Econometría I. 3º LADE
Tema 3. Regresores estocásticos no estacionarios

1 Introducción

Todos los procesos analizados anteriormente fueron estudiados bajo la hipótesis de estacionariedad. Sin embargo, la mayoría de las variables utilizadas en estudios de series temporales no presentan esta propiedad, sino que se caracterizan mejor como procesos integrados. Especialmente ilustrativo en este aspecto fue el trabajo de Nelson y Plosser (1982) donde se demuestra que variables como Producto Interior Bruto, el Consumo o la Inversión de los USA se pueden considerar como integrados. Este resultado puede extenderse a la mayoría de los países desarrollados. Para el caso español también ocurre lo mismo, como se puede apreciar en el Grá co 4.1. En el se presenta la evolución del PIB español a precios de mercado en pesetas constantes del año 1990. Se observa en este grá co la trayectoria creciente de esta variable en el periodo XXX-XXX. Por tanto, no se puede aceptar que esta variable tenga media constante a lo largo de todo el periodo, lo que hace que la hipótesis de no estacionariedad no tenga gran soporte desde el punto de vista empírico.

Al margen de la propia evidencia empírica, desde el punto de vista de la Teoría económica también es coherente la no estacionaridad de este tipo de variables. Debemos pensar que si la mayoría de modelos macroeconómicos admiten que las variables pueden crecer a lo largo del tiempo. Si además tenemos en cuenta que en la mayoría de los modelos asumen que este crecimiento es constante a lo largo del tiempo, estamos diciendo que estas variable pueden representarse de forma adecuada como procesos integrados de primer orden.

No obstante, en otras ocasiones la no estacionariedad no tiene tanta argumentación desde el punto de vista económico. Ejemplos de ellos pueden ser el tipo de cambio real entre dos monedas o el tipo de interés real de una economía. Ambas magnitudes se consideran estacionarias en la mayoría de los modelos teóricos.

Un ejemplo de proceso no estacionario es el denominado paseo aleatorio con deriva. Este modelo se de ne como sigue:

$$y_t = \delta + y_{t-1+} u_t$$

Sus propiedades fueron estudiadas en el tema anterior, donde vimos que los momentos dependen de t. Por tanto, no es un proceso estacionario. Este proceso se puede generalizar en un modelo $ARIMA(p, d, q)$. Este proceso se representa de la siguiente manera:

$$\left(1 - \phi_1 L - ... - \phi_p L^p\right)(1-L)^d\, y_t = \delta + \left(1 + \theta_1 L + ... + \theta_q L^p\right) u_t$$

donde p y q son parámetros que nos son conocidos y d es el parámetro que determina el orden de integración de la variable. Una forma alternativa de presentar el modelo anterior es la siguiente:

$$\left(1 - \phi_1 L - ... - \phi_p L^p\right) z_t = \delta + \left(1 + \theta_1 L + ... + \theta_q L^p\right) u_t$$

donde $z_t = (1-L)^d y_t$. Por tanto, observamos que la variable su cientemente diferenciada sigue un proceso ARMA(p,q). No obstante, debemos tener en cuenta que esto es sólo una forma de interpretar el modelo anterior ya que la variable que estamos analizando es y_t y no z_t.

Si consideramos que existe un componente estacional, entonces podemos hablar de procesos multiplicativos $ARIMA(p,d,q) \times ARIMA_s(P,D,Q)$

$$\Phi_P(L^s)\,\phi_p(L)\,(1-L)^d(1-L)^d\, y_t = \delta + \Theta_Q(L^s)\theta_q(L^s)\, u_t$$

donde D es el orden de integración de la parte estacional que no tiene por qué coincidir con d. Dado que el proceso no es estacionario, debemos en primer lugar conseguir que sea estacionario diferenciado cuantas veces haga falta. A continuación, será posible determinar el valor de los parámetro p, q, P y Q. En la siguientes secciones vamos a discutir los diversos métodos que existen para determinar el orden de integración de una variable.

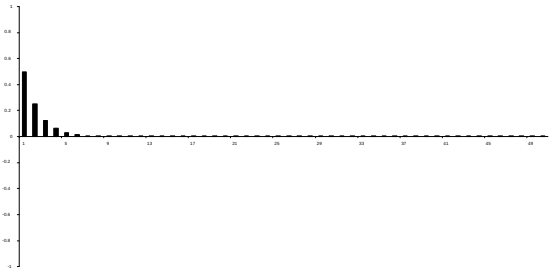
2 Métodos no parámetricos para determinar el orden de integración de un proceso

El orden de integración se puede determinar de diversas formas. Una de ella es el uso de métodos no paramétricos. Estos métodos tienen la ventaja de que no es necesario formular ninguna hipótesis ni calcular la distribución de un estadístico bajo dicha hipótesis. Por contra, suelen ser bastante subjetivos y no tan ables como los métodos paramétricos. En cualquier caso, siempre pueden darnos pistas sobre el orden de integración de la variable. En esta sección vamos a considerar tres métodos alternativos: estudio de la función de autocorrelación, sobrediferenciacón de la serie y estudio de la varianza para diversos órdenes de integración. No es conveniente utilizarlos de forma separada, sino que lo más habitual es usarlos conjuntamente. En lo que sigue vamos a considerar que el proceso no tiene parte estacional, aunque los resultados que vamos a obtener son directamente trasladables al caso estacional.

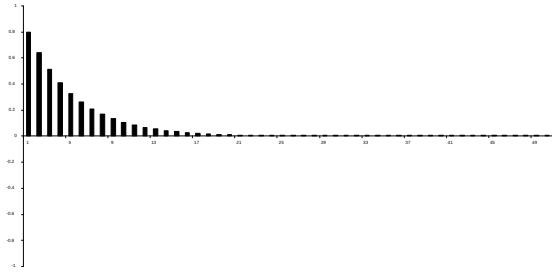
2.1 Análisis de la función de autocorrelación

Supongamos que tenemos un proceso autorregresivo de orden 1. Su función de autocorrelación es igual a $\rho(k) = \phi_1^k$, $\forall k = 1, 2, ...$ Esta función de autocorrelación decae hacia 0, pero sólo se anula en el in nito. Entonces, el decaimiento que se observa en la función de autocorrelación de la variable depende del valor del parámetro ϕ_1. Si este toma, por ejemplo, un valor igual a 0.5, entonces para k=5 el coe ciente de la función de autocorrelación es igual a 0.03. Por contra, si $\phi_1 = 0.8$, $\rho(5) = 0.33$, mientras que la función de autocorrelación toma el valor 0.03 cuando $k = 16$. Esto conlleva que cuanto mayor es el valor del parámetro autorregresivo más lenta es la convergencia hacia 0 de la función de autocorrelación. Si consideramos el caso límite, $\phi_1 = 1$, en el que el proceso no es estacionario, la función de autocorrelación debería expresar un decaimiento muy lento hacia 0, con abundantes valores distintos de 0.

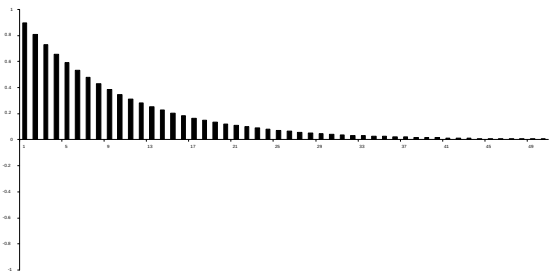
Grá co 4.1. FAC de un AR(1)



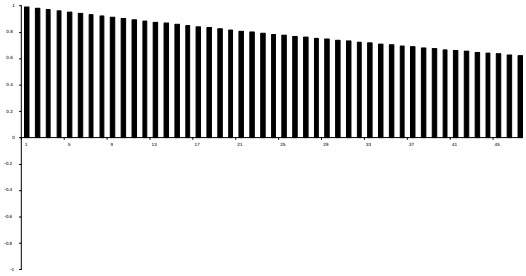
ϕ_1=0.5.



ϕ_1=0.8.



ϕ_1=0.9.



ϕ_1=0.99

En el grá co 4.1 se observa este fenómeno. A medida que se incrementa el

valor del parámetro ϕ_1 el valor del primer coe ciente de se aproxima hacia la unidad y decaimiento de la función de autocorrelación es más lento. El caso extremo es $\phi_1 = 0.99$, donde inlcuso para k=50 la función de autocorrelación toma valores alejados de 0.

El problema que presenta este procedimiento es su subjetividad. No queda claro cuando podemos considerar que el primer valor del coe ciente de autocorrelación está su cientemente próximo a la unidad, ¿basta con 0.7 o tiene que ser igual a 0.9999?, cuando el valor está próximo a 0, ni cuanto valores de la función de autocorrelación deben ser distintos de 0 para considerar que el decaimiento hacia 0 es lento.

2.2 Sobrediferenciación

La justi cación de este método es la siguiente. Supongamos que tenemos un proceso estacionario, un ruido blanco por ejemplo: $y_t = u_t$. Si tomamos primeras diferencias, entonces el modelo se convierte en:

$$(1 - L)\ y_t = (1 - L)\ u_t$$

Esto lo podemos interpretar diciendo que la variable Δy_t sigue un proceso MA(1) no invertible con parámetro $\theta_1 = -1$. Esta no invertibilidad nos está indicado que el proceso está sobre diferenciado. Si tomamos en consideración ahora la forma que adopta la función de autocorrelación de un MA(1), es directo comprobar que el primer valor de esta función será igual a:

$$\rho(1) = \frac{\theta_1}{1 + \theta_1^2} = \frac{-1}{1 + 1} = -0.5$$

y el resto de los valores serán iguales a 0. A partir de este resultado particular, podemos sacar la conclusión para el caso general y decir que si la serie que estamos analizando está sobrediferenciada, su función de autocorrelación exhibe un comportamiento de medias móviles con su primer valor de la función de autocorrelación próximo a -0.5.

Entonces, el proceso a seguir es el siguiente. En primer lugar debemos estudiar la función de autocorrelación de la serie en niveles. Si tenemos dudas sobre el decaimiento lento hacia 0, entonces debemos calcular la función de autocorrelación para la variable diferenciada. Si ésta presenta un comportamiento similar al de un MA(1) con su primer coe ciente próximo a -0.5 podemos concluir que está sobrediferenciada y concluir que la variable es estacionaria en niveles. En caso contrario, es posible concluir que la serie es integrada.

2.3 Estudio de la varianza

Un método asociado al caso anterior es el estudio de la varianza de la variable para diversos órdenes de integración La justi cación del método es la siguiente. Supongamos que la variable está generada por un paseo aleatorio $y_t = y_{t-1} + u_t$. La varianza de la variable y_t es igual a $Var(y_t) = t\,\sigma^2$. POr tanto, es una varianza que tiene hacia in nito. Si tomamos primeras diferencias, el modelo anterior nos queda de la siguiente manera:

$$y_t = y_{t-1} + u_t \Rightarrow \Delta y_t = u_t$$

Por tanto, es inmediato comprobar que la varianza de la variable en primeras diferencias es $Var(\Delta y_t) = \sigma^2$. Por último, si sobrediferenciamos la variable anterior, el modelo queda así:

$$\Delta^2 y_t = \Delta u_t \Rightarrow \Delta^2 y_t = u_t - u_{t-1}$$

De lo que se desprende que $Var(\Delta y_t) = Var(u_t - u_{t-1}) = Var(u_t) + Var(u_{t-1}) = 2\,\sigma^2$.

En virtud de estos resultados, el procedimiento a seguir sería obtener la varianza de la variable que estamos analizando para diversos valores del parámetro d. El orden de integración seleccionado sería aquel para el que la varianza de la variable se minimiza.

3 Contrastes de raíz unitaria

Los métodos anteriores nos pueden proporcionar información útil sobre cuál es el orden de integración de la variable que estamos analizando. No obstante, en la mayoría de las ocasiones no son conclusiones. Incluso, pueden llevarnos a tomar decisiones erróneas. Por ello, resulta aconsejable utilizar otro tipo de métodos que nos indiquen de forma más fehaciente cuál es el verdadero orden de integración de la variable que estamos analizando. Estos métodos están basados en un contraste de hipótesis. Esto supone que es necesario tomar una hipótesis nula, diseñar el estadístico de prueba y calcular su distribución bajo la hipótesis nula y, con posterioridad, comparar el valor del estadístico en nuestra muestra con los valores teóricos de la distribución.

El primer problema que se nos presenta es la elección de la hipótesis nula. La familia de estadísticos que vamos a estudiar en este apartado considera la hipótesis nula de raíz unitaria. Debemos recordar que si el polinomio autorregresivo de retardos presenta una raíz unitaria, entonces el proceso no es estacionario ya que al menos una raíz no está fuera del círculo unidad. Por

tanto, bajo la hipótesis nula, la variable es integrada. Entonces, si aceptamos la hipótesis nula, podemos decir que la variable no es estacionaria. Sólo si rechazamos esta hipótesis podemos admitir que existe evidencia en contra de esta hipótesis y, por tanto, considerar que la variable es estacionaria. A continuación vamos a analizar diversos estadísticos que estudian la hipótesis nula de raíz unitaria.

3.1 Caso A. No autocorrelación

El estadístico más sencillo es el propuesto en Dickey y Fuller (1979). Supongamos que tenemos una variable que sigue un proceso AR(1). Este modelo lo podemos expresar como sigue:

$$\begin{aligned} (1-\rho\,L)\,y_t &= u_t \Rightarrow \\ y_t &= \rho\,y_{t-1} + u_t \end{aligned} \tag{1}$$

De acuerdo a lo que vimos en apartados precedentes si ρ es igual a 1, entonces el proceso no es estacionario, mientras que si $|\rho| < 1$ entonces la variable es estacionario. El estudio de la hipótesis nula de estacionariedad resulta complicado, por cuanto tan estacionaria es la serie cuando $\rho = 0.5$, como cuando $\rho = 0.9$. POr tanto, no resulta sencillo determinar cuando hay evidencia su ciente para considerar que la variable es estacionaria. Por contra, más sencillo resulta analizar cuando la variable no es estacionaria ya que, eliminando los casos explosivos poco probables desde el punto de vista económico, la variable no es estacionaria cuando $\rho = 1$. Esto supone que bajo la hipótesis nula la variable es integrada, no estacionaria. El contraste de esta hipótesis aparentemente es sencillo por cuanto es su ciente estimar el modelo (1) y, con posterioridad, obtener calcular el valor del t-ratio

$$\tau = \frac{\hat{\rho} - 1}{\hat{\sigma}_{\hat{\rho}}}$$

No obstante, esto tiene una di cultad y es que, dado que bajo la hipótesis nula la variable no es estacionaria, no es cierto que el estadístico t converga hacia una distribución normal. Al contrario, este estadístico converge hacia la siguiente distribución:

$$\tau = \frac{\hat{\rho} - 1}{\hat{\sigma}_{\hat{\rho}}} \Rightarrow \frac{\int_0^1 W(r)\; dW(r)}{\sqrt{\int_0^1 W(r)^2\; dr}}$$

donde W(r) es un proceso de Wiener. La distribución anterior es actualmente bien conocida y esta tabulada en Fuller (1976) para algunos tamaños muestrales. Con posterioridad, algunos autores han estimado funciones de respuesta para obtener los valores críticos de esta distribución para cualquier tipo de tamaño muestral (Ver Davidson or Cheung).

La especi cación del modelo () no recoge elmente el comportamiento de la mayoría de las variables macroeconómicas. Debemos recordar que la mayoría de ellas presenta un componente tendencial o, al menos, se admite la presencia de un término independiente en el modelo. En este tipo de casos, el estudio de la hipótesis nula de raíz unitaria a partir del modelo () nos llevaría a tomar decisiones de forma errónea. Un ejemplo de cómo pueden afectar estos errores de especi cación en los contrastes de raíz unitaria aparece en Clemente et al. (1995), donde se demuestra que es posible la aparición de raíces explosiva cuya naturaleza es meramente espúrea.

Para corregir este supuesto, podemos utilizar las especi caciones alternativas propuestas en el trabajo original de Dickey y Fuller (1976). Estas son las siguientes:

$$y_t = \mu + \rho\, y_{t-1} + u_t \tag{2}$$
$$y_t = \mu + \beta\, t + \rho\, y_{t-1} + u_t \tag{3}$$

A partir de aquí podemos de nir los estadísticos tm y tt. Ambos se obtienen de la misma manera. El estadístico tm se de ne como:

$$\tau_\mu = \frac{\hat{\rho} - 1}{\hat{\sigma}_{\hat{\rho}}}$$

donde $\hat{\rho}$ y $\hat{\sigma}_{\hat{\rho}}$ se obtienen mediante la estimación del modelo (2). De forma similar, tt se de ne como:

$$\tau_t = \frac{\hat{\rho} - 1}{\hat{\sigma}_{\hat{\rho}}}$$

pero en esta ocasión $\hat{\rho}$ y $\hat{\sigma}_{\hat{\rho}}$ son el resultado de estimar del modleo (3). Al igual que ocurre con t, ninguno de estos dos estadísticos sigue una distribución con la que hayamos trabajado previamente. Por contra, ambos estadísticos convergen hacia las siguientes distribuciones:

$$\tau_\mu \Rightarrow \frac{\int_0^1 W^*(r)\; dW^*(r)}{\sqrt{\int_0^1 W^*(r)^2\; dr}}$$

$$\tau_\tau \;\Rightarrow\; \frac{\int_0^1 \tilde{W}(r)\; d\tilde{W}(r)}{\sqrt{\int_0^1 \tilde{W}(r)^2\; dr}}$$

donde W^* y $\tilde{W}$ representan, respectivamente, la proyección de un proceso de Wiener W sobre el espacio {1}y {1,t}. Esto quiere decir que son procesos de Wiener a los que se les ha eliminado el término independiete y, en el segundo caso, la tendencia determinística. Estas dos distribuciones están tabuladas en Fuller (1976). Davidson y Cheung también proporcionan super cies de respuesta que permiten obtener los valores críticos para distintos tamaños muestrales.

3.2 Caso B. Autocorrelación

Los estadísticos presentados en el apartado anterior son válidos siempre que la variable presente exclusivamente un comportamiento AR(1). Sin embargo, no siempre esto va a ocurrir. Es posible que las variables presenten esquemas de autocorrelación diferente. En ese caso, los estadísticos anteriores no son válidos. Para comprender mejor los problemas a los que nos enfrentamos, consideremos quie la variable yt viene generada por un proceso AR(2). Entonces, yt se de ne de la siguiente manera:

$$\begin{aligned}(1-\phi_1 L-\phi_2 L^2)\, y_t &= v_t \;\Rightarrow \\ y_t &= \phi_1 y_{t-1}+\phi_2 y_{t-2}+v_t \;\Rightarrow \\ y_t &= \rho\, y_{t-1}+u_t \qquad (4)\end{aligned}$$

donde ut= $\phi_2\, y_{t-2}+v_t$. La estimación del modelo anterior presentaría problemas de autocorrelación, por cuanto es fácil demostrar que E(ut ut-2)$\neq$ 0, por lo que se incumple una hipótesis básicas del modelo. En consecuencia, no es posible realizar inferencia. Además, los valores críticos tabulados en Fuller (1976) son válidos sólo para el caso de no autocorrelación. En consecuencia, es necesario llevar a cabo un procedimiento alternativo para poder contrastar la hipótesis nula de raíz unitaria. Para ello, vamos a realizar una pequeña transformación del modelo anterior de tal forma que queda de la siguiente manera:

$$\begin{aligned}y_t &= \phi_1 y_{t-1}+\phi_2 y_{t-2}+v_t \\ &= \phi_1 y_{t-1}+\phi_2 y_{t-2}+\phi_2\Delta y_{t-1}-\phi_2\Delta y_{t-1}+v_t\end{aligned}$$

$$
\begin{aligned}
&= (\phi_1 + \phi_2)\, y_{t-1} + (-\phi_2)\, \Delta y_{t-1} + v_t \;\Rightarrow \\
y_t &= \rho\, y_{t-1} + \varphi_1\, \Delta y_{t-1} + v_t \qquad (5)
\end{aligned}
$$

Entonces, la variable yt es integrada siempre que $\rho = 1$. Entonces, para contrastar la hipótesis nula de raíz unitaria se puede calcular el estadístico

$$\tau = \frac{\hat{\rho} - 1}{\hat{\sigma}_{\hat{\rho}}}$$

donde $\hat{\rho}$ y $\hat{\sigma}_{\hat{\rho}}$ se obtienen a partir de la estimación mínimo cuadrático ordinaria del modelo (5) . Said y Dickey (1984) demuestran que la distribución asíntótica de este estadístico coincide con la expresada en ().

Este resultado lo podemos extender al caso más general. Así, supongamos que yt viene dado por la siguiente relación:

$$A\,(L)\, y_t = B(L)\, e_t.$$

donde $A(L)$ es un polinomio de retardos de orden p que puede presentar una raíz unitaria. Además, también se cumple que A(1) = 0. Este polinomio se puede representar también de esta forma $A(L) = (1 - L)A^*(L)$ donde $A^*(L)$ es un polinomio de orden $p - 1$ que tiene todas sus raíces fuera del círculo unidad. Por ejemplo, si consideramos que el polinomio $A\,(L)$ es igual a $1 - 1.5L + 0.5L^2$, lo podemos expresar como producto de dos polinomios $(1 - L) \times (1 - 0.5L)$, por lo que $A^*(L)$ sería en este caso $(1 - 0.5L)$.

Del mismo modo, $B(L)$ es un polinomio de retardos de orden q cuyas raíces están todas fuera del círculo unidad. De esta manera, excluimos el caso de procesos no invertibles. Si dividimos a ambos lados de la igualdad por $B(L)$, el modelo anterior queda de la siguiente manera:

$$\frac{A\,(L)}{B(L)}\, y_t = \Psi(L)\, y_t = (1 - L)\Psi^*(L)\, e_t.$$

donde $\Psi(L) = \frac{A(L)}{B(L)} = (1 + \psi_1 L + \psi_2 L + ...)$ es un polinomio de un orden posiblemente in nito. Será de orden nito siempre y cuando B(L) = 1, en cuyo caso la variable no presenta componente de medias móviles. En otro caso, al tener parte de medias móviles, el proceso lo podemos representar mediante un AR(∞), de ahí que el orden de $\Psi(L)$ pueda ser in nito. Además, al contener una raíz unitaria, entonces se cumple que $\Psi(1) = 0$. Por lo tanto, la suma de los in nitos coe cientes de este polinomio debe ser igual a la unidad. Esto lo podemos utilizar para llevar a cabo el contraste, ya que transformando el modelo anterior tenemos que:

$$\begin{aligned}\Psi(L)\, y_t &= (1+\psi_1 L+\psi_2 L+....)\, y_t = e_t \Rightarrow \\ y_t &= \sum_{i=1}^{\infty} \psi_i\, y_{t-i} + e_t \\ &= \rho\, y_{t-1} + \Sigma_{j=2}^{p} \varphi_j \Delta y_{t-j+1} + e_t\end{aligned}$$

donde debemos tener en cuenta que

$$\rho = \sum_{i=1}^{\infty} \psi_i$$

y

$$\varphi_i = -. \sum_{j=i}^{\infty} \psi_j$$

El problema que se nos presenta ahora es el que los polinomios son de orden in nito por lo que este modelo no es aplicable en la práctica. No obstante, Said y Dickey demuestran que si elegimos un k su cientemente grande, entonces podemos estimar el siguiente modelo

$$y_t = \rho\, y_{t-1} + \Sigma_{j=2}^{\infty} \varphi_j \Delta y_{t-j+1} + e_t$$

Entonces, el contrate de la hipótesis nula de raíz unitaria es sencillo ya que basta con estimar este modelo por mínimos cuadrados ordinarios para, a continuación, calcular el estadístico t, de nido como

Si, como es habitual, consideramos que la especi cación anterior no es su ciente de cara a recoger los elementos determinístico de la variable yt, entonces podemos considerar los modelos alternativos

$$y_t = \mu + \rho\, y_{t-1} + \Sigma_{j=2}^{\infty} \varphi_j \Delta y_{t-j+1} + e_t$$

$$y_t = \mu + \beta\, t + \rho\, y_{t-1} + \Sigma_{j=2}^{\infty} \varphi_j \Delta y_{t-j+1} + e_t$$

4 Contrastes de Phillips-Perron

Una propuesta alternativa a los estadísticos de la familia Dickey-Fuller son los propuestos en Phillips y Perron (1988). La intuición de estos estadísticos es la siguiente. Supongamos que la variable es generada por el siguiente modelo: $y_t = y_{t-1} + u_t$, donde ahora ut no es un ruido blanco sino que cumple una serie de supuestos que permiten la existencia de cierto grado de autocorrelación

y heteroscedasticidad (ver Phillips, 1987). Estas condiciones permiten que la perturbación del modelo haya sido generada por los modelos ARIMA. En estas circunstancias, si estimamos el modelo (), es sencillo demostrar que el pseudo t-ratio converge hacia la siguiente distribución:

$$\tau \Rightarrow \frac{\sigma_u \int_0^1 W(r)\ dW(r)}{\sigma \sqrt{\int_0^1 W(r)^2\ dr}}$$

siendo $\sigma_u^2 = \lim_{T\to\infty} T^{-1} \sum_{i=1}^{T} E\left(e_i^2\right)$ mientras que $\sigma^2 = \lim_{T\to\infty} T^{-1} E\left[\left(\sum_{i=1}^{T} e_i\right)^2\right]$. Ese último elemento representa la varianza de largo plazo de la variable y está íntimamente relacionado con la frecuencia espectral en 0.

Phillips y Perron (1988) propone corregir no paramérticamente el pseudo t-ratio para que el estadístico resultante converga hacia la distribución tabulada en Fuller. Para ello proponen los siguientes estadísticos:

a) Para el caso del modelo (1)

$$Z_t = \frac{s_e}{s}\tau - \frac{1}{2}\frac{s^2 - s_e^2}{s\sqrt{T^{-2}\sum_{i=1}^{T} y_{t-1}^2}}$$

b) para el caso del modelo (2)

$$Z_t = \frac{s_e}{s}\tau - \frac{1}{2}\frac{s^2 - s_e^2}{s\sqrt{T^{-2}\sum_{i=1}^{T} \left(y_{t-1} - \bar{y}_{t-1}\right)^2}}$$

donde $\bar{y}_{t-1} = T^{-1} \sum_{i=1}^{T-1} y_t$

c) Para el caso del modelo (3)

$$Z_t = \frac{s_e}{s}\tau - \frac{1}{2}\frac{T^3\left(s^2 - s_e^2\right)}{4\, s\, D_X\sqrt{3}}$$

donde $\mathrm{D}_X = Det(X'X)$, con X=$\{1, t, y_{t-1}\}$. El estadístico $s_e^2 = T^{-1} \sum_{i=1}^{T} \hat{e}_i^2$, siendo $\sum_{i=1}^{T} \hat{e}_i^2$ la suma residual de cada uno de los anteriores modelos. POr último, s^2 es un estimador consistente de σ^2. Es este punto tenemos varias alternativas. La primera, defendida por los propios autores en su trabajo original, es usar el estimador de Newey-West

$$s^2_{NW} = T^{-1}\sum_{i=1}^{T}\hat{e}_i^2 + 2\,T^{-1}\sum_{j=1}^{\ell}\omega_{j\ell}\sum_{i=j+1}^{T}\hat{e}_t\ \hat{e}_{t-j}$$

donde $\omega_{j\ell} = 1 - \frac{j}{\ell+1}$. Sin embargo, existen otras posibilidades, como el uso de las ventanas de Parzen o cuadrático espectral.

La elección del parámetro ℓ resulta muy importante de cara al uso de estos estadísticos. Una primera posibilidad es usar el valor $\sqrt[4]{T}$. Otra, más adecuada, es usar métodos para determinar automáticamente este valor.

5 Orden de integración en variables estacionales

Hasta el momento hemos supuesto que las variables sometidas a estudio no presentan componente estacional. En esos casos, los estadísticos presentados en el apartado anterior serían los que deberíamos utilizar para determinar el orden de integración de una variable. Sin embargo, no siempre esto va a ser así. Es bastante frecuente que se dispongan de datos de frecuencia inferior al año en el estudio aplicado. En estos casos debemos estudiar no sólo el orden de integración de la parte regular, sino también el orden de integración del componente estacional. El objetivo de este tema es presentar los diferentes estadísticos que nos permiten estudiar el orden de integración en presencia de estacionalidad. Vamos a considerar dos tipos de estadísticos. El primero es una extensión de los contrastes de la familia de Dickey-Fuller al caso estacional. El segundo responde a una losofía diferente y analiza la presencia de raíces unitarias para las diversas frecuencias.

A largo de los siguientes apartados vamos a considerar que la variable analizada viene generada por (1-ls)yt = ut, donde s es el parámetro que nos da la frecuencia con la que se obtienen los datos. Podemos considerar que s=4, por lo que los datos son de tipo trimestral. La extensión a otros tipos de frecuencias es inmediato.

5.1 Contraste de Hasza-Fuller

Este estadístico pertenece a la familia de estadísticos de Dickey-Fuller. Para su obtención debemos estimar el siguiente modelo:

$$y_t = \sum_{i=1}^{s} d_i \; S_{it} + \; \rho_s \, y_{t-s} + u_t \tag{6}$$

donde S_{it} son unas variables dummies de tipo estacional. Entonces, el estadístico de Hasza-Fuller se de ne de forma análoga al resto de los estadísticos de esta familia:

$$\tau_{HF} = \frac{\hat{\rho}_s - 1}{\hat{\sigma}_{\hat{\rho}_s}}$$

donde $\hat{\rho}_s$ y $\hat{\sigma}_{\hat{\rho}_s}$ proceden de la estimación mínimo cuadrático ordinaria de (6). Como es obvio este estadístico sigue distribución que depende de procesos de Wiener, aunque en este caso son de tipo estacional.

La extensión a casos en los que se incluyen comportamientos determinísticos es inmediata. Así, podemos considerar los siguientes modelos alternativos:

$$y_t = \; \rho_s \, y_{t-s} + u_t \tag{7}$$

$$y_t = \; \rho_s \, y_{t-s} + u_t \tag{8}$$

5.2 Contraste HEGY

Los estadísticos anteriores tienen el grave problema que no son capaces de distinguir entre aquellas frecuencias que presentan un componente estacionaria de las que no es estacionario. Para entender este punto, tomemos como caso de estudio s=4. Entonces, se demuestra fácilmente que el poliómio (1-L4) puede expresarse como sigue

$$1 - L^4 = (1 - L) \; (1 + L) \, (1 + L^2)$$

lo que supone que el polinomio $1 - L^4$ admite la presencia de 4 raíces unitarias distintas. Es evidente que no todas ellas están relacionadas con la frecuencia estacional. Así, el factor $(1 - L)$ está claramente vinculado a la presencia de una raíz unitaria en la parte regular de la variable. El resto de los polinomios sí están relacionados con raíces estacionales. $(1 + L)$ implica la presencia de una raíz unitaria en la frecuenciam bi-anual (dos ciclos por año), mientras que $(1 + L^2)$ implica una raíz unitaria en la frecuencia anuak (un sólo ciclo por año).

Este mismo esquema se puede extender a otro tipo de frecuencias. Por ejemplo, en el caso mensual tenemos que:

$$
\begin{aligned}
1-L^{12} &= (1-L)(1+L)(1-iL)(1+iL) \\
&\times \left[1+\frac{\left(i+\sqrt{3}\right)L}{2}\right] \left[1+\frac{\left(\sqrt{3}-i\right)L}{2}\right] \\
&\times \left[1-\frac{\left(i+\sqrt{3}\right)L}{2}\right] \left[1-\frac{\left(\sqrt{3}-i\right)L}{2}\right] \\
&\times \left[1+\frac{\left(i+\sqrt{3}\right)L}{2}\right] \left[1-\frac{\left(\sqrt{3}-i\right)L}{2}\right] \\
&\times \left[1-\frac{\left(i+\sqrt{3}\right)L}{2}\right] \left[1+\frac{\left(\sqrt{3}-i\right)L}{2}\right]
\end{aligned}
$$

donde el polinomio original se trasnforma en producto de diversos polinomios.

Para comprobar la presencia de alguna o varias raíces unitarias en la frecuencia estacional, el procedimiento a seguir es el siguiene. Tomemos como caso de referencia el caso trimestral. Entonces, Hylleberg et al. (1990) proponen llevar a cabo la estimación del siguiente modelo:

$$
y_{4t} = \sum_{i=1}^{4} \mu_i S_{it} + \gamma t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \epsilon_t, \qquad (9)
$$

donde $y_{1t} = (1+L+L^2+L^3)y_t$, $y_{2t} = (-1+L-L^2+L^3)y_t$, $y_{3t} = (-1+L^2)y_t$ y $y_{4t} = \Delta_4 y_t = y_t - y_{t-4}$. La inclusión de una tendencia determinística depende del tipo de variable, por lo que es opcional. NO ocurre lo mismo con la inclusión de las variables dummies, a pesar de que en el trabajo originial de HEGY estas no se incluyen. SIn embargo, estudios de MOnte Carlo de XXX indican la convenencia de su inclusión. Una vez que hemos estimado este modelo por MCO podemos realizar los siguientes contrastes

H_o	estadístico
$\pi_1 = 0$	$t_{\pi_1} = \frac{\hat{\pi}_1}{\hat{\sigma}_{\hat{\pi}_1}}$
$\pi_2 = 0$	$t_{\pi_2} = \frac{\hat{\pi}_2}{\hat{\sigma}_{\hat{\pi}_2}}$
$\pi_3 = 0$	$t_{\pi_3} = \frac{\hat{\pi}_3}{\hat{\sigma}_{\hat{\pi}_3}}$
$\pi_4 = 0$	$t_{\pi_4} = \frac{\hat{\pi}_4}{\hat{\sigma}_{\hat{\pi}_4}}$
$\pi_3 = \pi_4 = 0$	$F_{34} = t_{\pi_3}\ t_{\pi_4}$
$\pi_2 = \pi_3 = \pi_4 = 0$	
$\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$	

6 Persistencia

Un concepto muy relacionado con el de raíz unitaria o con el de integración es el de persistencia. Para comprender este concepto resulta conveniente retomar un paseo aleatorio. Como ha quedado demostrado con anterioridad, este proceso se puede de nir como suma de in nitos shocks. Entonces el efecto que un shock en el periodo i-ésimo tiene una duración in nita. Por contra, si el proceso generador de los datos es un modelo estacionario, el efecto de un shock tiene una duración nita, en su caso extremo, sólamente un periodo. Si además tenemos en cuenta que los contrastes de raíz unitaria tienen escasa potencia en muestras pequeñas, esto ha llevado a algunos autores a cuestionar los métodos utilizados con anteriodidad para determinar el orden de integración de una variable. Por contra, sugieren utilizar diversos estadísticos que nos den una medida de la in uencia que tiene un shock sobre el comportamiento de una variable. Cuanto mayor es el valor de este estadístico, mayor es la persistencia del shock y, por tanto, más cerca estamos de considerar que la variable es integrada. Al contrario, si los estadísticos toman un valor próximo a 0, entonces el shock tiene una in uencia limitada y, en consecuencia, la hipótesis de estacionariedad de la variable parece más verosímil.

Una medida del grado de persistencia de la variable es la proporcionada por Cochrane (1986). Este estadístico, denominado el ratio varianza (VR), se de ne como sigue:

$$VR_k = \frac{V_k}{V_1} = \frac{\frac{Var(y_t - y_{t-k})}{k}}{\frac{Var(y_t - y_{t-1})}{1}} = \frac{Var\left(y_t - y_{t-k}\right)}{k\,Var\left(y_t - y_{t-1}\right)}$$

Se puede demostrar que este estadístico es igual a:

$$VR_k = 1 + 2\sum_{i=1}^{k}\left(1 - \frac{i}{k+1}\right) r_i$$

donde r_i es el coe ciente i-estimo de la función de autocorrelación de la variable Δy_t. Dado que los componentes del estadístico son momentos poblacionales, no son conocidos, por lo que es necesario su estimación. Cochrane (1986) propone la siguiente estimación:

$$\hat{VR}_k = \frac{s_k^2}{k\,s_1^2}\,\frac{T}{T-k+1}$$

donde s_i^2 es la varianza muestral de la variable $y_t - y_{t-i}$. El segundo elemento de la expresión anterior es un factor de corrector para muestras

nitas. Se demuestra que cuando T, $k \to \infty$ y $k/T \to \infty$, entonces $\hat{VR}_k$ converge hacia una distribución normal de media VR_k y varianza $4k\frac{VR_k^2}{\cdot 3T}$.

En la práctica es habitual calcular el estadístico anterior para diversos valores de k, con un k que debe ser elevado. El problema es que cuando disponemos de pocas observaciones, no es posible tomar un k elevado. Por tanto, tenemos problemas parecidos a los que aparecen cuando realizamos el contraste de raíz unitaria.

Otra medida de persistencia que es usada en trabajos empíricos es la vida media de un shock (HL). Este estadístico mide la duración de un shock a partir de la siguiente formula:

$$HL = \frac{\ln 0.5}{\ln \hat{\rho}}$$

donde $\hat{\rho}$ es el estimador del coe ciente de correlación entre y_t e y_{t-1}. Cuanto mayor es el valor de esta medida, mayor es el grado de persistencia y, en consecuencia, más cerca estamos de considerar que esta variable no es estacionaria.