ECONOMETRÍA

JOSÉ ALBERTO MAURICIO

2

REGRESIÓN LINEAL MÚLTIPLE I

2.3 MCO - PROPIEDADES ESTADÍSTICAS

Departamento de Análisis Económico y Economía Cuantitativa
Universidad Complutense de Madrid

EctrGr-JAM-2-3.pdf

ucm.randomshock.com/ectrgr

Versión 2.4 - Enero 2024

MATERIAL AUXILIAR Y COMPLEMENTARIO

El instrumental matemático que se emplea en esta Sección 2.3 es el mismo que el utilizado en las Secciones 2.1 - 2.2. El instrumental estadístico que se emplea en esta Sección 2.3 tiene que ver con variables aleatorias (escalares, vectoriales y matriciales), distribuciones de probabilidad (univariantes y multivariantes), e inferencia estadística (estimación y contrastes de hipótesis). Se puede repasar todo ello en Hill, Griffiths, Lim (2018): Apéndices A - D, y en Wooldridge (2020): Apéndices A - D. En el Apéndice 2 de esta Sección 2.3 se presenta un breve resumen sobre variables aleatorias y distribuciones de probabilidad.

Adicionalmente, en esta Sección 2.3 se hace referencia a los Ejercicios 2 - 4 disponibles en la página de la asignatura en Internet. A su vez, en esos ejercicios se mencionan varias secciones de la guía *Introducción al Uso de EViews 4.1* de cara a la resolución de las preguntas que se plantean en ellos.

BIBLIOGRAFÍA PARA TODO EL TEMA 2

Hayashi (2000): Capítulo 1.
Hill, Griffiths, Lim (2018): Capítulos 2 - 6.
Wooldridge (2020): Capítulos 2 - 6 y Apéndice E.

PARTE 1 - INTRODUCCIÓN

El criterio MCO es un criterio razonable para estimar los parámetros de cualquier modelo RLM. No obstante, la cuestión central para establecer la calidad de un modelo estimado (más allá de su grado de ajuste medido a través de algunos coeficientes de determinación) es la que tiene que ver con cómo evaluar el grado de proximidad entre lo que realmente valen los parámetros de un modelo (cuyos verdaderos valores no se conocen) y las estimaciones numéricas de dichos parámetros calculadas por MCO. La respuesta a esa cuestión puede obtenerse, en cierto sentido, a través de las **propiedades estadísticas** del criterio MCO, que dependen de las **hipótesis** que puedan asumirse sobre los elementos (datos y modelos) implicados en dicho criterio. Como se muestra en las páginas siguientes, las propiedades estadísticas del criterio MCO son óptimas bajo un conjunto de hipótesis que pueden no ser asumibles en algunos casos, lo que implica que la evaluación crítica de dichas hipótesis es un paso crucial en cualquier análisis econométrico aplicado.

PARTE 2 - LAS CINCO HIPÓTESIS CLÁSICAS A TRAVÉS DE UN EJEMPLO

Pretendemos estimar el efecto causal de las horas de asistencia a clase (HA) sobre las notas finales (NF) en una población de estudiantes de cierta asignatura en un curso académico determinado.

Datos:

$$\begin{bmatrix} nf_1 & ha_1 \\ nf_2 & ha_2 \\ \vdots & \vdots \\ nf_N & ha_N \end{bmatrix} = [\mathbf{y}, \mathbf{x}_2]. \qquad [1]$$

Los datos se interpretan como una muestra de N observaciones (puntos muestrales, o filas) $[nf_i, ha_i]$ $(i = 1, 2, ..., N)$, referida a N estudiantes de la población considerada.

Modelo (RLS):

$$NF = \beta_1 + \beta_2 HA + U, \qquad [2]$$

donde la perturbación U incluye todos aquellos factores que determinan (explican) las notas finales más allá de las HA, como (entre otros muchos) las horas de estudio fuera de clase (HF) y la calidad académica general (preparación, interés, responsabilidad, disciplina de estudio, …) de los estudiantes (CA).

Suponemos que los datos en [1] son una realización particular de una colección $\mathbf{W}$ de variables aleatorias (Apéndice 2),

$$\mathbf{W} = \begin{bmatrix} NF_1 & HA_1 \\ NF_2 & HA_2 \\ \vdots & \vdots \\ NF_N & HA_N \end{bmatrix} = [\mathbf{Y}, \mathbf{X}_2], \qquad [3]$$

que satisfacen las cinco “hipótesis clásicas” siguientes (considerar las NF y las HA como variables aleatorias es perfectamente asumible porque los valores de dichas características son inciertos hasta que se concretan en unos datos determinados):

[HC1] Linealidad (con respecto a los parámetros β_1, β_2)

$$NF_i = \beta_1 + \beta_2 HA_i + U_i\,,\ i = 1, 2, ..., N\,, \text{ o bien} \qquad [4]$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}\,, \text{ con} \qquad [5]$$

$$\mathbf{Y} = \begin{bmatrix} NF_1 \\ NF_2 \\ \vdots \\ NF_N \end{bmatrix},\ \mathbf{X} = [\mathbf{i}, \mathbf{X}_2] = \begin{bmatrix} 1 & HA_1 \\ 1 & HA_2 \\ \vdots & \vdots \\ 1 & HA_N \end{bmatrix},\ \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix},\ \mathbf{U} = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}. \qquad [6]$$

Cada perturbación U_i contiene, entre otros muchos factores, las HF y la CA específicas de cada estudiante (que también son variables aleatorias), de manera que el contenido del vector $\mathbf{U}$ puede ser, en principio, bastante heterogéneo. La heterogeneidad en el contenido del vector $\mathbf{U}$ se restringe considerablemente en las hipótesis [HC3] y [HC4].

[HC2] Ausencia de Multicolinealidad Exacta

$$\text{rango}(\mathbf{X}) = K\ (= 2) < N. \qquad [7]$$

En el modelo [2], esta hipótesis implica que cualesquiera que sean los datos contenidos en el vector $\mathbf{x}_2 = [ha_1, ha_2, ..., ha_N]'$, no son todos ellos iguales entre sí, de manera que la varianza muestral de $\mathbf{x}_2$ es positiva y la matriz $\mathbf{X}'\mathbf{X}$ es no singular.

[HC3] Exogeneidad Estricta

[HC3.1] $$\text{E}[U_i\,|\,\mathbf{X}] = \text{E}[U_i] \text{ para todo } i = 1, 2, ..., N. \qquad [8]$$

Como la primera columna de $\mathbf{X}$ es un vector de unos (constantes), el conjunto de información que condiciona en la primera esperanza de U_i en [8] es el formado por las HA de TODOS los estudiantes (el vector $\mathbf{X}_2$ en [3] o en [6]).

La igualdad entre las esperanzas condicional y marginal (incondicional) de U_i en [8] implica que no existe ningún tipo de correlación entre U_i y $\mathbf{X}$ ([A2.49]), de manera que:

[A] Las HF y la CA de cualquier estudiante no están correlacionadas con sus propias HA (una posibilidad bastante cuestionable).

[B] Las HF y la CA de cualquier estudiante no están correlacionadas con las HA de otros estudiantes (una posibilidad quizás razonablemente asumible).

En general, no es posible “demostrar” consideraciones del estilo de [A]-[B] en cualquier modelo econométrico. Pero sí se debe razonar cuidadosamente al respecto para intentar concluir si dichas consideraciones son asumibles o no.

[HC3.2] $$\text{E}[U_i] \text{ es la misma para todo } i = 1, 2, ..., N. \qquad [9]$$

Esta hipótesis implica que el efecto esperado de las HF y de la CA sobre las NF es el mismo para todos los estudiantes, de manera que (al menos en la muestra considerada) los estudiantes son en cierto sentido homogéneos en cuanto a sus HF y su CA.

Para compatibilizar las restricciones impuestas por las hipótesis [8]-[9] con lo que la realidad puede sugerir, es muy importante especificar modelos de tal manera que sus perturbaciones contengan un tipo de información (fundamentalmente factores difíciles o imposibles de observar) sobre la que se puedan asumir razonablemente dichas restricciones. Una mejora obvia en esta dirección en relación con el modelo [2] consistiría en incluir como variable explicativa adicional las HF, que dejarían así de formar parte de U.

[HC3.3] $$\mathrm{E}[U_i] = 0 \text{ para todo } i = 1, 2, ..., N. \qquad [10]$$

Esta hipótesis es perfectamente asumible (una vez asumidas las anteriores) en cualquier modelo con término constante. Por ejemplo, si $\mathrm{E}[U_i] = \delta \neq 0$ en [4], entonces es inmediato que $\mathrm{E}[U_i'] = 0$ en $NF_i = \beta_1' + \beta_2 HA_i + U_i'$, donde $\beta_1' = \beta_1 + \delta$, $U_i' = U_i - \delta$.

En conjunto, [8]-[10] se pueden expresar como

$$\mathrm{E}[U_i \,|\, \mathbf{X}] = 0 \text{ para todo } i = 1, 2, ..., N, \qquad [11]$$

o bien como $$\mathrm{E}[\mathbf{U} \,|\, \mathbf{X}] = \mathbf{0}. \qquad [12]$$

Las hipótesis [5] y [12] implican en conjunto que

$$\mathrm{E}[\mathbf{Y} \,|\, \mathbf{X}] = \mathrm{E}[\mathbf{X\beta} + \mathbf{U} \,|\, \mathbf{X}] = \underbrace{\mathrm{E}[\mathbf{X\beta} \,|\, \mathbf{X}]}_{\mathbf{X\beta}} + \underbrace{\mathrm{E}[\mathbf{U} \,|\, \mathbf{X}]}_{\mathbf{0}} = \mathbf{X\beta} \Rightarrow \hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \hat{\mathrm{E}}[\mathbf{Y} \,|\, \mathbf{X}]. \qquad [13]$$

Por ejemplo, un valor ajustado en el modelo [2] estimado por MCO se puede interpretar como una estimación de la NF esperada de un estudiante con unas HA determinadas.

[HC4] Perturbaciones Esféricas: Homoscedasticidad y Ausencia de Autocorrelación

$$\mathrm{E}[\mathbf{UU}' \,|\, \mathbf{X}] = \sigma^2 \mathbf{I}, \text{ con } \sigma^2 > 0. \qquad [14]$$

Junto con [12], [14] implica en primer lugar que las perturbaciones son “esféricas” porque

$$\mathrm{Var}[\mathbf{U} \,|\, \mathbf{X}] = \mathrm{E}[(\mathbf{U} - \mathrm{E}[\mathbf{U} \,|\, \mathbf{X}])(\mathbf{U} - \mathrm{E}[\mathbf{U} \,|\, \mathbf{X}])' \,|\, \mathbf{X}] = \mathrm{E}[\mathbf{UU}' \,|\, \mathbf{X}] = \sigma^2 \mathbf{I} \qquad [15]$$

es una matriz “escalar” (diagonal con sus elementos diagonales iguales entre sí),

$$\mathrm{Var}[\mathbf{U} \,|\, \mathbf{X}] = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix},$$

de manera que

$$\mathrm{Var}[U_i \,|\, \mathbf{X}] = \sigma^2 \text{ (constante) para todo } i = 1, 2, ..., N, \qquad [16]$$

$$\mathrm{Cov}[U_i, U_j \,|\, \mathbf{X}] = 0 \text{ para cualesquiera } i, j = 1, 2, ..., N \text{ con } i \neq j. \qquad [17]$$

Adicionalmente, las expresiones [5], [13] y [14] implican que

$$\mathrm{Var}[\mathbf{Y}\,|\,\mathbf{X}] = \mathrm{E}[\underbrace{(\mathbf{Y}-\mathrm{E}[\mathbf{Y}\,|\,\mathbf{X}])}_{\mathbf{X}\boldsymbol{\beta}+\mathbf{U}-\mathbf{X}\boldsymbol{\beta}}(\mathbf{Y}-\mathrm{E}[\mathbf{Y}\,|\,\mathbf{X}])'\,|\,\mathbf{X}] = \mathrm{E}[\mathbf{U}\mathbf{U}'\,|\,\mathbf{X}] = \sigma^2\mathbf{I}. \qquad [18]$$

Por lo tanto,

$$\mathrm{Var}[NF_i\,|\,\mathbf{X}] = \sigma^2 \text{ (constante) para todo } i = 1, 2, ..., N, \qquad [19]$$

$$\mathrm{Cov}[NF_i, NF_j\,|\,\mathbf{X}] = 0 \text{ para cualesquiera } i, j = 1, 2, ..., N \text{ con } i \neq j. \qquad [20]$$

Las expresiones [16] y [19] (homoscedasticidad, o varianza constante) implican, por ejemplo, que la dispersión o el grado de variación de las NF alrededor de sus valores esperados ($\beta_1 + \beta_2 HA_i$) es constante e independiente de las HA.

Las expresiones [17] y [20] (ausencia de autocorrelación) implican, por ejemplo, que las HF, la CA y, en consecuencia, las NF de dos estudiantes cualesquiera, no están correlacionadas entre sí, sean cuales sean sus HA y las de los demás estudiantes.

Las implicaciones de [14] pueden no ser asumibles en muchas circunstancias, pero esos casos (al contrario que cuando [12] es inasumible) son relativamente sencillos de tratar.

[HC5] Normalidad

$$\mathbf{U}\,|\,\mathbf{X} \sim \text{Normal}. \qquad [21]$$

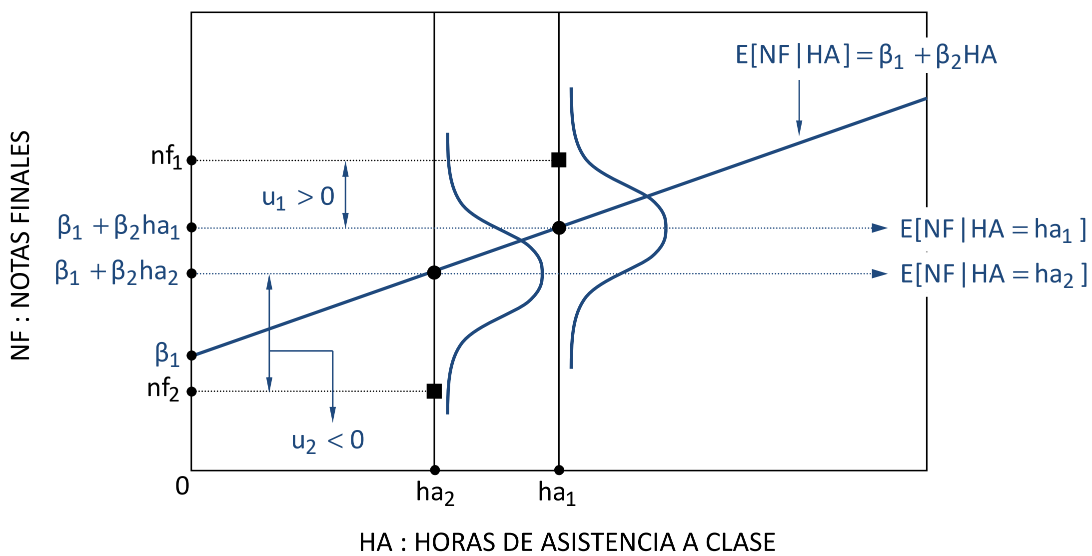
Esta última hipótesis es más "conveniente" que "necesaria" (en el sentido de que su incumplimiento tiene implicaciones que se pueden solventar en muchos casos en virtud de algún Teorema Central del Límite). En cualquier caso, [12], [14] y [21] implican en conjunto que

$$\mathbf{U}\,|\,\mathbf{X} \sim \text{Normal}\left(\mathbf{0}, \sigma^2\mathbf{I}\right), \qquad [22]$$

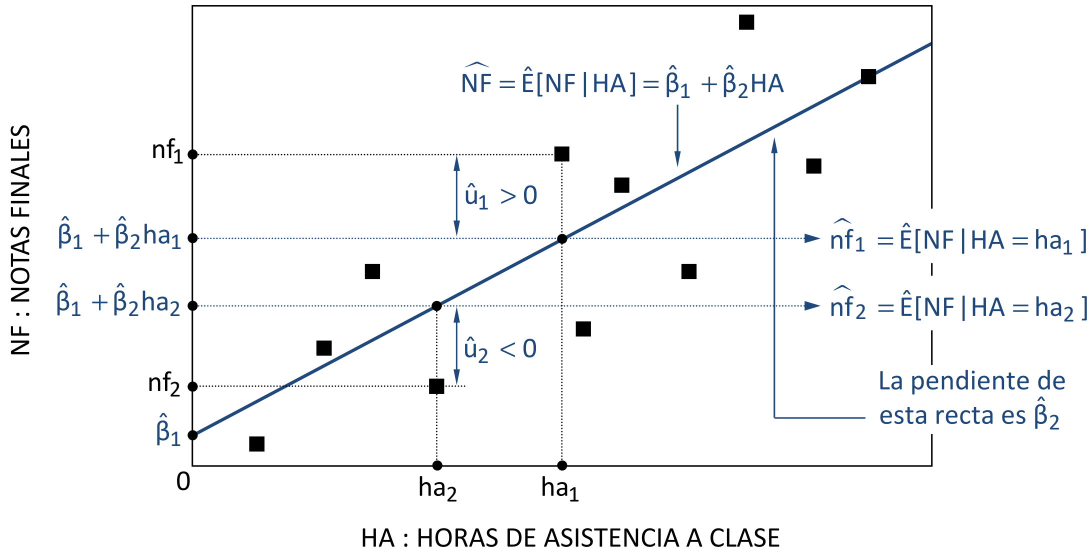
mientras que [5], [13], [18] y [21] implican en conjunto que

$$\mathbf{Y}\,|\,\mathbf{X} \sim \text{Normal}\left(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}\right). \qquad [23]$$

Por ejemplo, en el modelo [2]:



Una vez estimado, el modelo [2] podría quedar como en la figura siguiente:



PARTE 3 - EL MODELO RLM CLÁSICO Y EL ESTIMADOR MCO DE β

Para estimar los K parámetros $\beta_1, \beta_2, ..., \beta_K$ de un modelo RLM,

Modelo
$$Y = \beta_1 + \beta_2 X_2 + ... + \beta_K X_K + U, \tag{24}$$

disponemos de una colección de datos (muestra) de N observaciones sobre cada una de las variables $Y, X_2, ..., X_K$:

Datos
$$\begin{bmatrix} y_1 & x_{12} & \cdots & x_{1K} \\ y_2 & x_{22} & \cdots & x_{2K} \\ \vdots & \vdots & & \vdots \\ y_N & x_{N2} & \cdots & x_{NK} \end{bmatrix} = [\mathbf{y}, \mathbf{x}_2, ..., \mathbf{x}_K]. \tag{25}$$

Suponemos que los datos son una realización particular de una colección $\mathbf{W}$ de variables aleatorias,

$$\mathbf{W} = \begin{bmatrix} Y_1 & X_{12} & \cdots & X_{1K} \\ Y_2 & X_{22} & \cdots & X_{2K} \\ \vdots & \vdots & & \vdots \\ Y_N & X_{N2} & \cdots & X_{NK} \end{bmatrix} = [\mathbf{Y}, \mathbf{X}_2, ..., \mathbf{X}_K], \tag{26}$$

que satisfacen las cinco "hipótesis clásicas" siguientes:

[HC1] Linealidad (con respecto a los parámetros $\beta_1, \beta_2, ..., \beta_K$)

En cada observación o punto muestral,

$$Y_i = \beta_1 + \beta_2 X_{i2} + ... + \beta_K X_{iK} + U_i, \; i = 1, 2, ..., N, \text{ o bien} \tag{27}$$

$$\mathbf{Y} = \mathbf{X\beta} + \mathbf{U}, \text{ con} \tag{28}$$

$$\underset{N\times 1}{\mathbf{Y}} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}, \quad \underset{N\times K}{\mathbf{X}} = [\mathbf{i}, \mathbf{X}_2, ..., \mathbf{X}_K] = \begin{bmatrix} 1 & X_{12} & \cdots & X_{1K} \\ 1 & X_{22} & \cdots & X_{2K} \\ \vdots & \vdots & & \vdots \\ 1 & X_{N2} & \cdots & X_{NK} \end{bmatrix}, \quad \underset{K\times 1}{\boldsymbol{\beta}} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix}, \quad \underset{N\times 1}{\mathbf{U}} = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}, \qquad [29]$$

de manera que [27] y [28]-[29] son, en efecto, equivalentes:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} 1 & X_{12} & \cdots & X_{1K} \\ 1 & X_{22} & \cdots & X_{2K} \\ \vdots & \vdots & & \vdots \\ 1 & X_{N2} & \cdots & X_{NK} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}.$$

[HC2] Ausencia de Multicolinealidad Exacta

$$\text{rango}(\mathbf{X}) = K < N. \qquad [30]$$

Las K columnas de $\mathbf{X}$ son linealmente independientes, de manera que los datos sobre cada variable explicativa contienen información específica no redundante en el sentido de que no son una combinación lineal exacta de los datos sobre otra/s variable/s explicativa/s.

[HC3] Exogeneidad Estricta

$$\text{E}[\mathbf{U} \,|\, \mathbf{X}] = \mathbf{0}. \qquad [31]$$

[HC4] Perturbaciones Esféricas: Homoscedasticidad y Ausencia de Autocorrelación

$$\text{E}[\mathbf{U}\mathbf{U}' \,|\, \mathbf{X}] = \sigma^2 \mathbf{I}, \text{ con } \sigma^2 > 0\,. \qquad [32]$$

[HC5] Normalidad

$$\mathbf{U} \,|\, \mathbf{X} \sim \text{Normal}. \qquad [33]$$

En conjunto, [HC1]-[HC5] implican que $\mathbf{Y} \,|\, \mathbf{X} \sim \text{Normal}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ con $\text{rango}(\mathbf{X}) = K < N$, de manera que, en particular,

$$\text{E}[Y_i \,|\, \mathbf{X}] = \beta_1 + \beta_2 X_{i2} + ... + \beta_K X_{iK} \text{ para todo } i = 1, 2, ..., N\,.$$

EL ESTIMADOR MCO DE β

El estimador MCO de $\boldsymbol{\beta}$ es el vector $K \times 1$ de variables aleatorias

$$\hat{\boldsymbol{\beta}}_{\mathbf{W}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\,, \qquad [34]$$

donde $\mathbf{X}$ e $\mathbf{Y}$ son la matriz y el vector de variables aleatorias definidos en [29].

Concebir un **estimador** como un vector de variables aleatorias permite, a través del estudio de sus **propiedades estadísticas**, obtener información sobre cualquier **estimación** (realización particular del estimador) en cualquier análisis aplicado en el que se puedan asumir razonablemente las **hipótesis** que garantizan las propiedades del estimador. En cierto sentido, el estimador MCO de $\boldsymbol{\beta}$ representa cualquier estimación de $\boldsymbol{\beta}$ que se pueda obtener en la práctica en cualquier modelo RLM con cualquier colección de datos.

PARTE 4 - PROPIEDADES ESTADÍSTICAS DEL ESTIMADOR MCO DE β

Propiedades Estadísticas del Estimador MCO $\hat{\boldsymbol\beta}_{\mathbf{W}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ en el Modelo RLM Clásico

[HC1] $\mathbf{Y} = \mathbf{X}\boldsymbol\beta + \mathbf{U}$

Linealidad con respecto a $\beta_1, \beta_2, \ldots, \beta_K$

[HC2] $\text{rango}(\mathbf{X}) = K < N$

Ausencia de Multicolinealidad Exacta

[HC3] $\text{E}[\mathbf{U} \,|\, \mathbf{X}] = \mathbf{0}$

Exogeneidad Estricta

[HC1]–[HC3] $\Rightarrow \text{E}[\hat{\boldsymbol\beta}_{\mathbf{W}}] = \boldsymbol\beta$

[HC4] $\text{E}[\mathbf{U}\mathbf{U}' \,|\, \mathbf{X}] = \sigma^2\mathbf{I}$, con $\sigma^2 > 0$

Perturbaciones Esféricas:
Homoscedasticidad y Ausencia de Autocorrelación

[HC1]–[HC4] $\Rightarrow$ $\text{Var}[\hat{\boldsymbol\beta}_{\mathbf{W}} \,|\, \mathbf{X}] = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$; Teorema de Gauss-Markov; ELIO (BLUE)*

[HC5] $\mathbf{U} \,|\, \mathbf{X} \sim \text{Normal}$

Normalidad

[HC1]–[HC5] $\Rightarrow \hat{\boldsymbol\beta}_{\mathbf{W}}$ eficiente

* Estimador Lineal Insesgado Óptimo (Best Linear Unbiased Estimator)

$$\hat{\boldsymbol\beta}_{\mathbf{W}} = \underbrace{(\mathbf{X}'\mathbf{X})^{-1}}_{\substack{\uparrow \\ \text{Existe} \\ \text{[HC2]}}} \mathbf{X}'\mathbf{Y} \underset{\substack{\uparrow \\ \text{[HC1]}}}{=} (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol\beta + \mathbf{U})$$

$$= \underbrace{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}}_{\substack{\uparrow \\ \mathbf{I}}}\boldsymbol\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U} = \boldsymbol\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U} \quad [\Rightarrow \hat{\boldsymbol\beta}_{\mathbf{W}} - \boldsymbol\beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}].$$

$$\text{E}[\hat{\boldsymbol\beta}_{\mathbf{W}} \,|\, \mathbf{X}] = \text{E}[\boldsymbol\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U} \,|\, \mathbf{X}] = \text{E}[\boldsymbol\beta \,|\, \mathbf{X}] + \text{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U} \,|\, \mathbf{X}]$$

$$= \boldsymbol\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\underbrace{\text{E}[\mathbf{U} \,|\, \mathbf{X}]}_{\substack{\uparrow \\ \mathbf{0} \\ \text{[HC3]}}} = \boldsymbol\beta + (\mathbf{X}'\mathbf{X})^{-1}\underbrace{\mathbf{X}'\mathbf{0}}_{\substack{\uparrow \\ \mathbf{0}}} = \boldsymbol\beta \;\; (= \text{E}[\hat{\boldsymbol\beta}_{\mathbf{W}}] \text{ incondicionalmente}).$$

$$\text{Var}[\hat{\boldsymbol\beta}_{\mathbf{W}} \,|\, \mathbf{X}] = \text{E}[(\hat{\boldsymbol\beta}_{\mathbf{W}} - \underbrace{\text{E}[\hat{\boldsymbol\beta}_{\mathbf{W}} \,|\, \mathbf{X}]}_{\substack{\uparrow \\ \boldsymbol\beta \\ \text{[HC1]-[HC3]}}})(\hat{\boldsymbol\beta}_{\mathbf{W}} - \text{E}[\hat{\boldsymbol\beta}_{\mathbf{W}} \,|\, \mathbf{X}])' \,|\, \mathbf{X}] = \text{E}[\underbrace{(\hat{\boldsymbol\beta}_{\mathbf{W}} - \boldsymbol\beta)}_{\substack{\uparrow \\ (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U} \\ \text{[HC1]-[HC2]}}}(\hat{\boldsymbol\beta}_{\mathbf{W}} - \boldsymbol\beta)' \,|\, \mathbf{X}]$$

$$= \text{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U}\mathbf{U}'\mathbf{X}\underbrace{(\mathbf{X}'\mathbf{X})^{-1\prime}}_{\substack{\uparrow \\ (\mathbf{X}'\mathbf{X})^{-1} \\ \text{simétrica}}} \,|\, \mathbf{X}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\underbrace{\text{E}[\mathbf{U}\mathbf{U}' \,|\, \mathbf{X}]}_{\substack{\uparrow \\ \sigma^2\mathbf{I} \\ \text{[HC4]}}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\sigma^2\mathbf{I})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2\underbrace{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}}_{\substack{\uparrow \\ \mathbf{I}}}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}.$$

TEOREMA DE GAUSS-MARKOV: Bajo [HC1]-[HC4], si $\bar{\boldsymbol{\beta}}_{\mathbf{W}}$ es cualquier estimador de $\boldsymbol{\beta}$ lineal en $\mathbf{Y}$ con $\mathrm{E}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] = \boldsymbol{\beta}$ (insesgado), entonces $\mathrm{Var}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] - \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}]$ es una matriz semidefinida positiva, por lo que el estimador MCO $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$ es el ELIO (BLUE) de $\boldsymbol{\beta}$.

Demostración: Cualquier estimador de $\boldsymbol{\beta}$ que sea lineal en $\mathbf{Y}$ se puede expresar como

$$\bar{\boldsymbol{\beta}}_{\mathbf{W}} = \mathbf{CY}, \qquad [35]$$

donde $\mathbf{C}$ es una matriz $K \times N$ cuyo contenido puede depender de $\mathbf{X}$ pero no de $\mathbf{Y}$. Por otro lado, [HC1] implica que [35] se puede escribir como

$$\bar{\boldsymbol{\beta}}_{\mathbf{W}} = \mathbf{CX}\boldsymbol{\beta} + \mathbf{CU}. \qquad [36]$$

Por lo tanto, teniendo en cuenta [HC3],

$$\mathrm{E}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] = \mathbf{CX}\boldsymbol{\beta} + \mathbf{C}\,\mathrm{E}[\mathbf{U} \mid \mathbf{X}] = \mathbf{CX}\boldsymbol{\beta}, \qquad [37]$$

de manera que $\mathrm{E}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] = \boldsymbol{\beta}$ sí y sólo sí

$$\mathbf{CX} = \mathbf{I}. \qquad [38]$$

Las expresiones [35] y [38] caracterizan a cualquier estimador lineal e insesgado de $\boldsymbol{\beta}$ [en particular, al estimador MCO cuando $\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$]. Teniendo en cuenta [HC4], la matriz de varianzas-covarianzas de un estimador de este tipo es la siguiente:

$$\begin{aligned}\mathrm{Var}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] &= \mathrm{E}[(\bar{\boldsymbol{\beta}}_{\mathbf{W}} - \underbrace{\mathrm{E}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}]}_{\substack{\uparrow \\ \boldsymbol{\beta} \\ [37]\ [38]}})(\bar{\boldsymbol{\beta}}_{\mathbf{W}} - \mathrm{E}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}])' \mid \mathbf{X}] = \mathrm{E}[\underbrace{(\bar{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})}_{\substack{\uparrow \\ \mathbf{CU} \\ [36]\ [38]}}(\bar{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})' \mid \mathbf{X}] \\ &= \mathrm{E}[\mathbf{CUU}'\mathbf{C}' \mid \mathbf{X}] = \mathbf{C}\,\mathrm{E}[\mathbf{UU}' \mid \mathbf{X}]\,\mathbf{C}' = \mathbf{C}(\sigma^2\mathbf{I})\mathbf{C}' = \sigma^2\mathbf{CC}'.\end{aligned} \qquad [39]$$

En consecuencia,

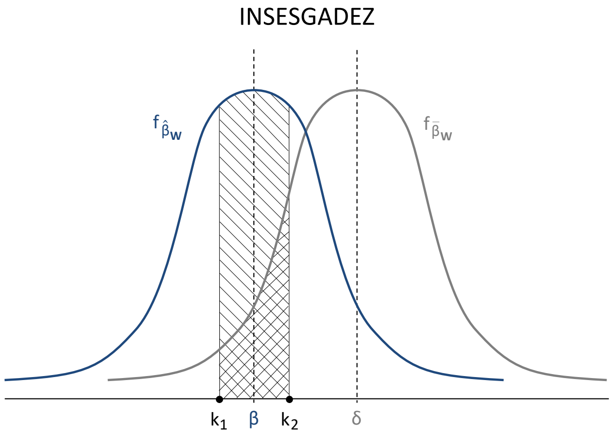
$$\begin{aligned}\mathrm{Var}[\bar{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] - \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] &= \sigma^2[\mathbf{CC}' - (\mathbf{XX})^{-1}] = \sigma^2[\mathbf{CC}' - \underbrace{\mathbf{CX}}_{\substack{\mathbf{I} \\ [38]}}(\mathbf{XX})^{-1}\underbrace{\mathbf{X}'\mathbf{C}'}_{\substack{\mathbf{I} \\ [38]}}] \\ &= \sigma^2\mathbf{C}\left[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\mathbf{C}' = \sigma^2\mathbf{CMC}',\end{aligned} \qquad [40]$$

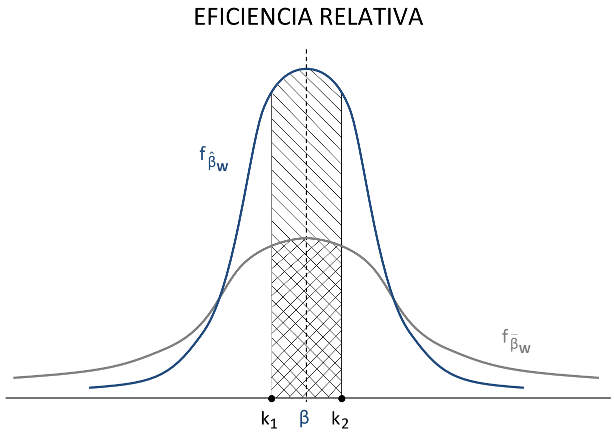
donde $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ es una matriz $N \times N$ simétrica ($\mathbf{M}' = \mathbf{M}$) e idempotente ($\mathbf{MM} = \mathbf{M}$). Por lo tanto, para cualquier vector $\mathbf{d}$ de orden $N \times 1$,

$$\mathbf{d}'(\mathbf{CMC}')\mathbf{d} = \mathbf{d}'(\mathbf{CMM}'\mathbf{C}')\mathbf{d} = (\mathbf{M}'\mathbf{C}'\mathbf{d})'(\mathbf{M}'\mathbf{C}'\mathbf{d}) \geq 0$$

(una suma de cuadrados), lo que prueba que [40] es una matriz semidefinida positiva. ■

El Teorema de Gauss-Markov implica, en particular, que todos los elementos en la diagonal principal de [40] son no negativos, de manera que $\mathrm{Var}[\bar{\beta}_{j\mathbf{W}} \mid \mathbf{X}] \geq \mathrm{Var}[\hat{\beta}_{j\mathbf{W}} \mid \mathbf{X}]$, por lo que $\hat{\beta}_{j\mathbf{W}}$ (MCO) es el estimador con varianza mínima entre todos los estimadores lineales e insesgados de β_j ($j = 1, 2, ..., K$) en cualquier modelo que satisfaga [HC1]-[HC4].





A igualdad de otras propiedades, $\text{Prob}[k_1 \le \hat{\beta}_{\mathbf{W}} \le k_2] > \text{Prob}[k_1 \le \overline{\beta}_{\mathbf{W}} \le k_2]$, de manera que la probabilidad de que un estimador insesgado, o relativamente eficiente, proporcione estimaciones próximas a lo que queremos estimar, es mayor que en el caso de un estimador sesgado, o relativamente ineficiente. En otros términos, el Teorema de Gauss-Markov nos da en la práctica cierta confianza en que las estimaciones MCO $\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_K$ estén relativamente próximas a lo que realmente valen los parámetros $\beta_1, \beta_2, ..., \beta_K$.

PARTE 5 - ESTIMACIÓN DE LA MATRIZ DE VARIANZAS-COVARIANZAS DE $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$

EMV.1: Bajo [HC1]-[HC4], si

$$\hat{\mathbf{U}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{W}}, \quad [41]$$

entonces

$$\text{E}[\hat{\mathbf{U}}\,|\,\mathbf{X}] = \mathbf{0},\;\; \text{Var}[\hat{\mathbf{U}}\,|\,\mathbf{X}] = \sigma^2\mathbf{M},\;\; \mathbf{M} = \mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'. \quad [42]$$

Demostración:

[HC1] [HC2] $$\hat{\mathbf{U}} = \mathbf{Y} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{MY} = \mathbf{M}(\mathbf{X}\boldsymbol{\beta} + \mathbf{U}) = \mathbf{MU}. \quad [43]$$

[HC3] $$\text{E}[\hat{\mathbf{U}}\,|\,\mathbf{X}] = \text{E}[\mathbf{MU}\,|\,\mathbf{X}] = \mathbf{M}\,\text{E}[\mathbf{U}\,|\,\mathbf{X}] = \mathbf{0}. \quad [44]$$

[HC4] $$\text{Var}[\hat{\mathbf{U}}\,|\,\mathbf{X}] = \text{E}[\hat{\mathbf{U}}\hat{\mathbf{U}}'\,|\,\mathbf{X}] = \text{E}[\mathbf{MUU}'\mathbf{M}'\,|\,\mathbf{X}] = \mathbf{M}\,\text{E}[\mathbf{UU}'\,|\,\mathbf{X}]\,\mathbf{M}' = \sigma^2\mathbf{M}, \quad [45]$$

donde (como en la parte final de la demostración del Teorema de Gauss-Markov) se ha tenido en cuenta que $\mathbf{M}$ es simétrica ($\mathbf{M}' = \mathbf{M}$) e idempotente ($\mathbf{MM} = \mathbf{M}$). ■

EMV.2: Bajo [HC1]-[HC4], si

$$\hat{\sigma}^2_{\mathbf{W}} = \frac{\hat{\mathbf{U}}'\hat{\mathbf{U}}}{N-K}, \quad [46]$$

entonces

$$\text{E}[\hat{\sigma}^2_{\mathbf{W}}\,|\,\mathbf{X}] = \sigma^2 \;\; (= \text{E}[\hat{\sigma}^2_{\mathbf{W}}] \text{ incondicionalmente}), \quad [47]$$

por lo que $\hat{\sigma}^2_{\mathbf{W}}$ es un estimador insesgado del parámetro σ^2 (la varianza constante de las perturbaciones en [HC4]).

Demostración: Teniendo en cuenta [45],

$$\begin{aligned}\mathrm{E}[\hat{\mathbf{U}}'\hat{\mathbf{U}}\,|\,\mathbf{X}] &= \mathrm{E}[\mathrm{traza}(\hat{\mathbf{U}}'\hat{\mathbf{U}})\,|\,\mathbf{X}] = \mathrm{E}[\mathrm{traza}(\hat{\mathbf{U}}\hat{\mathbf{U}}')\,|\,\mathbf{X}] \\ &= \mathrm{traza}\left(\mathrm{E}[\hat{\mathbf{U}}\hat{\mathbf{U}}'\,|\,\mathbf{X}]\right) = \sigma^2 \times \mathrm{traza}(\mathbf{M}),\end{aligned} \qquad [48]$$

con

$$\begin{aligned}\mathrm{traza}(\mathbf{M}) &= \mathrm{traza}\left[\,\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\,\right] \\ &= \mathrm{traza}(\mathbf{I}_N) - \mathrm{traza}\left[\,\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\,\right] \\ &= \mathrm{traza}(\mathbf{I}_N) - \mathrm{traza}\left[\,\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\,\right] \\ &= \mathrm{traza}(\mathbf{I}_N) - \mathrm{traza}(\mathbf{I}_K) = N - K.\end{aligned} \qquad [49]$$

De [48]-[49] se deduce directamente [47]. ■

EMV.3: Bajo [HC1]-[HC4], $\hat{\sigma}^2_{\mathbf{W}}(\mathbf{X}'\mathbf{X})^{-1}$ es un estimador insesgado de $\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}\,|\,\mathbf{X}]$.

Demostración: Es inmediata a partir de [47] porque $\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}\,|\,\mathbf{X}] = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. ■

Notación: En adelante no se vuelve a emplear (en general) la notación "$\cdot\,|\,\mathbf{X}$" ("condicional con respecto a $\mathbf{X}$", o "condicionado/a por $\mathbf{X}$"). No obstante, cualquier momento (esperanza, varianza, covarianza) o, en general, cualquier distribución, deberá entenderse siempre en sentido condicional. Si un momento o una distribución no depende de $\mathbf{X}$, entonces también podrá entenderse como un momento o una distribución incondicional o marginal.

Cuando las variables aleatorias en $\mathbf{X}$ e $\mathbf{Y}$ se remplazan por los datos correspondientes a una aplicación concreta (nótese que se emplea el mismo símbolo $\mathbf{X}$ para representar tanto la matriz de variables aleatorias como la matriz de datos referidos a las variables explicativas), se obtienen las cantidades numéricas siguientes: [A] $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ (la estimación MCO de $\boldsymbol{\beta}$), [B] $\hat{\sigma}^2 = (\hat{\mathbf{u}}'\hat{\mathbf{u}})/(N-K)$ (la estimación MCO de σ^2, cuyo numerador es la SCR), y [C] $\widehat{\mathrm{Var}}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$ (la matriz de varianzas-covarianzas estimadas del estimador MCO $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$), cuyos elementos son los siguientes:

$$\widehat{\mathrm{Var}}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \begin{bmatrix} \widehat{\mathrm{Var}}[\hat{\beta}_{1\mathbf{W}}] & \widehat{\mathrm{Cov}}[\hat{\beta}_{1\mathbf{W}}, \hat{\beta}_{2\mathbf{W}}] & \cdots & \widehat{\mathrm{Cov}}[\hat{\beta}_{1\mathbf{W}}, \hat{\beta}_{K\mathbf{W}}] \\ \widehat{\mathrm{Cov}}[\hat{\beta}_{2\mathbf{W}}, \hat{\beta}_{1\mathbf{W}}] & \widehat{\mathrm{Var}}[\hat{\beta}_{2\mathbf{W}}] & \cdots & \widehat{\mathrm{Cov}}[\hat{\beta}_{2\mathbf{W}}, \hat{\beta}_{K\mathbf{W}}] \\ \vdots & \vdots & & \vdots \\ \widehat{\mathrm{Cov}}[\hat{\beta}_{K\mathbf{W}}, \hat{\beta}_{1\mathbf{W}}] & \widehat{\mathrm{Cov}}[\hat{\beta}_{K\mathbf{W}}, \hat{\beta}_{2\mathbf{W}}] & \cdots & \widehat{\mathrm{Var}}[\hat{\beta}_{K\mathbf{W}}] \end{bmatrix}. \qquad [50]$$

La raíz cuadrada de $\hat{\sigma}^2 = (\hat{\mathbf{u}}'\hat{\mathbf{u}})/(N-K)$ se denomina el **error estándar de la regresión**. La raíz cuadrada de la varianza estimada $\widehat{\mathrm{Var}}[\hat{\beta}_{j\mathbf{W}}]$ se denomina el **error estándar del estimador** $\hat{\beta}_{j\mathbf{W}}$ $(1 \le j \le K)$ (que es simplemente la desviación típica estimada del estimador MCO de β_j). Otras cantidades que suelen incluirse en el resumen de un modelo estimado son el R^2 (centrado), el $\bar{R}^2$ (ajustado), y el tamaño muestral N.

EJ2 HASTA E | EJ3 PR2

En un modelo RLM del tipo $Y = \beta_1 + \beta_2 X_2 + \ldots + \beta_K X_K + U$ con $K \ge 2$, la matriz de varianzas-covarianzas estimadas del estimador MCO $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$ $[\widehat{\mathrm{Var}}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}]$ se puede expresar y calcular (Apartados 5.1 y 5.5 de la Sección 2.2) como

$$\hat{\text{Var}}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \hat{\sigma}^2 \begin{bmatrix} \frac{1}{N} + \bar{\mathbf{x}}'(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\bar{\mathbf{x}} & -\bar{\mathbf{x}}'(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1} \\ -(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\bar{\mathbf{x}} & (\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1} \end{bmatrix}, \qquad [51]$$

de manera que, en particular,

$$\hat{\text{Var}}[\hat{\boldsymbol{\beta}}_{\mathrm{b}\mathbf{W}}] = \hat{\sigma}^2(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}, \qquad [52]$$

con

$$\hat{\boldsymbol{\beta}}_{\mathrm{b}\mathbf{W}} = \begin{bmatrix} \hat{\beta}_{2\mathbf{W}} \\ \hat{\beta}_{3\mathbf{W}} \\ \vdots \\ \hat{\beta}_{K\mathbf{W}} \end{bmatrix}, \; \tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}} = \begin{bmatrix} \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 & \cdots & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_K \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3 & \cdots & \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_K \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{x}}_K'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_K'\tilde{\mathbf{x}}_3 & \cdots & \tilde{\mathbf{x}}_K'\tilde{\mathbf{x}}_K \end{bmatrix}. \qquad [53]$$

EJ3 PR2

EJEMPLO - MODELO RLM Y = $\beta_1 + \beta_2 X_2 + \beta_3 X_3$ + U

$$\hat{\boldsymbol{\beta}}_{\mathrm{b}\mathbf{W}} = \begin{bmatrix} \hat{\beta}_{2\mathbf{W}} \\ \hat{\beta}_{3\mathbf{W}} \end{bmatrix}, \; \tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}} = \begin{bmatrix} \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3 \end{bmatrix}.$$

$$(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1} = \frac{1}{(\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2)(\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3) - (\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3)(\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2)} \begin{bmatrix} \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3 & -\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 \\ -\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 \end{bmatrix}.$$

En este caso, el elemento en la posición (2,2) de [51] (posición (1,1) de [52]) es

$$\begin{aligned} \hat{\text{Var}}[\hat{\beta}_{2\mathbf{W}}] &= \hat{\sigma}^2 \times \frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3}{(\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2)(\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3) - (\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2)^2} = \hat{\sigma}^2 \times \frac{1}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 - \dfrac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3}\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2} \\ &= \frac{\hat{\sigma}^2}{\text{SCT}_2 - \text{SCE}_2} = \frac{\hat{\sigma}^2}{\text{SCR}_2} = \frac{\hat{\sigma}^2}{\text{SCT}_2(1 - R_2^2)}, \end{aligned} \qquad [54]$$

donde

$$\begin{aligned} \text{SCT}_2 &= \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2, \\ \text{SCE}_2 &= \frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3}\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2, \\ \text{SCR}_2 &= \text{SCT}_2 - \text{SCE}_2, \\ R_2^2 &= \frac{\text{SCE}_2}{\text{SCT}_2} = 1 - \frac{\text{SCR}_2}{\text{SCT}_2} \end{aligned}$$

son, respectivamente, la SCT, la SCE, la SCR y el R^2 (centrado) en la regresión con término constante de X_2 sobre X_3. Análogamente, el elemento en la posición (3,3) de [51] (posición (2,2) de [52]) es

$$\hat{\text{Var}}[\hat{\beta}_{3\mathbf{W}}] = \frac{\hat{\sigma}^2}{\text{SCR}_3} = \frac{\hat{\sigma}^2}{\text{SCT}_3(1 - R_3^2)}, \quad [55]$$

donde $\text{SCR}_3, \text{SCT}_3$ y R_3^2 son, respectivamente, la SCR, la SCT y el R^2 (centrado) en la regresión con término constante de X_3 sobre X_2.

En general, las expresiones [54]-[55] indican que en un modelo RLM con término constante del tipo $Y = \beta_1 + \beta_2 X_2 + ... + \beta_K X_K + U$ $(K \geq 3)$,

$$\hat{\text{Var}}[\hat{\beta}_{j\mathbf{W}}] = \frac{\hat{\sigma}^2}{\hat{\mathbf{r}}_j'\hat{\mathbf{r}}_j} = \frac{\hat{\sigma}^2}{\text{SCT}_j(1 - R_j^2)} \quad (2 \leq j \leq K), \quad [56]$$

donde $\hat{\mathbf{r}}_j$, SCT_j y R_j^2 son, respectivamente, los residuos MCO, la suma de cuadrados total $(\tilde{\mathbf{x}}_j'\tilde{\mathbf{x}}_j)$, y el coeficiente de determinación (centrado) en la regresión con término constante de X_j sobre las demás variables explicativas del modelo considerado.

EJ3 PR2

PARTE 6 - IMPLICACIONES DE LA HIPÓTESIS DE NORMALIDAD

IHN.1: Bajo [HC1]-[HC5]:

$$\hat{\boldsymbol{\beta}}_{\mathbf{W}} \sim \text{Normal}[\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}], \quad [57]$$

$$\hat{\sigma}_{\mathbf{W}}^2 \sim \left(\tfrac{\sigma^2}{N-K}\right) \times \chi^2(N-K), \quad [58]$$

los estimadores $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$ y $\hat{\sigma}_{\mathbf{W}}^2$ son independientes. [59]

Demostración (Apéndice A2.4): $\hat{\boldsymbol{\beta}}_{\mathbf{W}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U} \Rightarrow$ [57]. Por otro lado, $\mathbf{V} = \sigma^{-1}\mathbf{U} \sim \text{Normal}(\mathbf{0}, \mathbf{I}) \Rightarrow \sigma^{-2}(N-K)\hat{\sigma}_{\mathbf{W}}^2 = \mathbf{V}'\mathbf{M}\mathbf{V} \sim \chi^2(N-K) \Rightarrow$ [58]. $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{M} = \mathbf{0} \Rightarrow \sigma^{-1}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}$ y $\sigma^{-2}(N-K)\hat{\sigma}_{\mathbf{W}}^2 = \mathbf{V}'\mathbf{M}\mathbf{V}$ son independientes $\Rightarrow$ [59]. ■

IHN.2: Bajo [HC1]-[HC5], el estimador MCO $\hat{\boldsymbol{\beta}}_{\mathbf{W}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ es eficiente (insesgado y con varianza mínima entre TODOS los estimadores insesgados de $\boldsymbol{\beta}$).

Demostración: La demostración se basa en el Teorema de Cramér-Rao, que en el caso del modelo RLM clásico ([HC1]-[HC5]) establece que para cualquier estimador $\overline{\boldsymbol{\beta}}_{\mathbf{W}}$ de $\boldsymbol{\beta}$ que sea insesgado, $\text{Var}[\overline{\boldsymbol{\beta}}_{\mathbf{W}} \mid \mathbf{X}] - \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ es una matriz semidefinida positiva. ■

En relación con la eficiencia de $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$, suele de decirse que $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$ es el MVUE (*Minimum Variance Unbiased Estimator*) o el BUE (*Best Unbiased Estimator*) de $\boldsymbol{\beta}$. El Teorema de Gauss-Markov implica que $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$ es el BLUE (*Best **Linear** Unbiased Estimator*) o el ELIO (Estimador **Lineal** Insesgado Óptimo) de $\boldsymbol{\beta}$ (lo que **no** requiere la hipótesis de Normalidad [HC5]), pero **no** que $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$ sea eficiente: el Teorema de Gauss-Markov **no** excluye la posibilidad de que existan estimadores insesgados **no** lineales con menor varianza que $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$; por su parte, la hipótesis de Normalidad [HC5] (requerida por el Teorema de Cramér-Rao) **sí** excluye esa posibilidad. Para más detalles, ver Hayashi (2000): Sección 1.5.

IHN.3: Bajo [HC1]-[HC5], si $\mathbf{A}$ es una matriz $M \times K$ con rango$(\mathbf{A}) = M \leq K$, entonces

$$F_{\mathbf{W}}^{*} = \frac{[\mathbf{A}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})]'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}[\mathbf{A}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})]}{M\hat{\sigma}_{\mathbf{W}}^{2}} \sim F(M, N-K). \qquad [60]$$

Demostración (Apéndice A2.4): $\mathbf{V} = \sigma^{-1}\mathbf{U} \sim \text{Normal}(\mathbf{0}, \mathbf{I})$ implica que

1: $$\sigma^{-1}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V} \sim \text{N}[\mathbf{0}, (\mathbf{X}'\mathbf{X})^{-1}],$$

2: $$\sigma^{-1}\mathbf{A}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta}) = \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V} \sim \text{N}[\mathbf{0}, \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'],$$

3: $$\begin{aligned} Q_1 &= \sigma^{-2}[\mathbf{A}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})]'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}[\mathbf{A}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})] \\ &= [\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}]'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}] \\ &= \mathbf{V}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V} \sim \chi^2(M), \end{aligned}$$

4: $$Q_2 = \sigma^{-2}(N-K)\hat{\sigma}_{\mathbf{W}}^{2} = \mathbf{V}'\mathbf{M}\mathbf{V} \sim \chi^2(N-K),$$

5: Q_1 y Q_2 son independientes,

6: $$F_{\mathbf{W}}^{*} = \frac{N-K}{M} \times \frac{Q_1}{Q_2} \sim F(M, N-K).$$

IHN.4: La variable aleatoria $F_{\mathbf{W}}^{*}$ en [60] se puede expresar como

$$F_{\mathbf{W}}^{*} = M^{-1}[\mathbf{A}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})]'\widehat{\text{Var}}[\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}}]^{-1}[\mathbf{A}(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta})]. \qquad [61]$$

Demostración (Apéndice A2.3): $\widehat{\text{Var}}[\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \mathbf{A}\widehat{\text{Var}}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}]\mathbf{A}' = \hat{\sigma}_{\mathbf{W}}^{2}\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'$.

6.1 EL ESTADÍSTICO "F"

Bajo [HC1]-[HC5], si $\mathbf{A}$ es una matriz $M \times K$ con rango$(\mathbf{A}) = M \leq K$ y $\mathbf{c}$ es un vector $M \times 1$ tales que $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$, entonces la variable aleatoria $F_{\mathbf{W}}^{*}$ en [60]-[61] queda

$$\begin{aligned} F_{\mathbf{W}} &= \frac{(\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \mathbf{c})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \mathbf{c})}{M\hat{\sigma}_{\mathbf{W}}^{2}} \\ &= M^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \mathbf{c})'\widehat{\text{Var}}[\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}}]^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \mathbf{c}) \sim F(M, N-K), \end{aligned} \qquad [62]$$

que se denomina el **estadístico F** para contrastar $H_0: \mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ frente a $H_1: \mathbf{A}\boldsymbol{\beta} \neq \mathbf{c}$. El **valor calculado** del estadístico F en [62] es

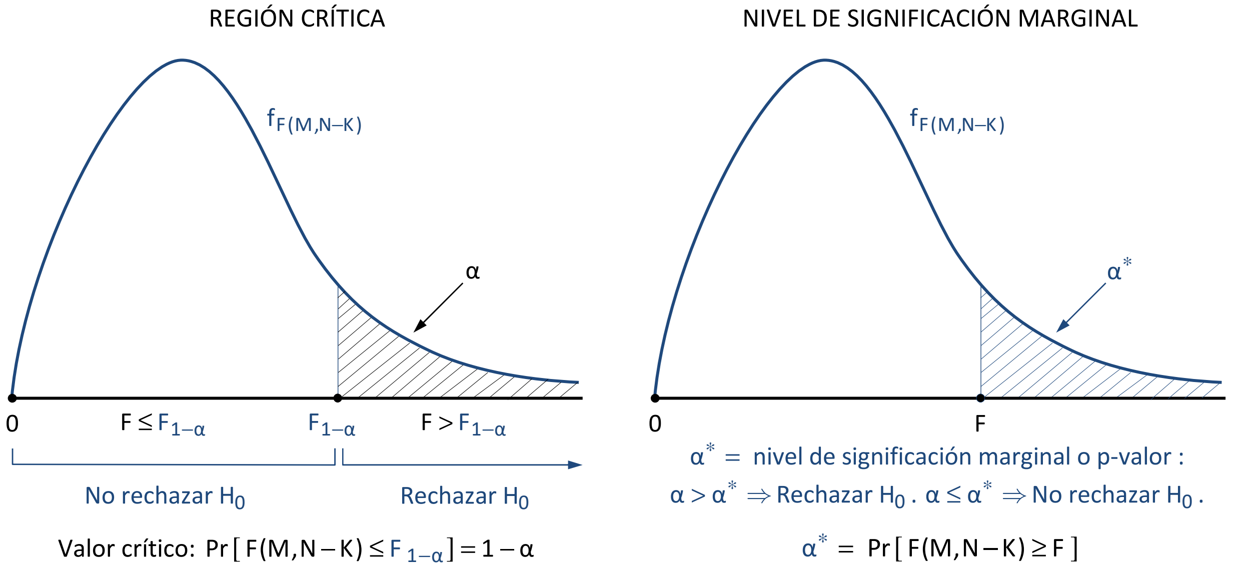
$$F = \frac{(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})}{M\hat{\sigma}^{2}} = \frac{(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})'\widehat{\text{Var}}[\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}}]^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})}{M}, \qquad [63]$$

con $\widehat{\text{Var}}[\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}}] = \hat{\sigma}^{2}\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'$, que también se puede expresar como

$$F = (\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})'\mathbf{Q}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c}), \text{ con } \mathbf{Q} = \frac{[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}}{M\hat{\sigma}^{2}} = \frac{\widehat{\text{Var}}[\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}}]^{-1}}{M}. \qquad [64]$$

Dado que $\mathbf{Q}$ es una matriz simétrica definida positiva, $F > 0$ ($F = 0$ sólo sí $\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c} = \mathbf{0}$)

se puede interpretar como una distancia ponderada entre $\mathbf{A}\hat{\boldsymbol{\beta}}$ (el lado izquierdo estimado de H_0) y $\mathbf{c}$ (el lado derecho de H_0), de manera que H_0 se debe rechazar en favor de H_1 cuando F es "suficientemente" grande (cuando $\mathbf{A}\hat{\boldsymbol{\beta}}$ es "suficientemente" distinto de $\mathbf{c}$), es decir, cuando la información contenida en los datos, resumida en un modelo estimado, sugiere que $\mathbf{A}\boldsymbol{\beta}$ (desconocido) es "suficientemente" distinto del valor propuesto para $\mathbf{A}\boldsymbol{\beta}$ (el vector $\mathbf{c}$) en H_0.



El nivel de significación marginal (p-valor, *p-value*) resume en un único número toda la información que proporcionan los datos, a través de un modelo estimado, sobre H_0 frente a H_1, en el sentido de que α^* puede interpretarse (en una escala 0-1) como una medida de la evidencia que contienen los datos en favor de H_0 con respecto a H_1.

EJ3 PR3

6.2 CÁLCULO DEL ESTADÍSTICO "F" MEDIANTE SUMAS DE CUADRADOS DE RESIDUOS

La estimación de Mínimos Cuadrados Restringidos (MCR) de $\boldsymbol{\beta}$ cuando en un modelo RLM se impone la restricción de que $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ [con $\mathbf{A}$ $M \times K$, rango$(\mathbf{A}) = M \le K$, $\mathbf{c}$ $M \times 1$] es el valor de $\mathbf{b}$ que resuelve el problema

$$\begin{aligned} &\text{minimizar } S(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) \\ &\text{sujeto a } \mathbf{A}\mathbf{b} = \mathbf{c}. \end{aligned} \tag{65}$$

FSC.1: La estimación MCR de $\boldsymbol{\beta}$ es

$$\hat{\boldsymbol{\beta}}^* = \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c}), \tag{66}$$

donde $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ es la estimación MCO (no restringida) de $\boldsymbol{\beta}$.

Demostración: La función lagrangiana asociada con el problema [65] se puede escribir como

$$L(\mathbf{b}, \mathbf{d}) = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) + \mathbf{d}'(\mathbf{A}\mathbf{b} - \mathbf{c}),$$

donde $\mathbf{d} = [d_1, ..., d_M]'$ es un vector $(M \times 1)$ de multiplicadores de Lagrange. Por lo tanto, asumiendo que rango$(\mathbf{X}) = K < N$, $\hat{\boldsymbol{\beta}}^*$ es la (única) solución de [65] sí y sólo sí

$$\text{[i] } \nabla_{\mathbf{b}} L(\hat{\boldsymbol{\beta}}^*, \mathbf{d}^*) = 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}^* - 2\mathbf{X}'\mathbf{y} + \mathbf{A}'\mathbf{d}^* = \mathbf{0}, \text{ y [ii] } \nabla_{\mathbf{d}} L(\hat{\boldsymbol{\beta}}^*, \mathbf{d}^*) = \mathbf{A}\hat{\boldsymbol{\beta}}^* - \mathbf{c} = \mathbf{0}.$$

Premultiplicando [i] por $\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}$ y haciendo uso de [ii] ($\Rightarrow \mathbf{A}\hat{\boldsymbol{\beta}}^* = \mathbf{c}$), se obtiene que $2(\mathbf{c} - \mathbf{A}\hat{\boldsymbol{\beta}}) = -\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\mathbf{d}^* \Rightarrow \mathbf{d}^* = 2 \times [\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})$. Sustituyendo esta expresión para $\mathbf{d}^*$ en [i] se obtiene [66]. ■

Observación I: Para repasar algunas cuestiones importantes sobre optimización matemática con restricciones, ver, por ejemplo, el Capítulo 11 en Luenberger, D.G., Ye, Y. (2021), *Linear and Nonlinear Programming (Fifth Edition)*, Springer.

Observación II: Las propiedades estadísticas del estimador MCR asociado con [66] se consideran en el Apéndice 1 de esta Sección 2.3.

FSC.2: El valor del estadístico F en [63] se puede calcular como

$$F = \frac{N-K}{M} \times \frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} = \frac{N-K}{M} \times \frac{\text{SCR}^* - \text{SCR}}{\text{SCR}}, \qquad [67]$$

donde $\hat{\mathbf{u}}_* = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*$ (residuos MCR, o restringidos), $\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ (residuos MCO, o no restringidos), $\text{SCR}^* = \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*$ (suma de cuadrados de residuos restringida), y $\text{SCR} = \hat{\mathbf{u}}'\hat{\mathbf{u}}$ (suma de cuadrados de residuos no restringida, o sin restringir).

Demostración: De acuerdo con [66], el numerador de la primera expresión para F en [63] se puede expresar y calcular como

$$\begin{aligned}
(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c}) &= (\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^*)'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^*) \\
&= (\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\hat{\boldsymbol{\beta}}^*)'(\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\hat{\boldsymbol{\beta}}^*) \\
&= [(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*) - (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})]'[(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*) - (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})] \\
&= (\hat{\mathbf{u}}_* - \hat{\mathbf{u}})'(\hat{\mathbf{u}}_* - \hat{\mathbf{u}}) = \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* + \hat{\mathbf{u}}'\hat{\mathbf{u}} - 2\hat{\mathbf{u}}_*'\hat{\mathbf{u}} \\
&= \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* + \hat{\mathbf{u}}'\hat{\mathbf{u}} - 2(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*)'\hat{\mathbf{u}} = \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* + \hat{\mathbf{u}}'\hat{\mathbf{u}} - 2\mathbf{y}'\hat{\mathbf{u}} \\
&= \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* + \hat{\mathbf{u}}'\hat{\mathbf{u}} - 2(\hat{\mathbf{y}} + \hat{\mathbf{u}})'\hat{\mathbf{u}} = \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* + \hat{\mathbf{u}}'\hat{\mathbf{u}} - 2\hat{\mathbf{u}}'\hat{\mathbf{u}} \\
&= \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}},
\end{aligned}$$

lo que, junto con que $\hat{\sigma}^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}}/(N-K)$ en el denominador de la primera expresión para F en [63], prueba lo indicado en [67]. ■

6.3 MCR MEDIANTE SUSTITUCIÓN DE RESTRICCIONES

El vector de residuos MCR en [67] se puede calcular (como alternativa equivalente al cálculo de $\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*$ a través de [66]) imponiendo $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ como restricción en el modelo en el que se desea contrastar H_0 (modelo original, no restringido, o sin restringir) y calculando los residuos MCO en el modelo resultante (modelo restringido bajo H_0). Por ejemplo, para contrastar $H_0\colon \beta_2 = 10\beta_3$ en $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + U$, el modelo restringido bajo H_0 es $Y = \beta_1 + \beta_3(X_3 + 10X_2) + U^*$, cuya estimación por MCO (sin restringir) proporciona directamente los residuos $\hat{u}_i^* = y_i - [\hat{\beta}_1^* + \hat{\beta}_3^*(x_{i3} + 10x_{i2})]$.

FSC.3: Si en relación con la estimación de un modelo RLM clásico $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}$ bajo la restricción de que $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ (problema [65]), consideramos las particiones

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{c} \Leftrightarrow [\ \underset{M \times M}{\underset{\uparrow}{\mathbf{A}_1}} \quad \underset{M \times (K-M)}{\underset{\uparrow}{\mathbf{A}_2}}\ ] \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} \begin{array}{l} \leftarrow M \\ \leftarrow K - M \end{array} = \mathbf{c}, \qquad [68]$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U} = [\ \underset{N \times M}{\underset{\uparrow}{\mathbf{X}_1}} \quad \underset{N \times (K-M)}{\underset{\uparrow}{\mathbf{X}_2}}\ ] \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} \begin{array}{l} \leftarrow M \\ \leftarrow K - M \end{array} + \mathbf{U}, \qquad [69]$$

donde $\mathbf{A}_1$ está formada por cualesquiera M columnas de $\mathbf{A}$ linealmente independientes entre sí ($\Rightarrow$ $\mathbf{A}_1^{-1}$ existe), $\mathbf{A}_2$ está formada por las $K - M$ columnas restantes de $\mathbf{A}$, $\boldsymbol{\beta}_1$ y $\boldsymbol{\beta}_2$ contienen los parámetros asociados con las columnas respectivas en $\mathbf{A}_1$ y $\mathbf{A}_2$, y $\mathbf{X}_1$ y $\mathbf{X}_2$ contienen las columnas de $\mathbf{X}$ asociadas con los parámetros respectivos en $\boldsymbol{\beta}_1$ y $\boldsymbol{\beta}_2$, entonces el vector de residuos MCR, $\hat{\mathbf{u}}_* = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*$, se puede calcular como

$$\hat{\mathbf{u}}_* = \mathbf{y}_* - \mathbf{X}_*\hat{\boldsymbol{\beta}}_2^*, \text{ con} \qquad [70]$$

$$\begin{gathered} \mathbf{y}_* = \mathbf{y} - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{c},\ \mathbf{X}_* = \mathbf{X}_2 - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{A}_2, \\ \hat{\boldsymbol{\beta}}_2^* = (\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\mathbf{y}_*. \end{gathered} \qquad [71]$$

Demostración: El punto central consiste en observar que la solución del problema [65] se puede calcular, como alternativa equivalente a lo expuesto en FSC.1, minimizando la suma de cuadrados restringida $S^*(\mathbf{b})$ que resulta de sustituir o incluir $\mathbf{Ab} = \mathbf{c}$ explícitamente en $S(\mathbf{b})$. Con respecto a dicha sustitución, de acuerdo con [68],

$$\mathbf{Ab} = \mathbf{c} \Leftrightarrow \mathbf{A}_1\mathbf{b}_1 + \mathbf{A}_2\mathbf{b}_2 = \mathbf{c} \Rightarrow \mathbf{b}_1 = \mathbf{A}_1^{-1}(\mathbf{c} - \mathbf{A}_2\mathbf{b}_2), \qquad [72]$$

de manera que, teniendo en cuenta [69], si $\mathbf{e}(\mathbf{b}) = \mathbf{y} - \mathbf{Xb} = \mathbf{y} - \mathbf{X}_1\mathbf{b}_1 - \mathbf{X}_2\mathbf{b}_2$ es el vector de residuos asociado con cualquier estimación posible $\mathbf{b}$ de $\boldsymbol{\beta}$, entonces

$$\mathbf{e}^*(\mathbf{b}) = \mathbf{y} - \mathbf{X}_1\mathbf{A}_1^{-1}(\mathbf{c} - \mathbf{A}_2\mathbf{b}_2) - \mathbf{X}_2\mathbf{b}_2 = \mathbf{e}^*(\mathbf{b}_2) \qquad [73]$$

es el vector de residuos asociado con cualquier estimación restringida $\mathbf{b}$ de $\boldsymbol{\beta}$ que satisface [72]. Por lo tanto, el problema [65] es equivalente a

$$\text{minimizar } S^*(\mathbf{b}) = \mathbf{e}^*(\mathbf{b})'\mathbf{e}^*(\mathbf{b}) = \mathbf{e}^*(\mathbf{b}_2)'\mathbf{e}^*(\mathbf{b}_2) = S^*(\mathbf{b}_2) \qquad [74]$$

con respecto a $\mathbf{b}_2$ sin restricciones de ningún tipo. Teniendo en cuenta que en [73]

$$\begin{aligned} \mathbf{e}^*(\mathbf{b}_2) &= \mathbf{y} - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{c} + \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{A}_2\mathbf{b}_2 - \mathbf{X}_2\mathbf{b}_2 \\ &= (\mathbf{y} - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{c}) - (\mathbf{X}_2 - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{A}_2)\mathbf{b}_2 = \mathbf{y}_* - \mathbf{X}_*\mathbf{b}_2, \end{aligned} \qquad [75]$$

con $\mathbf{y}_*$, $\mathbf{X}_*$ dados en la primera línea de [71], el problema [74] puede plantearse como

$$\text{minimizar } S^*(\mathbf{b}_2) = (\mathbf{y}_* - \mathbf{X}_*\mathbf{b}_2)'(\mathbf{y}_* - \mathbf{X}_*\mathbf{b}_2). \qquad [76]$$

La solución de [76] es la estimación MCO del vector $\boldsymbol{\beta}_2$ en la regresión lineal del vector $\mathbf{y}_*$ sobre la matriz $\mathbf{X}_*$, como se indica en la segunda línea de [71]. Adicionalmente, [72] implica que

$$\hat{\boldsymbol{\beta}}_1^* = \mathbf{A}_1^{-1}(\mathbf{c} - \mathbf{A}_2\hat{\boldsymbol{\beta}}_2^*). \qquad [77]$$

Por lo tanto,

$$\begin{aligned}\hat{\mathbf{u}}_* &= \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^* = \mathbf{y} - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1^* - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2^* = \mathbf{y} - \mathbf{X}_1[\mathbf{A}_1^{-1}(\mathbf{c} - \mathbf{A}_2\hat{\boldsymbol{\beta}}_2^*)] - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2^* \\ &= (\mathbf{y} - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{c}) - (\mathbf{X}_2 - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{A}_2)\hat{\boldsymbol{\beta}}_2^* = \mathbf{y}_* - \mathbf{X}_*\hat{\boldsymbol{\beta}}_2^*,\end{aligned}$$

como se indica en [70]. ■

Ejemplo 1: EJ3 PR4

$$\beta_2 - 10\beta_3 = 0 \ \ (M = 1), \text{ en}$$
$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + U \ \ (K = 3).$$

Observación: Cuando $M = 1$, o cuando $M \geq 2$ con restricciones que implican cada una de ellas a un único parámetro (como $\beta_2 = 1, \beta_3 = 0$), la manera más sencilla de calcular los residuos MCR y la suma de sus cuadrados (SCR* en [67]), consiste en introducir directamente las restricciones consideradas dentro el modelo original y estimar por MCO el modelo restringido resultante [lo que en este ejemplo, sustituyendo β_2 por $10\beta_3$ en el modelo original, daría lugar al modelo restringido $Y = \beta_1 + \beta_3(X_3 + 10X_2) + U^*$, es decir, la RLS de Y sobre $(X_3 + 10X_2)$]. No obstante, en este ejemplo y, especialmente, en el siguiente, se procede como se describe en FSC.3 para ilustrar las operaciones generales implicadas en el cálculo de $\hat{\boldsymbol{\beta}}^*$ y de $\hat{\mathbf{u}}_* = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*$ mediante sustitución de restricciones sin necesidad de recurrir a la fórmula [66].

En este ejemplo: $\mathbf{X} = [\mathbf{i}, \mathbf{x}_2, \mathbf{x}_3]$, $\boldsymbol{\beta} = \begin{bmatrix}\beta_1 \\ \beta_2 \\ \beta_3\end{bmatrix}$,

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{c} \Rightarrow [0, 1, -10]\begin{bmatrix}\beta_1 \\ \beta_2 \\ \beta_3\end{bmatrix} = 0.$$

Escoger un elemento de $\mathbf{A}$ para formar $\mathbf{A}_1$ cuadrada de orden $M = 1$ (un escalar) no singular (distinto de 0); los demás elementos (columnas) de $\mathbf{A}$ forman $\mathbf{A}_2$. Por ejemplo (en general, suele haber más de una opción),

$$\mathbf{A}_1 = a_2 = 1, \mathbf{A}_2 = [a_1, a_3] = [0, -10], \mathbf{c} = 0.$$

Esta partición de $\mathbf{A}$ determina de forma automática las siguientes:

$$\boldsymbol{\beta}_1 = \beta_2 \;:\; \text{parámetros asociados con las columnas de } \mathbf{A}_1,$$

$$\boldsymbol{\beta}_2 = \begin{bmatrix} \beta_1 \\ \beta_3 \end{bmatrix} \;:\; \text{parámetros asociados con las columnas de } \mathbf{A}_2,$$

$$\mathbf{X}_1 = \mathbf{x}_2 \;:\; \text{columnas de } \mathbf{X} \text{ asociadas con } \boldsymbol{\beta}_1,$$

$$\mathbf{X}_2 = [\mathbf{i}, \mathbf{x}_3] \;:\; \text{columnas de } \mathbf{X} \text{ asociadas con } \boldsymbol{\beta}_2.$$

$$\mathbf{X}_* = \mathbf{X}_2 - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{A}_2 = [\mathbf{i}, \mathbf{x}_3] - \mathbf{x}_2[0, -10] = [\mathbf{i}, \mathbf{x}_3 + 10\mathbf{x}_2],$$

$$\mathbf{y}_* = \mathbf{y} - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{c} = \mathbf{y}.$$

$$\hat{\boldsymbol{\beta}}_2^* = \begin{bmatrix} \hat{\beta}_1^* \\ \hat{\beta}_3^* \end{bmatrix} = (\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\mathbf{y}_*,$$

$$\hat{\boldsymbol{\beta}}_1^* = \hat{\beta}_2^* = \mathbf{A}_1^{-1}(\mathbf{c} - \mathbf{A}_2\hat{\boldsymbol{\beta}}_2^*) = -[0, -10]\begin{bmatrix} \hat{\beta}_1^* \\ \hat{\beta}_3^* \end{bmatrix} = 10\hat{\beta}_3^*.$$

$$\hat{\mathbf{u}}_* = \mathbf{y}_* - \mathbf{X}_*\hat{\boldsymbol{\beta}}_2^* \Rightarrow \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* = \text{SCR}^*.$$

La regresión lineal del vector $\mathbf{y}_*$ sobre la matriz $\mathbf{X}_*$ es la RLS de Y sobre $(X_3 + 10X_2)$, por lo que $\hat{\beta}_1^*$, $\hat{\beta}_3^*$, $\hat{\mathbf{u}}_*$, y SCR^*, son las estimaciones MCO del término constante y la pendiente, los residuos, y la suma de cuadrados de los residuos, en dicha RLS.

Ejemplo 2:

$$\beta_2 - 10\beta_3 = 0,\;\; 2\beta_1 + 2\beta_2 - 7\beta_3 = 50 \;\; (M = 2), \;\text{ en}$$
$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + U \;\; (K = 3).$$

$$\mathbf{X} = [\mathbf{i}, \mathbf{x}_2, \mathbf{x}_3],\; \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}.$$

$$\mathbf{A}\boldsymbol{\beta} = \mathbf{c} \Rightarrow \begin{bmatrix} 0 & 1 & -10 \\ 2 & 2 & -7 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 50 \end{bmatrix}.$$

Escoger $M = 2$ columnas de $\mathbf{A}$ para formar $\mathbf{A}_1$ cuadrada de orden $M = 2$ no singular; las demás columnas de $\mathbf{A}$ forman $\mathbf{A}_2$. Por ejemplo (hay otras opciones),

$$\mathbf{A}_1 = [\mathbf{a}_1, \mathbf{a}_2] = \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}, \mathbf{A}_2 = \mathbf{a}_3 = \begin{bmatrix} -10 \\ -7 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 0 \\ 50 \end{bmatrix}.$$

Esta partición de $\mathbf{A}$ determina de forma automática las siguientes:

$$\boldsymbol{\beta}_1 = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} : \text{parámetros asociados con las columnas de } \mathbf{A}_1,$$

$$\boldsymbol{\beta}_2 = \beta_3 : \text{parámetros asociados con las columnas de } \mathbf{A}_2,$$

$$\mathbf{X}_1 = [\mathbf{i}, \mathbf{x}_2] : \text{columnas de } \mathbf{X} \text{ asociadas con } \boldsymbol{\beta}_1,$$

$$\mathbf{X}_2 = \mathbf{x}_3 : \text{columnas de } \mathbf{X} \text{ asociadas con } \boldsymbol{\beta}_2.$$

$$\mathbf{X}_* = \mathbf{X}_2 - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{A}_2 = \mathbf{x}_3 - [\mathbf{i}, \mathbf{x}_2]\begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} -10 \\ -7 \end{bmatrix},$$

$$\mathbf{y}_* = \mathbf{y} - \mathbf{X}_1\mathbf{A}_1^{-1}\mathbf{c} = \mathbf{y} - [\mathbf{i}, \mathbf{x}_2]\begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 50 \end{bmatrix}.$$

$$\hat{\boldsymbol{\beta}}_2^* = \hat{\beta}_3^* = (\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\mathbf{y}_*,$$

$$\hat{\boldsymbol{\beta}}_1^* = \begin{bmatrix} \hat{\beta}_1^* \\ \hat{\beta}_2^* \end{bmatrix} = \mathbf{A}_1^{-1}(\mathbf{c} - \mathbf{A}_2\hat{\boldsymbol{\beta}}_2^*) = \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}^{-1} \left(\begin{bmatrix} 0 \\ 50 \end{bmatrix} - \begin{bmatrix} -10 \\ -7 \end{bmatrix} \hat{\beta}_3^* \right).$$

$$\mathbf{A}_1^{-1} = \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}^{-1} = -\tfrac{1}{2}\begin{bmatrix} 2 & -1 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0.5 \\ 1 & 0 \end{bmatrix}.$$

$$\hat{\mathbf{u}}_* = \mathbf{y}_* - \mathbf{X}_*\hat{\boldsymbol{\beta}}_2^* \Rightarrow \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* = \text{SCR}^*.$$

6.4 CÁLCULO DEL ESTADÍSTICO "F" MEDIANTE COEFICIENTES DE DETERMINACIÓN

FCD.1: Si las matrices $\mathbf{X}$ y $\mathbf{X}_*$ contienen ambas una columna de unos (tanto el modelo original como el restringido tienen término constante), y además $\mathbf{y}_* = \mathbf{y}$ (la variable dependiente es la misma en ambos modelos), entonces

$$F = \frac{N-K}{M} \times \frac{R^2 - R_*^2}{1 - R^2}, \qquad [78]$$

donde R^2 es el coeficiente de determinación (centrado) en el modelo original, y R_*^2 es el coeficiente de determinación (centrado) en el modelo restringido.

Demostración: La presencia de un término constante en ambos modelos implica que

$$R^2 = 1 - \frac{\text{SCR}}{\text{SCT}} = 1 - \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{\mathbf{y}'\mathbf{y} - N\bar{y}^2}, \quad R_*^2 = 1 - \frac{\text{SCR}^*}{\text{SCT}^*} = 1 - \frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*}{\mathbf{y}_*'\mathbf{y}_* - N\bar{y}_*^2},$$

o bien,

$$\text{SCR} = \text{SCT} \times (1 - R^2), \ \ \text{SCR}^* = \text{SCT}^* \times (1 - R_*^2),$$

de manera que el valor del estadístico F en [67] se puede calcular como

$$F = \frac{N-K}{M} \times \frac{\text{SCT}^* \times (1 - R_*^2) - \text{SCT} \times (1 - R^2)}{\text{SCT} \times (1 - R^2)},$$

de donde se obtiene [78] cuando $\mathbf{y}_* = \mathbf{y}$ $(\Rightarrow \text{SCT}^* = \text{SCT})$.

EJ3 PR4

FCD.2: En un modelo del tipo $Y = \beta_1 + \beta_2 X_2 + ... + \beta_K X_K + U$ $(K \geq 2)$, el estadístico F para contrastar la hipótesis de que $\beta_2 = 0, ..., \beta_K = 0$ (ausencia de significación conjunta de todas las pendientes $\beta_2, ..., \beta_K$) se puede calcular como

$$F = \frac{N-K}{K-1} \times \frac{R^2}{1-R^2}. \quad [79]$$

Demostración: El número de enunciados en la hipótesis nula es $M = K - 1$, y el modelo restringido es simplemente $Y = \beta_1 + U^*$, por lo que $R_*^2 = 0$. En consecuencia, [78] queda en este caso como se indica en [79].

EJ3 PR5

6.5 EL ESTADÍSTICO "t" - CONSIDERACIONES GENERALES

En relación con el contraste de una hipótesis nula general del tipo $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ (que consta, en general, de $M \geq 1$ filas o enunciados), consideramos el caso particular (muy frecuente en la práctica) de que dicha hipótesis conste de un único enunciado ($M = 1$) del tipo $a_1\beta_1 + a_2\beta_2 + ... + a_K\beta_K = c$, de manera que, en este caso, $\mathbf{A} = [a_1, a_2, ..., a_K] = \mathbf{a}'$ (un vector fila $1 \times K$), y $\mathbf{c} = c$ (un escalar 1×1). El estadístico F [62] referido a la hipótesis nula $\mathbf{a}'\boldsymbol{\beta} = c$ queda

$$\begin{aligned} F_{\mathbf{W}} &= \frac{(\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}} - c)'[\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}]^{-1}(\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}} - c)}{\hat{\sigma}^2_{\mathbf{W}}} \\ &= \frac{(\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}} - c)^2}{\hat{\sigma}^2_{\mathbf{W}}[\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}]} = \frac{(\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}} - c)^2}{\widehat{\text{Var}}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}]} \sim F(1, N-K), \end{aligned} \quad [80]$$

lo que, tomando raíces cuadradas en los dos últimos términos de [80], equivale a que

(Apéndice A2.4)
$$t_{\mathbf{W}} = \frac{\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}} - c}{\widehat{\text{Dvt}}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}]} \sim t(N-K), \quad [81]$$

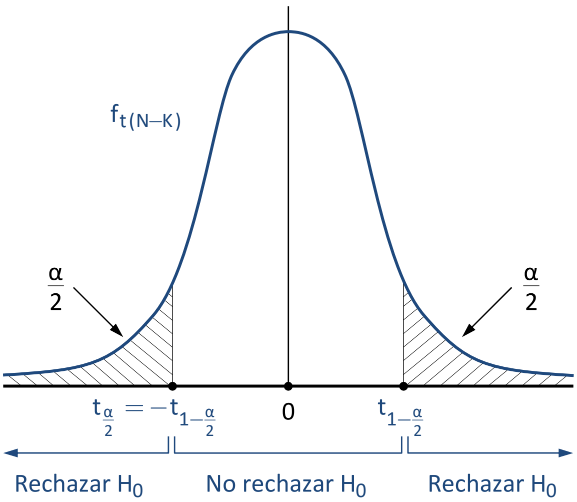
que se denomina el **estadístico t** referido a una hipótesis nula del tipo $H_0: \mathbf{a}'\boldsymbol{\beta} = c$. El **valor calculado** del estadístico t es

$$t = \frac{\mathbf{a}'\hat{\boldsymbol{\beta}} - c}{\widehat{\text{Dvt}}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}]} = \frac{\text{Lado izquierdo estimado de H}_0 \;-\; \text{Lado derecho de H}_0}{\text{Error estándar del estimador del lado izquierdo de H}_0}, \quad [82]$$

que se puede interpretar como una diferencia ponderada entre $\mathbf{a}'\hat{\boldsymbol{\beta}}$ (el lado izquierdo estimado de H_0) y c (el lado derecho de H_0). Por lo tanto:

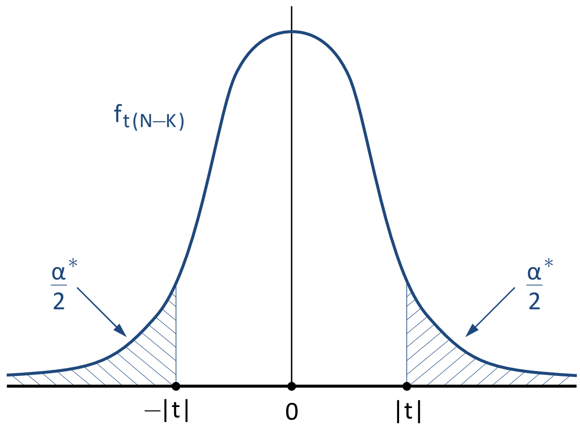
1. $H_0: \mathbf{a}'\boldsymbol{\beta} = c$ se debe rechazar en favor de $H_1: \mathbf{a}'\boldsymbol{\beta} \neq c$ cuando t es "suficientemente" positivo o negativo (cuando $|t|$ es "suficientemente" grande), es decir, cuando la información contenida en los datos y resumida en un modelo estimado sugiere que la cantidad $\mathbf{a}'\boldsymbol{\beta}$ (desconocida) está "suficientemente" lejos del valor propuesto para $\mathbf{a}'\boldsymbol{\beta}$ (el número c) en la hipótesis nula.

REGIÓN CRÍTICA H_1 $(\neq)$



Valor crítico: $\Pr[t(N-K) \leq t_{1-\frac{\alpha}{2}}] = 1 - \frac{\alpha}{2}$

NIVEL DE SIGNIFICACIÓN MARGINAL H_1 $(\neq)$



α^* = nivel de significación marginal o p-valor :

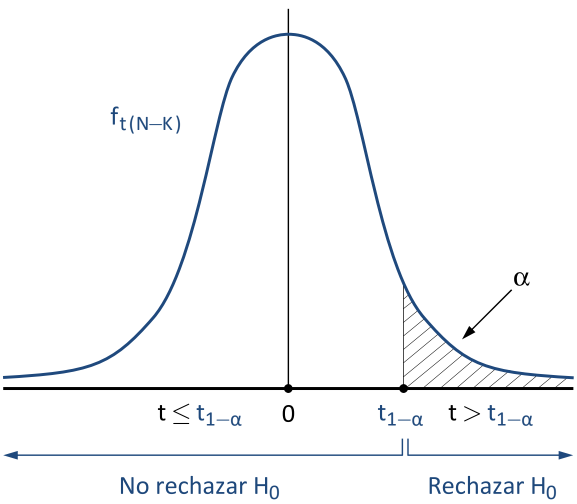
$\alpha > \alpha^* \Rightarrow$ Rechazar H_0 . $\alpha \leq \alpha^* \Rightarrow$ No rechazar H_0 .

$\alpha^* = 2 \times \Pr[t(N-K) \geq |t|]$

EJ3 PR6

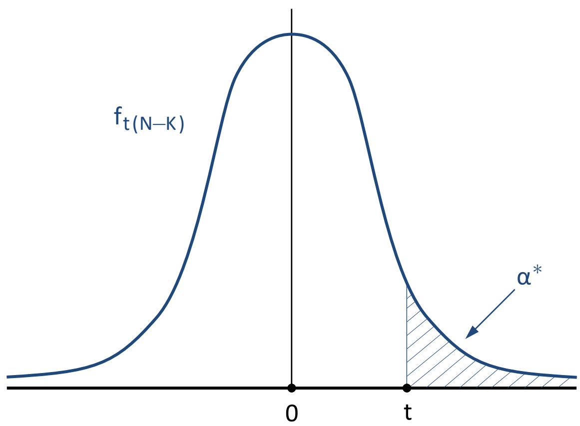
2. $H_0: \mathbf{a}'\boldsymbol{\beta} = c$ se debe rechazar en favor de $H_1: \mathbf{a}'\boldsymbol{\beta} > c$ cuando t es "suficientemente" positivo, es decir, cuando la información contenida en los datos y resumida en un modelo estimado sugiere que la cantidad $\mathbf{a}'\boldsymbol{\beta}$ (desconocida) está "suficientemente" por encima del valor propuesto para $\mathbf{a}'\boldsymbol{\beta}$ (el número c) en la hipótesis nula.

REGIÓN CRÍTICA H_1 $(>)$



Valor crítico: $\Pr[t(N-K) \leq t_{1-\alpha}] = 1 - \alpha$

NIVEL DE SIGNIFICACIÓN MARGINAL H_1 $(>)$

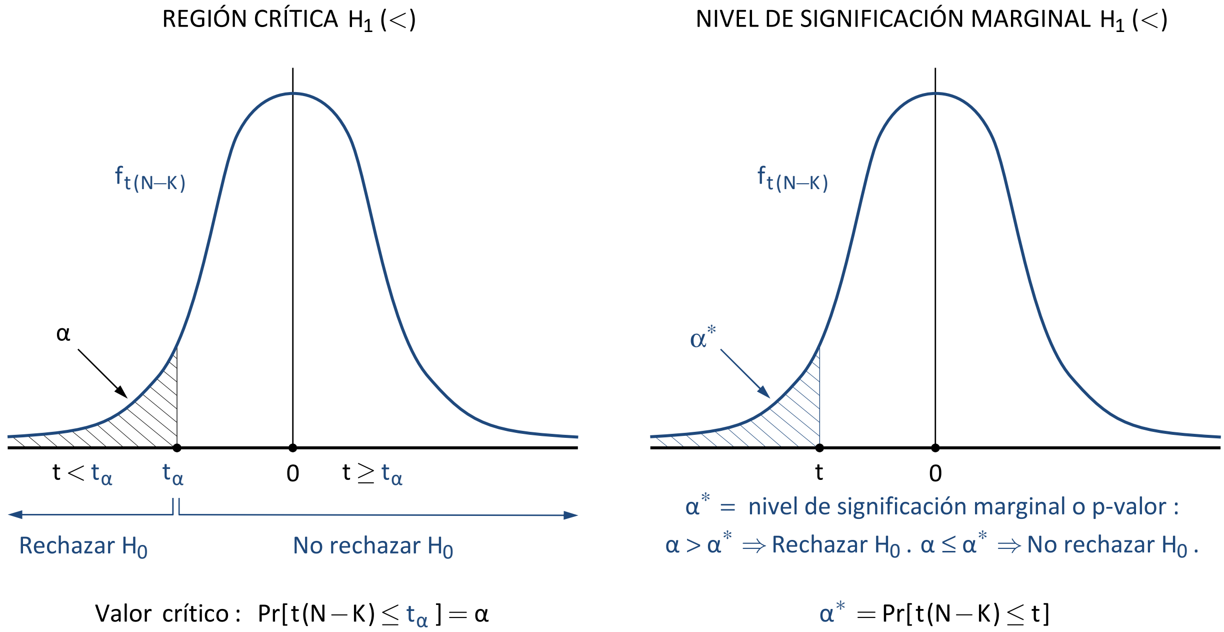


α^* = nivel de significación marginal o p-valor :

$\alpha > \alpha^* \Rightarrow$ Rechazar H_0 . $\alpha \leq \alpha^* \Rightarrow$ No rechazar H_0 .

$\alpha^* = \Pr[t(N-K) \geq t]$

3. $H_0: \mathbf{a}'\boldsymbol{\beta} = c$ se debe rechazar en favor de $H_1: \mathbf{a}'\boldsymbol{\beta} < c$ cuando t es "suficientemente" negativo, es decir, cuando la información contenida en los datos y resumida en un modelo estimado sugiere que la cantidad $\mathbf{a}'\boldsymbol{\beta}$ (desconocida) está "suficientemente" por debajo del valor propuesto para $\mathbf{a}'\boldsymbol{\beta}$ (el número c) en la hipótesis nula.



EJ3 PR7

6.6 EL ESTADÍSTICO "t" - CASOS PARTICULARES

El valor calculado [82] del estadístico [81] referido a una hipótesis nula como $H_0: \beta_j = c$ $(1 \le j \le K)$, se obtiene particularizando [82] cuando $\mathbf{a}' = [0, ..., 1, ..., 0]$ (un vector lleno de ceros excepto por un 1 en la posición j-ésima):

$$\frac{\hat{\beta}_j - c}{\widehat{\mathrm{Dvt}}[\hat{\beta}_{j\mathbf{W}}]} \quad \text{con } \widehat{\mathrm{Dvt}}[\hat{\beta}_{j\mathbf{W}}] = \sqrt{\widehat{\mathrm{Var}}[\hat{\beta}_{j\mathbf{W}}]}. \qquad [83]$$

Cuando $H_0: \beta_j = 0$, [83] queda

$$\frac{\hat{\beta}_j}{\widehat{\mathrm{Dvt}}[\hat{\beta}_{j\mathbf{W}}]} = \frac{\text{Estimación MCO de } \beta_j}{\text{Error estándar del estimador MCO de } \beta_j}, \qquad [84]$$

que se denomina el estadístico t para el **contraste de significación individual** de β_j (o, simplemente, **el estadístico t** de β_j), y se representa habitualmente como t_{β_j} o $t(\beta_j)$.

EJ3 PR8

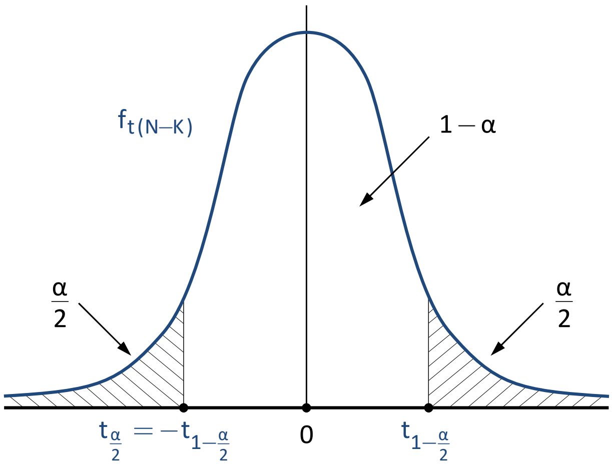
Observación: En modelos con término constante, el denominador de [83]-[84] se puede expresar y calcular (ver [56]) como $\widehat{\mathrm{Dvt}}[\hat{\beta}_{j\mathbf{W}}] = \hat{\sigma} \times [\mathrm{SCT}_j(1 - R_j^2)]^{-½}$. Por lo tanto, cualquier contraste basado en [83]-[84] será más propenso a generar valores de $|t|$ pequeños [y, por lo tanto, a no rechazar H_0 frente a una alternativa bilateral $(\neq)$] cuanto mayor sea el grado de relación muestral entre X_j y las demás variables explicativas (entre otros factores), es decir, cuanta menos información específica contengan los datos sobre X_j con respecto a la información que contienen las demás variables explicativas.

6.7 INTERVALOS DE CONFIANZA

En un modelo RLM clásico del tipo $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}$, un **intervalo de confianza** (IC) del $100(1-\alpha)\%$ para una combinación lineal de $\boldsymbol{\beta}$ del tipo $\mathbf{a}'\boldsymbol{\beta}$ (donde $\mathbf{a}'$ es un vector $1 \times K$ dado), es un subconjunto de $\mathbb{R}$ (el conjunto de todos los números reales) tal que para cualquier número b de dicho subconjunto, el resultado de contrastar $H_0: \mathbf{a}'\boldsymbol{\beta} = b$ frente a $H_1: \mathbf{a}'\boldsymbol{\beta} \neq b$ al $(100 \times \alpha)\%$ consiste en **no** rechazar H_0:

$$\mathrm{IC}_{1-\alpha}(\mathbf{a}'\boldsymbol{\beta}) = \left\{ b \in \mathbb{R} \; : \; |\, t \,| = \left| \frac{\mathbf{a}'\hat{\boldsymbol{\beta}} - b}{\mathrm{D\hat{v}t}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}]} \right| \leq t_{1-\frac{\alpha}{2}} \right\}. \qquad [85]$$

La cantidad $1-\alpha$ [o, en términos porcentuales, $100(1-\alpha)\%$] se denomina el **nivel de confianza** del IC [85].



Observación: Un IC contiene todos aquellos valores posibles de $\mathbf{a}'\boldsymbol{\beta}$ que son relativamente compatibles con la información contenida en los datos y resumida en un modelo estimado. La interpretación frecuente de un IC en los términos "$\Pr[\mathbf{a}'\boldsymbol{\beta} \in \mathrm{IC}_{1-\alpha}(\mathbf{a}'\boldsymbol{\beta})] = 1-\alpha$" es incorrecta porque $\mathbf{a}'\boldsymbol{\beta}$ es un número, no una variable aleatoria, de manera que el "suceso" que figura como argumento de la probabilidad anterior es o bien cierto ($\Rightarrow \Pr[\cdot] = 1$) o bien falso ($\Rightarrow \Pr[\cdot] = 0$).

IC.1: Un IC del $100(1-\alpha)\%$ para una combinación lineal de $\boldsymbol{\beta}$ del tipo $\mathbf{a}'\boldsymbol{\beta}$ se puede expresar y calcular como

$$\mathrm{IC}_{1-\alpha}(\mathbf{a}'\boldsymbol{\beta}) = \left[\mathbf{a}'\hat{\boldsymbol{\beta}} \mp \mathrm{D\hat{v}t}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}] \times t_{1-\frac{\alpha}{2}} \right]. \qquad [86]$$

Demostración: De acuerdo con [85], $\mathrm{IC}_{1-\alpha}(\mathbf{a}'\boldsymbol{\beta})$ es el conjunto formado por todos aquellos números reales $b \in \mathbb{R}$ que satisfacen la relación

$$\left| \mathbf{a}'\hat{\boldsymbol{\beta}} - b \right| \leq \mathrm{D\hat{v}t}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}] \times t_{1-\frac{\alpha}{2}}.$$

Esta desigualdad se puede escribir de manera equivalente como

$$-\mathrm{D\hat{v}t}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}] \times t_{1-\frac{\alpha}{2}} \leq \mathbf{a}'\hat{\boldsymbol{\beta}} - b \leq \mathrm{D\hat{v}t}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}] \times t_{1-\frac{\alpha}{2}},$$

o bien, finalmente, como

$$\mathbf{a}'\hat{\boldsymbol{\beta}} - \mathrm{D\hat{v}t}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}] \times t_{1-\frac{\alpha}{2}} \le b \le \mathbf{a}'\hat{\boldsymbol{\beta}} + \mathrm{D\hat{v}t}[\mathbf{a}'\hat{\boldsymbol{\beta}}_{\mathbf{W}}] \times t_{1-\frac{\alpha}{2}},$$

de donde resulta directamente [86]. ■

IC.2: Un IC del $100(1-\alpha)\%$ para para cualquier parámetro β_j $(1 \le j \le K)$ se puede expresar y calcular como

$$\mathrm{IC}_{1-\alpha}(\beta_j) = \left[\, \hat{\beta}_j \mp \mathrm{D\hat{v}t}[\hat{\beta}_{j\mathbf{W}}] \times t_{1-\frac{\alpha}{2}} \,\right]. \quad [87]$$

Demostración: Inmediata a partir de [86] con $\mathbf{a}' = [0, ..., 1, ..., 0]$ (un vector lleno de ceros excepto por un 1 en la posición j-ésima). ■

Observación I: [87] es un intervalo cerrado cuya amplitud depende (dado un nivel de confianza) del error estándar del estimador del parámetro β_j. En consecuencia, cuanto menor (mayor) sea dicho error estándar, menor (mayor) será el rango de valores posibles para β_j que son relativamente compatibles con los datos, y, por lo tanto, con mayor (menor) precisión o fiabilidad estará localizado el verdadero valor de β_j. Éste es el sentido que tiene interpretar el **error estándar** de $\hat{\beta}_{j\mathbf{W}}$ como una medida de su **precisión** o de la **fiabilidad** de la estimación puntual $\hat{\beta}_j$.

Observación II: De acuerdo con la forma en la que se han definido los IC [86]-[87], es inmediato comprobar la validez de las reglas siguientes:

1. En el contraste de $H_0: \mathbf{a}'\boldsymbol{\beta} = c$ frente a $H_1: \mathbf{a}'\boldsymbol{\beta} \ne c$, rechazar [no rechazar] H_0 en favor de H_1 al $(100 \times \alpha)\%$ cuando $c \notin \mathrm{IC}_{1-\alpha}(\mathbf{a}'\boldsymbol{\beta})$ [cuando $c \in \mathrm{IC}_{1-\alpha}(\mathbf{a}'\boldsymbol{\beta})$].

2. En el contraste de $H_0: \beta_j = c$ frente a $H_1: \beta_j \ne c$, rechazar [no rechazar] H_0 en favor de H_1 al $(100 \times \alpha)\%$ cuando $c \notin \mathrm{IC}_{1-\alpha}(\beta_j)$ [cuando $c \in \mathrm{IC}_{1-\alpha}(\beta_j)$].

EJ3 PR9

Resumen sobre Contrastes de Hipótesis en un Modelo RLM Clásico

H_0	H_1	Estadístico	Región Crítica con α dado	P-Valor
$\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$	$\mathbf{A}\boldsymbol{\beta} \ne \mathbf{c}$	F en [63] [67] [78] [79]	$F > F_{1-\alpha}$ con $\Pr[F(M, N-K) \le F_{1-\alpha}] = 1-\alpha$	$\Pr[F(M, N-K) \ge F]$
$\mathbf{a}'\boldsymbol{\beta} = \mathbf{c}$	$\mathbf{a}'\boldsymbol{\beta} \ne \mathbf{c}$	t en [82]	$\lvert t \rvert > t_{1-\frac{\alpha}{2}}$ con $\Pr[t(N-K) \le t_{1-\frac{\alpha}{2}}] = 1-\frac{\alpha}{2}$	$2 \times \Pr[t(N-K) \ge \lvert t \rvert]$
$\mathbf{a}'\boldsymbol{\beta} = \mathbf{c}$	$\mathbf{a}'\boldsymbol{\beta} > \mathbf{c}$	t en [82]	$t > t_{1-\alpha}$ con $\Pr[t(N-K) \le t_{1-\alpha}] = 1-\alpha$	$\Pr[t(N-K) \ge t]$
$\mathbf{a}'\boldsymbol{\beta} = \mathbf{c}$	$\mathbf{a}'\boldsymbol{\beta} < \mathbf{c}$	t en [82]	$t < t_{\alpha}$ con $\Pr[t(N-K) \le t_{\alpha}] = \alpha$	$\Pr[t(N-K) \le t]$

PARTE 7 - PREVISIÓN

De vuelta al ejemplo de la PARTE 2, una vez estimado el modelo RLS [2] (como en la figura de la página 5), se trataría de utilizar el modelo estimado para resolver estas cuestiones:

① Estimar el valor esperado de la nota final de un estudiante (la variable dependiente del modelo), es decir, calcular una **previsión puntual** de NF_f (una variable aleatoria), donde f hace referencia a un estudiante fuera de (no incluido en) la muestra empleada para estimar el modelo.

② Estimar algún **intervalo de confianza** $\mathrm{IC}_{1-\alpha}(NF_f)$ para NF_f con la propiedad de que $\Pr[NF_f \in \mathrm{IC}_{1-\alpha}(NF_f)] = 1-\alpha$ para algún α dado (típicamente 0.10, 0.05 ó 0.01).

③ Estimar **probabilidades** como $\Pr[NF_f \geq a]$, $\Pr[NF_f \leq b]$, o $\Pr[a \leq NF_f \leq b]$, con a y b dados (por ejemplo, $a = 5$, $b = 8$).

Todo ello utilizando un **valor numérico** ha_f **dado** (conocido o supuesto) para las horas de asistencia a clase (la variable explicativa del modelo).

Para resolver estas tres cuestiones, empezamos considerando, por analogía con [4], que

$$NF_f = \beta_1 + \beta_2 \times ha_f + U_f, \tag{88}$$

donde suponemos que U_f tiene las mismas propiedades que las perturbaciones que figuran en [4]. En particular,

$$\begin{gathered}\text{[A] } \mathrm{E}[U_f \,|\, \mathbf{X}] = 0, \ \text{[B] } \mathrm{Var}[U_f \,|\, \mathbf{X}] = \sigma^2, \\ \text{[C] } \mathrm{Cov}[U_f, U_j \,|\, \mathbf{X}] = 0 \ (j = 1, 2, ..., N), \ \text{[D] } U_f \,|\, \mathbf{X} \sim \text{Normal}.\end{gathered} \tag{89}$$

En relación con la cuestión ①, [88] junto con [A] en [89] implican que

$$\mathrm{E}[NF_f \,|\, \mathbf{X}] = \beta_1 + \beta_2 \times ha_f = \mathrm{E}[NF_f] \text{ (incondicionalmente).} \tag{90}$$

Haciendo uso de las estimaciones $\hat{\beta}_1, \hat{\beta}_2$ disponibles, [90] se puede estimar simplemente como $\hat{\beta}_1 + \hat{\beta}_2 \times ha_f$ (un valor ajustado "extramuestral"), que se denomina una **previsión puntual** (calculada) de la nota final NF_f y se representa como $\hat{nf}_f$:

$$\hat{nf}_f = \hat{\mathrm{E}}[NF_f] = \hat{\beta}_1 + \hat{\beta}_2 \times ha_f. \tag{91}$$

Así queda resuelta la cuestión ①. Para resolver las cuestiones ② y ③, expresamos, en primer lugar, [88] como

$$NF_f = \beta_1 + \beta_2 \times ha_f + U_f = \mathbf{x}'_f \boldsymbol{\beta} + U_f, \tag{92}$$

con $\mathbf{x}'_f = [1, ha_f]$, $\boldsymbol{\beta} = [\beta_1, \beta_2]'$, y definimos, en segundo lugar, el **estimador** (una variable aleatoria) asociado con la estimación [91] (un número) como

$$\hat{NF}_f = \hat{\beta}_{1\mathbf{W}} + \hat{\beta}_{2\mathbf{W}} \times ha_f = \mathbf{x}'_f \hat{\boldsymbol{\beta}}_{\mathbf{W}}, \tag{93}$$

con $\mathbf{x}'_f = [1, ha_f]$, $\hat{\boldsymbol{\beta}}_{\mathbf{W}} = [\hat{\beta}_{1\mathbf{W}}, \hat{\beta}_{2\mathbf{W}}]'$.

El **error de previsión** (teórico) asociado con [92] y [93] es

$$\mathcal{E}_f = NF_f - \hat{NF}_f = U_f - \mathbf{x}_f'(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta}), \tag{94}$$

cuyas propiedades estadísticas fundamentales bajo [HC1]-[HC5] y [89] son las siguientes:

$$\mathrm{E}[\mathcal{E}_f \mid \mathbf{X}] = 0, \tag{95}$$

$$\mathrm{Var}[\mathcal{E}_f \mid \mathbf{X}] = \sigma^2 \times [1 + \mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_f] = \sigma^2 + \mathrm{Var}[\hat{NF}_f \mid \mathbf{X}], \tag{96}$$

$$\frac{\mathcal{E}_f}{\hat{V}_f} \sim t(N-K), \tag{97}$$

donde

$$\hat{V}_f = \sqrt{\hat{\sigma}^2_{\mathbf{W}} \times [1 + \mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_f]} \tag{98}$$

es el estimador MCO de la **desviación típica** del error de previsión. En relación con las cuestiones ② y ③, [97] implica, por un lado, que (ver figura en página 23)

$$\Pr\left[-t_{1-\frac{\alpha}{2}} \le \frac{NF_f - \hat{NF}_f}{\hat{V}_f} \le t_{1-\frac{\alpha}{2}}\right] = 1-\alpha \Rightarrow \mathrm{IC}_{1-\alpha}(NF_f) = \left[\hat{NF}_f \mp \hat{V}_f \times t_{1-\frac{\alpha}{2}}\right],$$

y, por otro, que

$$\Pr[NF_f \ge a] = \Pr\left[\frac{NF_f - \hat{NF}_f}{\hat{V}_f} \ge \frac{a - \hat{NF}_f}{\hat{V}_f}\right] = \Pr\left[t(N-K) \ge \frac{a - \hat{NF}_f}{\hat{V}_f}\right].$$

Reemplazando en estas expresiones los estimadores $\hat{NF}_f$ y $\hat{V}_f$ por las estimaciones numéricas $\hat{nf}_f$ y $\hat{v}_f$ correspondientes, quedan resueltas las cuestiones ② y ③:

$$\hat{\mathrm{IC}}_{1-\alpha}(NF_f) = \left[\hat{nf}_f \mp \hat{v}_f \times t_{1-\frac{\alpha}{2}}\right], \tag{99}$$

$$\hat{\Pr}[NF_f \ge a] = \Pr\left[t(N-K) \ge \frac{a - \hat{nf}_f}{\hat{v}_f}\right], \tag{100}$$

y, análogamente,

$$\hat{\Pr}[NF_f \le b] = \Pr\left[t(N-K) \le \frac{b - \hat{nf}_f}{\hat{v}_f}\right], \tag{101}$$

$$\hat{\Pr}[a \le NF_f \le b] = \Pr\left[\frac{a - \hat{nf}_f}{\hat{v}_f} \le t(N-K) \le \frac{b - \hat{nf}_f}{\hat{v}_f}\right], \tag{102}$$

donde

$$\hat{v}_f = \sqrt{\hat{\sigma}^2 + \hat{\mathrm{Var}}[\hat{NF}_f \mid \mathbf{X}]} \tag{103}$$

es la estimación MCO de la **desviación típica** del error de previsión (la desviación típica estimada, o el error estándar, del error de previsión). Las fórmulas [91] (previsión puntual) y [99]-[103] son fácilmente generalizables al caso de que las operaciones correspondientes estén referidas a la previsión de la variable dependiente Y de un modelo RLM clásico en un punto extramuestral f utilizando unos valores dados $x_{f2}, ..., x_{fK}$ para las variables explicativas $X_2, ..., X_K$. En este caso, $Y_f = \beta_1 + \beta_2 \times x_{f2} + ... + \beta_K \times x_{fK} + U_f$, con

$$\hat{y}_f = \hat{\mathrm{E}}[Y_f] = \hat{\beta}_1 + \hat{\beta}_2 \times x_{f2} + ... + \hat{\beta}_K \times x_{fK}, \quad [104]$$

$$\hat{\mathrm{IC}}_{1-\alpha}(Y_f) = \left[\, \hat{y}_f \mp \hat{v}_f \times t_{1-\frac{\alpha}{2}} \,\right], \quad [105]$$

$$\hat{\Pr}[\, Y_f \geq a \,] = \Pr\left[t(N-K) \geq \frac{a - \hat{y}_f}{\hat{v}_f} \right], \quad [106]$$

$$\hat{\Pr}[\, Y_f \leq b \,] = \Pr\left[t(N-K) \leq \frac{b - \hat{y}_f}{\hat{v}_f} \right], \quad [107]$$

$$\hat{\Pr}[\, a \leq Y_f \leq b \,] = \Pr\left[\frac{a - \hat{y}_f}{\hat{v}_f} \leq t(N-K) \leq \frac{b - \hat{y}_f}{\hat{v}_f} \right], \quad [108]$$

donde

$$\hat{v}_f = \sqrt{\hat{\sigma}^2 + \hat{\mathrm{Var}}[\hat{Y}_f \,|\, \mathbf{X}]}, \quad [109]$$

$$\hat{\mathrm{Var}}[\hat{Y}_f \,|\, \mathbf{X}] = \hat{\mathrm{Var}}[\hat{\beta}_{1\mathbf{W}} + x_{f2} \times \hat{\beta}_{2\mathbf{W}} + ... + x_{fK} \times \hat{\beta}_{K\mathbf{W}} \,|\, \mathbf{X}]. \quad [110]$$

EJ3 PR10

Demostración de [95]-[98]:

El error de previsión en [94] se puede expresar (ver página 7) como

$$\begin{aligned} \mathcal{E}_f &= U_f - \mathbf{x}_f'(\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \boldsymbol{\beta}) = U_f - \mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{U} \\ &= [-\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}', 1] \begin{bmatrix} \mathbf{U} \\ U_f \end{bmatrix} = \mathbf{x}_*'\mathbf{U}_*, \end{aligned} \quad [111]$$

con $\mathrm{E}[\mathbf{U}_* \,|\, \mathbf{X}] = \mathbf{0}$ y $\mathrm{Var}[\mathbf{U}_* \,|\, \mathbf{X}] = \sigma^2\mathbf{I}$ (por [HC1]-[HC4] y [A]-[C] en [89]). Por lo tanto,

$$\mathrm{E}[\mathcal{E}_f \,|\, \mathbf{X}] = \mathrm{E}[\mathbf{x}_*'\mathbf{U}_* \,|\, \mathbf{X}] = \mathbf{x}_*'\mathrm{E}[\mathbf{U}_* \,|\, \mathbf{X}] = \mathbf{x}_*'\mathbf{0} = 0, \quad [112]$$

lo que prueba [95], y

$$\begin{aligned} \mathrm{Var}[\mathcal{E}_f \,|\, \mathbf{X}] &= \mathrm{Var}[\mathbf{x}_*'\mathbf{U}_* \,|\, \mathbf{X}] = \mathbf{x}_*'\mathrm{Var}[\mathbf{U}_* \,|\, \mathbf{X}]\mathbf{x}_* = \sigma^2\mathbf{x}_*'\mathbf{x}_* \\ &= \sigma^2[-\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}', 1] \begin{bmatrix} -\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_f \\ 1 \end{bmatrix} \\ &= \sigma^2 \times [\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_f + 1]. \end{aligned} \quad [113]$$

Adicionalmente, [93] implica que

$$\mathrm{Var}[\hat{NF}_f \,|\, \mathbf{X}] = \mathrm{Var}[\mathbf{x}_f'\hat{\boldsymbol{\beta}}_{\mathbf{W}} \,|\, \mathbf{X}] = \mathbf{x}_f'\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\mathbf{W}} \,|\, \mathbf{X}]\mathbf{x}_f = \sigma^2\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_f, \quad [114]$$

lo que, junto con [113], prueba [96]. Por otro lado, una manera de comprobar [97]-[98]

consiste en demostrar que (Apéndice A2.4)

$$\frac{\mathcal{E}_f}{\hat{V}_f} = \frac{W_1}{[Z_2/(N-K)]^{½}} \qquad [115]$$

con $W_1 \sim \text{Normal}(0,1)$ y $Z_2 \sim \chi^2(N-K)$ independientes. Con esta intención, definimos

$$W_1 = \frac{\mathcal{E}_f - \text{E}[\mathcal{E}_f \,|\, \mathbf{X}]}{\text{Var}[\mathcal{E}_f \,|\, \mathbf{X}]^{½}} = \frac{\mathbf{x}_*'\mathbf{U}_*}{\sigma \times [1 + \mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_f]^{½}} \sim \text{Normal}(0,1), \qquad [116]$$

$$\begin{aligned} Z_2 &= \sigma^{-2}\hat{\mathbf{U}}'\hat{\mathbf{U}} = \sigma^{-2}\mathbf{U}'\mathbf{M}\mathbf{U} = \sigma^{-2}[\mathbf{U}', U_f]\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix}\begin{bmatrix} \mathbf{U} \\ U_f \end{bmatrix} \\ &= \sigma^{-2}\mathbf{U}_*'\mathbf{M}_*\mathbf{U}_* \sim \chi^2(N-K), \end{aligned} \qquad [117]$$

donde $\sigma^{-1}\mathbf{U}_* \,|\, \mathbf{X} \sim \text{Normal}(\mathbf{0}, \mathbf{I})$ y $\mathbf{M}_*$ es una matriz $(N+1)\times(N+1)$ simétrica e idempotente con $\text{rango}(\mathbf{M}_*) = \text{rango}(\mathbf{M}) = \text{traza}(\mathbf{M}) = N-K$. Las variables aleatorias en [116]-[117] (una transformación lineal y una transformación cuadrática de $\sigma^{-1}\mathbf{U}_*$) son independientes porque

$$\mathbf{x}_*'\mathbf{M}_* = [-\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}', 1]\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix} = \begin{bmatrix} -\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{M} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix}, \text{ con}$$

$$\begin{aligned} -\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{M} &= -\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \\ &= -\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' + \mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\ &= -\mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' + \mathbf{x}_f'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{0}. \end{aligned}$$

Sustituyendo [116]-[117] en el lado derecho de [115] se obtiene el resultado propuesto. ■

APÉNDICE 1 - PROPIEDADES ESTADÍSTICAS DEL ESTIMADOR MCR DE β

De acuerdo con [66], el estimador MCR de $\boldsymbol{\beta}$ es el vector $K \times 1$ de variables aleatorias

$$\hat{\boldsymbol{\beta}}_{\mathbf{W}}^* = \hat{\boldsymbol{\beta}}_{\mathbf{W}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}}_{\mathbf{W}} - \mathbf{c}),$$

donde $\hat{\boldsymbol{\beta}}_{\mathbf{W}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ es el estimador MCO (no restringido) de $\boldsymbol{\beta}$. Bajo [HC1]-[HC4], el estimador MCR de $\boldsymbol{\beta}$ tiene las propiedades estadísticas siguientes:

MCR.1: Si $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ ($\mathbf{A}\boldsymbol{\beta} \neq \mathbf{c}$), entonces $\text{E}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}^*] = \boldsymbol{\beta}$ ($\text{E}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}^*] \neq \boldsymbol{\beta}$).

Demostración:

$$\begin{aligned} \text{E}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}^*] &= \text{E}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\text{E}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}] - \mathbf{c}) \\ &= \boldsymbol{\beta} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\boldsymbol{\beta} - \mathbf{c}). \end{aligned}$$ ■

MCR.2: $\text{Var}[\hat{\boldsymbol{\beta}}_{\mathbf{W}}^*] = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} - \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}$, tanto en el caso de que $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ como en el caso de que $\mathbf{A}\boldsymbol{\beta} \neq \mathbf{c}$.

Demostración:

$$\hat{\boldsymbol\beta}^*_{\mathbf W} - \mathrm{E}[\hat{\boldsymbol\beta}^*_{\mathbf W}] = (\hat{\boldsymbol\beta}_{\mathbf W} - \boldsymbol\beta) - (\mathbf X'\mathbf X)^{-1}\mathbf A'[\mathbf A(\mathbf X'\mathbf X)^{-1}\mathbf A']^{-1}\mathbf A(\hat{\boldsymbol\beta}_{\mathbf W} - \boldsymbol\beta)$$
$$= \mathbf D(\hat{\boldsymbol\beta}_{\mathbf W} - \boldsymbol\beta), \text{ con}$$
$$\mathbf D = \mathbf I - (\mathbf X'\mathbf X)^{-1}\mathbf A'[\mathbf A(\mathbf X'\mathbf X)^{-1}\mathbf A']^{-1}\mathbf A.$$

$$\mathrm{Var}[\hat{\boldsymbol\beta}^*_{\mathbf W}] = \mathrm{E}[\mathbf D(\hat{\boldsymbol\beta}_{\mathbf W} - \boldsymbol\beta)(\hat{\boldsymbol\beta}_{\mathbf W} - \boldsymbol\beta)'\mathbf D'] = \mathbf D\mathrm{Var}[\hat{\boldsymbol\beta}_{\mathbf W}]\mathbf D' = \sigma^2\mathbf D(\mathbf X'\mathbf X)^{-1}\mathbf D', \text{ con}$$
$$\mathbf D(\mathbf X'\mathbf X)^{-1}\mathbf D' =$$
$$= \left(\mathbf I - (\mathbf X'\mathbf X)^{-1}\mathbf A'[\mathbf A(\mathbf X'\mathbf X)^{-1}\mathbf A']^{-1}\mathbf A\right)(\mathbf X'\mathbf X)^{-1}\left(\mathbf I - (\mathbf X'\mathbf X)^{-1}\mathbf A'[\mathbf A(\mathbf X'\mathbf X)^{-1}\mathbf A']^{-1}\mathbf A\right)'$$
$$= (\mathbf X'\mathbf X)^{-1} - (\mathbf X'\mathbf X)^{-1}\mathbf A'[\mathbf A(\mathbf X'\mathbf X)^{-1}\mathbf A']^{-1}\mathbf A(\mathbf X'\mathbf X)^{-1}.$$

MCR.3: $\mathrm{Var}[\hat{\boldsymbol\beta}_{\mathbf W}] - \mathrm{Var}[\hat{\boldsymbol\beta}^*_{\mathbf W}]$ es una matriz semidefinida positiva.

Demostración:

$$\mathrm{Var}[\hat{\boldsymbol\beta}_{\mathbf W}] - \mathrm{Var}[\hat{\boldsymbol\beta}^*_{\mathbf W}] = \sigma^2(\mathbf X'\mathbf X)^{-1}\mathbf A'[\mathbf A(\mathbf X'\mathbf X)^{-1}\mathbf A']^{-1}\mathbf A(\mathbf X'\mathbf X)^{-1},$$

donde $\sigma^2 > 0$, $(\mathbf X'\mathbf X)^{-1}$ $K\times K$ definida positiva [$\Leftarrow \mathrm{rango}(\mathbf X) = K < N$], y $\mathbf A$ $M\times K$ con $\mathrm{rango}(\mathbf A) = M \le K$ ($\Rightarrow \mathbf A'[\mathbf A(\mathbf X'\mathbf X)^{-1}\mathbf A']^{-1}\mathbf A$ $K\times K$ semidefinida positiva).

La reducción en varianza asociada con el estimador MCR con respecto al estimador MCO puede interpretarse como el "valor" de la información contenida en las restricciones (sean ciertas o no, las restricciones contienen información no incluida en el estimador MCO).

MCR.4: Considerando las mismas particiones que en FSC.3, la matriz $\mathrm{Var}[\hat{\boldsymbol\beta}^*_{\mathbf W}]$ en MCR.2 se puede expresar por bloques como

$$\underset{K\times K}{\mathrm{Var}[\hat{\boldsymbol\beta}^*_{\mathbf W}]} = \begin{bmatrix} \underset{M\times M}{\mathrm{Var}[\hat{\boldsymbol\beta}^*_{1\mathbf W}]} & \underset{M\times(K-M)}{\mathrm{Cov}[\hat{\boldsymbol\beta}^*_{1\mathbf W}, \hat{\boldsymbol\beta}^*_{2\mathbf W}]} \\ \underset{(K-M)\times M}{\mathrm{Cov}[\hat{\boldsymbol\beta}^*_{2\mathbf W}, \hat{\boldsymbol\beta}^*_{1\mathbf W}]} & \underset{(K-M)\times(K-M)}{\mathrm{Var}[\hat{\boldsymbol\beta}^*_{2\mathbf W}]} \end{bmatrix}, \text{ con}$$

$$\mathrm{Var}[\hat{\boldsymbol\beta}^*_{1\mathbf W}] = \sigma^2\mathbf A_1^{-1}\mathbf A_2(\mathbf X_*'\mathbf X_*)^{-1}\mathbf A_2'\mathbf A_1^{-1\prime},$$
$$\mathrm{Var}[\hat{\boldsymbol\beta}^*_{2\mathbf W}] = \sigma^2(\mathbf X_*'\mathbf X_*)^{-1},$$
$$\mathrm{Cov}[\hat{\boldsymbol\beta}^*_{1\mathbf W}, \hat{\boldsymbol\beta}^*_{2\mathbf W}] = \sigma^2\mathbf A_1^{-1}\mathbf A_2(\mathbf X_*'\mathbf X_*)^{-1},$$
$$\mathrm{Cov}[\hat{\boldsymbol\beta}^*_{2\mathbf W}, \hat{\boldsymbol\beta}^*_{1\mathbf W}] = \sigma^2(\mathbf X_*'\mathbf X_*)^{-1}\mathbf A_2'\mathbf A_1^{-1\prime},$$

donde $\mathbf A_1, \mathbf A_2, \mathbf X_*$ son las matrices dadas en [68]-[71].

Demostración: Inmediata a partir de [71] y [77].

Observación: Con los archivos de programa para EViews MCR1.PRG y MCR2.PRG se pueden llevar a cabo paso a paso, a partir de cualquier modelo RLM estimado por MCO, todas las operaciones referidas al cálculo de estadísticos F con las fórmulas [63] y [67], incluyendo el cálculo de $SCR^* = \hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*$ a través de $\hat{\mathbf{u}}_* = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}^*$ con $\hat{\boldsymbol{\beta}}^*$ dado en [66] (MCR1.PRG) y a través de [68]-[71] y [77] mediante sustitución de restricciones (MCR2.PRG).

EJ4

APÉNDICE 2 - DISTRIBUCIONES DE PROBABILIDAD

Para más detalles sobre lo que se resume en este apéndice, ver Mittelhammer, R.C. (2013), *Mathematical Statistics for Economics and Business (Second Edition)*, Springer.

A2.1 VARIABLES ALEATORIAS - DISTRIBUCIONES UNIVARIANTES

Una variable aleatoria X es una característica concreta de una entidad observable tal que:

1. Es susceptible de tomar distintos valores numéricos (reales).
2. Todos los valores que puede tomar son perfectamente conocidos.
3. El valor que realmente toma no se puede prever o anticipar con absoluta certeza.

Cada uno de los valores que puede tomar X se denomina un **valor posible** o una **realización posible** de X. El conjunto de todos los valores posibles de X es el **rango** $R(X)$ de X. Un valor que realmente toma (o que se supone que toma) X es un **valor observado** o una **realización particular** de X.

Variable Aleatoria Discreta:

[VAD.1] $R(X)$ es un conjunto finito o infinito numerable,

[VAD.2] existe una función real $f_X(x) \geq 0$ para todo $x \in (-\infty, +\infty)$ tal que

$$\text{[i]}\ f_X(x) = \Pr[X = x] \text{ para todo } x \in (-\infty, +\infty),$$

$$\text{[ii]}\ x \notin R(X) \Rightarrow f_X(x) = 0\,,\ \text{[iii]}\ \textstyle\sum_{x \in R(X)} f_X(x) = 1\,.$$

Variable Aleatoria Continua:

[VAC.1] $R(X)$ es un conjunto infinito no numerable,

[VAC.2] existe una función real $f_X(x) \geq 0$ para todo $x \in (-\infty, +\infty)$ tal que

$$\text{[i]}\ \textstyle\int_{x \in A} f_X(x)dx = \Pr[X \in A] \text{ para todo } A \subseteq (-\infty, +\infty),$$

$$\text{[ii]}\ x \notin R(X) \Rightarrow f_X(x) = 0\,,\ \text{[iii]}\ \textstyle\int_{-\infty}^{+\infty} f_X(x)dx = 1\,.$$

Función de Densidad (PDF): $$f_X(x). \qquad \text{[A2.1]}$$

$$\Pr[x_1 \leq X \leq x_2] = \textstyle\sum_{x \in B} f_X(x),\ \text{ con } B = \{\, b \in R(X) : x_1 \leq b \leq x_2 \,\}, \qquad \text{[A2.2]}$$

$$\Pr[x_1 \leq X \leq x_2] = \textstyle\int_{x_1}^{x_2} f_X(x)dx. \qquad \text{[A2.3]}$$

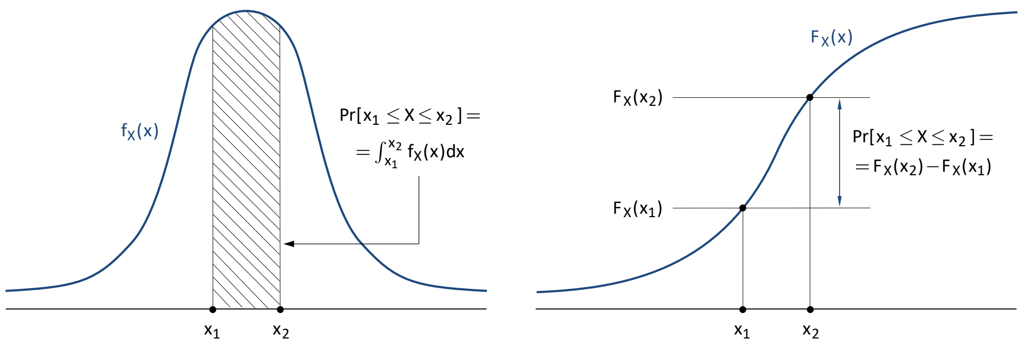
Función de Distribución (CDF):

$$F_X(x) = \Pr[X \leq x] \text{ para todo } x \in (-\infty, +\infty). \qquad \text{[A2.4]}$$

$$F_X(x) = \sum_{t \in B} f_X(t), \text{ con } B = \{ b \in R(X) : b \leq x \}, \qquad [A2.5]$$

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \qquad [A2.6]$$

$$\Pr[x_1 < X \leq x_2] = F_X(x_2) - F_X(x_1). \qquad [A2.7]$$



Esperanza: $$\mathrm{E}[X] = \sum_{x \in R(X)} x f_X(x), \ \mathrm{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)dx. \qquad [A2.8]$$

$$\mathrm{E}[g(X)] = \sum_{x \in R(X)} g(x) f_X(x), \ \mathrm{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x)dx. \qquad [A2.9]$$

Propiedades de la Esperanza:

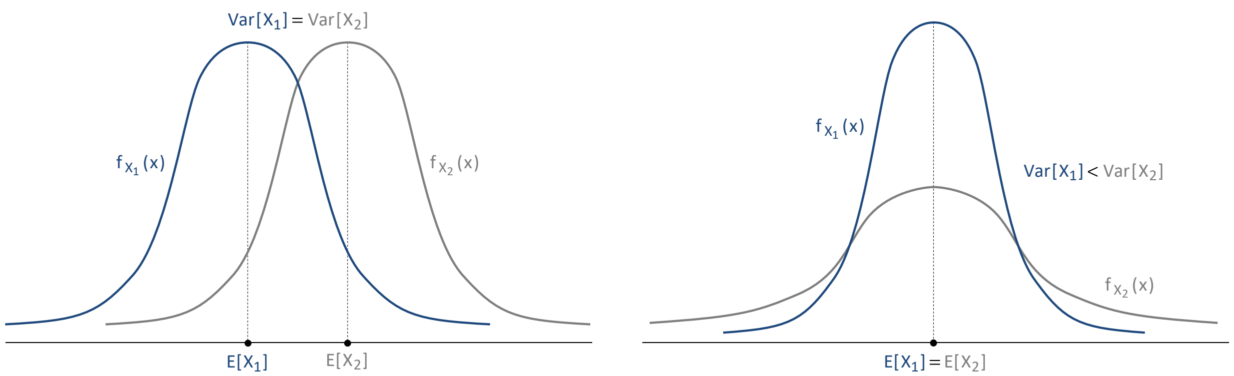
$$\mathrm{E}[c] = c. \qquad [A2.10]$$

$$\mathrm{E}[cX] = c\mathrm{E}[X]. \qquad [A2.11]$$

$$\mathrm{E}\left[\sum_{i=1}^{n} g_i(X) \right] = \sum_{i=1}^{n} \mathrm{E}[g_i(X)]. \qquad [A2.12]$$

$$\mathrm{E}[c_1 + c_2 X] = c_1 + c_2 \mathrm{E}[X]. \qquad [A2.13]$$

Varianza: $$\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]. \qquad [A2.14]$$



Propiedades de la Varianza:

$$\mathrm{Var}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2. \qquad [A2.15]$$

$$\begin{gathered} \mathrm{Var}[c] = 0, \\ \mathrm{Var}[c_1 + c_2 X] = c_2^2 \mathrm{Var}[X]. \end{gathered} \qquad [A2.16]$$

Desviación Típica: $$\mathrm{Dvt}[X] = \sqrt{\mathrm{Var}[X]}. \qquad [A2.17]$$

A2.2 DISTRIBUCIONES BIVARIANTES CONTINUAS

Función de Densidad (PDF) Conjunta: $f_{UX}(u,x)$, [A2.18]

$$f_{UX}(u,x) \geq 0,\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{UX}(u,x)dxdu = 1. \quad [A2.19]$$

$$\Pr[u_1 \leq U \leq u_2, x_1 \leq X \leq x_2] = \int_{u_1}^{u_2}\int_{x_1}^{x_2} f_{UX}(u,x)dxdu. \quad [A2.20]$$

Esperanza: $\mathrm{E}[g(U,X)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(u,x) f_{UX}(u,x)dxdu.$ [A2.21]

PDFs Marginales:
$$f_U(u) = \int_{-\infty}^{+\infty} f_{UX}(u,x)dx,$$
$$f_X(x) = \int_{-\infty}^{+\infty} f_{UX}(u,x)du. \quad [A2.22]$$

Esperanzas Marginales:
$$\mathrm{E}[U] = \int_{-\infty}^{+\infty} u f_U(u)du,$$
$$\mathrm{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)dx. \quad [A2.23]$$

Varianzas Marginales:
$$\mathrm{Var}[U] = \mathrm{E}[(U - \mathrm{E}[U])^2],$$
$$\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]. \quad [A2.24]$$

Covarianza:
$$\begin{aligned}\mathrm{Cov}[U,X] &= \mathrm{E}[(U-\mathrm{E}[U])(X-\mathrm{E}[X])] \\ &= \mathrm{E}[UX] - \mathrm{E}[U]\,\mathrm{E}[X] = \mathrm{Cov}[X,U].\end{aligned} \quad [A2.25]$$

Correlación: $\mathrm{Corr}[U,X] = \dfrac{\mathrm{Cov}[U,X]}{\mathrm{Dvt}[U]\,\mathrm{Dvt}[X]},\ -1 \leq \mathrm{Corr}[U,X] \leq 1.$ [A2.26]

Propiedades:
$$\mathrm{E}\left[\textstyle\sum_{i=1}^n g_i(U,X)\right] = \textstyle\sum_{i=1}^n \mathrm{E}[g_i(U,X)]. \quad [A2.27]$$

$$\mathrm{E}[c_1U + c_2X] = c_1\mathrm{E}[U] + c_2\mathrm{E}[X]. \quad [A2.28]$$

$$\mathrm{Var}[c_1U + c_2X] = c_1^2\mathrm{Var}[U] + c_2^2\mathrm{Var}[X] + 2c_1c_2\mathrm{Cov}[U,X]. \quad [A2.29]$$

Independencia [1]: $f_{UX}(u,x) = f_U(u) \times f_X(x).$ [A2.30]

$$Q_1 = q_1(U) \text{ y } Q_2 = q_2(X) \text{ son independientes}, \quad [A2.31]$$

$$\mathrm{E}[UX] = \mathrm{E}[U] \times \mathrm{E}[X] \Rightarrow \mathrm{Cov}[U,X] = \mathrm{Corr}[U,X] = 0. \quad [A2.32]$$

PDFs Condicionales:
$$f_{U|X}(u,x) = \frac{1}{f_X(x)} \times f_{UX}(u,x),$$
$$f_{X|U}(u,x) = \frac{1}{f_U(u)} \times f_{UX}(u,x). \quad [A2.33]$$

A partir de cualquiera de las dos PDFs condicionales en [A2.33], se pueden considerar todas las características "univariantes" de la distribución condicional correspondiente (análogas a las del Apéndice A2.1) y algunas más. En particular, condicionando por X,

$$\mathrm{E}[U \mid X = x] = \int_{-\infty}^{+\infty} u f_{U\mid X}(u,x)du = \mu_1(x), \qquad [A2.34]$$

$$\mathrm{E}[g(U,X) \mid X = x] = \int_{-\infty}^{+\infty} g(u,x) f_{U\mid X}(u,x)du = \mu_2(x) \qquad [A2.35]$$

son funciones que dependen de x (algún número real), mientras que

$$\mathrm{E}[U \mid X] = \mu_1(X), \qquad [A2.36]$$

$$\mathrm{E}[g(U,X) \mid X] = \mu_2(X) \qquad [A2.37]$$

son funciones que dependen de X (una variable aleatoria), por lo que las dos esperanzas condicionales en [A2.36] y [A2.37] son variables aleatorias.

Propiedades de la Esperanza Condicional:

$$\mathrm{E}[c \mid X] = c. \qquad [A2.38]$$

$$\mathrm{E}[cU \mid X] = c\mathrm{E}[U \mid X]. \qquad [A2.39]$$

$$\mathrm{E}\left[\sum_{i=1}^{n} g_i(U) \mid X \right] = \sum_{i=1}^{n} \mathrm{E}[g_i(U) \mid X]. \qquad [A2.40]$$

$$\mathrm{E}[c_1 + c_2 U \mid X] = c_1 + c_2 \mathrm{E}[U \mid X]. \qquad [A2.41]$$

$$\mathrm{E}[h(X) \mid X] = h(X). \qquad [A2.42]$$

$$\mathrm{E}[g_1(X) \times U + g_2(X) \mid X] = g_1(X) \times \mathrm{E}[U \mid X] + g_2(X). \qquad [A2.43]$$

$$\mathrm{E}[g(U,X)] = \mathrm{E}[\mathrm{E}[g(U,X) \mid X]]. \qquad [A2.44]$$

Varianza Condicional:

$$\mathrm{Var}[U \mid X] = \mathrm{E}[(U - \mathrm{E}[U \mid X])^2 \mid X] = \mathrm{E}[U^2 \mid X] - \mathrm{E}[U \mid X]^2. \qquad [A2.45]$$

Independencia [2]:

$$f_{UX}(u,x) = f_U(u) \times f_X(x) \Leftrightarrow f_{U\mid X}(u,x) = f_U(u). \qquad [A2.46]$$

$$f_{U\mid X}(u,x) = f_U(u) \Rightarrow \mathrm{E}[g(U,X) \mid X] = \mathrm{E}[g(U,X)], \qquad [A2.47]$$

$$\mathrm{E}[U \mid X] = \mathrm{E}[U],\ \mathrm{Var}[U \mid X] = \mathrm{Var}[U]. \qquad [A2.48]$$

$$\mathrm{E}[U \mid X] = \mathrm{E}[U] \Rightarrow \mathrm{Cov}[U,X] = \mathrm{Corr}[U,X] = 0. \qquad [A2.49]$$

De hecho, $\mathrm{E}[U \mid X] = \mathrm{E}[U] \Rightarrow U$ está incorrelacionada con cualquier función de X.

$$\mathrm{E}[U \mid X] = 0 \Rightarrow \mathrm{E}[U] = 0 \text{ y } \mathrm{Corr}[U,X] = 0. \qquad [A2.50]$$

Demostración de [A2.42]:

[A2.37] con $g(U,X) = h(X) \Rightarrow \mathrm{E}[h(X) \mid X] = \mu_2(X)$. Para encontrar la forma funcional de $\mu_2(\cdot)$, [A2.35] con $g(U,X) = h(X)$ proporciona lo siguiente:

$$\mathrm{E}[h(X) \mid X = x] = \int_{-\infty}^{+\infty} h(x) f_{U\mid X}(u,x)du = h(x)\int_{-\infty}^{+\infty} f_{U\mid X}(u,x)du = h(x),$$

donde se ha tenido en cuenta que, por [A2.33] y [A2.22],

$$\int_{-\infty}^{+\infty} f_{U|X}(u,x)du = \int_{-\infty}^{+\infty} \frac{1}{f_X(x)} f_{UX}(u,x)du = \frac{1}{f_X(x)} \int_{-\infty}^{+\infty} f_{UX}(u,x)du = \frac{f_X(x)}{f_X(x)} = 1,$$

como ocurre, por otro lado, con cualquier función de densidad continua de cualquier tipo (condicional, marginal, o conjunta). En consecuencia, $\mu_2(x) = h(x)$ y $\mathrm{E}[h(X)\,|\,X] = \mu_2(X) = h(X)$.

Demostración de [A2.44]:

[A2.21] $\Rightarrow$ $\mathrm{E}[g(U,X)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(u,x) f_{UX}(u,x)dxdu$

[A2.33] $\Rightarrow$ $= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(u,x) f_{U|X}(u,x) f_X(x)dxdu$

Reagrupando términos: $= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} g(u,x) f_{U|X}(u,x)du\right] f_X(x)dx$

[A2.35] $\Rightarrow$ $= \int_{-\infty}^{+\infty} m_2(x) f_X(x)dx$

[A2.9] $\Rightarrow$ $= \mathrm{E}[m_2(X)]$

[A2.37] $\Rightarrow$ $= \mathrm{E}[\mathrm{E}[g(U,X)\,|\,X]]$.

Demostración de [A2.49]:

$$\begin{aligned}\mathrm{Cov}[U,X] &= \mathrm{E}[UX] - \mathrm{E}[U]\times\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[UX\,|\,X]] - \mathrm{E}[U]\times\mathrm{E}[X] \\ &= \mathrm{E}[\mathrm{E}[U\,|\,X]\times X] - \mathrm{E}[U]\times\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[U]\times X] - \mathrm{E}[U]\times\mathrm{E}[X] \\ &= \mathrm{E}[U]\times\mathrm{E}[X] - \mathrm{E}[U]\times\mathrm{E}[X] = 0.\end{aligned}$$

La demostración es idéntica si se reemplaza X por cualquier función $h(X)$.

Demostración de [A2.50]:

[A2.44] $\Rightarrow$ $\mathrm{E}[U] = \mathrm{E}[\mathrm{E}[U\,|\,X]] = \mathrm{E}[0] = 0.$

[A2.49] $\Rightarrow$ $\mathrm{E}[U\,|\,X] = \mathrm{E}[U]\ (=0) \Rightarrow \mathrm{Cov}[U,X] = \mathrm{Corr}[U,X] = 0.$

A2.3 VECTORES DE ESPERANZAS Y MATRICES DE VARIANZAS-COVARIANZAS

Si $\mathbf{W}$ y $\mathbf{X}$ son matrices $N\times M$ y $N\times K$, respectivamente, de variables aleatorias, entonces los símbolos $\mathrm{E}[\mathbf{W}]$ y $\mathrm{E}[\mathbf{W}\,|\,\mathbf{X}]$ representan matrices de orden $N\times M$ cuyos elementos en las posiciones ij ($i = 1,2,...,N$; $j = 1,2,...,M$) son $\mathrm{E}[W_{ij}]$ y $\mathrm{E}[W_{ij}\,|\,\mathbf{X}]$, respectivamente. En particular, si $\mathbf{V} = [V_1, V_2, ..., V_N]'$ es un vector $N\times 1$ (columna) de variables aleatorias, entonces

$$\mathrm{E}[\mathbf{V}] = \begin{bmatrix}\mathrm{E}[V_1]\\ \mathrm{E}[V_2]\\ \vdots \\ \mathrm{E}[V_N]\end{bmatrix},\ \mathrm{E}[\mathbf{V}\,|\,\mathbf{X}] = \begin{bmatrix}\mathrm{E}[V_1\,|\,\mathbf{X}]\\ \mathrm{E}[V_2\,|\,\mathbf{X}]\\ \vdots \\ \mathrm{E}[V_N\,|\,\mathbf{X}]\end{bmatrix} \qquad \text{[A2.51]}$$

son los **vectores de esperanzas** (medias) incondicionales y condicionales, respectivamente, de $\mathbf{V}$. Por otro lado,

$$(\mathbf{V}-\mathrm{E}[\mathbf{V}])(\mathbf{V}-\mathrm{E}[\mathbf{V}])' = \begin{pmatrix} V_1 - \mathrm{E}[V_1] \\ V_2 - \mathrm{E}[V_2] \\ \vdots \\ V_N - \mathrm{E}[V_N] \end{pmatrix} (V_1 - \mathrm{E}[V_1], V_2 - \mathrm{E}[V_2], ..., V_N - \mathrm{E}[V_N]) =$$

$$= \begin{bmatrix} (V_1 - \mathrm{E}[V_1])^2 & (V_1 - \mathrm{E}[V_1])(V_2 - \mathrm{E}[V_2]) & \cdots & (V_1 - \mathrm{E}[V_1])(V_N - \mathrm{E}[V_N]) \\ (V_2 - \mathrm{E}[V_2])(V_1 - \mathrm{E}[V_1]) & (V_2 - \mathrm{E}[V_2])^2 & \cdots & (V_2 - \mathrm{E}[V_2])(V_N - \mathrm{E}[V_N]) \\ \vdots & \vdots & \ddots & \vdots \\ (V_N - \mathrm{E}[V_N])(V_1 - \mathrm{E}[V_1]) & (V_N - \mathrm{E}[V_N])(V_2 - \mathrm{E}[V_2]) & \cdots & (V_N - \mathrm{E}[V_N])^2 \end{bmatrix}.$$

Por lo tanto, teniendo en cuenta que $\mathrm{E}[(V_i - \mathrm{E}[V_i])^2] = \mathrm{Var}[V_i]$ $(i = 1, 2, ..., N)$, y que $\mathrm{E}[(V_i - \mathrm{E}[V_i])(V_j - \mathrm{E}[V_j])] = \mathrm{Cov}[V_i, V_j]$ $(i, j = 1, 2, ..., N; i \neq j)$, aplicando el operador $\mathrm{E}[\cdot]$ a la matriz $(\mathbf{V}-\mathrm{E}[\mathbf{V}])(\mathbf{V}-\mathrm{E}[\mathbf{V}])'$ anterior resulta que

$$\mathrm{E}[(\mathbf{V}-\mathrm{E}[\mathbf{V}])(\mathbf{V}-\mathrm{E}[\mathbf{V}])'] = \begin{bmatrix} \mathrm{Var}[V_1] & \mathrm{Cov}[V_1, V_2] & \cdots & \mathrm{Cov}[V_1, V_N] \\ \mathrm{Cov}[V_2, V_1] & \mathrm{Var}[V_2] & \cdots & \mathrm{Cov}[V_2, V_N] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[V_N, V_1] & \mathrm{Cov}[V_N, V_2] & \cdots & \mathrm{Var}[V_N] \end{bmatrix},$$

que se denomina la **matriz de varianzas-covarianzas** (o la matriz de varianzas, o la matriz de covarianzas) incondicionales de $\mathbf{V}$:

$$\mathrm{Var}[\mathbf{V}] = \mathrm{E}[(\mathbf{V}-\mathrm{E}[\mathbf{V}])(\mathbf{V}-\mathrm{E}[\mathbf{V}])']. \qquad \text{[A2.52]}$$

La matriz $\mathrm{Var}[\mathbf{V}]$ es **cuadrada** $(N \times N)$ y **simétrica** $(\mathrm{Var}[\mathbf{V}]' = \mathrm{Var}[\mathbf{V}])$; en su diagonal principal figuran las varianzas de los componentes individuales de $\mathbf{V}$; los elementos fuera de su diagonal principal son las covarianzas entre cada par de componentes de $\mathbf{V}$:

$$\mathrm{Var}[\mathbf{V}]_{ii} = \mathrm{Var}[V_i] \text{ para } i = 1, 2, ..., N,$$

$$\mathrm{Var}[\mathbf{V}]_{ij} = \mathrm{Cov}[V_i, V_j] = \mathrm{Cov}[V_j, V_i] = \mathrm{Var}[\mathbf{V}]_{ji} \text{ para } i, j = 1, 2, ..., N \text{ con } i \neq j.$$

Análogamente, la matriz de varianzas-covarianzas condicionales de $\mathbf{V}$ es

$$\mathrm{Var}[\mathbf{V} \mid \mathbf{X}] = \mathrm{E}[(\mathbf{V}-\mathrm{E}[\mathbf{V} \mid \mathbf{X}])(\mathbf{V}-\mathrm{E}[\mathbf{V} \mid \mathbf{X}])' \mid \mathbf{X}]. \qquad \text{[A2.53]}$$

En los enunciados siguientes, los elementos de $\mathbf{V}$ $(N \times 1)$ y $\mathbf{W}$ $(N \times M)$ son variables aleatorias, y los de $\mathbf{b}$ $(P \times 1)$, $\mathbf{c}$ $(N \times 1)$, $\mathbf{A}$ $(P \times N)$ y $\mathbf{B}$ $(M \times Q)$ son números reales (constantes). Todos son directamente generalizables al caso de distribuciones condicionales.

Esperanza de una Combinación Lineal:

$$\mathrm{E}\left[\textstyle\sum_{i=1}^{N} c_i V_i \right] = \mathrm{E}[\mathbf{c}'\mathbf{V}] = \textstyle\sum_{i=1}^{N} c_i \mathrm{E}[V_i] = \mathbf{c}'\mathrm{E}[\mathbf{V}]. \qquad \text{[A2.54]}$$

Varianza de una Combinación Lineal:

$$\operatorname{Var}\left[\sum_{i=1}^{N} c_i V_i\right] = \operatorname{Var}[\mathbf{c}'\mathbf{V}] = \mathbf{c}'\operatorname{Var}[\mathbf{V}]\mathbf{c}. \tag{A2.55}$$

Esperanza de un Vector de Combinaciones Lineales:

$$\operatorname{E}[\mathbf{AV} + \mathbf{b}] = \mathbf{A}\operatorname{E}[\mathbf{V}] + \mathbf{b}. \tag{A2.56}$$

Esperanza de una matriz de Combinaciones Lineales:

$$\operatorname{E}[\mathbf{AW}] = \mathbf{A}\operatorname{E}[\mathbf{W}],\;\; \operatorname{E}[\mathbf{WB}] = \operatorname{E}[\mathbf{W}]\mathbf{B},\;\; \operatorname{E}[\mathbf{AWB}] = \mathbf{A}\operatorname{E}[\mathbf{W}]\mathbf{B}. \tag{A2.57}$$

Matriz de Varianzas de un Vector de Combinaciones Lineales:

$$\operatorname{Var}[\mathbf{AV} + \mathbf{b}] = \mathbf{A}\operatorname{Var}[\mathbf{V}]\mathbf{A}'. \tag{A2.58}$$

Demostración:

$$\begin{aligned}
\operatorname{Var}[\mathbf{AV}+\mathbf{b}] &= \operatorname{E}[(\mathbf{AV}+\mathbf{b}-\operatorname{E}[\mathbf{AV}+\mathbf{b}])(\mathbf{AV}+\mathbf{b}-\operatorname{E}[\mathbf{AV}+\mathbf{b}])'] \\
&= \operatorname{E}[(\mathbf{AV}+\mathbf{b}-\mathbf{A}\operatorname{E}[\mathbf{V}]-\mathbf{b})(\mathbf{AV}+\mathbf{b}-\mathbf{A}\operatorname{E}[\mathbf{V}]-\mathbf{b})'] \\
&= \operatorname{E}[(\mathbf{AV}-\mathbf{A}\operatorname{E}[\mathbf{V}])(\mathbf{AV}-\mathbf{A}\operatorname{E}[\mathbf{V}])'] \\
&= \operatorname{E}[\mathbf{A}(\mathbf{V}-\operatorname{E}[\mathbf{V}])(\mathbf{V}-\operatorname{E}[\mathbf{V}])'\mathbf{A}'] \\
&= \mathbf{A}\operatorname{E}[(\mathbf{V}-\operatorname{E}[\mathbf{V}])(\mathbf{V}-\operatorname{E}[\mathbf{V}])']\mathbf{A}' = \mathbf{A}\operatorname{Var}[\mathbf{V}]\mathbf{A}'.
\end{aligned}$$

■

Matriz de Correlaciones:

$$\begin{aligned}
\operatorname{Corr}[\mathbf{V}] &= \begin{bmatrix} 1 & \operatorname{Corr}[V_1, V_2] & \cdots & \operatorname{Corr}[V_1, V_N] \\ \operatorname{Corr}[V_2, V_1] & 1 & \cdots & \operatorname{Corr}[V_2, V_N] \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{Corr}[V_N, V_1] & \operatorname{Corr}[V_N, V_2] & \cdots & 1 \end{bmatrix} \\
&= \mathbf{D}^{-1}\operatorname{Var}[\mathbf{V}]\mathbf{D}^{-1},
\end{aligned} \tag{A2.59}$$

donde

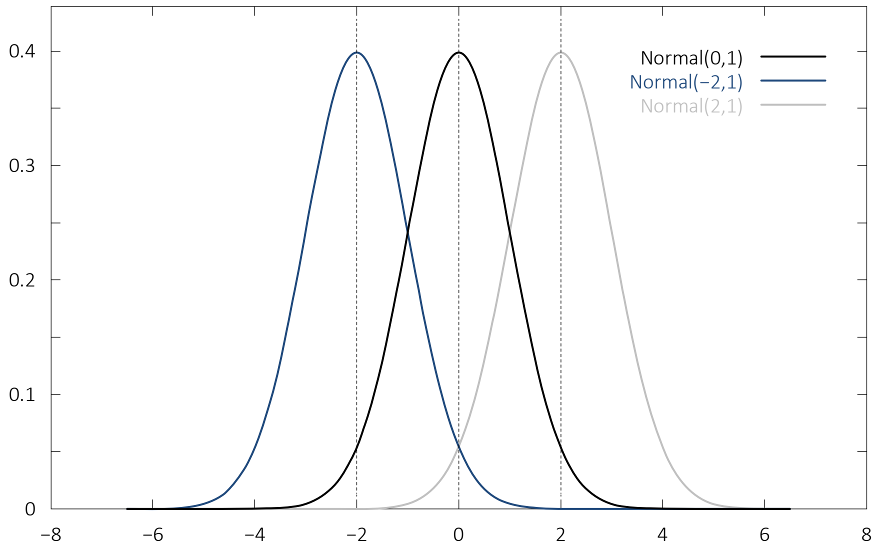
$$\mathbf{D} = \begin{bmatrix} \operatorname{Dvt}[V_1] & 0 & \cdots & 0 \\ 0 & \operatorname{Dvt}[V_2] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \operatorname{Dvt}[V_N] \end{bmatrix}. \tag{A2.60}$$

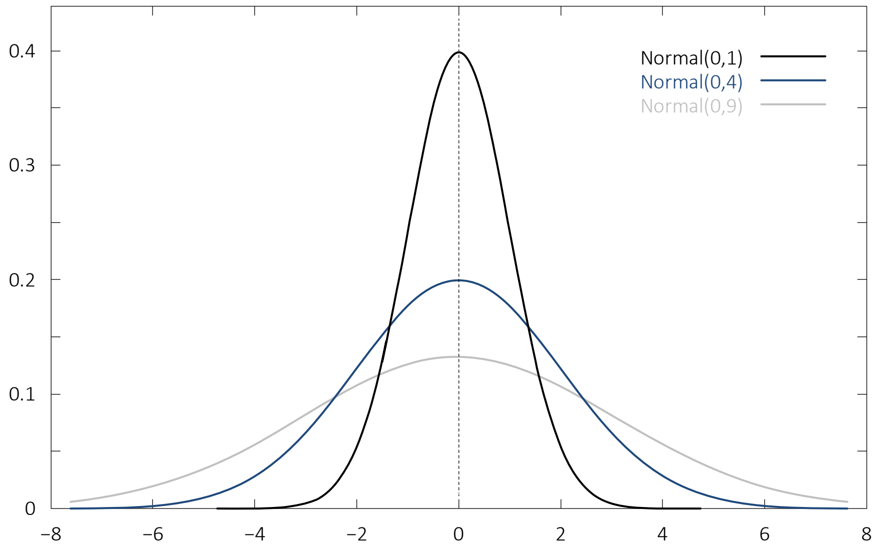
A2.4 DISTRIBUCIONES DERIVADAS DE LA DISTRIBUCIÓN NORMAL

Un vector $\mathbf{V} = [V_1, V_2, ..., V_N]'$ de variables aleatorias sigue una distribución Normal o Gaussiana multivariante, con vector de medias $\operatorname{E}[\mathbf{V}] = \boldsymbol{\mu}$ y matriz de varianzas-covarianzas $\operatorname{Var}[\mathbf{V}] = \boldsymbol{\Sigma}$, sí y sólo sí su función de densidad (PDF) es

$$f(\mathbf{v}) = (2\pi)^{-N/2}\,|\boldsymbol{\Sigma}|^{-1/2}\exp\left[-\tfrac{1}{2}(\mathbf{v}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{v}-\boldsymbol{\mu})\right],$$

lo que suele representarse como $\mathbf{V} \sim \operatorname{Normal}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.





DN.1 Transformaciones Lineales

Si $\mathbf{V} \sim \text{Normal}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces $\mathbf{Z} = \mathbf{QV} + \mathbf{q} \sim \text{Normal}(\mathbf{Q}\boldsymbol{\mu} + \mathbf{q}, \mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}')$, con $\mathbf{V}$ (variables aleatorias) $N \times 1$, $\mathbf{Q}$ (constantes) $M \times N$, $\mathbf{q}$ (constantes) $M \times 1$.

DN.2 Transformaciones Cuadráticas I: Distribución Chi-Cuadrado I

Si $\mathbf{V} \sim \text{Normal}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces $Z = (\mathbf{V} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{V} - \boldsymbol{\mu}) \sim \chi^2(N)$, con $\mathbf{V}$ (variables aleatorias) $N \times 1$.

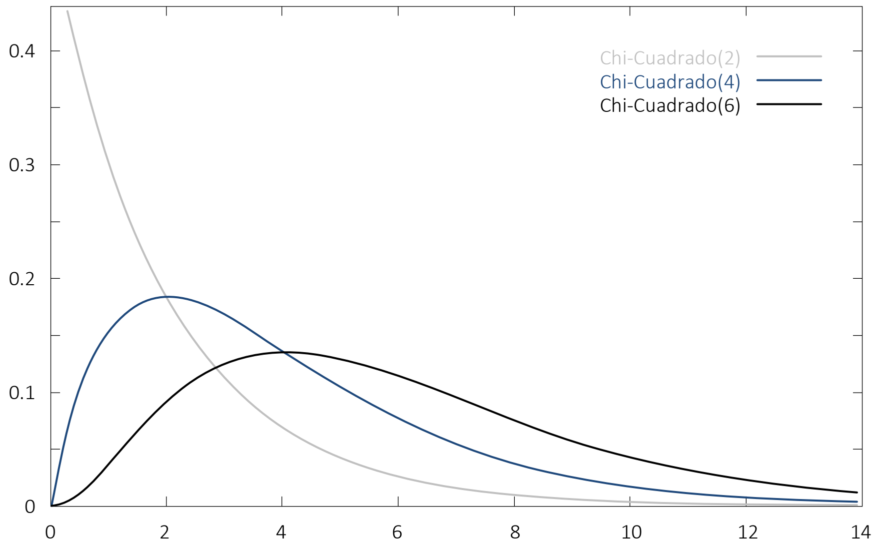
DN.3 Transformaciones Cuadráticas II: Distribución Chi-Cuadrado II

Si $\mathbf{V} \sim \text{Normal}(\mathbf{0}, \mathbf{I})$, entonces $Z = \mathbf{V}'\mathbf{PV} \sim \chi^2(P)$, con $\mathbf{V}$ (variables aleatorias) $N \times 1$, $\mathbf{P}$ (constantes) $N \times N$ simétrica e idempotente con rango$(\mathbf{P}) = P$ $(1 \leq P \leq N)$.

Observación: Una matriz $\mathbf{P}$ de orden $N \times N$ (cuadrada) es **simétrica** cuando $\mathbf{P}' = \mathbf{P}$, y es **idempotente** cuando $\mathbf{PP} = \mathbf{P}$. En cualquier matriz simétrica e idempotente ocurre que

$$\text{rango}(\mathbf{P}) = \text{traza}(\mathbf{P}),$$

donde traza$(\mathbf{P})$ es la **traza** de la matriz $\mathbf{P}$ (la suma de los elementos en su diagonal principal). La relación



anterior se debe a que (i) el rango de cualquier matriz simétrica coincide con su número de **autovalores** distintos de cero, (ii) la traza de cualquier matriz simétrica coincide con la suma de sus autovalores, y (ii) cada autovalor de una matriz simétrica e idempotente es igual a uno o a cero.

En general, en relación con la traza de una matriz $\mathbf{A}$ de orden $M \times M$ (cuadrada), $\text{traza}(\mathbf{A}) = \sum_{i=1}^{M} a_{ii}$, es inmediato comprobar que

$$\text{traza}(\mathbf{A}') = \text{traza}(\mathbf{A}),$$

$$\text{traza}(b\mathbf{A}) = b \times \text{traza}(\mathbf{A}) \text{ para cualquier número } b.$$

Adicionalmente, si $\mathbf{A}$ y $\mathbf{B}$ son matrices de orden orden $M \times M$ (cuadradas), entonces

$$\text{traza}(\mathbf{A} \pm \mathbf{B}) = \text{traza}(\mathbf{A}) \pm \text{traza}(\mathbf{B}).$$

Por último, si $\mathbf{A}$ y $\mathbf{B}$ son matrices de orden orden $M \times N$ y $N \times M$, respectivamente, entonces

$$\text{traza}(\mathbf{AB}) = \text{traza}(\mathbf{BA}).$$

DN.4 Independencia entre Transformaciones Lineales y Cuadráticas

Si $\mathbf{V} \sim \text{Normal}(\mathbf{0}, \mathbf{I})$, entonces $\mathbf{Z}_1 = \mathbf{QV}$ y $Z_2 = \mathbf{V}'\mathbf{PV}$ son independientes sí y sólo sí $\mathbf{QP} = \mathbf{0}$, con $\mathbf{V}$ (variables aleatorias) $N \times 1$, $\mathbf{Q}$ (constantes) $M \times N$, $\mathbf{P}$ (constantes) $N \times N$ simétrica e idempotente.

DN.5 Independencia entre Transformaciones Cuadráticas

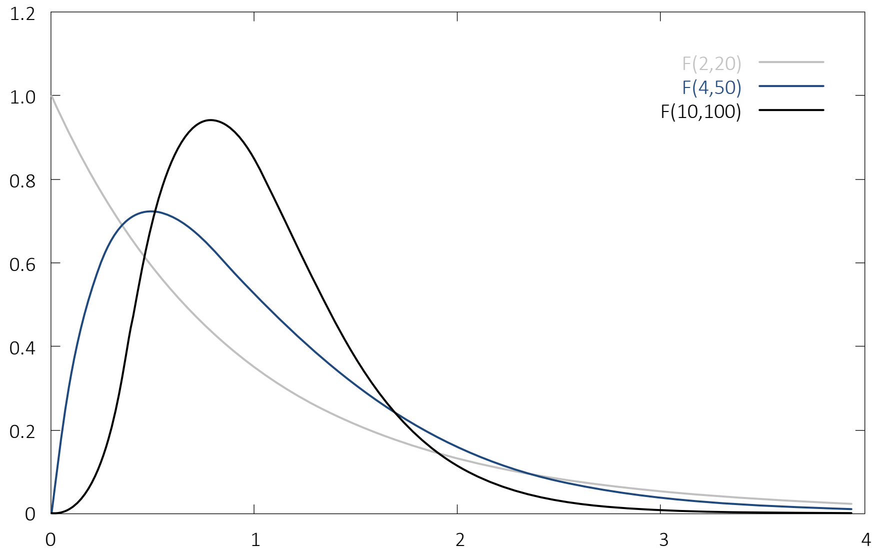
Si $\mathbf{V} \sim \text{Normal}(\mathbf{0}, \mathbf{I})$, entonces $Z_1 = \mathbf{V}'\mathbf{PV}$ y $Z_2 = \mathbf{V}'\mathbf{QV}$ son independientes sí y sólo sí $\mathbf{PQ} = \mathbf{0}$, con $\mathbf{V}$ (variables aleatorias) $N \times 1$, $\mathbf{P}$ y $\mathbf{Q}$ (constantes) $N \times N$ simétricas e idempotentes.

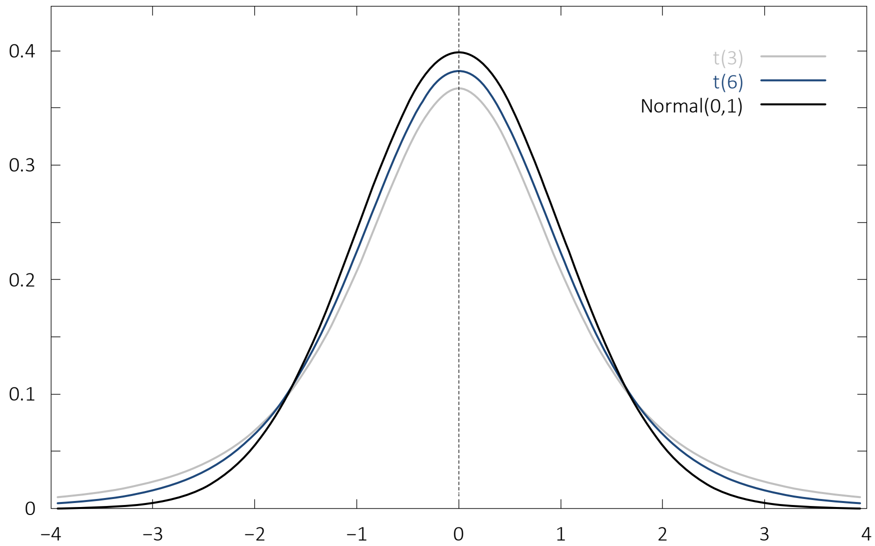
DN.6 Si $Z_1 \sim \chi^2(G_1)$ y $Z_2 \sim \chi^2(G_2)$ son independientes, entonces

Distribución F
$$F = \frac{Z_1/G_1}{Z_2/G_2} = \frac{G_2}{G_1} \times \frac{Z_1}{Z_2} \sim F(G_1, G_2).$$

DN.7 Si $W_1 \sim \text{Normal}(0,1)$ y $Z_2 \sim \chi^2(G)$ son independientes, entonces

Distribución t de Student
$$T = \frac{W_1}{(Z_2/G)^{½}} \sim t(G).$$





DN.8 Si $F_* \sim F(1,G)$, entonces $T_* = F_*^{1/2} \sim t(G)$.

Demostración: De acuerdo con DN.6,

$$F_* \sim F(1,G) \Rightarrow F_* = \frac{Z_1/1}{Z_2/G} = \frac{Z_1}{Z_2/G},$$

con $Z_1 \sim \chi^2(1)$ y $Z_2 \sim \chi^2(G)$ independientes. Por otro lado, de acuerdo con DN.2,

$$Z_1 \sim \chi^2(1) \Rightarrow Z_1 = W_1^2,$$

con $W_1 \sim \text{Normal}(0,1)$ independiente de $Z_2 \sim \chi^2(G)$. En consecuencia, por DN.7,

$$F_* = \frac{W_1^2}{Z_2/G} \Rightarrow T_* = F_*^{1/2} = \frac{W_1}{(Z_2/G)^{1/2}} \sim t(G).$$ ■

Los resultados considerados en este Apéndice A2.4 son de particular relevancia para la obtención de algunas propiedades estadísticas del criterio MCO en el contexto del Análisis de Regresión Lineal, y para el uso de modelos RLM como herramientas para contrastar hipótesis y para calcular previsiones. En relación con estas dos posibilidades, las tablas que se ofrecen a continuación aún se utilizan en algunos análisis econométricos aplicados.

Tabla A2.1: Valores de la Función de Distribución (CDF) Normal Estándar

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.0	0.001350	0.001306	0.001264	0.001223	0.001183	0.001144	0.001107	0.001070	0.001035	0.001001
-2.9	0.001866	0.001807	0.001750	0.001695	0.001641	0.001589	0.001538	0.001489	0.001441	0.001395
-2.8	0.002555	0.002477	0.002401	0.002327	0.002256	0.002186	0.002118	0.002052	0.001988	0.001926
-2.7	0.003467	0.003364	0.003264	0.003167	0.003072	0.002980	0.002890	0.002803	0.002718	0.002635
-2.6	0.004661	0.004527	0.004396	0.004269	0.004145	0.004025	0.003907	0.003793	0.003681	0.003573
-2.5	0.006210	0.006037	0.005868	0.005703	0.005543	0.005386	0.005234	0.005085	0.004940	0.004799
-2.4	0.008198	0.007976	0.007760	0.007549	0.007344	0.007143	0.006947	0.006756	0.006569	0.006387
-2.3	0.010724	0.010444	0.010170	0.009903	0.009642	0.009387	0.009137	0.008894	0.008656	0.008424
-2.2	0.013903	0.013553	0.013209	0.012874	0.012545	0.012224	0.011911	0.011604	0.011304	0.011011
-2.1	0.017864	0.017429	0.017003	0.016586	0.016177	0.015778	0.015386	0.015003	0.014629	0.014262
-2.0	0.022750	0.022216	0.021692	0.021178	0.020675	0.020182	0.019699	0.019226	0.018763	0.018309
-1.9	0.028717	0.028067	0.027429	0.026803	0.026190	0.025588	0.024998	0.024419	0.023852	0.023295
-1.8	0.035930	0.035148	0.034380	0.033625	0.032884	0.032157	0.031443	0.030742	0.030054	0.029379
-1.7	0.044565	0.043633	0.042716	0.041815	0.040930	0.040059	0.039204	0.038364	0.037538	0.036727
-1.6	0.054799	0.053699	0.052616	0.051551	0.050503	0.049471	0.048457	0.047460	0.046479	0.045514
-1.5	0.066807	0.065522	0.064255	0.063008	0.061780	0.060571	0.059380	0.058208	0.057053	0.055917
-1.4	0.080757	0.079270	0.077804	0.076359	0.074934	0.073529	0.072145	0.070781	0.069437	0.068112
-1.3	0.096800	0.095098	0.093418	0.091759	0.090123	0.088508	0.086915	0.085343	0.083793	0.082264
-1.2	0.115070	0.113139	0.111232	0.109349	0.107488	0.105650	0.103835	0.102042	0.100273	0.098525
-1.1	0.135666	0.133500	0.131357	0.129238	0.127143	0.125072	0.123024	0.121000	0.119000	0.117023
-1.0	0.158655	0.156248	0.153864	0.151505	0.149170	0.146859	0.144572	0.142310	0.140071	0.137857
-0.9	0.184060	0.181411	0.178786	0.176186	0.173609	0.171056	0.168528	0.166023	0.163543	0.161087
-0.8	0.211855	0.208970	0.206108	0.203269	0.200454	0.197663	0.194895	0.192150	0.189430	0.186733
-0.7	0.241964	0.238852	0.235762	0.232695	0.229650	0.226627	0.223627	0.220650	0.217695	0.214764
-0.6	0.274253	0.270931	0.267629	0.264347	0.261086	0.257846	0.254627	0.251429	0.248252	0.245097
-0.5	0.308538	0.305026	0.301532	0.298056	0.294599	0.291160	0.287740	0.284339	0.280957	0.277595
-0.4	0.344578	0.340903	0.337243	0.333598	0.329969	0.326355	0.322758	0.319178	0.315614	0.312067
-0.3	0.382089	0.378280	0.374484	0.370700	0.366928	0.363169	0.359424	0.355691	0.351973	0.348268
-0.2	0.420740	0.416834	0.412936	0.409046	0.405165	0.401294	0.397432	0.393580	0.389739	0.385908
-0.1	0.460172	0.456205	0.452242	0.448283	0.444330	0.440382	0.436441	0.432505	0.428576	0.424655
-0.0	0.500000	0.496011	0.492022	0.488034	0.484047	0.480061	0.476078	0.472097	0.468119	0.464144
0.0	0.500000	0.503989	0.507978	0.511966	0.515953	0.519939	0.523922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555670	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621720	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802337	0.805105	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823814	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.879000	0.881000	0.882977
1.2	0.884930	0.886861	0.888768	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903200	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935745	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959070	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965620	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999

Ejemplos: Si $X \sim \text{Normal}(0,1)$, entonces $\Pr(X \leq -1.96) = 0.024998$ y $\Pr(X \leq 2.00) = 0.977250$.

Fuente: Tabla elaborada con la función @CNORM del programa EViews.

Tabla A2.2: Valores Críticos de la Distribución Chi-Cuadrado

	CDF =	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
	1	0.000039	0.000157	0.000982	0.003932	0.0158	2.71	3.84	5.02	6.63	7.88
	2	0.0100	0.0201	0.0506	0.10	0.21	4.61	5.99	7.38	9.21	10.60
	3	0.0717	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34	12.84
	4	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28	14.86
	5	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09	16.75
	6	0.68	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
	7	0.99	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
G	8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.95
R	9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
A	10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
D											
O	11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72	26.76
S	12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
	13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
D	14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
E	15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
	16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
L	17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41	35.72
I	18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
B	19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58
E	20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
R											
T	21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
A	22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
D	23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
	24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
	25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
	26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
	27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
	28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
	29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
	30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67

Ejemplos: Si $X \sim \chi^2(20)$, entonces $\Pr(X \leq 9.59) = 0.025$ y $\Pr(X \leq 34.17) = 0.975$.

Fuente: Tabla elaborada con la función @QCHISQ del programa EViews.

Tabla A2.3: Valores Críticos al 10% (CDF = 0.90) de la Distribución F

		GRADOS DE LIBERTAD DEL NUMERADOR									
		1	2	3	4	5	6	7	8	9	10
	1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19
	2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39
	3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23
G	4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92
R	5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30
A	6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94
D	7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70
O	8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54
S	9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42
	10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32
D											
E	11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25
	12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19
L	13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14
I	14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10
B	15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06
E	16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03
R	17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00
T	18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98
A	19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96
D	20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94
D	21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92
E	22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90
L	23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89
	24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88
D	25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87
E	26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86
N	27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85
O	28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84
M	29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83
I	30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82
N											
A	40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76
D	60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71
O	90	2.76	2.36	2.15	2.01	1.91	1.84	1.78	1.74	1.70	1.67
R	120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65
	INF	2.71	2.30	2.08	1.95	1.85	1.77	1.72	1.67	1.63	1.60

Ejemplos: Si $X \sim F(2, 20)$, entonces $\Pr(X \leq 2.59) = 0.90$. Si $X \sim F(3, n)$ con n suficientemente grande, entonces $\Pr(X \leq 2.08) \approx 0.90$.

Fuente: Tabla elaborada con la función @QFDIST del programa EViews.

Tabla A2.4: Valores Críticos al 5% (CDF = 0.95) de la Distribución F

		GRADOS DE LIBERTAD DEL NUMERADOR									
		1	2	3	4	5	6	7	8	9	10
	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
G	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
R	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
A	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
D	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
O	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
S	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
D											
E	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
L	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
I	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
B	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
E	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
R	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
T	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
A	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
D	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
D	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
E	22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
L	23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
D	25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
E	26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
N	27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
O	28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
M	29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
I	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
N											
A	40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
D	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
O	90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
R	120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91
	INF	3.84	3.00	2.61	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Ejemplos: Si $X \sim F(2,20)$, entonces $\Pr(X \leq 3.49) = 0.95$. Si $X \sim F(3,n)$ con n suficientemente grande, entonces $\Pr(X \leq 2.61) \approx 0.95$.

Fuente: Tabla elaborada con la función @QFDIST del programa EViews.

Tabla A2.5: Valores Críticos al 1% (CDF = 0.99) de la Distribución F

		GRADOS DE LIBERTAD DEL NUMERADOR									
		1	2	3	4	5	6	7	8	9	10
	1	4052	5000	5403	5625	5764	5859	5928	5981	6022	6056
	2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40
	3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23
G	4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55
R	5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05
A	6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
D	7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
O	8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
S	9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26
	10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
D											
E	11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
	12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
L	13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
I	14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94
B	15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80
E	16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
R	17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59
T	18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
A	19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
D	20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
D	21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31
E	22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
L	23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21
	24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17
D	25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13
E	26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09
N	27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06
O	28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03
M	29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00
I	30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
N											
A	40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
D	60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
O	90	6.93	4.85	4.01	3.53	3.23	3.01	2.84	2.72	2.61	2.52
R	120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47
	INF	6.64	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32

Ejemplos: Si $X \sim F(2,20)$, entonces $\Pr(X \leq 5.85) = 0.99$. Si $X \sim F(3,n)$ con n suficientemente grande, entonces $\Pr(X \leq 3.78) \approx 0.99$.

Fuente: Tabla elaborada con la función @QFDIST del programa EViews.

Tabla A2.6: Valores Críticos de la Distribución t de Student

	CDF =	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
						NIVEL DE SIGNIFICACIÓN					
	$H_1 \neq$	-	-	-	-	-	0.200	0.100	0.050	0.020	0.010
	$H_1 >$	-	-	-	-	-	0.100	0.050	0.025	0.010	0.005
	$H_1 <$	0.005	0.010	0.025	0.050	0.100	-	-	-	-	-
	1	-63.657	-31.821	-12.706	-6.314	-3.078	3.078	6.314	12.706	31.821	63.657
	2	-9.925	-6.965	-4.303	-2.920	-1.886	1.886	2.920	4.303	6.965	9.925
	3	-5.841	-4.541	-3.182	-2.353	-1.638	1.638	2.353	3.182	4.541	5.841
	4	-4.604	-3.747	-2.776	-2.132	-1.533	1.533	2.132	2.776	3.747	4.604
	5	-4.032	-3.365	-2.571	-2.015	-1.476	1.476	2.015	2.571	3.365	4.032
	6	-3.707	-3.143	-2.447	-1.943	-1.440	1.440	1.943	2.447	3.143	3.707
	7	-3.499	-2.998	-2.365	-1.895	-1.415	1.415	1.895	2.365	2.998	3.499
	8	-3.355	-2.896	-2.306	-1.860	-1.397	1.397	1.860	2.306	2.896	3.355
	9	-3.250	-2.821	-2.262	-1.833	-1.383	1.383	1.833	2.262	2.821	3.250
	10	-3.169	-2.764	-2.228	-1.812	-1.372	1.372	1.812	2.228	2.764	3.169
	11	-3.106	-2.718	-2.201	-1.796	-1.363	1.363	1.796	2.201	2.718	3.106
G	12	-3.055	-2.681	-2.179	-1.782	-1.356	1.356	1.782	2.179	2.681	3.055
R	13	-3.012	-2.650	-2.160	-1.771	-1.350	1.350	1.771	2.160	2.650	3.012
A	14	-2.977	-2.624	-2.145	-1.761	-1.345	1.345	1.761	2.145	2.624	2.977
D	15	-2.947	-2.602	-2.131	-1.753	-1.341	1.341	1.753	2.131	2.602	2.947
O	16	-2.921	-2.583	-2.120	-1.746	-1.337	1.337	1.746	2.120	2.583	2.921
S	17	-2.898	-2.567	-2.110	-1.740	-1.333	1.333	1.740	2.110	2.567	2.898
	18	-2.878	-2.552	-2.101	-1.734	-1.330	1.330	1.734	2.101	2.552	2.878
D	19	-2.861	-2.539	-2.093	-1.729	-1.328	1.328	1.729	2.093	2.539	2.861
E	20	-2.845	**-2.528**	-2.086	-1.725	-1.325	1.325	**1.725**	2.086	2.528	2.845
L	21	-2.831	-2.518	-2.080	-1.721	-1.323	1.323	1.721	2.080	2.518	2.831
I	22	-2.819	-2.508	-2.074	-1.717	-1.321	1.321	1.717	2.074	2.508	2.819
B	23	-2.807	-2.500	-2.069	-1.714	-1.319	1.319	1.714	2.069	2.500	2.807
E	24	-2.797	-2.492	-2.064	-1.711	-1.318	1.318	1.711	2.064	2.492	2.797
R	25	-2.787	-2.485	-2.060	-1.708	-1.316	1.316	1.708	2.060	2.485	2.787
T	26	-2.779	-2.479	-2.056	-1.706	-1.315	1.315	1.706	2.056	2.479	2.779
A	27	-2.771	-2.473	-2.052	-1.703	-1.314	1.314	1.703	2.052	2.473	2.771
D	28	-2.763	-2.467	-2.048	-1.701	-1.313	1.313	1.701	2.048	2.467	2.763
	29	-2.756	-2.462	-2.045	-1.699	-1.311	1.311	1.699	2.045	2.462	2.756
	30	-2.750	-2.457	-2.042	**-1.697**	-1.310	1.310	**1.697**	2.042	2.457	2.750
	40	-2.704	-2.423	-2.021	-1.684	-1.303	1.303	**1.684**	2.021	2.423	**2.704**
	60	-2.660	-2.390	-2.000	-1.671	-1.296	1.296	1.671	2.000	2.390	2.660
	90	-2.632	-2.368	-1.987	-1.662	-1.291	1.291	1.662	1.987	2.368	2.632
	120	-2.617	-2.358	-1.980	-1.658	-1.289	1.289	1.658	1.980	2.358	2.617
	INF	-2.576	-2.327	**-1.960**	-1.645	-1.282	1.282	1.645	1.960	**2.327**	2.576

Ejemplos: Si $X \sim t(20)$, entonces $\Pr(X \leq -2.528) = 0.01$ y $\Pr(X \leq 1.725) = 0.95$. Si $X \sim t(n)$ con n suficientemente grande, entonces $\Pr(X \leq -1.96) \approx 0.025$ y $\Pr(X \leq 2.327) \approx 0.99$. Los valores críticos bilaterales al 10% y al 1% con 40 grados de libertad son, respectivamente, 1.684 y 2.704. El valor crítico unilateral por la derecha [izquierda] al 5% con 30 grados de libertad es 1.697 [−1.697].

Fuente: Tabla elaborada con la función @QTDIST del programa EViews.