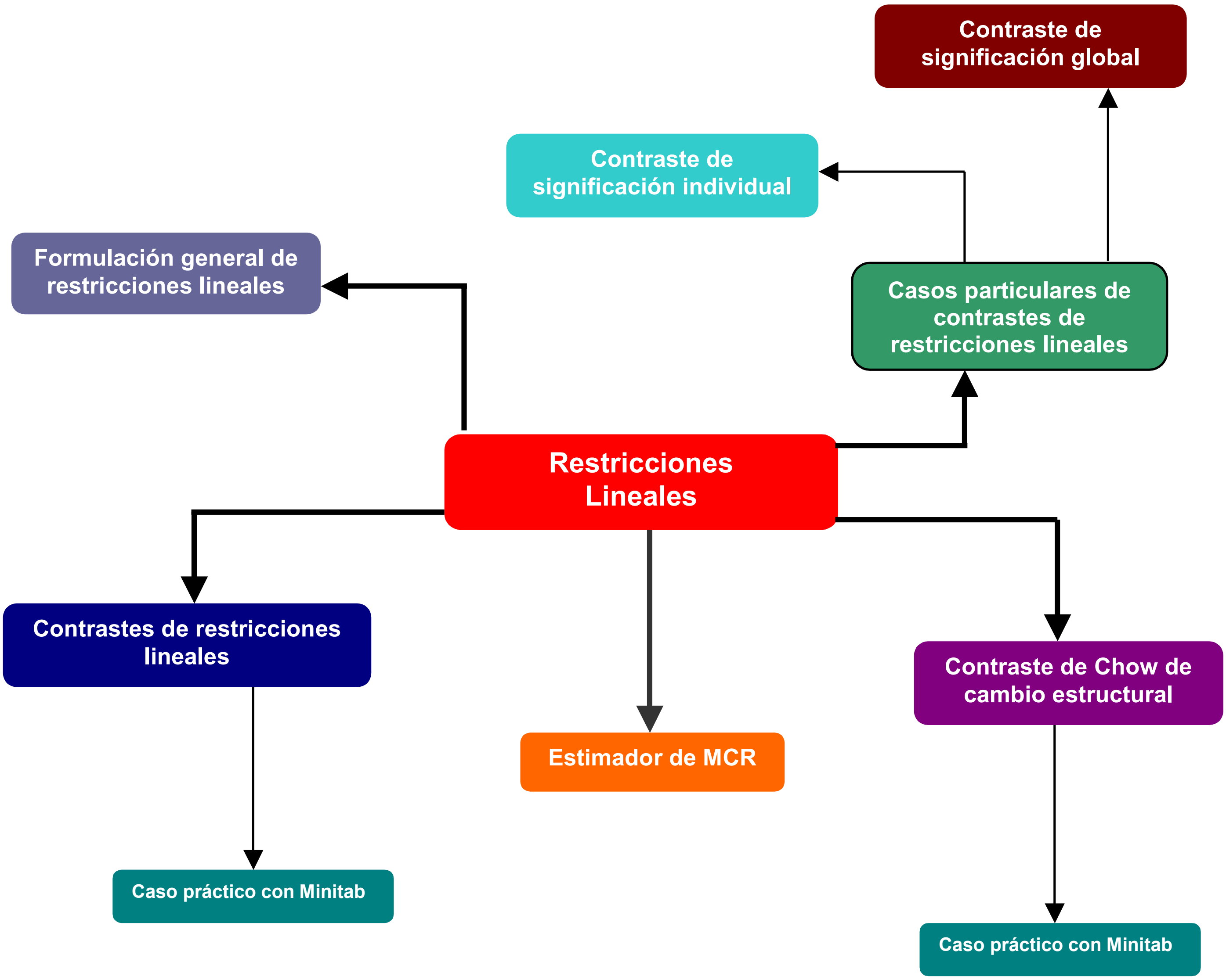
RESTRICCIONES LINEALES

Autores: Renatas Kizys (rkizys@uoc.edu), Ángel A. Juan (ajuanp@uoc.edu).

ESQUEMA DE CONTENIDOS



INTRODUCCIÓN

Es sabido que las contrastes de hipótesis constituyen una potente herramienta de la inferencia en el modelo lineal. El *math-block* sobre el MRLM considera los contrastes de significación individual y global, y los estadísticos de contraste correspondientes. No obstante, a menudo estamos interesados en contrastar determinados supuestos alternativos, postulados por la teoría económica. Dichos supuestos suelen ser mucho más complejos que la significación estadística de las variables. En este caso hemos de recurrir a la formulación general de las restricciones lineales.

El objetivo de este *math-block* es familiarizarse con un enfoque general de los contrastes de hipótesis. Nos centraremos únicamente en las hipótesis más habituales, que tratan las restricciones lineales. Además, analizaremos cual es la manera más eficiente de estimar los modelos bajo las restricciones lineales. Finalmente, estudiaremos los contrastes de cambio estructural.

OBJETIVOS

- Plantear y resolver correctamente los contrastes de hipótesis mediante las restricciones lineales expresadas en forma matricial.
- Obtener los estimadores de mínimos cuadrados restringidos (MCR) del modelo de regresión lineal múltiple (MRLM).
- Ser capaz de plantear y realizar el contraste de Chow de cambio estructural.
- Introducirse en el uso de Minitab para realizar los contrastes de hipótesis sobre las restricciones lineales.

CONOCIMIENTOS PREVIOS

Aparte de estar iniciado en el uso del paquete estadístico Minitab, resulta muy conveniente haber leído con profundidad los siguientes *math-blocks* relacionados con Estadística e Introducción a la Econometría:

- Intervalos de confianza y contraste de hipótesis para 1 y 2 poblaciones
- Modelo de Regresión Lineal Múltiple

CONCEPTOS FUNDAMENTALES

Formulación de las restricciones lineales

Sea un modelo de regresión lineal múltiple en forma matricial:

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{B} + \mathbf{U}$$

Consideremos ahora un conjunto de restricciones lineales de la forma [4]:

$$\begin{aligned}
&R_{11} \cdot \beta_1 + R_{12} \cdot \beta_2 + \ldots + R_{1k} \cdot \beta_k = r_1 \\
&R_{21} \cdot \beta_1 + R_{22} \cdot \beta_2 + \ldots + R_{2k} \cdot \beta_k = r_2 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&R_{q1} \cdot \beta_1 + R_{q2} \cdot \beta_2 + \ldots + R_{qk} \cdot \beta_k = r_q
\end{aligned}$$

que pueden resumirse en una única ecuación,

$$\mathbf{R} \cdot \mathbf{B} = \mathbf{r};$$

donde:

La matriz no aleatoria R, de dimensión q x k, donde q representa el número de restricciones lineales que queremos contrastar y k, el número de parámetros del modelo. Por tanto, la matriz R tiene tantas columnas cuantos parámetros tenga el modelo y tantas filas cuantas restricciones se hayan impuesto bajo la hipótesis nula. Además, R tiene rango de fila completo, de modo que $\rho(R) = q$ que deber ser menor o igual que k.

El vector no aleatorio r, de dimensión q x 1, que está formado por los términos independientes de las restricciones lineales. Así pues, al igual que la matriz R, el vector r tendrá tantas filas como restricciones.

Ejemplo de formulación de restricciones: consideremos el siguiente modelo de inversión [4]:

$$INV_i = \beta_1 + \beta_2 \cdot Y_i + \beta_3 \cdot I_i + \beta_4 \cdot \Pi_i + u_i, t = 1,\ldots,N.$$

que afirma que los inversores reaccionan al producto interno bruto, al tipo de interés y a la tasa de inflación, respectivamente. Una teoría alternativa afirma que *"los inversores se preocupan por el tipo de interés real".* Por lo tanto, el modelo alternativo sería:

$$INV_i = \beta_1 + \beta_2 \cdot Y_i + \beta_3 \cdot (I_i - \Pi_i) + \beta_4 \cdot \Pi_i + u_i, t = 1,\ldots,N.$$

Aunque este nuevo modelo tiene efectivamente en cuenta la teoría, la ecuación todavía incluye tanto la tasa de inflación como el tipo de interés nominal. La teoría no tiene una implicación contrastable para nuestro modelo. Ahora bien, consideremos la hipótesis más fuerte: *"los inversores sólo tienen en cuenta el tipo de interés real"*. La ecuación resultante, por tanto, es:

$$INV_i = \beta_1 + \beta_2 \cdot Y_i + \beta_3 \cdot (I_i - \Pi_i) + u_i, t = 1,\ldots,N$$

que ahora está restringida; pues en el contexto del primer modelo, la restricción es:

$\beta_3 + \beta_4 = 0$, o bien, $\beta_3 = -\beta_4$, o bien, en forma matricial, tenemos: $\mathbf{R \cdot B = r}$.

siendo R = (0 0 1 1) y r = (0).

El contraste de hipótesis en el modelo anterior puede llevarse a cabo desde dos puntos de vista diferentes:

- Una vez obtenidos los parámetros estimados, podemos preguntarnos si dichos parámetros pueden, razonablemente, satisfacer las restricciones que implican las hipótesis.

- El imponer las restricciones que presupone la teoría nos lleva a una pérdida de ajuste. Entonces, podemos averiguar si esta pérdida de ajuste se debe sencillamente a un error muestral, o a algo más profundo, como para tener dudas sobre la validez de las restricciones.

Todos los tipos de contrastes que estudian el comportamiento de determinadas restricciones lineales entre los parámetros del modelo, se conocen como *contrastes de restricciones lineales*.

Contrastes de restricciones lineales

Estadístico de contraste en forma matricial

Una vez especificadas las restricciones en forma de matrices, procedemos a la realización de los contrastes de hipótesis sobre los parámetros del modelo. A tales efectos, el procedimiento a seguir es el siguiente:

Las restricciones lineales se admiten como ciertas bajo la hipótesis nula, mientras que, bajo la alternativa, los parámetros de la población no satisfacen estas restricciones:

$$H_0 : \mathbf{R \cdot B = r}$$
$$H_A : \mathbf{R \cdot B \neq r}.$$

Una vez definida la hipótesis nula, el paso siguiente es obtener un estadístico de contraste que permite rechazarla o no rechazarla. Para ello, hemos de especificar el MRLM y estimarlo por MCO. Podéis comprobar que el estadístico de contraste para contrastar las restricciones lineales sobre los parámetros es el siguiente [1]:

$$F_0 = \frac{(R \cdot \hat{B} - r)' \cdot \left[R \cdot (X' \cdot X)^{-1} \cdot R'\right]^{-1} (R \cdot \hat{B} - r)}{e' \cdot e} \cdot \frac{n-k}{q} \sim F_{q;n-k}.$$

A partir del valor numérico que tome el estadístico de contraste es posible determinar si la diferencia entre R·B y r es estadísticamente significativa o no lo es. La regla de decisión es la siguiente:

- Si $F_0 \geq F_{q,n-k;\alpha}$, el estadístico de contraste se encuentra fuera de la región de aceptación, lo cual nos lleva a rechazar la hipótesis nula. Por tanto, las restricciones lineales no son ciertas en el ámbito de la población.

- Si $F_0 < F_{k-1,n-k;\alpha}$, el estadístico de contraste cae dentro de la región de aceptación, con lo cual no podemos rechazar la hipótesis nula. En consecuencia, podemos afirmar que las restricciones son ciertas en el ámbito de la población.

Ejemplo de contraste de restricciones lineales: Una cierta compañía de producción de juguetes de repente se ve amenazada por la entrada de las compañías competidoras. El directivo se encarga a investigar el comportamiento de la variable el volumen de las ventas mensuales mediante el siguiente modelo econométrico:

$$Y_t = \beta_1 + \beta_2 \cdot P_t + \beta_3 \cdot Z_t + u_t, t = 1,...,T;$$

siendo Y_t el volumen de las ventas mensuales (en millones de €), P_t el precio esperado de la competencia en € en el mes t y Z_t los gastos mensuales en publicidad expresados en millones de €. Supondremos que el modelo cumple las hipótesis del MRLM con errores normales. Con 60 observaciones, al estimar por MCO el modelo queda:

$$\hat{Y}_t = 4{,}5 + 0{,}34 \cdot P_t + 3{,}7 \cdot Z_t, t = 1,...,60.$$

Supongamos R^2 = 47,6%, y la siguiente estimación de la varianza de los parámetros estimados:

$$\hat{\sigma}^2 (X'X)^{-1} = \begin{pmatrix} 5{,}2 & 0 & 0 \\ 0 & 0{,}025 & -0{,}95 \\ 0 & -0{,}95 & 4{,}85 \end{pmatrix}.$$

Se trata de contrastar la hipótesis nula de que una disminución en el precio esperado de la competencia en un € y un aumento en 100,000 de € en publicidad no tienen, conjuntamente, ningún efecto sobre las ventas esperadas de juguetes frente a una alternativa bilateral. En primer lugar, especificamos la hipótesis nula

H_0: $\beta_2 = 0{,}1\cdot\beta_3$ frente a H_A: $\beta_2 \neq 0{,}1\cdot\beta_3$.

La restricción lineal bajo la hipótesis nula la podemos rescribir en forma matricial, es decir,

H_0: $\mathbf{R\cdot B = r}$ frente a H_A: $\mathbf{R\cdot B \neq r}$,

siendo R = (0 –1 0,1), B = (β_1 β_2 β_3)’ y r = 0.

A continuación calculamos estadístico de contraste:

$$F_0 = \frac{\left(R\cdot\hat{B}-r\right)'\left[R\cdot(X'\cdot X)^{-1}\cdot R'\right]^{-1}\left(R\cdot\hat{B}-r\right)}{e'\cdot e}\cdot\frac{n-k}{q} \sim F_{q;n-k}.$$

Sustituyendo los valores respectivos, queda F_0 = 0,0034. A partir de las tablas de una distribución F de Snedecor, tenemos que $F_{q;n-k;\alpha} = F_{1;57;0,05} = 4{,}0099$. Vemos claramente que $F_0 = 0{,}0034 \in$ R.A.{0; 4,0099}, con lo que se acepta la hipótesis nula para nivel de significación de 5%. Por tanto, una disminución en el precio esperado de la competencia en un € y un aumento en 100,000 de € en publicidad no tienen, conjuntamente, ningún efecto sobre las ventas esperadas de juguetes.

Estadístico de contraste en sumas cuadráticas residuales

Un método alternativo permite comparar la suma de cuadrados de los errores del modelo restringido (SCE_R o $e_R'\cdot e_R$), con la del modelo sin restringir, o ampliado (SCE_A o $e_A'\cdot e_A$). Podéis comprobar que el estadístico de contraste es el siguiente [1]:

$$F_0 = \frac{SCE_R - SCE_A}{SCE_A}\cdot\frac{n-k}{q} = \frac{SCE_R - SCE_A}{\hat{\sigma}^2}\cdot\frac{1}{q} \sim F_{q;n-k},$$

donde $\hat{\sigma}^2 = \frac{e'\cdot e}{n-k} = \frac{SCE_A}{n-k}$, la varianza estimada del término de error.

El modelo restringido es el que incorpora la hipótesis nula, es decir, el modelo resultante de considerar la restricción lineal como cierta, mientras que el modelo sin restringir es el modelo original, que no incorpora restricción lineal.

Entre los contrastes de restricciones lineales también se encuentran los contrastes de significación individual y de significación global.

Contraste de significación individual del modelo:

Supongamos que queremos contrastar la significación individual individual de un parámetro β_j del modelo.

En primer lugar, especificamos las hipótesis nula y alternativa:

$H_0 : \beta_j = 0$

$H_A : \beta_j \neq 0.$

Por consiguiente, el modelo sin restringir, o ampliado queda:

$$Y_i = \beta_1 + \beta_2 \cdot X_{2i} + \ldots + \beta_k \cdot X_{ki} + u_i ; i = \ldots., n.$$

Si incorporamos la hipótesis nula, el modelo resultante que denominaremos modelo restringido, es:

$$Y_i = \beta_1 + \beta_2 \cdot X_{2i} + \ldots + \beta_{j-1} \cdot X_{j-1,i} + + \beta_{j+1} \cdot X_{j+1,i} + \beta_k \cdot X_{ki} + u_i ; i = \ldots., n.$$

Una vez que hemos especificado el modelo restringido y sin restringir, calculamos las sumas de cuadrados de los errores, SCE_R y SCE_A, respectivamente, y construimos el estadístico de contraste teniendo en cuenta los grados de libertad [1]:

$$F_0 = \frac{SCE_R - SCE_A}{SCE_A} \cdot (n-k) = \frac{SCE_R - SCE_A}{\sigma^2} \sim F_{1;n-k}.$$

Finalmente, aplicamos la regla de decisión que permite rechazar o no la hipótesis nula.

Contraste de significación global del modelo:

Al igual que en el caso de significación global, primero especificamos las hipótesis nula y alternativa:

$H_0 : \beta_2 = \beta_3 = \ldots = \beta_k = 0$
H_A : No H_0.

El mismo estadístico construido a partir de las sumas residuales, puede utilizarse para contrastar también la significación global del modelo. El modelo de regresión sin restringir es:

$$Y_i = \beta_1 + \beta_2 \cdot X_{2i} + \ldots + \beta_k \cdot X_{ki} + u_i ; i = \ldots., n;$$

mientras que el modelo restringido, es decir, al incorporar la hipótesis nula, queda:

$$Y_i = \beta_1 + v_i ; i = \ldots., n.$$

Es fácil ver que el estimador de mínimos cuadrados restringidos no es sino la media muestral de la variable dependiente:

$$\hat{\beta}_1 = \hat{\beta}_{MCR} = \bar{Y};$$

siendo $\bar{Y}$ la media muestral de la variable dependiente. De modo que el residuo del modelo restringido viene determinado por:

$$e_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_i = Y_i - \bar{Y}.$$

Por tanto, la SCE del modelo restringido viene representado por:

$$SCE_R = e_R{}' e_R = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = SCT.$$

Puesto que el modelo tiene término independiente, la variabilidad total de la variable endógena puede descomponerse de la forma SCT = SCE + SCR, en cuyo caso la SCR es igual a cero. Planteamos el estadístico de contraste de significación global:

$$F_0 = \frac{SCE_R - SCE_A}{SCE_A} \cdot \frac{n-k}{k-1} = \frac{SCT - SCE_A}{SCE_A} \cdot \frac{n-k}{k-1}.$$

Por otro lado, la igualdad $SCR_A = SCT - SCE_A$ es cierta. Por tanto, la expresión del estadístico F_0 viene dada por [1]

$$F_0 = \frac{SCR_A}{SCE_A} \cdot \frac{n-k}{k-1} \sim F_{k-1;n-k}.$$

Estimador de mínimos cuadrados restringidos

Un aspecto básico de la inferencia estadística que se lleva a cabo en la economía de la empresa es que el directivo sólo contrasta hipótesis en cuya validez está dispuesto a creer a priori, de modo que si su contraste no las rechaza, entonces pasa a imponerlas en la representación estructural que está considerando.

Por tanto, sería interesante disponer de un procedimiento para estimar de nuevo el modelo, pero imponiendo el conjunto de hipótesis que hemos contrastado y no rechazado. La idea de eficiencia está ligada a la utilización óptima de toda la información disponible. Si se cree que los coeficientes de modelo satisfacen ciertas restricciones, entonces se ganaría eficiencia introduciendo dichas restricciones en el proceso de estimación.

El método de estimación de mínimos cuadrados restringidos (MCR) consiste en minimizar la suma cuadrática de los errores, pero teniendo en cuenta las restricciones lineales, es decir:

$$\text{Min}_B \ \mathbf{e}'\cdot\mathbf{e}$$

$$\text{sujeto a } \mathbf{R}\cdot\mathbf{B} = \mathbf{r}.$$

Al resolver el problema de optimización, obtenemos el estimador MCR de los parámetros B del modelo cuya expresión viene dada por [1]:

$$\hat{B}_{MCR} = \hat{B}_{MCO} + (X' \cdot X)^{-1} \cdot R' \cdot [R \cdot (X' \cdot X) \cdot R']^{-1} (r - R \cdot \hat{B}_{MCO}).$$

La interpretación de la expresión anterior es que el estimador MCR es una corrección del estimador MCO sin restringir. El tamaño de dicha corrección viene dado por el segundo sumando de la expresión. Así pues, si el estimador MCO satisface las restricciones lineales, la discrepancia $r - R \cdot \hat{B}_{MCO}$ es cero, de manera que el estimador MCR coincide con el estimador MCO.

Ejemplo de estimación de mínimos cuadrados restringidos: Supongamos que para presentar la tecnología existente en el sector aeronáutico de los Estados Unidos se especifica una función de producción Cobb-Douglas estocástica:

$$\log(Y_t) = \beta_1 + \beta_2 \cdot \log(L_t) + \beta_3 \cdot \log(K_t) + u_t, t = 1,\ldots,T;$$

siendo Y_t la producción (en millones de dólares), L_t es el nivel de empleo (que representaremos a través del agregado de las nóminas (en millones de dólares) y K_t es el nivel de capital utilizado (en millones de dólares). Supondremos que esta relación satisface las hipótesis el MRLM con normalidad del término de error. A partir de los datos anuales correspondientes a 1958-1996 que se disponen en el *math-block* relativo al MRLM, el modelo anterior se ha estimado obteniéndose:

$$\log(Y_t) = -1{,}61 + 0{,}591 \cdot \log(L_t) + 0{,}622 \cdot \log(K_t) + e_t, t = 1,\ldots,39.$$

La empresa conjetura que la industria aeronáutica presenta rendimientos constantes a escala, es decir que $\beta_2 + \beta_3 = 1$. Suponiendo cierta la hipótesis nula, procedemos a determinar el estimador de mínimos cuadrados restringidos (MCR). Bajo la hipótesis nula, la expresión del estimador MCR de los parámetros B del modelo viene dada por:

$$\hat{B}_{MCR} = \hat{B}_{MCO} + (X' \cdot X)^{-1} \cdot R' \cdot [R \cdot (X' \cdot X) \cdot R']^{-1} (r - R \cdot \hat{B}_{MCO}).$$

Especificamos R = (0 1 1) y r = 1 tal que **R·B = r**. Luego, estimar el modelo por mínimos cuadrados restringidos equivale a sustituir los valores respectivos en la expresión anterior. Primero, calculamos:

$$X' X^{-1} = \begin{pmatrix} 5{,}548 & -0{,}11 & -0{,}448 \\ -0{,}11 & 3{,}2 & -2{,}644 \\ -0{,}448 & -2{,}644 & 2{,}237 \end{pmatrix}$$

Sustituyendo, obtenemos:

$$B_{MCR} = \begin{pmatrix} -1{,}61 \\ 0{,}591 \\ 0{,}622 \end{pmatrix} + \begin{pmatrix} 5{,}548 & -0{,}11 & -0{,}448 \\ -0{,}11 & 3{,}2 & -2{,}644 \\ -0{,}448 & -2{,}644 & 2{,}237 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot 6{,}7114 \cdot (-0{,}213),$$

siendo $[R \cdot (X' \cdot X) \cdot R']^{-1} = 6{,}7114$ y $r - R \cdot \hat{B}_{MCO} = -0{,}213$. Así, obtenemos

$$B_{MCR} = \begin{pmatrix} -1{,}61 \\ 0{,}591 \\ 0{,}622 \end{pmatrix} + \begin{pmatrix} 0{,}798 \\ -0{,}795 \\ 0{,}582 \end{pmatrix} = \begin{pmatrix} -0{,}812 \\ -0{,}204 \\ 1{,}204 \end{pmatrix}.$$

Es interesante observar que la restricción $B_2 + B_3 = 1$ se respeta dentro del vector de parámetros estimados.

Propiedades del estimador MCR

A. El estimador MCR es insesgado sólo si las restricciones bajo las que se ha obtenido son ciertas, o sea si se cumple que R·B = r.

B. Una segunda propiedad es que el estimador MCR siempre satisfará las restricciones R·B = r. Eso es así tanto si las restricciones son ciertas como son falsas.

C. La matriz de covarianzas del estimador MCR es siempre inferior a la matriz de covarianzas del estimador MCO, incluso si las restricciones no son ciertas. Eso se debe al hecho de que al imponer las restricciones limitamos la región del espacio paramétrico en la que buscamos el estimador mínimo cuadrático, por lo que podremos estimarlo con una mayor precisión, aunque quizá con sesgo. Más formalmente [1],

$$Var(\hat{B}_{MCR}) = Var(\hat{B}_{MCO}) - \sigma^2 (X' \cdot X)^{-1} \cdot R' \cdot [R \cdot (X' \cdot X) \cdot R']^{-1} \cdot R \cdot (X' \cdot X)^{-1}.$$

En efecto, el segundo sumando de la expresión viene dado por una matriz semidefinida positiva.

Análisis estructural

Una de las hipótesis básicas del MRLM es la hipótesis de permanencia estructural de los parámetros poblacionales. La hipótesis significa que todas las submuestras han sido generadas por una misma estructura económica.

¿Qué quiere decir el incumplimiento de esta hipótesis? La respuesta depende del tipo de datos que tratamos:

- Si trabajamos con una serie temporal, el cambio estructural se produce cuando se tiene cierta información acerca de una variación estructural que ocurrió en algún momento durante el período muestral y que dicha variación fue suficientemente importante como para generar cambios en los coeficientes del modelo.
- Si trabajamos con un corte transversal, el cambio estructural implica que hay dos o más grupos de individuos caracterizados por comportamientos diferentes entre si.

Contraste de Chow de cambio estructural

Un contraste que tiene especial importancia por su interés, así como por la frecuencia con que aparece en aplicaciones empíricas, es el contraste de cambio estructural que se suele denominarse como test de Chow.

El procedimiento de test de Chow de consiste de los siguiente pasos:

A. Dividimos la muestra en dos submuestras y especificamos el modelo sin restringir para cada una de ellas:

$$Y_i = X_i' B^{(1)} + u_i, i = 1, \ldots, n_1;$$
$$Y_i = X_i' B^{(2)} + u_i, i = n_1 + 1, \ldots, n;$$

donde $n = n_1 + n_2$. Así pues, el modelo consta de una regresión diferente para cada una de las dos submuestras.

B. Ahora postulamos la hipótesis nula:

$$H_0 : \beta_1^{(1)} = \beta_1^{(2)}, \beta_2^{(1)} = \beta_2^{(2)}, \beta_3^{(1)} = \beta_3^{(2)}, \ldots, \beta_k^{(1)} = \beta_k^{(2)}$$

A fin de contrastar frente a la hipótesis alternativa:

$$H_A : \beta_1^{(1)} \neq \beta_1^{(2)} \text{ ó } \beta_2^{(1)} \neq \beta_2^{(2)} \text{ ó } \beta_3^{(1)} \neq \beta_3^{(2)} \text{ ó } \ldots \text{ ó } \beta_k^{(1)} \neq \beta_k^{(2)}.$$

C. Al imponer las restricciones bajo la hipótesis nula en el paso anterior sobre el modelo sin restringir, el modelo resulta ser:

$$Y_i = X_i' B + u_i, i = 1, \ldots, n_1, n_1 + 1, \ldots, n.$$

De modo que, el modelo restringido presupone que la estructura de relación entre la variable endógena y las variable explicativas permanece estable dentro de la muestra.

D. Por tanto, tenemos k restricciones lineales. Para contrastar la hipótesis nula, utilizaremos el estadístico F en función de las sumas cuadráticas residuales del modelo original (o ampliado) y del modelo restringido. La suma de cuadrados de los errores del modelo no restringido (o ampliado), SCE_A, es el agregado de las sumas de cuadrados de los errores de cada una de las dos regresiones:

$$SCE_A = SCE_{A1} + SCE_{A2}.$$

E. A continuación, calculamos el estadístico de contraste [8]:

$$F_0 = \frac{SCE_R - SCE_A}{SCE_A} \cdot \frac{n-2k}{q} = \frac{SCE_R - (SCE_{A1} + SCE_{A2})}{SCE_{A1} + SCE_{A2}} \frac{n-2k}{q} \sim F_{q;n-2k},$$

donde:

q representa el número de restricciones; si estamos contrastando el cambio estructural entre dos submuestras, entonces tenemos tantas restricciones cuantos parámetros a restringir, i.e. q = k.

$n = n_1 + n_2$ es el agregado del número de vacaciones en cada una de las submuestras.

n – 2k indica el número de grados de libertad del modelo sin restringir para el caso de 2 submuestras, ya que se estiman k parámetros con las primeras n_1 observaciones muestrales y otros k parámetros con las restantes $n_2 = n - n_1$. En total, n observaciones y 2k parámetros estimados.

$n = n_1 + n_2$ es el agregado del número de vacaciones en cada una de las submuestras.

F. Si $F_0 \geq F_{k,n-2k;\alpha}$, el estadístico de contraste es superior al valor crítico en las tablas de la F de Snedecor, lo cual nos lleva a rechazar la hipótesis nula. Concluimos, por tanto, que ha habido cambio estructural en el ámbito de la población.

G. Si $F_0 < F_{k,N-2k;\alpha}$, el estadístico de contraste se sitúa dentro de la región de aceptación; no podemos rechazar la hipótesis nula. En consecuencia, podemos afirmar que no ha producido cambio estructural en el ámbito de la población.

Ejemplo de contraste de Chow: Para examinar las diferencias de la relación entre renta (R_t) y consumo (C_t) entre los países europeos y países africanos un economista dispone de 22 datos; los 12 primeros datos corresponden a países europeos y los 10 últimos a países africanos. Para ello, el economista estima, mediante MCO, las siguientes ecuaciones de regresión:

$$C_t = \beta_1 + \beta_2 \cdot R_t + u_t, t = 1,...,22; (1)$$
$$C_t = \gamma_1 + \gamma_2 \cdot R_t + v_t, t = 1,...,12; (2)$$
$$C_t = \delta_1 + \delta_2 \cdot R_t + w_t, t = 13,...,22. (3)$$

Las sumas de cuadrados de los errores, obtenidas en cada uno de estos modelos, fueron, respectivamente, SCE_1 = 15,7, SCE_2 = 11,5, SCE_3 = 2,3. Los modelos (1), (2) y (3) cumplen las hipótesis del MRLM con normalidad, siendo v_l independiente de w_m para cualesquiera l, m. Se pide contrastar, tomando 0,05 como nivel de significación, la hipótesis nula de que $\gamma_1 = \delta_1$ y $\gamma_2 = \delta_2$ frente a la hipótesis alternativa: $\gamma_1 \neq \delta_1$ ó $\gamma_2 \neq \delta_2$. Para ello, utilizaremos el estadístico F en función de las sumas cuadráticas residuales del modelo original (o ampliado) y del modelo restringido. La suma de cuadrados de los errores del modelo no restringido (o ampliado), SCE_A, es el agregado de las sumas de cuadrados de los errores de cada una de las dos regresiones, $SCE_A = SCE_{A1} + SCE_{A2} = 11,5 + 2,3 = 13,8$. Haremos el uso de la siguiente expresión del estadístico de contraste:

$$F_0 = \frac{SCE_R - SCE_A}{SCE_A} \cdot \frac{n-2k}{q} = \frac{SCE_R - (SCE_{A1} + SCE_{A2})}{SCE_{A1} + SCE_{A2}} \frac{n-2k}{q} \sim F_{q;n-2k},$$

Sustituyendo los valores respectivos, obtenemos el siguiente resultado:

$$F_0 = \frac{15,7 - 13,8}{13,8} \cdot \frac{18}{2} = 1,239.$$

La tabla de distribución de F de Snedecor nos proporciona el valor crítico $F_{k;T-2k;\alpha} = F_{2,18,0,05} = 3,5546$. Vemos que $F_0 = 1,239 \in$ R.A. (0; 3,5546), con lo que aceptamos la hipótesis nula de permanencia estructural de los parámetros del modelo.

CASOS PRÁCTICOS CON SOFTWARE

Contraste de restricciones lineales

Ejemplo 1: Consideremos el caso práctico sobre la industria aeronáutica de los Estados Unidos que está disponible dentro del math-block "MRLM". Recordemos que el caso trataba de estimar la función de producción Cobb-Douglas estocástica aumentada por la variable el avance tecnológico:

$$\log(Y_t) = \beta_1 + \beta_2 \cdot \log(L_t) + \beta_3 \cdot \log(K_t) + \beta_4 \cdot \log(A_t) + u_t, t = 1,\dots,T;$$

donde Y_t es la producción (en millones de dólares), L_t es el nivel de empleo (que representaremos a través del agregado de las nóminas (en millones de dólares), K_t es el nivel de capital utilizado (en millones de dólares), y A_t es el avance tecnológico, representado por la proporción del PIB de las empresas tecnológicas en el PIB total en la economía americana (en tanto por ciento). Supondremos que esta relación satisface las hipótesis el MRLM con normalidad del término de error. Se dispone de datos anuales correspondientes a 1958-1996.

Supongamos una empresa de investigación de mercados se propone a contrastar la hipótesis nula de que la función de producción Cobb-Douglas presenta rendimientos constantes a escala. Los rendimientos constantes a escala se caracterizan por $\beta_2 + \beta_3 + \beta_4 = 1$. Por tanto, la hipótesis nula es H_0: $\beta_2 + \beta_3 + \beta_4 = 1$, mientras que la alternativa bilateral es H_A: $\beta_2 + \beta_3 + \beta_4 \neq 1$. A fin de contrastar la hipótesis nula, utilizaremos el estadístico F, cuya expresión es:

$$F_0 = \frac{SCE_R - SCE_A}{SCE_A} \cdot \frac{T-k}{q}.$$

En primer lugar, a fin de determinar el modelo restringido, hemos de incorporar la restricción, según la hipótesis nula, al modelo. Para ello, despejamos $\beta_4 = 1 - \beta_2 - \beta_3$ y sustituimos en la ecuación anterior:

$$\log(Y_t) = \beta_1 + \beta_2 \cdot \log(L_t) + \beta_3 \cdot \log(K_t) + (1 - \beta_2 - \beta_3) \cdot \log(A_t) + u_t, t = 1,\dots,T.$$

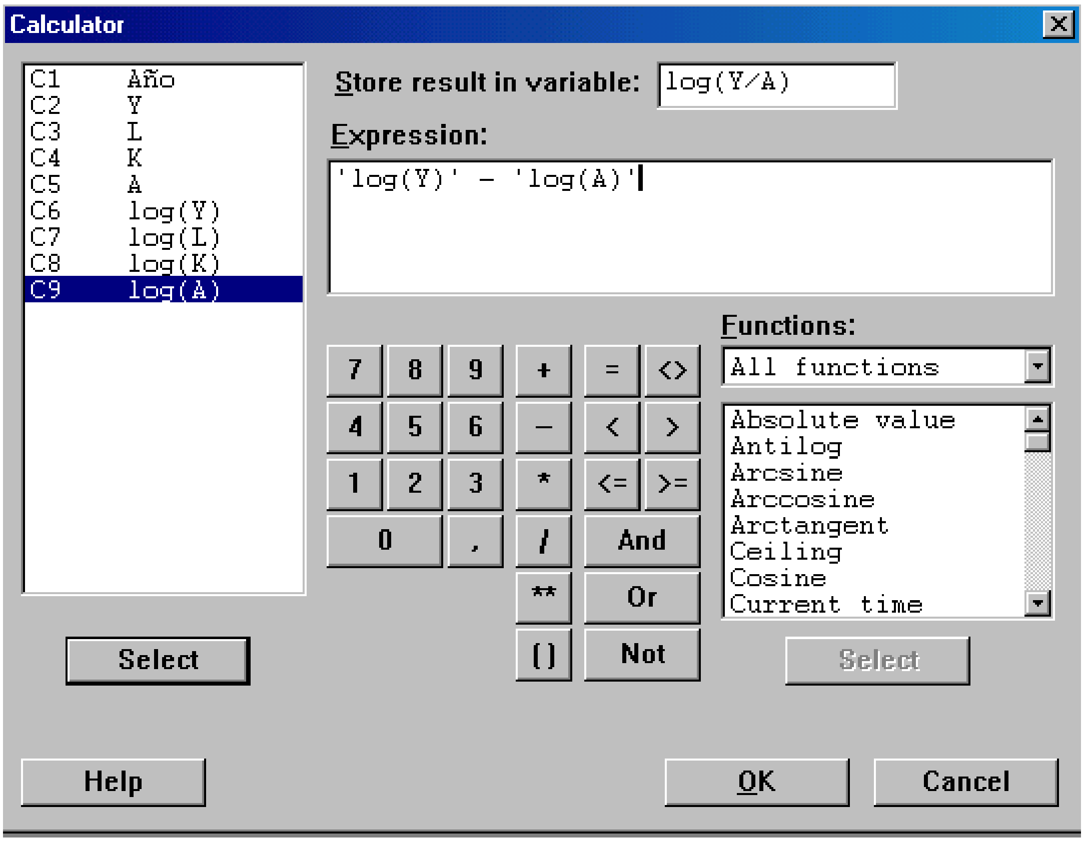
De aquí, reagrupando los términos, queda:

$$\log(Y_t) - \log(A_t) = \beta_1 + \beta_2 \cdot [\log(L_t) - \log(A_t)] + \beta_3 \cdot [\log(K_t) - \log(A_t)] + u_t, t = 1,\dots,T.$$

O bien,

$$\log\left(\frac{Y_t}{A_t}\right) = \beta_1 + \beta_2 \cdot \log\left(\frac{L_t}{A_t}\right) + \beta_3 \cdot \log\left(\frac{K_t}{A_t}\right) + u_t, t = 1,\dots,T.$$

Así, habremos obtenido el modelo en variables transformadas. Recordemos que el Minitab permite transformarlas mediante la opción **`Calc > Calculator`**: La primera etapa del estudio consiste en estimar el modelo por MCO mediante el Minitab. Para ello, seleccionamos **`Stat > Regression > Regression:`**



Así mismo, transformamos las variables L y K. Una vez transformadas todas las variables del modelo, procedemos a estimar el modelo restringido, siguiendo las indicaciones disponibles en el *math-block* correspondiente al MRLM. Los resultados de estimación se presentan en el siguiente cuadro:

Regression Analysis

```
The regression equation is
log(Y/A) = 0,843 + 0,001 log(L/A) + 0,850 log(K/A)

Predictor        Coef       StDev          T        P
Constant       0,8426      0,4246       1,98    0,055
log(L/A)       0,0010      0,1087       0,01    0,993
log(K/A)       0,8504      0,1280       6,65    0,000

S = 0,09345     R-Sq = 95,2%     R-Sq(adj) = 95,0%

Analysis of Variance

Source            DF          SS          MS         F         P
Regression         2      6,2593      3,1296    358,39     0,000
Residual Error    36      0,3144      0,0087
Total             38      6,5736
```

La estimación nos proporciona SCE_R = 0,3144. Por otro lado, a partir de la estimación del modelo Cobb-Douglas no restringido (*math-block* "MRLM") SCE_A = 0,1595. Calculamos F_0 = 33,991 y lo comparamos con el valor crítico distribuido con una distribución F de Snedecor con 1 y 35 grados de libertad, para nivel de significación α = 0,05:

$$F_0 = 33{,}991 > F_{1,35;0,05} = 4{,}1213.$$

Puesto que el valor de estadístico descansa fuera de la región de aceptación, rechazamos la hipótesis nula. En conclusión, la función de producción no presenta rendimientos de escala constantes.

Contraste de cambio estructural

Ejemplo 2: Se desea estudiar la relación del ahorro privado (Y_t) e ingreso privado (X_t) en la economía de los Estados Unidos, ambos expresados en miles de millones de dólares. Se dispone de 96 datos trimestrales, correspondientes a 1977:1T – 2000:4T que aparecen en el archivo **AhorroEEUU.mtw**. Dividiremos la muestra en tres períodos en función del partido político en poder (partido demócrata o partido republicano). El primer período (1977:1T – 1980:4T) corresponde a presidencia demócrata (J. Carter). El segundo período (1981:1T – 1992:4T) corresponde a presidencia republicana (R. Reagan y G. Bush). El tercer período (1993:1T – 2000:4T) corresponde a presidencia demócrata (B. Clinton).

Consideremos el siguiente modelo:

$$Y_t = \beta_1 + \beta_2 \cdot X_t + u_t, t = 1,...,T.$$

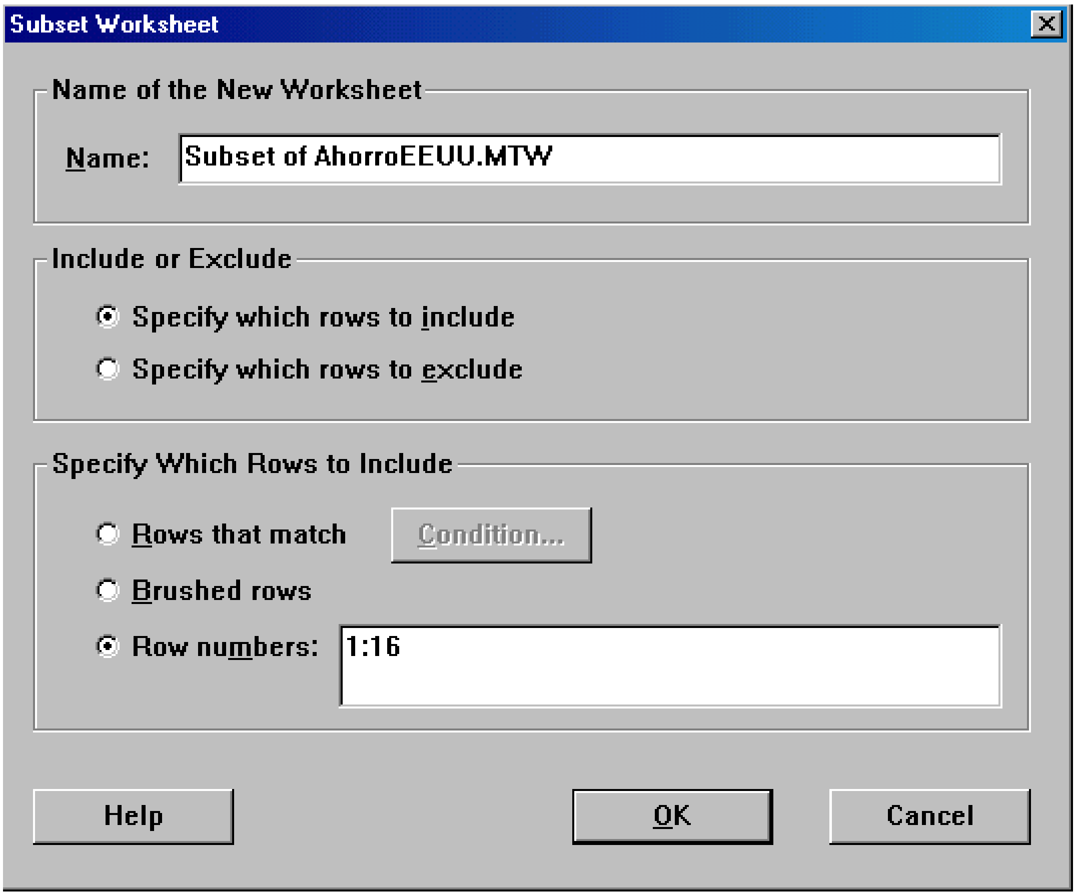
Suponiendo que el modelo satisface las hipótesis del MRLM con normalidad del término de error, estimaremos los parámetros de esta ecuación de regresión, utilizando toda la información disponible. De esta manera, habremos imputado la restricción de igualdad de parámetros entre los distintos períodos de según el partido de presidencia. Recordemos que la estimación del modelo por MCO mediante el Minitab se realiza seleccionando **`Stat > Regression > Regression`***:*

Regression Analysis

```
The regression equation is
Y = 230 + 0,00648 X

Predictor        Coef       StDev          T        P
Constant       229,70       21,51      10,68    0,000
X            0,006483    0,004337       1,49    0,138

S = 82,16       R-Sq = 2,3%      R-Sq(adj) = 1,3%

Analysis of Variance

Source            DF          SS          MS        F        P
Regression         1       15085       15085     2,23    0,138
Residual Error    94      634516        6750
Total             95      649600
```

Observamos que, en contra a lo esperado, el modelo explica tan sólo un 2,3% de la variación del ahorro privado. Este resultado nos lleva pensar que los hogares adoptan políticas de presupuesto familiar diferentes dependiendo del partido político en poder. A fin de contestar a esta pregunta, recurriremos al contraste de Chow de cambio estructural para contrastar si puede admitirse que la relación ahorro-ingreso depende del partido político en el poder. A efectos de contrastes de hipótesis, mantendremos un 5% nivel de significación. Para dividir una muestra en 3 submuestras, seguiremos los siguientes pasos:

Seleccionamos **`Manip > Subset Worksheet…`** :



Así, obtenemos la primera submuestra de 16 observaciones. Para obtener la segunda (filas 17 a 64) y tercera (filas 65 a 96) submuestras, reiteramos el mismo procedimiento. A continuación estimamos el modelo para cada una de las tres submuestras. Para ello, se recomienda seguir las indicaciones propuestas en el *math-block* relativo al MRLM. El software nos proporciona las siguientes estimaciones:

Primer período (presidencia demócrata, J. Carter):

Regression Analysis

```
The regression equation is
Y = - 62,6 + 0,113 X

Predictor        Coef       StDev          T        P
Constant       -62,56       13,14      -4,76    0,000
X            0,113130    0,006603      17,13    0,000

S = 7,001       R-Sq = 95,4%     R-Sq(adj) = 95,1%

Analysis of Variance

Source          DF          SS          MS        F        P
Regression       1       14387       14387   293,52    0,000
Residual Error  14         686          49
Total           15       15073
```

Segundo período (presidencia republicana, R. Reagan, G. Bush):

Regression Análisis

```
The regression equation is
Y = 103 + 0,0493 X

Predictor        Coef        StDev          T        P
Constant       103,09        21,53       4,79    0,000
X            0,049333     0,005352       9,22    0,000

S = 33,47       R-Sq = 64,9%     R-Sq(adj) = 64,1%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1       95190       95190     84,95    0,000
Residual Error    46       51543        1121
Total             47      146733
```

Tercer período (presidencia demócrata, B. Clinton):

Regression Analysis

```
The regression equation is
Y = 841 - 0,0860 X

Predictor        Coef        StDev          T        P
Constant       840,79        67,64      12,43    0,000
X           -0,085958     0,009807      -8,76    0,000

S = 49,76       R-Sq = 71,9%     R-Sq(adj) = 71,0%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      190216      190216     76,82    0,000
Residual Error    30       74282        2476
Total             31      264497
```

Vemos que las estimaciones del modelo son totalmente distintas para cada una de las tres muestras, lo cual sugiere que los cambios de partido político en poder podrían haber provocado un cambio estructural en la relación ahorro-ingreso. Especificamos la hipótesis nula: H_0 : $\beta_1^{(1)} = \beta_1^{(2)} = \beta_1^{(3)}$; $\beta_2^{(1)} = \beta_2^{(2)} = \beta_2^{(3)}$ a fin de contrastarla frente a la hipótesis alternativa H_A : No H_0. Observamos que eso se traduce en 4 restricciones. Haremos uso del estadístico de contraste F, que tiene la siguiente expresión:

$$F_0 = \frac{SCE_R - SCE_A}{SCE_A} \cdot \frac{n-3k}{q} = \frac{SCE_R - (SCE_{A1} + SCE_{A2} + SCE_{A3})}{SCE_{A1} + SCE_{A2} + SCE_{A3}} \cdot \frac{n-3k}{q} \sim F_{q;n-3k},$$

$$F_0 = \frac{634516 - (686 + 51543 + 74282)}{686 + 51543 + 74282} \cdot \frac{90}{4} = 90{,}35.$$

La tabla de distribución de F de Snedecor nos proporciona el valor crítico $F_{q;T-3k;\alpha} = F_{4,90,0,05} =$ 2,4729. Vemos que $F_0 = 90{,}35 \notin$ R.A. (0; 2,4729), con lo cual rechazamos la hipótesis nula de permanencia estructural. Por tanto, podemos admitir que la relación ahorro-ingreso depende del partido político en el poder. Es más, el ingreso pasa a ser estadísticamente significativo para cada uno de los tres períodos. En los dos primeros períodos, el ingreso influye positivamente en la predicción del ahorro, mientras que en el tercero, durante la presidencia de B. Clinton, el ingreso provoca un efecto negativo. Podemos pensar que, en el último período, los hogares se forman expectativas pesimistas sobre los ingresos futuros, con lo que tienden a destinar una mayor proporción de los ingresos a sus cuentas de ahorro.

BIBLIOGRAFÍA

[1] Artís, M.; Suriñach, J.; et al (2001): "Introducción a la econometría". Ed. Fundació per a la Universitat Oberta de Catalunya. Barcelona.

[2] Carter, R.; Griffiths, W.; Judge, G. (2000): "Using Excel for Undergraduate Econometrics". ISBN: 0-471-41237-6

[3] Doran, H. (1989): "Applied Regression Analysis in Econometrics". Ed. Marcel Dekker, Inc. ISBN: 0-8247-8049-3

[4] Green, W.H. (1999): "Análisis Econométrico". Prentice Hall Iberia. ISBN: 84-8322-007-5

[5] Gujarati, D. (1997): "Econometría básica". McGraw-Hill. ISBN 958-600-585-2

[6] Johnston, J. (2001): "Métodos de econometría". Ed. Vicens Vives. Barcelona. ISBN 84-316-6116-X

[7] Kennedy, P. (1998): "A Guide to Econometrics". Ed. MIT Press. ISBN: 0262611406

[8] Novales, A. (1993): "Econometría". McGraw-Hill. ISBN 84-481-0128-6

[9] Pulido, A. (2001): "Modelos econométricos". Ed. Pirámide. Madrid. ISBN 84-368-1534-3

[10] Uriel, E. (1990): "Econometría: el modelo lineal". Ed. AC. Madrid. ISBN 84-7288-150-4

[11] Wooldridge, J. (2001): "Introducción a la Econometría: un enfoque moderno". Ed. Thomson Learning. ISBN: 970-686-054-1

ENLACES

- http://www.feweb.vu.nl/econometriclinks/index.html
  The Econometrics Journal On-Line
- http://www.elsevier.com/hes/books/02/menu02.htm
  Libro on-line: Handbook of Econometrics Vols. 1-5
- http://elsa.berkeley.edu/users/mcfadden/discrete.html
  Libro on-line: Structural Analysis of Discrete Data and Econometric Applications
- http://www.oswego.edu/~kane/econometrics/stud_resources.htm
  Online Resources for Econometric Students
- http://www.econ.uiuc.edu/~morillo/links.html
  Econometric Sources: a collection of links in econometrics and computing. University of Illinois
- http://www.econometrics.net/
  Econometrics, Statistics, Mathematics, and Forecasting
- http://ideas.uqam.ca/EDIRC/ectrix.html
  Economics Departments, Institutes and Research Centers in the World: Econometrics, Mathematical Economics