} \\ \vdots \\ Y_T - \overline{Y} \end{bmatrix}$$

De modo que

$$\mathbf{Ay} = \mathbf{y} - \mathbf{ii'y} = \begin{bmatrix} Y_1 - \overline{Y} \\ Y_2 - \overline{Y} \\ \vdots \\ Y_T - \overline{Y} \end{bmatrix}$$

Lo mismo se puede realizar con cada una de las variables exógenas del modelo. Por lo que el modelo estimado se puede escribir como:

$$\mathbf{Ay} = \mathbf{AX_2}\hat{\boldsymbol{\beta}}_2 + \mathbf{e}$$

Observación: Para demostrar esta última igualdad se parte del hecho de que el estimador MCO $\hat{\boldsymbol{\beta}}$ y el vector de residuos están ligados por $\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$.

Si se realiza la partición de la matriz $\mathbf{X}$ como

$$\mathbf{X} = [\mathbf{x}_1 | \mathbf{X}_2]$$

Donde

$\mathbf{x}_1$, es un vector columna de unos

$\mathbf{X}_2$, es la matriz $\mathbf{Tx(k-1)}$ de observaciones de las variables $X_2, X_3, \ldots, X_k$

Entonces se puede reescribir la relación entre el estimador y el vector de residuos de la siguiente manera

$$\mathbf{y} = \mathbf{x}_1\hat{\beta}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{e}$$

Con lo que, $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$

Premultiplicando por $\mathbf{A}$ da,

$$\mathbf{Ay} = \mathbf{AX}_2\hat{\boldsymbol{\beta}}_2 + \mathbf{e}$$

Con lo que queda demostrado. En la última expresión, se ha usado el resultado general: cualquier vector cuyos elementos sean idénticos premultiplicado por $\mathbf{A}$ es igual al vector nulo.

Con este resultado a la vista se puede reescribir el modelo en forma de desviaciones como

$$\mathbf{y}_d = \mathbf{X}_d\hat{\boldsymbol{\beta}}_2 + \mathbf{e}$$

Donde los subíndices indican que el modelo esta expresado en forma de desviaciones con respecto a la media. Como $\mathbf{X'e} = \mathbf{0}$, resulta que $\mathbf{X'}_d\,\mathbf{e} = \mathbf{0}$. Por lo que premultiplicando el modelo en forma de desviaciones por $\mathbf{X'}_d$ se obtiene

$$\mathbf{X'}_d\,\mathbf{y}_d = (\mathbf{X'}_d\,\mathbf{X}_d)\hat{\boldsymbol{\beta}}_2 \qquad [23]$$

Que son las conocidas *ecuaciones normales*, excepto que los datos están en forma de desviaciones y que el vector de estimadores incluye solo los coeficientes de la pendiente y excluye el término independiente. Para obtener este último, luego del proceso de estimación, se puede premultiplicar $\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$ por $\frac{1}{T}\mathbf{i'}$,

$$\frac{1}{T}\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}\begin{bmatrix}Y_1\\ Y_2\\ \vdots\\ Y_T\end{bmatrix} = \frac{1}{T}\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}\begin{bmatrix}1 & X_{21} & X_{31} & \cdots & X_{k1}\\ 1 & X_{22} & X_{32} & \cdots & X_{k2}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & X_{2T} & X_{3T} & \cdots & X_{kT}\end{bmatrix}\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\\ \vdots\\ \hat{\beta}_T\end{bmatrix} + \frac{1}{T}\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}\begin{bmatrix}e_1\\ e_2\\ \vdots\\ e_T\end{bmatrix}$$

Da lugar a

$$\overline{Y} = \begin{bmatrix} 1 & \overline{X}_2 & \overline{X}_3 & \cdots & \overline{X}_k \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

Que puede expresarse como

$$\overline{Y} = \hat{\beta}_1 + \hat{\beta}_2 \overline{X}_2 + \hat{\beta}_3 \overline{X}_3 + \cdots + \hat{\beta}_k \overline{X}_k$$

donde, $\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2 \overline{X}_2 - \hat{\beta}_3 \overline{X}_3 - \cdots - \hat{\beta}_k \overline{X}_k$ [24]

También se puede expresar la descomposición de la *suma de cuadrados* como,

$$\begin{aligned} \mathbf{y'_d\, y_d} &= \mathbf{\hat{\beta}'_2\, X'_d\, X_d \hat{\beta}_2} + \mathbf{e'\, e} \\ SCT &= SCE + SCR \end{aligned}$$

El *coeficiente de correlación múltiple,* R, se define como la raíz cuadrada positiva de

$$R^2 = \frac{SCE}{SCT} = \frac{\mathbf{\hat{\beta}'_2\, X'_d\, X_d \hat{\beta}_2}}{\mathbf{y'_d\, y_d}} \qquad [25]$$

Ejemplo 16.3 Con los siguientes datos muestrales, en forma de desviaciones, realice la estimación y obtenga los coeficientes de determinación y de correlación múltiple y los coeficientes de correlación parcial del modelo

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \varepsilon_t; \quad \forall t = 1,\ldots,5$$

$$y_d = \begin{bmatrix} -1 \\ -3 \\ 4 \\ -1 \\ 1 \end{bmatrix}; \quad X_d = \begin{bmatrix} 0 & 0 \\ -2 & -1 \\ 2 & 1 \\ -1 & -1 \\ 1 & 1 \end{bmatrix}$$

Demuestre que las ecuaciones normales son

$$\begin{bmatrix} 10 & 6 \\ 6 & 4 \end{bmatrix} \cdot \begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 16 \\ 9 \end{bmatrix}$$

Luego obtenga la solución para el vector de estimadores.

Partiendo del vector $\mathbf{y}_d$ compruebe que la suma de cuadrados totales es igual a 28. Obtenga la suma de cuadrados residual, la suma de cuadrados explicada, el coeficiente de determinación corregido y el coeficiente de correlación múltiple.

Las correlaciones parciales cobran importancia en caso de dos o más regresores. Si se trabajan los datos en forma de desviaciones, se puede calcular el residuo parcial de la regresión entre la variable dependiente y, por ejemplo, X_3, de la siguiente manera: $e_{yx_3} = y - \hat{\beta}_{yx_3} x_3$,

Donde $\hat{\beta}_{yx_3} = \frac{\sum yx_3}{\sum x_3^2}$,

se denomina coeficiente de regresión parcial, en este caso entre Y y X_3.

El coeficiente de correlación parcial, entre Y y X_3, se define como el cociente de correlación entre ambos conjunto de residuos. Se indica como $r_{YX_3|X_2.}$.

Su cálculo se realiza mediante la siguiente expresión:

$$r_{YX_3|X_2.} = \frac{r_{YX_2} - r_{YX_3} r_{X_2X_3}}{\sqrt{1-r_{YX_2}^2}\sqrt{1-r_{X_3X_2}^2}}$$

De manera similar se puede calcular $r_{YX_2|X_3.}$.

El primero mide la asociación entre Y y X_3 una vez eliminada la influencia ejercida por X_2, mientras que el segundo mide la asociación entre X_2 y X_3 cuando desaparece cualquier efecto que pueda ejercer la variable endógena.

Los coeficientes de correlación simple como $r_{YX2}, r_{YX3}, r_{X_2X_3}$ se suelen denominar coeficientes de orden cero; mientras que, los coeficientes de correlación parcial reciben el nombre de coeficientes de primer orden.

Realice el cálculo con los datos del ejemplo. ¿Podría llegarse al mismo resultado si en lugar de los coeficientes de correlación simple se usaran los residuos parciales?. Compruébelo.

Con los datos del ejemplo, también se puede calcular la suma de cuadrados totales en forma secuencial. De la siguiente manera:

Fuente de variación	Suma de cuadrados	Fuente de variación	Suma de cuadrados
X_2	$r^2_{YX_2}\Sigma y^2$	X_3	$r^2_{YX_3}\Sigma y^2$
Incremento debido a X_3	$r^2_{YX_3 \mid X_2 \cdot}\left(1-r^2_{YX_2}\right)\Sigma y^2$	Incremento debido a X_2	$r^2_{YX_2 \mid X_3 \cdot}\left(1-r^2_{YX_3}\right)\Sigma y^2$
X_2 y X_3	$R^2\Sigma y^2$	X_2 y X_3	$R^2\Sigma y^2$
Residuos	$(1-R^2)\sum y^2$	Residuos	$(1-R^2)\sum y^2$

Reemplace las fórmulas por números y obtenga el resultado correspondiente. También compruebe que las sumas de cuadrados explicada, total y residual coinciden con los resultados anteriores.

Cuando hay dos o más variables explicativas, no existe modo de determinar la importancia relativa que cada una de las variables tiene para explicar las variaciones de Y.

Kruskal (1987) considera varios métodos para evaluar la importancia de las distintas variables explicativas. Su propuesta se centra en el promedio de los cuadrados de los coeficientes de correlación simple y parcial sobre los distintos momentos posibles de introducir las variables explicativas. En cada etapa, los coeficientes de correlación al cuadrado relevantes indican la proporción de varianza explicada por una variable X específica. Con los datos del ejemplo, se tiene

Proporción media para $X_2 = \left(r^2_{YX_2} + r^2{}_{YX_3|X_2}.\right)/2$

Proporción media para $X_3 = \left(r^2_{YX_3} + r^2{}_{YX_2|X_3}.\right)/2$

Obtenga dichos valores y demuestre que según los coeficientes medios de Kruskal, en nuestro ejemplo, el papel de X_2 es más importante que el de X_3 a la hora de determinar Y.

Una forma alternativa, de ver las contribuciones individuales, fue introducida por Tinbergen en su diagrama utilizado en el estudio de los ciclos de negocios. Trabajando con los datos del ejemplo y las variables en forma de desvíos realice cuatro gráficos. En el primero dibuje lo valores de y con los valores de $\hat{y}$ (compare estos gráficos con los que hubiera obtenido desde las relativas cíclicas e irregulares); en el segundo dibuje $\hat{\beta}_2 x_2$; en el tercero $\hat{\beta}_3 x_3$ y, finalmente, grafique los residuos de la regresión. Llega Tinbergen a la misma conclusión que Kruskal. ¿Porqué?.

16.4. Predicción en el modelo lineal

Una vez que mediante los métodos econométricos de estimación se ha asignado valores numéricos a los parámetros, el modelo puede utilizarse con dos objetivos: a. Descripción de la economía, de la que procede la información muestral (Familia, Región, País, etc.), y b.Predicción.

La fiabilidad de la Predicción dependerá de:

- el horizonte de predicción
- la constancia de los valores paramétricos estimados a lo largo del horizonte de predicción
- la calidad de las estimaciones de los parámetros del modelo
- que el modelo utilizado sea apropiado y que, en particular, esté especificado correctamente.

Ejemplo 16.11. Un análisis de predicción es fundamental para hacer cualquier estudio de política económica. Si el Banco Central quiere hacer un análisis del posible efecto inflacionario de una expansión monetaria; se usarán predicciones de los tipos de interés (que explican los posibles gastos de consumo) y de los gastos de consumo utilizando un determinado supuesto sobre el crecimiento de la oferta monetaria.

Para que el modelo estimado sea adecuado para predecir valores futuros de Y, se ha de suponer que la relación lineal entre Y y X se mantiene también en el período de predicción.

Bajo este supuesto de estabilidad y dados unos valores conocidos de las variables exógenas en el período de predicción, $\mathbf{x_P}$, la predicción por punto de Y_P, será:

$$\hat{Y}_p = \mathbf{x}_\mathbf{p}' \hat{\boldsymbol{\beta}} \qquad [26]$$

Este predictor es lineal, ya que es una combinación lineal de las observaciones Y.

El error de predicción correspondiente es:

$$e_p = Y_\mathbf{p} - \hat{Y}_\mathbf{p} = \mathbf{x}_\mathbf{p}' \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right) + u_p \qquad [27]$$

Este error de predicción incluye dos componentes, uno relacionado con el error en la estimación de $\boldsymbol{\beta}$ y otro inherente a la parte estocástica del modelo.

Bajo los supuestos habituales se tiene que el error de predicción sigue una distribución normal con media:

$$E(e_p) = E\left(\mathbf{x}_\mathbf{p}' (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + u_p\right) = 0 \qquad [28]$$

Por lo que el predictor $\hat{Y}_P$ es insesgado.

En cuanto a su varianza:

$$\sigma_e^2 = \sigma^2 \left(1 + \mathbf{x}_\mathbf{p}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_\mathbf{p}\right) \qquad [29]$$

El intervalo de confianza $1-\alpha$ para Y_P, será:

$$P\left[\mathbf{x}_\mathbf{p}' \hat{\boldsymbol{\beta}} \pm t_{\alpha/2(T-k)} \left(\hat{\sigma}^2 \left(1 + \mathbf{x}_\mathbf{p}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_\mathbf{p}\right)\right)^{1/2}\right] = 1-\alpha \qquad [30]$$

Este intervalo de confianza es aleatorio, ya que depende de los estimadores $\hat{\boldsymbol{\beta}}$ y $\hat{\sigma}^2$.

Si se contara con 100 muestras diferentes, se podría construir 100 intervalos de confianza, de los cuales $1-\alpha$ contendrían el verdadero valor de Y_P.

Ejemplo 16.1.e Para los datos suministrados en el Ejemplo 15.5, se quiere un intervalo de confianza del 95 por ciento para $\hat{Y}_{t+1}$ dado $X_{2,t+1} = 8$, $X_{3,t+1} = 4$; el intervalo sería:

$$[1 \quad 8 \quad 4]\begin{bmatrix} 2.4185 \\ 1.0336 \\ 0.34 \end{bmatrix} \pm 4.30\sqrt{0.0429}\sqrt{1 + [1 \quad 8 \quad 4]\begin{bmatrix} 0.8854 & 0.0554 & -0.08 \\ 0.0554 & 0.0084 & 0 \\ -0.08 & 0 & 0.02 \end{bmatrix}\begin{bmatrix} 1 \\ 8 \\ 4 \end{bmatrix}}$$

El intervalo se encuentra entre $[11.0647 \;\; 13.0293]$

A veces no se está interesado en predecir tanto el valor futuro de la variable endógena Y_P como su valor $E(Y_P)$. La predicción por punto es, operacionalmente, similar a la expresada en [26]:

$$\hat{E}(Y_{\mathbf{p}}) = \mathbf{x}_{\mathbf{p}}^{'}\,\hat{\boldsymbol{\beta}} \qquad [31]$$

el error de predicción asociado es

$$V_p = E(Y_p) - \hat{E}(Y_p) = \mathbf{x}_{\mathbf{p}}^{'}\,\boldsymbol{\beta} - \mathbf{x}_{\mathbf{p}}^{'}\,\hat{\boldsymbol{\beta}} = \mathbf{x}_{\mathbf{p}}^{'}\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right) \qquad [32]$$

y la varianza

$$\sigma_V^2 = \hat{\sigma}^2\,\mathbf{x}_{\mathbf{p}}^{'}\left(\mathbf{X}^{'}\mathbf{X}\right)^{-1}\mathbf{x}_{\mathbf{p}} \qquad [33]$$

Este error de predicción es debido, solamente, al error en la estimación de $\boldsymbol{\beta}$, por lo que $\sigma_V^2 \le \sigma_e^2$ y el intervalo de predicción para $E(Y_P)$:

$$P\left[\mathbf{x}_p^{'}\ \hat{\boldsymbol{\beta}} \pm t_{\alpha/2\,(T-k)}\ \left(\hat{\sigma}^2\ \mathbf{x}_p^{'}\left(\mathbf{X}^{'}\ \mathbf{X}\right)^{-1}\mathbf{x}_p\right)^{1/2}\right] = 1-\alpha \qquad [34]$$

Va a ser más estrecho que el correspondiente a la variable endógena, Y_P.

Ejemplo 16.1.f Para los datos del Ejemplo 15.5, calcula el intervalo de confianza

CASOS DE ESTUDIO, PREGUNTAS Y PROBLEMAS

Caso 16.1: Cálculo de R^2

Utilice la información que se suministra para el cálculo de R^2.

$\sum X = 1700 \quad \sum Y = 1110 \quad n = 100$

$\sum xy = 205500 \quad \sum x^2 = 322000 \quad \sum y^2 = 132100$

Resuelva calculando el coeficiente de correlación y elevando al cuadrado; o, especificando el modelo en forma de desviaciones, por el cociente entre la $SCE = \sum \hat{y}^2 = \beta \sum xy$ y la $SCT = \sum y^2$.

Caso 16.2: Modelo de Inversión de empresas españolas entre 1959-1971

Sobre la base de la información que se adjunta, se pide:

a) Cálculo del desvío estándar de la regresión
b) Cálculo del desvío de $\hat{\alpha}_1$
c) Cálculo del coeficiente de determinación
d) Contraste $\alpha_4 = -55$, nivel de confianza de 0.95
e) Se supone que la utilización de la capacidad productiva en 1972 se ubicará en sus niveles medios; mientras que se espera una reducción del 10% en los niveles de cash flow y capital, y del 5% en los niveles de renta y rendimiento de las obligaciones. En este contexto, ¿cuál es el valor esperado para la inversión?
f) Contraste la hipótesis conjunta de que $\alpha_1 = -0{,}7 \quad \alpha_2 = 0{,}04 \quad \alpha_3 = 0{,}15 \quad \alpha_4 = -40 \quad \alpha_5 = 0{,}20$

Tabla de datos

Obs	CAPITAL	CASHFLOW	INVER	RENDOBLI	RENTA	UCP
1958	989.2	88.4		6.7583	964.6	
1959	1032.9	82.8	87.3	6.9422	943.8	78
1960	1088.4	90.2	98.8	6.95	951.8	74.7
1961	1154.6	103.9	114.2	6.945	1059.3	80
1962	1240.4	115.3	136.6	5.9379	1161.9	83.2
1963	1345.3	115.7	159.3	6.1638	1272.1	85.21
1964	1461.3	126.6	176.2	5.9756	1355.3	84.9
1965	1611.7	138.6	220.2	5.9787	1457.9	84
1966	1778.9	155	250.3	6.2188	1569.1	83
1967	1943.6	150.9	261.2	6.3087	1639	80
1968	2111.3	178.9	271.8	6.4047	1733.2	80.5
1969	2297.4	192.6	303.6	7.1207	1849.6	83.75
1970	2484.7	210.9	317.3	7.558	1957.6	83.75
1971	2647.8	234.4	305	8.6722	2053.3	82.2

Descripción de las variables

	CAPITAL	CASHFLOW	INVER	RENDOBLI	RENTA	UCP
Mean	1656.250	141.7286	207.8308	6.709614	1426.321	81.78538
Median	1536.500	132.6000	220.2000	6.581500	1406.600	83.00000
Maximum	2647.800	234.4000	317.3000	8.672200	2053.300	85.21000
Minimum	989.2000	82.80000	87.30000	5.937900	943.8000	74.70000
Std. Dev.	561.7621	47.73313	83.21946	0.753005	384.3025	3.031231
Skewness	0.436316	0.532956	-0.116699	1.245159	0.174604	-1.003863
Kurtosis	1.841288	2.162788	1.518098	4.252310	1.715203	3.251134
Jarque-Bera	1.227392	1.071638	1.219026	4.532479	1.034046	2.217600
Probability	0.541346	0.585190	0.543616	0.103701	0.596293	0.329955
Sum	23187.50	1984.200	2701.800	93.93460	19968.50	1063.210
Sum Sq. Dev.	4102496.	29619.87	83105.75	7.371223	1919949.	110.2603
Observations	14	14	13	14	14	13

Especificación del modelo:

$$I_t = \alpha_0 + \alpha_1 C_t + \alpha_2 Y_{t-1} + \alpha_3 CF_{t-1} + \alpha_4 RO + \alpha_5 K_{t-1} + \mu_t$$

donde:

$I_t = INVER =$ Inversión en miles de millones de pesetas de 1969

$C = UCP =$ Utilización de la capacidad productiva del periodo en %

$Y = RENTA =$ Renta en miles de millones de pesetas de 1969

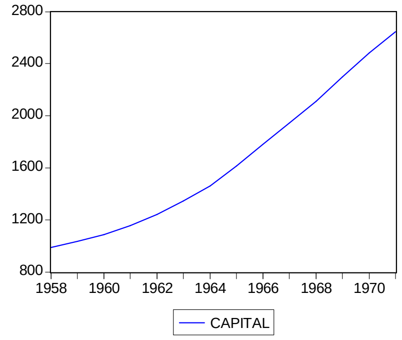
$CF = CASHFLOW =$ Cash-flow de las empresas en miles de millones de pesetas de 1969

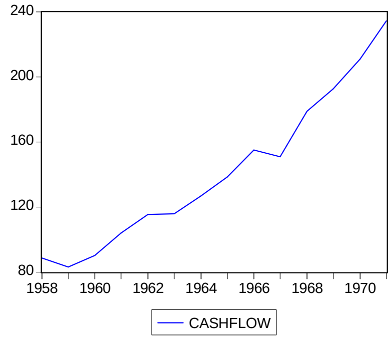
$RO = RENDOBLI =$ Rendimiento de las obligaciones en %

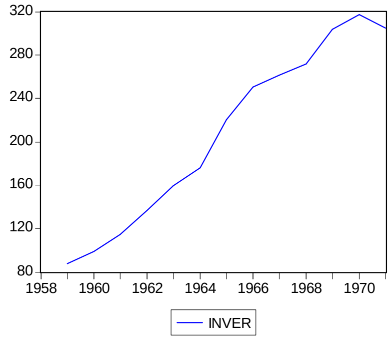
$K = CAPITAL =$ Stock de capital en miles de millones de pesetas de 1969

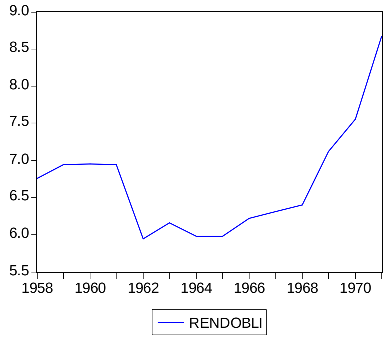
Matriz inversa de X´X

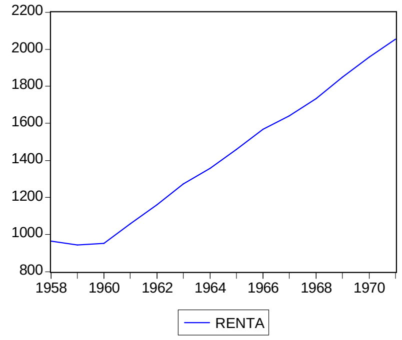
		CAPITAL	CASHFLOW	RENDOBLI	RENTA	UCP
	305,581292	0,02607461	0,9307289	-14,0134991	-0,12074648	-2,58732386
CAPITAL	0,02607461	5,8854E-05	-0,00012317	-0,00380121	-6,5333E-05	0,00015946
CASHFLOW	0,9307289	-0,00012317	0,0058293	-0,04112714	-0,00042482	-0,00806403
RENDOBLI	-14,0134991	-0,00380121	-0,04112714	0,92270685	0,00879191	0,08866388
RENTA	-0,12074648	-6,5333E-05	-0,00042482	0,00879191	0,00013049	0,00051801
UCP	-2,58732386	0,00015946	-0,00806403	0,08866388	0,00051801	0,02582799

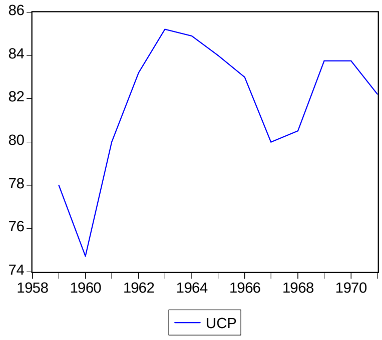












Matriz de varianzas y covarianzas de los coeficientes

	C	UCP	RENTA(-1)	CASHFLOW (-1)	RENDOBLI	CAPITAL(-1)
C	55260.180	-467.88199	-21.8353	168.30954	-2534.1489	4.71523
UCP	-467.8819	4.6706	0.09367	-1.45827	16.03364	0.028837
RENTA(-1)	-21.8353	0.0937	0.02359	-0.07682	1.58989	-0.011815
CASHFLOW(-1)	168.3095	-1.4583	-0.07682	1.05415	-7.437279	-0.022273
RENDOBLI	-2534.1489	16.0336	1.58989	-7.437279	166.8588	-0.687395
CAPITAL(-1)	4.7152	0.0288	-0.01181	-0.02227	-0.68739	0.01064

Estimación del modelo de inversión

Dependent Variable: INVER
Method: Least Squares
Date: 09/04/06 Time: 10:54
Sample(adjusted): 1959 1971
Included observations: 13 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	142.6521	235.0748	0.606837	0.5631
UCP	-0.634780	2.161166	-0.293721	0.7775
RENTA(-1)	0.075908	0.153615	0.494142	0.6363
CASHFLOW(-1)	0.281594	1.026718	0.274267	0.7918
RENDOBLI	-30.10861	12.91739	-2.330860	0.0525
CAPITAL(-1)	0.111703	0.103164	1.082771	0.3148

R-squared	0.984768	Mean dependent var	207.8308
Adjusted R-squared	0.973888	S.D. dependent var	83.21946
S.E. of regression	13.44754	Akaike info criterion	8.339507
Sum squared resid	1265.854	Schwarz criterion	8.600253
Log likelihood	-48.20679	F-statistic	90.51270
Durbin-Watson stat	2.130300	Prob(F-statistic)	0.000003

Caso 16.3: Modelo de Inversión para Argentina

Especifica y estima un modelo para estudiar la inversión en Argentina.

BIBLIOGRAFIA

- **Johnston, J. y Dinardo, J. .***Métodos De Econometría.* Barcelona: Editorial Vicens Vives, 2001.
- **Kruskal, W.** "Relative Importante by Averagin over Orderings." *The American Statiscian*, 1987.
- **Novales, A.***Econometría*. Madrid: McGraw Hill, 1993.
- **Pulido San Román, Antonio.***Modelos Econométricos*. Madrid: Editorial Pirámide, 1993.
- **Tinbergen, J. .** "Bussiness Cycles in the United Status of America, 1919 – 1932."*League of Nations*, 1939.