

CAPÍTULO 4

ANÁLISIS UNIVARIANTE. COINTEGRACIÓN.

4.1. INTRODUCCIÓN

Suponemos que a partir del marco teórico se han especificado una serie de relaciones y un conjunto de variables. También suponemos que existen fuentes estadísticas a partir de las cuales se dispone de datos que reflejan la evolución seguida por las variables.

Una vez elaborado el banco de datos, el siguiente paso es caracterizar individualmente a cada una de las variables. Esto requiere, en primer lugar, establecer una tipología de series y, en segundo lugar, desarrollar un grupo de técnicas mediante las cuales sea posible asociar una serie cualquiera a uno de los elementos de la tipología.

Esta caracterización univariante de los modelos es importante por, al menos, dos tipos de razones:

- Condiciona la forma en que se especifican los modelos.
- Las propiedades de los estadísticos que sirven para estimar los parámetros y contrastar hipótesis sobre los mismos varían según sea el tipo de variable que se utiliza en su definición.

En la primera parte de la Sección 4.2 se van a contemplar los diferentes tipos de series prestando especial atención a una tipología de series no estacionarias. Los tipos estudiados no pretenden ser exhaustivos aunque sí cubren una amplia gama de las variables que uno encuentra en el trabajo aplicado. La segunda parte de esta sección esta dedicada a estudiar tres tipos de instrumentos: no paramétricos, atípicos y análisis paramétrico.

La Sección 4.3 está dedicada a estudiar algunos de los procedimientos propuestos en la literatura para contrastar el orden de integración de una variable. Estos procedimientos son muy numerosos por lo que limitaremos nuestra atención a los más

utilizados.

El Capítulo termina definiendo la cointegración y presentando los contrastes que se han propuesto para su determinación, prestando una atención especial al método de Johansen.

4.2 ANÁLISIS UNIVARIANTE.

Comenzaremos esta sección estableciendo una tipología de series, distinguiendo entre series estacionarias y no estacionarias. Entre las no estacionarias distinguiremos aquellas que tienen tendencia determinista, las que tienen tendencia estocástica y las que tienen ambos tipos de tendencia.

Serie Estacionaria.

Una serie estacionaria es la realización observada de un proceso estocástico estacionario. Decimos que y_t es un proceso estocástico estacionario en sentido débil cuando:

$$\begin{aligned} E y_t &= \mu \quad \forall t \\ \text{Var}(y_t) &= \sigma^2 \quad \forall t \\ \text{Cov}(y_t, y_{t-n}) &= \gamma_n \quad \forall t \end{aligned}$$

Por tanto, un proceso estocástico es estacionario en sentido débil cuando la media y la varianza son las mismas en todas las observaciones y la autocovarianza sólo depende de la distancia entre el par de elementos del proceso considerados.

Serie No Estacionaria Con Tendencia Determinista.

Una serie no estacionaria con tendencia determinista es aquella cuyo proceso generador de datos viene dado por:

$$y_t = \alpha + \beta t + u_t \quad (4.1)$$

en donde:

$$u_t = \Psi(L)\varepsilon_t \quad (4.2)$$

con $\sum_{j=0}^{\infty} |\Psi_j| < \infty$, las raíces del polinomio $\Psi(L)=0$ están fuera del círculo unitario y $\{\varepsilon_t\}$ es un ruido blanco con media cero y varianza σ^2 .

Los momentos de la distribución de probabilidad de y_t vienen dados por:

$$E y_t = \alpha + \beta t \quad (4.3)$$

$$\text{Var}(y_t) = \sigma_u^2 \quad (4.4)$$

teniendo en cuenta que u_t es estacionario se ve que la no estacionariedad de y_t queda reflejada en la esperanza matemática. El resultado se puede generalizar, suponiendo que $y_t = f(t) + u_t$, siendo $f(\cdot)$ cualquier función del tiempo.

Serie No Estacionaria Con Tendencia Estocástica.

Una serie no estacionaria con tendencia estocástica es aquella cuyo proceso generador de datos viene dado por:

$$y_t = y_{t-1} + u_t \quad (4.5)$$

en donde u_t ha sido definido en (4.2).

Una forma equivalente de (4.5) es:

$$y_t = y_0 + \sum_{i=1}^t u_i \quad (4.6)$$

En este caso los momentos de la distribución de probabilidad vienen dados por:

$$E y_t = y_0 \quad (4.7)$$

$$\text{Var}(y_t) = \sigma_u^2 t \quad (4.8)$$

Suponiendo que y_0 no es estocástico. Notar que la no estacionariedad se refleja en la varianza.

Serie No Estacionaria Con Ambas Tendencias.

La serie no estacionaria con ambas tendencias es la realización de un proceso estocástico no estacionario cuyo proceso generador de datos es:

$$y_t = \alpha + y_{t-1} + u_t \quad (4.9)$$

Una expresión equivalente de (4.9) es la siguiente:

$$y_t = y_0 + \alpha t + \sum_{i=1}^t u_i \quad (4.10)$$

Cuyos momentos son:

$$E y_t = y_0 + \alpha t \quad (4.11)$$

$$\text{Var}(y_t) = \sigma_u^2 t \quad (4.12)$$

Destacar que la no estacionariedad se manifiesta tanto en la esperanza como en la varianza.

Establecidos los tipos de series pasemos ahora a examinar los instrumentos disponibles para determinar en cada caso concreto si una serie es de un tipo u otro.

Los procedimientos propuestos van desde el simple análisis gráfico de la serie individual y, en su caso, de algunos de las posibles transformaciones, hasta procesos más sofisticados de análisis de series temporales en los que se incluyen instrumentos como el correlograma, la función de autocorrelación parcial y los contrastes paramétricos para determinar el orden de integración de una variable.

En esta sección vamos a prestar atención a tres enfoques:

- 1) Enfoque No Paramétrico.
- 2) Análisis de Atípicos.
- 3) Enfoque Paramétrico.

1) Enfoque No Paramétrico.

El enfoque no paramétrico está basado, por un lado, en la visualización del

gráfico de la serie original y, en su caso, de alguna de sus transformaciones como la primera y segunda diferencias, y, por otro, en el uso de ciertos estadísticos como el de Daniel y el de Kruskal-Wallis.

Nosotros vamos a prestar atención al enfoque gráfico. El tratamiento de los estadísticos mencionados puede verse en el Capítulo 2 de Aznar y Trívez (1993).

A priori, las pautas que cabría esperar del análisis gráfico serían las que aparecen en el Cuadro 4.1

A continuación, vamos a presentar los gráficos de algunas series.

Las Figuras 4.1 y 4.2 se refieren a series estacionarias, la segunda muy próxima a la no estacionariedad. La perturbación u_t es un ruido blanco con media cero y varianza igual a 0,01.

Las seis figuras siguientes recogen los gráficos de tres series no estacionarias sin deriva y con diferentes esquemas autorregresivos en la parte estacionaria. En la Figura 4.3 la perturbación es el mismo ruido blanco de las Figuras 4.1 y 4.2; en la Figura 4.5 la perturbación es el proceso representado en 4.1 y, en la Figura 4.7, el proceso de la perturbación es el representado en la Figura 4.2.

Un hecho a destacar es la presencia de importantes subperiodos con la misma pauta de crecimiento o decrecimiento. Es como si existieran tendencias segmentadas a lo largo de todo el periodo. Las primeras diferencias en todos los casos giran en torno a una constante que, en este caso, es cero por ser este valor la media de todos los procesos estacionarios que corresponden a la perturbación aleatoria.

CUADRO 4.1. Pautas Esperadas a priori del Analisis Gráfico.

<u>SERIE</u>	<u>NIVELES</u>	<u>PRIMERA DIFERENCIA</u>
<u>Estacionaria</u>	Los valores giran en torno a una constante según la estructura autorregresiva de u_t .	Los valores giran en torno a una constante.
	La dispersión es constante.	La dispersión es constante y mayor que la de niveles.
<hr/>		
<u>No Estacionaria</u>		
Tendencia Determinista	Los valores giran en torno a una función del tiempo.	Los valores giran en torno a una constante.
	La dispersión es constante	La dispersión es constante pero mayor que la de niveles.
Tendencia Estocástica	Los valores giran en torno a una constante, aunque pueden aparecer subperiodos grandes con pauta creciente o decreciente.	Los valores giran en torno a una constante.
	La dispersión es creciente	La dispersión es constante y menor que la de niveles.
Ambas Tendencias	Los valores giran en torno a una función lineal del tiempo.	Los valores giran en torno a una constante.
	La dispersión es creciente	La dispersión es constante y menor que la de niveles.

FIGURA 4.1. $y_t = 0.5y_{t-1} + u_t$

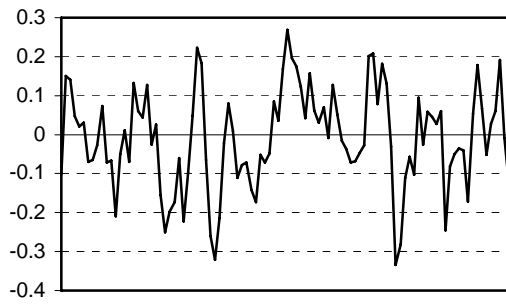


FIGURA 4.2. $y_t = 0.9y_{t-1} + u_t$

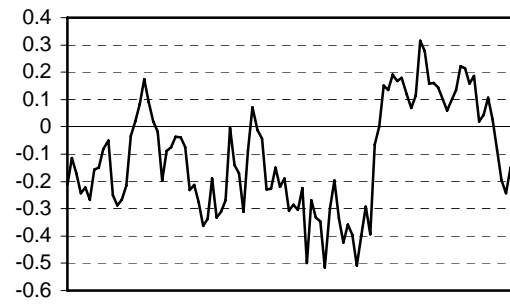


FIGURA 4.3. $y_t = y_{t-1} + u_t$

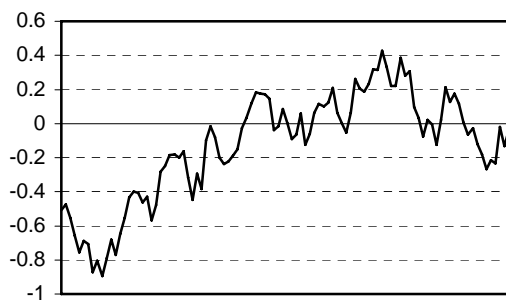


FIGURA 4.4. Primeras Diferencias

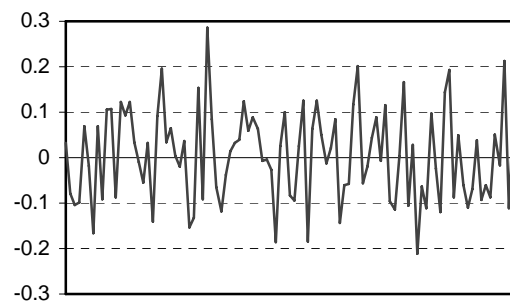


FIGURA 4.5. $y_t = y_{t-1} + z_t$

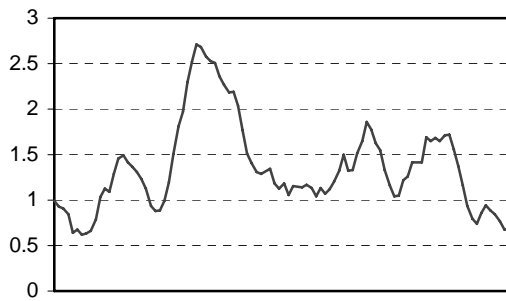


FIGURA 4.6. Primeras Diferencias

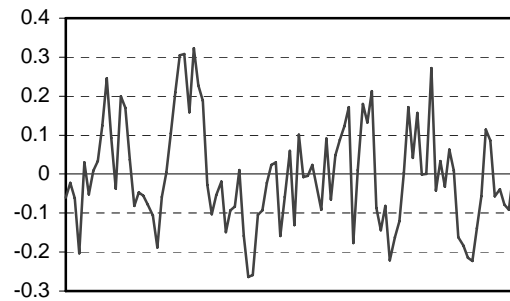


FIGURA 4.7. $y_t = y_{t-1} + x_t$

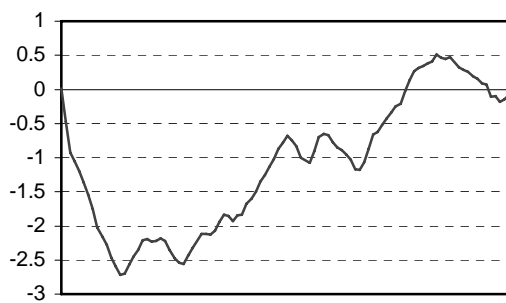


FIGURA 4.8. Primeras Diferencias

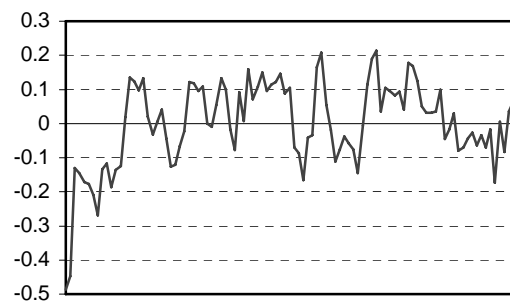


FIGURA 4.9. $y_t = 1 + y_{t-1} + u_t$

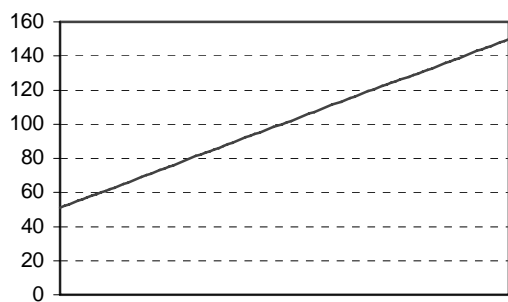


FIGURA 4.10. Primeras Diferencias

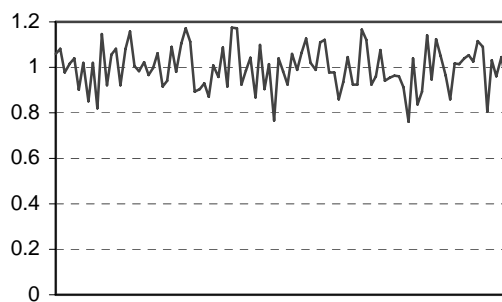


FIGURA 4.11. $y_t = 1 + y_{t-1} + z_t$

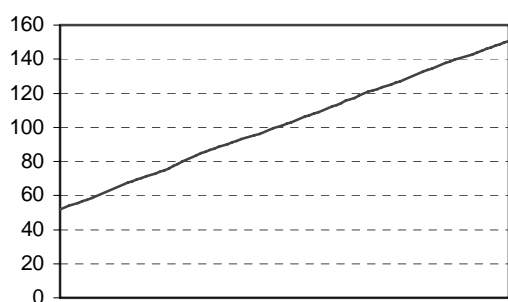


FIGURA 4.12. Primeras Diferencias

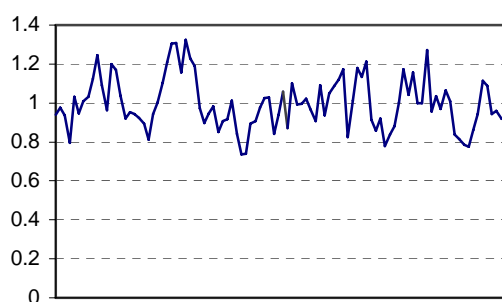


FIGURA 4.13. $y_t = 1 + y_{t-1} + x_t$

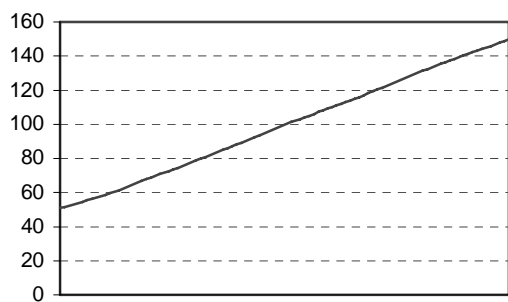


FIGURA 4.14. Primeras Diferencias

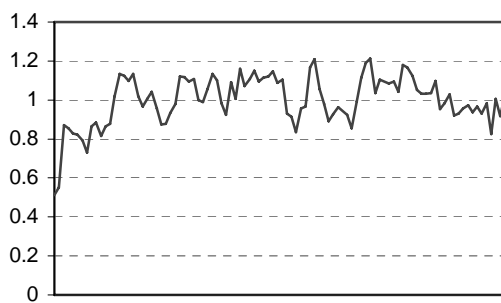


FIGURA 4.15. $y_t = 1 + y_{t-1} + u_t$

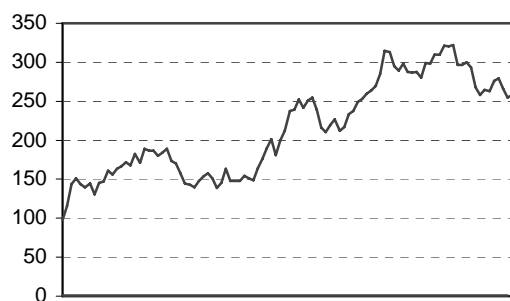
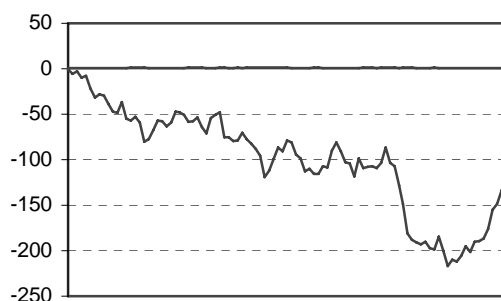


FIGURA 4.16. $y_t = y_{t-1} + u_t$



Las siguientes seis figuras corresponden a un proceso no estacionario que tienen los dos tipos de tendencia. En los tres casos se ve que la serie mantiene una pauta constante de variación -en este caso creciente- que se mantiene a lo largo de todo el periodo. Las perturbaciones son las mismas que las comentadas para las tres series anteriores. Las primeras diferencias de las tres series giran en torno a la unidad que es la deriva del proceso.

De todas formas, las ilustraciones anteriores son ejemplos estándar y no nos deberían hacer pensar en una regla mecánica de clasificación. Se ha visto que en el caso sin deriva la forma que adoptaba la serie dependía de la estructura de la perturbación aleatoria. En el caso de series con deriva, las diferencias parecían ser menores pero eso no debe llevarnos a engaño a la hora de clasificar una serie como de un tipo u otro. Este escepticismo se justifica observando los gráficos 4.15 y 4.16. El primero, representa el mismo proceso que el de la Figura 4.9 pero suponiendo que la varianza de la perturbación es 100 en lugar de 0,01. Se ve como se sigue manteniendo una pauta de crecimiento pero no tan clara como era en la Figura 4.9. De hecho, podrían distinguirse segmentos de tendencia e, incluso, intervalos con variaciones negativas.

La Figura 4.16 representa un proceso como el que corresponde a la Figura 4.3. En este caso las variaciones no son tan relevantes porque el carácter descendente o ascendente puede deberse al tramo considerado de toda la realización potencial del proceso.

Pese a este espíritu escéptico alguna conclusión puede extraerse. Cuando la pauta de crecimiento o decrecimiento es constante e inequívoca como en las Figuras 4.9, 4.11 y 4.13, claramente apunta a un proceso con deriva. Cuando se observa un gráfico con tendencias segmentadas con cambios más o menos rápidos en el signo del cambio experimentado entonces la conclusión es más difícil; o bien puede explicarse mediante un proceso sin deriva pero con una estructura particular del proceso estacionario del modelo o bien con un proceso con deriva pero con una varianza muy grande de la perturbación.

2) Análisis de Atípicos

Para detectar observaciones atípicas en el marco del análisis univariante se han propuesto diversos métodos en la literatura que ya se incorporan en algunos paquetes de ordenador como son el SCA y el TRAMO. Pero todos estos procedimientos requieren una identificación de un modelo en el origen que, a su vez, está condicionada por la presencia de las observaciones atípicas. En nuestra opinión, es interesante disponer de un procedimiento que no requiera la especificación de ninguna hipótesis a priori para poder detectar las observaciones que escapan a la pauta normal.

Además, este análisis hay que llevarlo a cabo al comienzo de todo el ejercicio de la caracterización del comportamiento individual de las series porque la presencia de estas observaciones atípicas puede afectar significativamente la aplicación de técnicas posteriores en el proceso. Ver, por ejemplo, Trívez (1994).

Nuestro punto de partida son las T observaciones de una variable. Distinguiremos varios casos según sea o no estacionaria la serie y si tiene o no componente estacional.

Comenzando con una serie estacionaria sin componente estacional, el procedimiento que se propone consta de los siguientes pasos:

1. Se ordenan de menor a mayor valor las observaciones de la serie.
2. Se eliminan el 10% de las observaciones; el 5% de las mayores y el 5% de las menores.
3. Con el resto de las observaciones se calcula la media (μ) y la desviación típica (σ) y se forma el intervalo: $\mu \pm c.\sigma$, en donde c es una constante que, en general, supondremos que es un valor entre 3 y 4.
4. Se toma cada una de las observaciones eliminadas en 2) y se decide que una de esas observaciones es atípica cuando su valor cae fuera del intervalo calculado en 3).
5. Para cada observación calificada como atípica su valor se sustituye por la media de los valores anterior y posterior que no sean atípicos o cualquier otro procedimiento ad hoc utilizado en la literatura.

6. La serie corregida se obtiene de la siguiente forma: las observaciones que no son atípicas mantienen su valor y las que son atípicas lo cambian por el valor calculado en 5).

Si la serie es estacionaria con componente estacional se aplica el mismo proceso ya comentado pero introduciendo las siguientes variaciones:

E-1. Previo a la ordenación en 1) se resta a cada observación la media de la estación a la que pertenece.

E-2. A la hora de obtener la serie corregida, al valor obtenido en 6) se le suma la media de la estación a la que pertenece dicha observación calculando la media de cada estación utilizando sólo las observaciones no calificadas como atípicas.

Si la serie no es estacionaria se aplica el procedimiento descrito para la serie estacionaria después de aplicar a la serie original las transformaciones necesarias para convertirla en una serie estacionaria.

Veamos ahora cómo aplicar este procedimiento a una serie no estacionaria con una observación atípica tipo aditivo. Los datos aparecen en el Cuadro 4.2.

CUADRO 4.2. Serie No Estacionaria-Atípico Aditivo

Observación	serie: y_t	$\Delta y_t = y_t - y_{t-1}$	y_t corregida
1	2		2
2	4	2	4
3	7	3	7
4	8	1	8
5	9	1	9
6	12	3	12
7	25	13	15
8	18	-7	18
9	20	2	20
10	24	4	24
11	28	4	28

En la primera columna aparece el número de la observación; en la segunda, los valores de la serie; en la tercera, los valores de las primeras diferencias, apareciendo al final la media de esas primeras diferencias; el mayor valor absoluto corresponde a las

observaciones 7 y 8 y éstas son las que se eliminan. Con las 8 restantes se calculan la media y la varianza que resultan ser: $\mu = 2,5$ y $\sigma = 1,11$.

A partir de estos valores, los extremos del intervalo son: $\mu \pm 3\sigma = (-0,83; 5,83)$. Comparando los valores de las observaciones excluidas con los extremos del intervalo se concluye con que las dos observaciones, 7 y 8, son atípicas.

La serie corregida se obtiene modificando los mínimos valores posibles de la serie original de forma que los incrementos resultantes queden dentro de los extremos del intervalo definido como estándar.

Hay que destacar que el procedimiento descrito sirve para un atípico tipo aditivo. Se ha visto que en la serie diferenciada genera dos valores atípicos de signo contrario.

Podría pensarse también en procedimientos para identificar atípicos tipo escalón. En este caso, el carácter distintivo es que en la serie diferenciada sólo se observa un atípico. Podría completarse el análisis con un cálculo de medias para diferentes subperiodos calculando series diferenciadas con diferentes retardos.

En nuestra opinión, la aplicación de este procedimiento permitirá detectar las observaciones atípicas más significativas que, de no ser tratadas adecuadamente en el punto de partida podían condicionar decisivamente, en las etapas posteriores, la aplicación de otros contrastes utilizados para describir los rasgos básicos de la estructura univariante de cada serie.

3) Análisis Paramétrico de Series Temporales

Por Análisis Paramétrico de una serie se entiende lo que en la metodología Box-Jenkins corresponde a la etapa de Identificación, especialmente lo que se refiere al Análisis de Estacionariedad. Se sigue prestando atención a la visualización del gráfico de la serie pero, ahora, se añaden nuevos instrumentos como son el correlograma y el gráfico de la Función de Autocorrelación Parcial. Es bien conocido que a una serie no estacionaria le corresponde un correlograma que se amortigua muy lentamente y una función de autocorrelación parcial con el primer valor próximo a la unidad.

Todo el procedimiento está descrito en el Capítulo 10 de Aznar y Trávez (1993) por lo que aquí no insistiremos sobre el tema.

4.3. CONTRASTE DEL ORDEN DE INTEGRACIÓN.

El Análisis del Orden de Integración se refiere a procedimientos alternativos a los ya comentados para determinar si una serie es o no estacionaria. Estos nuevos procedimientos se han desarrollado dentro de una gran corriente asociada con el tratamiento de modelos con variables no estacionarias. Toda esta literatura arranca a partir de los trabajos de Granger y Newbold (1974, 1977), Fuller (1976), Dickey y Fuller (1979, 1981) y Nelson y Plosser (1982). Un estudio comparado reciente, bastante comprensivo, puede verse en Maddala y Kim (1998).

Comenzaremos introduciendo el concepto del orden de integración de una serie tomando la definición propuesta por Engle y Granger (1987):

Definición 4.1: "Se dice que una serie es integrada de orden d , y la denotaremos por $y_t \sim I(d)$, cuando no tiene ningún componente determinista y después de ser diferenciada d veces resulta una representación ARMA estacionaria e invertible".

Una forma alternativa, aunque equivalente, de dar cuenta de este concepto es la que proporciona Johansen (1995) en las dos definiciones siguientes:

Definición 4.2: El proceso lineal, $y_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}$ en donde ε_t es un ruido blanco, se dice que es $I(0)$ si $\Psi(1) = \sum \Psi_i \neq 0$.

Definición 4.3: Un proceso estocástico, y_t , se dice que es integrado de orden $d, I(d)$, si $\Delta^d(y_t - E(y_t))$ es $I(0)$.

Supongamos que se propone que la serie ha sido generada por el siguiente modelo:

$$y_t = \phi y_{t-1} + u_t \quad (4.13)$$

en donde u_t es un ruido blanco. En el caso en que $\phi = 1$, entonces y_t es un paseo

aleatorio y la serie no es estacionaria. Entonces parece lógico pensar en contrastar la siguiente hipótesis nula:

$$H_0: \phi = 1 \quad (4.14)$$

frente a la alternativa:

$$H_1: |\phi| < 1 \quad (4.15)$$

Si se acepta la hipótesis nula se concluye con que la serie no es estacionaria, o equivalente, que es integrada de orden 1, $I(1)$. Si se rechaza la hipótesis nula entonces se acepta que la serie es estacionaria, es decir, $I(0)$.

Para contrastar esta hipótesis nula se propone el t-ratio definido como:

$(\hat{\phi} - 1) / \hat{\sigma}_{\hat{\phi}}$ en donde $\hat{\phi}$ y $\hat{\sigma}_{\hat{\phi}}$ son los estimadores MCO de los correspondientes parámetros ϕ y σ_{ϕ} . Pero la utilización de este estadístico plantea los siguientes problemas:

1°. Bajo la hipótesis nula, el estadístico t-ratio no sigue una distribución t - Student por lo que se hace necesario obtener nuevos puntos críticos diferentes a los habitualmente utilizados.

2°. Tanto en lo que respecta al PGD como al modelo que se utiliza para llevar a cabo el contraste es posible pensar en especificaciones alternativas a la considerada llegándose a numerosos escenarios que hay que tener en cuenta a la hora de diseñar una estrategia de contraste que sea operativa.

Considerar los siguientes tres modelos de contraste:

$$M1: y_t = \phi y_{t-1} + u_t$$

$$M2: y_t = \alpha + \phi y_{t-1} + u_t$$

$$M3: y_t = \alpha + \beta t + \phi y_{t-1} + u_t$$

Estos tres modelos pueden utilizarse para definir el t-ratio y contrastar la

hipótesis nula $H_0: \phi = 1$. Por ejemplo, si el gráfico indica utilizar el modelo M3 entonces estamos pensando en un PGD del siguiente tipo: $y_t = \alpha + \beta t + v_t$ con $v_t = \phi v_{t-1} + u_t$. Transformando se obtiene: $y_t - \alpha - \beta t = \phi(y_{t-1} - \alpha - \beta(t-1)) + u_t$ llegándose al siguiente modelo: $y_t = (1 - \phi)\alpha + \beta\phi + (1 - \phi)\beta t + \phi y_{t-1} + u_t$. Bajo H_0 obtenemos el modelo M2. Luego si utilizamos el modelo M3 la distribución del t-ratio bajo H_0 hay que derivarla asumiendo el modelo M2. Similares comentarios se pueden hacer si utilizamos el modelo M2; la distribución del t-ratio hay que derivarla asumiendo el modelo M1.

Por lo tanto, podemos considerar tres procesos generadores de datos. El primer proceso corresponde al caso en que los valores de la variable giran en torno a cero. Bajo la hipótesis nula de que la serie tiene una tendencia estocástica el PGD puede modelizarse como,

$$H_0 : \text{PGD1} : y_t = y_{t-1} + u_t \quad y_0 = 0$$

Bajo la hipótesis alternativa de ser un proceso estacionario, el modelo es

$$H_1 : y_t = \phi y_{t-1} + u_t \quad |\phi| < 1$$

El segundo proceso corresponde al caso en que los valores de la variable giran en torno a una constante. Bajo la hipótesis nula de que la variable tiene una tendencia estocástica, la modelización es,

$$H_0 : \text{PGD2} : y_t = y_{t-1} + u_t \quad y_0 \neq 0$$

Bajo la hipótesis alternativa de ser un proceso sin tendencia estocástica, el modelo es,

$$H_1 : y_t = \alpha + \phi y_{t-1} + u_t \quad |\phi| < 1$$

Por último, vamos a analizar el caso en que los valores de la variable mantienen una pauta creciente o decreciente a lo largo de toda la muestra. El modelo que consideramos suponiendo que la serie tiene tendencia estocástica es el siguiente,

$$H_0 : \text{PGD3} : y_t = \alpha + y_{t-1} + u_t$$

Bajo la hipótesis alternativa de no existencia de tendencia estocástica el modelo es,

$$H_1 : y_t = \alpha + \beta t + \phi y_{t-1} + u_t \quad |\phi| < 1$$

El resumen de las posibilidades que se presentan puede verse en el Cuadro 4.3. Las columnas corresponden al modelo utilizado en el contraste y las filas al modelo que, se supone, ha generado los datos. Cada casilla corresponde al cruce de un modelo utilizado y un modelo que supondremos genera los datos. En cada casilla aparecen, en la parte superior, el número del escenario y en la inferior, entre paréntesis, la identificación de los puntos críticos que le corresponden. La notación Bloque 1, Bloque 2 y Bloque 3 corresponde a los bloques distinguidos en el Cuadro 4.4. Un primer aspecto a destacar es la distinción entre PGD1 y PGD2 no es relevante desde el punto de vista teórico pero sí que lo es desde el punto de vista práctico. Otro punto a destacar es que el escenario 4 solo tiene relevancia desde el punto de vista teórico.

CUADRO 4.3. Contraste del Orden de Integración Escenarios

		MODELO UTILIZADO		
		M₁	M₂	M₃
PGD	PGD1	ESCENARIO 1	ESCENARIO 3	ESCENARIO 5
	PGD2	(BLOQUE 1)	(BLOQUE 2)	(BLOQUE 3)
		ESCENARIO 2	ESCENARIO 4	ESCENARIO 6
	PGD3	(NO HAY DISTRIBUCION DEFINIDA)	(PUNTOS CRITICOS ESTANDAR)	(BLOQUE 3)

Al final de esta sección se presenta una estrategia de contraste distinguiendo tres situaciones diferentes que son las asociadas con los escenarios 1, 2, y 6 que son los que tienen una relevancia mayor en el trabajo aplicado.

CUADRO 4.4. Puntos Críticos según Escenario

BLOQUES	Tamaño Muestral	Nivel de Significación		
		1%	5%	10%
BLOQUE 1	25	-2,66	-1,95	-1,60
	50	-2,62	-1,95	-1,61
	100	-2,60	-1,95	-1,61
	250	-2,58	-1,95	-1,62
BLOQUE 2	25	-3,75	-3,00	-2,63
	50	-3,58	-2,93	-2,60
	100	-3,51	-2,89	-2,58
	250	-3,46	-2,88	-2,57
BLOQUE 3	25	-4,38	-3,60	-3,24
	50	-4,15	-3,50	-3,18
	100	-4,04	-3,45	-3,15
	250	-3,99	-3,43	-3,13

Analicemos ahora los seis escenarios presentados en el Cuadro 4.3.

Escenario 1. En este caso se utiliza el modelo M1 para el contraste pero suponiendo que los datos han sido generados por PGD1 ó PGD2.

Para derivar las propiedades del estimador MCO de ϕ cuando $\phi = 1$, hay que tener en cuenta que:

$$T(\hat{\phi} - 1) = \frac{\sum y_{t-1} u_t / T}{\sum y_{t-1}^2 / T^2}$$

Veamos ahora las propiedades del estimador MCO cuando $\phi = 1$. Podemos escribir la estimación como:

$$T(\hat{\phi} - 1) = \frac{\sum y_{t-1} u_t / T}{\sum y_{t-1}^2 / T^2} \rightarrow \frac{\frac{1}{2} \sigma^2 \chi^2(1) - \frac{1}{2} \sigma^2}{\sigma^2 \int_0^1 (W(r))^2 dr} = \frac{\frac{1}{2} (\chi^2(1) - 1)}{\int_0^1 (W(r))^2 dr} \quad (4.16)$$

Utilizando los apartados d) y f) del Resultado 2.6. También podemos escribir:

$$\hat{\phi} = 1 + \frac{1}{T} \frac{\frac{1}{2}(\chi^2(1) - 1)}{\int_0^1 (W(r))^2 dr} \quad (4.17)$$

Se ve como $(\hat{\phi} - 1)$ es $O_p(T^{-1})$; es decir, el estimador es igual al valor del parámetro que se estima más una expresión bien definida asintóticamente dividida por T . En este caso, la tasa de convergencia es mayor que la que correspondía cuando $|\phi| < 1$, ya que ahora se multiplica por T y no por \sqrt{T} . Esta es la razón del término "superconsistente" con el que se designa al estimador que cumple (4.17). Es decir, el estimador MCO de ϕ , definido a partir de M1, cuando y_t es un paseo aleatorio es superconsistente.

Esta propiedad de superconsistencia se puede juzgar como algo positivo para un estimador. Pero el problema que afecta al estimador MCO cuando $\phi = 1$, tal como puede verse en (4.16), es que tras normalizar con su orden de probabilidad no resulta una distribución normal como ocurría en el caso estacionario para $\sqrt{T}(\hat{\phi} - \phi)$. Podemos decir que, en el marco no estacionario el estimador MCO de ϕ , $\hat{\phi}$, se acerca más rápidamente al parámetro que en el caso estacionario pero siguiendo un camino menos aceptable que el que se sigue en el marco estacionario que garantiza la simetría en torno a cero.

El hecho de que $T(\hat{\phi} - 1)$ no siga una distribución Gausiana lleva a que el t-ratio no tenga una distribución estándar. En concreto, la convergencia de este t-ratio es la siguiente (Ver expresión (17.4.12) de Hamilton (1994)).

$$t_T \rightarrow \frac{(1/2) \{ [W(1)]^2 - 1 \}}{\left\{ \int_0^1 [W(r)]^2 dr \right\}^{1/2}}$$

Se ve como tiende a una distribución de probabilidad bien definida pero diferente a las estándar. Sus puntos críticos se han derivado mediante ejercicios de

simulación y los valores concretos son los que aparecen en el Bloque 1 del Cuadro 4.4.

Escenario 2. En este caso, se utiliza para el contraste el modelo M1 pero el que ha generado los datos es el PGD3.

El estimador MCO de ϕ puede escribirse como:

$$\hat{\phi} = \frac{\sum y_t y_{t-1}}{\sum y_{t-1}^2} = \alpha \frac{\sum y_{t-1}}{\sum y_{t-1}^2} + 1 + \frac{\sum y_{t-1} u_t}{\sum y_{t-1}^2}$$

Por lo que:

$$\hat{\phi} - 1 = \alpha \frac{\sum y_{t-1}}{\sum y_{t-1}^2} + \frac{\sum y_{t-1} u_t}{\sum y_{t-1}^2}$$

Teniendo en cuenta el Resultado 2.6, podemos escribir:

$$\begin{aligned} \sum_{t=1}^T y_{t-1} &= \sum_{t=1}^T y_0 + \sum_{t=1}^T \alpha(t-1) + \sum_{t=1}^T y_{t-1}^* \\ &O_p(T) \quad O_p(T^2) \quad O_p(T^{3/2}) \end{aligned}$$

en donde se ha utilizado que $\sum_{t=1}^T y_0 = T y_0$. Para obtener una expresión bien definida

asintóticamente tenemos que dividir ambos lados por T^2 , resultando:

$$T^{-2} \sum y_{t-1} \rightarrow 0 + \frac{\alpha}{2} + 0 \equiv NE$$

(NE \equiv expresión no estocástica)

Por lo tanto:

$$\sum_{t=1}^T y_{t-1} \equiv O_p(T^2)$$

De la misma forma:

$$\sum_{t=1}^T y_{t-1}^2 = \sum_{t=1}^T (y_0 + \alpha(t-1) + y_{t-1}^*)^2 =$$

$$\begin{aligned} & \sum y_0^2 + \sum \alpha^2 (t-1)^2 + \sum y_{t-1}^{*2} + \\ & \quad O_p(T) \quad O_p(T^3) \quad O_p(T^2) \\ & + 2 \sum y_0 \alpha (t-1) + 2 \sum y_0 y_{t-1}^* + 2 \sum \alpha (t-1) y_{t-1}^* \\ & \quad O_p(T^2) \quad O_p(T^{3/2}) \quad O_p(T^{5/2}) \end{aligned}$$

Para obtener una expresión asintóticamente bien definida tenemos que dividir los dos lados por T^3 resultando:

$$T^{-3} \sum y_{t-1}^2 \rightarrow 0 + \frac{\alpha^2}{3} + 0 + 0 + 0 + 0 \equiv NE$$

por lo que se puede escribir:

$$\sum y_{t-1}^2 \equiv O_p(T^3)$$

Por último:

$$\begin{aligned} \sum y_{t-1} u_t &= \sum_{t=1}^T y_0 u_t + \sum_{t=1}^T \alpha (t-1) u_t + \sum_{t=1}^T y_{t-1}^* u_t \\ & \quad O_p(T^{1/2}) \quad O_p(T^{3/2}) \quad O_p(T) \end{aligned}$$

por lo que:

$$\sum y_{t-1} u_t \equiv O_p(T^{3/2})$$

Hay que destacar que $T^{-3/2} \sum y_{t-1} u_t$ sigue una distribución normal con media cero y varianza igual a $\sigma^2/3$.

Utilizando estos resultados podemos escribir:

$$\begin{aligned}
T(\hat{\phi} - 1) &= \alpha \frac{T(\sum y_{t-1} / T^3)}{\sum y_{t-1}^2 / T^3} + \frac{T(\sum y_{t-1} u_t / T^3)}{\sum y_{t-1}^2 / T^3} = \\
&= \alpha \frac{\sum y_{t-1} / T^2}{\sum y_{t-1}^2 / T^3} + \frac{\sum y_{t-1} u_t / T^2}{\sum y_{t-1}^2 / T^3} \rightarrow \alpha \frac{\frac{\alpha}{2}}{\frac{\alpha^2}{3}} + 0 = \frac{3}{2}
\end{aligned}$$

Se demuestra que el estimador MCO, $\hat{\phi}$, es superconsistente pero que no hay posibilidad de normalizar para derivar una distribución asintótica.

Escenario 3. Se utiliza el modelo M2 para el contraste generando los datos PGD1 ó PGD2.

A partir de M2 los estimadores MCO se definen como,

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\phi} \end{bmatrix} = \begin{bmatrix} T & \sum y_{t-1} \\ - & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum y_t y_{t-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} T & \sum y_{t-1} \\ & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum u_t \\ \sum y_{t-1} u_t \end{bmatrix}$$

Esta expresión también puede escribirse como,

$$\begin{bmatrix} \hat{\alpha} - 0 \\ \hat{\phi} - 1 \end{bmatrix} = \begin{bmatrix} T^{-\frac{1}{2}} & 0 \\ 0 & T^{-1} \end{bmatrix} \begin{bmatrix} \frac{T}{T} & \frac{\sum y_{t-1}}{T^{3/2}} \\ - & \frac{\sum y_{t-1}^2}{T^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\sum u_t}{\sqrt{T}} \\ \frac{\sum y_{t-1} u_t}{T} \end{bmatrix}$$

Utilizando el Resultado 2.6 se obtiene,

$$\begin{bmatrix} \sqrt{T}(\hat{\alpha} - 0) \\ T(\hat{\phi} - 1) \end{bmatrix} \xrightarrow{d} \begin{bmatrix} 1 & \int W(r) dr \\ - & \int W(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} W(1) \\ \frac{1}{2}(W(1)^2 - 1) \end{bmatrix}$$

De donde se obtiene que

$$T(\hat{\phi} - 1) \rightarrow \frac{(1/2)\{[W(1)]^2 - 1\} - W(1) \int W(r) dr}{\int [W(r)]^2 dr - [\int W(r) dr]^2}$$

En cuanto al t-ratio hay que tener en cuenta que,

$$t = \frac{(\hat{\phi} - 1)}{a_{22}^{1/2} \sigma} = \frac{T(\hat{\phi} - 1)}{\sigma(Ta_{22})^{\frac{1}{2}}}$$

en donde ,

$$a_{22} = \frac{T}{T \sum y_{t-1}^2 - (\sum y_{t-1})^2} = \frac{\frac{1}{T^2}}{\frac{\sum y_{t-1}^2}{T^2} - (\frac{\sum y_{t-1}}{T})^2}$$

por ser la matriz de varianzas y covarianzas

$$\sigma^2 (X'X)^{-1} = \sigma^2 \begin{pmatrix} T & \sum y_{t-1} \\ - & \sum y_{t-1}^2 \end{pmatrix}^{-1} = \sigma^2 \begin{pmatrix} a_{11} & a_{12} \\ - & a_{22} \end{pmatrix}$$

La distribución límite del t-ratio es

$$t_T \rightarrow \frac{1/2 \{ [W(1)]^2 - 1 \} - W(1) \int W(r) dr}{\left\{ \int [W(r)]^2 dr - \left[\int W(r) dr \right]^2 \right\}^{1/2}}$$

Esta distribución no se corresponde con ninguna de las ya tabuladas pero sus puntos críticos se han derivado mediante ejercicios de simulación. Los valores concretos son los que aparecen en el Bloque 2 del Cuadro 4.4.

Escenario 4. En este caso hay coincidencia entre el modelo que se utiliza en el contraste y el que genera los datos; se trata del modelo M2.

Los estimadores MCO de α y ϕ pueden escribirse como:

$$\begin{aligned} \begin{bmatrix} \hat{\alpha} \\ \hat{\phi} \end{bmatrix} &= \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum y_{t-1} y_t \end{bmatrix} = \\ &= \begin{bmatrix} \alpha \\ \phi \end{bmatrix} + \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum u_t \\ \sum y_{t-1} u_t \end{bmatrix} \end{aligned}$$

y utilizando el Resultado 2.7 se tiene que:

$$\begin{bmatrix} \hat{\alpha} - \alpha \\ \hat{\phi} - 1 \end{bmatrix} = \begin{bmatrix} O_p(T) & O_p(T^2) \\ O_p(T^2) & O_p(T^3) \end{bmatrix}^{-1} \begin{bmatrix} O_p(T^{1/2}) \\ O_p(T^{3/2}) \end{bmatrix}$$

Teniendo en cuenta los elementos de la diagonal de la matriz inversa, la matriz de normalización que adoptamos es:

$$V_T = \begin{bmatrix} T^{1/2} & 0 \\ 0 & T^{3/2} \end{bmatrix}$$

Premultiplicando por esta matriz se obtiene:

$$\begin{bmatrix} T^{1/2}(\hat{\alpha} - \alpha) \\ T^{3/2}(\hat{\phi} - 1) \end{bmatrix} \xrightarrow{d} \begin{bmatrix} 1 & \alpha/2 \\ \alpha/2 & \alpha^2/3 \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2} \sum u_t \\ T^{-3/2} \sum y_{t-1} u_t \end{bmatrix}$$

y teniendo en cuenta que:

$$\begin{bmatrix} T^{-1/2} \sum u_t \\ T^{-3/2} \sum y_{t-1} u_t \end{bmatrix} \rightarrow N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & \frac{\alpha}{2} \\ \frac{\alpha}{2} & \frac{\alpha^2}{3} \end{pmatrix} \right]$$

El resultado al que se llega es que:

$$\begin{bmatrix} T^{1/2}(\hat{\alpha} - \alpha) \\ T^{3/2}(\hat{\phi} - 1) \end{bmatrix} \rightarrow N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & \frac{\alpha}{2} \\ \frac{\alpha}{2} & \frac{\alpha^2}{3} \end{pmatrix}^{-1} \right]$$

Es decir, el estimador $\hat{\phi}$ es superconsistente y, además, después de normalizar sigue una distribución normal con momentos finitos. El correspondiente t-ratio seguirá una distribución t de Student y no se presenta nada nuevo respecto a la práctica tradicional.

Maddala y Kim (1998) indican que este resultado tiene un escaso interés porque el comportamiento de y_t , bajo la hipótesis nula, tiene tendencia pero no bajo la hipótesis alternativa.

Desde nuestro punto de vista el resultado sí que tiene interés porque, bajo la hipótesis nula, $(\hat{\phi} - 1)$ sigue una distribución asintótica que es la calculada y, bajo la alternativa, $(\hat{\phi} - 1)$ sigue una distribución también gaussiana. Suponiendo que el PGD es M2 con $|\phi| < 1$ y utilizando el Resultado 2.5 y (2.16) se tiene que, asintóticamente:

$$\sqrt{T}(\hat{\phi} - \phi) \xrightarrow{d} N[0, (1 - \phi^2)]$$

A partir de aquí se tiene que:

$$\sqrt{T}(\hat{\phi} - 1) \xrightarrow{d} N[(\phi - 1), (1 - \phi^2)]$$

Por lo que la distribución de $(\hat{\phi} - 1)$ se centrará en torno a otro valor y puede servir para el contraste. Lógicamente, será más potente cuanto mayor sea la distancia $(\phi - 1)$.

Escenarios 5 y 6. El modelo que se utiliza para el contraste es el M3 pero los datos los genera cualquiera de los tres PGDs.

Suponiendo que los datos han sido generados por PGD3, siguiendo la misma línea de razonamiento que en el Escenario 3, para los estimadores MCO definidos a partir de M3 se obtiene,

$$\begin{bmatrix} \hat{\alpha} - 1 \\ \hat{\beta} - 0 \\ \hat{\phi} - 1 \end{bmatrix} = \begin{bmatrix} T^{-1/2} & 0 & 0 \\ 0 & T^{-3/2} & 0 \\ 0 & 0 & T^{-1} \end{bmatrix} \begin{bmatrix} \frac{T}{T} & \frac{\sum t}{T^2} & \frac{\sum y_{t-1}}{T^{3/2}} \\ - & \frac{\sum t^2}{T^3} & \frac{\sum ty_{t-1}}{T^{3/2}} \\ - & - & \frac{\sum y_{T-1}^2}{T^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\sum u_t}{T^{1/2}} \\ \frac{\sum tu_t}{T^{3/2}} \\ \frac{\sum y_{t-1}u_t}{T} \end{bmatrix}$$

Como puede verse a partir de la expresión (17.4.53) de Hamilton (1994) el límite de $T(\hat{\phi} - 1)$ adopta una forma más compleja pero está bien definida y un resultado importante es que no depende del valor que toma α ; por eso, la distribución es la misma para ambos escenarios.

Por otra parte, el t-ratio tiende a una distribución no-estándar cuyos puntos

críticos han sido generados mediante ejercicios de simulación y son los que aparecen en el Bloque 3 del Cuadro 4.4. En este caso el t-ratio toma la forma siguiente,

$$t = \frac{\hat{\phi} - 1}{\sigma(a_{33})^{1/2}}$$

en donde a_{33} es el elemento (3,3) de la matriz de varianzas y covarianzas de los estimadores MCO.

A la vista de estos resultados cabría pensar en una estrategia de contraste del siguiente tipo:

Caso 1: La serie gira en torno a cero. En este caso, se define el t-ratio a partir de M1 y se utilizan los puntos críticos del Bloque 1.

Caso 2: La serie gira en torno a una constante. En este caso, se define el t-ratio a partir del Modelo 2 (M2) y se utilizan los puntos críticos del Bloque 2. Podría pensarse en una segunda vía consistente en utilizar el contraste de la F para la hipótesis nula $\alpha = 0$ y $\phi = 1$. Diferentes estudios han demostrado que los resultados a los que se llega utilizando este segundo método son similares a los obtenidos con el primer método.

Caso 3: La serie sigue una pauta creciente o decreciente a lo largo de todo el periodo muestral. Se define el t-ratio utilizando el Modelo 3 (M3) y se contrasta la hipótesis nula utilizando los puntos críticos del Bloque 3. En este caso también cabría pensar en utilizar el contraste de la F para la Hipótesis nula $\beta = 0$ y $\phi = 1$. Destacar que, para llevar a cabo este contraste, los puntos críticos son diferentes a los de la distribución F standard. Los resultados obtenidos con ambos métodos son muy similares.

La práctica normal en la literatura ha consistido en utilizar no los modelos como se han escrito anteriormente sino cambiando la variable dependiente en niveles por su incremento. En este caso se contrasta la hipótesis nula: $\phi^* = 0$ en lugar de $\phi = 1$, en donde $\phi^* = \phi - 1$. Por ejemplo, el modelo M2 tomaría la forma siguiente:

$$\Delta y_t = \alpha + \phi^* y_{t-1} + u_t$$

Hasta este momento hemos supuesto que la perturbación aleatoria está libre de

autocorrelación. En el caso en que haya autocorrelación entonces lo que se hace es plantear una regresión en la que a los regresores anteriores se añaden incrementos retardados de la variable dependiente:

$$\Delta y_t = \alpha + \phi^* y_{t-1} + \sum_{j=1}^p \gamma_j \Delta y_{t-j} + u_t$$

Lo mismo se haría para los modelos M1 y M3. La hipótesis nula que se contrasta es la misma: $H_0: \phi^* = 0$. El estadístico que se utiliza es el t-ratio correspondiente a ϕ^* . Este contraste se llama: "Contraste de Dickey-Fuller Ampliado" o DFA. Los puntos críticos son los mismos ya vistos para el caso en que no habría autocorrelación.

Los contrastes DF y DFA con el propuesto por Phillips y Perron (1988) son los más utilizados en el trabajo aplicado. Esto no significa que estén libres de problemas. En Maddala y Kim (1998) puede encontrarse un planteamiento crítico sobre los mismos. Nosotros, sin embargo, seguiremos recomendando su uso por dos razones: primero, porque no existe un contraste que sea óptimo en todas las situaciones y, segundo, en una primera etapa del proceso de selección basta con una aproximación simple que acompañe a otros procedimientos, ya comentados, para extraer la evidencia en una dirección u otra.

4.4. COINTEGRACIÓN. CONTRASTES

En esta Sección vamos a prestar atención al concepto de Cointegración y a los procedimientos propuestos en la literatura para contrastar su presencia.

En la última sección del Capítulo 2 hemos mostrado la relevancia de la cointegración para derivar las propiedades de los estimadores y los estadísticos de contraste. Es, por lo tanto, importante disponer de procedimientos fiables para diagnosticar la existencia y forma de la cointegración.

Comenzaremos con la definición contenida en el trabajo de Engle y Granger (1987):

Definición 4.4: "Los componentes de un vector y_t se dice que están cointegrados de

orden d , b , y lo denotamos por: $y_t \sim CI(d, b)$, si (i) todos los componentes de y_t son $I(d)$, (ii) existe un vector $\alpha (\neq 0)$ tal que $Z_t = \alpha' y_t \sim I(d-b)$, $b > 0$. El vector α es el vector de cointegración".

Lo más habitual ha consistido en tomar $d = b = 1$, de forma que todos los elementos de y_t son $I(1)$ y Z_t es $I(0)$, es decir, estacionaria.

Considerar el siguiente ejemplo de Banerjee et al. (1993). Las relaciones del modelo son:

$$y_{1t} + \lambda y_{2t} = u_{1t}$$

$$y_{1t} + \beta y_{2t} = u_{2t}$$

en donde:

$$u_{1t} = u_{1t-1} + \varepsilon_{1t}$$

$$u_{2t} = \rho u_{2t-1} + \varepsilon_{2t} \quad |\rho| < 1$$

$$E\varepsilon_{1t} = E\varepsilon_{2t} = 0 \quad \forall t \quad \text{y} \quad \text{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) = \sigma_{12}$$

Resolviendo y_{1t} e y_{2t} a partir de este sistema se obtiene:

$$y_{1t} = \beta(\beta - \lambda)^{-1} u_{1t} - \lambda(\beta - \lambda)^{-1} u_{2t}$$

$$y_{2t} = -(\beta - \lambda)^{-1} u_{1t} + (\beta - \lambda)^{-1} u_{2t}$$

Por ser u_{1t} un paseo aleatorio y u_{2t} una variable estacionaria, tanto y_{1t} como y_{2t} son $I(1)$. Además, la combinación $\{y_{1t} + \beta y_{2t}\}$ es estacionaria por lo que ambas variables están cointegradas.

Para contrastar si existe o no cointegración se han propuesto dos tipos de métodos: Uniecuacionales y Multiecuacionales. Los primeros están basados en el vector de residuos MCO de las posibles regresiones de cointegración, después de adoptar una normalización. Los multiecuacionales consideran el modelo en su conjunto y adoptan

otro tipo de normalización. El método más conocido de estos últimos es el contraste de Johansen.

Un análisis comparado de la mayor parte de los contrastes hoy disponibles puede encontrarse en Maddala y Kim (1998).

El punto de partida de los métodos uniecuacionales es el vector de residuos MCO obtenidos a partir de la relación de cointegración definidos como:

$$\hat{u}_t = y_{1t} - \hat{\beta}y_{2t}$$

en donde $\hat{\beta}$ es el estimador MCO. Los contrastes más utilizados en el trabajo aplicado han sido los siguientes: CRDW, DF y DFA.

El contraste CRDW se basa en el estadístico de Durbin-Watson:

$$CRDW = \frac{\sum (\hat{u}_t - \hat{u}_{t-1})^2}{\sum \hat{u}_{t-1}^2}$$

Aproximadamente, este estadístico puede escribirse como:

$$CRDW \approx 2(1 - \hat{\phi})$$

en donde:
$$\hat{\phi} = \frac{\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_{t-1}^2}$$

La hipótesis nula de que no hay cointegración se mantiene cuando ϕ está próxima a 1 o, equivalentemente, cuando CRDW toma un valor muy pequeño. Los puntos críticos pueden verse en la tercera columna del Cuadro 4.5. Así, por ejemplo, si el número de variables es 3 y el número de observaciones es 100 entonces se rechaza la existencia de cointegración si se cumple que $CRDW < 0,48$.

Para el contraste Dickey-Fuller (DF) se parte del modelo:

$$\Delta \hat{u}_t = \phi \hat{u}_{t-1} + e_t$$

Como ya se ha comentado para el análisis univariante, la hipótesis nula de no

cointegración toma la forma: $H_0: \phi = 0$. Para ello, se define el t-ratio como:

$$t = \frac{\hat{\phi}}{\hat{\sigma}_{\hat{\phi}}} \quad (4.18)$$

y la hipótesis nula se rechaza para un nivel de significación del 5% cuando este estadístico toma un valor menor que el correspondiente punto crítico que aparece en la cuarta columna del Cuadro 4.5.

Para el contraste DFA se parte del siguiente modelo:

$$\Delta \hat{u}_t = \phi \hat{u}_{t-1} + \sum_{i=1}^p \delta_i \Delta \hat{u}_{t-i} + e_t \quad (4.19)$$

Se define el t-ratio como en (4.18) y la hipótesis nula de no cointegración se rechaza, utilizando un nivel de significación del 5%, cuando este estadístico t-ratio toma un valor inferior al punto crítico correspondiente que aparece en la última columna del Cuadro 4.5.

Así, por ejemplo, siguiendo con el caso anterior de 3 variables y 100 observaciones, concluiremos que no hay cointegración cuando el valor del estadístico DF es superior a -3,93 y cuando el valor del estadístico DFA es superior a -3,62.

Un análisis de las características de estos contrastes puede verse en Banerjee et al. (1993).

Contraste de Johansen

Comenzaremos reescribiendo la versión MCE del modelo contenida en la Sección 5.3 como: $\Delta y_t = \pi y_{t-1} + \Gamma \Delta y_{t-1} + v_t$

en donde: $\Delta y_t = \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix}, y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}, v_t = \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$

$$\pi = \begin{bmatrix} \alpha_1 - \alpha_1 \beta \\ \alpha_2 - \alpha_2 \beta \end{bmatrix}; \Gamma = \begin{bmatrix} \phi_{111}^* & \phi_{121}^* \\ \phi_{211}^* & \phi_{221}^* \end{bmatrix}$$

La generalización al caso en que se tienen p variables y k retardos en el modelo en niveles es:

$$\Delta y_t = \pi y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + v_t \quad (4.20)$$

Esta es la versión utilizada por Johansen (1995) añadiendo o no términos deterministas - constante y/o variables ficticias estacionales.

La versión más restringida de este modelo puede escribirse como:

$$\Delta y_t = \pi y_{t-1} + v_t \quad (4.21)$$

CUADRO 4.5. Puntos Críticos de los Contrastes de Cointegración (Nivel Significación 5%)

Número de Variables	Tamaño Muestral	CRDW	DF	DFA
2	50	0,72	-3,67	-3,29
	100	0,38	-3,37	-3,17
	200	0,20	-3,37	-3,25
3	50	0,89	-4,11	-3,75
	100	0,48	-3,93	-3,62
	200	0,25	-3,78	-3,78
4	50	1,05	-4,35	-3,98
	100	0,58	-4,22	-4,02
	200	0,30	-4,18	-4,13
5	50	1,19	-4,76	-4,15
	100	0,68	-4,58	-4,36
	200	0,35	-4,48	-4,43

Fuente: Para el CRDW obtenidos de Sargan y Barghava (1983) y para DF y DFA de Engle y Yoo (1987)

A continuación, vamos a derivar el contraste de cointegración de Johansen utilizando este modelo. La derivación del contraste para marcos menos restringidos es inmediata y puede encontrarse en la referencia ya hecha de Johansen (1995).

La hipótesis nula de que no existe cointegración se formula como:

$$H_0: \pi = 0.$$

Bajo esta hipótesis nula se tiene que (suponiendo que $y_{10} = y_{20} = 0$):

$$y_{1t} = y_{1t-1} + v_{1t} = \sum_{i=1}^t v_{1i}$$

$$y_{2t} = y_{2t-1} + v_{2t} = \sum_{i=1}^t v_{2i}$$

generalizando los resultados vistos en el Capítulo 2, podemos escribir:

$$T^{-1/2} \sum_{t=1}^T v_t \Rightarrow B(r) \quad r \in [0,1]$$

en donde $B(r)$ es un vector Browniano con:

$$B(r) = \begin{pmatrix} B_1(r) \\ B_2(r) \end{pmatrix} = BM(\Omega)$$

Ω es la matriz de varianzas y covarianzas del vector Browniano que, en este caso, es:

$$\Omega = \begin{pmatrix} \sigma_1^2 + \beta^2 \sigma_2^2 + 2\beta\sigma_{12} & \beta\sigma_2^2 + \sigma_{12} \\ \beta\sigma_2^2 + \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

El contraste de Johansen es un contraste de Razón de Verosimilitud (LR). Como hemos indicado en el Capítulo 3, este contraste está basado en el estadístico:

$$LR = -2 \log \frac{L_{H_0}^{MAX}}{L_{H_1}^{MAX}}$$

y la hipótesis nula se rechaza cuando el estadístico toma un valor grande. $L_{H_1}^{MAX}$ denota el valor máximo que toma la función de verosimilitud bajo H_1 . Nosotros consideramos como alternativa la hipótesis de que existen dos relaciones de cointegración.

En el Capítulo 6 de su libro, Johansen demuestra que:

$$\left(L_{H_0}^{MAX} \right)^{-2/T} = |S_{00}|$$

$$\left(L_{H_1}^{MAX} \right)^{-2/T} = |S_{00}| \prod_{i=1}^2 (1 - \tilde{\lambda}_i)$$

en donde $\tilde{\lambda}_i$, $i = 1, 2$ son las raíces características de la ecuación:

$$\left| \tilde{\lambda} S_{11} - S_{10} S_{00}^{-1} S_{01} \right| = 0$$

siendo:

$$\begin{aligned} S_{00} &= \frac{1}{T} \begin{bmatrix} \sum \Delta y_{1t}^2 & \sum \Delta y_{1t} \Delta y_{2t} \\ \sum \Delta y_{1t} \Delta y_{2t} & \sum \Delta y_{2t}^2 \end{bmatrix} \\ S_{10} &= \frac{1}{T} \begin{bmatrix} \sum y_{1t-1} \Delta y_{1t} & \sum y_{1t-1} \Delta y_{2t} \\ \sum y_{2t-1} \Delta y_{1t} & \sum y_{2t-1} \Delta y_{2t} \end{bmatrix} \\ S_{11} &= \frac{1}{T} \begin{bmatrix} \sum y_{1t-1}^2 & \sum y_{1t-1} y_{2t-1} \\ \sum y_{2t-1} y_{1t-1} & \sum y_{2t-1}^2 \end{bmatrix} \end{aligned}$$

El estadístico Razón de Verosimilitud puede escribirse como:

$$LR = -T \sum_{i=1}^2 \log (1 - \tilde{\lambda}_i) \quad (4.22)$$

Para derivar la distribución asintótica de este estadístico considerar la expansión de Taylor de $\log (1 - \tilde{\lambda}_i)$ en torno a $\tilde{\lambda}_i = 0$ que puede escribirse como:

$$\log(1 - \tilde{\lambda}_i) \approx \log(1) - \frac{1}{1 - \tilde{\lambda}_i} \bigg|_{\tilde{\lambda}_i=0} (\tilde{\lambda}_i - 0) = -\tilde{\lambda}_i$$

Por lo que el estadístico de la Razón de Verosimilitud asintóticamente puede escribirse como:

$$LR \approx T \sum_{i=1}^2 \tilde{\lambda}_i \quad (4.23)$$

Teniendo en cuenta que las raíces de $\left| \tilde{\lambda} S_{11} - S_{10} S_{00}^{-1} S_{01} \right| = 0$ coinciden con las de $\left| \tilde{\lambda} I - S_{11}^{-1} S_{10} S_{00}^{-1} S_{01} \right| = 0$ y que la suma de las raíces de una matriz coincide con la traza de esa matriz se tiene que:

$$LR = T \text{tr} \{ S_{11}^{-1} S_{10} S_{00}^{-1} S_{01} \} = \text{tr} \left\{ \left(\frac{S_{11}}{T} \right)^{-1} S_{10} S_{00}^{-1} S_{01} \right\}$$

Teniendo en cuenta el Resultado 2.8 se tiene que:

$$\frac{1}{T} S_{11} \rightarrow \int_0^1 BB' dr$$

$$S_{10} \rightarrow \int_0^1 B(dB')$$

$$S_{00} \rightarrow \Omega$$

Utilizando estos resultados, el estadístico LR toma la forma siguiente:

$$LR = \text{tr} \left\{ \left(\int_0^1 BB' dr \right)^{-1} \int_0^1 B(dB') \Omega^{-1} \int_0^1 (dB)B' \right\}$$

en donde:

$$\int_0^1 \mathbf{B}\mathbf{B}'dr = \begin{bmatrix} \int_0^1 B_1^2 dr & \int_0^1 B_1 B_2 dr \\ 0 & 0 \\ \int_0^1 B_2 B_1 dr & \int_0^1 B_2^2 dr \\ 0 & 0 \end{bmatrix}$$

$$\int_0^1 \mathbf{B}(dB)' = \begin{bmatrix} \int_0^1 B_1 dB_1 & \int_0^1 B_1 dB_2 \\ 0 & 0 \\ \int_0^1 B_2 dB_1 & \int_0^1 B_2 dB_2 \\ 0 & 0 \end{bmatrix}$$

Si definimos el vector de variables Brownianas estandarizadas como:

$\mathbf{W} = \Omega^{-1/2} \mathbf{B}$ el estadístico puede escribirse como:

$$\begin{aligned} \text{LR} &= \text{tr} \left\{ \left(\int_0^1 \Omega^{-1/2} \mathbf{B}\mathbf{B}' \Omega^{-1/2} dr \right)^{-1} \Omega^{-1} \int_0^1 \mathbf{B}(dB)' \Omega^{-1} \int_0^1 (dB)\mathbf{B}' \right\} = \\ &= \text{tr} \left\{ \left(\int_0^1 \mathbf{W}\mathbf{W}' dr \right)^{-1} \int_0^1 \Omega^{-1/2} \mathbf{B}(dB\Omega^{-1/2})' \int_0^1 (\Omega^{-1/2} dB)(\mathbf{B}'\Omega^{-1/2}) \right\} = \\ &= \text{tr} \left\{ \int_0^1 \mathbf{W}(d\mathbf{W})' \left[\int_0^1 \mathbf{W}\mathbf{W}' dr \right]^{-1} \int_0^1 (d\mathbf{W})\mathbf{W}' \right\} \end{aligned}$$

Al no depender \mathbf{W} de ningún parámetro desconocido se puede hallar la distribución de probabilidad y los correspondientes puntos críticos mediante un ejercicio de simulación.

Hay que destacar que todos estos resultados se han derivado asumiendo el marco más simple escrito en (4.21). Pero tal como puede verse en Johansen (1995), estos resultados pueden extenderse con facilidad a marcos menos restringidos.

La extensión natural de (4.20) es el siguiente modelo:

$$\Delta y_t = \alpha \beta' y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \mu_t + v_t \quad (4.24)$$

en donde $\pi = \alpha\beta'$ siendo β la matriz de las relaciones de cointegración. μ_t suponemos que es una tendencia lineal: $\mu_t = \mu_0 + \mu_1 t$.

Johansen (1995) contempla los siguientes cinco escenarios.

Escenario	μ_t	Restricción	Estructura de las y_{it}
1	$\mu_t = 0$	-	No hay deriva
2	$\mu_t = \mu_0$	-	Tendencia lineal
3	$\mu_t = \mu_0$	$\alpha_{\perp}' \mu_0 = 0$	No tienen Tendencia Lineal
4	$\mu_t = \mu_0 + \mu_1 t$	-	Tendencia Cuadrática
5	$\mu_t = \mu_0 + \mu_1 t$	$\alpha_{\perp}' \mu_1 = 0$	Tendencia Lineal

α_{\perp} representa el espacio ortogonal a α .

Johansen (1995) propone calcular tres estadísticos alternativos.

Estadístico N° 1. En este caso, las matrices S_{00} , S_{10} y S_{11} se definen a partir de los residuos de Δy_t e y_{t-1} sobre $\Delta y_{t-1}, \dots, \Delta y_{t-k+1}$ y μ_t . Sean:

R_{0t} : residuos de la regresión de Δy_t sobre

$\Delta y_{t-1}, \dots, \Delta y_{t-k+1}$ y μ_t

R_{1t} : residuos de la regresión de y_{t-1} sobre

$$\Delta y_{t-1}, \dots, \Delta y_{t-k+1} \text{ y } \mu_t$$

Entonces las matrices se definen como:

$$S_{ij} = \frac{1}{T} \begin{bmatrix} \sum R_{1it} R_{1jt} & \sum R_{1it} R_{2jt} \\ \sum R_{2it} R_{1jt} & \sum R_{2it} R_{2jt} \end{bmatrix}$$

El estadístico será el escrito en (4.22) con las raíces calculadas a partir de estas matrices.

Johansen (1995) deriva la distribución de probabilidad de este estadístico para los escenarios 1, 2 y 4. Esta distribución y los correspondientes puntos críticos dependen de la diferencia entre el número de variables y el número de relaciones de cointegración supuesto para la hipótesis nula. La hipótesis alternativa siempre asume que este número coincide con el de variables. Todos los resultados se han derivado para un tamaño muestral de 400.

Algunos de los puntos críticos derivados por Johansen pueden verse en el Cuadro 4.6.

CUADRO 4.6. Puntos Críticos del Contraste de la Razón de Verosimilitud de Johansen. (Nivel de significación: 5%)

n-r	Escenario 1	Escenario 2	Escenario 4
1	4,14	3,84	3,84
2	12,21	15,34	18,15
3	24,08	29,38	34,56
4	39,71	47,21	54,11
5	59,24	68,68	77,79
Fuente: Tablas 15.1, 15.3 y 15.5 de Johansen (1995)			

En nuestro ejemplo con dos variables si se contrasta que no hay cointegración frente a que hay dos relaciones de cointegración entonces la hipótesis nula se rechaza cuando el Estadístico nº 1 toma un valor superior a 18,15. Si toma un valor entre 12,21 y 18,15, la conclusión dependerá de cual sea el escenario más plausible.

En general, cuando el número de variables es superior a 2, es necesario pensar en procesos secuenciales para determinar el rango del espacio de cointegración. Supongamos un caso con tres variables. Se comienza contrastando la hipótesis nula de que no hay cointegración utilizando los puntos críticos correspondientes a la fila $n-r = 3$. Si se rechaza la no existencia de cointegración, se contrasta la hipótesis nula de que existe una relación de cointegración, utilizando los puntos críticos escritos en la fila $n - r = 2$. Y así se seguiría el proceso.

Estadístico Nº 2. En este caso las matrices S_{00} , S_{01} , S_{11} se definen a partir de los residuos de las regresiones de Δy_t e $(y_{t-1}, 1)$ sobre $\Delta y_{t-1}, \dots, \Delta y_{t-k+1}$. El resto del proceso sigue siendo el mismo.

Johansen (1995) deriva la distribución de este estadístico asumiendo el Escenario 3. Los puntos críticos aparecen en la segunda columna del Cuadro 4.7.

CUADRO 4.7. Puntos Críticos del Contraste de Johansen (Nivel de Significación 5%)

n-r	Escenario 3	Escenario 5
1	9,13	12,39
2	19,99	25,47
3	34,80	42,20
4	53,42	62,61
5	75,74	86,96

Fuente: Tablas 15.2 y 15.4 de Johansen (1995).

El mensaje sería el siguiente: Cuando hay mucha evidencia a favor de incluir una constante en el modelo y las variables no parecen incorporar una tendencia lineal entonces calcula el Estadístico nº2 y compáralo con los puntos críticos escritos en la segunda columna del Cuadro 4.7.

Estadístico N° 3. En este caso las matrices S_{00} , S_{01} , S_{11} se definen a partir de los residuos de las regresiones de Δy_t e (y_{t-1}, t) sobre $\Delta y_{t-1}, \dots, \Delta y_{t-k+1}, 1$. El resto de los cálculos siguen siendo los mismos.

Los puntos críticos aparecen en la tercera columna del Cuadro 4.7. Ahora el mensaje sería del siguiente tipo: Cuando hay mucha evidencia a favor de incluir una tendencia lineal en el modelo pero las variables no sugieren seguir una tendencia cuadrática, entonces calcula el Estadístico nº 3 y compáralo con los valores de la tercera columna del Cuadro 4.7.

EJERCICIOS

4.1). Sean:

$$y_{1t} = \sum_{i=1}^t u_{1i} + u_{2t}$$
$$y_{2t} = \frac{1}{2} \sum_{i=1}^t u_{1i} + u_{3t}$$
$$y_{3t} = u_{2t}$$

en donde u_{1t}, u_{2t}, u_{3t} , son ruidos blancos.

- 1). Demostrar que y_{1t} e y_{2t} son I(1).
- 2). Demostrar que el vector $y'_t = (y_{1t}, y_{2t}, y_{3t})$ es I(1).
- 3). Demostrar que y_{3t} es I(0), pero que Δy_{3t} no es I(0).
- 4). Demostrar que existen dos relaciones de cointegración y especificar los vectores de cointegración respectivos.

4.2). Sean :

$$y_{1t} = \sum_{i=1}^t \sum_{j=1}^i u_{1j} + \sum_{i=1}^t u_{2i} + u_{3t}$$
$$y_{2t} = \sum_{i=1}^t \sum_{j=1}^i u_{1j} - \sum_{i=1}^t u_{2i} + u_{3t}$$
$$y_{3t} = \sum_{i=1}^t u_{1i} + \sum_{i=1}^t u_{2i} + u_{3t}$$

en donde u_{1t}, u_{2t}, u_{3t} son ruidos blancos.

- 1). Demostrar que y_{1t} e y_{2t} son I(2) mientras que y_{3t} es I(1).
- 2). Demostrar que el vector $y'_t = (y_{1t}, y_{2t}, y_{3t})$ es I(2).
- 3). Demostrar que $y_{1t} - y_{2t}$ es I(1) y que $y_{1t} - y_{2t} - 2y_{3t}$ es I(1).

4.3). Suponer que el proceso generador de datos viene dado por:

$$y_t = \rho y_{t-1} + u_t \quad \rho = 1, y_0 = 0, t = 1, 2, 3, \dots$$

u_t es una secuencia de ruidos blancos. A partir del modelo:

$$y_t = \alpha + \beta t + v_t$$

se definen los estimadores MCO de α y β . Se pide:

- 1). Derivar las propiedades asintóticas de los dos estimadores.

- 2). Derivar la distribución asintótica del t-ratio para la hipótesis nula $H_0: \beta = 0$.
- 3). Comentar las implicaciones de los resultados anteriores.

4.4). Suponer el siguiente modelo:

$$y_t = x_t' \beta + u_t$$

$$\text{con } x_t' = (1, t) \quad \text{y} \quad \beta = \begin{pmatrix} \alpha \\ \delta \end{pmatrix}$$

Definidos los estimadores MCO de α y δ , se pide:

- 1). Estudiar el comportamiento asintótico de dichos estimadores.
- 2). Adoptando la normalización adecuada derivar la distribución asintótica de los dos estimadores.
- 3). Derivar la distribución asintótica del t-ratio definido para contrastar la hipótesis nula $H_0: \delta = \delta_0$.

4.5). Considerar el siguiente modelo:

$$\Delta y_{1t} = \varepsilon_{1t}$$

$$\Delta y_{2t} = y_{1t-1} + \varepsilon_{2t}$$

en donde:

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim \text{i.i.d.} N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix} \right]$$

Se pide:

- 1). Obtener la esperanza y varianza de y_{1t} e y_{2t} y derivar el orden de integración de ambas variables.
- 2). ¿Puede haber cointegración entre y_{1t} e y_{2t} ?.
- 3). ¿Puede haber cointegración entre y_{1t} e Δy_{2t} ?. Si la respuesta es afirmativa escribir el modelo en forma de mecanismo de corrección de error.

4.6). Suponer que u_t es un vector con dos elementos distribuido idéntica e independientemente con media cero y matriz de varianzas y covarianzas igual a Ω .

Suponer, a continuación, el siguiente modelo:

$$\Delta x_t = \Pi x_{t-1} + u_t$$

en donde Π es una matriz 2×2 de coeficientes. Se va a contrastar la hipótesis nula de que no hay cointegración ($\Pi = 0$) frente a la alternativa de que el rango de cointegración

es igual a dos.

1). Demostrar que el contraste de la razón de verosimilitud para contrastar esa hipótesis nula puede escribirse como:

$$LR = -T \sum_{i=1}^2 \log(1 - \tilde{\lambda}_i)$$

en donde $\tilde{\lambda}_1$ y $\tilde{\lambda}_2$ son raíces de la matriz $S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}$ con

$$S_{00} = \frac{1}{T} \sum \Delta x_t \Delta x_t' ; S_{10} = \frac{1}{T} \sum x_{t-1} \Delta x_t' \\ S_{11} = \frac{1}{T} \sum x_{t-1} x_{t-1}'$$

2). Derivar la distribución de probabilidad, bajo H_0 , del contraste LR

4.7). Suponer que u_t es un vector con dos elementos distribuido idéntica e independientemente con media cero y matriz de varianzas y covarianzas igual a Ω .

Suponer, a continuación, el siguiente modelo:

$$\Delta x_t = \Pi x_{t-1} + \Gamma_1 \Delta x_{t-1} + u_t$$

en donde Π y Γ_1 son matrices 2×2 de coeficientes. Se va a contrastar la hipótesis nula de que no hay cointegración ($\Pi=0$) frente a la alternativa de que el rango de cointegración es igual a dos.

1). Demostrar que el contraste de la razón de verosimilitud para esa hipótesis nula puede escribirse como:

$$LR = -T \sum_{i=1}^2 \log(1 - \tilde{\lambda}_i)$$

en donde $\tilde{\lambda}_1$ y $\tilde{\lambda}_2$ son raíces de la matriz $S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}$ con

$$S_{ij} = M_{ij} - M_{i2}M_{22}^{-1}M_{2j}, \quad i, j = 0, 1.$$

$$M_{00} = \frac{1}{T} \sum \Delta x_t \Delta x_t' ; M_{01} = \frac{1}{T} \sum \Delta x_t x_{t-1}' ; M_{02} = \frac{1}{T} \sum \Delta x_t \Delta x_{t-1}'$$

$$M_{11} = \frac{1}{T} \sum x_{t-1} x_{t-1}' ; M_{12} = \frac{1}{T} \sum x_{t-1} \Delta x_{t-1}' ; M_{22} = \frac{1}{T} \sum \Delta x_{t-1} \Delta x_{t-1}'$$

2). Derivar la distribución de probabilidad, bajo H_0 , del contraste LR

4.8) Suponer el siguiente modelo:

$$\Delta y_{1t} = -\frac{1}{2}(y_{1t-1} - y_{2t-1}) + \varepsilon_{1t} \\ \Delta y_{2t} = \varepsilon_{2t}$$

ε_{1t} y ε_{2t} son ruidos blancos independientes entre sí. Suponer que el modelo está escrito en forma MCE. Se pide:

- 1). Encontrar una expresión de y_t en función de $y_0, \varepsilon_1, \dots, \varepsilon_t$.
- 2). Encontrar la perturbación de la relación de cointegración y derivar su media, varianza y su función de autocorrelación.
- 3). Escribir el modelo en la forma:

$$y_t = \pi_1 y_{t-1} + \dots + \pi_p y_{t-p} + \varepsilon_t$$

Identificar los diferentes elementos del modelo y determinar el rango de la matriz:

$$I - \pi_1 \dots - \pi_p$$

- 4). Escribir el modelo en la forma:

$$\Delta y_t = \alpha \beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t$$

e identificar α, β y las matrices Γ_i .
