
CAPITULO 15

SERIES DE TIEMPO ESTACIONARIAS

15.1. INTRODUCCIÓN

Una serie de tiempo es una secuencia de datos numéricos, cada uno de los cuales se asocia con un instante específico del tiempo. Podemos citar como ejemplos de series de tiempo al índice mensual de inflación, al tipo de cambio diario, al producto bruto interno trimestral, al índice de desempleo anual, etc. Estas series poseen como característica que los lapsos de tiempo, para cada una de ellas, son homogéneos; es decir, si se utiliza una serie con datos semanales, este mismo tipo de datos (misma frecuencia) se deberá mantener durante toda la serie.

El análisis de una secuencia de datos se conoce como un *análisis de series de tiempo de una variable*. En caso que se esté estudiando un conjunto de series para la misma secuencia de tiempo, este análisis es denominado *análisis múltiple de series de tiempo*. El objetivo de este análisis es estudiar la dinámica o estructura temporal de la información. A través del estudio y comprensión del comportamiento de las series económicas, se puede llegar a predecir utilizando la información que se tiene hasta el momento utilizando extrapolación.

En el capítulo Modelos de Rezagos Distribuidos se introdujeron rezagos en el modelo. Estos principalmente correspondían a las variables explicativas, aunque también se emplearon los rezagos de la variable explicada. Sin embargo, el empleo de estas no se realizó de forma sistemática ni se buscó determinar la estructura temporal de la variable a explicar.

La importancia del estudio de las Series de Tiempo radica en que empíricamente el empleo de modelos estructurales (que incluyen valores presentes y pasados de varias variables exógenas y algunos rezagos de las variables endógenas) no ha podido alcanzar la misma calidad de resultados que se obtienen a partir del análisis univariado. En este capítulo nos abocaremos al

estudio univariado de las series, utilizando para esto los rezagos de la variable (su componente autorregresivo) junto con términos estocásticos presentes y rezagados (componente de medias móviles)¹.

15.2. SERIES DE TIEMPO ESTACIONARIAS Y NO ESTACIONARIAS

Las series de tiempo económicas generalmente están compuestas por una tendencia, un comportamiento estacional, uno cíclico y un componente irregular, tal como lo muestra la figura 15.1.

Tal como se puede observar, el componente tendencial y el componente estacional² son de más fácil cálculo, aunque aún así el investigador puede toparse con algunas dificultades ya que pueden incorporar en ellos algún tipo de componente estocástico (tendencias estocásticas o bien irregularidad en la frecuencia o amplitud de los ciclos). El componente estacional presenta un comportamiento bastante predecible. Los ciclos se van repitiendo cada 17 unidades de tiempo, lapso de tiempo en que se repiten los picos. La presencia de la tendencia provoca que la media de la serie vaya cambiando constantemente. La ciclicidad de las series se marcan más profundamente por la coincidencia de puntos altos del componente irregular y de picos estacionales mientras que los puntos de depresión son provocados por la coincidencia de puntos valle junto con puntos profundos del componente estacional.

Los componentes de la serie “ST” pueden, por ejemplo estar descritos por las siguientes relaciones:

$$\begin{array}{ll} \text{Tendencia:} & T = 1 + 0.07*t \\ \text{Estacionalidad:} & E = 0.6*\sin(t\pi/6) \\ \text{Componete Irregular:} & I = 0.6*I_{t-1} + \varepsilon_t \end{array}$$

donde ε_t es el elemento que le introduce la estocasticidad al componente I. Es decir, el comportamiento de este componente se ve continuamente afectado por shocks aleatorios.

Lo que ahora nos interesa es el estudio del componente irregular. Como se puede ver en la Figura 15.1, el comportamiento de este componente presenta puntos bajos seguidos por puntos bajos y puntos altos continuados con puntos altos. La predicción de corto plazo aprovecha este tipo de relaciones. Sin embargo, si se quiere realizar una estimación de largo plazo, los investigadores tienen que limitarse con seguir la media que presenta este componente irregular ya que de esta manera podemos tener la única certidumbre que en promedio estamos obteniendo una buena estimación. Esto está condicionado a que la media de la serie permanezca más o menos constante a lo largo de la muestra porque si fuese muy inestable, entonces el investigador no tendría certidumbre de su medición.

Teóricamente una serie de tiempo puede ser vista como una colección de variables aleatorias Y_t ³. Es por este motivo que una colección de este tipo de datos se le denomina proceso

¹ Ambos componentes son denominados componente AR y componente MA, respectivamente.

² Para los fines prácticos de este libro, no haremos diferenciación entre comportamiento cíclico y tendencial. Generalmente el primero involucra el comportamiento de la serie para un mayor lapso de tiempo (de 32 a 6 trimestres), mientras que el segundo representa un comportamiento recurrente en un menor plazo de tiempo (aprox. un año). Una crítica que se le hace al análisis de covarianzas que se empleará en este capítulo es que no permite determinar con exactitud los movimientos de alta o baja frecuencia de la serie, impidiendo realizar una adecuada selección de los componentes de la serie. Para este objetivo, el análisis espectral presenta mejores resultados.

³ Se dice que una variable es aleatoria si para cualquier número real r , existe una probabilidad $p(Y \leq r)$ de que y tome valores menores o iguales a r . Esto implica que al menos existe un valor de r por el cual se cumple que $0 < \text{prob}(Y=r)$

estocástico. Cada una de estas observaciones es una realización del proceso estocástico subyacente. Es una tarea de la teoría económica el desarrollar modelos que capturen el verdadero proceso generador de datos (PGD).

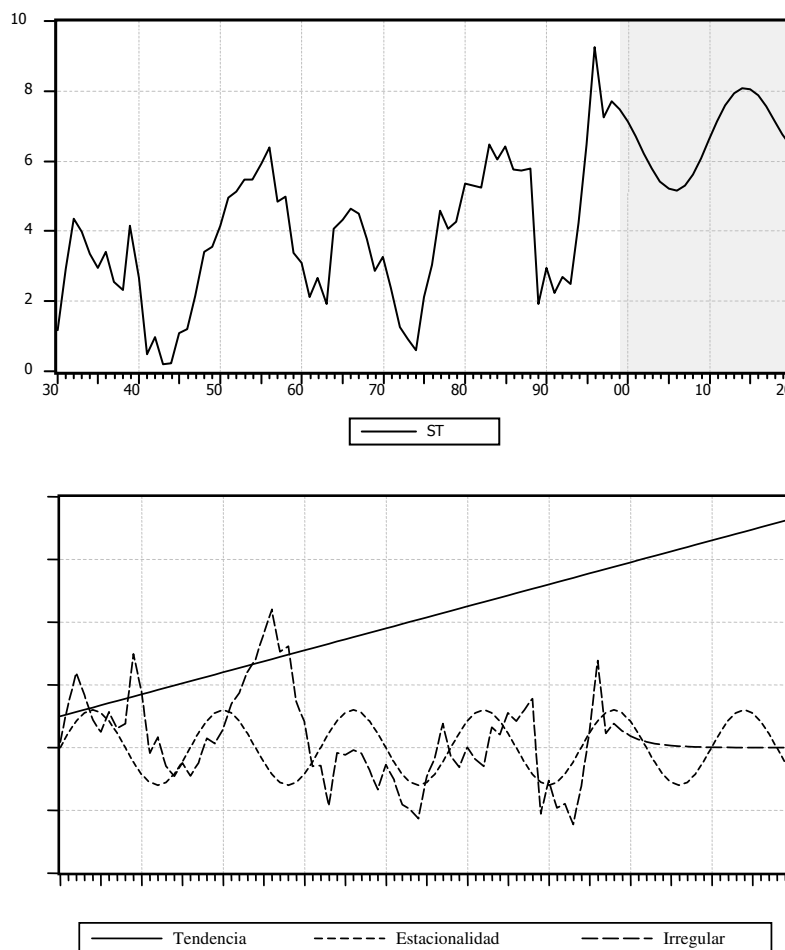


Figura 15.1
 Serie de Tiempo y sus componentes:
 Tendencia, estacionalidad y el componente
 irregular

Una serie de tiempo puede ser continua o discreta. Un ejemplo de una serie continua vendría a estar dado por el registro de un electrocardiograma. Series discretas pueden ser encontradas en grandes cantidades si buscamos en cualquier compendio estadístico. Ejemplo de éstas pueden ser el índice de desempleo, el PBI agrícola, la oferta monetaria, etc. Con el mejoramiento de la

<1. Si existiese alguna r para el cual $p(Y=r)=1$, y es determinístico antes que aleatorio. Ender Walter, Applied Econometric Time Series, John Wiley & Sons, pp. 64

tecnología y el boom que se ha dado en el sector financiero, las series discretas de cotización de acciones que hace unos años eran publicadas semanal o diariamente, ahora son registradas en cada segundo por lo que podemos decir que contamos con nuevas series continuas o sino fíjense en la cotización de Microsoft en el Nasdaq (se asemeja mucho a un electrocardiograma ¿será el electrocardiograma de Bill Gates?)⁴.

Estacionariedad Estricta y Débil

Un proceso estocástico puede ser descrito de acuerdo a la distribución conjunta de las variables Y_t . Sin embargo, y para evitar hacer tan complicado el análisis, al momento de definir un proceso bastará con determinar el primer y el segundo momento de las variables Y_t . Estos son:

1. La media: $\mu(t) = E(Y_t)$
2. La varianza $\sigma^2(t) = \text{var}(Y_t)$
3. Las autocovarianzas $\gamma(t_1, t_2) = \text{cov}(Y_{t_1}, Y_{t_2})$

Una clase muy importante de procesos estocásticos son los estacionarios. Una serie se define como estacionaria cuando no presenta tendencia y su desarrollo corriente se encuentra alrededor de su media. Cualquier shock que sufra en cualquier momento en el tiempo no tendrá efectos permanentes y sólo la alejará temporalmente de su equilibrio. En caso que la serie sea no estacionaria, el camino que recorre a través del tiempo está determinado por los shocks que percibe durante su trayectoria y son estos los que determinan íntegramente su recorrido. No presentan una media determinada.

Durante todos los capítulos anteriores se había estado trabajando bajo el supuesto que todas las series eran estacionarias. De lo contrario, la comprobación de las pruebas de hipótesis basada en el empleo de las pruebas t , la F y la chi-cuadrado y otras pruebas similares, perdería su confiabilidad. Esto último se debe al hecho que la variable no responde a una media constante y por ende su distribución estaría indefinida.

La estacionariedad puede ser definida de dos maneras: estacionariedad fuerte y estacionariedad débil.

Se define un proceso estocástico discreto (PED), en el sentido estricto, a aquel PED que, dada una m -tupla (t_1, t_2, \dots, t_m) y número entero K , posee un vector de variables $(Y_{t_1}, Y_{t_2}, Y_{t_3}, \dots, Y_{t_m})$ que tiene la misma distribución de probabilidad conjunta que el vector $(Y_{t_1+k}, Y_{t_2+k}, Y_{t_3+k}, \dots, Y_{t_m+k})$. Esta definición es válida para todos los valores de m . Si determinamos que el conjunto de variables fuese 1 ($m=1$), obtenemos un $\mu(t)=\mu$ constante y $\sigma^2(t) = \sigma^2$ constante para todo t (independiente del tiempo). En el caso en que sustituyéramos $m=2$, obtendríamos una distribución conjunta para Y_{t_1} e Y_{t_2} igual a la distribución de Y_{t_1+k} y Y_{t_2+k} . Es decir, la distribución debe ser independiente del tiempo. Por ende, las covarianzas y el coeficiente de correlación no varían ante cambios del tiempo si mantenemos el número de rezago (k) constante.

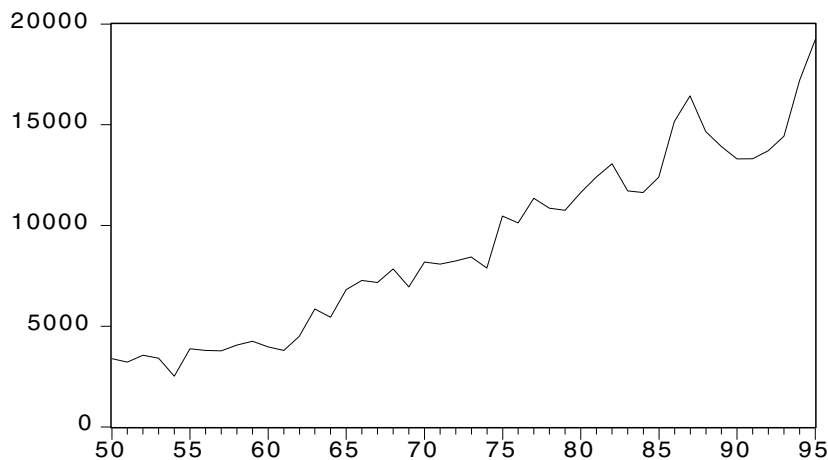
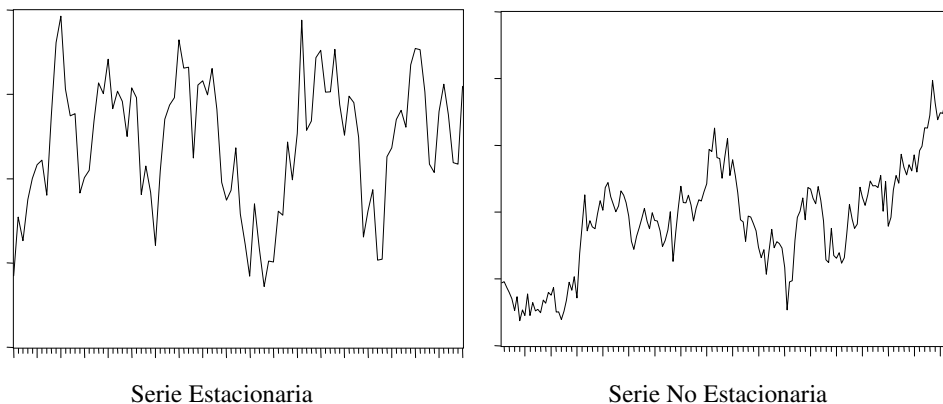
La definición estricta de estacionariedad es demasiado rigurosa como para que realmente se pueda comprobar en la práctica. La misma implica que la distribución de Y_t y de cualquier combinación de éstas, involucrados todos sus momentos, es independiente del tiempo. Si se relaja un poco los supuestos y se limita esta independencia al primer (media) y segundo momento (varianza y autocovarianza), entonces estamos definiendo débilmente la

⁴ En el análisis de series de tiempo se emplean series discretas.

estacionariedad. Este tipo de estacionariedad también es denominado estacionariedad de segundo orden, estacionariedad en sentido amplio, estacionariedad de covarianza.

No estacionariedad

En el próximo capítulo se realizará el análisis de series de tiempo no estacionarias. Estas se diferencian de las estacionarias en que presentan una media que sí depende del tiempo. Además puede ser que no mantengan constantes el resto de los momentos de su distribución.



Serie No Estacionaria con tendencia determinística.
Consumo Privado en el Perú (1950-1995)

Figura 15.2
Series estacionarias y no estacionarias

En la Figura 15.2 se puede apreciar una serie estacionaria, una no estacionaria y una no estacionaria con tendencia determinística. La primera de ellas presenta un comportamiento irregular pero se puede distinguir cierto patrón de comportamiento: las observaciones altas vienen acompañadas por otras observaciones altas. Lo mismo sucede con las observaciones bajas. Además, estas oscilaciones se presentan alrededor de una media constante. La serie no estacionaria también presenta un comportamiento irregular pero su media va cambiando

conforme pasa el tiempo. En este tipo de series los shocks mantienen su efecto permanentemente hasta que otro shock de mayor magnitud produzca un cambio en su comportamiento⁵. Finalmente se cuenta con una serie no estacionaria pero con tendencia determinística. Esta serie, en comparación a la anterior, presenta un comportamiento relacionado con el tiempo. En este caso se ha graficado la serie del consumo privado en el Perú de 1950 hasta 1995. Como la mayoría de las series macroeconómicas, el factor tiempo es bastante importante y determina el comportamiento de largo plazo de este tipo de series.

15.3. EL CÍRCULO UNITARIO Y LAS CONDICIONES DE ESTACIONARIEDAD

En esta sección del capítulo estableceremos cuáles son las características que deben presentar las series de tiempo para determinar si son estacionarias o no. Para hacerlo utilizaremos la teoría que se encuentra detrás del desarrollo de ecuaciones en diferencia. Estas últimas dejan de parecerse a una serie de tiempo por el único hecho de no incluir el término estocástico. Generalizando, una serie de tiempo puede estar definida de la siguiente manera:

$$y_t = a_0 + \sum_{i=1}^n a_i y_{t-i} + X_t \quad (15.1)$$

en donde X_t es conocido como el *proceso generador* y está compuesto por los valores presentes y pasados de otras variables así como de variables estocásticas y del tiempo.

Si como ejemplo definimos que X_t sea igual a un término estocástico (ε_t) y $n=1$, obtendremos la siguiente expresión:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \quad (15.2)$$

Al igual que en la resolución de una ecuación en diferencia, al trabajar con series de tiempo necesitamos encontrar una expresión para Y_t que no incluya a sus rezagos. ¿Cuál es la solución para la ecuación anterior?. Para hallarla necesitamos rezagar esta expresión varias veces:

$$\begin{aligned} y_{t-1} &= a_0 + a_1 y_{t-2} + \varepsilon_{t-1} \\ y_{t-2} &= a_0 + a_1 y_{t-3} + \varepsilon_{t-2} \\ &\vdots \end{aligned}$$

Si reemplazamos Y_{t-1} en Y_t y luego Y_{t-2} en Y_t y así sucesivamente, entonces obtendremos:

$$y_t = a_0 [1 + a_1 + a_1^2 + \dots + a_1^m] + \sum_{i=0}^m a_1^i \varepsilon_{t-i} + a_1^{m+1} y_{t-m-1} \quad (15.3)$$

¿Es esta la solución?. No, porque aún tenemos al lado derecho de la ecuación la presencia de los rezagos (en este caso un rezago) de la variable Y_t .

Podemos seguir incrementando el valor de m (donde “ m ” representa el número de veces en que rezago la variable y la vuelvo a reemplazar) y vamos a observar que continúa la presencia de las Y_t en el lado derecho. Como al parecer, si seguimos rezagando las Y_t no vamos a mejorar

⁵ Este tipo de serie presentan raíces unitarias. Este concepto será ampliamente discutido en el siguiente capítulo.

nuestro resultado, asumiremos algunas condiciones o supuestos. El primero de ellos es asumir que el valor absoluto de a_1 es menor a 1 ($|a_1| < 1$). Con esto estamos indicando que cada vez que rezaguemos más la expresión (15.3), el efecto marginal de los nuevos errores será geoméricamente cada vez menor. La segunda condición para conseguir una solución a la ecuación (15.3) es suponer que se rezaga infinitas veces. Con esta última condición se logrará que el valor de Y_t converja a algún valor fijo en la medida que m va aumentando. Este valor es el siguiente:

$$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \tag{15.4}$$

Esta expresión corresponde a la solución de la expresión (15.2)⁶. Sin embargo, el valor que adopte a_1 va a producir diferentes resultados:

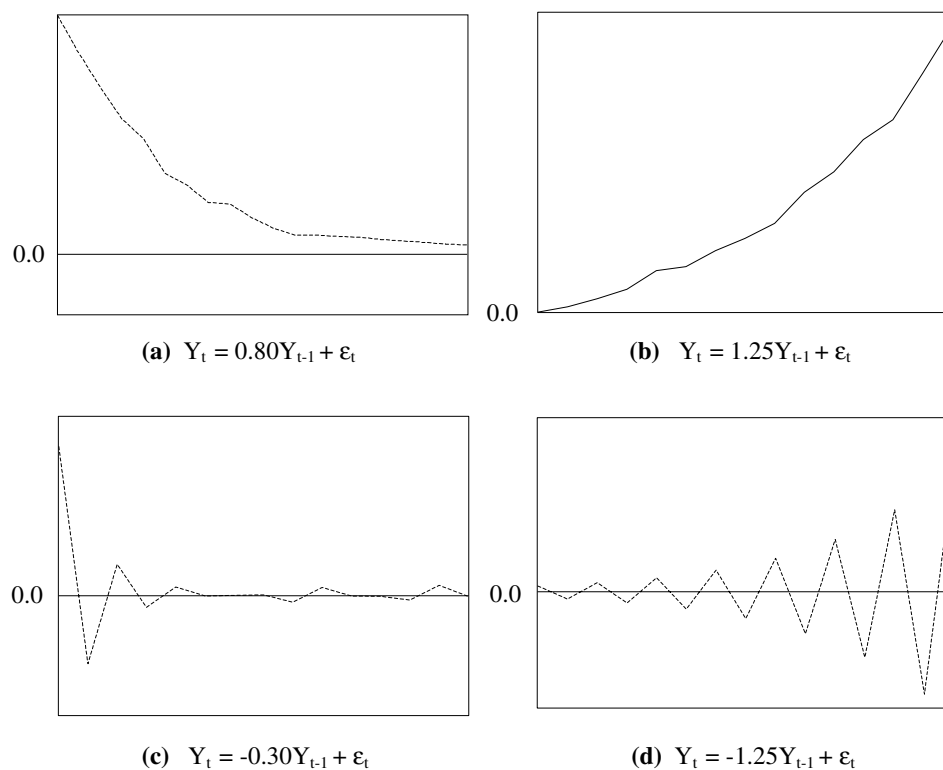


Figura 15.3
Diferentes trayectorias que puede optar una serie que presenta un único rezago y un término estocástico

⁶ Es importante resaltar que ante la presencia de un valor inicial de Y_t (Y_0), la solución para la expresión (15.2) se podría calcular sin incorporar las dos restricciones ($|a_1| < 1$ y $m \rightarrow \infty$):

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0 + \sum_{i=0}^{t-1} a_1^i \varepsilon_{t-i}$$

Y_t está expresado en función de constantes, de términos aleatorios, del tiempo y de un valor conocido de Y_t (Y_0).

La solución de la expresión (15.2) ha sido sencilla; sin embargo, el investigador se topará con expresiones (procesos estocásticos) un tanto más complejos por lo que se requiere contar con una metodología de resolución.

Tal como hemos indicado al inicio de esta sección, las series de tiempo pueden ser tratadas como ecuaciones en diferencia. Si el lector recuerda cómo desarrollarlas estará en mejor posición para asimilar la teoría que impartiremos en esta sección. Sin embargo, aquí se desarrollará los puntos necesarios para hacer entendible esta teoría y se destacará cuál es su importancia para este y el resto de los capítulos.

Como se ha indicado, nos interesa conocer la solución de Y_t ya que de esta manera podemos observar su comportamiento en el tiempo. La solución representada por la expresión (15.4) nos indica que la condición para que esta serie sea estable es que, $|a_1| < 1$, tal como lo observamos en los gráficos (15.3). Sin embargo, podemos “caracterizar” el comportamiento de cualquier proceso a partir de sus raíces características. La condición sobre la cual descansa la convergencia de cualquier serie radica en el valor de sus raíces características. Todas ellas deben ser menores a 1 en valor absoluto⁷.

Ahora surge la pregunta natural: ¿cómo se calculan las raíces características?. Su cálculo está ligado al cálculo de las soluciones homogéneas de las ecuaciones en diferencia. Con un ejemplo trataremos de facilitar el entendimiento de estos puntos.

Si tuviésemos la expresión:

$$y_t = 0.30y_{t-1} + 0.40y_{t-2} + 2 \quad (15.5)$$

Para calcular las soluciones homogéneas (es decir, encontrar una expresión para Y_t que haga que la suma de los componentes de Y_t de la expresión anterior sea nula) se procede a reemplazar $Y_t = A\alpha^t$ y obtenemos⁸:

$$A\alpha^t - 0.30A\alpha^{t-1} - 0.40A\alpha^{t-2} = 0$$

si dividimos esta expresión por $A\alpha^{t-2}$ obtendremos:

$$\alpha^2 - 0.30\alpha - 0.40 = 0$$

Esta expresión es denominada ecuación característica y su resolución nos da los valores de las raíces características:

$$\begin{aligned} \alpha_1 &= 0.80 \\ \alpha_2 &= -0.50 \end{aligned}$$

Ambas raíces son menores a 1 en valor absoluto, por lo que la expresión (15.5) es estable y converge a su valor medio. ¿Cuál es la trayectoria de convergencia?. Para el caso de la expresión (15.2), las distintas formas de convergencias se limitaban a 2: o bien se acercaban “suavemente” a la media o bien lo hacían “oscilando” (Figura 15.3 a y c).

⁷ La condición de estabilidad para la expresión (15.2) era que $|a_1| < 1$ ya que este coeficiente era a su vez la única raíz característica. El número de raíces características depende del grado de la expresión.

⁸ Las constantes y variables diferentes a Y_t son anuladas.

Sin embargo, para expresiones del tipo (15.5), la forma de convergencia (o no convergencia) es mucho más variable. Para conocerlas tenemos que generalizar la expresión (15.5):

$$y_t - a_1 y_{t-1} - a_2 y_{t-2} = 0$$

(ya se ha eliminado el término constante y el término generador).

Luego se construye la ecuación característica:

$$\alpha^2 - a_1 \alpha - a_2 = 0$$

en donde las raíces características adoptan los valores de:

$$\begin{aligned} \alpha_1, \alpha_2 &= (a_1 \pm \sqrt{a_1^2 + 4a_2})/2 \\ &= (a_1 \pm \sqrt{d})/2 \end{aligned}$$

donde “d” es el discriminante. El comportamiento de la serie de tiempo (su convergencia o no convergencia) va a depender del valor que adopte el discriminante. De tal forma que se pueden presentar tres casos:

CASO 1: $d > 0$, entonces, $a_1^2 + 4a_2 > 0$

Se puede ejemplificar este caso utilizando la expresión (15.5):

$$Y_t = 0.30Y_{t-1} + 0.40Y_{t-2} + 2$$

en donde los coeficientes eran $a_1 = 0.30$ y $a_2 = 0.40$. El valor del discriminante es de:

$$d = (0.30)^2 + 4(0.40) = 1.69 > 0$$

La ecuación característica es igual a: $\alpha^2 - 0.30\alpha - 0.40 = 0$; y sus raíces son: $\alpha_1 = 0.80$ y $\alpha_2 = -0.50$. Ambas son menores a la unidad en valor absoluto por lo que su trayectoria será convergente a su valor medio⁹. Se puede observar en la Figura (15.4a).

⁹ Con tan sólo observar el valor del discriminante y el de las raíces características podemos generalizar la forma de convergencia o no convergencia de la serie. Sin embargo, si construimos la solución homogénea vamos a determinar con mayor exactitud esta trayectoria. Esta solución tiene la siguiente estructura:

$$Y_t^h = A_1(\alpha_1)^t + A_2(\alpha_2)^t$$

Para calcular las constantes A_1 y A_2 es necesario contar dos condiciones adicionales (dos valores dados de Y_t). La trayectoria mostrada en la Figura (15.4a) es elaborada asumiendo que el valor que adoptan estos coeficientes es igual a la unidad por lo que toma esta forma de comportamiento.

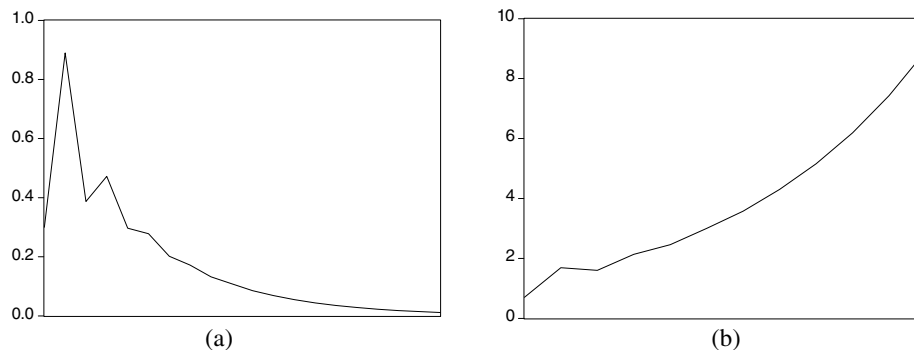


Figura 15.4: Caso $d > 0$

En la parte (a), el valor absoluto de las raíces unitarias no exceden a la unidad. En la parte (b) no se cumple esto por lo que la serie no converge.

Como segundo ejemplo se emplea la siguiente expresión:

$$Y_t = 0.70Y_{t-1} + 0.60Y_{t-2} - 6$$

los coeficientes tienen los siguientes valores: $a_1 = 0.70$ y $a_2 = 0.60$ por lo que el discriminante es mayor a cero ($d=2.89$). La ecuación característica es igual a: $\alpha^2 - 0.70\alpha - 0.60 = 0$, por lo que sus raíces adoptan los siguientes valores: $\alpha_1 = 1.20$ y $\alpha_2 = -0.50$. Como se puede apreciar en la Figura (15.4 b), el hecho de presentar raíces unitarias mayores a la unidad (en valor absoluto) provoca que la serie no sea estacionaria.

CASO 2: $d = 0$, entonces, $a_1^2 + 4a_2 = 0$

Si el valor del discriminante es cero, las raíces características son todas iguales a $a_1 / 2$. Como ya se ha dicho, la condición de estabilidad considera que las raíces deben ser menores a 1, y para este caso específico, esta condición equivale a decir que $|a_1| < 2$.

Ejemplo: $Y_t = 1.2Y_{t-1} - 0.36Y_{t-2} - 0.40$

Rápidamente podemos comprobar que el discriminante es igual a cero y que ambas raíces características adoptan el valor de 0.60 ¹⁰.

¹⁰ Para el caso en que el discriminante sea igual a cero, la solución homogénea en la mayoría de los casos está determinada por la siguiente estructura:

$$Y_t^h = A_1(\alpha_1)^t + A_2 t (\alpha_2)^t$$

El segundo componente del lado derecho de esta expresión es quien provoca el comportamiento creciente y luego decreciente de Y_t .

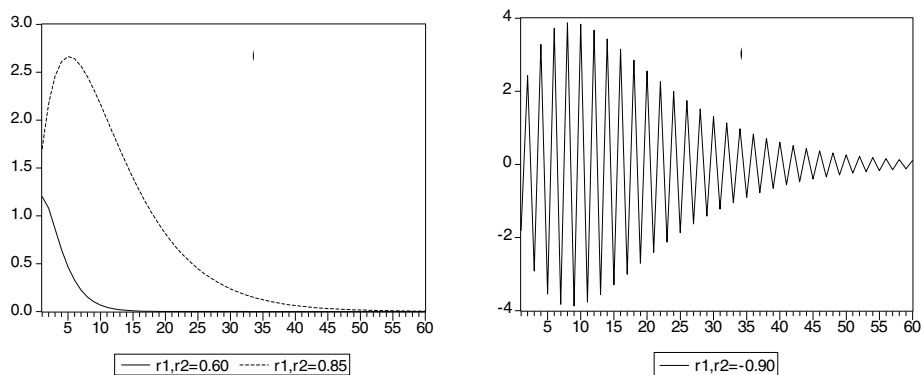


Figura 15.5: Caso d=0

Diferentes trayectorias que puede adoptar la serie dependiendo de los valores que tomen las raíces características. En la parte (a), los valores de las mismas son de 0.60 en un caso y de 0.85 en otro caso. En la parte (b), las raíces adoptan un valor negativo (-0.90) por lo que la serie oscila hasta alcanzar la convergencia.

CASO 3: $d < 0$, entonces, $a_1^2 + 4a_2 < 0$

En el caso que el discriminante sea inferior a cero, las raíces características tendrán un componente imaginario:

$$\alpha_{1,2} = (a_1 \pm i\sqrt{-d}) / 2$$

donde $i = \sqrt{-1}$.

Ejemplo: $Y_t = 1.2Y_{t-1} - 0.90Y_{t-2} - 0.22$

En este caso, el discriminante es menor a cero por lo que las raíces serán complejas, las mismas que se pueden comportar de la siguiente manera¹¹:

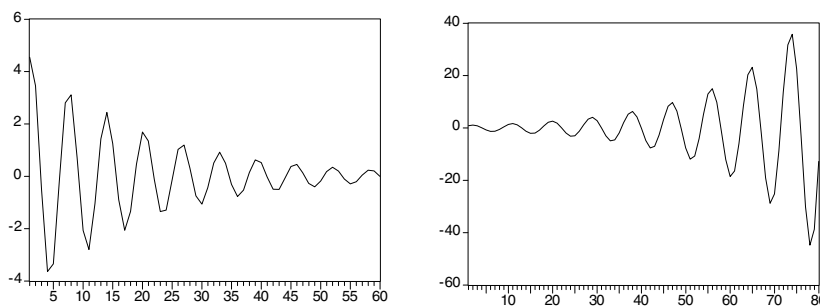


Figura 15.6: Caso d<0

La presencia de raíces complejas provoca un trayectoria sinusoidal decreciente o creciente según sea el caso.

¹¹ La trayectoria sinusoidal es propia de este caso ($d < 0$). En los casos anteriores se ha visto que las series oscilan aparentando una trayectoria sinusoide; sin embargo, esto sólo era aparentemente.

La convergencia en este tipo de series está determinada por la siguiente condición:

$$r = \sqrt{-a_2} < 1$$

En la parte (a) de la figura esta condición se cumple por lo que converge, mientras que en la parte (b) esto no se cumple¹².

El Círculo Unitario

Una vez que hemos visto las diferentes formas que pueden adoptar las raíces características, procederemos a precisar qué es lo que se entiende por “la condición de estabilidad del círculo unitario”.

El círculo unitario es un círculo de radio igual a 1 cuyo eje horizontal está compuesto por un segmento de la recta de número reales, mientras que el eje vertical está compuesto por los números imaginarios. El punto de intersección de ambas rectas está dado en el punto (0,0).

Podemos fácilmente ubicar las raíces características para los casos 1 y 2 ya que en ambos casos sólo se está tomando posesión de un punto de la recta horizontal. Para ubicar a las raíces características del caso 3 es necesario recordar:

$$\alpha_{1,2} = \frac{a_1}{2} \pm i \frac{\sqrt{-d}}{2}$$

Si a_1 fuese positivo, entonces las raíces características están ubicado en los cuadrantes I y IV de la Figura (15.7). La abscisa va a ser la misma para ambas raíces ($a_1/2$); mientras que la ordenada adopta los valores de $\pm i \frac{\sqrt{-d}}{2}$.

¹² La solución homogénea para un proceso de este tipo es la siguiente:

$$Y_t^h = A_1 r^t \cos(\theta t + A_2)$$

donde: $r = \sqrt{-a_2}$
 $\cos(\theta) = a_1 / [2\sqrt{-a_2}]$

La amplitud de las oscilaciones depende del término “r”. Otra vez, A_1 y A_2 son constantes.

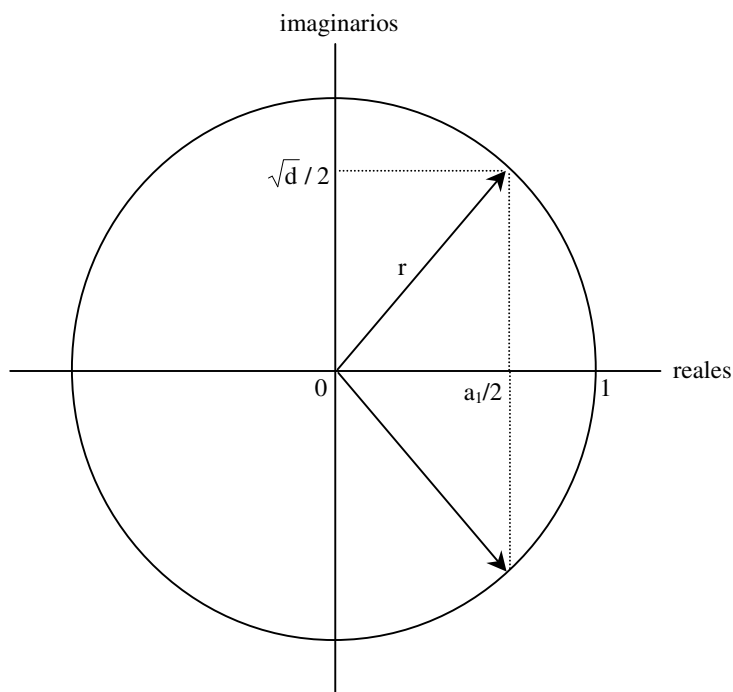


Figura 15.7: El Círculo Unitario

El eje de las abscisas está constituido por los números reales; mientras que el eje de las ordenadas, por número imaginarios. El radio de este círculo es igual a la unidad y en él se pueden representar los tres tipos de raíces características que un proceso puede presentar.

Sin embargo, en ambos casos, la extensión del radio es igual a:

$$r = \sqrt{(a_1 / 2)^2 + (i\sqrt{d} / 2)}$$

Si se desarrolla se va a llevar a que el radio es igual a: $r = \sqrt{-a_2}$, y de acuerdo con el caso #3, $r < 1$ para que exista la convergencia.

De esta manera, para que cualquier tipo de serie sea estacionaria, todas sus raíces características deben situarse dentro de un círculo de radio igual a 1.

Generalizando: si tenemos procesos de mayor grado, tal como el siguiente:

$$Y_t = \sum_{i=1}^n a_i Y_{t-i}$$

se tienen que calcular sus “n” raíces características procedentes de esta ecuación:

$$\alpha^n - a_1\alpha^{n-1} - \dots - a_n = 0$$

la estabilidad de la serie Y_t requiere que todas las raíces caigan dentro del círculo unitario. Para el caso particular de contar con sólo raíces reales, esta condición se restringe a que las raíces sean en valor absoluto menores a 1.

Al ser muy difícil calcular las raíces unitarias de grados mayores al cúbico, especificaremos algunas condiciones necesarias o suficientes que deben cumplir los coeficientes para que la serie sea estacionaria.

- Condición necesaria: $\sum_{i=1}^n a_i < 1$
- Condición suficiente: $\sum_{i=1}^n |a_i| < 1$

En el caso que $\sum_{i=1}^n a_i = 1$, esto significa que existe una raíz igual a 1 por lo que el proceso llevará el nombre de “proceso de raíz unitaria”¹³.

Antes de terminar esta sección debemos advertirle a nuestros lectores que es necesario que refresquen su memoria con respecto al uso del operador de rezagos. Si bien esta herramienta ya ha sido empleada con anterioridad, es necesario conocer todas sus características y propiedades. Para facilitarles la tarea, en el Apéndice 15.1 se especifican algunas propiedades de este operador que serán de bastante utilidad.

15.4. MODELOS DE SERIES DE TIEMPO

Durante el desarrollo de la econometría se han elaborado diferentes métodos que buscaban dotar a los investigadores de herramientas para poder realizar estimaciones acerca de los valores futuros de las variables bajo análisis. Uno de los primeros métodos consistía en el empleo de los modelos de regresión uniecuacionales. Estos buscan recoger de la teoría económica las relaciones a establecer entre las variables que van a conformar el modelo. Por ejemplo, la demanda por dinero debería estar explicada por el nivel de ingreso de los agentes, la tasa de interés y la inflación esperada. Luego, con el empleo de las distintas series se estima cuáles son las variables significativas y sus coeficientes. Una vez determinada las variables relevantes y todos los parámetros, el modelo puede ser utilizado para predecir la demanda por dinero futuro. Sin embargo, el grado de ajuste de este tipo de modelo se iba deteriorando marcadamente conforme se alejase en el futuro.

Otra técnica bastante utilizada es el sistema de ecuaciones. Estas fueron empleadas con el objetivo de predecir variables macroeconómicas a través de una compleja gama de ecuaciones que buscaba recoger toda relación existente entre las variables bajo análisis. Sin embargo, tras la pésima performance que presentaron durante la crisis del petróleo y la crítica de Lucas, estos modelos fueron dejados de lado. La crítica se centra principalmente en que los parámetros calculados por el modelo dependen de la política prevaleciente en el momento en que se realiza la estimación. Esto implica que ante un cambio de la misma, la probabilidad de que los parámetros dejen de ser significativos es alta.

¹³ Este tipo de procesos son denominados procesos con tendencias estocásticas, cómo ya se mencionó en la sección 15.2, estos serán analizados en el próximo capítulo.

A finales de la década de los setenta, Box y Jenkins desarrollaron una nueva herramienta de predicción denominado como modelo autorregresivo integrado de media móvil (ARIMA). Este método antes de tratar de realizar un análisis económico, se concentra en las propiedades probabilísticas de las series de tiempo. En su forma más general está representado de la siguiente manera:

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_q \varepsilon_{t-q}$$

Un modelo ARIMA estacionario es denominado ARMA. A continuación vamos a ir explicando las diferentes formas estructurales que un modelo ARIMA puede adoptar dependiendo de la inclusión de los rezagos de la variable Y (componente autorregresivo-AR), de los rezagos de los errores (componente de medias móviles-MA) o bien de ambos a la vez (modelos ARMA).

Sin embargo, antes de empezar a describir cada uno de los modelos es necesario explicar el comportamiento del elemento básico de los modelos estocásticos discretos de series de tiempo: **el proceso ruido blanco** (*white-noise process*). Una serie $\{\varepsilon_t\}$ es denominada ruido blanco si su media es igual a cero, su varianza es constante y los elementos de la serie no se correlacionan entre sí.

$$E(\varepsilon_t) = E(\varepsilon_{t-1}) = \dots = 0$$

$$E(\varepsilon_t^2) = E(\varepsilon_{t-1}^2) = \dots = \sigma^2 \quad \left[\text{o bien } \text{var}(\varepsilon_t) = \text{var}(\varepsilon_{t-1}) = \dots = \sigma^2 \right]$$

y para todo j

$$E(\varepsilon_t \varepsilon_{t-s}) = E(\varepsilon_{t-j} \varepsilon_{t-j-s}) = 0 \quad \left[\text{o bien } \text{cov}(\varepsilon_t, \varepsilon_{t-s}) = \text{cov}(\varepsilon_{t-j}, \varepsilon_{t-j-s}) = 0 \right]$$

De la forma en que se ha definido el proceso ruido blanco es fácil determinar que este cumple con las características de estacionariedad débil ya que mantiene constante sus dos primeros momentos a lo largo de la serie.

Otro proceso estocástico discreto es el denominado **paseo aleatorio** (*random walk*). Se caracteriza porque presenta como series de primeras diferencia a valores aleatorios. Si suponemos que $\{\varepsilon_t\}$ es una serie puramente aleatoria (*white-noise*) con media μ y varianza σ^2 , definimos que la serie $\{Y_t\}$ sigue un *random walk* si:

$$Y_t = Y_{t-1} + \varepsilon_t$$

Si suponemos que la primera observación, Y_0 , es igual a cero, entonces el proceso estocástico sigue el siguiente comportamiento:

$$Y_1 = \varepsilon_1$$

$$Y_2 = Y_1 + \varepsilon_2 = \varepsilon_1 + \varepsilon_2$$

en forma general:

$$Y_t = \sum_{i=1}^t \varepsilon_i$$

De esta manera podemos definir a $E(Y_t) = t\mu$ y $\text{var}(Y_t) = t\sigma^2$. Esta serie al no presentar una media ni una varianza constante, entonces no es estacionaria, aunque su primera diferencia sí lo sea.

A continuación presentamos los procesos estocásticos discretos (PED) que servirán como estructuras a asociar con las series de datos económicas que se estén empleando en un estudio. Si conocemos las características de los principales PED, entonces una vez hecha la asociación con una serie de datos, el investigador obtendrá las características de esta última.

15.4.1 Modelos Autorregresivos (AR)

Si se vuelve a suponer que $\{\varepsilon_t\}$ es un proceso aleatorio con media 0 y varianza σ^2 , entonces el proceso $\{Y_t\}$, definido de la siguiente manera:

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_p Y_{t-p} + \varepsilon_t \quad (15.6)$$

se lo denomina como un proceso autorregresivo de orden p y se le denota como $AR(p)$. En términos del operador de rezagos L , la expresión anterior puede ser representada como

$$Y_t = (a_1 L + a_2 L^2 + \dots + a_p L^p) Y_t + \varepsilon_t$$

o bien

$$(1 - a_1 L - a_2 L^2 - \dots - a_p L^p) Y_t = \varepsilon_t \quad (15.7)$$

despejando Y_t obtenemos:

$$\begin{aligned} Y_t &= \frac{1}{(1 - a_1 L - a_2 L^2 - \dots - a_p L^p)} \varepsilon_t \\ &= \frac{1}{(1 - \pi_1 L)(1 - \pi_2 L) \dots (1 - \pi_p L)} \varepsilon_t \end{aligned}$$

donde $\pi_1, \pi_2, \dots, \pi_p$ son las raíces de la ecuación

$$(\pi)^t - a_1 (\pi)^{t-1} - a_2 (\pi)^{t-2} - \dots - a_p (\pi)^{t-p} = 0$$

como ya se ha indicado en la sección 15.3, la condición para que la varianza de esta serie sea finita, es decir, que sea estacionaria, implica que las raíces características sean menores a uno en valor absoluto ($|\pi_i| < 1$) o bien que el término L sea mayor a 1 en valor absoluto¹⁴.

Podemos emplear otro método más intuitivo y operativo para determinar las condiciones para que un proceso $AR(1)$ sea estacionario¹⁵. Si definimos al proceso $AR(1)$ de la siguiente forma:

$$Y_t = a_0 + a_1 Y_{t-1} + \varepsilon_t \quad (15.8)$$

donde $\{\varepsilon_t\}$ es un ruido blanco. Si suponemos que en el periodo de tiempo inicial, la serie $\{Y_t\}$ adopta el valor y_0 , entonces la solución de esta ecuación sería:

¹⁴ Que es lo mismo que decir que caigan dentro y fuera del círculo unitario, respectivamente: $L_i = 1/\pi_i$

¹⁵ Esto también se vio en la sección anterior. En el caso que el lector no domine bien estos métodos, les recomendamos que efectúe una rápida revisión.

$$Y_t = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0 + \sum_{i=0}^{t-1} a_1^i \varepsilon_{t-i} \tag{15.9}$$

Tomando el valor esperado de la expresión (15.9), obtenemos:

$$E(Y_t) = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0 \tag{15.10}$$

el valor esperado de la serie “k” periodos después será igual a:

$$E(Y_{t+k}) = a_0 \sum_{i=0}^{t+k-1} a_1^i + a_1^{t+k} y_0 \tag{15.11}$$

Si comparamos las expresiones (15.10) y (15.11), es evidente que ambas nos están presentando diferentes medias. La media para un proceso AR(1) no es independiente del tiempo. Sin embargo, si se tratase de un periodo de tiempo lo suficientemente prolongado y además que $|a_1| < 1$, entonces la expresión $(a_1)^t$ convergerá a cero y la sumatoria $a_0[1+a_1+(a_1)^2+(a_1)^3+...]$ en el límite será igual a $a_0/(1-a_1)$. De esta manera, cuando $t \rightarrow \infty$ y $|a_1| < 1$,

$$\lim y_t = a_0/(1-a_1) + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \tag{15.12}$$

y si ahora tomamos el valor esperado a la expresión (15.12) para valores de t lo suficientemente grande, obtendremos que $E(y_t) = a_0/(1-a_1)$. De esta manera el valor de la media de y_t es finita e independiente del tiempo: $E(y_t) = E(y_{t-s}) = \mu$. En lo que se refiere al valor que la varianza alcanzará en el límite, este está dado por la siguiente expresión:

$$\begin{aligned} E(y_t - \mu)^2 &= E\left[\left(\varepsilon_t + a_1 \varepsilon_{t-1} + (a_1)^2 \varepsilon_{t-2} + \dots\right)^2\right] \\ &= \sigma^2 \left[1 + (a_1)^2 + (a_1)^4 + \dots\right] = \sigma^2 / \left[1 - (a_1)^2\right] \end{aligned}$$

que al igual que la media es finito e independiente del tiempo. Finalmente, y para determinar si el proceso AR(1) es estacionario o no, haremos referencia al comportamiento de los valores de la autocovarianza. Estos son finitos e independientes del tiempo. Estos dependen de la extensión del rezago “k”:

$$\begin{aligned} E\left[(y_t - \mu)(y_{t-k} - \mu)\right] &= E\left\{\left[\varepsilon_t + a_1 \varepsilon_{t-1} + (a_1)^2 \varepsilon_{t-2} + \dots\right] \left[\varepsilon_{t-k} + a_1 \varepsilon_{t-k-1} + (a_1)^2 \varepsilon_{t-k-2} + \dots\right]\right\} \\ &= \sigma^2 (a_1)^k \left[1 + (a_1)^2 + (a_1)^4 + \dots\right] \\ &= \sigma^2 (a_1)^k / \left[1 - (a_1)^2\right] \end{aligned} \tag{15.13}$$

De esta manera, el proceso AR(1) propuesto en la expresión (15.8) es estacionario si contamos con un t lo suficientemente grande y que $|a_1| < 1$. Sin embargo, si la muestra es generada por un proceso que recién ha empezado, su realización no sería estacionaria.

Recordemos que es muy importante que un proceso sea estacionario en media ya que en caso que esta característica no se presentase, el investigador no podría realizar estimaciones “confiables”¹⁶ de la serie. Lo mismo sucedería si la varianza presentase un comportamiento heterocedástico. Lo peor que pudiese pasar al no ser estacionaria una serie es que no existiese una distribución detrás de los datos que tenemos.

Si ahora quisiéramos determinar cuáles son las condiciones que debe cumplir un proceso AR(2) para que sea estacionario, el empleo del método iterativo resultaría muy engorroso por lo que se emplea las soluciones homogéneas. Si definimos el proceso estocástico AR(2) de la siguiente manera:

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \varepsilon_t \quad (15.15)$$

para calcular sus raíces características lo tendríamos que desarrollar aplicando el método de ecuaciones en diferencia, es decir, calcular las soluciones homogéneas: $Y_t^h = A(\pi)^t$. Reemplazando esta expresión en (15.15) obtendremos:

$$A(\pi)^t - a_1 A(\pi)^{t-1} - a_2 A(\pi)^{t-2} = 0 \quad (15.16)$$

dividiendo todo entre $A(\pi)^{t-2}$, obtenemos la siguiente ecuación:

$$\pi^2 - a_1 \pi - a_2 = 0$$

de donde se puede inferir que para que los valores de las raíces características sean en valor absoluto menores a 1 ($|\pi_i| < 1$) implica que:

$$\left| \frac{a_1 \pm \sqrt{a_1^2 + 4a_2}}{2} \right| < 1$$

Esto nos da las condiciones que debe cumplir los coeficientes del modelo AR(2) para que sean estacionarios:

$$\begin{aligned} a_2 + a_1 &< 1 \\ a_2 - a_1 &< 1 \\ |a_2| &< 1 \end{aligned}$$

15.4.2 Modelos de Medias Móviles (MA)

Un ejemplo citado en muchos libros para indicar al lector cómo es que se producen procesos tipo MA es empleando un juego de suma cero. Por ejemplo, aquel que involucra tirar una moneda y ganar S./1 si sale cara o perder el mismo monto si sale sello. En promedio la ganancia a obtener sería representada como un ruido blanco (ε_t). Si lanzamos la moneda cuatro veces, tu ganancia (Y_t) sería igual a:

¹⁶ Hemos visto en la Figura (15.1) que el comportamiento del componente irregular puede ser predicho siempre que las series sean estacionarias.

$$Y_t = \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \beta_3 \varepsilon_{t-3}$$

donde β_i son las probabilidades de ocurrencia. Para $i \leq 3$, este coeficiente es igual a 0.25; para $i > 3$, es igual a cero. Sin embargo, a pesar que el proceso MA(q) está integrado por puros ε_t (no incluiremos la constante), esto no significa que un proceso MA(q) también sea ruido blanco.

Generalizando, suponiendo que $\{\varepsilon_t\}$ es un proceso aleatorio con media 0 y varianza σ^2 , entonces el proceso $\{Y_t\}$, definido de la siguiente manera

$$Y_t = \varepsilon_t - \beta_1 \varepsilon_{t-1} - \dots - \beta_q \varepsilon_{t-q} \quad (15.17)$$

es conocido como proceso de medias móvil de orden q, y se denota como MA(q). La característica de este tipo de proceso es que es estacionario por definición. La esperanza de Y_t es igual a cero ya que la esperanza de cada uno de los elementos de $\{\varepsilon_t\}$ es nula. Asimismo, la varianza de nuestro proceso es determinada independientemente del tiempo. Esta depende del número de términos que posea Y_t , es decir, de su rango, conjuntamente con el valor que adopten los coeficientes:

$$\text{var}(Y_t) = \left(1 + \sum_{i=1}^q \beta_i^2 \right) \sigma^2$$

El proceso MA(q) puede escribirse utilizando el operador de rezagos:

$$Y_t = (1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q) \varepsilon_t = \beta(L) \varepsilon_t \quad (15.18)$$

Si bien sabemos que el proceso Y_t es estacionario, la expresión (15.18) se parece mucho a la (15.7). En esta última, la variable Y_t estaba multiplicada por un polinomio de rezagos y nos importaba contar con raíces características que caigan dentro del círculo unitario. Ahora la figura es la misma pero la variable que se va a explicar es ε_t . De esta manera, si las raíces características de la ecuación (15.18), que puede ser replanteada así,

$$Y_t = (1 + \pi_1 L)(1 - \pi_2 L) \dots (1 - \pi_q L) \varepsilon_t \quad (15.19)$$

son menores a uno en valor absoluto ($|\pi_i| < 1$), entonces el proceso MA es invertible. Esto implica que se puede transformar un proceso MA(q) en un proceso AR(∞): $\varepsilon_t = [\beta(L)^{-1} Y_t]$.

Podemos determinar con un ejemplo las condiciones de la invertibilidad. Para esto usaremos un proceso MA(1):

$$Y_t = \varepsilon_t - \beta_1 \varepsilon_{t-1}$$

si despejamos el error de esta expresión y la vamos rezagando, obtenemos:

$$\begin{aligned} \varepsilon_t &= Y_t + \beta_1 \varepsilon_{t-1} \\ \varepsilon_{t-1} &= Y_{t-1} + \beta_1 \varepsilon_{t-2} \\ \varepsilon_{t-2} &= Y_{t-2} + \beta_1 \varepsilon_{t-3} \end{aligned}$$

Si se reemplaza ε_{t-1} y ε_{t-2} en la expresión de ε_t , obtendremos como resultado:

$$\begin{aligned}\varepsilon_t &= Y_t + \beta_1(Y_{t-1} + \beta_1\varepsilon_{t-2}) = Y_t + \beta_1 Y_{t-1} + \beta_1^2\varepsilon_{t-2} \\ \varepsilon_t &= Y_t + \beta_1 Y_{t-1} + \beta_1^2(Y_{t-2} + \beta_1\varepsilon_{t-3}) \\ &= Y_t + \beta_1 Y_{t-1} + \beta_1^2 Y_{t-2} + \beta_1^3\varepsilon_{t-3}\end{aligned}$$

luego, al despejar Y_t :

$$Y_t = -\beta Y_{t-1} - \beta^2 Y_{t-2} - \beta^3 \varepsilon_{t-3} + \varepsilon_t$$

En el caso que se hubiese continuado reemplazado los errores rezagados, se llegaría a una expresión como esta:

$$Y_t = -\sum_{k=1}^{\infty} \beta_1^k Y_{t-k} + \varepsilon_t$$

Para llevar a esta expresión se elimina siempre el error más rezagado, lo cual da paso a un nuevo rezago de la variable Y_t . Este proceso puede ser extendido hasta el infinito. Para que esta “inversión” de un proceso MA en un proceso AR sea finita y tenga sentido, el coeficiente β_1 debe cumplir con: $|\beta_1| < 1$. Es finito porque en un punto de esta sumatoria va a empezar a presentarse rezagos de Y_t que al estar multiplicado por un coeficiente tan cercano a cero, dejarán de ser significativos. Además, es más lógico pensar que los rezagos más recientes tengan mayor importancia (influyen más) en el valor corriente de Y_t . Por consiguiente, cuando $|\beta_1| < 1$, entonces un proceso MA(1) es invertible¹⁷.

Si tomamos como ejemplo un proceso MA de orden 2, $Y_t = (1 + \beta_1 L + \beta_2 L^2)\varepsilon_t$, las raíces características vendrían a estar dadas a partir de la ecuación:

$$\pi^2 + \beta_1 \pi + \beta_2 = 0$$

y para que se pueda dar la invertibilidad del proceso MA, se debe cumplir con esta relación: $|\pi_i| < 1$ por lo que el valor de los coeficientes está restringido a cumplir con:

$$\left| \frac{-\beta_1 \pm \sqrt{\beta_1^2 - 4\beta_2}}{2} \right| < 1$$

desarrollando obtenemos:

$$\begin{aligned}\beta_1 + \beta_2 &> -1 \\ \beta_2 + \beta_1 &> -1 \\ |\beta_2| &< 1\end{aligned}$$

¹⁷ La importancia de la invertibilidad de un proceso MA radica en que para hacer predicciones es necesario contar con una estructura autorregresiva.

Procesos de media móvil puede también presentarse como resultado de la eliminación de la tendencia de una serie. Generalmente para limpiar una serie de su tendencia es necesario tomar sucesivas diferencias a la misma. Si tuviésemos el siguiente modelo:

$$Y_t = a_0 + a_1 t + a_2 t^2 + \varepsilon_t \quad (15.20)$$

y quisiésemos eliminar su tendencia, entonces la expresión (15.20) tendría que ser rezaga y luego restada del modelo original. El resultado obtenido ahora ha dejado de presentar una tendencia cuadrática y ahora presenta una tendencia lineal. Además, el modelo resultante está compuesto por un promedio móvil que podría indicar un ciclo.

15.4.3 Modelos Autorregresivos de promedio móvil (ARMA)

No es frecuente encontrar modelos que sean puramente AR o MA. Muchas series económicas presentan ambas estructuras. A este tipo de modelos, que cuentan con un componente autorregresivo (AR) y uno de medias móvil (MA), se los conocen como ARMA (p,q) y se definen de la siguiente manera:

$$Y_t = a_1 Y_{t-1} + \dots + a_p Y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} \quad (15.21)$$

Un proceso ARMA (p,0) representa a un proceso AR(p); de igual forma, un proceso ARMA(0,q) no es más que un proceso MA(q) puro.

El más sencillo proceso ARMA es el ARMA(1,1), y es representado de la siguiente manera:

$$Y_t = a_1 Y_{t-1} + \varepsilon_t - \beta_1 \varepsilon_{t-1} \quad (15.22)$$

Si se empieza a rezagar a Y_t y a la vez se reemplaza en la expresión (15.22), se llega a:

$$Y_t = (a_1 - \beta_1) \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i-1} + \varepsilon_t \quad (15.23)$$

La esperanza de esta expresión es igual a cero y su varianza es:

$$\begin{aligned} \text{var}(Y_t) &= (a_1 - \beta_1)^2 \sum_{i=0}^{\infty} a_1^{2i} \sigma_\varepsilon^2 + \sigma_\varepsilon^2 \\ &= \frac{(\beta_1^2 + 1 - 2a_1\beta_1)}{(1 - a_1^2)} \sigma_\varepsilon^2 \end{aligned} \quad (15.24)$$

Tanto la expresión (15.23) como la (15.24) nos indican que la condición necesaria para que la serie Y_t sea estacionaria es que el coeficiente a_1 sea menor a 1 en valor absoluto. Para determinar la condición de invertibilidad emplearemos la misma forma que se utilizó con el MA:

La expresión (15.22) debe ser reordenada de la siguiente forma:

$$\varepsilon_t = Y_t - a_1 Y_{t-1} + \beta_1 \varepsilon_{t-1}$$

y si lo rezagamos continuamente podemos obtener:

$$\varepsilon_t = Y_t - (a_1 - \beta_1) \sum_{i=0}^{\infty} \beta_1^i Y_{t-i-1}$$

por consiguiente, la condición para su invertibilidad se centra en que β_i cumpla con: $|\beta_i| < 1$.

Un proceso ARMA también puede ser representado utilizando el operador de rezagos (L) y obtener:

$$\phi(L)Y_t = \theta(L)\varepsilon_t$$

donde $\phi(L)$ y $\theta(L)$ son polinomios de órdenes p y q, respectivamente, y están definidos de la siguiente forma:

$$\begin{aligned} \phi(L) &= 1 - a_1L - a_2L^2 - \dots - a_pL^p \\ \theta(L) &= 1 - \beta_1L - \beta_2L^2 - \dots - \beta_qL^q \end{aligned} \tag{15.22}$$

Un proceso ARMA(p,q) es estacionario si lo es su componente AR, y es invertible si lo es su componente MA. Técnicamente hablando, para que un proceso ARMA sea estacionario se requiere que los valores que tomen los operadores de rezagos de $\phi(L)=0$ queden fuera del círculo unitario¹⁸. Similares requisitos son necesarios para que el componente MA sea invertible (claro está que se debe emplear las raíces características del polinomio $\theta(L)$). Si las raíces características de la expresión (15.22) caen dentro del círculo unitario, entonces el modelo es denominado autorregresivo con promedio móvil. En caso en que una o más raíces características no cumplan con las restricciones anteriormente mencionadas, la serie $\{Y_t\}$ es un proceso integrado y recibe el nombre de modelo autorregresivo integrado de promedios móvil (ARIMA).

15.4.4 Modelos Autorregresivos de promedio móvil integrados (ARIMA)

La mayor parte de las series de tiempo presentan un importante factor tendencial en su composición. Con el paso del tiempo el consumo, el tráfico, la demanda por trigo, la demanda por dinero, etc. se incrementa y esto no responde a shocks temporales sino a la normal evolución y crecimiento de la población. De esta manera las series pueden presentar una media cambiante con el paso del tiempo y esto no implicaría que este comportamiento se deba a alguna desviación producida por un shock. Para eliminar este comportamiento tendencial de las series es necesario tomar las primeras diferencias y de esta manera obtener una serie estacionaria. Otro origen de no estacionalidad se debe a que por la naturaleza de la serie, un shock produzca tenga un efecto permanente en el tiempo. A este proceso se le denomina raíz unitaria y se analizará en el próximo capítulo. Otra vez, la solución a este problema se obtiene diferenciando la serie.

Si observamos que para diferentes intervalos de tiempo, la serie a analizar presenta diferentes patrones de inestabilidad (volatilidad), entonces podemos concluir que no es estacionaria. Esto se debe a que la varianza de las observaciones cambia con el tiempo (presencia de heterocedasticidad). Ante este fenómeno, el investigador podrá aplicar logaritmos

¹⁸ O bien que las raíces características de ese polinomio caigan dentro del círculo unitario.

para estabilizar a la serie. Es importante recalcar que la aplicación de logaritmos no va a corregir el problema de heterocedasticidad sino que representa un método sencillo y efectivo para estabilizar a la serie. Luego de su aplicación, el investigador podrá continuar con los procedimientos de identificación y tratamiento descritos en este capítulo¹⁹.

Si, por ejemplo, analizáramos la serie del consumo privado, es evidente que su composición cuenta con un componente tendencial, de tal manera que podría ser representada de la siguiente manera:

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + a_3 Y_{t-3} + a_4 t + \varepsilon_t \quad (15.23)$$

muy al margen de las condiciones que deben cumplir sus raíces características para que este proceso autorregresivo sea estacionario, la presencia de la tendencia (t) alteraría constantemente la media de Y_t . Para eliminar este componente es necesario aplicarle la primera diferencia:

$$\Delta Y_t = a_1 \Delta Y_{t-1} + a_2 \Delta Y_{t-2} + a_3 \Delta Y_{t-3} + a_4 + \varepsilon_t - \varepsilon_{t-1} \quad (15.24)$$

En caso que la serie Y_t tuviese una tendencia cuadrática, sería necesario obtener la segunda diferencia ($\Delta^2 Y_t$). Generalizando, una serie que tenga que diferenciarse “d” veces para obtener una serie estacionaria ($\Delta^d Y_t$) es denominado un proceso ARIMA (p,d,q). Este tipo de series es denominada integrada de orden (d). Un modelo integrado de orden 1, I(1), es aquel que luego de una diferencia se convierte en estacionario. Si la serie es I(2), entonces será necesario hasta dos diferenciaciones para volverlo una serie I(0), es decir, estacionaria.

Antes del empleo de la metodología Box-Jenkins²⁰ se debe contar con la serie estacionaria. En el caso que esta no lo fuese, se deberá diferenciar cuantas veces sea necesario. Esta técnica busca establecer cual es el comportamiento de las series para de esta manera poder realizar sus predicciones. Si las series son estacionarias se puede confiar que las características que ha presentado hasta hoy se mantendrán durante el futuro.

¹⁹ En el Capítulo 17, Procesos de Heterocedasticidad Condicional Autorregresiva, se procede a determinar como modelar la varianza de la serie.

²⁰ Esta metodología es la más ampliamente utilizada en el análisis de series de tiempo. La misma que será tratada en la sección 15.6.

CAPITULO 16

MODELOS DE SERIES DE TIEMPO NO ESTACIONARIAS

16.1. INTRODUCCIÓN

El análisis estadístico tradicional de series de tiempo, presentado en capítulo anterior, supone la existencia de la propiedad estadística de la estacionariedad; sin embargo, muchas de las series económicas que observamos no cumplen con esta característica; al contrario presentan una tendencia creciente a lo largo del tiempo o una clara inclinación a permanecer durante largos periodos por encima o por debajo de su media.

Como se verá en el presente capítulo, la no estacionariedad en las series de tiempo se puede deber a la presencia de una tendencia determinística, al problema de las raíces unitarias o a la presencia de una varianza heterocedástica. Dependiendo del caso en que nos encontremos, la no estacionariedad tendrá determinadas implicancias teóricas y estadísticas, así como determinados procedimientos para transformar las series en estacionarias. Los investigadores han buscado formas para reconocer este tipo de series. Con este fin se han creado diversos tests estadísticos, algunos de los cuales se presentan al final del presente capítulo.

16.2. PROCESOS ESTACIONARIOS EN TENDENCIA Y ESTACIONARIOS EN DIFERENCIA

Muchos de los lectores deben haber notado que la mayoría de las series macroeconómicas presentan una tendencia a través del tiempo. Estas series con tendencia, claramente no son estacionarias, porque su valor medio cambia a través del tiempo. Si consideramos a y_t como la serie en análisis y ε_t un error ruido blanco, tenemos:

$$y_t = \alpha_0 + \alpha_1 t + \gamma(L)\varepsilon_t \quad (16.1)$$

donde α_t representa el componente tendencial y $\gamma(L)\varepsilon_t$ es el componente no tendencial (o serie removida la tendencia). Esta descomposición tradicional asume que $\gamma(L)\varepsilon_t$ es un proceso estacionario. Por consiguiente, la serie y_t se conoce como una variable *estacionaria en tendencia (TS)*¹, ya que a pesar de tener un componente tendencial que hace que su media varíe en el tiempo, las desviaciones de su tendencia determinística son estacionarias. Por varias razones es conveniente eliminar esta tendencia de las series (una de las cuales es volver la serie estacionaria). Tradicionalmente los investigadores removían la tendencia extrayendo de la serie la *tendencia determinística* expresada por $\alpha_{1,t}$, donde t es el índice del tiempo. Este procedimiento tradicional elimina el componente tendencial de serie y deja como resultado el componente cíclico o no tendencial.

Aproximadamente desde 1982, fecha en que se publica el trabajo de Nelson y Plosser², muchos investigadores han preferido usar un modelo alternativo para explicar algunas series económicas, el cual propone la siguiente formulación

$$\Delta y_t = \beta + A(L)\varepsilon_t \quad (16.2)$$

donde $\Delta y_t = y_t - y_{t-1}$, el componente $A(L)\varepsilon_t$ es estacionario y β representa el promedio de cambio por periodo de la variable y_t . Bajo esta formulación, la serie y_t es conocida como una variable *estacionaria en diferencia (DS)*³, ya que la serie necesita ser diferenciada para ser estacionaria.

La principal diferencia entre las variables estacionarias en tendencias (TS) y las estacionarias en diferencias (DS) es que la primera tiende a fluctuar alrededor de una tendencia determinística y la segunda no (salvo un caso específico). Como el componente $\gamma(L)\varepsilon_t$ en la ecuación (16.1) es estacionario con media cero, la serie y_t fluctúa alrededor de la función de tendencia $\alpha_{1,t}$. En cambio, en la formulación (16.2), la serie y_t crece en una cantidad β (constante) con respecto al periodo anterior, pero como esta cantidad es afectada por el componente estacionario $A(L)\varepsilon_t$, la medida a la que crece variará con el tiempo por lo que la serie no fluctuará alrededor de una tendencia.

El trabajo de Nelson y Plosser (1982)⁴ presenta evidencia de que muchas de las series económicas de E.E.U.U., inclusive la serie del PBI, son de la clase DS. Sobre la base de esto, argumentan que muchos de los ciclos económicos que presentó la economía norteamericana fueron consecuencia de shocks reales, en oposición al enfoque monetario, el cual propone que son los shocks no reales (monetarios) la principal fuente de los ciclos económicos. Si este fuera el caso, los shocks monetarios tendrían importancia secundaria en la formación de los ciclos. Estos argumentos no fueron convincentes para muchos macroeconomistas, pero la hipótesis de que muchas series económicas reflejan un proceso DS fue aceptada por varios años.

Las implicancias de que las series sean de clase DS y de TS, así como las consecuencias estadísticas de eliminar la no estacionariedad de las series serán discutidas en las siguientes líneas.

¹ Por sus siglas en inglés *Trend-Stationary*.

² Nelson, Charles R. y Plosser, Charles I., "Trends and Random Walks in Macroeconomic Time Series", *Journal of Monetary Economics*, North Holland P. C., 1982

³ Por sus siglas en inglés *Difference-Stationary*.

⁴ Nelson, Charles R. y Plosser, Charles I., op. cit.

16.3. TENDENCIAS DETERMINÍSTICAS O PROCESO ESTACIONARIO EN TENDENCIAS (TS)

Como se mencionó en la anterior sección, las series económicas pueden presentar algún tipo de tendencia determinística. La tendencia más simple y la más común es la tendencia lineal. Una serie que presenta tendencia lineal tiene el supuesto implícito que la serie posee una tasa de crecimiento de largo plazo; todas las desviaciones con respecto a su tendencia serán transitorias y la serie siempre regresará a su tendencia. Un ejemplo de tendencia lineal es la formulación (16.1) de la sección anterior. Se puede apreciar un ejemplo de tendencia lineal en la Figura 16.1.

De manera general, se pueden encontrar series que presenten tendencias polinomiales:

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \dots + \alpha_n t^n + \varepsilon_t \quad (16.3)$$

Algunas series económicas pueden presentar una tendencia exponencial. Una característica de la tendencia exponencial es que refleja un crecimiento proporcional constante. Si planteamos que :

$$y_t = e^{\delta t} \quad (16.4)$$

entonces $dy/dt = \delta y_t$. Por ejemplo, la población puede presentar un crecimiento proporcional constante. La causa de esto puede ser que el número de nacimientos sea una fracción constante de la población corriente.

Dado que este comportamiento ha sido observado en muchas variables económicas, muchos investigadores han supuesto que las variables económicas no estacionarias tienen una tendencia exponencial.

Debemos notar que las tendencias exponenciales, por su naturaleza, al tomarle el logaritmo natural se transforman en una tendencia lineal:

$$\text{Log}(y_t) = \delta t \quad (16.5)$$

Por ello es común tomar logaritmos a los datos antes de describirlos como si fueran de tendencia exponencial.

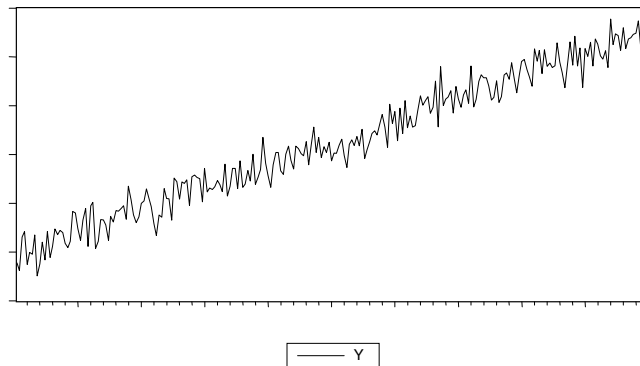


Figura 16.1
($y_t = 0.37 * t + \varepsilon_t$)

16.3.1 Removiendo las tendencias

Existen dos formas posibles de remover una tendencia determinística: diferenciando la serie o eliminando la tendencia. Eliminar la tendencia implica regresionar la serie contra una función que involucre el tiempo. Esta función puede ser simplemente una tendencia lineal o tomar otras formas como las que se presentaron en la sección anterior.

A continuación se explican los procedimientos necesarios para remover una tendencia de una serie y el por qué de su efectividad. Asimismo, al final de la presente sección se comparan ambos procedimientos para determinar sus posibles ventajas y desventajas.

Diferenciando

Considerando la presencia de una tendencia lineal, donde se asume que el componente de error es un ruido blanco:

$$y_t = \alpha + \beta_0 t + \varepsilon_t$$

Rezagando un periodo y restando del modelo inicial se obtiene:

$$y_t - y_{t-1} = \beta_0 + \varepsilon_t - \varepsilon_{t-1}$$

Como resultado del procedimiento se obtiene una ecuación diferenciada. Por consiguiente esta ecuación se compone de una constante más la resta de dos procesos estacionarios. La nueva serie diferenciada es un proceso estacionario. Para comprobarlo hallaremos su media, varianza y autocovarianza:

$$\begin{aligned} E(\Delta y_t) &= E(\beta_0 + \varepsilon_t - \varepsilon_{t-1}) = \beta_0 \\ \text{Var}(\Delta y_t) &= E(\Delta y_t - \beta_0)^2 = E(\varepsilon_t - \varepsilon_{t-1})^2 = 2\sigma_\varepsilon^2 \\ \text{Cov}(\Delta y_t, \Delta y_{t-s}) &= E(\Delta y_t - \beta_0)(\Delta y_{t-s} - \beta_0) = E(\varepsilon_t - \varepsilon_{t-1})(\varepsilon_{t-s} - \varepsilon_{t-s-1}) = 0 \\ &\forall s \neq 0 \end{aligned}$$

Como la media y varianza son constantes y la autocovarianza solo depende de s , podemos concluir que es un proceso estacionario.

Hay que recalcar que este procedimiento para remover la tendencia de la serie, la vuelve estacionaria pero imposibilita que pueda ser invertible. El proceso de medias móviles de la serie diferenciada tiene una raíz unitaria, lo cual imposibilita que pueda ser representado por un proceso autorregresivo⁵.

Generalizando, una tendencia polinomial de orden “ d ” debe ser diferenciada “ d ” veces para limpiar a la serie del polinomio de tendencia que presente.

Eliminando la tendencia

Como se vio en la sección anterior, el mecanismo de la diferenciación es efectivo para remover las tendencias de la serie, pero en algunos casos imposibilita que la serie pueda invertirse a un proceso autorregresivo. Como un procedimiento alternativo, se puede eliminar

⁵ La raíz unitaria es el coeficiente unitario del primer rezago del error. En el caso que el lector no recuerde la propiedad de invertibilidad de los procesos MA, es recomendable que revise el capítulo anterior.

la tendencia lineal del modelo realizando la siguiente regresión $y_t = \alpha + \beta_0 t + \varepsilon_t$. Luego se halla el valor estimado de y_t , el cual se resta del valor verdadero y se extrae el residuo $\{\varepsilon_t\}$ que sería la serie estacionaria.

De manera más general, la serie puede presentar una tendencia polinomial:

$$y_t = \alpha + \beta_0 t + \beta_1 t^2 + \dots + \beta_n t^n + \varepsilon_t$$

donde $\{\varepsilon_t\}$ es el proceso estacionario.

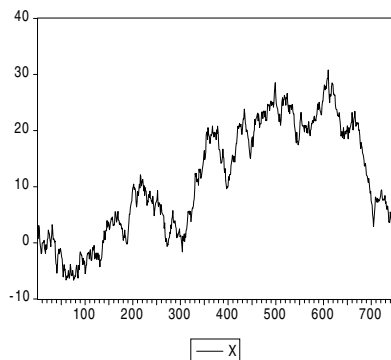
Eliminar la tendencia implicaría regresionar la serie contra la tendencia polinomial y extraer el residuo estacionario. El grado del polinomio se puede determinar mediante tests F o T y/o usando criterios estadísticos como el de Akaike o Schwartz. Comúnmente se regresiona usando el máximo grado posible del polinomio y se eliminan los grupos de coeficientes no significativos; luego se procede a utilizar los criterios de AIC y SBC para confirmar el grado. Una ayuda adicional la proporciona el análisis gráfico para determinar el grado del polinomio. Una vez que las series sean estacionarias, la serie puede ser modelada utilizando los métodos tradicionales.

16.4. PROCESOS ESTACIONARIOS EN DIFERENCIAS (DS)

Tanto los procesos estacionarios como los procesos de tendencia determinística tienen un valor de largo plazo determinado por su media o por su tendencia. Esto implica que desviaciones de su valor de largo plazo serán temporales, en otras palabras, shocks que afecten a estas series tendrán un efecto transitorio. Esta característica se conoce como reversión a la media.

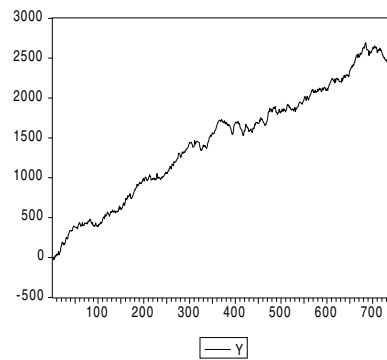
Pero existen series que no cumplen con esta importante característica y que en cambio presentan una tendencia estocástica. Este tipo de tendencia no presenta una tasa de crecimiento constante como es el caso de la tendencia lineal determinística, al contrario, su tasa de crecimiento va cambiando en el tiempo de manera aleatoria. Es por ello que una serie estacionaria no puede presentar una tendencia estocástica. Las series que presentan este tipo de tendencia tienen el problema de raíz unitaria, es decir, el polinomio característico presenta una o más raíces unitarias.

Figura 16.2



$$X_t = X_{t-1} + \varepsilon_t$$

(Camino aleatorio)



$$y_t = 3 + y_{t-1} + \varepsilon_t$$

(Camino aleatorio con constante)

Un proceso de raíces unitarias se puede presentar bajo la siguiente formulación:

$$y_t = y_{t-1} + \varepsilon_t \quad (16.6)$$

o lo que es igual

$$(I - L)y_t = \varepsilon_t \quad (16.7)$$

$$A(L)y_t = \varepsilon_t \quad (16.8)$$

donde L es el operador de rezagos, A(L) es el polinomio característico y ε_t es un error ruido blanco.

Un modelo presenta una raíz unitaria, si al igualar su polinomio característico a cero, $A(L)=0$, el valor de alguna de las raíces que cumple con esta ecuación es uno. Los modelos ARMA del capítulo anterior presentan un polinomio característico en el que ninguna de sus raíces es unitaria; sino que tienen un valor superior a uno⁶. Aplicándolo a la formulación anterior:

$$(1 - \beta L) = 0$$

$$L = \frac{1}{\beta}$$

Si el valor del coeficiente es menor a uno, entonces el valor de L es superior a uno y el modelo es estacionario. Pero como en el modelo (16.7) el valor de β es unitario, L es una raíz unitaria y el modelo no es estacionario.

Tal como se ha visto en el capítulo anterior, esta formulación particular de raíz unitaria, se conoce en econometría como *camino aleatorio* (Gráfico 17.2). Como se notará, este modelo no es más que una regresión del tipo AR(1), en la cual el coeficiente asociado al primer rezago de la variable y_t (β) es unitario y el término constante (α) es cero.

Como se puede apreciar en la ecuación (16.7), para un proceso de una raíz unitaria, los cambios en la serie tienen una representación estacionaria. Es por ello que a una serie que tiene raíces unitarias también se le conoce como estacionario en diferencias (DS)⁷.

A continuación, vamos a analizar las características de la media, varianza y covarianza del camino aleatorio.

Dado un valor inicial igual a y_0 , la solución general de una ecuación diferencial de primer orden representada por el modelo de camino aleatorio es:

$$y_t = y_0 + \sum_{i=1}^t \varepsilon_i$$

Si tomamos el valor esperado de esta expresión obtenemos $E(y_t) = E(y_{t-s}) = y_0$, de manera que la media de un camino aleatorio es constante. A pesar de esto, todos los shocks aleatorios pasados tienen efectos permanentes sobre la serie. Esto se observa en que el valor de la serie en el periodo t es explicado por la sumatoria de los errores. Si obtenemos la media condicional de

⁶ O su equivalente que todas las raíces características de la ecuación son menores a 1

⁷ Esto no implica que todas las series que no siendo estacionarias sigan un proceso estacionario luego de diferenciarlas, tengan procesos de raíces unitarias. Este es el caso de una serie con tendencia lineal.

y_{t+1} , dada la información hasta el periodo t de ε_t obtenemos (lo que es igual a obtener la media condicional):

$$E_t y_{t+1} = E_t (y_t + \varepsilon_{t+1}) = y_t$$

De lo cual concluimos que la mejor predicción de la serie en el periodo $t+1$ es el valor que tomó en el periodo t .

De igual forma, la media condicional de y_{t+s} (para cualquier $s>0$) se puede obtener de la siguiente forma:

$$y_{t+s} = y_t + \sum_{i=1}^s \varepsilon_{t+i}$$

$$E_t y_{t+s} = y_t + E_t \sum_{i=1}^s \varepsilon_{t+i} = y_t$$

De manera que la media condicional para el periodo $t+s$, dado cualquier valor de $s>0$, es igual al valor de la serie en el periodo t , el último en el que se tiene información.

Se debe notar, que a pesar que la media condicional no varía, la varianza si cambia con el tiempo.

$$\text{Var}(y_t) = \text{var}(\varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_1) = t\sigma^2$$

$$\text{Var}(y_{t-s}) = \text{var}(\varepsilon_{t-s} + \varepsilon_{t-s-1} + \dots + \varepsilon_1) = (t-s)\sigma^2$$

Dado que la varianza no es constante en el tiempo [$\text{var}(y_t) \neq \text{var}(y_{t-s})$], el camino aleatorio es un proceso no estacionario. Cuanto más se acerque t al infinito, la varianza de la serie también se acercará al infinito.

Ahora procedemos a calcular la covarianza entre y_t y y_{t-s} . Dado que la media es constante, se puede desarrollar la covarianza de la siguiente forma:

$$\begin{aligned} E[(y_t - y_0)(y_{t-s} - y_0)] &= E[(\varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_1)(\varepsilon_{t-s} + \varepsilon_{t-s-1} + \dots + \varepsilon_1)] \\ &= E[(\varepsilon_{t-s})^2 + (\varepsilon_{t-s-1})^2 + \dots + (\varepsilon_1)^2] \\ &= (t-s)\sigma^2 \end{aligned}$$

Como se observa el valor de la covarianza varía con el tiempo; este es otro motivo para que la serie no sea estacionaria. Será útil hallar la función de autocorrelación simple (FAS) del modelo; este se obtiene dividiendo la covarianza entre las desviaciones estándar de y_t y de y_{t-s} .

$$\begin{aligned} \rho_s &= (t-s) / \sqrt{(t-s)t} \\ &= [(t-s)/t]^{0.5} \end{aligned}$$

Este resultado juega un importante rol en la detección de series no estacionarias. Para las primeras correlaciones el ratio $(t-s)/t$ tendrá un valor cercano a uno, pero conforme aumenta el valor de s , el valor de la FAS irá disminuyendo.

El problema de utilizar este método para detectar la no estacionariedad en las series es que no se puede distinguir entre un proceso de raíz unitaria en el que $\beta = 1$ o de un proceso en el cual el valor de β es sólo cercano a uno. En general este será un problema que acompañará a la mayoría de tests para la detección de raíces unitarias.

Otro ejemplo de modelo que presenta una raíz unitaria es el camino aleatorio con constante (*random walk with drift*):

$$y_t = \alpha + y_{t-1} + \varepsilon_t.$$

(Representado en el Gráfico 16.2) Las características de esta serie serán analizadas en la siguiente sección, en la cual se compararán los procesos TS con los procesos DS.

16.5. COMPARACIONES ENTRE PROCESOS ESTACIONARIOS EN TENDENCIAS Y PROCESOS DE RAÍCES UNITARIAS

En la presente sección vamos a comparar las principales características e implicancias de la formulación de una serie como un proceso estacionario en tendencias (TS) o como un proceso estacionario en diferencias (DS) o raíz unitaria.

1. Acerca de las predicciones de cada modelo

Si tenemos un proceso TS como el siguiente:

$$y_t = \alpha + \beta t + \gamma(L)\varepsilon_t \quad (16.9)$$

Esta formulación tiene como componente determinístico a la tendencia y como componente estacionario al polinomio de rezagos del error ruido blanco $\gamma(L)\varepsilon_t$. Como se analizó en el capítulo anterior⁸, la predicción de un proceso estacionario tiende a hacia su media cuanto más alejado está el periodo $(t+s)$ del último periodo con información disponible (t) .

Debido a esto, la predicción del proceso TS tiende a converger a su tendencia. Como el proceso estacionario converge a su media, la cual es cero por ser un error ruido blanco, la predicción de la serie converge a su componente determinístico $\alpha + \beta t$.

Para el análisis de una variable con raíces unitarias, se empleará un proceso camino aleatorio con constante:

$$y_t = \alpha + y_{t-1} + \varepsilon_t \quad (16.10)$$

Si se busca realizar una predicción con este modelo, lo primero que debemos notar es que el valor de y_t estará dado por la suma acumulada de los errores pasados (basta reemplazar sucesivamente en el proceso y_{t-1} por $\alpha + y_{t-2} + \varepsilon_{t-1}$, luego y_{t-2} , y así sucesivamente). Suponiendo que el valor inicial de la serie es y_0 , la solución general para la serie y_t es:

$$y_t = y_0 + \alpha t + \sum_{i=1}^t \varepsilon_i$$

Ahora, si queremos hallar el valor de la serie en el periodo $t + s$:

⁸ En el capítulo anterior se analizó la predicción de los modelos ARMA así como la notación asociada al polinomio de rezagos $\gamma(L)\varepsilon_t$ y operador de rezagos.

$$\begin{aligned}
 y_{t+s} &= y_0 + \alpha(t+s) + \sum_{i=1}^{t+s} \varepsilon_i \\
 &= y_t + \alpha s + \sum_{i=1}^s \varepsilon_{t+i}
 \end{aligned}$$

Luego tomamos la esperanza condicional de y_{t+s} , ya que suponemos que sólo hay información hasta el periodo t .

$$E_t y_{t+s} = y_t + \alpha s$$

Como se puede observar, para predecir el valor de la serie en el periodo $t + s$, se necesita agregar a y_t s veces la constante y también los errores futuros. Pero como desde el periodo $t + 1$ hasta $t + s$ la esperanza de los errores es cero, ya que no se tiene información acerca de ellos en el periodo t , tendremos que la variable y_t crecerá a partir de su valor en el periodo t a una tasa α hasta el periodo $t + s$.

Puede verse que en ambas especificaciones, la variable y_t crece a una tasa constante. Sin embargo existe una diferencia muy clara entre ambas especificaciones, que tiene que ver con el punto a partir del cual crecen o hacia donde convergen. En el caso del proceso TS, la serie converge a la línea de tendencia $\alpha + \beta t$ independientemente del valor que tome y_t . Por ello se dice que para un proceso TS las condiciones iniciales no importan, siempre se regresará a la tendencia determinística. En el caso de un proceso con raíz unitaria, lo anterior no se cumple.

La variable crece a la tasa constante pero a partir del punto y_t . Además, no existe convergencia a una línea de tendencia. En este caso el valor inicial sí importa para realizar la predicción.

2. Acerca de los errores de predicción y sus varianzas

En el caso de un proceso TS, si se considera el polinomio $\psi(L)$, se verá que una característica de un proceso estacionario es que si este polinomio es de orden infinito (como sería el caso de un proceso AR(1)), los parámetros que conforman dicho polinomio cumplen con la siguiente condición

$$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

lo que implica que la varianza de la predicción tienen la característica que tiende a estabilizarse conforme el periodo de predicción s , tiende a infinito. Aquí debe recordarse que a pesar de que los errores ruidos blancos tienen como predicción su esperanza, la cual es cero, presentan una varianza constante y finita, por lo que la varianza del error de predicción será la multiplicación de σ^2 por la expresión anterior, que tiende a estabilizarse. Por ello, la varianza del error irá aumentando conforme pronostiquemos periodos más distantes; pero a partir de cierto punto esta varianza tenderá a estabilizarse.

En el caso de un proceso de raíz unitaria, la varianza de la predicción irá en aumento conforme se estiman periodos más distantes. La predicción de la varianza de la serie en el periodo $t + s$, siguiendo con la formulación (16.10); adoptará la siguiente representación:

$$Var(y_{t+s}/t) = var(\varepsilon_{t+s} + \varepsilon_{t+s-1} + \dots + \varepsilon_{t+1}) = s\sigma^2$$

La varianza de la serie en el periodo $t + s$ va ser igual a la suma de las varianzas de los errores ruido blanco desde el periodo $t + 1$ hasta el $t + s$, por lo tanto, conforme s tienda a

infinito, la varianza del error de predicción tenderá a infinito. Es más, en este último caso la varianza sería infinito. Por ello se suele decir que la varianza del error de predicción en un proceso DS no está acotada. En conclusión, la confiabilidad del error de pronóstico de un proceso DS irá descendiendo conforme se amplíe el horizonte de predicción.

3. Acerca del efecto de un shock temporal

En el caso de una variable TS, se tiene que el efecto de un shock temporal se desvanece en el tiempo y la variable tiende a regresar a su tendencia determinística. Esto puede quedar claro si se especifica el siguiente modelo:

$$y_t = \alpha + \delta t + \rho y_{t-1} + \varepsilon_t$$

con $\rho < 1$. En este caso, si el ruido blanco toma un valor distinto de su media en el periodo t ($\varepsilon_t \neq 0$) se obtiene para el periodo $t + 1$ lo siguiente:

$$y_{t+1} = \alpha + \delta t + \rho y_t$$

Como se aprecia, el valor de y_t (que se supone se salió de la tendencia) es multiplicado por ρ que es menor a uno y por lo tanto el efecto del shock ocurrido en t será menor en $t + 1$. Bajo este razonamiento para periodos posteriores, se deduce que en el periodo $t + s$ el valor de y_t estará multiplicado por ρ^s , el cual tenderá a ser muy pequeño. Para un s grande, el efecto del shock habrá desaparecido.

En verdad, el efecto del shock depende del valor del parámetro ρ . Si este es muy cercano a uno, entonces el efecto tenderá a persistir por buen tiempo; sin embargo, inexorablemente, éste se desvanecerá.

En el caso de un proceso con raíz unitaria (DS), siguiendo con la formulación de la ecuación (16.10), el parámetro que multiplica a y_{t-1} es precisamente uno. Si tomamos esto en cuenta y repetimos el ejercicio anterior (suponer que errores de periodos posteriores al t toman su valor medio), veremos que el efecto del shock en el periodo no se desvanece, sino que se mantiene perpetuamente. La única forma de que el shock desapareciera es mediante la presencia de otro shock de igual magnitud y de signo contrario.

Lo anterior sirve para entender el por qué los procesos de raíces unitarias presentan, durante varios periodos, se encuentra observaciones por encima de su media o tendencia; mientras que durante otros periodos, por debajo de la misma. Lo que ocurre es que si se produce un shock positivo que aleja la variable de su tendencia o media, esta desviación se mantendrá por mucho tiempo hasta que ocurra otro shock negativo que lleve a la serie por debajo de su media u otro shock positivo que lo aleje más de ella. En resumen, no existe nada que garantice que la variable retornará a su media o tendencia. Por ello el nombre de camino aleatorio, dado que no se puede predecir con exactitud la trayectoria de la serie. Además, anteriormente se vio que la varianza del error va creciendo conforme aumentaba el periodo de predicción. Este punto está muy relacionado con lo que se menciona aquí. Como no se sabe donde estará la serie, la varianza aumenta indicando que si se producen shocks en el futuro, la variable podría tomar cualquier valor.

4. Acerca de la forma para alcanzar la estacionariedad de la serie

A una serie que tiene efectivamente un proceso TS se le remueve la tendencia mediante una regresión (aunque hemos visto que diferenciarla también la vuelve estacionaria con algunos inconvenientes). Los errores de dicha regresión es la serie estacionaria lista para modelar.

En el caso de procesos con raíces unitarias o integrados de orden d , la mecánica a seguir consiste en diferenciar la serie d veces y esta nueva serie será estacionaria y lista para aplicarle Box y Jenkins.

Comparaciones entre procesos estacionarios en tendencias y procesos de raíces unitarias

	Estacionario en tendencias (TS)	Raíz Unitaria (DS) Camino aleatorio con constante
Pronóstico	Tiende a converger a la tendencia determinística y crece a una tasa constante dada por el coeficiente de la tendencia	No converge a ninguna tendencia, el pronóstico crece a una tasa dada por la constante y crece a partir del último valor observado a partir del cual se quiere pronosticar.
Varianza del pronóstico	Tiende a estabilizarse	Crece sin límite
Efectos de shocks transitorios	Los efectos tienden a desvanecerse más o menos rápido en función del parámetro ρ	El efecto es permanente
Formas de alcanzar la estacionariedad	Remover la tendencia	Diferenciar la serie hasta hacerla estacionaria

16.6. REGRESIONES ESPÚREAS

Granger y Newbold (1974) introdujeron este término para describir el hecho de que si tenemos dos variables que son independientes entre sí, y por lo tanto con una correlación igual a cero; y se estima un modelo haciendo depender una de otra, los resultados pueden mostrar que el modelo es estadísticamente significativo. Este caso se presenta cuando se toman variables no estacionarias.

Una regresión espúrea tiene un alto R^2 , el estadístico T aparece significativo, pero los resultados de la regresión no representan la verdadera relación entre las variables, la cual es nula.

Granger y Newbold (1974) consideraron el siguiente proceso generador de datos (PGD) para examinar las consecuencias de las regresiones espúreas :

$$y_t = y_{t-1} + \varepsilon_{yt} \quad (16.11)$$

$$x_t = x_{t-1} + \varepsilon_{xt} \quad (16.12)$$

donde ε_{yt} y ε_{xt} son errores ruido blanco independientes y por lo tanto y_t y x_t son caminos aleatorios no correlacionados. Como nada de y_t afecta a x_t , si corremos la siguiente regresión :

$$y_t = \alpha + \beta x_t + e_t \quad (16.13)$$

Se esperaba que el coeficiente β converjan en probabilidad a cero reflejando la ausencia de correlación entre ellos y también esperaríamos que el R^2 sea cercano a cero. Sin embargo, esto no ocurre necesariamente. En su análisis de Montecarlo, Granger y Newbold generaron muchas series con los modelos (16.11) y (16.12) y con ellas estimaron regresiones siguiendo el modelo (16.13). Ellos encontraron que al 5% de significancia, rechazaron la hipótesis nula ($\beta = 0$) en aproximadamente el 75% de las veces que realizaron el experimento.

Como una explicación intuitiva de los resultados de Montecarlo se puede observar que en la regresión (16.13) tanto, la hipótesis nula $\beta = 0$ como la hipótesis alternativa $\beta \neq 0$ conllevan a modelos falsos. Desde este punto de vista no debe sorprender que la hipótesis nula sea rechazada y que la autocorrelación del camino aleatorio y_t tienda a proyectarse sobre x_t , otro camino aleatorio fuertemente autocorrelacionado.

Otros estudios indican que las estimaciones de mínimos cuadrados ordinarios del modelo planteado llevan a resultados que no pueden ser interpretados dentro de los procedimientos tradicionales de inferencia estadística. Para empezar, los estadísticos t calculados para α y β no tienen distribuciones t , no presentan distribución alguna y de hecho, divergen en distribución conforme el número de observaciones aumenta. De esta manera, para un valor crítico fijo, las tasas de rechazo de la nula tienden a incrementarse con el tamaño de la muestra.

Además, los supuestos acerca de los errores de la regresión son transgredidos. Para examinar las propiedades de e_t eliminaremos, por simplicidad, la constante del modelo (16.13) y lo reescribiremos como :

$$e_t = y_t - \beta x_t$$

Si reemplazamos y_t y x_t por sus soluciones generales, bajo el supuesto que sus condiciones iniciales son $y_0 = x_0 = 0$, tenemos :

$$e_t = \sum_{i=1}^t \varepsilon_{y_i} - \beta \sum_{i=1}^t \varepsilon_{x_i}$$

Podemos ver que la varianza de e_t se vuelve infinita conforme t aumenta. Más aún, el residuo estimado tiene un alto grado de autocorrelación. Dado que los supuestos de los usuales tests de hipótesis han sido violados, las pruebas t y F o el valor del R^2 no son confiables.

Si bien el ejemplo se ha desarrollado para variables que son todas $I(1)$, los problemas son similares cuando tenemos otros ordenes de integración. Así, en una regresión de una variable $I(2)$ contra una $I(1)$, el problema suele agravarse (se rechazan con más probabilidad las H_0 cuando en verdad deberían aceptarse), mientras que el caso de una regresión de una $I(1)$ contra una $I(0)$, el problema suele reducirse. En todo caso, conforme aumenta el orden de integración lo problemas tienden a agravarse.

Para dar una idea de lo anterior, Banerjee et. el. Presentan un cuadro del cual extraemos algunos datos. Estos datos fueron calculados para una simulación de Montecarlo que se hicieron para 100 datos. Debe tenerse en cuenta que el primer orden de integración corresponde a la variable dependiente del modelo estimado.

Tipo	$[d_1 + d_2]$	$\Pr(k(\beta = 0) \geq 2)$
I(0), I(1)	0	0.0493
I(1), I(1)	2	0.7570
I(2), I(2)	4	0.9406
I(0), I(1)	1	0.0458
I(1), I(0)	1	0.0486
I(2), I(1)	3	0.8530
I(1), I(2)	3	0.8444

La última columna es la más importante para nuestro análisis. Ahí se aprecia que la probabilidad de que los tests calculados alcancen un valor mayor a dos en valor absoluto y, por consiguiente, se rechace la hipótesis nula, aumenta conforme aumenta el orden de integración.

Otro caso en el cual el modelo puede “verse bien”, aunque no exista una relación verdadera entre las variables, sucede cuando se relaciona las tendencias de las variables. Dos variables no estacionarias pueden ser totalmente independientes entre sí, pero una regresión de uno con respecto del otro puede presentar a los regresores como significativos y el R^2 alto, debido a que las series provienen de un proceso TS y la regresión recoge la relación directa o indirecta de sus tendencias. Para evitar esto, se debe remover la tendencia de la serie antes de realizar la regresión.

Granger y Newbold sugirieron que una forma de detectar el problema de regresiones espúreas era comparando el valor del estadístico Durbin-Watson y el R^2 . Si el primero era menor al segundo era evidencia del problema. Estudios recientes muestran que cuando se realizan regresiones espúreas, el DW tiende a cero conforme el tamaño de la muestra crece. Eso evidencia que existe un error que no es estacionario y que presenta una fuerte autocorrelación.

En resumen, la discusión anterior nos muestra que el problema de inferencia es serio en regresiones que presentan variables integradas. Los tests no pueden interpretarse de manera habitual.

16.7. PRUEBAS DE DETECCIÓN DE RAÍCES UNITARIAS

Saber si la no estacionariedad en los datos se debe a una tendencia determinística o a una raíz unitaria es un asunto muy importante. En este sentido, muchos investigadores estarán interesados en saber si un shock (positivo o negativo) sobre el producto presentará un efecto permanente o transitorio sobre dicha variable.

Aunque es muy importante saber si una serie de tiempo contiene una raíz unitaria, muchos artículos recientes arguyen que la cuestión no tiene respuesta aparente sobre la base de observaciones de muestras finitas. El argumento que sirve de base para lo anterior es que para todo proceso de raíz unitaria existe un proceso estacionario que será imposible de distinguir de la representación de raíces unitarias para cualquier realización muestral de tamaño T . Tal proceso estacionario es hallado de manera suficientemente fácil si es que se fija uno de los valores característicos muy cercano a uno. Por ejemplo supongamos que la muestra incluya $T = 10,000$ observaciones que fueron generadas por un random walk puro:

$$y_t = y_{t-1} + \varepsilon_t \quad \text{modelo verdadero (raíz unitaria)}$$

consideremos ahora tratar de diferenciarlo del siguiente proceso estacionario:

$$y_t = \phi y_{t-1} + \varepsilon_t \quad |\phi| < 1 \quad \text{modelo falso (estacionario)}$$

el pronóstico de s períodos hacia adelante del modelo verdadero es:

$$\hat{y}_{t+s/t} = y_t$$

con un error cuadrático medio:

$$E(y_{t+s} - \hat{y}_{t+s/t})^2 = s\sigma^2$$

El pronóstico correspondiente para el modelo falso es:

$$\hat{y}_{t+s/t} = \phi^s y_t$$

Con un error cuadrático medio

$$E(y_{t+s} - \hat{y}_{t+s|t})^2 = (1 + \phi^2 + \phi^4 + \dots + \phi^{2(s-1)})\sigma^2$$

Claramente se observa que existe un valor de ϕ lo suficientemente cercano a 1, tal que la representación gráfica de la serie estacionaria es muy similar a la del proceso verdadero (no estacionario).

La proposición inversa también es verdadera. Para cualquier proceso estacionario y un tamaño muestral dado T , existe un proceso con raíz unitaria que será imposible de distinguir de la representación estacionaria.

Los procesos con raíces unitarias y estacionarios difieren en sus implicaciones en horizontes temporales infinitos, pero para cualquier número dado de observaciones finitas, existe un representante de cada uno de los procesos que puede dar cuenta de todas las características observadas de los datos. Por lo tanto, necesitamos ser cuidadosos con la elección de alguna prueba de raíz unitaria o bien en la forma de testear si las innovaciones tienen un efecto permanente sobre el nivel de la serie. Aunque esto es interesante, en la práctica es imposible de hacer.

No obstante, existen cuestiones que sí tienen respuesta. Por ejemplo, si se tienen suficientes datos, podríamos preguntarnos si las innovaciones tienen un efecto significativo sobre el nivel de la serie dentro de un horizonte finito especificado. Para un período de 3 años, existe un tamaño muestral (digamos 50 observaciones) tal que podemos preguntar propiamente si $\delta y_{t+s} / \delta \varepsilon_t$ es cercano a cero. No podemos decir si los datos fueron generados por un proceso con raíz unitaria o por un proceso estacionario, pero podemos medir si las innovaciones tienen mucha o poca persistencia sobre un intervalo fijo de tiempo.

También podemos llegar a hipótesis verificables si estamos dispuestos a restringir aún más las clases de procesos considerados. Supongamos que la dinámica de una muestra dada $\{y_1, \dots, y_T\}$ es modelada usando una autorregresión de orden p . Supongamos, por ejemplo, que estamos considerando usar un proceso AR(1):

$$y_t = \phi y_{t-1} + \varepsilon_t$$

dentro de estos modelos, la restricción $H_0: \phi = 0$, es verificable. Si bien es cierto que existen alternativas locales (tal como $\phi = 0.9999$) contra las cuales un test no tendrá mucho poder, esto es cierto para muchos tests de hipótesis. También existen alternativas (tal como $\phi = 0.3$) que llevarían a un rechazo de la H_0 dado un número suficiente de observaciones. La hipótesis “ $\{y_t\}$ es un proceso AR(1) con una raíz unitaria” es potencialmente refutable, mientras que la hipótesis “ $\{y_t\}$ es un proceso general de orden ARMA desconocido o infinito con una raíz unitaria” no lo es.

Existen buenas razones para restringirnos a considerar sólo representaciones autorregresivas de orden pequeño. Los modelos parsimoniosos frecuentemente se comportan mejor y las autorregresiones son mucho más fáciles de estimar y pronosticar que procesos de medias móviles, particularmente procesos MA con una raíz cercana a 1.

Si estamos dispuestos a describir los datos con autorregresiones de bajo orden, saber si la restricción adicional de una raíz unitaria debe ser impuesta puede ser importante por dos razones. La primera razón tiene que ver con el trade-off usual entre eficiencia y consistencia. Si una restricción (en este caso una raíz unitaria) es verdadera, se obtendrán estimadores más eficientes al imponerla. Los estimados de los otros coeficientes y los multiplicadores dinámicos

serán más precisos y las predicciones serán mejores. Si la restricción es falsa, los estimadores no serán adecuados no importando el tamaño de la muestra. Los investigadores difieren en su consejo de cómo comportarse frente a este trade-off. Una guía práctica es estimar el modelo tanto con la restricción como sin la restricción. Si las principales inferencias relacionadas al modelo son similares, será mucho mejor porque podremos comparar. Si las principales inferencias son distintas se deberá intentar una explicación para ello.

Adicionalmente al trade-off entre eficiencia y consistencia, la decisión de imponer o no una raíz unitaria sobre una autorregresión también implica asuntos relacionados con la teoría de distribución asintótica que se utiliza para verificar la hipótesis del proceso. En general, las distribuciones para verificar la hipótesis no son las tradicionales (F y t) sino que presentan ciertas características que están relacionadas con lo que se mencionó para el caso de regresiones espúreas. Este tipo de hipótesis llevó a plantear otro tipo de distribuciones que son las que se utilizan en la práctica. Intuitivamente, para solucionar el problema de regresiones espúreas al momento de estimar una autorregresión, y disminuir la posibilidad de rechazar la H_0 cuando es verdadera, se introducen ciertos cambios a las distribuciones tradicionales. Este es el espíritu que guía a procedimientos como el Dickey-Fuller para el testeo de hipótesis relacionadas con raíces unitarias y que se usa mucho en la práctica.

16.8. TESTS DE RAÍCES UNITARIAS

Los tests para comprobar la presencia de raíces unitarias parten de la hipótesis nula de un modelo AR(p) que contiene una raíz característica igual a 1. En todo caso lo que se verifica es que el parámetro ϕ de una autorregresión del tipo:

$$y_t = \phi y_{t-1} + \varepsilon_t$$

es igual a 1. Sin embargo, muchos de estos modelos se reparametrizan de la siguiente manera:

$$\Delta y_t = \phi y_t + \varepsilon_t$$

De esta forma, la hipótesis a verificar en este modelo es que el coeficiente de y_{t-1} es cero, por lo que pueden aplicarse los tests del tipo t. Sin embargo, como ya hemos mencionado, las distribuciones que se utilizan para estas pruebas serán diferentes de los tradicionales. Para ilustrar lo anterior, mencionaremos que la distribución asintótica del parámetro ϕ , en el modelo anterior cuando éste es menor a 1, es de la siguiente forma:

$$\sqrt{T}(\hat{\phi}_T - \phi) \xrightarrow{D} N(0, (1 - \phi^2))$$

Si se parte de la hipótesis nula que $\phi = 1$, entonces la anterior distribución colapsa a una distribución degenerada, ya que:

$$\sqrt{T}(\hat{\phi}_T - \phi) \xrightarrow{P} 0$$

como es obvio, esta expresión no es adecuada para realizar inferencia sobre los parámetros. Por ello, la presencia de raíces unitarias implica el uso de otro tipo de teoría asintótica, que implica otros procesos estocásticos como el movimiento Browniano o Proceso de Wiener, cuyo estudio escapa del alcance del presente capítulo.⁹

⁹ Para una revisión de la intuición de esta teoría se puede consultar Hamilton (1994) y Benerjee, et. al (1993)

16.8.1 El test de CRDW

Este test se basa en la fuerte correlación que presentan series con raíces unitarias. Para esto emplea el test Durbin Watson (DW) usado en los test de autocorrelación de los errores:

$$DW = 2(1 - \rho)$$

donde ρ es el parámetro del primer rezago de los errores. La hipótesis nula a comprobar es que el parámetro ρ es igual a 1, por lo que el estadístico DW tenderá a cero. Por ello si es que se observa que el estadístico DW es significativamente distinto de cero, entonces no se podrá aceptar la hipótesis nula, por lo que la serie presenta raíz unitaria.

Engle y Granger presentan una tabla con los valores críticos para este test que aquí reproducimos:

Nivel de Significación	0.01	0.05	0.10
Valor Crítico	0.511	0.386	0.322

Entonces si el valor calculado del test es menor que los valores críticos mencionados, tendremos que se acepta la H_0 . Si el valor calculado es mayor se tenderá a no aceptar la H_0 . Un problema que presenta este test es la existencia de una zona de indeterminación que varía según el modelo que se analiza, Aunque se necesita alterarse ante la inclusión de una constante o tendencia dentro del modelo a estimar.

En la práctica, este test se realiza estimando una regresión en cualquier paquete econométrico donde la variable dependiente será la serie en cuestión y el regresor será una constante. Al momento de observar el output de esta regresión, se muestra siempre el estadístico DW calculado para la regresión. Este se usará para compararlo con los valores críticos.

16.8.2 Tests de Dickey-Fuller

Dickey y Fuller, en una serie de artículos publicados entre 1979 y 1981, establecieron los tests más populares para verificar la presencia de raíz unitaria. Estos Tests están diseñados para incluir una constante y una tendencia dentro del modelo a estimar. Los tests que desarrollaron fueron de tres tipos: los t, los F y otro que se especifica de la siguiente manera $T(\hat{\alpha} - \alpha)$. En los dos primeros las distribuciones originales son corregidas para recoger el efecto de no estacionariedad. El último de los tests es explicado a profundidad en el apéndice.

Los tests que se plantean son de una sola cola, dado que la posibilidad que la raíz sea mayor a uno no interesa ya que la serie mantendría un comportamiento explosivo y bajo la nula los regresores no serían todos estacionarios.

Existen muchas versiones de los tests de este tipo. En el presente trabajo nos remitiremos a la sistematización que hace Hamilton¹⁰. Este autor presenta el siguiente cuadro resumen:

¹⁰ Hamilton, James, "Time Series Analysis", Princeton University Press, U.S.A., 1994.

Resumen de los Tests de Dickey Fuller en Ausencia de Correlación Serial
<p>Caso 1: Regresión estimada : $y_t = \rho y_{t-1} + \mu_t$</p> <p>Proceso Verdadero: $y_t = y_{t-1} + \mu_t$ $\mu_t \approx \text{i.i.d.N}(0, \sigma^2)$</p> <p>$T(\hat{\rho}_T - 1)$ tiene la distribución descrita bajo el caso 1 en la Tabla A.1</p> <p>$(\hat{\rho}_T - 1) / \hat{\sigma}_{\rho_T}$ tiene la distribución descrita bajo el caso 1 en la Tabla A.2</p>
<p>Caso 2: Regresión estimada : $y_t = \alpha + \rho y_{t-1} + \mu_t$</p> <p>Proceso Verdadero: $y_t = y_{t-1} + \mu_t$ $\mu_t \approx \text{i.i.d.N}(0, \sigma^2)$</p> <p>$T(\hat{\rho}_T - 1)$ tiene la distribución descrita bajo el caso 2 en la Tabla A.1</p> <p>$(\hat{\rho}_T - 1) / \hat{\sigma}_{\rho_T}$ tiene la distribución descrita bajo el caso 2 en la Tabla A.2</p> <p>Test F de MCO de la hipótesis conjunta que $\alpha = 0$ y $\rho = 1$ tiene la distribución descrita bajo el caso 2 de la Tabla A.3</p>
<p>Caso 3: Regresión estimada : $y_t = \alpha + \rho y_{t-1} + \mu_t$</p> <p>Proceso Verdadero: $y_t = \alpha + y_{t-1} + \mu_t$ $\alpha \neq 0$, $\mu_t \approx \text{i.i.d.N}(0, \sigma^2)$</p> <p>$(\hat{\rho}_T - 1) / \hat{\sigma}_{\rho_T} \xrightarrow{D} N(0,1)$</p>
<p>Caso 4: Regresión estimada : $y_t = \alpha + \rho y_{t-1} + \delta t + \mu_t$</p> <p>Proceso Verdadero: $y_t = \alpha + y_{t-1} + \mu_t$ cualquier α $\mu_t \approx \text{i.i.d.N}(0, \sigma^2)$</p> <p>$T(\hat{\rho}_T - 1)$ tiene la distribución descrita bajo el caso 4 en la Tabla A.1</p> <p>$(\hat{\rho}_T - 1) / \hat{\sigma}_{\rho_T}$ tiene la distribución descrita bajo el caso 4 en la Tabla A.2</p> <p>Test F de MCO de la hipótesis conjunta que $\delta = 0$ y $\rho = 1$ tiene la distribución descrita bajo el caso 4 de la Tabla A.3</p>

La principal interrogante en este punto es la determinación del caso a emplear para testear la hipótesis nula de presencia de raíz unitaria.

Si el analista tiene una hipótesis nula específica acerca del proceso que generó la data esta será obviamente la guía para efectuar el test. En la ausencia de tal guía, un principio general puede ser formar una especificación que sea una buena descripción de la data bajo ambas hipótesis, la nula y la alternativa. Este principio puede sugerir utilizar el caso 4 del test para series con tendencia y el caso 2 para series con una tendencia no significativa.

Por ejemplo, considerando el caso en que se desee probar la presencia de raíz unitaria en la tasa nominal de interés. No existe ninguna teoría económica que sugiera una tendencia en las tasas de interés, de manera que la hipótesis nula natural es que el verdadero proceso es un camino aleatorio sin tendencia, la hipótesis alternativa más acertada es que la serie sigue un proceso AR(1), con un coeficiente asociado al primer rezago menor a uno. Se podría argumentar también que el proceso es estacionario y con una media positiva, por lo cual la tasa de interés se debería modelar con una constante, bajo la hipótesis nula de que el verdadero proceso no contiene un término constante. De manera que el caso 2 sería el enfoque más adecuado para la data.

Se puede realiza otro ejemplo considerando la inflación mensual peruana (1992:01 al 1998:07), la cual es presentada en la Figura 16.3. La hipótesis nula que planteamos en este caso es que la inflación tiene un proceso raíz unitario sin constante, es decir, nos estamos refiriendo al caso 1. Como hipótesis alternativa se consideraría que la series en un proceso AR(1) sin constante. Los resultados obtenidos son los siguientes:

$$y_t = 0.946691y_{t-1} \\ (0.004921)$$

donde la desviación estándar está entre paréntesis. El tamaño muestral es de $T=80$. Aplicando el test de Dickey Fuller:

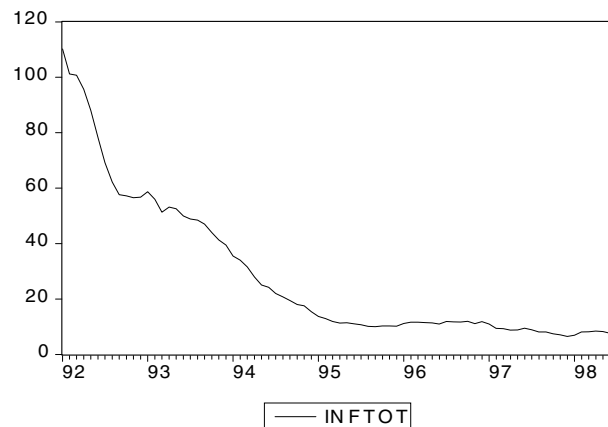
$$T(\hat{\rho} - 1) = 80(0.946691 - 1) = -4.26472$$

Dado que $-4.26472 > -7.9$, la hipótesis nula de que la inflación esta caracterizada por un camino aletaorio sin constante se acepta al 5% de confianza.

Finalmente, debemos tomar en cuenta que estos tests son válidos cuando no se presenta autocorrelación serial de los errores. Como sabemos, en modelos que presentan una variable endógena rezagada, si el error presenta autocorrelación, el estimador MCO nos da parámetros inconsistentes. Más adelante regresaremos a este tema cuando consideremos pruebas que permiten que los errores tengan autocorrelación.

Un resultado interesante de estos tests es la normalidad de las pruebas cuando estamos en el caso 3. No se va a profundizar en detalles, pero el lector interesado puede consultar West (1988)¹¹.

Figura 16.3



Debe tenerse en cuenta que para estas pruebas, el planteamiento es el siguiente:

$$H_0 : \rho = 1$$

$$H_1 : \rho < 1$$

Si se considera el caso de una distribución t, esta hipótesis se verificaría con la cola del lado izquierdo, tomando en cuenta los valores negativos de la distribución. Sin embargo, dada la naturaleza de la distribución cuando $\rho = 1$, se modifica esta distribución ampliando la cola de la

¹¹ West, Keneth D. (1987) "Asymptotic Normality when regressors have a unit root". *En*: *Econometrica*, 56: 1397-1418.

izquierda, por lo que comparado con los valores de una distribución t estándar tendremos que los valores de rechazo estarán muy por debajo del valor usual -2. Por consiguiente, si el valor calculado está por encima de -2 no se podría rechazar la hipótesis nula.

Tal como se aprecia en el cuadro anterior, los tests se calculan para un proceso AR(1). Usualmente, el estimador MCO para el caso estacionario tiende a subestimar al parámetro ρ en muestras pequeñas; por lo que la diferencia $(\rho - 1)$ tenderá a ser negativa. En este sentido el test F se refiere a la prueba donde se comparan las sumas de residuos al cuadrado de las estimaciones restringidas y no restringidas. En la estimación restringida se imponen las restricciones que se mencionan.

Tal como se mencionó anteriormente, las pruebas descritas hasta el momento, requieren que no exista autocorrelación en los errores. Para resolver este problema, surgieron varias propuestas. Sin embargo, citaremos las pruebas más usadas que permiten ciertos cambios y correcciones a las pruebas antes descritas, de tal forma que toma en cuenta la posible presencia de autocorrelación de los errores.

Uno de los caminos adoptados es calcular los estadísticos arriba mencionados, pero corregir por ciertas expresiones. Este es el espíritu de las pruebas propuestas por Phillips y Perron¹². La otra alternativa es extender el orden de la autorregresión de las ecuaciones estimada incluyendo rezagos adicionales de las diferencias de la variable que se está analizando. Ese es el espíritu de las pruebas llamadas Dickey Fuller Aumentadas (DFA). En esta sección haremos simplemente una descripción de las pruebas y nos referiremos a la intuición que está detrás de ellas.

16.8.3 Test de Phillips y Perron

En el primer caso Phillips y Perron sugieren calcular los estadísticos $T(\hat{\rho}_T - 1)$ y t usuales y someterlos a las siguientes correcciones:

Para el primer caso la expresión a calcular sería:

$$Z_{\rho} = T(\hat{\rho} - 1) - (1/2)(T^2 \cdot \hat{\sigma}_{\rho}^2 / s^2) \cdot (\hat{\lambda}^2 - \hat{\gamma}_0)$$

mientras que para la prueba t :

$$\left(\hat{\gamma}_0 / \hat{\lambda}^2 \right)^{1/2} t - \left\{ (1/2)(\hat{\lambda}^2 - \hat{\gamma}_0)(T \cdot \hat{\sigma}_{\rho} / s) \div \hat{\lambda} \right\}$$

Si bien las expresiones pueden aparentar ser un poco complicadas, contienen elementos que son calculables sobre la base de los resultados de una regresión que incluya los valores de la variable dependiente rezagada tanto sin constante, con constante y/o con tendencia.

Por ejemplo, el término $\hat{\lambda}^2$ que se incorpora en la corrección de las pruebas habituales se aproxima por el estimador de Newey-West y es como sigue:

¹² Phillips (1987) y Phillips y Perron (1988) son las referencias obligadas para este tipo de tests.

$$\hat{\lambda}^2 = \hat{\gamma}_0 + 2 \sum_{j=1}^q [1 - j/(q+1)] \hat{\gamma}_j$$

$$\text{donde } \hat{\gamma}_j = T^{-1} \sum \hat{\mu}_t \hat{\mu}_{t-j}$$

La última expresión es la autocovarianza de orden j del error de la ecuación estimada según sea el caso la variable. q se refiere al número de autocovarianzas consideradas; S^2 se refiere a la varianza de los residuos de la ecuación estimada. Debe tenerse cuidado que esta expresión difiere de $\hat{\gamma}_0$ dado que se divide la suma de errores al cuadrado de la ecuación estimada entre $T - k$, donde k es el número de parámetros estimados en la ecuación y no entre T .

La idea es que las correcciones introducidas permiten resolver el problema de inconsistencia de los parámetros cuando el error presenta autocorrelación en todas las ecuaciones estimadas. La corrección anterior se basa, como ya se mencionó, en aquella propuesta por Newley y West (1987). Esta corrección es arbitraria y como ella existen otras ponderaciones posibles; pero estas correcciones parecen funcionar bien dado que permiten obtener una prueba robusta y consistente.

Un resumen de las pruebas de este tipo, las presentamos en el siguiente cuadro:

Resumen de los Tests de Phillips y Perron en Presencia de Correlación Serial
<p>Caso 1: Regresión estimada : $y_t = \rho y_{t-1} + \mu_t$</p> <p>Proceso Verdadero: $y_t = y_{t-1} + \mu_t$</p> <p>Z_ρ tiene la distribución asintótica descrita bajo el caso 1 en la Tabla A.1</p> <p>Z_t tiene la distribución asintótica descrita bajo el caso 1 en la Tabla A.2</p>
<p>Caso 2: Regresión estimada : $y_t = \alpha + \rho y_{t-1} + \mu_t$</p> <p>Proceso Verdadero: $y_t = y_{t-1} + \mu_t$ $\mu_t \approx \text{i.i.d.N}(0, \sigma^2)$</p> <p>$Z_\rho$ tiene la distribución asintótica descrita bajo el caso 2 en la Tabla A.1</p> <p>Z_t tiene la distribución asintótica descrita bajo el caso 2 en la Tabla A.2</p>
<p>Caso 4: Regresión estimada : $y_t = \alpha + \rho y_{t-1} + \delta t + \mu_t$</p> <p>Proceso Verdadero: $y_t = \alpha + y_{t-1} + \mu_t$ cualquier α $\mu_t \approx \text{i.i.d.N}(0, \sigma^2)$</p> <p>$Z_\rho$ tiene la distribución asintótica descrita bajo el caso 4 en la Tabla A.1</p> <p>Z_t tiene la distribución asintótica descrita bajo el caso 4 en la Tabla A.2</p>

16.8.4 Test de Dickey y Fuller

Dickey y Fuller tomaron otro camino, introduciendo rezagos de las diferencias de la variable dentro de la ecuación a estimar. El orden de la autorregresión debe ser lo suficientemente grande de tal forma de poder tener alguna seguridad de que los errores de las ecuaciones estimadas son ruido blanco. Si lo anterior se cumple entonces podremos emplear las tablas

usuales de los tests Dickey-Fuller. Estos tests son los llamados tests DFA (Dickey-Fuller aumentados).

En estos tests se supone que la variable a analizar tiene la siguiente representación:

$$y_t = \sum_{i=1}^p \rho_i y_{t-i} + \varepsilon_t$$

Con lo cual se puede construir un test a partir del siguiente modelo:

$$y_t = \rho y_{t-1} + \sum_{i=1}^{p-1} \xi_i \Delta y_{t-i} + \mu_t$$

Este tipo de regresión se puede generalizar de tal forma que incluya un término constante o una tendencia. Uno de los resultados más importantes es que presentará las mismas distribuciones mencionadas para los casos de ausencia de autocorrelación. El siguiente cuadro nos da un resumen de estos tests.

Resumen de los Tests de Dickey Fuller Aumentado (DFA)
<p>Caso 1: Regresión estimada : $y_t = \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \dots + \xi_{p-1} \Delta y_{t-p+1} + \rho y_{t-1} + \mu_t$ Proceso Verdadero: La misma especificación que la regresión estimada con $\rho = 1$. Cualquier test t o F que involucre a los parámetros ξ puede ser comparado con las tablas usuales para un test asintóticamente válido. Z_{DF} tiene la misma distribución asintótica que la variable descrita bajo el caso 1 en la Tabla A.1 $(\hat{\rho}_T - 1) / \hat{\sigma}_{\rho_T}$ tiene la distribución descrita bajo el caso 1 en la Tabla A.2</p>
<p>Caso 2: Regresión estimada : $y_t = \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \dots + \xi_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \mu_t$ Proceso Verdadero: La misma especificación que la regresión estimada con $\alpha = 0$ y $\rho = 1$. Cualquier test t o F que involucre a los parámetros ξ puede ser comparado con las tablas usuales para un test asintóticamente válido. Z_{DF} tiene la distribución descrita bajo el caso 2 en la Tabla A.1 $(\hat{\rho}_T - 1) / \hat{\sigma}_{\rho_T}$ tiene la distribución descrita bajo el caso 2 en la Tabla A.2 Test F de MCO de la hipótesis conjunta que $\alpha = 0$ y $\rho = 1$ tiene la distribución descrita bajo el caso 2 de la Tabla A.3</p>
<p>Caso 3: Regresión estimada : $y_t = \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \dots + \xi_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \mu_t$ Proceso Verdadero: La misma especificación que la regresión estimada con $\alpha \neq 0$ y $\rho = 1$. Cualquier test t o F que involucre a cualquiera de los parámetros puede ser comparado con las tablas usuales para un test asintóticamente válido.</p>
<p>Caso 4: Regresión estimada : $y_t = \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \dots + \xi_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \delta t + \mu_t$ Proceso Verdadero: La misma especificación que la regresión estimada con α con cualquier valor, $\delta = 0$ y $\rho = 1$. Cualquier test t o F que involucre a los parámetros ξ puede ser comparado con las tablas usuales para un test asintóticamente válido. Z_{DF} tiene la distribución descrita bajo el caso 4 en la Tabla A.1 $(\hat{\rho}_T - 1) / \hat{\sigma}_{\rho_T}$ tiene la distribución descrita bajo el caso 4 en la Tabla A.2 Test F de MCO de la hipótesis conjunta que $\delta = 0$ y $\rho = 1$ tiene la distribución descrita bajo el caso 4 de la Tabla A.3</p>

En este caso, el número p debe ser grande para obtener buenos resultados. Una sugerencia para los tests de Phillips Perron y DFA es analizar los residuos de las ecuaciones estimadas y verificar que son ruido blanco con test de autocorrelación generales como el Breusch-Godfrey.

Muchos paquetes econométricos incluyen tests de raíces unitarias. Por ejemplo el EViews, realiza una prueba Dickey Fuller Aumentada. Como ejemplo procedemos a efectuar el test del paquete a la serie de Inflación mensual peruana (1992:01 al 1998:07) cuyos resultados son los siguientes:

ADF Test Statistic	-4.174110	1% Critical Value*	-2.5929
		5% Critical Value	-1.9445
		10% Critical Value	-1.6180

*MacKinnon critical values for rejection of hypothesis of a unit root.

Tal como se puede ver se incluyen valores estadísticos de un trabajo de MacKinnon que han demostrado ser útiles en estos casos. En estos casos, si el valor calculado es menor en valor absoluto a los valores de tabla no se puede rechazar la hipótesis nula de presencia de raíz unitaria, es decir, acepta la presencia de raíz unitaria. En el cuadro anterior, la hipótesis nula de la presencia de raíz unitaria no se acepta ni al 1% de significación estadística. La regresión que estima este paquete es la siguiente:

Augmented Dickey-Fuller Test Equation				
LS // Dependent Variable is D(INFTOT)				
Sample(adjusted): 1992:03 1998:07				
Included observations: 77 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
INFTOT(-1)	-0.029356	0.007033	-4.174110	0.0001
D(INFTOT(-1))	0.382840	0.097133	3.941383	0.0002
R-squared	0.520271	Mean dependent var	-1.217882	
Adjusted R-squared	0.513875	S.D. dependent var	2.186316	
S.E. of regression	1.524358	Akaike info criterion	0.868777	
Sum squared resid	174.2750	Schwarz criterion	0.929655	
Log likelihood	-140.7062	F-statistic	81.33839	
Durbin-Watson stat	1.555230	Prob(F-statistic)	0.000000	

Debe notarse que en regresión estimada, la variable dependiente no es la serie en niveles sino la serie en diferencias. Esta reparametrización presenta la ventaja de que el coeficiente de la primera diferencia ya es $\rho - 1$, por lo que al dividirla entre su desviación estándar nos da un a prueba t de significación estadística del parámetro en cuestión que en este caso debe ser cero. Tal como se aprecia, es justamente la prueba t que arroja el programa para el coeficiente de y_{t-1} la que se utiliza para compararla con los valores de tablas.

Entre las opciones, este programa permite introducir el número de rezagos que sea necesario, introducir una constante y/o una tendencia. La significancia de los rezagos de la variable Inflación (INFTOT) en diferencias es un indicador de la autocorrelación de los errores. Esta autocorrelación hace que el estimador hallado al aplicar el test de Dickey Fuller sin presencia de autocorrelación sea inconsistente y por lo tanto, no sean de utilidad los resultados hallados. Luego de eliminar los rezagos no significativos y la constante y tendencia si fuera el caso, se procede a verificar que los errores son ruido blanco.

16.8.5 Raíces Unitarias Múltiples¹³

Existen algunas series que pueden presentar más de una raíz unitaria. Este puede ser el caso de los no inventarios de empresas. Cuando se desea verificar si una serie es integrada de orden

¹³ Esta sección es tomada de Banerjee et. el. (1993)

mayor a 1, la secuencia que parecería ser lógica seguir sería testear la presencia de raíces unitarias en las series sin diferenciar; y luego (si esta primera hipótesis no puede ser rechazada), proceder a evaluar las series en diferencias para estudiar la presencia de una segunda raíz unitaria. Sin embargo, este procedimiento no es estadísticamente válido porque todos los tests de raíces unitarias considerados aquí toman como alternativa la ausencia absoluta de raíces unitarias.

Dickey y Pantula (1987) sugieren una secuencia del test más natural que toma en cuenta el número más grande de raíces unitarias bajo consideración como la primera hipótesis a mantener y entonces decrece el orden de diferenciación cada vez que la hipótesis nula es rechazada. Esto continúa hasta que la hipótesis nula no puede ser rechazada.

Ilustremos el caso para $d = 2$. consideremos el modelo AR(2) de la siguiente manera:

$$(1 - \rho_1 L)(1 - \rho_2 L)y_t = \mu_t$$

Este modelo puede ser reparametrizado de la siguiente manera:

$$\Delta^2 y_t = \beta_1 \Delta y_{t-1} + \beta_2 y_{t-2} + \mu_t$$

donde $\beta_1 = (\rho_1 \rho_2 - 1)$ y $\beta_2 = -(1 - \rho_1)(1 - \rho_2)$. EL test consiste de los siguientes dos pasos:

1. Verificar la nula de dos raíces unitarias contra la alternativa de una sola raíz unitaria. Bajo la hipótesis nula $\beta_1 = \beta_2 = 0$ y puede usarse un test F para probar lo anterior. Sin embargo, tal test no toma en cuenta la naturaleza de una cola del verdadero test. Una manera más poderosa se obtiene notando que, bajo ambas hipótesis $\beta_2 = 0$. Sin embargo, $\beta_1 = 0$ bajo la nula pero es menor que cero bajo la alternativa. Así un test más poderoso es obtenido estimando la regresión de $\Delta^2 y_t$ sobre Δy_{t-1} , computando el ratio t del parámetro de esta última variable y realizando el test de una cola utilizando las tablas de Dickey Fuller.
2. Si la nula del punto 1. es rechazada, procedemos a verificar la nula de una raíz unitaria versus la alternativa de estacionariedad. Aquí la nula y la alternativa están dadas por $\beta_1 < 0, \beta_2 = 0$ y $\beta_1 < 0, \beta_2 < 0$, respectivamente. Así un test t de una sola cola involucra estimar la regresión de $\Delta^2 y_t$ sobre Δy_{t-1} e y_{t-1} y computar el ratio t del coeficiente de esta última variable y compararlo con los valores de las tablas D-F

Este procedimiento puede generalizarse para verificar la existencia de tres o más raíces unitarias. El artículo de Dickey y Pantula (1987) contiene los resultados de estudios de simulación. su conclusión general es que el procedimiento secuencial es más poderoso que el test F.

16.9. QUIEBRE ESTRUCTURAL Y RAÍCES UNITARIAS

Desde el trabajo de Nelson y Plosser (1982) se empezó a considerar a muchas series económicas y financieras como series no estacionarias debido a la presencia de raíces unitarias probada por estos tests. De alguna forma, esta manera de modelar las series implicó dar algún sustento empírico a la teoría de los ciclos económicos reales, ya que cambios tecnológicos

podrían alterar la trayectoria del producto. Debe recordarse que el artículo de Kydland y Prescott se publicó en *Econometría* el mismo año.

Sin embargo, no todos estaban de acuerdo con esta hipótesis. Una era la existencia de cambios estructurales en la serie. Un cambio estructural afecta la parte determinística de una serie dado que la evolución o tendencia de largo plazo puede modificarse. Ante eventos como cambios institucionales, guerras o grandes crisis pueden explicar este tipo de comportamientos. Esta variación de la estructura de la serie es independiente de los efectos, sean persistentes o no, de los shocks. Sin embargo, los tests de raíz unitaria no consideraban estas situaciones.

Por este motivo, surgieron artículos cuestionando la visión de las raíces unitarias. Entre los más famosos destacan las investigaciones de Perron (1989) y Rappoport y Reichlin (1989)¹⁴. Ambos artículos sostienen que el poder de los tests de raíces unitarias disminuye cuando se permite cambios exógenos en las series debido a hechos conocidos. Así, Perron sostiene que la gran crisis de 1929 implicó un cambio de nivel de la serie permanente, o mientras que la crisis del petróleo de 1973 implicó un cambio en la pendiente de la tendencia a partir de dicho año, lo que explicaría la reducción de las tasas de crecimiento de la economía norteamericana.

Estos tests consideraron que el investigador debe tener información a priori de las fechas de los cambios estructurales.

En su artículo original, Perron considera como hipótesis nula series que presentan una raíz unitaria y posiblemente un drift diferente de cero. Sin embargo, este enfoque es modificado de tal forma que incluya un cambio en la estructura en el período T_B , ($1 < T_B < T$). Las posibles formas de quiebres que se consideran bajo la hipótesis nula son las siguientes: uno que permite un cambio exógeno en el nivel de la serie, (Modelo A), uno que permite un cambio exógeno en la tasa de crecimiento (Modelo B) y uno que permite cambios tanto en nivel como en tasa de crecimiento (Modelo C).

Estas hipótesis se parametrizan de la siguiente manera:

Hipótesis nula:

$$\text{Modelo (A): } y_t = \mu + dD(TB)_t + y_{t-1} + e_t$$

$$\text{Modelo (B): } y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1)DU_t + e_t$$

$$\text{Modelo (C): } y_t = \mu_1 + y_{t-1} + dD(TB)_t + (\mu_2 - \mu_1)DU_t + e_t$$

donde: $dD(TB)_t = 1$ si $t > T_B + 1$, 0 de otro modo

$$DU_t = 1 \quad \text{si } t > T_B, \quad 0 \text{ de otro modo}$$

$$A(L)e_t = B(L)v_t$$

$$v_t \sim N(0, \sigma^2)$$

con $A(L)$ y $B(L)$ polinomios de orden p y q respectivamente. La innovación $\{e_t\}$ se toma como si siguiera un proceso ARMA (p,q) con órdenes p y q posiblemente desconocidos. Esta forma permite que la serie y_t represente un proceso más general.

¹⁴ Perron, Pierre (1989) "The Great Crash, the oil price shock and the unit root hypothesis" EN: *Econometrica* Vol. 57 No. 6 pp 1361-1401

Rappoport, Peter y Lucrezia Reichlin (1989) "Segmented trends and non-stationary time series" EN: *The Economic Journal* Vol 89 pp. 168-177

En vez de considerar la hipótesis alternativa que y_t es un proceso estacionario alrededor de un tendencia determinística con parámetros que no cambian en el tiempo, analiza los siguientes tres modelos:

Hipótesis alternativas:

$$\text{Modelo (A): } y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)DU_t + e_t$$

$$\text{Modelo (B): } y_t = \mu_1 + \beta_1 t + (\beta_2 - \beta_1)DT_t^* + e_t$$

$$\text{Modelo (C): } y_t = \mu_1 + \beta_1 t + (\mu_2 - \mu_1)DU_t + (\beta_2 - \beta_1)DT_t + e_t$$

donde: $DT_t^* = t - T_B$ y $DT_t = t$ si $t > T_B$ y 0 de otro modo

A manera de resumen, los resultados obtenidos por Perron arrojan que si cualquiera de los cambios estructurales es significativo, es imposible rechazar la hipótesis de raíz unitaria cuando se utiliza los tests habituales (Dickey Fuller o ADF). Cuanto más fuerte es el cambio estructural, la distribución del estadístico t se acerca más a cero, sesgando el resultado del test a aceptar la hipótesis de raíz unitaria. Cuando existen cambios en la pendiente de la tendencia el test se vuelve inconsistente; en los demás casos el parámetro del valor rezagado de la pendiente tiende a 1 cuando t tiende a infinito, a pesar de que en el verdadero proceso este parámetro es menor a uno. Así se encuentra que los cambios [estructurales] introducidos en la serie crean una raíz unitaria espúrea que no desaparecerá ni siquiera cuando aumentan el número de observaciones.

Por ello, Perron plantea las siguientes tablas para testear las hipótesis nulas planteadas en los modelos A, B y C. Se debe notar que el parámetro λ se refiere al tiempo relativo con respecto a toda la muestra en que ocurre el cambio estructural así $\lambda = T_B/T$. (Ver Tabla A.2).

Con estas tablas encuentra que de las 14 series consideradas por Nelson y Plosser en su estudio, 11 no presentan raíz unitaria sino quiebres en su tendencia.

Los anteriores fueron los principales resultados de Perron. Sin embargo, para resaltar lo que está en juego con estas pruebas nos permitimos reproducir los comentarios finales de su artículo.

“Cuando verificamos la presencia de una raíz unitaria en una serie de tiempo contra la hipótesis de fluctuaciones estacionarias alrededor de un tendencia determinística, el uso de una gran cantidad de datos tiene ventajas definitivas. Ello permite testear con mayor poder, en muchos casos, en comparación con datos de muestras de menor tamaño, aún si ésta última permite tener más observaciones. El problema, sin embargo, es que el conjunto de datos de gran tamaño tiene más oportunidades de incluir un evento mayor el cual uno podría considerar como un outlier o como exógeno dada su importancia relativa. El argumento en este artículo descansa en el postulado que dos de tales eventos han ocurrido en el siglo XX: la crisis de 1929 y la reducción del crecimiento desde el shock del petróleo de 1973. Por lo tanto, consideramos como una alternativa relevante una función de tendencia con un cambio de intercepto en 1929 y un cambios de pendiente después de 1973.”

“Lo que tenemos en mente cuando especificamos nuestras hipótesis mantenidas puede ser parametrizada como sigue:

$$y_t = \eta_t + Z_t, \quad \eta_t = \mu_t + \beta_t t$$

donde $A(L)Z_t = B(L)e_t$; $e_t \sim i.i.d. (0, \sigma^2)$; $\mu_t = \mu_{t-1} + V(L)v_t$ y $\beta_t = \beta_{t-1} + W(L)\omega_t$. Aquí los Z 's son desviaciones (no necesariamente estacionarias) de la función de tendencia η_t . El intercepto y la pendiente de la función de tendencia (μ_t y β_t) son por sí mismas variables aleatorias, modeladas como procesos integrados con $W(L)$ y $V(L)$ polinomios estacionarios e invertibles. Sin embargo, la diferencia importante es que el periodo de ocurrencia de los shocks [que afectan a los coeficientes de la tendencia] v_t y ω_t son raros, relativos a la secuencia de innovaciones $\{e_t\}$ [que afectan a las desviaciones Z]. La idea intuitiva dentro de este tipo de modelación es que los coeficientes de la función de tendencia son determinados por variables fundamentales económicas de largo plazo (i.e la estructura de la organización económica, crecimiento poblacional, etc.) y que dichos fundamentos cambian muy raramente. En nuestros ejemplos v_t no es cero en 1929 (la gran depresión) y ω_t es diferente de cero en 1973 (el shock del precio del petróleo).

En ese sentido, nuestro supuesto de exogeneidad acerca de los cambios en la función de tendencia es un hecho que permite tomar dichos shocks fuera de la función de ruido (los errores) $[Z]$ y dentro de la función de tendencia sin una modelación específica de la naturaleza estocástica del comportamiento de μ_t y β_t . Es en este sentido que nuestro enfoque no provee una representación incondicional de las propiedades de series de tiempo de las variables”.

“Dentro del enfoque mencionado, el propósito del artículo es testear si Z_t es un proceso integrado [de orden uno] o no, lo cual implica testear que los shocks e_t tienen efectos persistentes que no se desvanecen en un horizonte largo. Nuestro enfoque es remover de la función de ruido dos eventos que ocurren en dos fechas donde creemos acontecieron shocks positivos del tipo de v_t y ω_t y los modelamos como parte de la función de tendencia. El hecho que modelemos dichos cambios como exógenos implica que nuestros resultados son condicionales. Esto es, condicionadas en la presencia de un cambio en la función de tendencia en 1929 y 1973, las fluctuaciones son estacionarias (i.e Z_t es estacionaria).

Una importante dirección para investigaciones futuras es hacer de este resultado condicional un planteamiento incondicional. Esto puede, en principio, ser alcanzado por un test directo de cambio estructural en la función de tendencia. En un sentido, nuestro procedimiento permite tal test, pero condicional en un cambio ocurrido en una fecha conocida fija. Así problemas de pre -testeo y ”data mining” podrían aparecer si tomamos en cuenta el rol de mirar, ex-post, los datos sobre la elección de la fecha. De acuerdo con lo anterior, lo que es necesario es un test de cambios estructurales en la función de tendencia que ocurre en una fecha desconocida. El problema, sin embargo, es que debe tenerse cuidado en asegurar que el test tenga un tamaño importante bajo las hipótesis de raíz unitaria y estacionariedad en tendencias. Ningún test está disponible actualmente en la literatura. Esperamos, reportar, en un futuro cercano, desarrollos en esta área y aplicaciones en este contexto”.

Finalmente, hay que resaltar las implicancias de los resultados de Perron en el debate acerca de la efectividad de políticas. Si la mayoría de las series económicas presentaran cambios estructurales pero no raíces unitarias, se debería buscar cambios estructurales profundos en la economía, de manera que alteren la función de tendencia de las series; y luego que la esta se ha alterado, no habría espacio para política activa de corto plazo, dado que a la larga la serie regresaría a su valor de tendencia. Y si este fuera el caso, sería importante no introducir ruido adicional al proceso, con políticas sorpresivas que podrían ser contraproducentes.

16.10. NUEVOS AVANCES

Tal como lo sostuvo Perron la agenda de investigación estaba abierta y muy recientemente se han producido avances en ese sentido. Tal como el texto lo indicaba, Hamilton había planteado procesos probabilísticos en donde la función de tendencia cambia en el tiempo. Recientemente, Hamilton (1994)¹⁵ ha avanzado en dicho sentido planteando que los cambios en los regímenes de las series se pueden modelar suponiendo probabilidades para diferentes estados de la economía. La herramienta que usa para ello son las cadenas de Markov. No vamos a desarrollar este tema porque implica el manejo de conceptos un poco más elaborados.

Sin embargo, en otra dirección que sí se ha avanzado es en el perfeccionamiento de los tests del tipo de Perron. Así Perron y otros autores publicaron en 1992 un número especial del *Journal of Business & Economic Statistics* acerca de puntos de quiebre y raíces unitarias. Allí están los trabajos de Perron y Voglensang, Banerjee, Lumsdaine y Stock y un paper de Andrews y Zivot.

Nos limitaremos a desarrollar empíricamente dos de los artículos mencionado y el test propuesto en ellos. Una de las limitaciones del test de Perron era la exogeneidad de los cambios en la tendencia que se suponían conocidos. Los artículos de Perron y Voglensang y Andrews y Zivot van más allá permitiendo determinar endógenamente la fecha de quiebre y verificando si existe raíz unitaria tomando en cuenta la fecha de quiebre determinada.

Perron y Voglensang desarrollaron un test para verificar la presencia de una raíz unitaria permitiendo un posible cambio en el nivel de la serie en una fecha que es desconocida. Ellos parten suponiendo que la fecha de quiebre se denota por T_B donde $1 < T_B < T$, donde T es el tamaño muestral. Sobre la base de lo anterior consideran dos modelos.

16.10.1 El additive outlier model

Este modelo supone que el cambio ocurre instantáneamente. En particular, el efecto del cambio sobre el nivel de la serie y_t no depende de la dinámica exhibida por la estructura de correlación de y_t . Bajo la nula de raíz unitaria, este modelo puede ser parametrizado como:

$$y_t = \delta D(TB)_t + y_{t-1} + w_t$$

$$\text{donde } D(TB)_t = 1 \text{ si } t = T_B + 1 \text{ y } 0 \text{ de otro modo}$$

Se supone que $y_1 = y(1)$ es tanto una constante o una variable aleatoria¹⁶. En este caso se supone que w_t sigue un proceso ARMA(p,q). Específicamente se supone que $A^*(L)w_t = B(L)e_t$ donde e_t es i.i.d. $(0, \sigma^2)$ con cuarto momento finito¹⁷. Cada uno de los polinomios se supone que tiene raíces estrictamente fuera del círculo unitario. Bajo el modelo anterior, la media de la serie es $y(1)$ hasta el período T_b e $y(1) + \delta$ en adelante.

Bajo la hipótesis alternativa, la serie $\{y_t\}$ no contiene una raíz unitaria y puede ser caracterizada por la siguiente especificación:

¹⁵ Hamilton, James (1994) "Time Series Analysis" capítulo 22 'Modelando series de tiempo con cambios en regímenes'. Princeton University Press. U.S.A.

¹⁶ Valor que toma la variable "y" en el periodo 1. Es una constante en tanto este sea considerado el valor inicial de donde parte la variable, y es aleatoria en tanto proviene, por su estructura, de la suma de errores aleatorios hasta el periodo 1.

¹⁷ $A(L) = (1 - L)A^*(L)$, donde $A(L)$ es un polinomio con $p+1$ rezagos.

$$y_t = c + \delta DU_t + v_t \quad t = 2, \dots, T$$

donde $DU_t = 1$ si $t > T_b$ y 0 de otro modo

La secuencia de errores $\{v_t\}$ es un proceso ARMA (p+1,q) de la forma $A(L)v_t = B(L)e_t$. En este caso, la media de la serie es c hasta el período T_b y $c + \delta$ de allí en adelante. La hipótesis nula especificada por la primera ecuación es un caso especial de la especificación realizada en la segunda ecuación cuando $c = y(1)$ y $A(L) = (1 - L)A^*(L)$.

De esta manera, debe notarse que ambas especificaciones son iguales salvo por la presencia de δDU_t . Por ello, una estrategia para poder llevar a cabo el test es remover un estimado de la parte determinística de la serie y testear si el ruido remanente es caracterizado por la presencia de una raíz unitaria.

Así para un valor de la fecha de quiebre T_b fija, tenemos el siguiente proceso en dos etapas:

1. Se remueve la parte determinística de la serie utilizando los estimados de la regresión:

$$y_t = \mu + \delta DU_t + \tilde{y}_t$$

2. Se corre entonces la siguiente regresión:

$$\tilde{y}_t = \sum_{i=0}^k \omega_i D(TB)_{t-i} + \alpha \tilde{y}_{t-1} + \sum_{i=1}^k c_i \Delta \tilde{y}_{t-i} + e_t$$

$$t = k + 2, \dots, T$$

y en esta ecuación se plantea la hipótesis de que $\alpha = 1$ usando el estadístico t .

La inclusión de las variables dummy $D(TB)_{t-i}$ ($i = k+2, \dots, T$) en la segunda regresión se hacen con el fin de que esta regresión sea comparable a otra que veremos más adelante y la misma tabla se pueda aplicar a ambas regresiones (o pruebas). Nótese que se incluyen k rezagos de la diferencia de la variable filtrada para asegurar que la prueba sea del tipo de los ADF. Si bien hemos considerado que los errores siguen un proceso ARMA (p,q), se utiliza la estrategia planteada por Said y Dickey (1984) de aproximar este polinomio por un AR(k). Otro aspecto que hay que notar es que la fecha de quiebre T_b , se debe determinar de algún modo. Una forma razonable de hacer ello, será calculando recursivamente la regresión (esto es, para cada período se calcula la regresión) y se tomará como punto de quiebre T_b , aquel donde el estadístico t de α sea el menor en valor absoluto. Con esto se trata de realzar el hecho de que la serie aún en el punto más bajo de este estadístico cumple con la hipótesis nula de raíz unitaria. Así se le está pidiendo al modelo que cumpla con la hipótesis nula en todos los períodos.

Otra estrategia para seleccionar el punto de quiebre es tratar de maximizar en valor absoluto, el estadístico t de δ , con lo cual se le estaría dando mayor énfasis a la alternativa de un quiebre estructural. Con ello se le pide a la regresión que la hipótesis de raíz unitaria se cumpla aún en el período en donde la alternativa de quiebre estructural es más fuerte.

Con esto debe quedar claro que la hipótesis planteadas por el Test de Perron y Volgensang son las siguientes:

- H_0 : La serie tiene raíz unitaria más drift con un gran shock exógeno.
 H_1 : La serie es estacionaria con un cambio en la tendencia.

El test permitirá sólo un quiebre estructural dentro de la muestra por ello se escogen los valores extremos de las pruebas t respectivas (maximizar el valor absoluto δ o minimizar el valor absoluto de α).

El estadístico t para testear $\alpha = 1$ en la regresión del segundo paso, Perron y Volgensang lo denotan como $t_{\alpha}(AO, T_b, k)$, donde T_b denota el período de quiebre y k el orden de los rezagos de las diferencias de la variable filtrada incluida en la regresión.

16.10.2 El Innovational Outlier Model

Este modelo llamado IO supone que el cambio tiene lugar pero gradualmente, lo que implica que hay un período de transición. Aunque en principio, el efecto dinámico podría tomar cualquier forma, una manera natural y simple de modelar dicha transición es suponer que la economía responde a un shock a la función de tendencia (aquí un cambio en la media) de la misma manera que reacciona a otros shocks.

El supuesto anterior nos lleva a la siguiente especificación bajo la nula de raíz unitaria:

$$y_t = y_{t-1} + \psi(L)(e_t + \theta D(TB)_t) \\ t = 2, \dots, T$$

donde $\psi(L) = A^*(L)^{-1}B(L)$ define una representación de medias móviles con $A^*(L)$ y $B(L)$ definidos como en el caso anterior. El efecto impacto del cambio en la media es θ y el impacto de largo plazo está dado por $\psi(1)\theta$.

Bajo la hipótesis alternativa de fluctuaciones estacionarias, el modelo es representado por:

$$y_t = a + \phi(L)(e_t + \delta D(TB)_t) \\ t = 2, \dots, T$$

donde $\phi(L) = A(L)^{-1}B(L)$ con $A(L)$ definida como en el caso anterior. El impacto inmediato sobre el cambio de la media es δ y el impacto de largo plazo es $\delta\phi(1)$. Los dos modelos descritos en esta sección pueden ser anidados por un modelo de orden finito autorregresivo:

$$y_t = \mu + \delta DU_t + \theta D(TB)_t + \alpha y_{t-1} + \sum c_i \Delta y_{t-i} + e_t \\ t = k + 2, \dots, T$$

bajo la hipótesis nula de una raíz unitaria, $\alpha = 1$ (lo cual implica también que $\delta = 0$). Así la estrategia adecuada es usar el t estadístico de $\alpha = 1$ cuando esta última regresión estimada por MCO. Ese test estadístico, obtenido para valores dados de T_b y k será denotado $t_{\alpha}^{\wedge}(IO, T_b, k)$.

Para ambos estadísticos tanto T_b como k son desconocidos, por tanto hay que buscar formas de determinarlos. Tal como se mencionó anteriormente existen dos estrategias para determinar el punto de quiebre, uno referido al t del parámetro α (el mínimo calculado para todos los períodos) y otro referido al t máximo en valor absoluto en favor de la hipótesis del quiebre estructural (referida a la significación de δ). Esto implica tomar en cuenta la fecha en que dicho

estadístico es el máximo posible. Este procedimiento es el mismo para ambos modelos y por tanto se toma en cuenta las regresiones planteadas.

Para el caso de k (número de rezagos de las variables en diferencia respectivamente para cada modelo) existen varios métodos que describimos a continuación:

1. Uno que es exógenos a los datos en donde se supone que k es un valor fijo
2. Un método que es endógeno a los datos, implica que para un t_b dado, se escoge el número de rezagos en función a cierta significancia del último rezago. Así si el rezago de la diferencia rezagada en cuatro períodos (Δy_{t-4}) no es significativa al 10% con respecto al estadístico t, y el rezago en tres períodos sí lo es, seleccionamos $k=3$.
3. Un tercer método también dependiente de los datos es uno similar al anterior pero referente a un test F. Así se escoge el número k realizando la prueba de significancia del último rezago imponiendo la restricción de que el beta del último rezago es cero. Así es un test de mínimos cuadrados restringidos.

Las tablas que se utilizan para este test se presentan en un anexo.

16.10.3 El Test de Zivot y Andrews

Este test trata de verificar la hipótesis nula de presencia de raíz unitaria solamente. Esto contrasta con el test de Perron y Volgensang en que la nula es la presencia de raíz unitaria más un quiebre estructural. Este test es más general, dado que el anterior sólo se refería a la presencia de raíz unitaria o no, pero siempre con la presencia de un quiebre estructural.

Así estos autores plantean los siguientes modelos a estimar para tres casos distintos de hipótesis alternativas que plantean quiebre en niveles (A), quiebres en tendencia (B) y quiebres en tendencia y niveles (C).

Para cada una de estas alternativas proponen calcular las siguientes regresiones:

Test de Zivot y Andrews:	
Hipótesis nula : $y_t = \mu + y_{t-1} + e_t$	
<i>Modelo (A):</i> Alternativa: quiebre en niveles	Ecuación a estimar: $y_t = \mu + \theta DU_t + \beta t + \alpha y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + e_t$
<i>Modelo (B):</i> Alternativa: quiebre en la tendencia	Ecuación a estimar : $y_t = \mu + \beta t + \gamma DT_t + \alpha y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + e_t$
<i>Modelo (C):</i> Alternativa: quiebre en tendencia y niveles	Ecuación a estimar: $y_t = \mu + \theta DU_t + \beta t + \gamma DT_t + \alpha y_{t-1} + \sum_{j=1}^k c_j \Delta y_{t-j} + e_t$
En los tres casos: $DU_t = 1$ si $t > T_b$ y 0 de otro modo $DT_t = t - T_b$ si $t > T_b$ y 0 de otro modo	
Para seleccionar T_b se selecciona aquel periodo que da el t de α más bajo. Esto equivale a seleccionar el período que dé el menor t o lo que es lo mismo el valor t más elevado en valor absoluto. Deberá notarse que en las tablas que presentan estos autores también se consideran valores para T_b seleccionados a priori.	
Para la selección de k se utiliza el método 2 de Perron y Volgensang.	

En ambos tests se acepta la nula si el t de α del periodo T_b es mayor que el valor de tablas (o menor en valor absoluto). La hipótesis nula se rechaza si es que el t de α del período T_b es menor al de tablas (o mayor en valor absoluto).

Para el test de Zivot y Andrews se presentan los siguientes valores de tablas que se refieren a la selección de k bajo el método 2 de Perron y Volgensang.

Quiebres en:	1%	2.5%	5%	10%	50%	90%	95%	97.5%	99%
Niveles	-5.34	-5.02	-4.80	-4.58	-3.75	-2.99	-2.77	-2.56	-2.32
Tendencia	-4.93	-4.67	-4.42	-4.11	-3.23	-2.48	-2.31	-2.17	-1.97
Ambos	-5.57	-5.30	-5.08	-4.82	-3.98	-3.25	-3.06	-2.91	-2.72

*Los porcentajes se refieren al nivel de significancia

Para valores referidos a T_b determinado exógenamente, ver las tablas anexas

Los tests presentados y la discusión precedente son sólo una parte de la discusión. En la actualidad han aparecido otros tests que pretenden ser más exactos. Sin embargo, estos dos últimos tests resumen el estado del arte en cuanto a la verificación de presencias de raíces unitarias vs. las alternativas de series estacionarias con quiebres en tendencia. La ventaja del Test de Zivot y Andrews es que permite formular una alternativa en donde la serie es simplemente un random walk.

Si bien en la literatura no se han mostrado mucho estos tests, creemos que es relevante su estudio, dado el carácter cambiante de la economía y los mercados financieros durante los últimos años.

APENDICE 16.1**El Test de Zivot y Andrews
Programa en Econometric Views**

Para llevar a cabo el test de Z&A se tiene que seguir los siguientes pasos :

1. Determinar el modelo de hipótesis alternativa que se va a usar (A, B o C), lo cual implica conocer cual es el modelo que mejor explica a la serie ; con quiebre en media, en tendencia o ambas. En otras palabras, determinar el quiebre o conjunto de quiebres que presentan el mayor estadístico F .
2. Determinar el número de rezagos “k” de la diferencia de la serie, que garantiza que se elimina el problema de autocorrelación. Como sabemos, estamos usando el método 2 de Perron y Volgensang para hallarlo, es decir, se escoge el número de rezagos en función de cierta significancia del último rezago. Si el rezago cuarto no es significativo al 5% y el tercero si lo es, $k = 3$.
3. Correr la regresión de la hipótesis alternativa elegida recursivamente (probando cada año como punto de quiebre), considerando el menor valor del estadístico asociado a α como T_b y comparando este con el de tabla.

A continuación presentamos el programa.

Se genera un nueva serie denominada LSERIE con la cual se realizan todos los procedimientos. Se halla el máximo valor de k (!reg), el cual es la raíz cúbica del número total de observaciones. Se genera la tendencia.

```

genr lserie=log(pbi)
genr dlserie=d(lserie)
!reg=-@ceiling(!obs^(1/3))
genr t=@trend(1896)
!regfin=0

```

Se procede a determinar el número de rezagos “k” , el cual será guardado en la variable !regfin.

```

for !rreg=!reg to -1
equation temp.ls dlserie c lserie(-1) dlserie(-1 to !rreg) t
!mcoef=-!rreg+2
!tdist=@tdist(c(!mcoef)/sqr(@covariance(!mcoef,!mcoef)),temp.@regobs-temp.@ncoef)
if !regfin=0 and !tdist<0.05 then
!regfin=!rreg
genr regfin=!regfin
endif
next

```

Se procede a determinar la cota superior e inferior para las variables dummy. Esto debido a que no se puede regresionar contra una variable que tenga un sólo valor unitario el resto ceros, ya que generaría una matriz singular al ser una variable linealmente dependiente. Por ello, se eliminan el 15% de las observaciones superiores y 15% de las inferiores.

Se procede a generar las variables dummy en media y en tendencia dentro de las cotas.

```

!nui=1896
!nuf=1995
!cotau=!nuf-@ceiling(0.15*!obs)
!cotat=!nui+@ceiling(0.15*!obs)

```



```

FOR !num=!cotal to !cotau
  smpl !nui !num
  genr dum{!num}=0
  genr dut{!num}=0
  smpl !num+1 !nuf
  genr dum{!num}=1
  genr dut{!num}=t-!num+!nui
NEXT

```

Se realizan las regresiones de la hipótesis alternativa de Zivot para los tres modelos (A, B y C), diferenciando dos casos $k=0$ y $k>0$.

```

IF !regfin=0 then

  smpl !nui !nuf
  ls dlserie c lserie(-1) t dut{!num} dum{!num}
  smpl !num !num
  genr zivot=(c(2)/sqr(@covariance(2,2)))
  smpl !nui !nuf
  ls dlserie c lserie(-1) t dut{!num}
  smpl !num !num
  genr zivott=(c(2)/sqr(@covariance(2,2)))
  smpl !nui !nuf
  ls dlserie c lserie(-1) t dum{!num}
  smpl !num !num
  genr zivotm=(c(2)/sqr(@covariance(2,2)))

```

ENDIF

```

IF !regfin<>0 then

  smpl !nui !nuf
  ls dlserie c lserie(-1) dlserie(-1 to !regfin) t dut{!num} dum{!num}
  smpl !num !num
  genr zivot=(c(2)/sqr(@covariance(2,2)))
  smpl !nui !nuf
  ls dlserie c lserie(-1) dlserie(-1 to !regfin) t dut{!num}
  smpl !num !num
  genr zivott=(c(2)/sqr(@covariance(2,2)))
  smpl !nui !nuf
  ls dlserie c lserie(-1) dlserie(-1 to !regfin) t dum{!num}
  smpl !num !num
  genr zivotm=(c(2)/sqr(@covariance(2,2)))

```

ENDIF

NEXT

Se plotea los valores estadísticos de α , junto con los valores críticos, de manera que visualmente se determine si el menor valor de α cae por debajo de los valores críticos [es decir, es menor], con lo cual se aceptaría la hipótesis alternativa. Se utilizan los valores críticos al 5% de confianza.

```

smpl !nui !nuf
genr vcrit=-5.08
genr vcritm=-4.8
genr vcritt=-4.42

group ziv zivott zivotm zivot vcritt vcritm vcrit

```

plot ziv

Se plotea la serie de estadísticos F de la regresión con quiebre en media, tendencia o ambos para cada año y permite la deducción gráfica de donde se encuentra el mayor estadístico F, en quiebre en media, en tendencia o en ambos, de manera que se pueda decidir los resultados de que modelo del test de Z&A tomar en cuenta.

```

!bestf=0
!bestft=0
!bestfm=0

FOR !num=!cotal to !cotau
  smpl !nui !nuf
  ls lserie c t dut{!num} dum{!num}
  smpl !num !num
  genr f=@f
  !f=@f
  if !f>!bestf then
    !bestf=!f
    !fecha=!num
  endif
  smpl !nui !nuf
  ls lserie c t dut{!num}
  smpl !num !num
  genr ft=@f
  !ft=@f
  if !ft>!bestft then
    !bestft=!ft
    !fechat=!num
  endif
  smpl !nui !nuf
  ls lserie c t dum{!num}
  smpl !num !num
  genr fm=@f
  !fm=@f
  if !fm>!bestfm then
    !bestfm=!fm
    !fecham=!num
  endif

```

NEXT

```

smpl !nui !nuf
genr fecha=!fecha
genr fechat=!fechat
genr fecham=!fecham
group fstat f ft fm
plot fstat

```

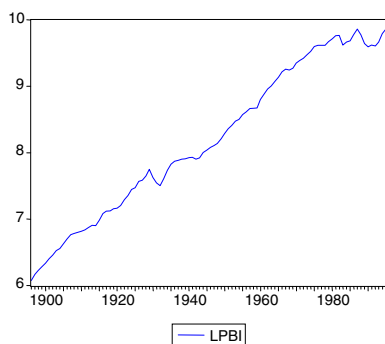
Fin del programa.

APENDICE 16.2

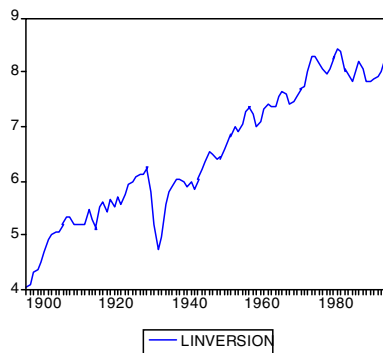
Equivalencia Observacional

Como el lector debe haber notado, las series económicas presentan una diversidad de trayectorias. Por su variedad, no todas ellas presentan una forma funcional evidente, correspondiente a alguna de las especificaciones conocidas hasta el momento. Al contrario, una serie puede presentar varias “posibles” especificaciones, a pesar de que sólo una de ellas es la correcta.

Se puede pensar que la simple vista gráfica de la serie puede eliminar cualquier duda al respecto, de manera que visualmente uno pueda diferenciar entre una serie que presenta raíz unitaria y otra que es estacionaria con quiebre en media, por ejemplo. Como ejercicio, tratemos de inferir acerca de la forma funcional de las siguientes series :



(a)



(b)

Gráfico 1

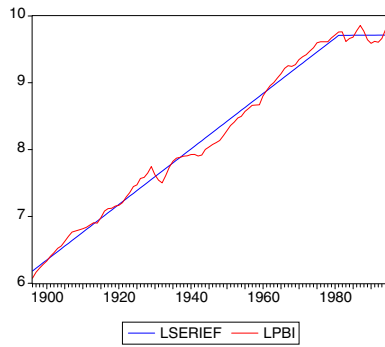
La serie (a) es el logaritmo del PBI peruano entre 1896 y 1995 y la serie (b) es el logaritmo la Inversión peruana en el mismo periodo.

El PBI puede ser modelado de dos maneras distintas : como una raíz unitaria con constante o como una serie estacionaria en tendencia con un quiebre en la tendencia en el año 1981.

La Inversión puede ser modelada como una raíz unitaria con constante o como una serie estacionaria en tendencia con un quiebre en la constante en el año 1929 (Gráfico 2).

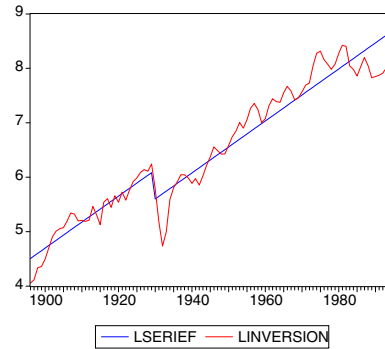
La razón por la que existe cierta duda en estas series es porque visualmente equivalen a un proceso u a otro. Este problema se conoce como *Equivalencia Observacional*.

Gráfico 2



(a)

¿Raíz unitaria con constante
o estacionario en tendencia con
quiebre en tendencia (1981)?



(b)

¿Raíz unitaria con constante
o estacionario en tendencia con
quiebre en constante (1929)?

La búsqueda de la verdadera especificación de las series económicas, debido en parte a sus implicancias en la formulación de políticas, ha llevado a que se desarrollen tests que discriminan entre dos alternativas de modelación de la serie. Así tenemos los tests de Dickey Fuller y de Perron que discriminan entre procesos con raíz unitaria y series estacionarias (con constante y/o tendencia). Para diferenciar entre series que presentan raíz unitaria con un cambio en la estructura en un periodo fijo y series estacionarias un cambio estructural en media y/o tendencia están los tests de Perron, Perron y Volgensang, y Zivot y Andrews. Como es de esperarse, las investigaciones continúan al respecto.