



ECONOMETRÍA DE SERIES DE TIEMPO

(PARTE 5, 2022)

Raimundo Soto

1

MODELOS UNIVARIADOS PARA SERIES NO ESTACIONARIAS



MODELOS UNIVARIADOS PARA SERIES NO ESTACIONARIAS

- El objetivo de este capítulo es introducir al lector en la econometría de variables no estacionarias.
- Para ello, lo primero que se hace es discutir las diferencias que tienen procesos estacionarios y no estacionarios.
- Posteriormente se desarrollan los elementos centrales de la teoría asintótica para procesos no estacionarios.
- Finalmente se discuten varios tests que permiten dilucidar empíricamente la naturaleza de las variables y el tratamiento adecuado que se debe darle.

PROCESOS NO ESTACIONARIOS

- Existen diferencias fundamentales entre un proceso estacionario y uno no estacionario en términos de
 - Predicciones
 - Incertidumbre (varianza) de las predicciones
 - Persistencia de las innovaciones (shocks)
- Para entender estas diferencias nos vamos a concentrar en la *tendencia de las variables económicas*

REPRESENTACIONES DE TENDENCIAS EN ANÁLISIS UNIVARIADO

- Hay dos formas básicas:
- La variable corresponde al caso en que hay una **tendencia determinística** de la forma:

$$y_t = \alpha + \beta t + \mu_t$$
$$y_t = \alpha + \beta t + \theta(L)\varepsilon_t$$

Donde $\theta(L)\varepsilon_t$ es una ARMA(p,q,) que refleja toda la dinámica de la variable

- La serie filtrada sería $z_t = y_t - \hat{\alpha} - \hat{\beta}t$
- Esto es lo que se debe usar en un VAR como los ya estudiados.

PROCESOS NO ESTACIONARIOS

- Hay series que tienen **tendencia estocástica** de la forma:

$$\begin{aligned}\Delta y_t &= \delta + \theta(L)\varepsilon_t \\ (1 - L)y_t &= \delta + ARMA(L)\varepsilon_t\end{aligned}$$

Donde Δy_t es la primera diferencia de la serie (tasa de crecimiento) refleja toda la dinámica de la variable

- La tasa de crecimiento *promedio* es δ pero en cada instante de tiempo el crecimiento observado es $\delta + \theta(L)\varepsilon_t$

DIFERENCIAS ENTRE ESTOS PROCESOS NO ESTACIONARIOS

- La naturaleza de los errores de predicción asociados a cada proceso.
- La varianza de dichos errores de predicción.
- La persistencia de las innovaciones de cada proceso.

DIFERENCIAS ENTRE ESTOS PROCESOS NO ESTACIONARIOS

- Consideremos que el proceso con tendencia determinística NO es estacionario, pero lo es alrededor de su tendencia (ET).
- Consideremos que el proceso estacionario en diferencias es NO estacionario, pero su tasa de crecimiento es estacionaria (ED).

¿EQUIVALENCIA ENTRE ESTOS PROCESOS?

- Consideremos el ARMA del modelo ET

- $ARMA(p, q)\mu_t =$

$$(1 - \phi_1 L - \dots - \phi_p L^p)\mu_t = (1 + \theta_1 L + \dots + \theta_q L^q)\varepsilon_t$$

$$\mu_t = \frac{(1 + \theta_1 L + \dots + \theta_q L^q)}{(1 - \phi_1 L - \dots - \phi_p L^p)} \varepsilon_t$$

$$\mu_t = \frac{(1 + \theta_1 L + \dots + \theta_q L^q)}{(1 - \lambda_1 L)(1 - \lambda_2 L) \dots (1 - \lambda_p L)} \varepsilon_t$$

¿EQUIVALENCIA ENTRE ESTOS PROCESOS?

$$\mu_t = \frac{(1 + \theta_1 L + \dots + \theta_q L^q)}{(1 - \lambda_1 L)(1 - \lambda_2 L) \dots (1 - \lambda_p L)} \varepsilon_t = \Psi(L) \varepsilon_t$$

- Cuando se cumple sumabilidad absoluta (absolute summability) $\sum_{j=0}^{\infty} |\Psi_j| < \infty$, entonces $|\lambda_i| < 1 \quad \forall i$ y la variable es estacionaria alrededor de la tendencia determinística β .
- Supongamos que $\lambda_1 = 1$ y $|\lambda_i| < 1 \quad \forall i > 1$ Entonces:

¿EQUIVALENCIA ENTRE ESTOS PROCESOS?

$$(1 - L)\mu_t = \frac{(1 + \theta_1 L + \dots + \theta_1 L^q)}{(1 - \lambda_2 L) \dots (1 - \lambda_p L)} \varepsilon_t = \Psi^*$$

- donde ahora se cumple sumabilidad absoluta para la secuencia de ψ^* .
- El proceso tiene una raíz unitaria.
- Aplicamos operador de rezagos al modelo ET

$$(1 - L)y_t = (1 - L)\alpha + (1 - L)\beta t + (1 - L)\mu_t$$

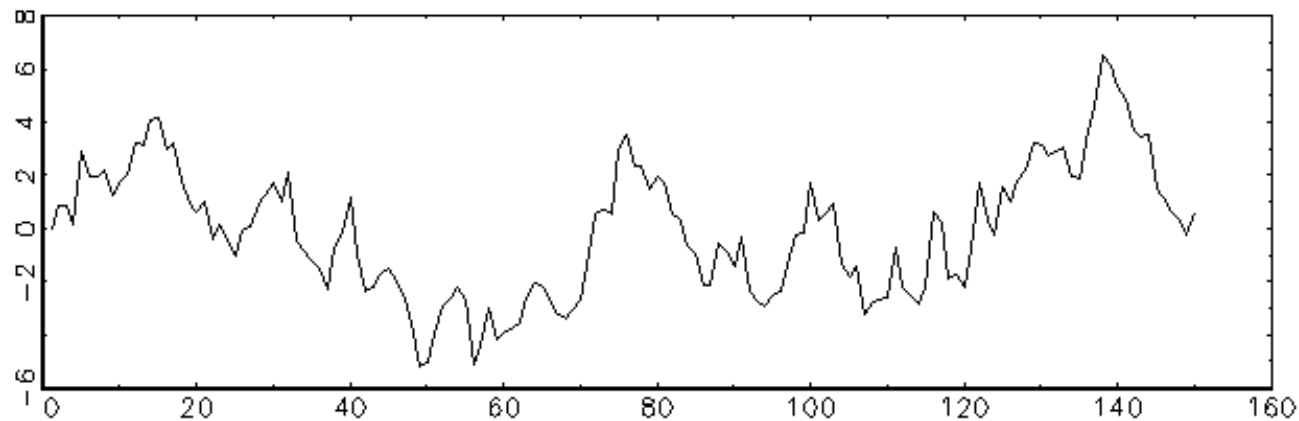
$$(1 - L)y_t = \beta + \Psi^*(L)\varepsilon_t$$

RAÍCES UNITARIAS

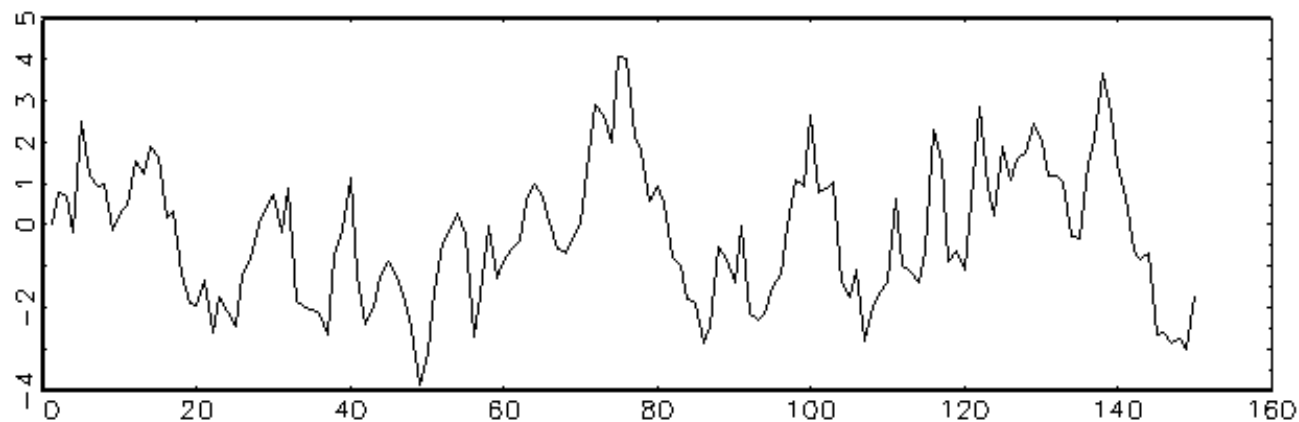
- Obviamente si $\lambda_1 = 1$ y $\lambda_2 = 1$, el proceso tiene dos raíces unitarias y $(1 - L)^2 y_t$ es estacionario en la tasa de crecimiento de la tasa de crecimiento.
- Resulta común denominar a las series de tiempo por su *nivel de integración* de forma tal que el proceso anterior es $I(2)$ y el proceso previo es $I(1)$, en tanto que el proceso ET es llamado $I(0)$.

¿COMO SE VE CADA TIPO DE PROCESO?

Proceso I(1)



Proceso I(0)



PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA

$$y_t = \alpha + \beta t + \psi(L)\varepsilon_t$$

- Veamos la predicción “s” pasos adelante

$$E[y_{t+s}] = E[\alpha + \beta(t + s) + \psi(L)\varepsilon_{t+s}] = \alpha + \beta(t + s)$$

- Es decir la predicción es que la serie va a estar sobre su tendencia determinística
- En el largo plazo las innovaciones (ocurridas antes de $t+1$) van a desaparecer y lo único que queda activo en la predicción es la parte determinística.

PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA

$$y_t = \alpha + \beta t + \psi(L)\varepsilon_t$$

- Veamos la predicción “s” pasos adelante

$$\begin{aligned}\hat{y}_{t+s|t} &= \alpha + \beta(t+s) + \psi(L)\varepsilon_{t+s} \\ &= \alpha + \beta(t+s) + \psi_s\varepsilon_t + \psi_{s-1}\varepsilon_{t+1} + \dots\end{aligned}$$

- Donde la predicción incluye la constante $-\alpha$ -, el crecimiento secular $-\beta(t+s)$ - y el efecto de las innovaciones que aún están presentes en el instante t (es decir, $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$) ponderadas por el efecto que tienen en el instante s ($\psi_s, \psi_{s-1}, \psi_{s-2}, \dots$)

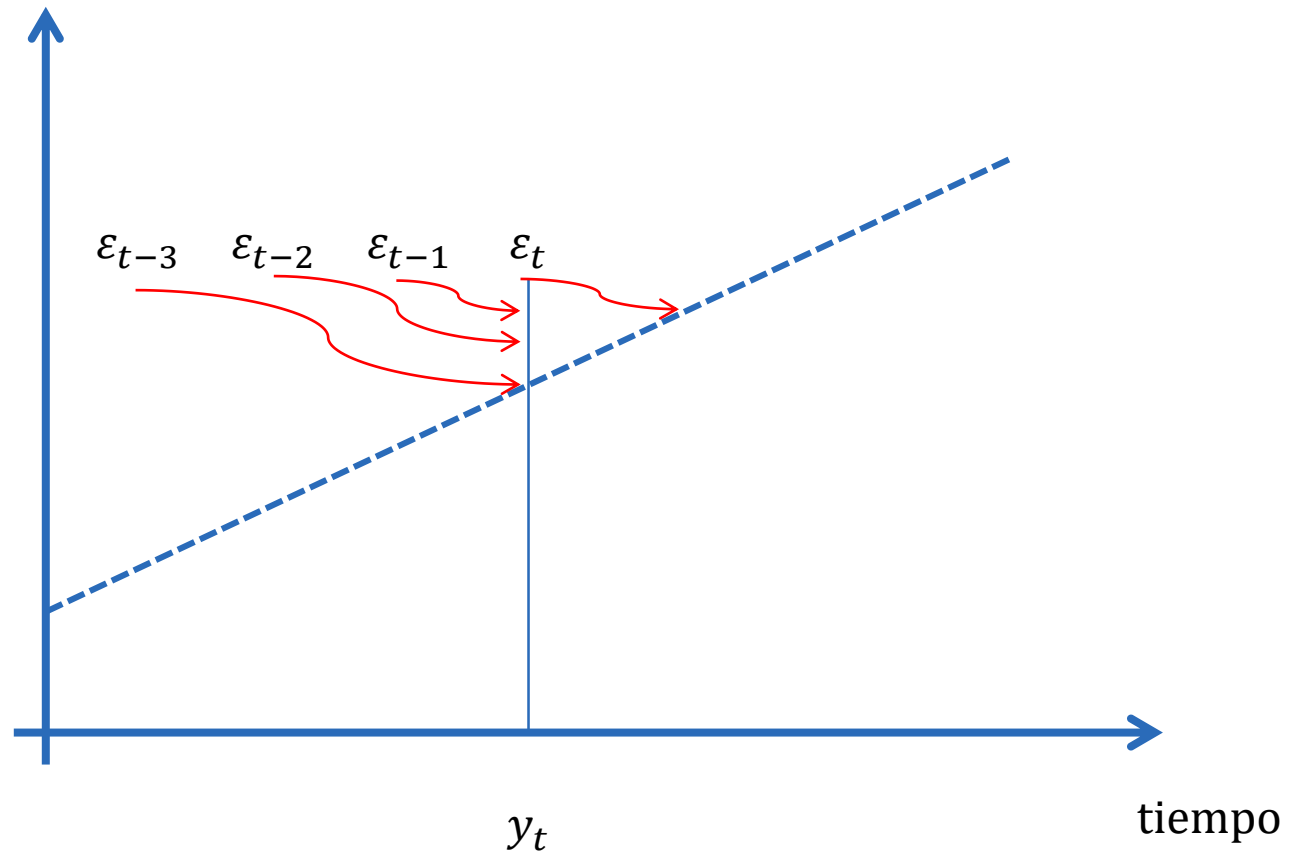
PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA

- Veamos la predicción “s” pasos adelante

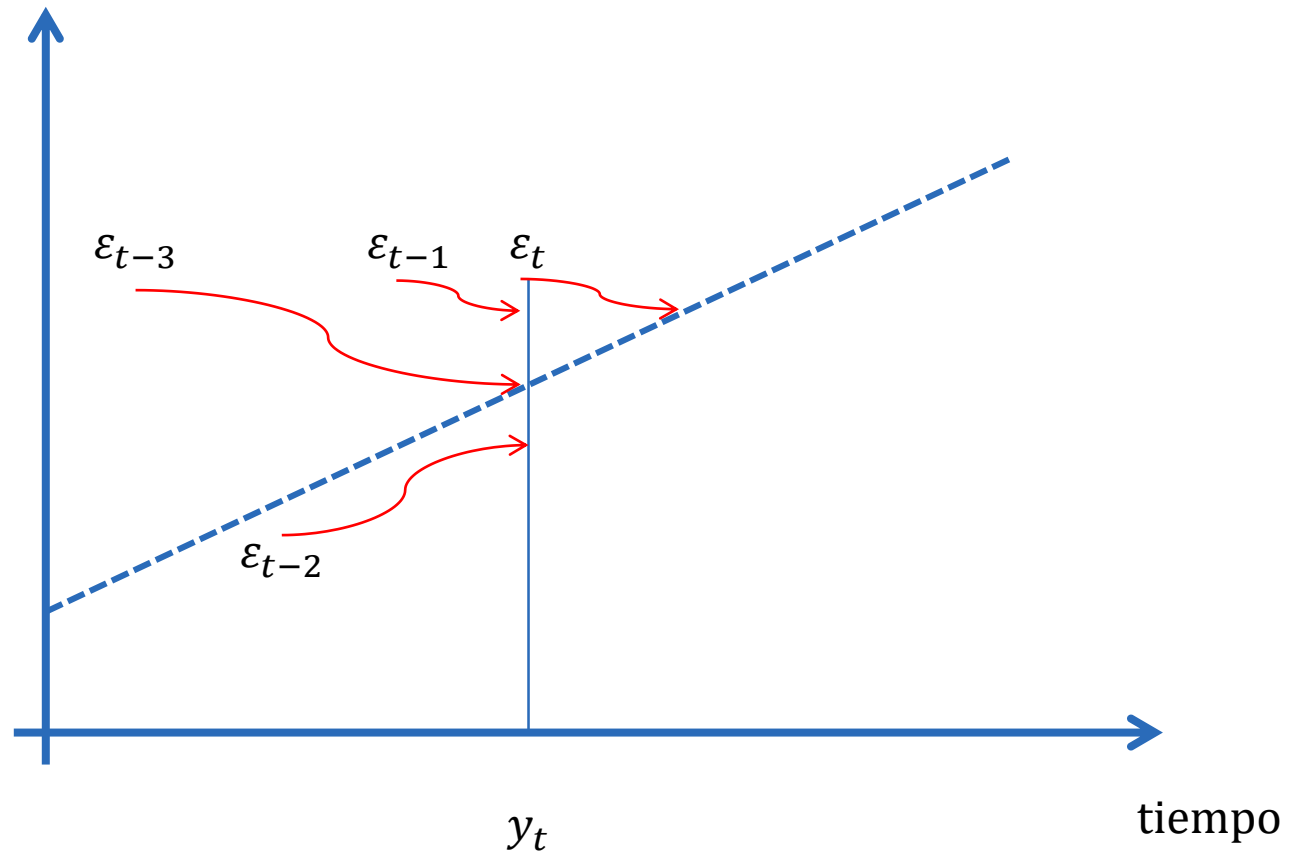
$$\begin{aligned} E[y_{t+s}] &= E[\alpha + \beta(t + s) + \psi(L)\varepsilon_{t+s}] \\ &= \alpha + \beta(t + s) \end{aligned}$$

- Es decir la predicción es que la serie va a estar sobre su tendencia determinística
- En el largo plazo las innovaciones (ocurridas antes de $t+1$) van a desaparecer y lo único que queda activo en la predicción es la parte determinística.

PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA



PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA



PROCESOS NO ESTACIONARIOS CON TENDENCIA ESTOCÁSTICA

$$\Delta y_t = \delta + \theta(L)\varepsilon_t$$

- Veamos la predicción “s” pasos adelante
- Nuestro modelo no predice y_{t+s} sino Δy_{t+s} .
- Para predecir y_{t+s} hay que predecir todas las tasas de crecimiento $\Delta y_{t+1}, \Delta y_{t+2} + \dots + \Delta y_{t+s}$
- Y luego “integramos” para construir:

$$y_{t+s} = y_t + \Delta y_{t+1} + \Delta y_{t+2} + \dots + \Delta y_{t+s}$$

PROCESOS NO ESTACIONARIOS CON TENDENCIA ESTOCÁSTICA

- Así, la predicción “s” pasos adelante es:

$$\begin{aligned}\hat{y}_{t+s|t} = & y_t + s\delta + (\psi_s + \psi_{s-1} + \dots + \psi_1)\varepsilon_t \\ & + (\psi_{s+1} + \psi_s + \dots + \psi_2)\varepsilon_{t-1} \\ & + (\psi_{s+2} + \psi_{s+1} + \dots + \psi_3)\varepsilon_{t-2} + \dots\end{aligned}$$

- Depende de la posición inicial de la serie y_t
 - El intercepto en modelos de variables con tendencia determinística es fijo (α).
- Si δ no es cero, la serie integrada tiene tasa de crecimiento promedio δ pero es estocástica.
- La historia de las innovaciones afecta permanentemente el nivel de la variable.

PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA

- Veamos ahora la incertidumbre de la predicción
- El error cometido al predecir y_{t+s} son producto de todos los shocks futuros que suceden entre t y $t+s$ y que no se han disipado

$$\hat{y}_{t+s} - y_{t+s} = \varepsilon_{t+s} + \varphi_1 \varepsilon_{t+s-1} + \varphi_2 \varepsilon_{t+s-2} + \dots$$

- Tomamos la varianza pero recuerde que todos los shocks tienen igual varianza y son independientes entre si: la varianza de la suma es la suma de las varianzas

$$V(\hat{y}_{t+s} - y_{t+s}) = V(\varepsilon_{t+s} + \varphi_1 \varepsilon_{t+s-1} + \varphi_2 \varepsilon_{t+s-2} + \dots)$$

PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA

$$\begin{aligned} V(\hat{y}_{t+s} - y_{t+s}) &= V(\varepsilon_{t+s} + \varphi_1 \varepsilon_{t+s-1} + \varphi_2 \varepsilon_{t+s-2} + \dots) \\ &= \sigma^2 (1 + \varphi_1^2 + \varphi_2^2 + \dots) \end{aligned}$$

- La varianza de la predicción aumenta con s pero converge a un valor finito porque los φ_i son una secuencia convergente (por eso los shocks se disipan)

PROCESOS NO ESTACIONARIOS CON TENDENCIA ESTOCÁSTICA

- El error de predicción es fácil de obtener pero largo

$$\hat{y}_{t+s} - y_{t+s|t} = \varepsilon_{t+s} + (1 + \psi_1)\varepsilon_{t+s-1} + (1 + \psi_1 + \psi_2)\varepsilon_{t+s-2} + \dots$$

- Encontrar la varianza del error de predicción es.

$$E[\hat{y}_{t+s} - y_{t+s|t}]^2 = \sigma^2[1 + (1 + \psi_1)^2 + (1 + \psi_1 + \psi_2)^2 + \dots]$$

- Aunque los ψ_i son una secuencia convergente, los 1 no se disipan.
- Así, la varianza crece con s y se va a infinito (aún si $\varphi_i = 0$).

PROCESOS NO ESTACIONARIOS CON TENDENCIA DETERMINÍSTICA Y ESTOCÁSTICA

- Finalmente, consideremos la persistencia de los shocks
- En una variable con tendencia determinística los shocks tienen que disiparse (si no lo hicieran no volverían a la tendencia)
- En una variable con tendencia estocástica los shocks no se disipan y por ello no vuelven a ninguna tendencia determinística

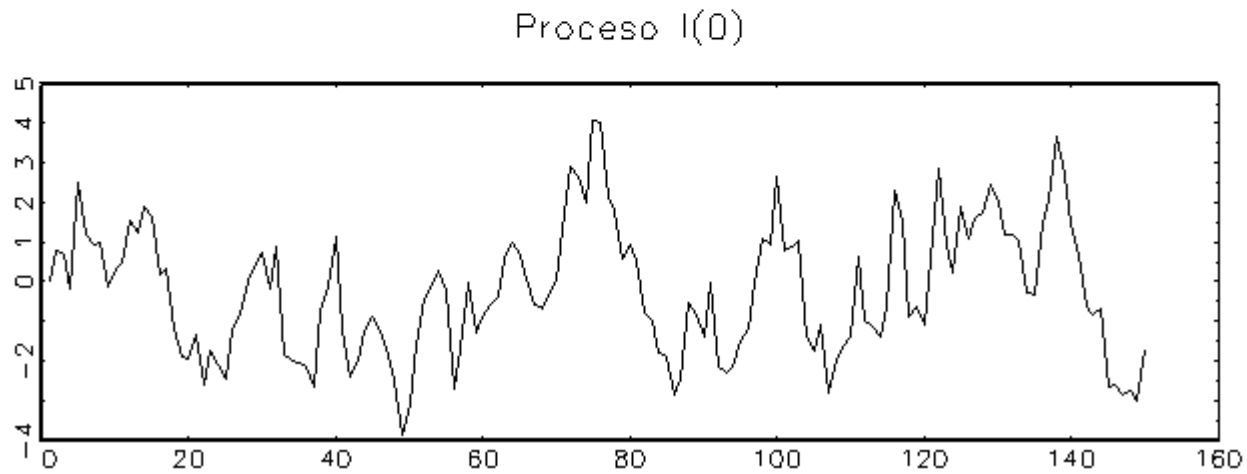
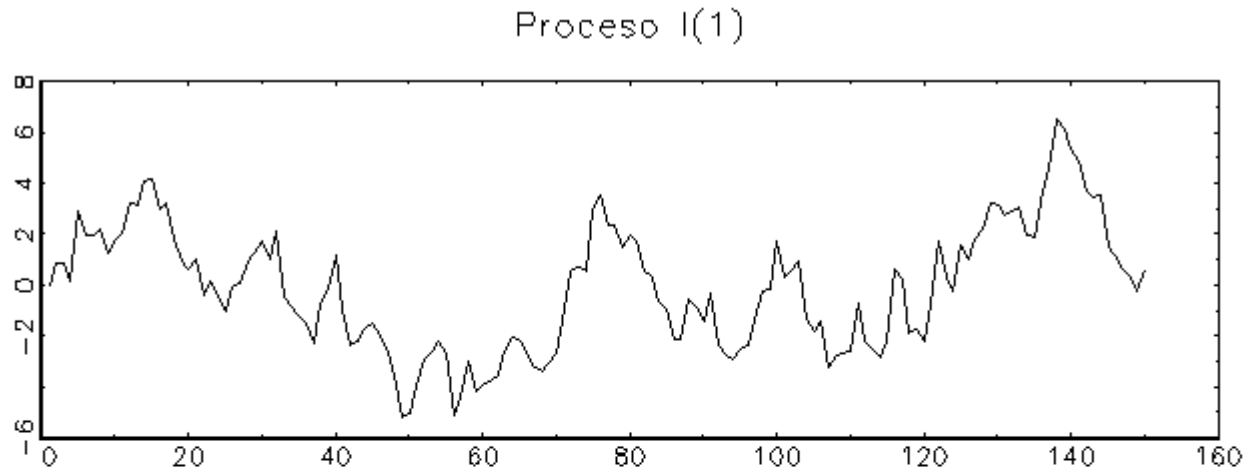
REMOCIÓN DE TENDENCIAS

- La forma que tome la tendencia de una variable –determinística o estocástica– produce notables diferencias para la estimación y el análisis.
- No resulta sorprendente que la manera de remover las tendencias de las series de tiempo sea un tema central en series de tiempo.
 - La manera correcta de remover la tendencia en series estacionarias en tendencia es quitar la tendencia determinística (pre-filtrado).
 - La manera correcta de remover la tendencia en series estacionarias en diferencias es tomar la diferencia de orden d , tal que la variable se vuelva estacionaria.

REMOCIÓN DE TENDENCIAS

- El principal problema es que de solo observar una serie no se sabe de cuál tipo es.

REMOCIÓN DE TENDENCIAS



REMOCIÓN DE TENDENCIAS

- ¿qué pasa si nos equivocamos?
 - Si el proceso es TD y tomamos diferencias, hemos convertido el proceso en uno con shocks permanentes (no invertible)
 - Si el proceso es TE y le quitamos una tendencia lineal, no hemos evitado el problema de tener shocks permanentes
- Equivocarse es además caro en dos dimensiones
 - Correlación espuria (shocks permanentes hacen creer que hay correlación donde no hay nada)
 - Se pierden los niveles al tomar diferencias, información que podría ser útil.

UN PROCESO ELEMENTAL NO ESTACIONARIO

- Consideremos el proceso

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

- Donde

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

- Estudiemos el estimador de mínimos cuadrados ordinarios de ρ en una muestra con T observaciones

$$\hat{\rho}_T = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2}$$

UN PROCESO ELEMENTAL NO ESTACIONARIO

- La distribución del estimador de mínimos cuadrados ordinarios de ρ es:

$$\hat{\rho}_T \rightarrow N(\rho, (1 - \rho^2)\sigma^2/T)$$

- Mientras $|\rho| < 1$, el estimador tiene distribución normal pero cuando $|\rho| = 1$, la distribución degenera porque tiene varianza cero.
- No hay inferencia posible

UN PROCESO ELEMENTAL NO ESTACIONARIO

- Estudiemos el proceso con $|\rho| = 1$

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

- Pero $y_{t-1} = y_{t-2} + \varepsilon_{t-1}$
- Entonces $y_t = y_{t-2} + \varepsilon_{t-1} + \varepsilon_t$
- Así: $y_t = \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \dots$
- Ahora y_t es la suma de todos sus shocks que no se disipan nunca

UN PROCESO ELEMENTAL NO ESTACIONARIO

- Y como $y_{t+s} = y_t + \varepsilon_{t+1} + \varepsilon_{t+2} + \dots + \varepsilon_{t+s}$
- La predicción de y_{t+s} es ... y_t
- Es decir es esencialmente impredecible.
- Y como cada shock es normal, independiente y con igual varianza entonces

$$y_t \sim N(0, t\sigma^2)$$

- La varianza crece con el tiempo (es un alien!!!)

UN PROCESO ELEMENTAL NO ESTACIONARIO

- Volvamos por un segundo al estimador mínimos cuadrados de ρ del modelo $y_t = \rho y_{t-1} + \varepsilon_t$

$$\hat{\rho} = \frac{\text{cov}(y_t, y_{t-1})}{\text{var}(y_{t-1})}$$

- Reemplazamos y_t por el modelo verdadero $y_t = y_{t-1} + \varepsilon_t$.
Obtenemos:

$$\hat{\rho} = \frac{\text{cov}(y_{t-1} + \varepsilon_t, y_{t-1})}{\text{var}(y_{t-1})} = \frac{\text{cov}(y_{t-1}, y_{t-1})}{\text{var}(y_{t-1})} + \frac{\text{cov}(\varepsilon_t, y_{t-1})}{\text{var}(y_{t-1})}$$

UN PROCESO ELEMENTAL NO ESTACIONARIO

$$\hat{\rho} = \frac{\text{cov}(y_{t-1} + \varepsilon_t, y_{t-1})}{\text{var}(y_{t-1})} = \frac{\text{cov}(y_{t-1}, y_{t-1})}{\text{var}(y_{t-1})} + \frac{\text{cov}(\varepsilon_t, y_{t-1})}{\text{var}(y_{t-1})}$$

$$\hat{\rho} = \frac{\text{cov}(y_{t-1} + \varepsilon_t, y_{t-1})}{\text{var}(y_{t-1})} = 1 + \frac{\text{cov}(\varepsilon_t, y_{t-1})}{\text{var}(y_{t-1})}$$

- La distribución de $\hat{\rho}$ dependerá del último término
- Veamos la parte de arriba en detalle:
 - Necesitamos $\text{cov}(\varepsilon_t, y_{t-1}) = \frac{1}{T} \sum_{i=1}^T y_{t-1} \varepsilon_t$

UN PROCESO ELEMENTAL NO ESTACIONARIO

- Pero recordemos que $y_t = y_{t-1} + \epsilon_t$

- Así,
$$y_t^2 = (y_{t-1} + \epsilon_t)^2 = y_{t-1}^2 + 2y_{t-1}\epsilon_t + \epsilon_t^2.$$

- Sumamos para todo t
$$\sum_{t=1}^T y_{t-1}\epsilon_t = \frac{1}{2}[y_T^2 - y_0^2] - \frac{1}{2}\sum_{t=1}^T \epsilon_t^2$$

- Obtenemos
$$y_{t-1}\epsilon_t = \frac{1}{2}[y_t^2 - y_{t-1}^2 - \epsilon_t^2]$$

- Si $y_0 = 0$, entonces
$$\frac{1}{T}\sum_{t=1}^T y_{t-1}\epsilon_t = (1/2)(1/T)y_T^2 - (1/2)(1/T)\sum_{t=1}^T \epsilon_t^2$$

UN PROCESO ELEMENTAL NO ESTACIONARIO

$$\frac{1}{T} \sum_{t=1}^T y_{t-1} \epsilon_t = (1/2)(1/T) y_T^2 - (1/2)(1/T) \sum_{t=1}^T \epsilon_t^2$$

- Dividimos por σ^2 y escribimos esto de manera conocida

$$\frac{1}{\sigma^2 T} \sum_{t=1}^T y_{t-1} \epsilon_t = (1/2) \underbrace{\left[\frac{y_T}{\sigma \sqrt{T}} \right]^2}_{\text{N}(0,1) \text{ al cuadrado}} - (1/2) \underbrace{(1/\sigma^2 T) \sum_{t=1}^T \epsilon_t^2}_{\text{Converge a 1}}$$

- Es decir

N(0,1) al cuadrado Converge a 1
Es una χ^2 Cuando T crece

$$\frac{1}{T} \sum_{t=1}^T y_{t-1} \epsilon_t \text{ converge a } \frac{1}{2} [\chi^2(1) - 1]$$

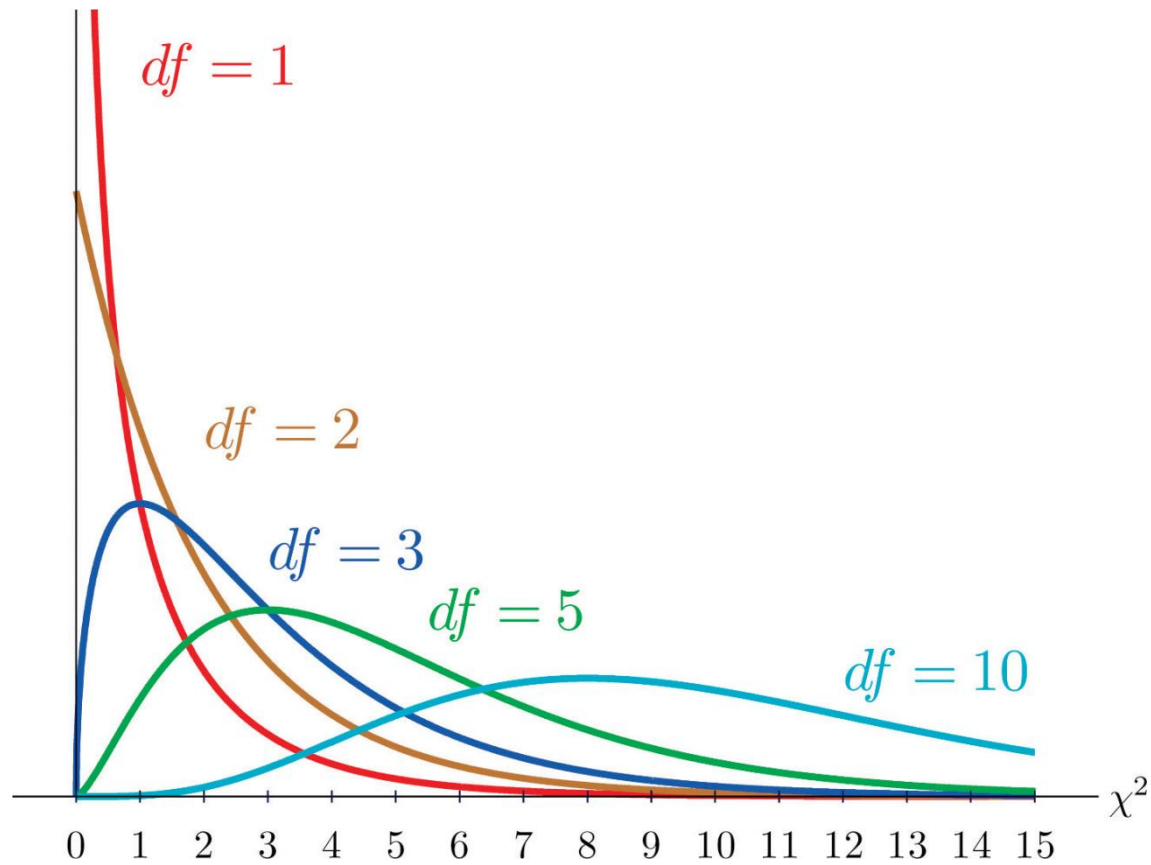
UN PROCESO ELEMENTAL NO ESTACIONARIO

- El denominador de nuestro estimador es:

$$E[y_{t-1}^2] = \sigma^2(t-1) \Rightarrow E \sum_{t=1}^T [y_{t-1}^2] = \sigma^2 \sum_{t=1}^T (t-1) = \sigma^2(T-1) \frac{T}{2}$$

- Parece difícil pero la interpretación es fácil
- Cuando estimas el modelo $y_t = \rho y_{t-1} + \varepsilon_t$ tu crees que $\hat{\rho}$ se distribuye de manera estándar (normal, simétrico, etc.)
- En verdad se distribuye de una manera rarísima
 - Es una χ^2 asimétrica, descentrada, y aplastada

LA DISTRIBUCIÓN χ^2



UN PROCESO ELEMENTAL NO ESTACIONARIO

- ¿y qué?
- Si haces un test $H_0: \hat{\rho} = 1$ usando $y_t = \rho y_{t-1} + \varepsilon_t$
 - No se distribuye normal (o t-Student)
 - La distribución del test no está centrada en 1
 - La distribución depende del número de observaciones (T)
- Si haces un test $H_0: \hat{\rho} \neq 1$ usando $y_t = \rho y_{t-1} + \varepsilon_t$ tienes el mismo problema
- Se debe desarrollar una metodología especial para hacer tests de raíces unitarias.

MOVIMIENTOS BROWNIANOS Y PROCESOS DE WIENER

- Tomemos un proceso tipo paseo aleatorio

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

- Consideremos la innovación $\varepsilon_t = y_t - y_{t-1}$ como proviniendo de una distribución $N(0,1)$.

- Entonces, $y_t = \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \dots$ y por lo tanto,

$$y_t \sim N(0, t)$$

MOVIMIENTOS BROWNIANOS Y PROCESOS DE WIENER

- Consideremos el siguiente intervalo:

$$[y_t - y_s] = \varepsilon_{t+1} + \varepsilon_{t+2} + \cdots + \varepsilon_s$$

entonces se cumplen las siguientes propiedades:

- $[y_t - y_s] \sim N(0, s - t)$
- $[y_t - y_s]$ independiente de $[y_r - y_q]$ si $r < q < s < t$
- Supongamos, ahora, que ε_t está compuesto de dos procesos Gaussianos independientes, donde cada uno es una realización de una distribución $N(0, 1/2)$:

$$\varepsilon_t = e_t^1 + e_t^2$$

MOVIMIENTOS BROWNIANOS Y PROCESOS DE WIENER

- Podemos pensar que e_t^1 corresponde a una parte de Δy :

$$e_t^1 = y_{t-\frac{1}{2}} - y_{t-1}$$
$$e_t^2 = y_T - y_{t-\frac{1}{2}}$$

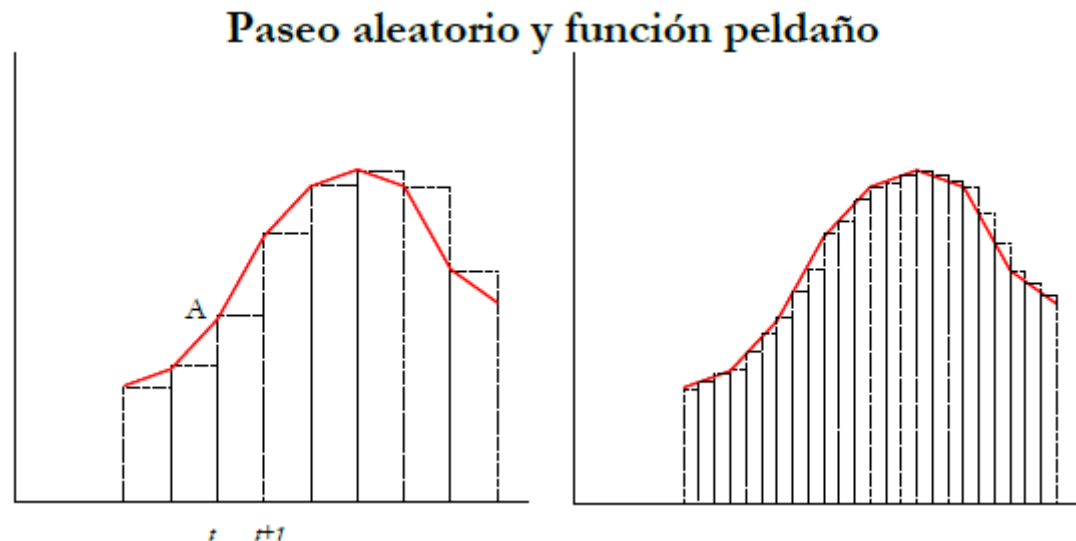
- El incremento del paseo aleatorio podría ser el incremento de medio ε_t primero y otro medio ε_t después, tal que cada “pedazo” es en sí un paseo aleatorio.
- Note que estos procesos e_t también verifican las propiedades descritas más arriba.
- Si hacemos el mismo ejercicio fraccionando el intervalo en k trozos, se obtienen los mismos resultados.

MOVIMIENTOS BROWNIANOS Y PROCESOS DE WIENER

- Si $k \rightarrow \infty$, obtenemos un proceso continuo, llamado Browniano estandar (*standard Brownian motion*) o proceso de Wiener y lo denominamos como $W(t)$.
- Es un paseo aleatorio, pero haciendo el intervalo infinitamente pequeño.

MOVIMIENTOS BROWNIANOS Y PROCESOS DE WIENER

- ¿qué implica este “fraccionamiento”?
 - Panel izquierdo: 8 trozos de igual ancho, pero de distinto alto.
 - Panel derecho: 24 trozos => ajuste muy superior.



- Si hacemos un fraccionamiento infinitamente angosto, pasamos de tiempo discreto a continuo.

MOVIMIENTOS BROWNIANOS Y PROCESOS DE WIENER

- Este proceso tiene tres características.
 - Es un proceso Markoviano. Es decir, la historia futura sólo depende la posición actual de y , pero no del pasado.
 - Los incrementos son independientes; en principio, ellos pueden ser i.i.d.
 - Los incrementos se distribuyen normal.
- En términos de la Figura, considere que la variable se ubica en el punto A en el instante t .
 - Su ubicación en el instante $t+1$ sólo dependerá de su posición inicial (el punto A) y de la magnitud de la innovación que recibe en el instante $t+1$.
 - Que el incremento sea normal sólo dice que, sobre $t+1$, los valores más alejados de la altura de A son menos probables.

MOVIMIENTOS BROWNIANOS Y PROCESOS DE WIENER

- Matemáticamente, se puede escribir el proceso de Wiener como:

$$\text{si } y_t \text{ es } W(t) \text{ entonces } \Delta y_t = \varepsilon_t \sqrt{\Delta t}$$

- Las siguientes características de este proceso son importantes:
 - Cuando $\Delta t \rightarrow 0$, el proceso es $dy_t = \varepsilon_t \sqrt{dt}$.
 - Nótese que $E[dy] = 0$ y $V[dy] = dt$.
- Sea $z(t) = \sigma W(t)$. Entonces esta transformación es también un movimiento Browniano, solo que sus incrementos son i.i.d. y $N(0, \sigma^2)$ en vez de $N(0, 1)$.
- Hemos pasado de procesos discretos a continuos. ¿por qué? Porque necesitamos ver límites cuando $t \rightarrow \infty$.

FUNCIONAL DEL TEOREMA DE LÍMITE CENTRAL

- Necesitamos una generalización del teorema de límite central que se aplique a funciones y no a variables aleatorias.
- Recordemos la versión más simple del TLC:
 - Si μ_t es i.i.d. con media 0 y varianza σ^2 , la media muestral satisface:

$$\sqrt{T}\bar{\mu}_T \rightarrow N(0, \sigma^2)$$

- Es importante notar que el TLC se refiere a la distribución de la media muestral que, en el caso de procesos estacionarios, es constante.

FUNCIONAL DEL TEOREMA DE LÍMITE CENTRAL

- En el caso de procesos no estacionarios, la media va cambiando en el tiempo, por lo que se enfrenta el problema de descubrir la distribución límite de una secuencia de medias que van cambiando con t .
- La solución consiste en mapear las medias de su intervalo creciente $[0, T]$ a un intervalo fijo $[0, 1]$.
- Supongamos un estimador curioso: “*calcule la media muestral con la mitad de los datos, bote el resto*”.
- El estimador $\bar{\mu}_{T/2}$ también cumple el TLC.
- Y es independiente del mismo tipo de estimador hecho con la otra mitad de los datos: el otro intervalo.

FUNCIONAL DEL TEOREMA DE LÍMITE CENTRAL

- Cada intervalo cumple el TCL
- Podemos generalizar esto para una fracción r de los datos, con $r \in [0,1]$ y definir una función tipo peldaño (*step*), es decir, la “media calculada con una fracción r de la muestra de T datos”:

$$X_T(r) = \frac{1}{T} \sum_{t=1}^{rT} \mu_t$$

- entonces, aun cuando la muestra va creciendo de tamaño ($T \rightarrow \infty$), X_T continua definida en el intervalo $[0,1]$.

FUNCIONAL DEL TEOREMA DE LÍMITE CENTRAL

- La función $X_T(r)$ es una función en r , que mapea la información en el intervalo $[0,T]$ en la recta $[0,1]$ de la siguiente forma:

$$X_T(r) = \left[\begin{array}{cc} 0 & 0 \leq r < \frac{1}{T} \\ \frac{\mu_1}{T} & \frac{1}{T} \leq r < \frac{2}{T} \\ \frac{(\mu_1 + \mu_2)}{T} & \frac{2}{T} \leq r < \frac{3}{T} \\ \vdots & \vdots \\ \frac{(\mu_1 + \mu_2 + \dots + \mu_T)}{T} & r = 1 \end{array} \right]$$

FUNCIONAL DEL TEOREMA DE LÍMITE CENTRAL

- Así:

$$\sqrt{T}X_T(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{rT} \mu_t = \frac{1}{\sqrt{T}} \frac{\sqrt{rT}}{\sqrt{rT}} \sum_{t=1}^{rT} \mu_t = \frac{\sqrt{rT}}{\sqrt{T}} \left[\frac{1}{\sqrt{rT}} \sum_{t=1}^{rT} \mu_t \right]$$

- Aplicando el TLC al paréntesis de la ecuación se obtiene que éste converge a $N(0, \sigma^2)$, en tanto que la fracción converge a \sqrt{r} .
- Así:

$$\sqrt{T}X_T(r) \rightarrow N(0, r\sigma^2)$$

$$\sqrt{T} \frac{X_T(r)}{\sigma} \rightarrow N(0, r)$$

- es decir, converge a un proceso Browniano.
- Este resultado es el *teorema del funcional del límite central*.

FUNCIONAL DEL TEOREMA DE LÍMITE CENTRAL

- Considere las diferencias entre:

$$\sqrt{T}\bar{\mu}_T \rightarrow N(0, \sigma^2)$$

$$\sqrt{T}X_T(r) \rightarrow N(0, r\sigma^2)$$

- Nótese que la segunda ecuación es un caso particular de la primera ecuación cuando $r=1$.

DISTRIBUCIÓN DE FUNCIONES DE VARIABLES NO ESTACIONARIAS

- Consideremos de nuevo el paseo aleatorio $y_t = y_{t-1} + \varepsilon_t$. Si $y_0 = 0$, entonces

$$y_t = \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \cdots + \varepsilon_1$$

- Definimos la función estocástica:

$$X_T(r) = \begin{bmatrix} 0 & 0 \leq r < 1/T \\ y_1/T & 1/T \leq r < 2/T \\ y_2/T & 2/T \leq r < 3/T \\ \vdots & \vdots \\ y_T/T & r=1 \end{bmatrix}$$

DISTRIBUCIÓN DE FUNCIONES DE VARIABLES NO ESTACIONARIAS

- La integral de la función es:

$$\int_0^1 X_T(r) dr = \frac{y_1}{T^2} + \frac{y_2}{T^2} + \dots \frac{y_T}{T^2}$$

- La función peldaño que definimos en la ecuación (4.45) señala que los incrementos en r son del tipo y_i/T y ocurren a distancia $1/T$.*

- Entonces,*
$$\int_0^1 X_T(r) dr = \sum_{i=1}^T \int_{(i-1)/T}^{iT/T} X_T(r) dr$$

- Pero en el intervalo $[(i-1)/T, iT/T]$ la función $X_T(r)$ no depende de r . Así,*

$$\sum_{i=1}^T \int_{(i-1)/T}^{iT/T} X_T(r) dr = \sum_{i=1}^T \frac{y_i}{T} \int_{(i-1)/T}^{iT/T} dr$$

- Evaluamos la integral que, trivialmente, es $1/T$.*

- Por lo tanto,*
$$\int_0^1 X_T(r) dr = \frac{1}{T} \sum_{i=1}^T \frac{y_i}{T} \blacksquare$$

DISTRIBUCIÓN DE FUNCIONES DE VARIABLES NO ESTACIONARIAS

- Multiplicando por \sqrt{T} :

$$\sqrt{T} \int_0^1 X_T(r) dr = \sqrt{T} \left[\frac{y_1}{T^2} + \frac{y_2}{T^2} + \dots \frac{y_T}{T^2} \right] = T^{-3/2} \sum_{t=1}^T y_{t-1}$$

- Pero el lado izquierdo converge a $\sigma \int_0^1 W(r) dr$

- Por lo tanto:

$$T^{-3/2} \sum_{t=1}^T y_{t-1} \rightarrow \sigma \int_0^1 W(r) dr$$

DISTRIBUCIÓN DE FUNCIONES DE VARIABLES NO ESTACIONARIAS

- Intuición: reescribamos

$$\begin{aligned} T^{\frac{-3}{2}} \sum_{t=1}^T y_{t-1} &= T^{\frac{-3}{2}} [\mu_1 + (\mu_1 + \mu_2) + (\mu_1 + \mu_2 + \mu_3) + \dots] \\ &= T^{\frac{-3}{2}} [\mu_1(T-1) + \mu_2(T-2) + \mu_3(T-3) + \dots] \\ &= T^{\frac{-3}{2}} \sum_{t=1}^T (T-t) \mu_t = T^{\frac{-1}{2}} \sum_{t=1}^T \mu_t - T^{\frac{-3}{2}} \sum_{t=1}^T t \mu_t \end{aligned}$$

- Usamos un resultado conocido

$$\begin{bmatrix} \sqrt{T} \sum_{t=1}^T \mu_t \\ T^{-3/2} \sum_{t=1}^T t \mu_t \end{bmatrix} \rightarrow N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}^{-1} \right)$$

DISTRIBUCIÓN DE FUNCIONES DE VARIABLES NO ESTACIONARIAS

- La ecuación anterior dice es que $T^{-3/2} \sum_{t=1}^T y_{t-1}$ es asintóticamente Gaussiano con media cero y varianza, $\sigma^2[1 - 2 \cdot \frac{1}{2} + \frac{1}{3}] = \sigma^2 \frac{1}{3}$.
- El FTLC describe $T^{-3/2} \sum_{t=1}^T y_{t-1}$ como una variable aleatoria (paseo aleatorio) con distribución $N(0, \sigma/3)$.
- En resumen, si y_t es un paseo aleatorio sin deriva, la media muestral diverge, pero $T^{-3/2} \sum_{t=1}^T y_{t-1}$ converge a una variable aleatoria que se describe por la integral de un proceso Browniano con varianza σ^2 .

Convergencia en procesos estocásticos

Modelo: $y_t = y_{t-1} + u_t$

Variable	Converge a
$T^{-\frac{1}{2}} \sum_{t=1}^T u_t$	$\sigma \cdot W(1)$
$T^{-5/2} \sum_{t=1}^T t y_{t-1}$	$\sigma \int_0^1 r W(r) dr$
$T^{-1} \sum_{t=1}^T u_t y_{t-1}$	$\frac{1}{2} \sigma^2 \cdot [W(1)^2 - 1]$
$T^{-2} \sum_{t=1}^T y_{t-1}^2$	$\sigma^2 \int_0^1 [W(r)]^2 dr$
$T^{-\frac{3}{2}} \sum_{t=1}^T y_{t-1}$	$\sigma \int_0^1 [W(r)]^2 dr$
$T^{-\frac{3}{2}} \sum_{t=1}^T t u_t$	$\sigma W(1) - \sigma \int_0^1 W(r) dr$
$T^{-3} \sum_{t=1}^T t y_{t-1}^2$	$\sigma \int_0^1 r [W(r)]^2 dr$
$T^{v+1} \sum_{t=1}^T t^v$	$\frac{1}{1+v}$

DISTRIBUCIÓN DE TESTS DE FUNCIONES PARA VARIABLES NO ESTACIONARIAS

- Usando el Cuadro 4.1 derivamos la distribución del test de raíces unitarias para distintos tipos de procesos estocásticos.
- Estudiamos qué le sucede al test t hecho sobre los estimadores de los parámetros de una regresión $y_t = \alpha + \beta t + \rho y_{t-1} + \mu_t$ cuando el verdadero proceso que genera y_t tiene o no tiene una raíz unitaria.
- Para derivar los resultados de convergencia del Cuadro 4.1 se utiliza el Teorema del Mapeo de Continuos (*continuous mapping theorem*), que señala que si $f(\cdot)$ es un funcional continuo en el intervalo $[0,1]$, entonces:

$$f(X_T(r)) \rightarrow f(W(r))$$

DISTRIBUCIÓN DE TESTS DE FUNCIONES PARA VARIABLES NO ESTACIONARIAS

- En este sentido, el “mapeo” toma la siguiente forma:

$$\begin{aligned}\Sigma &\Rightarrow \int \\ t &\Rightarrow r \\ \varepsilon_t &\Rightarrow dW(r) \\ y_t &\Rightarrow W(r)\end{aligned}$$

CASO 1. MODELO SIN CONSTANTE NI TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO

- Supongamos que hacemos un AR(1), $y_t = \rho y_{t-1} + \varepsilon_t$ y estimamos $\hat{\rho}$ por mínimos cuadrados ordinarios. ¿Qué pasa con el estimador de mínimos cuadrados ordinarios de ρ si el verdadero ρ es 1?

$$\hat{\rho}_T = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2}$$

Y se desvía del verdadero valor en:

$$(\hat{\rho}_T - 1) = \frac{\sum_{t=1}^T y_{t-1} \mu_t}{\sum_{t=1}^T y_{t-1}^2}$$

CASO 1. MODELO SIN CONSTANTE NI TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO

$$(\hat{\rho}_T - 1) = \frac{\sum_{t=1}^T y_{t-1} \mu_t}{\sum_{t=1}^T y_{t-1}^2}$$

- Numerador converge a $\frac{1}{2} \sigma^2 [W(1)^2 - 1]$
- Denominador converge a $\sigma^2 \int [W(r)]^2 dr$
 - Este último es un valor dado que no interviene de manera importante en la distribución del estimador $\hat{\rho}_T$.

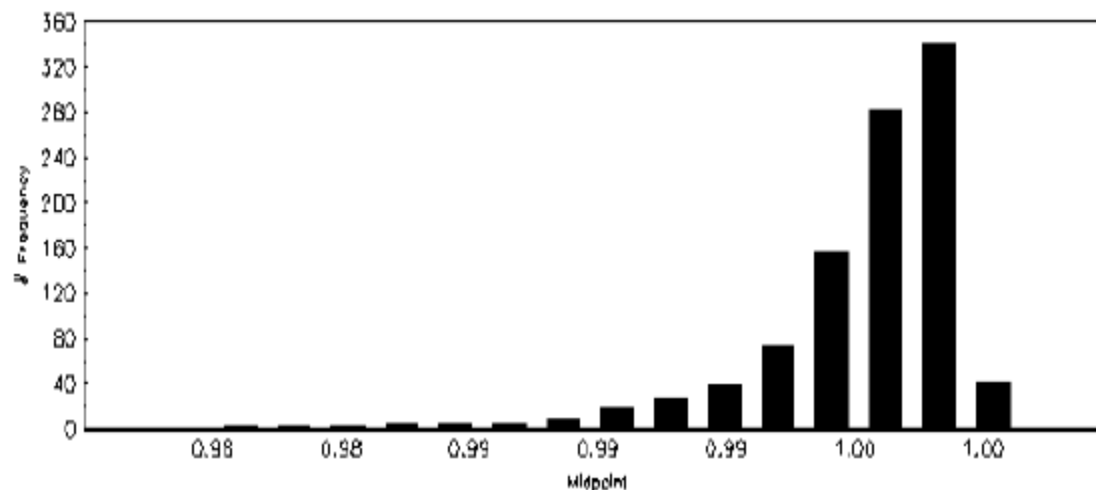
CASO 1. MODELO SIN CONSTANTE NI TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO

$$(\hat{\rho}_T - 1) \rightarrow \frac{\frac{1}{2} \sigma^2 [W(1)^2 - 1]}{\sigma^2 \int [W(r)]^2 dr}$$

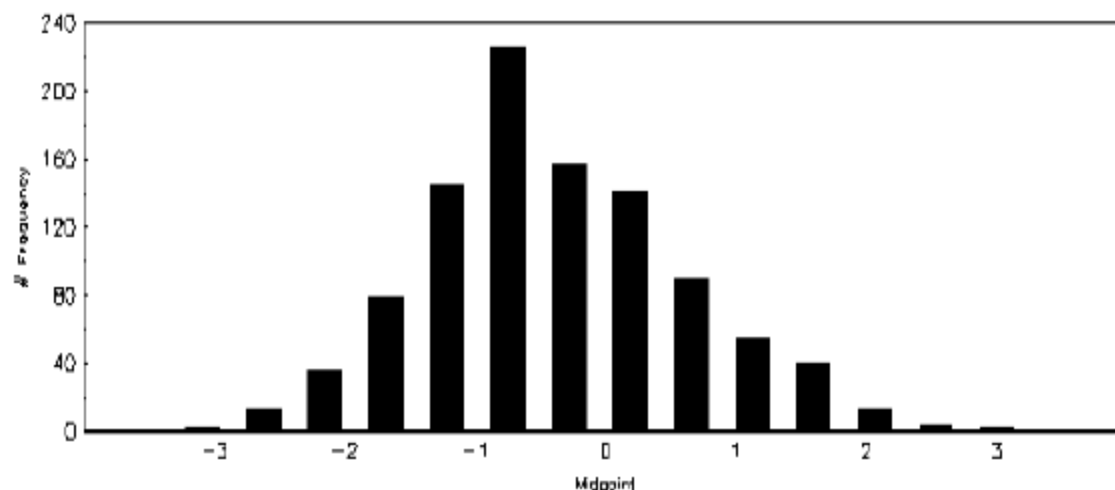
- La distribución de $T(\hat{\rho}_T - 1)$, y por ende la distribución de $\hat{\rho}_T$, depende de $W(1)^2$.
 - Como $W(1)$ es $N(0,1)$, entonces $W(1)^2$ es $\chi^2(1)$.
- Esto implica :
 - La probabilidad de obtener un valor menor que 1 en una $\chi^2(1)$ es de 68%. Es decir, en un 68% de muestras generadas por un paseo aleatorio, el $\hat{\rho}$ estimado por MCO será menor que 1 aun si el verdadero proceso tiene $\rho = 1$.
 - La distribución no es simétrica, sino que está sesgada hacia la izquierda.

Distribución del estimador y test t cuando se estima un AR(1) y el modelo es paseo aleatorio
Tamaño de Muestra = 1000, Número de Replicaciones = 1000

Distribución del Estimador Mínimos Cuadrados



Distribución del Test t



CASO 2. MODELO CON CONSTANTE PERO SIN TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO.

- Supongamos que hacemos un AR(1), $y_t = \alpha + \rho y_{t-1} + \varepsilon_t$ y estimamos $\hat{\rho}$ por mínimos cuadrados ordinarios.

$$\hat{\beta}_T = \begin{bmatrix} \hat{\alpha}_T \\ \hat{\rho}_T \end{bmatrix} = \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum y_{t-1} y_t \end{bmatrix}$$

- La hipótesis nula es $H_0: \alpha = 0, \rho = 1$, entonces obtenemos

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{\rho}_T - 1 \end{bmatrix} = \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum \mu_t \\ \sum y_{t-1} \mu_t \end{bmatrix}$$

CASO 2. MODELO CON CONSTANTE PERO SIN TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO.

- Es directo obtener

$$T(\hat{\rho}_T - 1) \rightarrow \frac{1/2 \cdot [W(1)^2 - 1] - W(1) \int_0^1 W(r) dr}{\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r) dr \right)^2}$$

- El estimador MCO de ρ no tiene como distribución límite un proceso Gaussiano.
- Tampoco se distribuye como el caso 1.
- No se puede usar los valores críticos de la tabla anterior, sino una nueva tabla.
- La distribución del test es aun mas *skewed* hacia la izquierda.

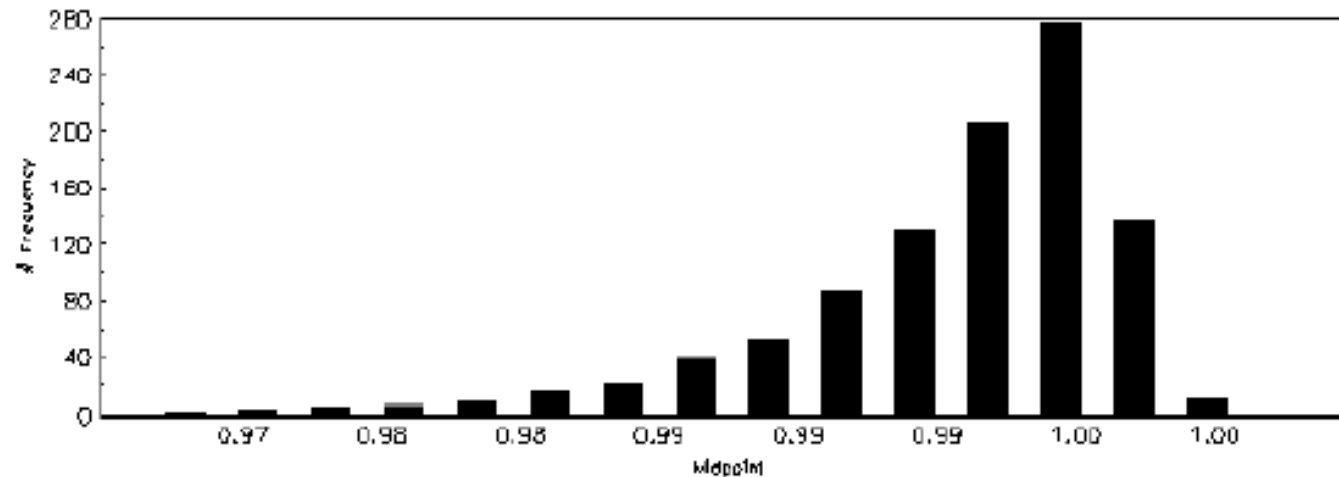
CASO 2. MODELO CON CONSTANTE PERO SIN TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO.

- Dickey y Fuller (1979) proponen un test tipo t . Este test se distribuye asintóticamente como:

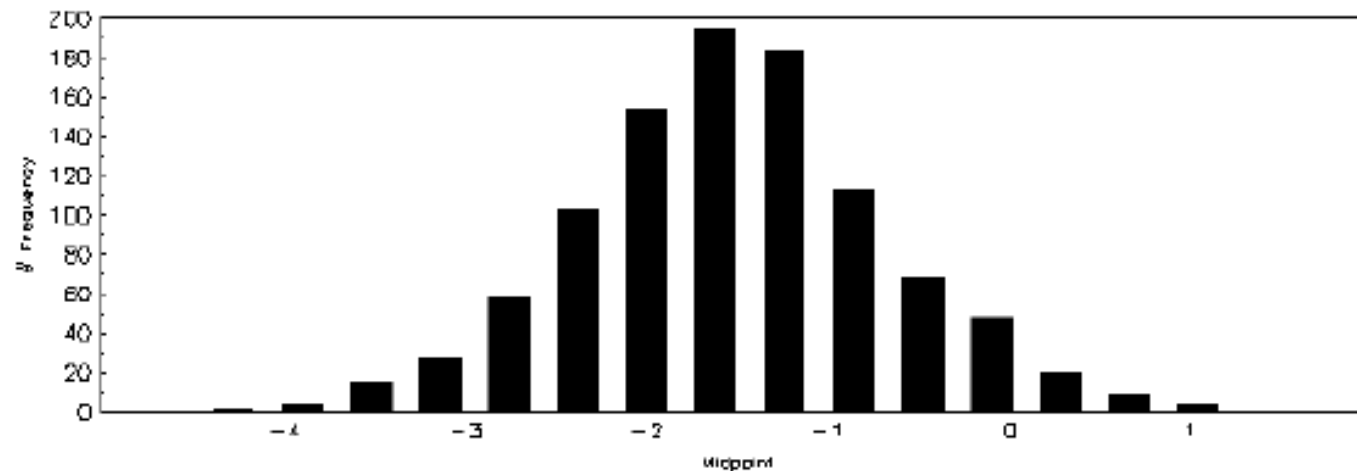
$$\frac{T(\hat{\rho}_T - 1)}{T \hat{\sigma}_{\rho_T}} \rightarrow \frac{1/2 \cdot [W(1)^2 - 1] - W(1) \int_0^1 W(r) dr}{\left[\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r) dr \right)^2 \right]^{1/2}}$$

Distribución del estimador y test t cuando se estima un AR(1) con constante y el verdadero modelo es paseo aleatorio
 Tamaño de Muestra = 1000, Número de Replicaciones = 1000

Distribucion del Estimador Minimas Cuadrados



Distribucion del Test t



CASO 3. MODELO CON CONSTANTE PERO SIN TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO CON TENDENCIA ESTOCÁSTICA.

- Supongamos que el verdadero proceso es $y_t = \alpha + y_{t-1} + \varepsilon_t$, y estimamos $\hat{\rho}$ en $y_t = \alpha + \rho y_{t-1} + \varepsilon_t$, por MCO cuando $\alpha \neq 0$.
- Demostraremos que la distribución de los estimadores cambia bastante con respecto a los casos anteriores
- El verdadero proceso puede ser descrito como:
$$y_t = t\alpha + y_0 + \sum \varepsilon_t = t\alpha + y_0 + \varphi_t$$
- Consideremos, ahora,

$$\sum_{t=1}^T y_{t-1} = \sum_{t=1}^T [y_0 + \alpha(t-1) + \psi_{t-1}]$$

CASO 3. MODELO CON CONSTANTE PERO SIN TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO CON TENDENCIA ESTOCÁSTICA.

$$\sum_{t=1}^T y_{t-1} = \sum_{t=1}^T [y_0 + \alpha(t-1) + \psi_{t-1}]$$

- el primer término es Ty_0 y converge si se divide por T ;
 - el segundo converge a $\alpha/2$ si se divide por T^2 ,
 - el tercero converge a $\sigma \int W(r)dr$ si se divide por $T^{3/2}$.
- Entonces, el segundo termino domina asintóticamente a los otros dos. Así:

$$\frac{1}{T^2} \sum_{t=1}^T y_{t-1} = \frac{1}{T} y_0 + \frac{1}{T^2} \sum_{t=1}^T \alpha(t-1) + \frac{1}{T^2} \sum_{t=1}^T \psi_{t-1} \rightarrow 0 + \frac{\alpha}{2} + 0$$

CASO 3. MODELO CON CONSTANTE PERO SIN TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO CON TENDENCIA ESTOCÁSTICA.

- Estudiamos $\sum_{t=1}^T y_{t-1}^2$

$$\begin{aligned} \sum y_{t-1}^2 = & \sum y_0^2 + \sum \alpha^2 (t-1)^2 + \sum \psi_{t-1}^2 \\ & + \sum 2 y_0 \alpha (t-1) + \sum 2 y_0 \alpha \psi_{t-1} + \sum 2 \alpha (t-1) \psi_{t-1} \end{aligned}$$

- el orden de probabilidad es $O(T)$, $O(T^3)$, $O(T^2)$, $O(T^2)$, $O(T^{3/2})$ y $O(T^{5/2})$, respectivamente.

- Entonces: $\frac{1}{T^3} \sum_{t=1}^T y_{t-1}^2 \rightarrow \frac{\alpha^2}{3}$

- Con el mismo raciocinio se deriva: $\frac{1}{T^{3/2}} \sum_{t=1}^T y_{t-1} \mu_t \rightarrow \frac{1}{T^{3/2}} \sum_{t=1}^T \alpha (t-1) \mu_t$

CASO 3. MODELO CON CONSTANTE PERO SIN TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO CON TENDENCIA ESTOCÁSTICA.

- De manera sintética

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

- Definimos

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

- Y obtenemos:

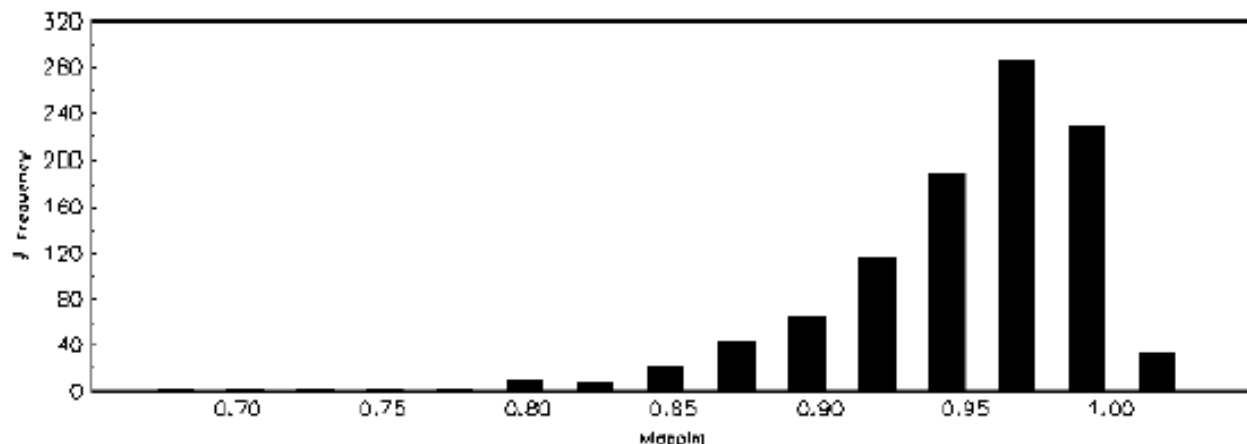
$$\begin{bmatrix} T^{1/2}(\hat{\alpha}_T - \alpha) \\ T^{3/2}(\hat{\rho}_T - 1) \end{bmatrix} \rightarrow N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & \alpha/2 \\ \alpha/2 & \alpha/3 \end{bmatrix}\right)$$

- es decir, obtuvimos lo mismo que en el caso del modelo con tendencia determinística. Por lo tanto se aplican los tests t y F directamente.

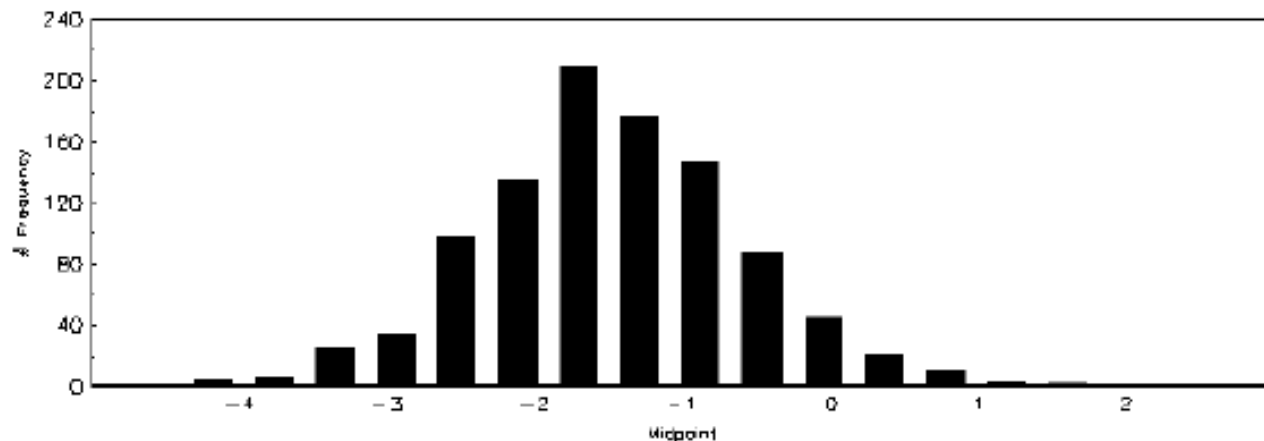
Distribución del estimador y test t cuando se estima un AR(1) con constante y el verdadero modelo es paseo aleatorio con tendencia aleatoria

Tamaño de Muestra = 1000, Número de Replicaciones = 1000

Distribucion del Estimador Mínimos Cuadrados



Distribucion del Test t



CASO 4. MODELO CON CONSTANTE Y TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO CON TENDENCIA ESTOCÁSTICA

- Considere el proceso $y_t = \alpha + \rho y_{t-1} + \delta t + \varepsilon_t$
- Se puede escribir como

$$\begin{aligned} y_t &= (1-\rho)\alpha + \rho[y_{t-1} - \alpha(t-1)] + (\delta + \rho\alpha)t + \varepsilon_t \\ &= \alpha^* + \rho^* \psi_{t-1} + \delta^* t + \varepsilon_t \end{aligned}$$

- Bajo la hipótesis nula $\psi_t = y_0 + \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_{t-1}$
- El estimador MCO es:

$$\begin{bmatrix} \hat{\alpha}_T^* - 0 \\ \hat{\rho}_T^* - 1 \\ \hat{\delta}_T^* - \alpha_0 \end{bmatrix} = \begin{bmatrix} T & \sum \psi_{t-1} & \sum t \\ \sum \psi_{t-1} & \sum \psi_{t-1}^2 & \sum \psi_{t-1} t \\ \sum t & \sum t \psi_{t-1} & \sum t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum \mu_t \\ \sum \psi_{t-1} \mu_t \\ \sum t \mu_t \end{bmatrix}$$

CASO 4. MODELO CON CONSTANTE Y TENDENCIA DETERMINÍSTICA Y EL VERDADERO PROCESO ES UN PASEO ALEATORIO CON TENDENCIA ESTOCÁSTICA

- Entonces

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

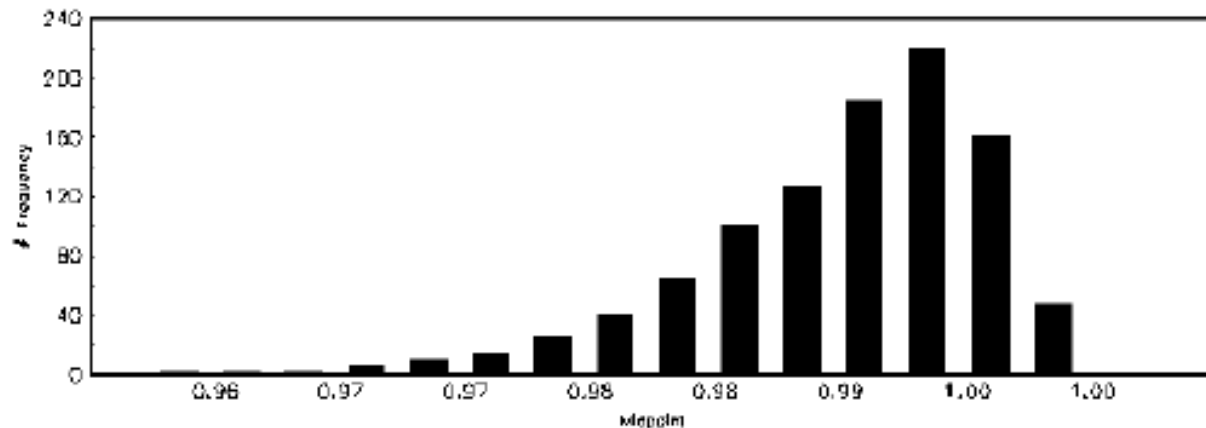
- La distribución del test es

$$\begin{bmatrix} T^{1/2} \hat{\alpha}_T^* \\ T(\hat{\rho}_T^* - 1) \\ T^{3/2}(\hat{\delta}_T^* - \alpha_0) \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \int W(r) dr & 1/2 \\ \int W(r) dr & \int W(r)^2 dr & \int rW(r) dr \\ 1/2 & \int rW(r) dr & 1/3 \end{bmatrix}^{-1} \cdot \begin{bmatrix} W(1) \\ 1/2(W(1)^2 - 1) \\ W(1) - \int W(r) dr \end{bmatrix}$$

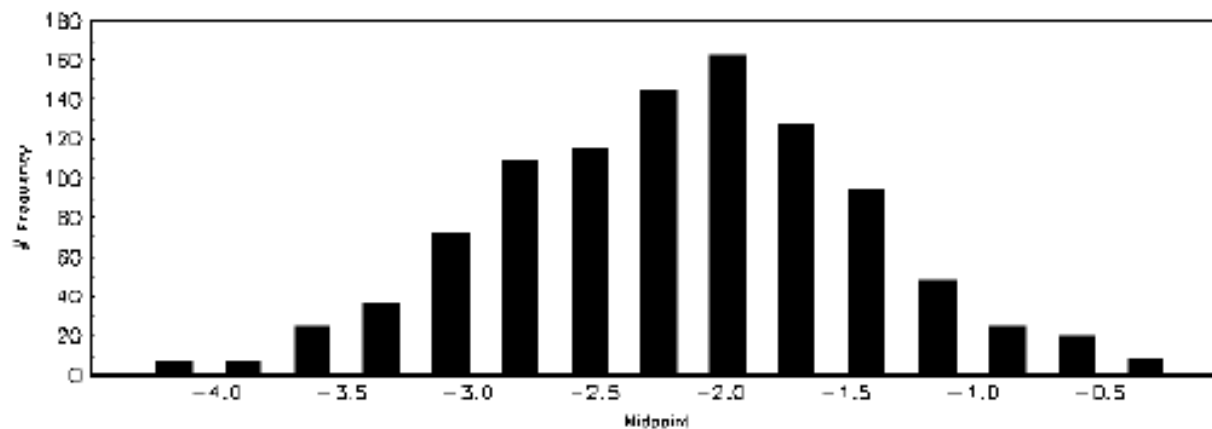
- La distribución de $T(\hat{\rho}_T - 1)$ no depende de α o σ , en particular no depende de que $\alpha = 0$
- La distribución de $(\hat{\rho}_T - 1)/\sigma_{\hat{\rho}_T}$ es más ancha que todas las anteriores .

Distribución del estimador y test t cuando se estima un AR(1) con tendencia determinística y el verdadero modelo es paseo aleatorio con tendencia aleatoria
 Tamaño de Muestra = 1000, Número de Replicaciones = 1000

Distribución del Estimador Mínimos Cuadrados



Distribución del Test t



TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS

- Existe una enorme variedad de tests de raíces unitarias, los que difieren en enfoque, poder y robustez frente a desvíos en las condiciones supuestas para derivarlos.
- Veremos primero los tests paramétricos de raíces unitarias.
- Para el modelo $y_t = \alpha + \rho y_{t-1} + \varepsilon_t$ la hipótesis nula de una raíz unitaria sería $H_0: \hat{\rho} = 1$. Un test de Wald de dicha hipótesis exigiría computar $\hat{\rho} = 1/\sigma_{\hat{\rho}}$.
- Resulta conveniente modificar ligeramente la hipótesis nula de modo que se pueda leer el resultado del test directamente de la regresión.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS

- Restamos y_{t-1} del modelo a cada lado

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

- Ahora la hipótesis nula es simplemente $H_0: \phi = 0$.
- La hipótesis alternativa del test es que el modelo es estacionario.
- Como hay varias alternativa de modelos estacionarios, habrán varias especificaciones de tests de raíces unitarias que debemos considerar.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS

- En principio, cuando la variable no tiene una raíz unitaria, esta podría
 - Oscilar alrededor de cero,
 - Oscilar alrededor de una constante,
 - Oscilar alrededor de una tendencia determinística lineal.
- Cada una de estos casos es una distinta hipótesis alternativa.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: FAMILIA DE TESTS DE DICKEY-FULLER

- Podemos anidar ambas hipótesis en un solo modelo.
- Habrá varias hipótesis alternativas, dependiendo del tipo de proceso que caracteriza a la serie cuando ésta es estacionaria.

$$(a) \Delta y_t = \phi y_{t-1} + \epsilon_t$$

$$(b) \Delta y_t = \alpha + \phi y_{t-1} + \epsilon_t$$

$$(c) \Delta y_t = \alpha + \beta t + \phi y_{t-1} + \epsilon_t$$

- La especificación (a) corresponde al caso de la hipótesis alternativa que el modelo es simplemente autoregresivo – ρ será menor que 1– con media cero y sin tendencia determinística.
- En la especificación (b) la hipótesis alternativa corresponde a un proceso con media distinta de cero pero sin tendencia determinística.
- La ultima especificación es el modelo mas general, en el sentido que permite que tanto la media como la tendencia determinística sean libres.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: FAMILIA DE TESTS DE DICKEY-FULLER

- Testeo de raíces unitarias debe hacerse de “general a particular”.
 - Se empieza por con el modelo (c) y se verifica que el estimador del parámetro β sea estadísticamente distinto de cero.
 - Si no se rechaza dicha hipótesis nula, se estima el modelo (b), verificándose que el parámetro el estimador del α sea cero.
 - Si no se rechaza dicha hipótesis nula, se estima la especificación (a).
- Nótese que la distribución de los estimadores de α y β es estándar (asintóticamente normal), en tanto que la del estimador de ρ es no estándar.
- Al cambiar de hipótesis alternativa, cambia la forma de la distribución del test debido a la presencia de parámetros incidentales.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: FAMILIA DE TESTS DE DICKEY-FULLER

- Así, habrá que construir tablas para cada tipo de test y nivel de significancia.
- Como puede verse, el tamaño del test depende del número de datos de la muestra y de la presencia de parámetros incidentales (*nuisance parameters*) los que afectan la ubicación de la distribución Dickey-Fuller y el ancho de los intervalos de confianza del test sobre el parámetro de interés.

Valores críticos de la distribución Dickey Fuller para el test tipo “t”

Basado en 10.000 replicaciones

	<i>Tamaño Muestra</i>	<i>Valor crítico al 95%</i>	<i>Valor crítico al 90%</i>
Modelo (a)	50	-1,91	-1,60
	100	-1,96	-1,62
	1000	-1,91	-1,95
Modelo (b)	50	-2,92	-2,60
	100	-2,85	-2,57
	1000	-2,88	-2,57
Modelo (c)	50	-3,51	-3,16
	100	-3,42	-3,13
	1000	-3,38	-3,11

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: FAMILIA DE TESTS DE DICKEY-FULLER AUMENTADO

- La estimador de la varianza de ε_t en presencia de autocorrelacion residual no es consistente.
- Se puede corregir usando rezagos de la variable endógena

$$(a) \Delta y_t = \phi y_{t-1} + \epsilon_t + \sum_{p=1} \theta_p \Delta y_{t-p}$$

$$(b) \Delta y_t = \alpha + \phi y_{t-1} + \epsilon_t + \sum_{p=1}^k \theta_p \Delta y_{t-p}$$

$$(c) \Delta y_t = \alpha + \beta t + \phi y_{t-1} + \epsilon_t + \sum_{p=1}^k \theta_p \Delta y_{t-p}$$

- Nuevamente, la distribución de los estimadores de α , β y θ se distribuyen asintóticamente normales, en tanto que el estimador de ρ se distribuye de manera no estándar.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: FAMILIA DE TESTS PHILLIPS-PERRON

- Phillips y Perron (1988) utilizan el material de convergencia vistos para desarrollar un test paramétrico de raíces unitarias para el caso que el residuo sigue cualquier tipo de correlación serial (es decir, no necesariamente AR(p)).
- Si la primera diferencia de y_t es estacionaria, $\mu_t = \psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, y ϵ_t es i.i.d. $(0, \sigma^2)$, entonces el **nivel** de y_t puede ser descrito como:

$$\begin{aligned} y_t &= \mu_1 + \dots + \mu_t + y_0 \\ &= \psi(1)(\epsilon_1 + \dots + \epsilon_t) + \eta_t - \eta_0 + y_0 \end{aligned}$$

- Donde $\psi(1) = \sum_{j=0}^{\infty} \psi_j$, $\eta_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$, $\alpha_j = -(\psi_{j+1} + \psi_{j+2} + \dots)$ y $\sum_{j=0}^{\infty} |\alpha_j| < \infty$

- Es decir, la combinación de un paseo aleatorio $\psi(1)(\epsilon_1 + \dots + \epsilon_t)$, condiciones iniciales η_0 y y_0 y un proceso estocástico, η_t .

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: FAMILIA DE TESTS PHILLIPS-PERRON

- Esta descomposición (desarrollada por Beveridge y Nelson) se usa para obtener el caso general de correlación serial.
- Para hacer el test,
 - 1. se estima un modelo AR(1) típico $y_t = \lambda y_{t-1} + \varepsilon_t$ y se calculan los estimadores de λ y σ y el error estándar de la regresión, s .
 - 2. se estima la función de autocorrelación de los residuos muestrales, la que se usa para construir el estimador de Newey-West de la persistencia de la serie $\hat{\lambda} = \hat{\psi}(1)\hat{\sigma}$ y de la varianza de los residuos, $\hat{\gamma}_0$.
 - 3. se transforman los estimadores obtenidos en 1. para obtener tests que tienen la distribución del caso (2). Las transformaciones son:

$$[\hat{\gamma}_0 \hat{\lambda}^2] t_T - 1/2 (\hat{\lambda}^2 - \hat{\gamma}_0) / \hat{\lambda} (T \hat{\sigma}_{\varepsilon} / s_T)$$

$$T(\hat{\rho}_T - 1) - 1/2 (T^2 \hat{\sigma}^2 \hat{\rho}_T^2 / s_T) (\hat{\lambda}^2 - \hat{\gamma}_0)$$

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: TESTS DICKEY-FULLER CON MAYOR PODER

- Elliott et al. (1996): variación del estimador MCG que consiste en pre-filtrar la variable de interés, es decir usan los residuos de la regresión de $[y_1, y_2 - \alpha y_1, \dots, y_T - \alpha y_{T-1}]$ en $[z_1, z_2 - \alpha z_1, \dots, z_T - \alpha z_{T-1}]$, donde z_t puede incluir constante o constante + tendencia determinística.
- Variable cuasi-diferenciadas usando $\alpha = 1 + \bar{c}T^{-1}$ donde $\bar{c} \in (-\infty, 0)$ es un parámetro que se elige considerando el poder de la envolvente (típicamente -7 si solo se incluye una constante en z_t y -13.5 si el test incluye constante y tendencia).
- Se hace un test de Dickey-Fuller estándar sobre la variable cuasi-diferenciada.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: TESTS DICKEY-FULLER CON MAYOR PODER

- Elliott (1999) señala que en el test anterior se supone que la primera observación es fija, siendo preferible suponer que hay estacionariedad en covarianzas.

- Por ello, si el modelo es de la forma

$$y_t = \alpha z_t + \rho y_{t-1} + \epsilon_t$$

- las desviaciones con respecto a la tendencia son una variable aleatoria con media cero y varianza no-condicional $\sigma^2(1-\rho^2)^{-1}$

- El test se hace usando los residuos de la regresión de

$$[(1-\alpha^2)^{1/2} y_1, y_2 - \alpha y_1, \dots, y_T - \alpha y_{T-1}] \text{ en } [(1-\alpha^2)^{1/2} z_1, z_2 - \alpha z_1, \dots, z_T - \alpha z_{T-1}]$$

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS:

FAMILIA DE TESTS DE PERRON PARA QUIEBRE ESTRUCTURAL

- Perron (1989) observa que la hipótesis nula de los tests anteriores supone que el proceso determinístico no cambia.
 - Pero puede haber quiebres estructurales en la media de la serie y/o en su tendencia, que pueden “engañar” al test de raíces unitarias, sesgando los resultados hacia la hipótesis nula de no estacionariedad.
- Un test general considera la hipótesis alternativa con variables mudas (*dummies*) para capturar los posibles quiebres. El test con ambas hipótesis anidadas es:

$$\Delta y_t = \alpha + \alpha_1 D_t^1 + \beta t + \beta_1 D_t^2 t + \rho y_{t-1} + \epsilon_t + \sum_{p=1}^k \theta_p \delta_{t-p}$$

- La distribución de los estimadores, en este caso, depende además de la ubicación del quiebre en la muestra. Hay tablas Perron.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS:

FAMILIA DE TESTS DE PERRON PARA QUIEBRE ESTRUCTURAL

- Endogeneizando la ubicación del quiebre.
 - El test de Perron para quiebre estructural requiere conocer la fecha en que ocurre el quiebre. Ello no es común. Banerjee et al. (1992) y Zivot y Andrews (1992) proponen tests en que, aparte de determinar la presencia de raíces unitarias, se determina la existencia y fecha de quiebres estructurales.
- Determinación del número de rezagos
 - La determinación del número de rezagos (k) no es inocente.
 - ¿bastaría con hacer que el residuo de la regresión? El problema radica en que el número de rezagos incluidos en la regresión afecta el poder del test. Alternativamente, el tamaño del test no es 5% si el número de rezagos es el incorrecto.
 - Ng y Perron (1995) muestran que si residuos siguen un $AR(p)$ el test RU tiene el tamaño correcto si el número de rezagos incluidos es mayor que p .

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS:

FAMILIA DE TESTS DE PERRON PARA QUIEBRE ESTRUCTURAL

- Determinación del número de rezagos
 - Desafortunadamente, para residuos que siguen un MA(q) no hay una regla útil. Schwert (1989) propone usar una regla dependiente del tamaño de muestra, del tipo $k = \text{Trunc}(c(T/100)^{1/d})$ y sugiere usar $c = 12$ y $d = 4$. El problema de este tipo de regla es que no hay una razón clara para esperar que esta regla sea óptima para todo AR(p) o MA(q).
 - Es común fijar k usando criterios informacionales como Akaike o Schwartz.
 - Alternativamente se puede partir de un número máximo de rezagos – kmax – y luego ir reduciendo el orden del rezago si el marginal no es estadísticamente significativo. De esta manera se procede de general a particular.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS:

TESTS CON HIPÓTESIS NULA DE ESTACIONARIEDAD

- Debido a los problemas de poder y tamaño de los test de raíces unitarias que tienen como hipótesis nula no estacionariedad, ha habido interés en desarrollar tests con la hipótesis nula opuesta, es decir, estacionariedad.
- Este tipo de test es menos común pero aún así es útil usarlo con el objetivo de comprobar las conclusiones del primer tipo de test.
- Uno de los tests más populares es el de Kwiatkosvky et al. (1992), usualmente llamado KPSS. Considere el siguiente modelo:

$$y_t = \delta t + \zeta_t + \varepsilon_t$$

donde ζ es un paseo aleatorio $\zeta_t = \zeta_{t-1} + \mu_t$ con $\mu_t \rightarrow N(0, \sigma^2)$

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: TESTS CON HIPÓTESIS NULA DE ESTACIONARIEDAD

- La hipótesis nula de estacionariedad corresponde a $H_0: \sigma^2 = 0$ es decir que ζ_t es constante. Es una versión del modelo de parámetros cambiantes de Nabeya y Tanaka (1988)

$$y_t = \beta_t x_t + \gamma z_t + \varepsilon_t$$
$$\beta_t = \beta_{t-1} + \mu_t$$

- Nabeya y Tanaka proponen el test $LM = \sum_t \frac{s_t^2}{\hat{\sigma}_e^2}$, donde e son los residuos de una regresión de y_t en una constante y tendencia lineal, $\hat{\sigma}_e^2$ es la varianza residual de dicha regresión, y $S_t = \sum_t e_t$. Este test fue derivado suponiendo que las innovaciones son i.i.d.

TESTS EMPÍRICOS ESTÁNDARES DE RAÍCES UNITARIAS: TESTS CON HIPÓTESIS NULA DE ESTACIONARIEDAD

- KPSS extiende el test para el caso en que las innovaciones no son i.i.d., lo que requiere usar un estimador de largo plazo de la varianza residual

$$\sigma^2 = \lim_{T \rightarrow \infty} T^{-1} E(S_T^2)$$

- Un estimador consistente de S_T^2 es:

$$\bar{S}_T = T^{-1} \sum_{t=1}^T e_t^2 + 2T^{-1} \sum_{\tau=1}^l w_{\tau l} \sum_{t=\tau+1}^T e_t e_{t-\tau}$$

- Donde $w_{\tau l} = 1 - \frac{\tau}{l+1}$

EQUIVALENCIA OBSERVACIONAL, PODER Y TESTS DE RAÍCES UNITARIAS

- Un problema común a cualquier test de raíz unitaria es el bajo poder derivado de la equivalencia observacional. Supongamos que el verdadero modelo es

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

- pero estimamos un AR(1) en una muestra finita de tamaño T).

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

- Consideremos ahora la predicción y ECM de ambos modelos.

	Predicción	ECM
Verdadero Modelo	$\hat{y}_{t+s t} = y_t$	$E(y_{t+s} - \hat{y}_{t+s t})^2 = s\sigma^2$
Modelo Estimado	$\hat{y}_{t+s t} = \rho^s y_t$	$E(y_{t+s} - \hat{y}_{t+s t})^2 = (1 + \rho^2 + \rho^4 + \dots)\sigma^2$

EQUIVALENCIA OBSERVACIONAL, PODER Y TESTS DE RAÍCES UNITARIAS

- Similarmente, si el VM es estacionario $y_t = \varepsilon_t$ y estimamos el modelo en primeras diferencias $(1 - L)y_t = (1 - \theta)\varepsilon_t$
- Consideremos ahora la predicción y ECM de ambos modelos.

	Predicción	ECM
Verdadero Modelo	$\hat{y}_{t+s t} = 0$	$E(y_{t+s} - \hat{y}_{t+s t})^2 = \sigma^2$
Modelo Estimado	$\hat{y}_{t+s t} = (1 - \theta)(\varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_1)$	$E(y_{t+s} - \hat{y}_{t+s t})^2 = (1 + (s - 1)(1 - \theta))\sigma^2$

- Si $\theta \rightarrow 1$, puede haber equivalencia observacional.

EQUIVALENCIA OBSERVACIONAL, PODER Y TESTS DE RAÍCES UNITARIAS

- La razón de la equivalencia observacional es que, en una muestra finita, los procesos que tienen mucha memoria suelen asemejarse a aquellos que tienen raíces unitarias porque no hay suficiente información para determinar si las innovaciones son transitorios o permanentes.
- ¿Qué efecto tiene esto sobre los tests de raíces unitarias?
 - Pérdida de poder
 - Búsqueda de tests no paramétricos.

TESTS NO PAREMÉTRICOS DE RAÍCES UNITARIAS

- Un test clásico no es el de Cochrane (1988).
- El argumento es bastante simple y se basa en evaluar qué tan persistentes son las innovaciones del proceso estocástico.
 - Si ellas son completamente persistentes, entonces existe una raíz unitaria.
 - Si ellas desaparecen en el tiempo, el proceso es estacionario en tendencia.
- Medimos la persistencia de las innovaciones usando el hecho que toda variable cuya primera diferencia es un proceso estacionario lineal, puede ser representada por una combinación de procesos estacionarios (c_t) y camino aleatorio (z_t), i.e., usamos la descomposición de Beveridge-Nelson.

TESTS NO PAREMÉTRICOS DE RAÍCES UNITARIAS

$$\begin{aligned}y_t &= \zeta_t + c_t \\ \zeta_t &= \mu + \zeta_{t-1} + a(1)\epsilon_t \\ a(1) &= \sum_{j=0}^{\infty} a(j) \quad c_t = a^*(L)\epsilon_t \quad a_k^* = - \sum_{j=k+1}^{\infty} a(j)\end{aligned}$$

- donde ϵ es ruido blanco, $a(1)$ es la parte de las innovaciones que no se disipa en el largo plazo, y μ es el deriva de la variable y_t .
- Cualquiera sea la forma en que se separen componentes transitorios y permanentes, la varianza de este ultimo es la misma y corresponde a:

$$\sigma_{\Delta z}^2 = \left(\sum a_j \right)^2 \sigma_{\epsilon}^2$$

TESTS NO PAREMÉTRICOS DE RAÍCES UNITARIAS

- Se testea usando la razón entre la varianza de z y aquella de la primera diferencia de y :

$$V = \frac{\sigma_{\Delta z}^2}{\sigma_{\Delta y}^2}$$

- Si es ET, $\forall \varepsilon_t$ no hay efecto permanente y $\sigma_{\Delta z}^2$ y V tenderán a cero.
- Si es ED, $\forall \varepsilon_t$ hay efectos permanentes que $\sigma_{\Delta z}^2 \approx \sigma_{\Delta y}^2$ y $V=1$.
- En la práctica

$$V = \frac{1}{k} \frac{\sigma_{(y_t - y_{t-k})}^2}{\sigma_{(y_t - y_{t-1})}^2}$$

TESTS NO PAREMÉTRICOS DE RAÍCES UNITARIAS

- Se puede demostrar que V es equivalente a la suma de las autocorrelaciones de la serie más uno. En muestras finitas se usa el estimador de la autocorrelación de la primera diferencia de la serie:

$$V^m = 1 + 2 \left[\sum_{j=1}^{m-1} \frac{m-j}{m} \hat{\rho}_{\Delta Y_j} \right]$$
$$\hat{\rho}_{\Delta Y_j} = \frac{T}{T-j} \frac{\text{cov}(\Delta Y_t, \Delta Y_{t-j})}{\text{var}(\delta Y_t)}$$

- Los términos $\frac{T}{T-j}$ y $\frac{m-j}{m}$ son correcciones por tamaño de muestra.

TESTS DE RAÍCES UNITARIAS EN COMPONENTE ESTACIONAL

- Se puede extender el análisis a los componentes estacionales.
- Largo e inconcluso debate sobre la conveniencia de modelar usando variables originales o desestacionalizadas.
 - Hansen y Sargent (1993) agentes toman decisiones usando datos sin desestacionalizar; remover la estacionalidad implica eliminar información útil para estimar parametrizar.
 - Sims (1993) es preferible usar datos desestacionalizados; al ajustar un modelo a los datos originales se da demasiada los efectos estacionales –que son secundarios– en desmedro de una mejor parametrización de las características principales del modelo.
- La mayoría de los investigadores prefiere usar datos desestacionalizados.
- Justificación: si bien estos movimientos son anticipables, no están directamente relacionados al objetivo del estudio y, por lo tanto, no resulta necesario modelarlos explícitamente.

TESTS DE RAÍCES UNITARIAS EN COMPONENTE ESTACIONAL

- Métodos más comunes para remover la estacionalidad:
 - variables mudas estacionales,
 - método de promedios móviles –desde “variación en x-periodos” hasta ARIMA X-12.
- Esta última es la metodología más popular tanto porque su nivel de desarrollo analítico es superior, como porque viene implementada como una opción en muchos programas econométricos.
- Este tipo de filtro asume que la estacionalidad es estocástica pero estacionaria.
- Pero si los componentes estacionales de una variable son estocásticos, es posible que tengan una raíz unitaria en los efectos estacionales.
- No reconocer la existencia raíces unitarias estacionales puede originar problemas de correlaciones espurias y parametrizaciones inestables.

TESTS DE RAÍCES UNITARIAS EN COMPONENTE ESTACIONAL

- Si las variables son estacionalmente integradas (shocks estacionales con componentes permanentes), desestacionalizarlas con métodos clásicos es inadecuado.
- La remoción de estacionalidad afecta el poder de los tests de raíces unitarias (y de cointegración).
- Si la estacionalidad es determinística, removerla con la ayuda de variables estacionales no tiene efecto alguno sobre el desempeño o resultado de los tests de raíz unitaria.
- Sin embargo, en presencia de efectos estacionales estocásticos, los filtros tradicionales pueden afectar de manera significativa el poder de los tests de raíz unitaria. Ghysels (1990), Olekalns (1994), Abeysinghe (1991).

TESTS DE RAÍCES UNITARIAS EN COMPONENTE ESTACIONAL

- Tests de raíces unitarias en el componente estacional:
 - Paramétricos: Dickey et al. (1984), Hylleberg, Engle, Granger y Yoo (1990) y Johansen y Schaumburg (1999).
 - No paramétricos: Geweke y Porter-Hudak (1983)
 - Bayesianos: Franses et al. (1997)
- Veremos la extensión del test de Dickey-Fuller para datos trimestrales (Hylleberg et al., 1990)
- Test HEGY es preferido porque procede de general a específico, tiende a ser más robusto en presencia de raíces unitarias no estacionales en las variables, y tiene mayor poder que especificaciones alternativas.

TESTS DE RAÍCES UNITARIAS EN EL COMPONENTE ESTACIONAL

- El crecimiento anual en frecuencia trimestral se descompone en:

$$(1 - L^4) = (1 - L)(1 + L)(1 - iL)(1 + iL)$$

donde $i = \sqrt{-1}$

- El test puede ser escrito como:

$$\begin{aligned} & \Delta_4 y_t \\ &= \sum_{i=1}^p \alpha_i D_{st} + \gamma T_t + \pi_1 y_{1t-1} + \pi_2 y_{2t-1} + \pi_3 y_{1t-2} + \pi_4 y_{3t-1} \\ &+ \sum_{k=1}^p \varphi_i \Delta_4 y_{t-k} + \varepsilon_t \end{aligned}$$

Donde

$$\begin{aligned} y_{1t} &= (1 + L + L^2 + L^3)y_t \\ y_{2t} &= -(1 - L + L^2 - L^3)y_t \\ y_{3t} &= -(1 - L^2)y_t \end{aligned}$$

TESTS DE RAÍCES UNITARIAS EN EL COMPONENTE ESTACIONAL

$$\begin{aligned} \Delta_4 y_t &= \sum_{i=1}^4 \alpha_i D_{st} + \gamma T_t + \pi_1 y_{1t-1} + \pi_2 y_{2t-1} + \pi_3 y_{1t-2} + \pi_4 y_{3t-1} \\ &+ \sum_{k=1}^p \varphi_i \Delta_4 y_{t-k} + \varepsilon_t \end{aligned}$$

- Si
 - $\pi_1 = 0$, la variable y_t tiene una raíz unitaria en la frecuencia anual.
 - $\pi_2 = 0$, la variable y_t tiene una raíz unitaria en la frecuencia semestral.
 - $\pi_3 = \pi_4 = 0$, la variable y_t tiene raíces unitarias en la frecuencia trimestral.
- Se debe modelar
 - $(1-L) y_t$ si $\pi_1 = 0$
 - $(1+L) y_t$ si $\pi_2 = 0$
 - $(1+L^2) y_t$ si $\pi_3 = \pi_4 = 0$

TESTS DE RAÍCES UNITARIAS EN EL COMPONENTE ESTACIONAL

- Bajo la hipótesis nula que hay raíces unitarias, los tests de hipótesis de los parámetros obtenidos no tienen la distribución estándar normal, por lo que deben usarse los valores críticos tabulados por HEGY.
- Éstos, a su vez, dependen de si la hipótesis nula incorpora regresores de intercepto, tendencia determinística y estacionalidad determinística.
- Como el test se basa en la hipótesis de no estacionariedad en los distintos componentes, su poder es limitado para distinguir procesos estacionarios con autocorrelación cercana, pero distinta, a uno.
- La inclusión de rezagos de la variable dependiente, sin embargo, permite controlar la eventual correlación de residuos y aumenta el poder del test.