



Tarea 4.

Oscar Herrera
Segundo Semestre de 2024.

Pregunta 1.1.a

De acuerdo a Engle & Granger (1987), la cointegración es descrita de la siguiente forma al inicio de su paper de la siguiente forma:

“Si cada elemento de un vector de series temporales \mathbf{x}_t alcanza la estacionariedad después de diferenciarse, pero una combinación lineal $\boldsymbol{\alpha}'\mathbf{x}_t$ ya es estacionaria, se dice que las series temporales \mathbf{x}_t están co-integradas con el vector de co-integración $\boldsymbol{\alpha}$. Puede haber varios de estos vectores de co-integración, de modo que $\boldsymbol{\alpha}$ se convierte en una matriz. Interpretando $\boldsymbol{\alpha}'\mathbf{x}_t = 0$ como un equilibrio de largo plazo, la co-integración implica que las desviaciones del equilibrio son estacionarias, con varianza finita, aunque las series en sí mismas no sean estacionarias y tengan varianza infinita.”

En palabras simples, sea un vector con dos o más variables $I(1)$, por ejemplo $\mathbf{x}_t = \begin{bmatrix} y_t \sim I(1) \\ z_t \sim I(1) \end{bmatrix}$, si existen combinaciones lineales de estas variables que den como resultado series estacionarias, entonces existe un vector de cointegración entre las variables.

Para explicar la metodología de test de hipótesis los autores comienzan ejemplificando con el siguiente modelo de cointegración:

$$\begin{aligned} x_{1t} + \beta x_{2t} &= u_{1t}, & u_{1t} &= u_{1t-1} + \varepsilon_{1t}, \\ x_{1t} + \alpha x_{2t} &= u_{2t}, & u_{2t} &= \rho u_{2t-1} + \varepsilon_{2t}, & |\rho| < 1. \end{aligned}$$

donde la hipótesis nula sería $\rho = 1$. En el caso de cumplirse la hipótesis nula, no habría co-integración.

Si $\boldsymbol{\alpha}$ fuera conocido, entonces el test se construiría sugiendo Dickey & Fuller (1979), en cuyo caso la distribución sería la distribución no estándar computada por Montecarlo por Dickey (1976). Si $\boldsymbol{\alpha}$ no es conocido, debe ser estimado con los datos, pero si $\rho = 1 \implies \boldsymbol{\alpha}$ no es identificable. Por ende, sólo si existe cointegración $\boldsymbol{\alpha}$ se puede

estimar α (con regresiones). Sin embargo, para poder testear, paso requerido en cualquier análisis estadístico, se debe construir la hipótesis nula de alguna forma. Es por esto, que en el paper proponen un set de siete estadísticos “especiales” para testear la nula de no co-integración, contra la alternativa de co-integración.

Para motivos de *testing*, los autores distinguen en dos casos: el caso en que el se sabe que el sistema es de orden 1 (1 lag) y el caso en que no.

En el primer caso, sea $\delta = \frac{1-\rho}{\alpha-\beta}$, diferenciando $x_{1t} + \beta x_{2t} = u_{1t}$ y $x_{1t} + \alpha x_{2t} = u_{2t}$ (casos de orden 1), se obtiene:

$$\begin{aligned}\Delta x_{1t} &= \beta \delta x_{1t-1} + \alpha \beta \delta x_{2t-1} + \eta_{1t}, \\ \Delta x_{2t} &= -\delta x_{1t-1} - \alpha \delta x_{2t-1} + \eta_{2t}\end{aligned}$$

donde η 's son combinaciones lineales de ε 's. Luego, la representación para la corrección de errores está dada por:

$$\begin{aligned}\Delta x_{1t} &= \beta \delta z_{t-1} + \eta_{1t} \\ \Delta x_{2t} &= -\delta z_{t-1} + \eta_{2t}\end{aligned}$$

Finalmente, con correlato directo en la expresión anterior, la hipótesis nula estará dada por:

$$\begin{aligned}\Delta y_t &= \varepsilon_{1t}, \\ \Delta x_t &= \varepsilon_{2t}, \\ \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} &\sim N(0, \Omega) \implies \rho = 1, \delta = 0\end{aligned}$$

En el segundo caso, se permite que el sistema sea de cualquier orden y no necesariamente tiene que ser de orden 1. Sin embargo, se impone el supuesto de que el sistema sea estacionario en primera diferencia. Para este caso se enlistan una serie de “*augmented*” tests, los cuales se comportan asintóticamente como Dickey & Fuller para el caso univariado. Para estos tests, se pueden calcular valores críticos por simulaciones de Montecarlo.

Los tests propuestos son los siguientes:

- CRDW: Co-integrating Regression Durbin-Watson statistic.
- DF: Dickey-Fuller test applied to residuals.
- ADF: Augmented Dickey-Fuller test with additional lags to account for autocorrelation.

- RVAR: Restricted Vector Autoregression.
- ARVAR: Augmented Restricted Vector Autoregression.
- UVAR: Unrestricted Vector Autoregression.
- AUVAR: Augmented Unrestricted Vector Autoregression.

Luego, la motivación y los casos particulares en los que es recomendable usar cada test se encuentran en la página 266, mientras que la especificación formal de cada test se encuentra en la Table I, página 268.

En resumen, la metodología para probar la hipótesis de cointegración de Engle & Granger se puede definir como:

- Identificar si las series involucradas son $I(1)$.
- Identificar si el VAR en cuestión corresponde a un $VAR(1)$ o a uno en que sólo se puede establecer estacionariedad en primera diferencia.
- En el primer caso, testear estacionariedad del residuo de cointegración.
- En el segundo caso, escoger entre uno de los tests propuestos por los autores.
- Una vez escogido el test, usar las tablas de los valores críticos de los autores (Table II, pg. 269).

Pregunta 1.1.b

Sea el siguiente modelo de Vector de Corrección de Errores (VECM):

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\delta} t + \boldsymbol{\varphi} (\boldsymbol{\beta}' \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t + \boldsymbol{\tau} t) + \boldsymbol{\theta}_1 \Delta \mathbf{y}_{t-1} + \cdots + \boldsymbol{\theta}_p \Delta \mathbf{y}_{t-p} + \boldsymbol{\mu}_t,$$

donde $\boldsymbol{\alpha}_0$, $\boldsymbol{\delta}$, $\boldsymbol{\varphi}$, $\boldsymbol{\beta}$, y $\boldsymbol{\tau}$ son constantes, y $\boldsymbol{\mu}_t$ y $\boldsymbol{\varepsilon}_t$ son ruidos blancos.

Desarrollando la expresión del modelo:

$$\mathbf{y}_t - \mathbf{y}_{t-1} = \boldsymbol{\alpha}_0 + (\boldsymbol{\delta} + \boldsymbol{\tau}) t + \boldsymbol{\varphi} \boldsymbol{\varepsilon}_t + \boldsymbol{\varphi} \boldsymbol{\beta}' \mathbf{y}_{t-1} + \boldsymbol{\theta}_1 \Delta \mathbf{y}_{t-1} + \cdots + \boldsymbol{\theta}_p \Delta \mathbf{y}_{t-p} + \boldsymbol{\mu}_t,$$

$$\mathbf{y}_t = \boldsymbol{\alpha}_0 + (\boldsymbol{\delta} + \boldsymbol{\tau}) t + \boldsymbol{\varphi} \boldsymbol{\varepsilon}_t + (\mathbf{I} + \boldsymbol{\varphi} \boldsymbol{\beta}') \mathbf{y}_{t-1} + \boldsymbol{\theta}_1 \Delta \mathbf{y}_{t-1} + \cdots + \boldsymbol{\theta}_p \Delta \mathbf{y}_{t-p} + \boldsymbol{\mu}_t,$$

Como sabemos que el sistema cointegra, entonces se debe cumplir que $\boldsymbol{\varphi} \boldsymbol{\beta}' \mathbf{y}_{t-1} = \mathbf{z}_{t-1}$ con estacionario. Esto, debido al Teorema de Representación de Granger: "Si un vector multivariado X_t tiene una representación para la corrección de errores si y sólo si es co-integrado" (Ogaki, 1998). Por lo tanto, como $\Delta \mathbf{y}_t$ representa el VECM, se cumple.

Luego, como Hansen (2005) prueba, el Teorema de Represantación de Granger también establece que un vector de co-integración autorregresivo se puede descomponer en: un random walk, un proceso estacionario, un componente

determinístico (tendencia determinística) y un término de condiciones iniciales. Es decir, se puede representar de la forma:

$$\mathbf{X}_t = \mathbf{C} \sum_{s=1}^t \boldsymbol{\varepsilon}_s + \mathbf{C}(L)\boldsymbol{\varepsilon}_t + \boldsymbol{\tau}(t) + \mathbf{C}(\mathbf{X}_0 - \boldsymbol{\Gamma}_1\mathbf{X}_{-1} - \cdots - \boldsymbol{\Gamma}_{k-1}\mathbf{X}_{-k+1})$$

Como es posible notar la expresión obtenida para el *VAR* en niveles es consistente con lo anterior, por lo tanto, **la expresión derivada conforma una expresión para el modelo en niveles, y este tiene tendencia lineal.**

Pregunta 1.1.c

La expresión es falsa. Se han identificado dos formas de demostrarlo: *i*) Encontrar el proceso representativo de un *ARIMA*(1, 1, 0), o *ii*) Encontrar el orden *ARIMA* que representa un *random walk with drift*. Por simplicidad, se adopta el segundo camino.

Un *random walk with drift model* está dado por:

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

Lo anterior tiene la siguiente forma recursiva (Enders, 2014):

$$y_t = y_0 + a_0 + \sum_{i=1}^t \varepsilon_i$$

Se puede apreciar que el componente *MA* de la anterior expresión posee raíz unitaria, por lo que el proceso al menos tiene orden de integración igual a 1.

Tomando primera diferencia:

$$\Delta y_t = a_0 + \varepsilon_t$$

Lo anterior corresponde a un proceso *w.n.* alrededor de una constante, por lo que por un lado el proceso es estacionario $\implies \sim I(1)$, y además, el ruido blanco con constante representa un *ARMA*(0, 0) (Enders, 2014). En definitiva, el *random walk with drift model* es un $ARIMA(0, 1, 0) \neq ARIMA(1, 1, 0)$.

Pregunta 1.2.i

Pregunta 1.2.i.a

Sea el *VAR*:

$$y_t = \alpha_{10} + \alpha_{11}y_{t-1} + \alpha_{12}z_{t-1} + \varepsilon_t^y$$

$$z_t = \alpha_{20} + \alpha_{21}y_{t-1} + \alpha_{22}z_{t-1} + \varepsilon_t^z$$

donde se asume que ε_t^y y ε_t^z son ruidos blancos no correlacionados, y α_{ij} son parámetros fijos.

Desarrollando z_t :

$$\begin{aligned} (1 - \alpha_{22}L)z_t &= \alpha_{20} + \alpha_{21}y_{t-1} + \varepsilon_t^z \\ (1 - \alpha_{22}L)z_{t-1} &= \alpha_{20} + \alpha_{21}y_{t-2} + \varepsilon_{t-1}^z \\ z_{t-1} &= \frac{1}{(1 - \alpha_{22}L)}(\alpha_{20} + \alpha_{21}y_{t-2} + \varepsilon_{t-1}^z) \end{aligned}$$

Luego:

$$\begin{aligned} y_t &= \alpha_{10} + \alpha_{11}Ly_t + \alpha_{12}Lz_t + \varepsilon_t^y \\ (1 - \alpha_{11}L)y_t &= \alpha_{10} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}(\alpha_{20} + \alpha_{21}y_{t-2} + \varepsilon_{t-1}^z) + \varepsilon_t^y \\ &= \alpha_{10} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}(\alpha_{20} + \alpha_{21}L^2y_t + L\varepsilon_t^z) + \varepsilon_t^y \\ &= \alpha_{10} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}\alpha_{20} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}\alpha_{21}L^2y_t + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}L\varepsilon_t^z + \varepsilon_t^y \\ (1 - \alpha_{11}L)y_t - \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}\alpha_{21}L^2y_t &= \alpha_{10} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}\alpha_{20} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}L\varepsilon_t^z + \varepsilon_t^y \\ \left(1 - \alpha_{11}L - \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}\alpha_{21}L^2\right)y_t &= \alpha_{10} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}\alpha_{20} + \alpha_{12}\frac{1}{(1 - \alpha_{22}L)}L\varepsilon_t^z + \varepsilon_t^y \\ \left(\frac{(1 - \alpha_{11}L)(1 - \alpha_{22}) - \alpha_{12}\alpha_{21}L^2}{(1 - \alpha_{22}L)}\right)y_t &= \frac{(1 - \alpha_{22}L)\varepsilon_t^y + (1 - \alpha_{22})\alpha_{10} + \alpha_{12}L\varepsilon_t^z + \alpha_{12}\alpha_{20}}{(1 - \alpha_{22}L)} \\ y_t &= \frac{(1 - \alpha_{22}L)\varepsilon_t^y + (1 - \alpha_{22})\alpha_{10} + \alpha_{12}L\varepsilon_t^z + \alpha_{12}\alpha_{20}}{(1 - \alpha_{11}L)(1 - \alpha_{22}) - \alpha_{12}\alpha_{21}L^2} \end{aligned}$$

■

Pregunta 1.2.i.b

Retomando la expresión para z_t :

$$(1 - \alpha_{22}L)z_t = \alpha_{20} + \alpha_{21}y_{t-1} + \varepsilon_t^z$$

$$z_t = \frac{\alpha_{20} + \alpha_{21}y_{t-1} + \varepsilon_t^z}{1 - \alpha_{22}L}$$

Reemplazando en la expresión para y_t :

$$(1 - \alpha_{11}L)y_t = \alpha_{10} + \alpha_{12}z_{t-1} + \varepsilon_t^y$$

Sustituyendo en la expresión de $y : t$:

$$(1 - \alpha_{22}L)z_t = \alpha_{20} + \alpha_{21} \left(\frac{\alpha_{10} + \alpha_{12}z_{t-1} + \varepsilon_t^y}{1 - \alpha_{11}L} \right) + \varepsilon_t^z$$

$$= \alpha_{20} + \frac{\alpha_{21}\alpha_{10}}{1 - \alpha_{11}L} + \frac{\alpha_{21}\alpha_{12}z_{t-1}}{1 - \alpha_{11}L} + \frac{\alpha_{21}\varepsilon_t^y}{1 - \alpha_{11}L} + \varepsilon_t^z$$

$$= \alpha_{20} + \frac{\alpha_{21}\alpha_{10}}{1 - \alpha_{11}L} + \frac{\alpha_{21}\alpha_{12}Lz_t}{1 - \alpha_{11}L} + \frac{\alpha_{21}\varepsilon_t^y}{1 - \alpha_{11}L} + \varepsilon_t^z$$

$$\left((1 - \alpha_{22}L) - \frac{\alpha_{21}\alpha_{12}L}{1 - \alpha_{11}L} \right) z_t = \alpha_{20}(1 - \alpha_{11}L) + \alpha_{21}\alpha_{10} + \alpha_{21}\varepsilon_t^y + (1 - \alpha_{11}L)\varepsilon_t^z$$

$$z_t = \frac{(1 - \alpha_{11}L)\varepsilon_t^z + (1 - \alpha_{11})\alpha_{20} + \alpha_{21}\alpha_{10} + \alpha_{21}L\varepsilon_t^y}{(1 - \alpha_{11}L)(1 - \alpha_{22}L) - \alpha_{12}\alpha_{21}L^2}$$

■

Pregunta 1.2.i.c

En línea con Enders (2014) (eq. 6.13), para que un vector multivariado tenga representación de error de corrección, debe ser poder expresado como:

$$\Delta x_t = \pi_0 + \pi x_{t-1} + \pi_1 \Delta x_{t-1} + \pi_2 \Delta x_{t-2} + \cdots + \pi_p \Delta x_{t-p} + \varepsilon_t$$

donde el término πx_{t-1} debe ser distinto de cero y además estacionario.

Reexpresando el VAR en forma matricial:

$$y_t = \alpha_{10} + \alpha_{11}y_{t-1} + \alpha_{12}z_{t-1} + \varepsilon_t^y$$

$$z_t = \alpha_{20} + \alpha_{21}y_{t-1} + \alpha_{22}z_{t-1} + \varepsilon_t^z$$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix}$$

Tomando primera diferencia:

$$\begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} = \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \end{bmatrix} + \left(\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \end{bmatrix} + \begin{bmatrix} \alpha_{11} - 1 & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - 1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix}$$

Podemos inferir que la matriz $\begin{bmatrix} \alpha_{11} - 1 & \alpha_{12} \\ \alpha_{21} & \alpha_{22} - 1 \end{bmatrix}$ corresponde al vector de cointegración debido a que se asumen las condiciones (6.19) y (6.21). Reescribiendo, obtenemos que la corrección de errores está presente y el lag del sistema es cero:

$$\Delta y_t = \alpha_{10} + (\alpha_{11} - 1)y_{t-1} + \alpha_{12}z_{t-1} + \varepsilon_t^y$$

$$\Delta z_t = \alpha_{20} + \alpha_{21}y_{t-1} + (\alpha_{22} - 1)z_{t-1} + \varepsilon_t^z$$

En contraste, las ecuaciones (6.23) y (6.24) son:

$$\Delta y_t = -[a_{12}a_{21}/(1 - a_{22})]y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt}$$

$$\Delta z_t = a_{21}y_{t-1} - (1 - a_{22})z_{t-1} + \varepsilon_{zt}$$

La diferencia, tal como lo mencionan los autores en la página 355 luego de plantear el sistema, se debe a que Enders ignora el intercepto por simplicidad.

Pregunta 1.2.ii

Sea x_t un proceso univariado que sigue un (“random walk with drift”):

$$x_t = \delta + x_{t-1} + \mu_t$$

donde μ_t es un ruido blanco con varianza σ_μ^2 . Sea x_0 un valor fijo (no estocástico). Sea, además:

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

donde ε_t es un ruido blanco con varianza σ_ε^2 .

Resolviendo para x_t (exógena):

$$x_t = \delta + x_{t-1} + \varepsilon_t$$

$$x_t = x_{t-2} + \delta + \varepsilon_t + a_0 + \varepsilon_{t-1}$$

$$x_t = x_{t-2} + 2\delta + \varepsilon_t + \varepsilon_{t-1}$$

$$x_t = x_{t-3} + 3\delta + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2}$$

⋮

$$x_t = x_0 + t \cdot \delta + \sum_{i=1}^t \varepsilon_{t-i}$$

Reemplazando de forma directa en y_t se obtiene:

$$\begin{aligned} y_t &= \alpha + \beta x_t + \varepsilon_t \\ &= \alpha + \beta \left(x_0 + t \cdot \delta + \sum_{i=1}^t \varepsilon_{t-i} \right) + \varepsilon_t \\ &= \alpha + \beta x_0 + \beta \delta t + \beta \sum_{i=1}^t \varepsilon_{t-i} + \varepsilon_t \end{aligned}$$

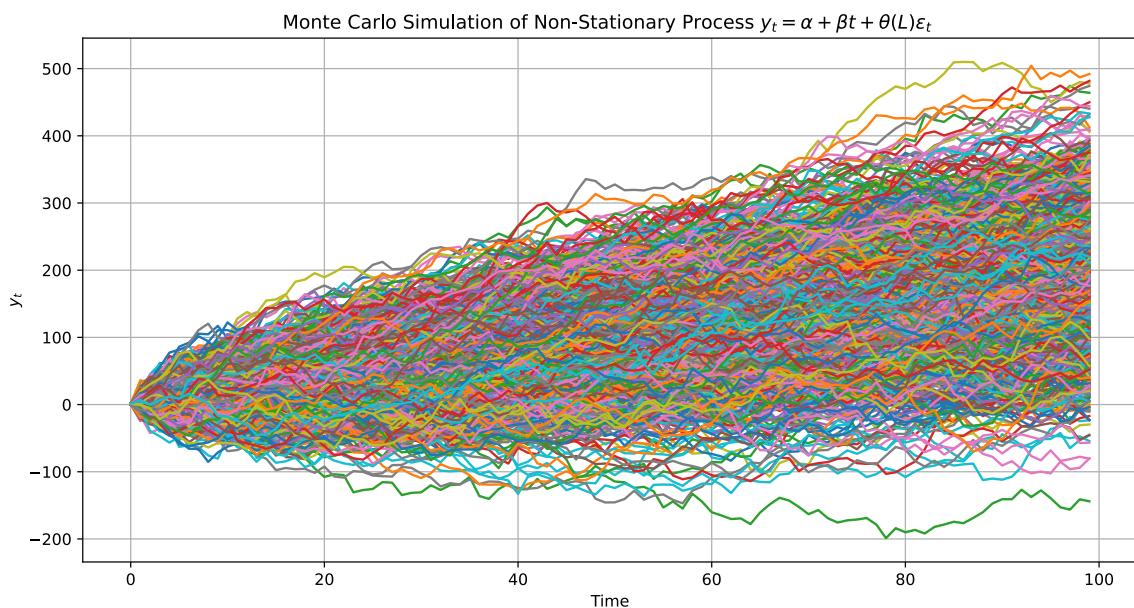
Luego, el cálculo de la esperanza incondicional de está dado por:

$$\begin{aligned}\mathbb{E}[y_t] &= \alpha + \beta x_0 + \beta \delta t + \beta \sum_{i=1}^t \mathbb{E}[\varepsilon_{t-i}] + \mathbb{E}[\varepsilon_t] \\ &= \alpha + \beta x_0 + \beta \delta t\end{aligned}$$

Luego, teniendo en consideración que los shocks no correlacionan (*strict exogeneity*), la varianza del proceso está dada por:

$$\begin{aligned}\text{Var}[y_t] &= \text{Var} \left(\alpha + \beta x_0 + \beta \delta t + \beta \sum_{i=1}^t \varepsilon_{t-i} + \varepsilon_t \right) \\ &= \beta^2 \text{Var} \left(\sum_{i=1}^t \varepsilon_{t-i} \right) + \text{Var}(\varepsilon_t) \\ &= \beta^2 \sum_{i=1}^t \sigma_\varepsilon^2 + \sigma_\varepsilon^2 \\ &= \beta^2 t \sigma_\varepsilon^2 + \sigma_\varepsilon^2 \\ &= (\beta^2 t + 1) \sigma_\varepsilon^2\end{aligned}$$

El proceso tiene esperanza en función de su tendencia principalmente, sin embargo la incertidumbre crecerá a un ritmo de la raíz del tiempo ponderado por β^2 . Como la incertidumbre crece, esto es consistente con un *random walk* en tendencia.



Pregunta 2.a & 2.b

Sean los siguientes modelos:

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

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \varepsilon_t,$$

$$\Delta Y_t = \alpha + \beta t + \phi Y_{t-1} + \varepsilon_t,$$

donde ε_t es un ruido blanco, $\{\alpha, \beta, \phi\}$ son parámetros a ser estimados, y t es una tendencia lineal.

Análogo a como muestra Enders (2014) en la sección *Deterministic and Stochastic Trends (Chapter 4.1)* para el caso univariado, cada uno de estos procesos en primera diferencia representa un comportamiento completamente único bajo la eventual presencia de raíces unitarias, siendo posible que algunas de estas series llegaran a mostrar comportamiento relativamente explosivo.

Es por esto que todos los parámetros son relevantes a la hora de evaluar la naturaleza del proceso mediante inferencia estadística ante la eventual presencia de raíz unitaria, porque en última instancia, también estos son parámetros incidentales afectan el comportamiento de la distribución de la inferencia estadística, como muestran Dickey & Fuller (1979). Esta es la razón por la cual se obtienen distintos valores críticos de la distribución de Dickey-Fuller cuando cambia la especificación del modelo (Hamilton, 1994, pp. 492 & 501).

En el caso del primer modelo, el cual no presenta ni tendencia ni constante, la hipótesis nula correspondería sólo a la presencia de raíz unitaria.

En términos informales, en todos los casos nos es de interés determinar lo siguiente:

H_0 : La serie es de raíz unitaria, i.e., no es estacionaria.

H_1 : La serie es estacionaria.

Notar que si el modelo que genera los datos es un *random walk* $Y_t = \rho \cdot Y_{t-1} + \varepsilon_t = Y_{t-1} + \varepsilon_t$. Tomando primera diferencia obtendríamos $\Delta Y_{t-1} = \varepsilon_t$. Por lo tanto, el testeo de raíz unitaria involucra determinar si el proceso $\Delta Y_t = \phi Y_{t-1} + \varepsilon_t$ es más bien caracterizado por $\phi = 0$.

En el caso en que $\phi \neq 0$ tenemos lo siguiente:

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

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

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

Recordar, a partir de la descripción informal de las hipótesis, lo que nos interesa en última instancia es la determinación de estacionariedad. Si $\phi \neq 0$, para forzar estacionariedad se requeriría:

$$\begin{aligned}
0 &\leq |(\phi + 1)| < 1 \\
0 \leq (\phi + 1) &< 1 \quad \vee \quad 0 \leq -(\phi + 1) < 1 \\
-1 \leq \phi &< 1 \quad \vee \quad 0 \geq (\phi + 1) > -1 \\
-1 \leq \phi &< 1 \quad \vee \quad -1 \geq (\phi + 1) > -2 \\
-1 \leq \phi &< 0 \quad \vee \quad -2 < \phi \leq -1 \\
-2 &< \phi < 0
\end{aligned}$$

Así, el diseño de las hipótesis del test en términos formales sería:

$$\begin{aligned}
H_0: \phi &= 0 \\
H_1: \phi &< 0
\end{aligned}$$

El segundo caso corresponde a un modelo en que la primera diferencia presenta una constante. Como muestra Enders (2014) y como se ha mostrado en preguntas anteriores, este modelo correspondería a un *random walk with drift*: $Y_t = \alpha + Y_{t-1} + \varepsilon_t = \alpha \cdot t + Y_0 + \sum_{i=0}^t \varepsilon_{t-i}$.

$$\begin{aligned}
\Delta Y_t &= \alpha + \phi Y_{t-1} + \varepsilon_t \\
Y_t - Y_{t-1} &= \alpha + \phi Y_{t-1} + \varepsilon_t \\
Y_t &= \alpha + (\phi + 1) Y_{t-1} + \varepsilon_t \\
\implies \text{Estacionariedad} &\Leftrightarrow -2 < \phi < 0, \quad \text{si } \phi \neq 0.
\end{aligned}$$

En este caso, la hipótesis nula involucraría tanto determinar si $\phi = 0$, como si $\alpha = 0$, tal Hamilton (1994, pp. 490 & 494). Así, el diseño de las hipótesis estaría dado por:

$$\begin{aligned}
H_0: \phi &= 0, \alpha = 0 \\
H_1: \phi &< 0, \alpha \neq 0
\end{aligned}$$

Con lo anterior, ya podemos establecer que la hipótesis alternativa no es igual en cada especiación.

El tercer caso corresponde a uno en el que el proceso en niveles presenta un caso especial de tendencia. Suponiendo que $\phi \neq 0$, notar que:

$$\Delta Y_t = \alpha + \beta t + \phi Y_{t-1} + \varepsilon_t$$

$$\begin{aligned}
Y_t - Y_{t-1} &= \alpha + \beta t + \phi Y_{t-1} + \varepsilon_t \\
Y_t &= \alpha + \beta t + (\phi + 1)Y_{t-1} + \varepsilon_t \\
\implies \text{Estacionariedad} &\Leftrightarrow -2 < \phi < 0, \quad \text{si } \phi \neq 0.
\end{aligned}$$

Resolviendo para Y_t :

$$\begin{aligned}
Y_t &= \alpha + \beta t + (\phi + 1)Y_{t-1} + \varepsilon_t \\
Y_t &= \alpha + \beta t + (\phi + 1)(\alpha + \beta t + (\phi + 1)Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
Y_t &= \alpha(1 + (\phi + 1)) + (\beta + \beta(\phi + 1))t + (\phi + 1)^2 Y_{t-2} + (\phi + 1)\varepsilon_{t-1} + \varepsilon_t \\
Y_t &= \alpha(1 + (\phi + 1)) + (\beta + \beta(\phi + 1))t + (\phi + 1)^2 Y_{t-2} + (\phi + 1)\varepsilon_{t-1} + \varepsilon_t \\
Y_t &= \alpha(1 + (\phi + 1)) + (\beta + \beta(\phi + 1))t + (\phi + 1)^2 (\alpha + \beta t + (\phi + 1)Y_{t-2} + \varepsilon_{t-1}) + (\phi + 1)\varepsilon_{t-1} + \varepsilon_t \\
Y_t &= \alpha(1 + (\phi + 1) + (\phi + 1)^2) + (\beta + \beta(\phi + 1) + \beta(\phi + 1)^2)t + (\phi + 1)^3 Y_{t-2} + (\phi + 1)^2 \varepsilon_{t-1} + (\phi + 1)\varepsilon_{t-1} + \varepsilon_t \\
&\vdots \\
Y_t &= (\phi + 1)^t Y_0 + \alpha \cdot \sum_{i=1}^t (\phi + 1)^i + t \cdot \sum_{i=0}^t \beta(\phi + 1)^i + \sum_{i=0}^t (\phi + 1)^i \varepsilon_{t-i}
\end{aligned}$$

En el caso en que $\phi = 0$ (raíz unitaria), se obtiene:

$$\begin{aligned}
Y_t &= Y_0 + \alpha t + t \cdot \sum_{i=0}^t \beta + \sum_{i=0}^t \varepsilon_{t-i} \\
Y_t &= Y_0 + \alpha t + t \cdot (t\beta) + \sum_{i=0}^t \varepsilon_{t-i} \\
Y_t &= Y_0 + \alpha t + t^2 \beta + \sum_{i=0}^t \varepsilon_{t-i}
\end{aligned}$$

A pesar de plantear especificaciones distintas, para cada uno de los modelos podría ser posible testear sólo que $\rho = 1$ asumiendo que todo el resto de parámetros son ceros, tal como plantea Hamilton (1994, p. 494, para. 3), pero en aquel caso no tendría sentido haber planteado las diversas especificaciones en primera instancia. Por ende, cuando se plantean especificaciones que extienden los modelos más simples, “resulta natural testear conjuntamente la presencia de raíz unitaria con la existencia de estos parámetros incidentales” (Hamilton, 1994). Así, las hipótesis en este caso estarían dadas por:

$$H_0: \phi = 0, \alpha = 0, \beta = 0$$

$$H_1 : \phi < 0, \alpha \neq 0, \beta \neq 0$$

Anidación de Hipótesis.

De acuerdo a lo planteado por Soto (2022), la idea de la anidación consistiría en partir del caso más general al más particular. En específico, el procedimiento se resumiría en el siguiente algoritmo:

- Partir por el tercer modelo y testear si β es distinto de cero.

$$\Delta Y_t = \alpha + \beta t + \phi Y_{t-1} + \varepsilon_t$$

$$H_0: \beta = 0$$

$$H_1 : \beta \neq 0$$

- Si no se rechaza dicha hipótesis nula, se estima el segundo modelo.
- Si se rechaza la hipótesis nula, podemos pasar a testear.

$$H_0: \phi = 0, \alpha = 0$$

$$H_1 : \phi < 0, \alpha \neq 0$$

Potencialmente, lo anterior también podría ser descompuesto en una sub-anidación de hipótesis.

- Una vez comprobado que estadísticamente $\beta = 0$, estimar el segundo modelo y testear las siguientes hipótesis:

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

$$H_0: \alpha = 0$$

$$H_1 : \alpha \neq 0$$

- Si no se rechaza dicha hipótesis nula, se estima el primer modelo.
- Si se rechaza la hipótesis nula, podemos pasar a testear.

$$H_0: \phi = 0$$

$$H_1 : \phi < 0$$

- Una vez comprobado que estadísticamente $\alpha = 0$ en el segundo modelo, estimar el primer modelo y testear las siguientes hipótesis:

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

$$H_0: \phi = 0$$

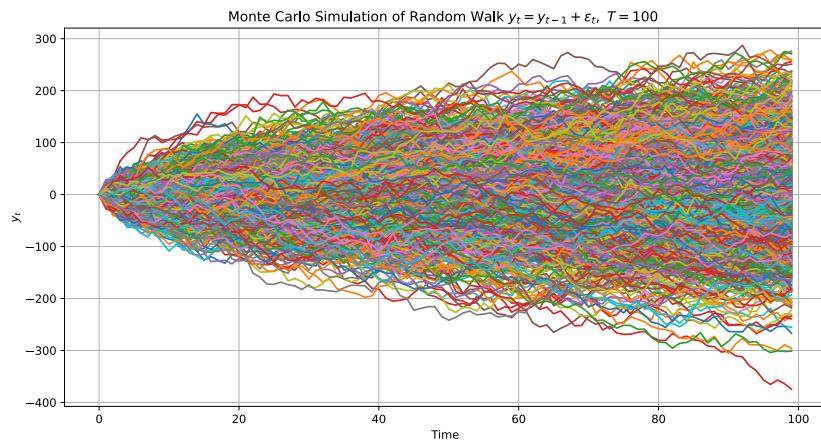
$$H_1: \phi < 0$$

Pregunta 2.c

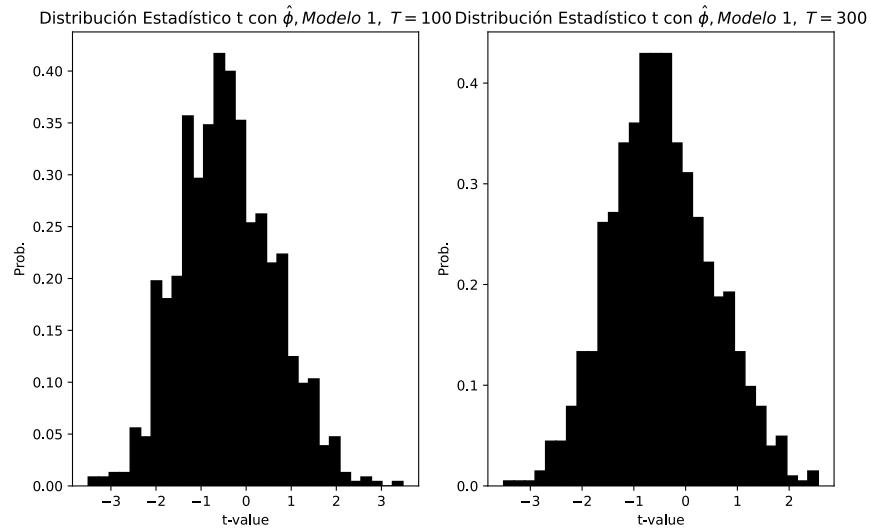
Para esta pregunta se ha usado la hipótesis nula (raíz unitaria, $\phi = 0 \implies \rho = 1$) como *baseline* para generar las simulaciones de Montecarlo que posteriormente se han usado para calcular estadísticos, sobre los cuales se computan distribuciones. La siguiente parametrización ha sido usada en los casos donde los coeficientes existen: $\{Y_0 = 0, \alpha = 1, \beta = 1, \sigma_\varepsilon = 10\}$.

A continuación se presentan las simulaciones de Montecarlo del primer modelo bajo hipótesis nula:

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



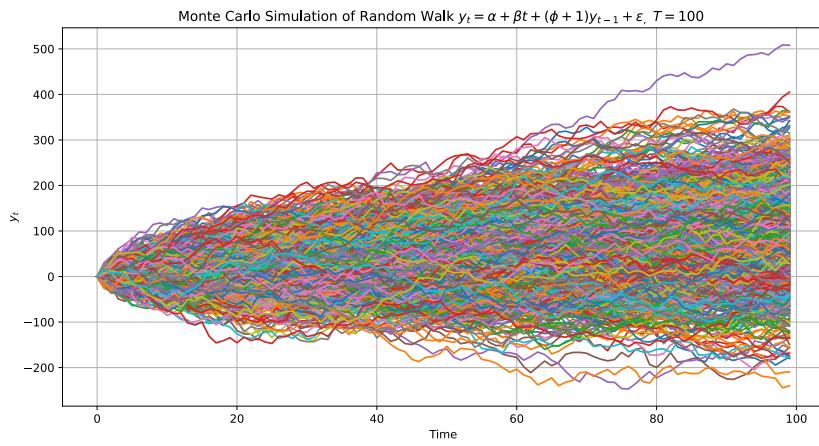
Luego de simular, el cálculo de estadísticos por cada simulación ha brindado la siguiente distribución empírica:



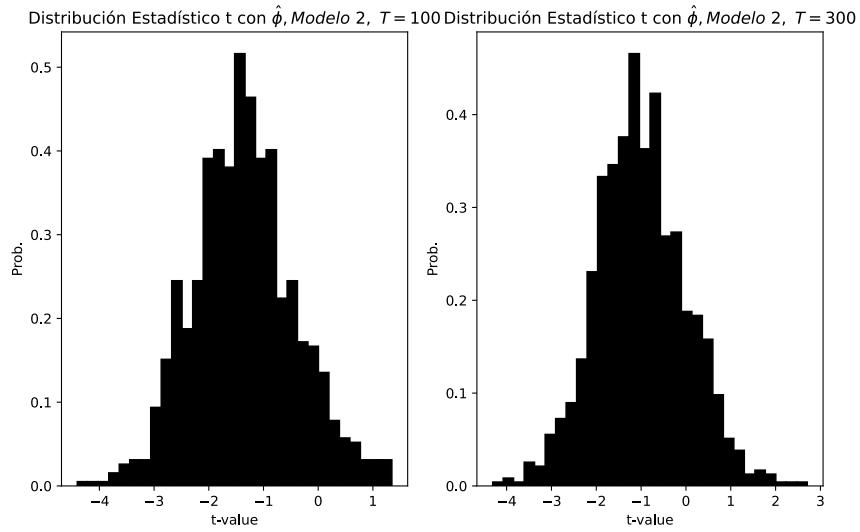
Podemos observar que los valores críticos convencionales llevarían a constantemente establecer que $\phi \neq 0$, cuando en realidad es cero.

A continuación se replica para los modelos 2 y 3.

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \varepsilon_t,$$

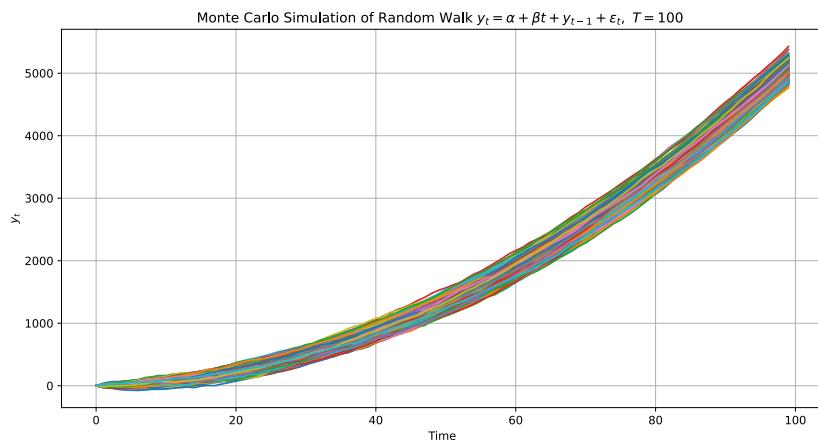


En línea con lo señalado en Enders (2014), visualmente, un *random walk* con deriva y sin constantes es difícil de distinguir de un *standard random walk* a simple vista.

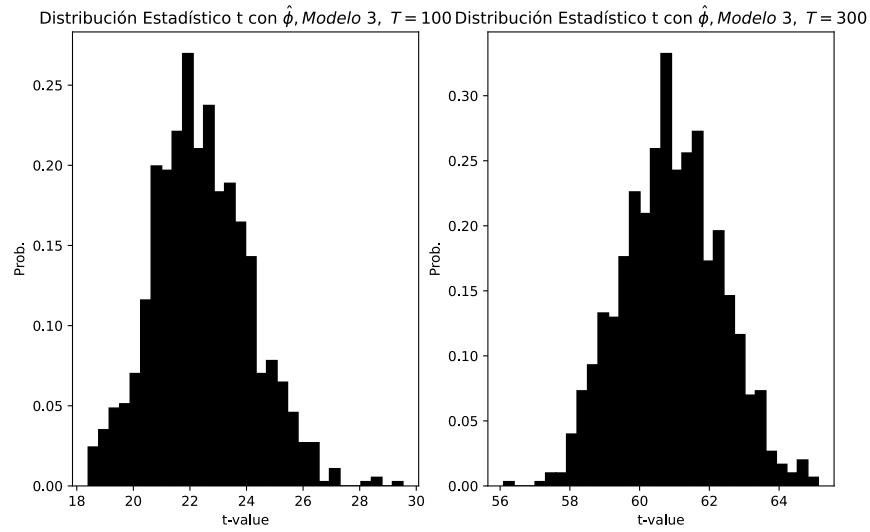


Adicionalmente, la distribución muestra mayor tendencia hacia valores negativos, alejándose aún más de una distribución normal estándar o de una distribución t estándar,

$$\Delta Y_t = \alpha + \beta t + \phi Y_{t-1} + \varepsilon_t,$$



El modelo con tendencia muestra un comportamiento en el que la tendencia es fuertemente dominante por sobre los componentes aleatorios alrededor de esta, incluso cuando se ha usado $\sigma_\varepsilon = 10$.



En definitiva, al usar valores críticos correctos, el último modelo parece expresarse *súper-consistente*.

A continuación se despliegan los valores críticos calculados para las distintas muestras:

Cuadro 1: Critical Values for Models 1 and 3 at T=100 and T=300

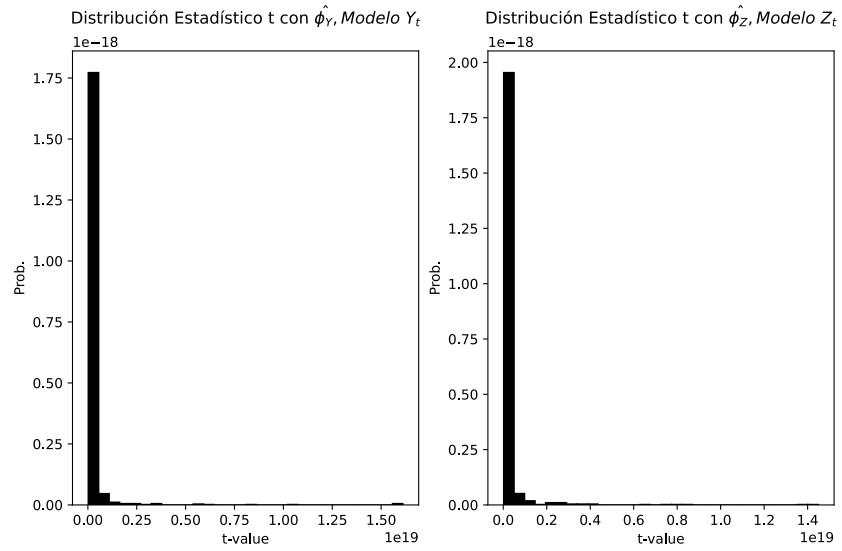
Sample Size (T)	Model	90 % Critical Value	95 % Critical Value
100	Model 1	-2.0065	-1.7303
	Model 2	-2.7740	-2.5625
	Model 3	19.7525	20.3980
300	Model 1	-2.0169	-1.6622
	Model 2	-2.6089	-2.1635
	Model 3	58.5990	59.0395

Se observa que el valor de significancia no altera mucho el valor crítico.

Pregunta 2.d

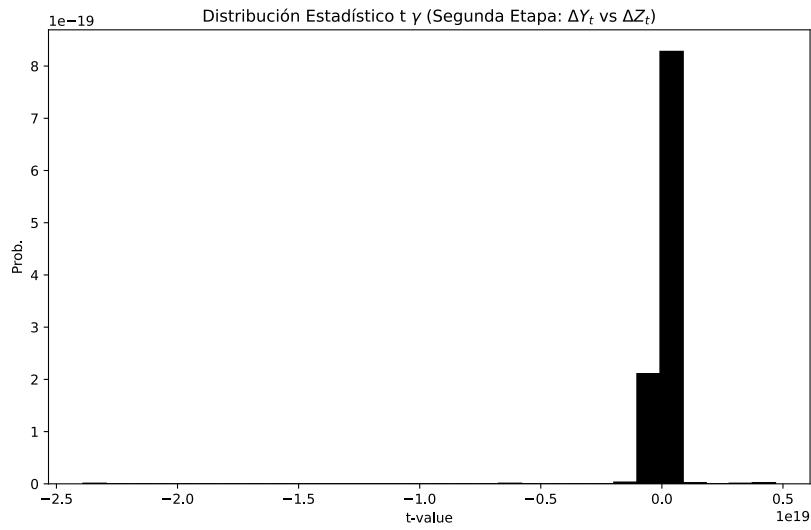
A continuación se despliegan las distribuciones del parámetro individual ϕ_X encontradas de los modelos modelo:

$$\begin{aligned} \Delta Y_t = \phi_Y Y_{t-1} + \varepsilon_t^Y &\iff Y_t = (\phi_Y + 1)Y_{t-1} + \varepsilon_t^Y, \quad \text{con } \phi_Y = 0,975 \\ \Delta Z_t = \phi_Z Z_{t-1} + \varepsilon_t^Z &\iff Z_t = (\phi_Z + 1)Z_{t-1} + \varepsilon_t^Z, \quad \text{con } \phi_Z = 0,975 \end{aligned}$$



A continuación se despliegan la distribución encontrada parámetro γ en el modelo:

$$\Delta Y_t = \gamma \Delta Z_t + \nu_t^Y,$$



Tal cómo establecen Granger, Hyung & Jeon (2001), las distribuciones no convergen a una normal estándar.

Pregunta 3.a

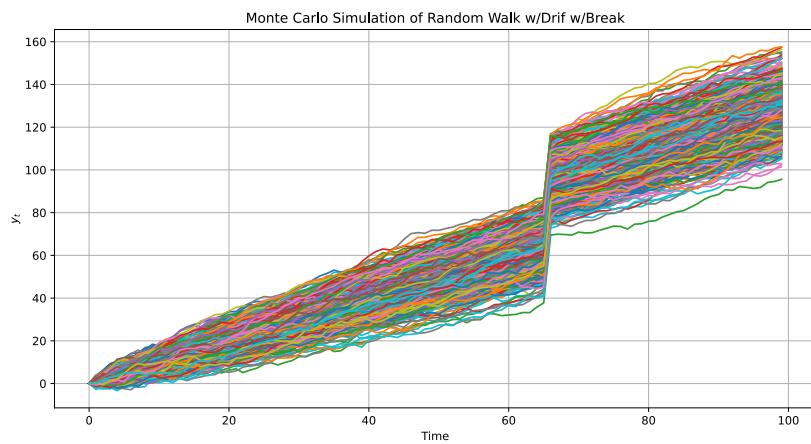
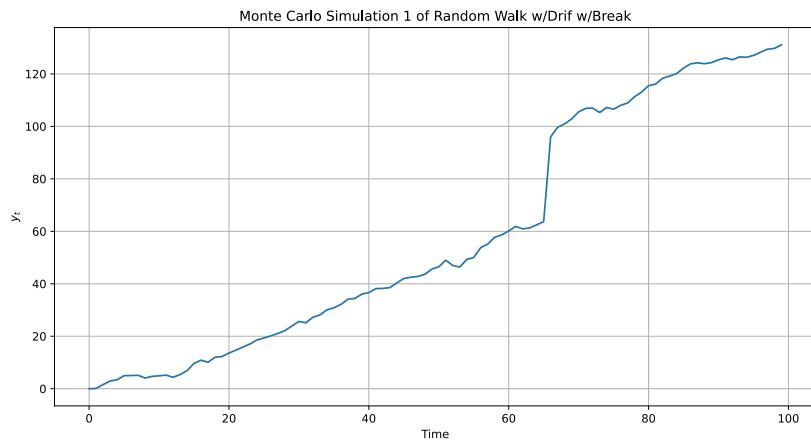
En línea con Enders (2014), *Chapter 4.8 “Structural Change”*, para las simulaciones de esta pregunta se usó el siguiente modelo con la parametrización:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \varepsilon_t + \mu_d D_\tau$$

$$\{\alpha = 1, \phi = 1, Y_0 = 0, \mu_d = 30\}$$

con $D_\tau = \{ \text{if } t = \tau; 0 \text{ else}\}$, la variable *dummy* de pulso y cambio estructural en $\tau = 0,66 \cdot T$.

A continuación, se despliega la primera observación simulada, y luego el conjunto de todas las simulaciones de Montecarlo.



Siguiendo Enders (2014), sección “**Perron’s Test for Structural Change**”, se han planteado las siguientes hipótesis:

$$H_1 : y_t = a_0 + y_{t-1} + \mu_1 D_P + \varepsilon_t$$

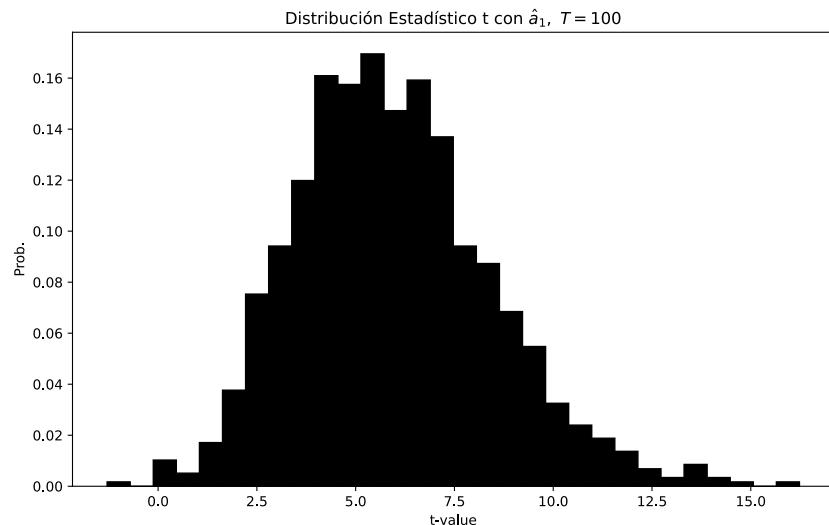
$$A_1 : y_t = a_0 + a_2 t + \mu_2 D_L + \varepsilon_t$$

donde $D_P = D_\tau$, y $D_L = \{1 \text{ if } t \geq \tau; 0 \text{ else}\}$.

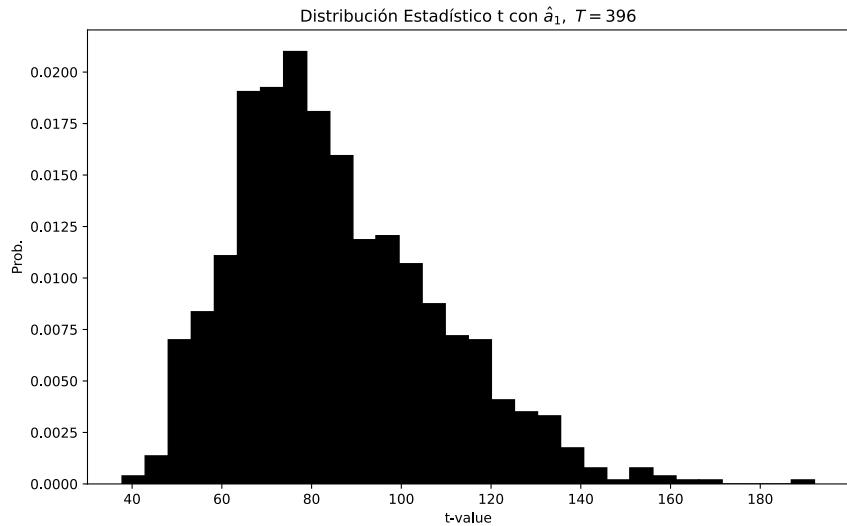
Luego, se implementaron los *Steps* de Enders para el cálculo de valor crítico de Perron, donde se testeó la significancia estadística para cada simulación del coeficiente estimado α_1 en la siguiente regresión aumentada:

$$y_t = a_0 + a_1 y_{t-1} + a_2 t + \mu_1 D_P + \mu_2 D_L + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t, \quad \text{con } p = 1.$$

La distribución del estadístico de α_1 para ambas muestras es la siguiente:



Se observa a continuación que el parámetro incidental T , tiene una gran influencia sobre la distribución, ya que al aumentar el tamaño de la muestra la distribución del estadístico muta considerablemente:



Los valores críticos encontrados son los siguientes:

Cuadro 2: Critical Values for Models 1 and 2 with $T=100$ and $T=396$ respectively

Model	90 % Critical Value	95 % Critical Value
Model 1	2.3344	3.0006
Model 2	54.2948	59.4269

Pregunta 3.b

Como se menciono en clases, las limitantes de esta prueba tienen que ver principalmente con la capacidad de captar quiebres estructurales cercanos al bordes de la muestra. Cuando los quiebres ocurren en un momento intermedio, la disponibilidad de datos para poder capturar los quiebres mediante inferencia no es un problema. Sin embargo, esto es una gran limitante cuando los shocks ocurren al final o al inicio, ya que no se es capaz de distinguir si este corresponde a un shock que se disipará, en el primer caso, o discernir este cambio respecto a los shocks corrientes en el segundo.

Referencias

- Dickey, D. A. (1976). “Estimation and Hypothesis Testing for Nonstationary Time Series.”, PhD. Thesis, Iowa State University, Ames. <https://doi.org/10.31274/rtd-180813-2848>
- Dickey, D. A., & Fuller, W. A. (1979). “Distribution of the Estimators for Autoregressive Time Series With a Unit Root”. *Journal of the American Statistical Association*, 74(366), 427–431. <https://doi.org/10.2307/2286348>
- Enders, W. (2014). “Applied Econometric Time Series, 4th Edition”. Wiley Series in Probability and Statistics. <https://www.wiley.com/en-us/Applied+Econometric+Time+Series,+4th+Edition-p-9781118808566>.
- Engle, R. F., & C. W. J. Granger. “Co-Integration and Error Correction: Representation, Estimation, and Testing.” *Econometrica*, vol. 55, no. 2, 1987, pp. 251–76. JSTOR, <https://doi.org/10.2307/1913236>.
- Granger, C. W. J., Hyung, N., & Jeon, Y. (2001). “Spurious regressions with stationary series. Applied Economics”, 33(7), 899–904. <https://doi.org/10.1080/00036840121734> <https://doi.org/10.1080/00036840121734>
- Hamilton, J. (1994), “Time Series Analysis”. Princeton University Press.
- Hansen, P. R. (2005). “Granger’s representation theorem: A closed-form expression for I(1) processes”. *The Econometrics Journal*, 8(1), 23–38. <http://www.jstor.org/stable/23114965>
- Perron, P. (1989). “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis”. *Econometrica*, 57(6), 1361–1401. <https://doi.org/10.2307/1913712>
- Ogaki, M. (1998). “On the Granger Representation Theorem: a counter example?” [https://doi.org/10.1016/S0165-1765\(98\)00091-3](https://doi.org/10.1016/S0165-1765(98)00091-3)
- Soto, R. (2022), “Econometría de Series de Tiempo (Parte 5, 2022)”[Slides Teoría Econométrica III]. Pontificia Universidad Católica de Chile.