

Σοηομετρία

Capítulo 1: Introducción

-Amparo Sancho-©
-Guadalupe Serrano-
-Bernardí Cabrer -

¿Qué es un modelo econométrico?

¿Qué son los estimadores por MCO y cuáles son sus propiedades?

¿Qué son los estimadores por MV y cuáles son sus propiedades?

¿Cuáles son los criterios para discriminar entre modelos?

Principales formas funcionales y su interpretación

Etimológicamente, *Σconometría* significa medición de la economía. Existen muchas definiciones para esta disciplina científica que van desde las más complejas, como “*el análisis cuantitativo de fenómenos económicos reales, basados en el desarrollo simultáneo de la teoría y la observación, relacionados mediante métodos apropiados de inferencia*”¹ hasta otras más simplificadoras como “*la determinación empírica de las leyes económicas*”². Como disciplina teórica, en su elaboración se unen las Matemáticas, la Estadística y la

Teoría Económica; como disciplina empírica se centra en la investigación social.

En este capítulo se realiza una breve reseña de los aspectos más importantes del proceso de estimación de los modelos lineales.

Para un análisis en profundidad de estos temas, se debe acudir a la bibliografía recomendada.

¿Qué es un modelo econométrico?

Los **MODELOS ECONÓMICOS** relacionan una variable dependiente con otras independientes o explicativas. Supone una relación exacta y determinista entre las variables.

$$y = f(x_1, x_2, x_3, \dots, x_k)$$

Sin embargo, a nivel empírico las relaciones no son deterministas. De hecho, si se especifica una relación a través de una función matemática para una

¹ P.A. Samuelson, T.C. Koopmans y J.R.N. Stone, “Report of the Evaluative Committee for Econometrica”, *Econometrica*, vol. 22, núm. 2, abril de 1954, pp.141-146.

² H. Theil, *Principles of Econometrics*, John Wiley & Sons, Nueva York, 1964, p.1.

muestra determinada con total seguridad habría más de una observación que no coincidiría con la función preestablecida.

Para considerar las relaciones inexactas entre las variables del mundo económico surgen los **MODELOS ECONOMETRICOS**. Éstos, además de relacionar una variable dependiente con otras independientes o explicativas (relación de comportamiento), introducen una **componente aleatoria** o término de error. Ésta tiene un comportamiento estocástico y representa factores determinantes del comportamiento de la variable endógena que los modelos no pueden recoger de forma explícita. Así, el comportamiento de la variable (y) viene explicado por un modelo o relación en la que se puede distinguir una parte determinista (integrada por las variables explicativas) y una parte aleatoria.

$$y = f(x_1, x_2, x_3, \dots, x_k, u)$$

Un sencillo ejemplo para una mejor distinción consiste en suponer una relación lineal entre una variable dependiente y y la independiente x . He aquí los modelos:

Modelo matemático

$$y = \alpha + \beta \cdot x$$

Modelo econométrico

$$y = \alpha + \beta \cdot x + u$$

Cuando se tiene una muestra de tamaño N de la población, el modelo queda especificado como $y_i = \beta_1 + \beta_2 \cdot x_i + u_i \quad i=1, \dots, N$.

¿Qué son los estimadores por mínimos cuadrados ordinarios (MCO) y cuáles son sus propiedades?

El **modelo de regresión lineal simple** es el más sencillo e incluye únicamente una variable independiente:

$$y_i = \beta_1 + \beta_2 \cdot x_i + u_i$$

En general, el problema que se plantea es la búsqueda del valor de los parámetros de la recta que mejor se ajusta a la nube de puntos que forma cada par de observaciones (y_i, x_i) . Para ello se minimiza la distancia entre la recta \hat{y} y todos los puntos observados y_i , es decir se minimizan la suma de errores al cuadrado para que los signos (+) y los (-) no se compensen:

$$\min \sum e_i^2 = \sum (y_i - \hat{\beta}_1 - \hat{\beta}_2 \cdot x_i)^2 = \sum (y_i - \hat{y}_i)^2$$

La minimización se realiza a través de las condiciones de primer orden del problema. Para ello hay que obtener las primeras derivadas de y_i respecto a los valores β_1 y β_2 e igualar a cero, llegando al siguiente **SISTEMA DE ECUACIONES NORMALES**:

$$\frac{\partial \sum e_i^2}{\partial \beta_1} = 0 \Rightarrow \sum y_i = N \cdot \hat{\beta}_1 + \hat{\beta}_2 \cdot \sum x_i \quad (1)$$

$$\frac{\partial \sum e_i^2}{\partial \beta_2} = 0 \Rightarrow \sum x_i \cdot y_i = \hat{\beta}_1 \cdot \sum x_i + \hat{\beta}_2 \cdot \sum x_i^2 \quad (2)$$

Despejando de (1) se obtiene:

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \cdot \bar{x}$$

y sustituyendo (1) en (2) se obtiene que ;

$$\hat{\beta}_2 = \frac{\sum x_i \cdot y_i - \bar{y} \cdot \sum x_i}{\sum x_i^2 - \bar{x} \cdot \sum x_i} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

Asimismo, el estimador de la varianza de las perturbaciones del modelo es:

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{N - 1}$$

El **modelo de regresión lineal múltiple** es la generalización del modelo de regresión lineal simple, para el caso de que el modelo tenga k-1 variables exógenas. Su resolución también pasa por un problema de minimización de residuos para hallar el vector de estimadores a partir del sistema de k ecuaciones normales.

FORMA EXTENDIDA	FORMA MATRICIAL
$y_i = \beta_1 + \beta_2 \cdot x_{2i} + \dots + \beta_k \cdot x_{ki} + u_i$ $\min \sum e_i^2 = \sum (y_i - \hat{y}_i)^2$	$Y = X \cdot \beta + u$ $\begin{matrix} N \times 1 & N \times k & k \times N & N \times 1 \end{matrix}$ $\min_{1 \times N \quad N \times 1} e' e = (Y - X\hat{\beta})'(Y - X\hat{\beta})$
Sistema de ecuaciones normales;	Sistema de ecuaciones normales;
$\sum y_i = N \cdot \hat{\beta}_1 + \hat{\beta}_2 \sum x_i$ $\sum x_i \cdot y_i = \hat{\beta}_1 \sum x_i + \hat{\beta}_2 \sum x_i^2$	$[x'x] \cdot \hat{\beta} = x'y$ $\begin{bmatrix} N & \sum_{i=1}^N x_{2i} & \dots & \sum_{i=1}^N x_{ki} \\ \sum_{i=1}^N x_{2i} & \sum_{i=1}^N x_{2i}^2 & \dots & \sum_{i=1}^N x_{2i} \cdot x_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^N x_{ki} & \sum_{i=1}^N x_{ki} \cdot x_{2i} & \dots & \sum_{i=1}^N x_{ki}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum x_2 y \\ \sum x_3 y \\ \vdots \\ \sum x_k y \end{bmatrix}$
<p>Vector de estimadores $\hat{\beta}_{MCO} = (X'X)^{-1} \cdot X'Y$</p> <p>Estimador de la varianza de u: $\hat{\sigma}_{MCO}^2 = \frac{(e'e)}{N - k}$</p>	

CUADRO 1: Hipótesis clásicas del modelo de regresión lineal

A continuación se exponen los supuestos básicos (también conocidos como clásicos) que se asumen como punto de partida del modelo de regresión lineal. En principio, dicho modelo se supone bien especificado, es decir, la ecuación es correcta y no falta ni sobra ninguna variable. Las hipótesis básicas son:

⇒ El valor esperado de la PERTURBACIÓN ALEATORIA es cero.

$$E(u_i) = 0 \quad E(u) = 0$$

⇒ HOMOCEDASTICIDAD ⇒ varianza constante de las perturbaciones a lo largo de la distribución.

$$E(u_i)^2 = \sigma_u^2 \quad E(u u') = \sigma_u^2 I$$

⇒ No AUTOCORRELACIÓN ⇒ las perturbaciones no están autocorrelacionadas a lo largo de la distribución.

$$E(u_i \cdot u_j) = 0, \forall i \neq j$$

El cumplimiento de los supuestos de No Autocorrelación y de Homocedasticidad recibe el nombre de Esfericidad. En este caso se habla de perturbaciones esféricas.

⇒ Las variables x son FIJAS, esto es, NO ALEATORIAS en el muestreo.

$$E(x_i \cdot u_i) = 0$$

$$E(x' \cdot u) = 0$$

⇒ AUSENCIA DE MULTICOLINEALIDAD entre las variables explicativas (modelo de rango pleno).

$$\rho(x'x) = k$$

Éstas son las hipótesis básicas sobre las que descansa el modelo de regresión lineal clásico. A lo largo del curso se va a ir viendo que son excesivamente rigurosas y en la mayoría de ocasiones dejan de cumplirse.

Si se establecen unas hipótesis de comportamiento en el modelo como las expuestas en el cuadro anterior, los Estimadores obtenidos por **Mínimos Cuadrados Ordinarios** (MCO) presentan las siguientes **PROPIEDADES PROBABILÍSTICAS**³:

⇒ LINEALES EN Y

$$\hat{\beta}_{MCO} = (X'X)^{-1} X'Y = f(Y)$$

y dado que las x son fijas en la muestra :

$$\hat{\beta}_{MCO} = (X'X)^{-1} X'Y = f(Y) = f(u)$$

⇒ INSESGADEZ

$$\begin{aligned} E(\hat{\beta}_{MCO}) &= E[(X'X)^{-1} \cdot X'Y] = E[(X'X)^{-1} \cdot X'(X\beta + u)] = \\ &= \underbrace{(X'X)^{-1} \cdot X'X}_{=I} \cdot \beta + E[(X'X)^{-1} \cdot X' \cdot u] \\ &= \beta + \underbrace{(X'X)^{-1} \cdot X' \cdot E(u)}_{=0} = \beta \end{aligned}$$

³ Para un desarrollo más completo puede acudir a URIEL, E. et al. (1990)

$$E(\hat{\beta}_{MCO}) = \beta$$

✓ Los estimadores por MCO son insesgados bajo el supuesto de que las x son fijas.

➤ OPTIMALIDAD Y EFICIENCIA

$$\begin{aligned} V(\hat{\beta}_{MCO}) &= E[(\hat{\beta} - \beta) \cdot (\hat{\beta} - \beta)] = E[(X'X)^{-1} \cdot X' \cdot u \cdot u' \cdot X \cdot (X'X)^{-1}] = \\ &= \sigma_u^2 E[(X'X)^{-1} (X'X) (X'X)^{-1}] = \sigma_u^2 \cdot (X'X)^{-1} \end{aligned}$$

$$V(\hat{\beta}_{MCO}) = \sigma_u^2 \cdot (X'X)^{-1}$$

Puede demostrarse, a través del Teorema de Gauss-Markov, que el resultado obtenido es de varianza mínima.

✓ Los estimadores por MCO son lineales insesgados y óptimos bajo los supuestos de homocedasticidad y no autocorrelación ($E(u_i \cdot u_j) = \sigma^2 \cdot I$).

¿Qué son los estimadores por Máxima Verosimilitud (MV) y cuáles son sus propiedades?

Se plantea un nuevo supuesto adicional a las hipótesis básicas planteadas anteriormente: La perturbación aleatoria sigue una distribución Normal de acuerdo con esta estructura:

$$u_i \approx N(0, \sigma_u^2)$$

Las perturbaciones aleatorias que así se distribuyen se conocen como **ruido blanco**.

Bajo este supuesto se podrían obtener un nuevo tipo de estimadores: los de Máxima Verosimilitud. Para ello, bastaría con resolver un problema de maximización de la función de probabilidad o verosimilitud. Se realizará esta demostración para el modelo de regresión simple.

La idea que subyace es la de encontrar aquellos valores de los parámetros que hacen máxima la probabilidad de que la muestra disponible proceda de una población caracterizada por dichos parámetros.

La FUNCIÓN DE PROBABILIDAD CONJUNTA puede descomponerse en el productorio de las individuales, dado que las perturbaciones aleatorias son independientes unas de otras (supuesto de no autocorrelación):

$$P(u_1, u_2, u_3, \dots, u_N) = P(u_1) \cdot P(u_2) \cdot P(u_3) \cdot \dots \cdot P(u_N)$$

↑
u's independientes

Como se ha asumido que la perturbación sigue una distribución Normal, su FUNCIÓN DE PROBABILIDAD INDIVIDUAL u_i viene definida por la conocida expresión:

$$P(u_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2} \cdot u_i^2}$$

Para hallar la función de probabilidad conjunta $P(u)$ se realiza el productorio de las distribuciones individuales; dado que es un vector.

$$\prod_{i=1} P(u_i) = P(u) = \frac{1}{(2\pi\sigma^2)^{N/2}} \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum u_i^2}$$

La perturbación aleatoria en el modelo de regresión simple puede ponerse en función de las variables y_i y x_i :

$$u_i = y_i - (\beta_1 + \beta_2 \cdot x_i)$$

De este modo se llega a la FUNCIÓN DE VEROSIMILITUD, que es la que se maximizará:

$$L = P(u/X, \sigma) = \frac{1}{(2\pi\sigma^2)^{N/2}} \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum (y_i - \beta_1 - \beta_2 \cdot x_i)^2}$$

Para ello es preciso realizar antes una transformación: linealizar la función de verosimilitud. Basta con tomar logaritmos:

$$\ell = \ln L = -\frac{N}{2} \cdot \ln 2\pi - \frac{N}{2} \cdot \ln \sigma_u^2 - \frac{1}{2\sigma_u^2} \cdot \sum (y_i - \beta_1 - \beta_2 \cdot x_i)^2$$

Para resolver cualquier problema de maximización, se obtienen las CONDICIONES DE PRIMER ORDEN, derivando la función respecto de las variables de control e igualando a 0.

En este caso, las variables de control del problema de maximización, son los parámetros para los que se busca maximizar la función; β_1, β_2 y σ^2 .

$$\frac{\partial \ln L}{\partial \beta_1} = + \frac{2}{2 \sigma_u^2} \cdot \sum (y_i - \beta_1 - \beta_2 \cdot x_i) = 0 \longrightarrow \sum y_i = N \cdot \tilde{\beta}_1 + \tilde{\beta}_2 \cdot \sum x_i$$

$$\boxed{\tilde{\beta}_{1MV} = \hat{\beta}_{1MCO}}$$

$$\frac{\partial \ln L}{\partial \beta_2} = + \frac{2}{2 \sigma_u^2} \cdot \sum x_i \cdot (y_i - \beta_1 - \beta_2 \cdot x_i) = 0 \longrightarrow \sum x_i \cdot y_i = \tilde{\beta}_1 \cdot \sum x_i + \tilde{\beta}_2 \cdot \sum x_i^2$$

$$\tilde{\beta}_{2MV} = \hat{\beta}_{2MCO}$$

$$\frac{\partial \ln L}{\partial \sigma_u^2} = - \frac{N}{2 \sigma_u^2} - \frac{1}{2} \cdot \sum ((y_i - \beta_1 - \beta_2 \cdot x_i)^2) \cdot (\sigma_u^2)^{-2} = 0$$

$$\boxed{\tilde{\sigma}_{uMV}^2 = \frac{e' \cdot e}{N} \neq \hat{\sigma}_{uMCO}^2 = \frac{e' \cdot e}{N - k}}$$

↑
varianza

↑
cuasivarianza

Así, se puede comprobar que los estimadores $\hat{\beta}_{MCO}$ por MCO son lineales, insesgados y óptimos bajo las hipótesis básicas establecidas:

$$E(u) = 0 \quad E(u_i \cdot u_j) = \sigma^2 \cdot I$$

Si suponemos además que $u_i \approx N(0, \hat{\sigma}_u^2)$, entonces coinciden con los estimadores de máxima verosimilitud del modelo de regresión lineal simple.

Los estimadores obtenidos por el método de **Máxima Verosimilitud (MV)** presentan unas **PROPIEDADES** muy deseables. A saber:

⇒ Siguen una distribución asintóticamente Normal ⇒ $\tilde{\beta}_{MV} \approx N$

⇒ Son consistentes

$$\lim_{N \rightarrow \infty} \text{prob } \tilde{\beta}_{MV} \longrightarrow \beta$$

⇒ Son asintóticamente eficientes ⇒ menor varianza asintótica que cualquier otro estimador consistente.

Los estimadores obtenidos por el método de Máxima Verosimilitud siguen una distribución asintóticamente Normal, son consistentes y asintóticamente eficientes.

De aquí se deduce la importancia del supuesto de que la perturbación aleatoria tenga una distribución normal, ya que ello conduce a que los estimadores por MCO tengan también las mismas propiedades que los estimadores de MV. Igualmente este supuesto permite establecer una función de distribución para los estimadores por MCO y con ello determinar intervalos de confianza y realizar contrastes de hipótesis.

Este supuesto de normalidad de las perturbaciones aleatorias hay que contrastarlo. Para ello se utiliza el test de **Jarque-Bera**, aplicado sobre los residuos del modelo estimado por MCO.

EJEMPLO 1: Test de Jarque-Bera o de Normalidad de los residuos

El Test de Jarque-Bera se formula bajo la hipótesis nula de Normalidad de los residuos. Se construye de la siguiente manera:

$$JB = N \cdot \left[\frac{s^2}{6} + \frac{(k-3)^2}{24} \right] \approx \chi^2_2$$

donde N es el tamaño muestral, s el coeficiente de asimetría y k la kurtosis o apuntamiento. Bajo la H_0 de normalidad, el estadístico JB se distribuye como una χ^2 de dos grados de libertad.

Si el valor obtenido del estadístico de Jarque-Bera es menor que el valor crítico tabulado, no se rechaza la hipótesis nula de normalidad. En caso contrario se rechaza la normalidad de la variable.

Asimismo, pese a que la estimación por MV resulte muy recomendable, debe someterse a pruebas de validación. Para ello se realizan diversos **CONTRASTES** sobre los parámetros y sobre el modelo en sí, que diferirán en cierta medida de los que se aplicaban habitualmente tras la estimación por MCO.

En cuanto a la inferencia, los contrastes de **SIGNIFICATIVIDAD INDIVIDUAL** de las variables explicativas se pueden realizar a partir de la distribución Normal (0,1), en muestras muy grandes. Ahora bien, en el caso de que se cumpla el supuesto de normalidad asintótica y /o se conozca la varianza del estimador es posible calcular el siguiente estadístico:

$$z - \text{stat} = \frac{\tilde{\beta}_i - \beta_i}{\sqrt{\text{Var } \tilde{\beta}_i}} \approx N(0,1)$$

Estos contrastes se emplean para analizar si una variable, por sí misma, explica parte del comportamiento de la variable endógena. De este modo, las hipótesis de este contraste bilateral o de dos colas son las siguientes:

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0$$

⇒ $|z - \text{stat}| < |N(0,1)| \Rightarrow$ No Rechazar $H_0 \Rightarrow$ la variable no es significativa.

⇒ $|z - \text{stat}| > |N(0,1)| \Rightarrow$ Rechazar $H_0 \Rightarrow$ la variable es significativa.

El análisis de la **SIGNIFICATIVIDAD GLOBAL** permite contrastar si todo el modelo en su conjunto es significativo para explicar el comportamiento de la variable endógena. Generalmente se considera la inclusión de todas las variables explicativas simultáneamente, comprendiendo todos los parámetros salvo la constante o término independiente. La hipótesis a contrastar, por tanto, será la siguiente:

$$H_0: \beta_2 = \beta_3 = \dots = \beta_k = 0$$

$$H_1: \text{Algún } \beta_i \neq 0$$

Para analizar la significatividad global de un modelo estimado por MCO se podía acudir a la distribución F de Snedecor. En la mayoría de ocasiones los

valores generados eran muy elevados, existiendo una tendencia irrevocable hacia el rechazo de la hipótesis nula o, lo que es lo mismo, la aceptación de la significatividad de los modelos estimados. Por ello, para el contraste de la significatividad conjunta de los modelos estimados por MV se realiza un test genérico: el **TEST DE LA RAZÓN DE VEROSIMILITUD** que parte de la expresión:

$$\lambda = \frac{L_R}{L_G} \longrightarrow \text{Ratio o razón de verosimilitud}$$

donde L_R se corresponde con la función de verosimilitud del modelo restringido y L_G con la función de verosimilitud del modelo general o sin restringir. En este caso se distinguen 2 modelos: el general y el restringido. El primero de ellos es el modelo que se está sometiendo a análisis; el segundo es aquel en el que se introduce la restricción, esto es, en el que se hace efectiva la hipótesis nula.

➤ **Modelo completo o general** \Rightarrow el modelo sin la restricción.

$$y_i = \beta_1 + \beta_2 \cdot x_{2i} + \beta_3 \cdot x_{3i} + \dots + \beta_k \cdot x_{ki} + u_i \text{ sería el modelo general.}$$

➤ **Modelo restringido** \Rightarrow el modelo en donde se cumple la hipótesis nula enunciada anteriormente.

$$\text{Si } H_0 \text{ es cierta } \Rightarrow y_i = \beta_1 \text{ sería el modelo restringido.}$$

A partir de la razón de verosimilitud se define el estadístico siguiente:

$$RV = -2 \cdot \log \lambda = -2 \cdot (\log L_R - \log L_G) \approx \chi^2_{\text{restricciones en la } H_0}$$

que permite contrastar las restricciones de la hipótesis nula de manera que:

➤ Si H_0 cierta \Rightarrow el modelo no es significativo en su conjunto.

➤ Si H_0 no es cierta \Rightarrow el modelo es significativo en su conjunto.

El test de razón de verosimilitud tiene otras utilidades y no es únicamente válido para contrastar la significatividad conjunta del modelo. Siempre va a poder emplearse el valor de la función de verosimilitud del modelo restringido y del modelo general para contrastar cualquier otra hipótesis sobre los parámetros que se considere relevante.

EJEMPLO 2: Test de omisión de variables o de errores de especificación (Test reset de Ramsey)

Este ejemplo pone de manifiesto uno de los múltiples usos del test de razón de verosimilitud. Suponer que se parte de un modelo de regresión lineal simple muy sencillo bajo la siguiente especificación:

$$y_i = \beta_1 + \beta_2 \cdot x_{2i} + u_i$$

Para analizar si el modelo está bien especificado, una vez realizada la estimación, se realizan los siguientes pasos

1. Estimar el modelo original tal cual estaba previsto:

$$y_i = \beta_1 + \beta_2 \cdot x_{2i} + u_i \quad (1)$$

2. Obtener la variable \hat{y}_i : $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 \cdot x_{2i}$

3. Incluir la variable \hat{y}_i en el modelo original (debe incluirse al cuadrado) como regresor adicional:

$$y_i = \alpha_1 + \alpha_2 \cdot x_{2i} + \alpha_3 \cdot \hat{y}_i^2 + u_i \quad (2)$$

4. Estimar el nuevo modelo (2) y contrastar la significatividad individual de la nueva variable incluida.

4.a Habitualmente, se recurre al estadístico t de Student:

$$\left. \begin{array}{l} H_0: \alpha_3 = 0 \\ H_1: \alpha_3 \neq 0 \end{array} \right\} t_{N-K}$$

Lógicamente, el hecho de que la inclusión de la nueva variable no fuera significativa implicaría una correcta especificación inicial del modelo.

- si $\alpha_3 = 0 \Rightarrow$ el modelo es correcto, es decir, está bien especificado.
- si $\alpha_3 \neq 0 \Rightarrow$ el modelo está mal especificado. En este caso, las predicciones explicarían mejor el modelo que el modelo en sí.

4.b Sin embargo, este contraste también se puede realizar a través de la razón de verosimilitud. La hipótesis a contrastar será también $\alpha_3 = 0$. El modelo (2) será el modelo general; a partir de su estimación podrá obtenerse el $\ln L_G$. El modelo (1), que incorpora la hipótesis nula de $\alpha_3 = 0$, será el modelo restringido; a partir de su estimación podrá obtenerse el $\ln L_R$.

$$RV = -2 \cdot (\log L_R - \log L_G) \approx \chi_1^2$$

Las conclusiones son exactamente idénticas a las anteriores. Si la inclusión de la variable no es significativa quiere decir que el modelo estaba correctamente especificado en su origen.

Este test detecta fallos en la forma funcional y omisión de variables.

¿Cuáles son los criterios para discriminar entre modelos?

La elección de un modelo entre varios que presenten la misma variable endógena se puede realizar a través de distintos estadísticos y/o criterios que se presentan a continuación:

Criterio R^2 o bien \bar{R}^2 : Se escogerá aquel con mayor valor del indicador, que es un estadístico de carácter descriptivo.

Con los criterios de Valor de la Función de Verosimilitud y del Logaritmo de la Función de Verosimilitud: Se escogerá el modelo que presenta mayor valor de ambos estadísticos, siendo habitual su utilización en el caso de muestras muy grandes, y también en el caso de modelos con distinta variable dependiente.

También para muestras grandes, los criterios de información de Akaike y Schwartz tendrán un valor menor para aquel modelo que presente un mejor ajuste.

El criterio de información de **Akaike (AIC)** se define como;

$$\text{AIC} = -\frac{2}{T} \ln L + \frac{2k}{T}$$

Siendo mejor el modelo que menor AIC presente, una vez realizadas las correlaciones pertinentes para tener en cuenta las posibles diferencias en la variables dependiente de los modelos. Sin embargo, se elegirá como mejor modelo el que presente mayor $\ln L$ (de nuevo realizando las correcciones pertinentes ante diferencias en la medida de la variable dependiente).

El criterio de **Schwartz** se utiliza también para comparar modelos como el AIC, pero además este criterio penaliza el tamaño muestral.

$$\text{Schwartz} = -\frac{2}{T} \ln L + \frac{k \cdot \ln T}{T}$$

Al igual que en el caso del AIC, se elegirá aquel modelo con menor criterio de Schwartz.

Los criterios antes descritos permitirán la elección de la forma funcional que mejor se adapte a la estructura de los datos, entre las diferentes propuestas para el modelo. Cuando se pretende discriminar entre un modelo con la variable endógena en niveles, y_i , y otro que presenta como variable dependiente el logaritmo neperiano de dicha variable, $\ln y_i$, el criterio a utilizar es el criterio

de información de Akaike corregido, que permite comparar modelos con formas funcionales distintas de la misma variable endógena.

$$AIC_C = AIC_L + 2 \cdot \overline{\ln Y}$$

siendo AIC_L el criterio AIC del modelo con $\ln y_i$ como variable dependiente

Principales formas funcionales y su interpretación

A menudo una mera línea recta (modelo lineal) no es la especificación funcional más adecuada para ajustar la nube de puntos de las observaciones muestrales. En ocasiones resulta mejor emplear formas funcionales alternativas. Esta es una de las razones por las que surge la necesidad de comparar modelos y elegir aquel que implique un mejor ajuste.

En el siguiente cuadro se recogen las especificaciones, formas funcionales o modelos más habituales, lineales o bien linealizados que se pueden estimar por MCO y MV. Para facilitar la interpretación de los coeficientes estimados en cada uno de los modelos se detallan la propensión marginal y la elasticidad de la variable dependiente ante variaciones de X .

Cuadro: Formas funcionales

	Modelo	Propensión Marginal en el punto medio muestral	Elasticidad
LINEAL	$y_i = \alpha + \beta \cdot x_i + u_i$	$\hat{\beta}$	$\hat{\beta} \cdot \frac{\bar{x}}{\bar{y}}$
INVERSO	$y_i = \alpha + \beta \cdot \frac{1}{x_i} + u_i$	$-\hat{\beta} \cdot \frac{1}{\bar{x}^2}$	$-\hat{\beta} \cdot \frac{1}{\bar{x} \cdot \bar{y}}$
SEMILOGARÍTMICO	$y_i = \alpha + \beta \cdot \ln x_i + u_i$	$\hat{\beta} \cdot \frac{1}{\bar{x}}$	$\hat{\beta} \cdot \frac{1}{\bar{y}}$
DOBLEMENTE LOGARÍTMICO	$\ln y_i = \alpha + \beta \cdot \ln x_i + u_i$	$\hat{\beta} \cdot \frac{\bar{y}}{\bar{x}}$	$\hat{\beta}$
LOGARÍTMICO LINEAL	$\ln y_i = \alpha + \beta \cdot x_i + u_i$	$\hat{\beta} \cdot \bar{y}$	$\hat{\beta} \cdot \bar{x}$
LOGARÍTMICO INVERSO	$\ln y_i = \alpha + \beta \cdot \frac{1}{x_i} + u_i$	$-\hat{\beta} \cdot \frac{\bar{y}}{\bar{x}^2}$	$-\hat{\beta} \cdot \frac{1}{\bar{x}}$

Los modelos logarítmicos del cuadro anterior tienen su origen en modelos exponenciales que se han linealizado posteriormente para su especificación. Conviene tener presente su especificación original:

	Modelo original	Modelo linealizado
DOBLEMENTE LOGARÍTMICO	$y_i = e^{\alpha} \cdot x_i^{\beta} \cdot e^{u_i}$	$\ln y_i = \alpha + \beta \cdot \ln x_i + u_i$
LOGARÍTMICO LINEAL	$y_i = e^{\alpha + \beta \cdot x_i + u_i}$	$\ln y_i = \alpha + \beta \cdot x_i + u_i$

Conclusiones

- Frente a los modelos matemáticos, los modelos econométricos se caracterizan por incluir un término de perturbación o de error, de carácter estocástico o aleatorio.
- Los estimadores de los parámetros de un modelo lineal por MCO, bajo las hipótesis básicas de homocedasticidad y no autocorrelación de las perturbaciones aleatorias y de variables explicativas fijas en el muestreo, son lineales en *los parámetros*, insesgados y óptimos (ELIO).
- Los estimadores por MV de los parámetros de un modelo lineal siguen una distribución Normal y son además consistentes y asintóticamente eficientes, bajo las hipótesis básicas establecidas.
- Existen numerosos estadísticos y criterios para escoger un modelo entre varias especificaciones. Cuando presentan la misma variable endógena, son el coeficiente de determinación ajustado y el criterio de información de Akaike, estadísticos que hay que maximizar y minimizar

respectivamente. Todos ellos permiten elegir la forma funcional que mejor explique el comportamiento de la variable dependiente del modelo.

➡ Cuando se quiere comparar un modelo que presenta la variable endógena en niveles frente a otro con el logaritmo neperiano de esa misma variable como variable dependiente, se debe utilizar el AIC de Akaike corregido para discriminar entre ambas especificaciones.

Lecturas propuestas

Al ser un capítulo introductorio, a continuación se exponen los manuales que mejor se adaptan a la estructura del curso. En cualquier caso, en cada capítulo se citarán los más apropiados para la materia que se esté tratando.

CABRER B., SANCHO, A.; SERRANO G: (2001): *Microeconometría y decisión*. Ed. Pirámide.

GREENE, W.(2001), *Econometría*, Madrid, Prentice Hall, , 3ª ed.

GUJARATI, D. N.,(2003): *Econometría*, Madrid, McGraw-Hill, , 4ª ed.

JOHNSTON, J. y DINARDO, J.(2001): *Métodos de econometría*, Barcelona, Vicens Vives,

NOVALES, A.,(1993): *Econometría*, McGraw-Hill, , 2ª ed.

URIEL, E. et al., (1990). *Econometría*, Madrid, Ed AC.