PUBLICACIONES DE 4er CURSO

Grado: Economía

Asignatura: ECONOMETRÍA III

DEFINICIONES Y CONCEPTOS BÁSICOS

Profesor: Antonio Aznar

Departamento de ANÁLISIS ECONÓMICO

Curso Académico 2015/16

Facultad de
Economía y Empresa
Universidad Zaragoza

Parte Primera

Definición 1.1. Experimentos Aleatorizados

Se refiere a una situación en la que todos los elementos de una población se pueden asignar aleatoriamente a dos grupos, A y B. Los elementos dentro de cada grupo son heterogéneos entre sí, pero podemos decir que, agregadamente, la composición de A es similar a la de B. A los componentes del grupo A se les somete a un estímulo que no reciben los de B. En ambos grupos se mide el nivel alcanzado por una determinada variable que puede verse afectada por el estímulo y que llamamos variable respuesta. Si los niveles alcanzados por esa variable son muy diferentes en ambos grupos podemos decir que la variable estímulo causa a la variable respuesta. Ejemplo: dos parcelas de terreno agrícola contiguas; a una de ellas se le aplica un fertilizante y a la otra no y se mide la diferencia en las cosechas.

Definición 1.2. Ideal Bacon–Descartes

Es una línea de pensamiento, dentro de la Filosofía de la Ciencia, que trata de establecer criterios para preferir una teoría frente a otras. Esta línea de pensamiento fue impulsada en la segunda mitad del siglo XX por el Neo-empirismo por autores como Lakatos, Radnitzky, Watkins,….Los criterios tienen carácter bipolar: por un lado, trata de elegir las teorías que son informativas, profundas y arriesgadas; por otro, trata de mantener teorías que se ajusten a los hechos.

Definición 1.3. Validación según el enfoque de Mill-Robbins

Para los autores en esta corriente, la ciencia económica descansa esencialmente en unas pocas proposiciones generales que son el resultado de la observación o la introspección y que cualquier hombre, tan pronto como oye de ellas, las admite como algo familiar, a partir de las cuales se derivan las conclusiones que serán verdad en ausencia de causas perturbadoras. Podemos decir que, en Economía, tenemos acceso directo a los elementos últimos de nuestras generalizaciones. Las teorías son deducciones a partir de una serie de postulados. Y la mayor parte de estos postulados son todos ellos hipótesis con un contenido simple e indiscutible sobre la experiencia diaria acerca de cómo se administra la escasez de los bienes económicos. No necesitamos ni experimentos controlados ni procedimientos de contraste de tipo estadístico-econométrico para establecer su validez: son algo tan familiar en nuestra experiencia diaria que basta que sean enunciados para que sean reconocidos como obvios. En cuanto a los seguidores de esta línea, además del propio John Stuart Mill (1806-1873), destacan otros economistas del siglo XIX, como Ricardo, Cairnes y Senior, que también pertenecen a esta escuela. El autor más importante de esta corriente fue Robbins quien, en 1932, escribió la obra "Un Ensayo sobre la Naturaleza y Significado de la Ciencia Económica". En fechas recientes los dos autores que han realizado aportaciones dentro de esta tradición han sido Hausman y Cartwright.

Definición 1.4. Validación según el enfoque de Hutchison-Friedman

Para los autores en esta línea, la validación de un sistema teórico hay que hacerla mediante el contraste empírico de alguna parte de ese sistema utilizando procedimientos de contraste de tipo estadístico-econométrico que sean replicables. No vale la introspección ni la evidencia inmediata. Hay que huir del componente subjetivo y tratar de utilizar procedimientos objetivos.

Hutchison escribe en 1938, con 28 años, un libro titulado "El Significado y los Postulados Básicos de la Teoría Económica" que es considerado por muchos como la primera formalización de esta segunda tradición. Milton Friedman escribe en 1953 un trabajo titulado "La Metodología de la Economía Positiva" que es, sin duda, el trabajo de metodología más citado del siglo XX.

Definición 1.5. Aspectos para saber como formular una predicción

A la hora de plantear un ejercicio de predicción es importante tener en cuenta los siguientes aspectos

- Quién va a utilizar la predicción y para qué.
- Qué nivel de detalle se requiere
- El horizonte temporal de predicción
- Nivel de exactitud requerido
- Recursos disponibles
- Urgencia y frecuencia con las que se necesitan las predicciones.
- Nivel de agregación deseado

- Todos aquellos aspectos que pueden ser relevantes para incorporar la predicción en el proceso de toma de decisiones.

Definición 1.6. Métodos de predicción. Tipos

Podemos distinguir tres grandes grupos de métodos de predicción:

A) Métodos Subjetivos

B) Estudios de Mercado

C) Métodos Objetivos

A) Métodos Subjetivos.

Son métodos basados en la información privilegiada que algunos agentes tienen del fenómeno que se va a predecir. Dentro de estos métodos podemos distinguir

- Jurado de Opinión
- Método de Escenarios
- Método Delphi
- Analogía Histórica

B) Estudios de Mercado

En este caso, es necesario recoger la información que está dispersa en una población mediante técnicas de muestreo.

C) Métodos Objetivos

- Modelos univariantes de series temporales
- Modelos Causales Multivariantes

Definición 1.7. Metapredicción

Hace referencia a los siguientes puntos:

1. La incertidumbre respecto al futuro siempre va a estar presente en cualquier ejercicio predictivo. La finalidad de éste no es tanto eliminar dicha incertidumbre como minimizarla.

2. Los economistas tienen que ser conscientes del tipo de realidad que la Economía estudia y trata de proyectar. Hutchison (1977) nos sitúa en el lugar apropiado: "Hay, sin embargo, generalizaciones útiles en la economía y las ciencias sociales que son descritas mejor como tendencias ya que en general no son tan precisas ni tan contrastables como las leyes propiamente dichas. Tendencias, y no leyes, es lo que el material de la economía y las ciencias sociales parecen proporcionar o han

proporcionado hasta el momento…. A falta de leyes, todo lo que los economistas tienen son tendencias y deben de procurar sacar el máximo de ellas”.

3. Cada situación hay que tratarla de forma específica y diferente. Hay que distinguir según sea el tipo de variable, el horizonte de predicción, la información disponible, la frecuencia y urgencia requeridas, los medios disponibles para calcular la predicción, etc.

4. Como ya hemos comentado, hay numerosos métodos de predicción y en cada situación hay que aplicar el que sea más apropiado.

5. Hay que ser cuidadosos sobre como se presentan los resultados de un ejercicio predictivo. En principio, es importante destacar que la predicción son dos números: el primero es el valor predicho y el segundo se refiere a la amplitud del intervalo formado por los valores que, en torno al valor predicho, se consideran como aceptables habiendo tomado un nivel de significación a priori.

Definición 1.8. La Ley de Esperanzas Iteradas

La media de Y es la media ponderada de la esperanza condicional de Y dado X, utilizando como ponderación la distribución marginal de X. Es decir,

$$E(Y) = \sum_{1}^{n} E(Y / X = x_i) P(X = x_i)$$

La ley de esperanzas iteradas implica que si la media condicional de Y dado X es cero, entonces la media de Y es cero.

Definición 1.9. Muestreo Aleatorio Simple y Variables Aleatorias i.i.d.

En un muestreo aleatorio simple se seleccionan aleatoriamente n objetos de una población y cada objeto tiene la misma probabilidad de ser seleccionado. El valor de la variable aleatoria Y para el objeto i-ésimo seleccionado aleatoriamente se expresa mediante Y_i. Como cada objeto tiene la misma probabilidad de ser seleccionado y la distribución de Y_i es la misma para todo i, las variables aleatorias $Y_1, Y_2, ... Y_n$ son independientes e idénticamente distribuidas (i.i.d.); es decir, la distribución de Y_i es la misma para todo i=1,2,….n e Y_i está independientemente distribuida de $Y_1, .. Y_{i-1}, Y_{i+1} ... Y_n$.

Definición 1.10. Estimadores Máximo-Verosímiles (MV)

Los estimadores MV son los que maximizan la función de verosimilitud de la muestra o, equivalentemente, su logaritmo. Se obtienen igualando a cero los elementos del gradiente. En general, son estimadores consistentes y asintóticamente eficientes.

Definición 1.11 Convergencia en Probabilidad, Consistencia y Ley de los Grandes Números.

La media muestra $\overline{Y}$ converge en probabilidad a μ_Y (o, de forma equivalente, $\overline{Y}$ es consistente) si la probabilidad de que $\overline{Y}$ se encuentre en el rango $\mu_Y - c$ a $\mu_Y + c$ se hace arbitrariamente cercana a 1 cuando n aumenta para cualquier constante c > 0. La convergencia en probabilidad de $\overline{Y}$ a μ_Y se expresa mediante $\overline{Y} \xrightarrow{p} \mu_Y$. La Ley de los Grandes Números establece que si Y_i, i=1,2…n son independientes e idénticamente distribuidas con $E(Y_i) = \mu_Y$ y, si los valores atípicos elevados resultan improbables, entonces $\overline{Y} \xrightarrow{p} \mu_Y$

Definición 1.12. Teorema Cental del Límite

Supongamos que $Y_1 ....Y_n$ son i.i.d. con $E(Y_i) = \mu_Y$ y $\operatorname{var}(Y_i) = \sigma_Y^2$ donde $0 < \sigma_Y^2 < \infty$. A medida que $n \to \infty$, la distribución $(\overline{Y} - \mu_Y)/\sigma_{\overline{Y}}$ (donde $\sigma_{\overline{Y}}^2 = \frac{\sigma_Y^2}{n}$) se aproxima arbitrariamente bien a la distribución Normal estándar.

Definición 1.13. Estimador Consistente y Asintóticamente Eficiente.

Un estimador es consistente si converge en probabilidad al parámetro que se quiere estimar o, lo que es lo mismo, si todos los posibles valores del estimador giran en torno al parámetro que se estima con una dispersión que va haciéndose menor conforme el tamaño muestral crece hasta llegar a ser cero cuando $T \to \infty$. Lo escribimos así: $\hat{\theta} \xrightarrow{p} \theta$.

Suponer que normalizamos multiplicando por $\sqrt{T}$ la diferencia entre el estimador y el valor del parámetro, y que la expresión normalizada converge a una distribución asintótica que tiene una varianza finita diferente de cero. Decimos que el estimador es asintóticamente eficiente si la varianza de la distribución asintótica es menor que la correspondiente a otros estimadores consistentes. Lo escribimos así: $\sqrt{T}\left(\hat{\theta} - \theta\right) \xrightarrow{d} D\left(0, \sigma_{\hat{\theta}}^2\right)$.

Definición 1.14. Contraste de Hipótesis (Teoría)

El contraste de hipótesis es un procedimiento estadístico que sirve para decidir si una hipótesis nula se rechaza o no. El proceso para definir un contraste consta de tres fases:

Primera, definición del estadístico de contraste.

Segunda, Derivación de la distribución de probabilidad del estadístico de contraste bajo la hipótesis nula.

Y tercera, Determinación de la región crítica del contraste utilizando los cuantiles de la distribución obtenida en la fase 2 y el nivel de significación especificado a priori.

Definición 1.15. Función de Potencia y Tamaño del Error Tipo 1.

La función de potencia de un contraste es una función que nos proporciona, para cada valor del parámetro, la probabilidad de rechazar la hipótesis nula. El tamaño del error tipo 1 es el valor que toma la función de potencia para el valor del parámetro que especifica la hipótesis nula. Alternativamente, podemos decir que es la probabilidad de rechazar la hipótesis nula cuando es cierta, es decir, es la probabilidad de tomar una decisión incorrecta cuando los datos son generados por la hipótesis nula. Se dice, también, que es el tamaño del contraste.

Definición 1.16. Función de Potencia y Potencia de un contraste.

La función de potencia de un contraste es una función que nos proporciona, para cada valor del parámetro, la probabilidad de rechazar la hipótesis nula. La potencia de un contraste es el valor que toma la función de potencia para valores del parámetro que caen bajo la hipótesis alternativa. Por lo tanto, la potencia es la probabilidad de tomar una decisión correcta cuando los datos son generados bajo la hipótesis alternativa.

Definición 1.17. Contraste Uniformemente más Potente de tamaño ε (UMP).

Decimos que un contraste es UMP de tamaño ε si cumple

(i) Tener un tamaño igual a ε

(ii) Su función de potencia toma siempre un valor superior a la de cualquier otro contraste que tenga el mismo tamaño.

Podemos sintetizar diciendo que el contraste UMP es aquel que entre todos los contrastes que se equivocan de la misma manera bajo la hipótesis nula, es el que acierta más para todos los valores del parámetro bajo la hipótesis alternativa.

Parte Segunda

Definición 2.1 Los Estimadores MCO, Valores Predichos, y Residuos (Modelo con un regresor)

Los estimadores MCO de los coeficientes son:

$$\hat{\beta}_1 = \frac{\sum_1^n \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_1^n \left(X_i - \overline{X}\right)^2}$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$$

Los valores MCO predichos, $\hat{Y}_i$, y los residuos, $\hat{u}_i$ son

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \quad \text{i=1,2,…n}$$

$$\hat{u}_i = Y_i - \hat{Y}_i \quad \text{i=1,2,…n}$$

Las estimaciones se calculan a partir de n observaciones de las dos variables.

Definición 2.2. Las hipótesis de los MCO. Un solo regresor no estocástico

Para el modelo

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \; i = 1,\ldots, n$$

Las hipótesis son

1. $E\left(u_i\right) = 0, i = 1,2,...,n$
2. Las $Y_i \; i = 1,2,..n \; son \; i.i.d.\left(Esfericidad\right)$
   - $EY_iY_j = 0 \Rightarrow Eu_iu_j \vee i, j, i \neq j \left(Noautocorrelación\right)$
   - $E\left(Y_i - EY_i\right)^2 = \sigma^2 \Rightarrow Eu_i^2 = \sigma^2 \; i = 1,2..n \left(Homocedasticidad\right)$
3. u_i se distribuye : $N\left(0,\sigma^2\right)$ (Normalidad).
4. No hay atípicos en ninguna variable.
5. Para muestras grandes se cumple:

$$\frac{1}{n}\sum_1^n \left(X_i - \overline{X}\right)^2 \rightarrow \sigma_x^2$$

Definición 2.3. Las hipótesis de los MCO. Un solo regresor estocástico. Heterocedasticidad

Para el modelo

$$Y_i = \beta_0 + \beta_1 X_i + u_i,\ i = 1,\ldots, n$$

Las hipótesis son

1.La distribución condicional de u dado X tiene media cero, esto es, E(u|X = x) = 0.

- Esto implica que $\hat{\beta}_1$ es insesgado

2.(Xi,Yi), i =1,…,n, son i.i.d.

- Esto es verdad si (X, Y) se recogen siguiendo el muestreo aleatorio simple

3.Se excluyen grandes atípicos de X y/o Y.

- Técnicamente, X e Y tienen momentos de cuarto orden finitos

Definición 2.4. Las hipótesis de los MCO. k regresores estocásticos. Homocedasticidad

Para el modelo

$$Y_i = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_k X_{ki} + u_i \quad i = 1, 2, ..n$$

Las hipótesis son

1. La media de la distribución condicional de u dadas las X es cero, esto es, $E\left(u_i / X_{1j} = x_{1j} .....X_{kj} = x_{kj}\right) = 0 \quad \vee i, j$.

2. $\left(X_{1i}, ...X_{ki}, Y_i\right) i = 1, 2, ..., n$ son i.i.d.. Es decir, la muestra se obtiene siguiendo el muestreo aleatorio simple (MAS)

3. Se excluyen grandes atípicos de las X y/o Y.

4. $Var\left(u_i / X_{1i}, ..., X_{ki}\right) = \sigma^2 \quad para\ todo\ i$.

5. No hay multicolinealidad exacta.

Definición 2.5. Ecuaciones Normales para un Modelo Lineal con una constante y Dos Regresores.

Para el modelo

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i \quad i = 1, 2, ..n$$

Las ecuaciones normales son:

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_1^n X_{1i} + \hat{\beta}_2 \sum_1^n X_{2i} = \sum_1^n Y_i$$

$$\hat{\beta}_0 \sum_1^n X_{1i} + \hat{\beta}_1 \sum_1^n X_{1i}^2 + \hat{\beta}_2 \sum_1^n X_{1i} X_{2i} = \sum_1^n X_{1i} Y_i$$

$$\hat{\beta}_0 \sum_1^n X_{2i} + \hat{\beta}_1 \sum_1^n X_{1i} X_{2i} + \hat{\beta}_2 \sum_1^n X_{2i}^2 = \sum_1^n X_{2i} Y_i$$

Notar que hay una ecuación normal por cada coeficiente que se estima.

Definición 2.6. Sesgo de Variable Omitida en el modelo con un solo regresor

El sesgo en el estimador MCO consecuencia de la omisión de un factor o variable, se llama sesgo de variable omitida. Para que se de este sesgo la variable omitida “Z” debe satisfacer dos condiciones:

(1) Z es un determinante de Y (i.e. Z es parte de u); y

(2) Z está correlacionada con el regresor X (i.e. corr(Z,X) $\neq$ 0)

Ambas condiciones deben de cumplirse para que la omisión de Z produzca el sesgo de variable omitida.

Definición 2.7. Validez Interna y Externa

- Validez Interna: Las inferencias estadísticas sobre los efectos causales son válidas para la población que se está estudiando.
- Validez Externa: Las inferencias estadísticas pueden ser generalizadas de la población y escenarios estudiados a otras poblaciones y escenarios, en donde escenarios se refiere a los entornos legal, histórico y político

Definición 2.8. Amenazas a la Validez Interna de un estudio de Regresión Múltiple

Existen cinco amenazas principales a la validez interna:

a. Sesgo de Variable Omitida
b. Forma funcional errónea
c. Sesgo por errores de observación.
d. Sesgo por datos ausentes y por selección muestral.
e. Sesgo de causalidad simultánea.

Todos estos implican que $E(u_i|X_{1i},\ldots,X_{ki}) \neq 0$ (o que no se cumple la hipótesis de la independencia en la media condicional) –en cuyo caso los MCO son sesgados e inconsistentes.
Adicionalmente, la presencia de autocorrelación y la heterocedasticidad que llevan a estimaciones incorrectas de los errores estándar constituyen también amenazas a la validez interna.
La aplicación de esta línea de amenazas a un estudio de regresión múltiple constituye un método sistemático de evaluar la validez interna del estudio.

Definición 2.9. Error de Especificación de la Forma Funcional

Este error aparece cuando la forma funcional de la regresión estimada difiere de la forma funcional de la función de regresión poblacional. Si la especificación es incorrecta entonces el estimador del efecto parcial de un cambio en una de las variables será, en general, sesgado.

Definición 2.10. Sesgo por errores en las variables

Este sesgo en los estimadores MCO se produce cuando una variable independiente se mide de forma imprecisa, lo que hace que la variable y el término de error estén correlacionados. Esta correlación hace que el estimador MCO sea sesgado e inconsistente. Este sesgo depende de la naturaleza del error de medida y persiste incluso si el tamaño de la muestra es grande.

Definición 2.11. Sesgo de Selección Muestral

Este sesgo se presenta cuando el proceso de selección de los elementos muestrales influye en le disponibilidad de los datos y el proceso está relacionado con la variable dependiente, además de depender de los regresores. Este proceso de selección induce correlación entre uno o más regresores y el término de error lo que da lugar a que los estimadores MCO sean sesgados e inconsistentes en muestras grandes.

Definición 2.12. Sesgo de Simultaneidad

Aparece en una regresion de Y sobre X, cuando, además del vínculo causal de interés que va desde X a Y, existe un vínculo causal desde Y hacia X. Esta causalidad inversa provoca que el regresor X de un modelo no sea estrictamente exógeno como consecuencia de que ese regresor depende de un grupo de variables estrictamente exógenas y de otros factores correlacionados con la perturbación del modelo. La consecuencia es que el regresor y la perturbación del modelo están correlacionados por lo que el estimador MCO no es consistente. Hay que utilizar estimadores de variable instrumental o estimadores en dos etapas para lograr estimadores consistentes.

Definición 2.13. Factor de Parsimonia del $\bar{R}^2$.

La región crítica de este criterio cuando se comparan dos modelos anidados de k_1 y k_2 regresores, respectivamente, es

$$\bar{R}_1^2 < \bar{R}_2^2$$

Sustituyendo en esta expresión la fórmula $\bar{R}_i^2 = 1 - \frac{T-1}{T-k_i} \frac{\hat{u}_i'\hat{u}_i}{\sum (y_t - \bar{y})^2}$ la región crítica puede escribirse como

$$\hat{u}_1'\hat{u}_1 > \hat{u}_2'\hat{u}_2 \frac{T-k_1}{T-k_2}$$

El factor de penalización es el término que multiplica a la suma de cuadrados de los residuos del modelo amplio, es decir

$$h\left(\bar{R}^2\right) = \frac{T-k_1}{T-k_2}$$

Definición 2.14. Factor de Parsimonia del AIC.

La región crítica de este criterio cuando se comparan dos modelos anidados de k_1 y k_2 regresores, respectivamente, es

$$AIC_1 > AIC_2$$

Sustituyendo en esta expresión la fórmula $AIC_i = \ln \tilde{\sigma}_i^2 + \frac{2k_i}{T}$ la región crítica puede escribirse como

$$\hat{u}_1'\hat{u}_1 > \hat{u}_2'\hat{u}_2 \exp\left(\frac{2(k_2 - k_1)}{T}\right)$$

El factor de penalización es el término que multiplica a la suma de cuadrados de los residuos del modelo amplio, es decir

$$h(AIC) = \exp\left(\frac{2(k_2-k_1)}{T}\right) \approx 1 + \frac{2(k_2-k_1)}{T}$$

Definición 2.15. Factor de Parsimonia del SBIC.

La región crítica de este criterio cuando se comparan dos modelos anidados de $k_1 \; y \; k_2$ regresores, respectivamente, es

$$SBIC_1 > SBIC_2$$

Sustituyendo en esta expresión la fórmula $SBIC_i = \ln \tilde{\sigma}_i^2 + \frac{\ln Tk_i}{T}$ la región crítica puede escribirse como

$$\hat{u}_1'\hat{u}_1 > \hat{u}_2'\hat{u}_2 \exp\left(\frac{\ln T\left(k_2 - k_1\right)}{T}\right)$$

El factor de penalización es el término que multiplica a la suma de cuadrados de los residuos del modelo amplio, es decir

$$h\left(SBIC\right) = \exp\left(\frac{\ln T(k_2 - k_1)}{T}\right) \approx \frac{\ln T\left(k_2 - k_1\right)}{T}$$

Definición 2.16. Contraste de la Razón de Verosimilitud (LR).

Suponemos que la función de verosimilitud depende de un vector de k parámetros que llamamos θ. Se trata de contrastar que los elementos de este vector cumplen r restricciones. El estadístico del contraste LR es

$$LR = 2\left[l\left(\tilde{\theta}\right) - l\left(\tilde{\theta}_R\right)\right]$$

En donde $l\left(\tilde{\theta}\right)$ es el valor que toma el logaritmo de la función de verosimilitud sustituyendo los k parámetro por sus estimadores MV sin restricciones; $l\left(\tilde{\theta}_R\right)$ es el valor que toma el logaritmo cuando los parámetros se sustituyen por los estimadores MV con restricciones. La distribución de probabilidad de este estadístico bajo la hipótesis nula es

$$LR \approx \chi^2(r)$$

En donde $\chi^2(r)$ es una chi-cuadrado con r grados de libertad. La región crítica del contraste viene dada por

$$LR > \chi_\varepsilon^2(r)$$

En donde ε es el nivel de significación adoptado a priori.

Definición 2.17. Contraste de los Multiplicadores de Lagrange (LM).

Suponemos que la función de verosimilitud depende de un vector de k parámetros que llamamos θ. Se trata de contrastar que los elementos de este vector cumplen r restricciones. El estadístico del contraste LM es

$$LM = d\left(\tilde{\theta}_R\right)' I\left(\tilde{\theta}_R\right)^{-1} d\left(\tilde{\theta}_R\right)$$

En donde $d\left(\tilde{\theta}_R\right)$ e $I\left(\tilde{\theta}_R\right)$ son, respectivamente, el gradiente y la matriz de información, ambos evaluados con los estimadores MV con restricciones.

La distribución de probabilidad de este estadístico bajo la hipótesis nula es

$$LM \approx \chi^2(r)$$

En donde $\chi^2(r)$ es una chi-cuadrado con r grados de libertad. La región crítica del contraste viene dada por

$$LM > \chi^2_\varepsilon(r)$$

En donde ε es el nivel de significación adoptado a priori.

Parte Tercera

Definición 3.1. Estacionariedad

Una serie temporal es estacionaria si su distribución de probabilidad no cambia en el tiempo; es decir, si la distribución conjunta de T elementos del proceso que comienza en el periodo s no depende de s. La definición se extiende a dos o más series temporales diciendo que son conjuntamente estacionarias. La estacionariedad requiere que el futuro sea como el pasado, al menos en probabilidad.

Definición 3.2. Proceso Estacionario en sentido Débil.

Decimos que una variable es estacionaria en sentido débil cuando se cumple lo siguiente

$$Ey_t = \mu \quad \forall t$$

$$Var(y_t) = \sigma^2 \quad \forall t$$

$$Cov(y_t, y_s) = \gamma_{t-s} \quad \forall t, s$$

Lo que significa esta caracterización de la estacionariedad es que todos los elementos del proceso, con independencia del periodo temporal en que nos encontremos, giran en torno a un mismo valor, con una dispersión en torno a ese valor medio que no cambia con el tiempo y con una dependencia temporal que tampoco cambia.

Definición 3.3. Variable Integrada de orden 1

Se dice que una serie es integrada de orden 1, y la denotaremos por yt ~ I(1), cuando no tiene ningún componente determinista y después de ser diferenciada una vez resulta una representación ARMA estacionaria e invertible.

Definición 3.4. Tendencia Estocástica

La tendencia estocástica es la suma de tantos ruidos blancos como indica el subíndice del elemento del proceso, es decir, $y_t = \sum_{i=1}^{t} \varepsilon_i$, en donde ε_i , i=1,2…t son ruidos blancos. Equivale a decir que el proceso es un paseo aleatorio sin deriva y con un valor de origen igual a cero; es decir

$$y_t = y_{t-1} + \varepsilon_t \qquad y_0 = 0$$

Definición 3.5. Contraste de Dickey y Fuller.

Es un procedimiento que se utiliza para rechazar, en su caso, la hipótesis nula de que la serie tiene una raíz unitaria o, equivalentemente, que la serie es integrada de orden 1. El estadístico de contraste es el t-ratio que se define como un cociente entre el estimador MCO del coeficiente de la variable retardada un periodo y la desviación típica de ese estimador. Lo forma que adoptan estos estimadores depende de la forma que toman los elementos deterministas del modelo. La distribución de probabilidad de ese estadístico no coincide con la t de Student.

Definición 3.6. Modelo CRD. Supuestos con Exogeneidad Estricta

Para el modelo

$$Y_t = \beta_0 + \beta_1 X_t + ... + \beta_{r-1} X_{t-r} + u_t$$

los supuestos son

1) $E\left(u_t / ...., X_{t-1}, X_t, X_{t+1}, ...\right) = 0$
2) (a) Y y X son estacionarias.
   (b) $\left(Y_t, X_t\right) y \left(Y_{t-j}, X_{t-j}\right)$ son independientes conforme j se hace grande.

3) Y y X tienen momentos de orden 8 diferentes de cero.

4) No hay multicolinealidad perfecta.

Definición 3.7. Modelo VAR(p).

El modelo VAR (p) es un modelo en el que todas las variables consideradas son función líneal de los p valores pasados de las variables. Para el caso de dos variables el modelo VAR(2) sería el siguiente:

$$y_t = \phi_{111} y_{t-1} + \phi_{112} y_{t-2} + \phi_{121} x_{t-1} + \phi_{122} x_{t-2} + u_{1t}$$

$$x_t = \phi_{211} y_{t-1} + \phi_{212} y_{t-2} + \phi_{221} x_{t-1} + \phi_{222} x_{t-2} + u_{2t}$$

u_{1t} *y* u_{2t} son perturbaciones aleatorias.

Definición 3.8 No causalidad en el sentido de Granger.

Decimos que x_t no causa a y_t en el sentido de Granger si la varianza del error de predicción de y_t, un periodo hacia delante, es la misma cuando se utiliza el pasado de la propia variable que cuando se utiliza el pasado de las dos variables. Dicho de otra manera, x_t no causa a y_t en el sentido de Granger, si la predicción de y_t no mejora si se añade el pasado de x_t.

Definición 3.9 Cointegración.

Se dice que dos variables están cointegradas cuando cumplen las dos condiciones siguientes: Primero, las dos variables tienen el mismo orden de integración; segundo, se puede encontrar una combinación lineal de las dos variables tal que el residuo resultante tenga un orden de integración inferior al de las dos variables. Se dice también que dos variables están cointegradas cuando la tendencia estocástica de una de ellas es explicada por la tendencia estocástica de la otra.

Definición 3.10 Contrastes de cointegración

Son procedimientos propuestos para contrastar la hipótesis nula de no cointegración frente a la alternativa de existencia de cointegración. Hay dos grupos de contrastes: uniecuacionales y multiecuacionales. Los primeros se basan en los residuos MCO de la relación de cointegración. Dentro del primer grupo, los más conocidos son el contraste

CRDW y el contraste Dickey-Fuller aplicado a los residuos. En el segundo grupo el más conocido es el contraste de Johansen que es el procedimiento de la razón de verosimilitud aplicado al modelo VAR definido para todas las variables.
El estadístico de contraste del CRDW es el estadístico del contraste Durbin-Watson para contrastar la autocorrelación. La distribución de probabilidad bajo la hipótesis nula es diferente a la obtenida por Durbin y Watson debido a la no estacionariedad bajo la hipótesis nula. Por último, la región crítica del contraste esta definida por aquellos valores del estadístico CRDW que superan un valor que depende del número de variables incluidas en la relación de cointegración y del nivel de significación adoptado previamente.

Definición 3.11. Correlación Espuria

La correlación espuria se refiere a aquellas situaciones en las que un uso incorrecto de ciertas técnicas estadísticas puede llevar a la conclusión de que dos variables están relacionadas entre sí cuando realmente no lo están. Por eso se dice que "correlación no es causación" o que "la correlación está en los datos y la causación en la mente". Si, en el marco no estacionario, tenemos dos variables que son I(1) y han sido generadas independientemente una de la otra y hacemos la regresión con MCO entonces tanto el t-ratio como el coeficiente de determinación nos van a decir que las dos variables están fuertemente relacionadas cuando sabemos que no lo están. El error, en este caso, reside en no haber contrastado previamente la hipótesis nula de no cointegración. Si hubiéramos prestado atención a los residuos, la conclusión habría sido muy diferente.

Definición 3.12. Modelo con Mecanismo de Corrección del Error (MCE)

El MCE es la versión restringida del modelo VAR (p) de variables en niveles que están cointegradas. Incorporando en el modelo VAR las restricciones asociadas con la cointegración se llega al MCE. En este modelo, cada uno de los incrementos contemporáneos de las variables se explica en función de los retardos de esos incrementos y del residuo de la relación de cointegración retardado un periodo. Para el caso de dos variables, el modelo VAR (2) en niveles toma la forma siguiente:

$$y_t = \phi_{111} y_{t-1} + \phi_{112} y_{t-2} + \phi_{121} x_{t-1} + \phi_{122} x_{t-2} + u_{1t}$$
$$x_t = \phi_{211} y_{t-1} + \phi_{212} y_{t-2} + \phi_{221} x_{t-1} + \phi_{222} x_{t-2} + u_{2t}$$

El correspondiente MCE es

$$\Delta y_t = \alpha_1 \left(y_{t-1} - \beta x_{t-1} \right) + \gamma_{111} \Delta y_{t-1} + \gamma_{121} \Delta x_{t-1} + v_{1t}$$
$$\Delta x_t = \alpha_2 \left(y_{t-1} - \beta x_{t-1} \right) + \gamma_{211} \Delta y_{t-1} + \gamma_{221} \Delta x_{t-1} + v_{2t}$$

Parte Cuarta

Definición 4.1 Datos Panel. Ejemplo.

Los datos de panel consisten en observaciones sobre las mismas n entidades individuales para dos o más periodos de tiempo, T. Si el conjunto de datos consta de las observaciones sobre las variables X e Y, entonces los datos se expresan como

$$\left(X_{it}, Y_{it} \right), i = 1,..,n \ y \ t = 1,..T$$

donde el primer subíndice, i, se refiere a la entidad individual que está siendo observada y el segundo subíndice, t, se refiere al periodo en el que se observa. El ejemplo queda abierto; cada uno elige el que le parece más ilustrativo.

Definición 4.2. Datos Panel. Modelo de Regresión de Efectos Fijos.

El modelo de Regresión de Efectos Fijos es

$$Y_{it} = \alpha_i + \beta_1 X_{1,it} + ... + \beta_k X_{k,it} + u_{it}$$

donde i = 1,..,n; t = 1,…T; $X_{1,it}$ es el valor del primer regresor para la entidad individual i en el periodo de tiempo t, $X_{2,it}$ es el valor del segundo regresor, y así sucesivamente; y $\alpha_1,....\alpha_n$ son los términos constantes específicos de cada entidad individual.

De manera equivalente, el modelo de regresión de efectos fijos puede expresarse en términos de una constante común, las X y n-1 variables binarias que representan a todas las entidades individuales excepto a una

$$Y_{it} = \beta_0 + \beta_1 X_{1,it} + ... + \beta_k X_{k,it} + \gamma_2 D2_i + ... + \gamma_n Dn_i + u_{it}$$

donde $D2_i = 1$ si i = 2 y $D2_i = 0$ en caso contrario, y así sucesivamente.

Definición 4.3. Datos Panel. Estimación MCO con variables en desviaciones.

Es un método de estimación aplicable al Modelo de Efectos Fijos. Se define utilizando un proceso en dos etapas. En la primera, se resta a cada observación la media temporal específica a cada entidad individual. En la segunda, se estiman los parámetros aplicando los MCO a las variables en desviaciones.

Considerando el Modelo de Efectos Fijos con un solo regresor,

$$Y_{it} = \alpha_i + \beta X_{it} + u_{it}$$

las etapas serían

1ª Etapa: Se calculan las medias individuales para cada entidad individual, $\bar{Y}_i = \frac{1}{T}\sum Y_{it}$, $\bar{X}_i = \frac{1}{T}\sum X_{it}$ y, a continuación, se obtienen las desviaciones, $\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$, $\tilde{X}_{it} = X_{it} - \bar{X}_i$.

2ª Etapa, Se aplican los MCO a estas desviaciones

$$\hat{\beta} = \frac{\sum_i^n \sum_t^T \tilde{X}_{it}\tilde{Y}_{it}}{\sum_i^n \sum_t^T \tilde{X}_{it}^2}$$

Definición 4.4 Variable Instrumental.

Es una variable correlacionada con un regresor endógeno (relevancia del instrumento) e incorrelacionada con el término de error de la regresión (exogeneidad del instrumento). La variable instrumental se utiliza cuando en un modelo el regresor no es estrictamente exógeno y, como consecuencia, el estimador MCO no es consistente.

Definición 4.5 Estimador de Variable Instrumental

Este estimador se propone para aquellos casos en que el regresor (x_t) no es estrictamente exógeno. Se parte de una variable, que es la variable instrumental (z_t), que está correlacionada con el regresor pero incorrelacionada con la perturbación aleatoria del modelo y se define la ecuación normal que puede escribirse como

$$\sum z_t y_t = \hat{\beta}_{VI} \sum z_t x_t$$

A partir de esta relación se despeja el estimador de variable instrumental que es un estimador consistente.

Definición 4.6 Estimador en dos etapas.

Este estimador se propone para aquellos casos en que el regresor (X_t) no es estrictamente exógeno debido a que puede escribirse como una función lineal de un grupo de variables estrictamente exógenas y de un perturbación que está correlacionada con la perturbación del modelo original. Esta correlación provoca que el estimador MCO sea inconsistente. El estimador en dos etapas es un estimador consistente. Se define así:

En la primera etapa, se hace la regresión del regresor sobre el grupo de variables estrictamente exógenas de las que depende y se define el regresor estimado.

En la segunda etapa, se sustituye el regresor original por el regresor estimado y se aplican MCO. El estimador resultante es el estimador en dos etapas y es consistente.

Definición 4.7 Contraste de Hausman

Es un procedimiento propuesto para contrastar la hipótesis nula de exogeneidad estricta. El estadístico del contraste es

$$m = \frac{\hat{q}^2}{\widehat{Var}\left(\hat{q}\right)}$$

El numerador es el cuadrado de $\hat{q} = \hat{\beta}_1 - \hat{\beta}_0$ en donde $\hat{\beta}_0$ es un estimador del parámetro del modelo que es consistente y asintóticamente eficiente bajo la hipótesis nula pero es inconsistente bajo la hipótesis alternativa y $\hat{\beta}_1$ es un estimador consistente bajo ambas hipótesis pero no es eficiente bajo la hipótesis nula.

La distribución de probabilidad de m bajo la hipótesis nula es una chi-cuadrado con un grado de libertad.

La región crítica del contraste viene dada por

$$m > \chi^2_\varepsilon(1)$$

En donde ε es el nivel de significación fijado a priori.