



Tarea 1.

Oscar Herrera
Segundo Semestre de 2024.

Pregunta 1.1.a

Series de tiempo.

Una serie de tiempo es una **secuencia** de una variable aleatoria, donde cada observación enumerada respresenta una realización de esta variable en un momento específico del tiempo. En otras palabras, una serie de tiempo $\{y_t\}_{t=0}^{T}$ representa un vector indexado en el tiempo que describe la evolución de una variable según sus realizaciones observadas.

Estacionariedad.

La **estacionariedad** es una propiedad de las series de tiempo que restringe la heterogeneidad de las observaciones para facilitar una válida inferencia. Podemos encontrar dos tipos de estacionariedad: fuerte o débil.

La **estacionariedad fuerte** establece que la distribución incondicional conjunta de un set determinado de observaciones, no cambia dependiendo del momento del tiempo en que consideramos el set. En términos formales:

$$(y_t, y_{t+j_1}, y_{t+j_2}, \dots, y_{t+j_n}) \sim f_{y_t, y_{t+j_1}, y_{t+j_2}, \dots, y_{t+j_n}}(y_t, y_{t+j_1}, y_{t+j_2}, \dots, y_{t+j_n}), \quad \forall t, j_i$$

Notar que la secuencia $\{j_1, j_2, \dots, j_n\}$ es arbitraria; esta puede tomar valores como: $\{1, 2, 3, \dots, n\}$, $\{1, 3, 5, 7, \dots, n\}$ o $\{2, 4, 6, 8, \dots, n\}$. Por lo tanto, lo que establece la estacionariedad fuerte es que, fijando ventanas de tiempo, para todo set que mantienen la misma ventana entre sí, sus distribuciones incondicionales conjuntas son la misma $\forall t$ (no depende del tiempo).

La **estacionariedad débil** establece que sólo la media incondicional, varianza incondicional y autocovarianzas incondicionales no dependen de t para una ventana fija. En términos formales, la **estacionariedad débil o estacionariedad en covarianza** establece que¹:

¹Notación inspirada en set de slides “Applied_macro_univariate.pdf”, Macroeconomía Aplicada, 16 de agosto de 2023.

- $\mathbb{E}[x_t] = \mu_t = \mu, \forall t$.
- $\mathbb{E}[(x_t - \mu_t)(x_{t-\tau} - \mu_{t-\tau})] = \mathbb{E}[(x_{t+s} - \mu_{t+s})(x_{t+s-\tau} - \mu_{t+s-\tau})], \forall t, s, \tau$

La segunda proposición hace referencia a que la covarianza entre dos momentos del tiempo con una ventana de tiempo τ arbitraria, es la misma para todos los intervalos cuya ventana es del mismo tamaño. Es decir, si tomamos una ventana τ , obtenemos la covarianza entre t y $t - \tau$, luego obtenemos la covarianza $t + s$ y $t + s - \tau$ (misma ventana pero en otra posición de la secuencia), el resultado siempre será el mismo en la medida que la distancia τ sea la misma. Lo anterior se cumple para toda la muestra: $\forall t, s, \tau$.

Intuitivamente, el concepto de estacionariedad débil se refiere a la propiedad de que los momentos estadísticos de la secuencia no dependen del tiempo, propiedad que conduce a que la serie exhiba un comportamiento de “*reversión a la media*”.

Ergodicidad.

La **ergodicidad** es la propiedad de una serie de tiempo $\{y_t\}_{t=-\infty}^{\infty}$ en la cual los momentos incondicionales **muestrales** de la serie convergen a sus momentos incondicionales del **Data Generating Process poblacional**, dado un “ t ” específico². Por ejemplo, si el promedio muestral incondicional converge a $\mathbb{E}[Y_t]$ para un “ t ” arbitrario (*ensemble average* en un momento específico).

Se dice que un proceso es ergódico en media si:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T y_t^{(1)} = \mathbb{E}[Y_t], \quad \forall t$$

donde $y_t^{(1)}$ corresponde a una de las posibles realizaciones de la secuencia $\{y_t\}_{t=1}^{t=T}$ y $\mathbb{E}[Y_t]$ corresponde al *ensemble average* de un momento t arbitrario.

Se dice que un proceso es ergódico en segundos momentos si:

$$\lim_{T \rightarrow \infty} \frac{1}{T-j} \sum_{t=j+1}^T (y_t - \mu)(y_{t-j} - \mu) = \gamma_j, \quad \forall t, j$$

o equivalentemente,

$$\frac{1}{T-j} \sum_{t=j+1}^T (y_t - \mu)(y_{t-j} - \mu) \xrightarrow{p} \gamma_j$$

²Hamilton (1994, p. 47).

Además de lo anterior, se dice que un proceso será ergódico en media si la covariación entre dos observaciones de la serie y_t y y_{t-j} tiende a cero **rápidamente** a medida que j tiende a infinito (condición suficiente si el proceso es estacionario)³. Es por esto que es común encontrar el concepto de ergodicidad definido como⁴:

$$\lim_{j \rightarrow \infty} Cov(y_t, y_{t-j}) = 0$$

Ergodicidad definida como la propiedad anterior restringe la interdependencia temporal en ventanas de tiempo grande, e intuitivamente es equivalente a decir que “la memoria de la serie desaparece en el largo plazo”.

Proceso Estacionario y No ergódico.

Sea el siguiente algoritmo para generar realizaciones de un proceso no ergódico $y_t^{(i)} = \mu^{(i)} + \varepsilon_t$ ⁵:

Algorithm 1 Algoritmo generador de procesos estacionarios no ergódicos.

```

Obtener una realización de:  $\mu^{(i)} \sim N(0, \lambda^2)$ 
Obtener una realización de la secuencia:  $\{y_t^{(i)}\}_{t=1}^T$ 
for  $t = 1$  to  $T$  do
     $y_t^{(i)} = \mu^{(i)} + \varepsilon_t$                                  $\triangleright$  Notar que  $\mu^{(i)}$  es v.a. con media 0
     $\varepsilon_t \sim N(0, \sigma^2)$ 
end for

```

El anterior proceso $\{y_t^{(i)}\}_{t=1}^T$ posee los siguientes momentos:

$$\begin{aligned}\mu_t &= \underbrace{\mathbb{E}[\mu^{(i)}]}_{\text{v.a.}} + \mathbb{E}[\varepsilon_t] \\ &= 0 + 0 = 0\end{aligned}$$

$$\begin{aligned}\gamma_{0t} &= \mathbb{E}[y_t^{(i)} - \mu_t]^2 \\ &= \mathbb{E}[\mu^{(i)} + \varepsilon_t]^2 \\ &= \lambda^2 + \sigma^2\end{aligned}$$

³Hamilton (1994, p. 47).

⁴Tal como en set de slides “Applied Macro Intro.pdf”, Macroeconomía Aplicada, 16 de agosto de 2023

⁵Siguiendo Hamilton (1994, p. 47).

$$\begin{aligned}
\gamma_{jt} &= \mathbb{E}[y_t^{(i)} - \mu_t][y_{t-j}^{(i)} - \mu_{t-j}] \\
&= \mathbb{E}[\mu^{(i)} + \varepsilon_t][\mu^{(i)} + \varepsilon_{t-j}] \\
&= \lambda^2,
\end{aligned}
\quad \text{para } j \neq 0$$

Lo anterior demuestra estacionariedad en covarianza.

Luego, notar que $\{y_t^{(i)}\}_{t=1}^T$ posee los siguientes momentos muestrales incondicionales (*time averages*):

$$\begin{aligned}
\frac{1}{T} \sum_{t=1}^T y_t^{(i)} &= \frac{1}{T} \sum_{t=1}^T (\mu^{(i)} + \varepsilon_t) \\
&= \mu^{(i)} + \underbrace{\frac{1}{T} \sum_{t=1}^T \varepsilon_t}_{\rightarrow 0} \\
&= \mu^{(i)}
\end{aligned}$$

Por lo tanto, el promedio muestral incondicional no converge a la media del proceso, lo que demuestra que el proceso estacionario Y_t no es ergódico.

Pregunta 1.1.b

Autocorrelograma Total.

Sea la función de autocovarianza γ_{jt} de un proceso $\{y_t\}_{t=-\infty}^{\infty}$:

$$\gamma_{jt} \equiv \mathbb{E}[y_t - \mu][y_{t-j} - \mu], \quad \forall t, j$$

entonces, la función de autocovarianza de un proceso débilmente estacionario está dada por:

$$\gamma_j = \mathbb{E}[y_t - \mu][y_{t-j} - \mu], \quad \forall t, j$$

Luego, si un proceso es estacionario, entonces su correlograma total estará dado por:

$$\rho_j = \frac{\gamma_j}{\gamma_0}, \quad \forall j$$

Autocorrelograma Parcial.

Siguiendo Shumway, R. & Stoffer, D. (2017), el autocorrelograma parcial de orden k $\phi_{k,k}$ se define como la correlación entre x_{t+k} y x_t que no está explicada por lags de orden intermedio. Es decir, la correlación luego de remover y controlar por dependencias lineales de x_{t+k} en $\{x_{t+k-1}, x_{t+k-2}, x_{t+k-3}, \dots, x_{t+1}\}$ y de x_t en $\{x_{t+k-1}, x_{t+k-2}, x_{t+k-3}, \dots, x_{t+1}\}$.

La remoción de correlaciones explicadas por rezagos intermedios se obtiene luego de extraer a x_{t+k} su proyección lineal en las variables $\{x_{t+k-1}, x_{t+k-2}, x_{t+k-3}, \dots, x_{t+1}\}$. Así, $\forall k = 1, 2, \dots$ la autocorrelación parcial está dada por (Shumway, R. & Stoffer, D., 2017):

$$\phi_{k,k} = \begin{cases} \text{corr}(x_{t+1}, x_t), & \text{si } k = 1, \\ \text{corr}(x_{t+k} - \hat{x}_{t+k}, x_t - \hat{x}_t), & \text{si } k \geq 2, \end{cases}$$

donde las proyecciones:

$$\begin{aligned} \hat{x}_{t+k} &= \beta_1 x_{t-k-1} + \beta_2 x_{t-k-2} + \dots + \beta_{k-1} x_{t+1} \\ \hat{x}_t &= \Gamma_1 x_{t-k-1} + \Gamma_2 x_{t-k-2} + \dots + \Gamma_{k-1} x_{t+1} \end{aligned}$$

son las que minizan los errores cuadráticos medios $\mathbb{E}[(x_{t+k} - \hat{x}_{t+k})^2]$ y $\mathbb{E}[(x_t - \hat{x}_t)^2]$ $\forall k$.

Equivalentemente, $\phi_{k,k}$ puede ser obtenido mediante la estimación del último coeficiente α_m en la siguiente ecuación⁶:

$$(x_t - \mu) = \alpha_1(x_{t-1} - \mu) + \dots + \alpha_m(x_{t-m} - \mu) \quad (1)$$

Notar que si x_t sigue un proceso $AR(p)$, entonces $\alpha_m = 0 \ \forall m > p$.

Autocorrelogramas, Proceso AR(2).

Sea un $AR(2)$:

$$x_t = \mu + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim w.n.$$

⁶Slide 18, “Applied_macro_univariate.pdf”, Macroeconomía Aplicada, 16 de agosto de 2023.

En primer lugar, por la ecuación (1), es fácil ver que el autocorrelograma parcial de un AR(2) es el siguiente:

$$\phi_{k,k} = \begin{cases} \phi_1, & \text{si } k = 1, \\ \phi_2, & \text{si } k = 2, \\ 0, & \text{si } k \geq 3, \end{cases}$$

Luego, partiendo por el cálculo de la varianza incondicional (para después obtener la función de autocovarianza para todo orden), suponiendo varianza y autocovarianzas constantes $\forall t$:

$$\begin{aligned} \gamma_0 &= \text{Var}(x_t) = \text{Var}(\phi_1 x_{t-1}) + \text{Var}(\phi_2 x_{t-2}) + \text{Var}(\varepsilon_t) \\ &\quad + 2 \text{Cov}(\phi_1 x_{t-1}, \phi_2 x_{t-2}) \\ &\quad + 2 \underbrace{\text{Cov}(\phi_1 x_{t-1}, \varepsilon_t)}_{=0} + 2 \underbrace{\text{Cov}(\phi_2 x_{t-2}, \varepsilon_t)}_{=0} \\ &= \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + \sigma_\varepsilon^2 + 2\phi_1 \phi_2 \gamma_1 \\ \implies \gamma_0(1 - \phi_1^2 - \phi_2^2) &= 2\phi_1 \phi_2 \gamma_1 + \sigma_\varepsilon^2 \\ \gamma_0 &= \frac{2\phi_1 \phi_2 \gamma_1 + \sigma_\varepsilon^2}{1 - \phi_1^2 - \phi_2^2} \end{aligned}$$

Por otro lado:

$$\begin{aligned} \gamma_1 &= \text{Cov}(c + \phi_1 x_{t-2} + \phi_2 x_{t-3} + \varepsilon_{t-1}, x_{t-2}) \\ &= \phi_1 \text{Var}(x_{t-2}) + \text{Cov}(\phi_2 x_{t-3}, x_{t-2}) \\ &= \phi_1 \text{Var}(x_{t-2}) + \phi_2 \text{Cov}(x_{t-3}, x_{t-2}) \\ &= \phi_1 \gamma_0 + \phi_2 \gamma_1 \\ \gamma_1(1 - \phi_2) &= \phi_1 \gamma_0 \\ \gamma_1 &= \frac{\phi_1}{1 - \phi_2} \gamma_0 \end{aligned}$$

Reemplazando:

$$\gamma_0 = \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + \sigma_\varepsilon^2 + 2\phi_1 \phi_2 \gamma_1$$

$$\begin{aligned}\gamma_0(1 - \phi_1^2 - \phi_2^2) &= 2\phi_1\phi_2\gamma_1 + \sigma_\varepsilon^2 \\ \gamma_0 &= \frac{2\phi_1\phi_2\gamma_1 + \sigma_\varepsilon^2}{1 - \phi_1^2 - \phi_2^2}\end{aligned}$$

$$\begin{aligned}\implies \gamma_0 &= \frac{2\phi_1\phi_2 \left(\frac{\phi_1}{1-\phi_2} \gamma_0 \right) + \sigma_\varepsilon^2}{1 - \phi_1^2 - \phi_2^2} \\ &= \frac{2\phi_1^2\phi_2 \left(\frac{\gamma_0}{1-\phi_2} \right)}{1 - \phi_1^2 - \phi_2^2} + \frac{\sigma_\varepsilon^2}{1 - \phi_1^2 - \phi_2^2} \\ &= \frac{2\phi_1^2\phi_2\gamma_0}{(1 - \phi_2)(1 - \phi_1^2 - \phi_2^2)} + \frac{\sigma_\varepsilon^2}{1 - \phi_1^2 - \phi_2^2}\end{aligned}$$

$$\begin{aligned}\gamma_0 \left(1 - \frac{2\phi_1^2\phi_2}{(1 - \phi_2)(1 - \phi_1 - \phi_2)} \right) &= \frac{\sigma_\varepsilon^2}{(1 - \phi_1 - \phi_2)} \\ \gamma_0 \frac{(1 - \phi_2)(1 - \phi_1 - \phi_2) - 2\phi_1^2\phi_2}{(1 - \phi_2)(1 - \phi_1 - \phi_2)} &= \frac{\sigma_\varepsilon^2}{(1 - \phi_1 - \phi_2)}\end{aligned}$$

$$\gamma_0 = \frac{\sigma_\varepsilon^2(1 - \phi_2)}{(1 - \phi_2)(1 - \phi_1 - \phi_2) - 2\phi_1^2\phi_2}$$

Luego, recursivamente:

$$\begin{aligned}\gamma_j &= \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} \\ \gamma_{j-1} &= \phi_1\gamma_{j-2} + \phi_2\gamma_{j-3} \\ \implies \gamma_j &= \phi_1[\phi_1\gamma_{j-2} + \phi_2\gamma_{j-3}] + \phi_2\gamma_{j-2} \\ \gamma_{j-2} &= \phi_1\gamma_{j-3} + \phi_2\gamma_{j-4} \\ \implies \gamma_j &= \phi_1^2\gamma_{j-2} + \phi_1\phi_2\gamma_{j-3} + \phi_2\gamma_{j-2} \\ &= (\phi_1^2 + \phi_2)\gamma_{j-2} + \phi_1\phi_2\gamma_{j-3} \\ &= (\phi_1^2 + \phi_2)[\phi_1\gamma_{j-3} + \phi_2\gamma_{j-4}] + \phi_1\phi_2\gamma_{j-3} \\ &= (\phi_1^2 + \phi_2)\phi_1\gamma_{j-3} + (\phi_1^2 + \phi_2)\phi_2\gamma_{j-4} + \phi_1\phi_2\gamma_{j-3} \\ &= \phi_1[\phi_1^2 + \phi_2 + \phi_2]\gamma_{j-3} + [(\phi_1^2 + \phi_2)\phi_2]\gamma_{j-4} \\ &= \phi_1(\phi_1^2 + 2\phi_2)\gamma_{j-3} + (\phi_1^2\phi_2 + \phi_2^2)\gamma_{j-4} \\ &\vdots\end{aligned}$$

Se puede apreciar que el autocorrelograma decaerá suavemente y no será cero para ordenes superiores, aunque tenderá a cero.

Generalizando γ_j para ordenes superiores bajo un proceso $AR(2)$:

$$\begin{aligned}
\text{Cov}(x_t, x_{t-j}) &= \text{Cov}(c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t, x_{t-j}) \\
&= \phi_1 \text{Cov}(x_{t-1}, x_{t-j}) + \phi_2 \text{Cov}(x_{t-2}, x_{t-j}) \\
&= \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2}
\end{aligned}$$

Yule-Walker equation (caso particular):

$$\begin{aligned}
\gamma_j &= \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} \\
\implies \rho_j &= \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2}
\end{aligned}$$

Con esto, hemos encontrado el autocorrelograma total teórico.

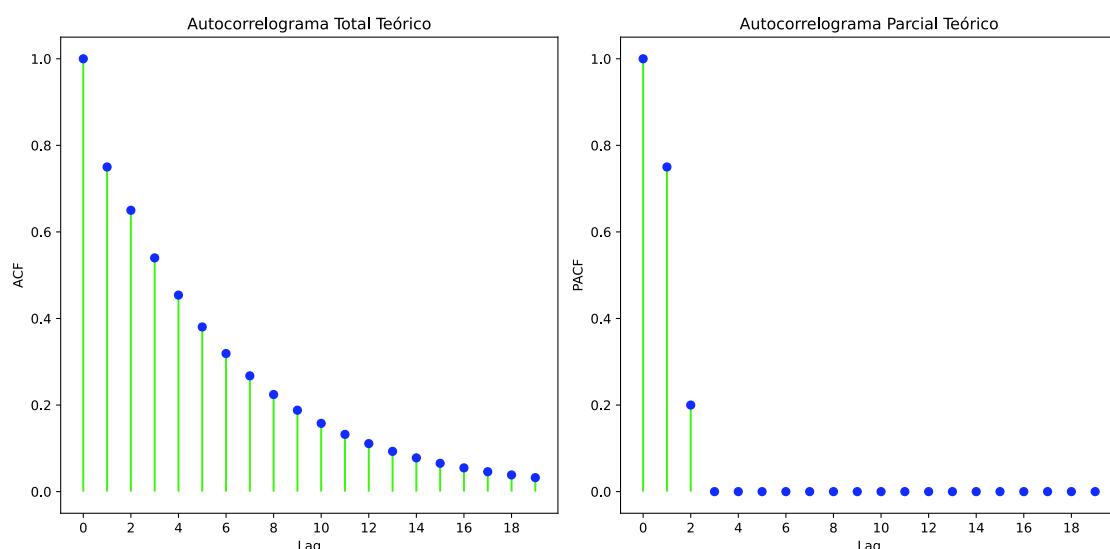
Ejemplo numérico.

Definiendo:

$$\phi_1 = 0,6$$

$$\phi_2 = 0,2$$

los autocorrelogramas teóricos numéricos son los siguientes:



Cuadro 1: ACF y PACF teóricos para un AR(2) con $\phi_1 = 0,6$ y $\phi_2 = 0,2$

Lag	ACF γ_j	PACF $\phi_{j,j}$
0	1.0000	1.0000
1	0.7500	0.7500
2	0.6500	0.2000
3	0.5400	0.0000
4	0.4540	0.0000
5	0.3804	0.0000
6	0.3190	0.0000
7	0.2675	0.0000
8	0.2243	0.0000
9	0.1881	0.0000
10	0.1577	0.0000
11	0.1322	0.0000
12	0.1109	0.0000
13	0.0930	0.0000
14	0.0780	0.0000
15	0.0654	0.0000
16	0.0548	0.0000
17	0.0460	0.0000
18	0.0385	0.0000
19	0.0323	0.0000

En línea con la teoría, el autocorrelograma total decae progresivamente, mientras que el autocorrelograma parcial alcanza cero en $\gamma_{3,3}$, tal como plantea el criterio Box-Jenkins⁷.

Pregunta 1.1.c

No toda secuencia de números $\{y_0, y_1, y_2, y_3, y_4, \dots\}$, con $y_0 = 1$ y cada número y_i con $i > 0$ siendo menor a 1 en valor absoluto⁸ es necesariamente un autocorrelograma de un proceso $AR(p)$ estacionario. A continuación, se procede a demostrar lo anterior por contraejemplo.

Sea una serie de tiempo $AR(2)$ estacionaria. Entonces, aparte de las restricciones sobre los parámetros del proceso, la matriz varianza-covarianza asociada al proceso debe ser definida positiva.

Consideremos una muestra de tamaño $T = 3$. Entonces, la matriz varianza-covarianza de la muestra es la siguiente:

⁷Códigos adjuntos en la entrega.

⁸Esto implica que $|x_i| < x_0$, $\forall i > 0$.

$$\Sigma = \begin{bmatrix} \text{Var}(y_t) & \text{Cov}(y_t, y_{t-1}) & \text{Cov}(y_t, y_{t-2}) \\ \text{Cov}(y_{t-1}, y_t) & \text{Var}(y_{t-1}) & \text{Cov}(y_{t-1}, y_{t-2}) \\ \text{Cov}(y_{t-2}, y_t) & \text{Cov}(y_{t-2}, y_{t-1}) & \text{Var}(y_{t-2}) \end{bmatrix} = \begin{bmatrix} \sigma_{y_t}^2 & \sigma_{y_t, y_{t-1}} & \sigma_{y_t, y_{t-2}} \\ \sigma_{y_t, y_{t-1}} & \sigma_{y_{t-1}}^2 & \sigma_{y_{t-1}, y_{t-2}} \\ \sigma_{y_t, y_{t-2}} & \sigma_{y_{t-1}, y_{t-2}} & \sigma_{y_{t-2}}^2 \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_2 & \gamma_1 & \gamma_0 \end{bmatrix}$$

La matriz de autocorrelogramas *teórico* asociado a aquella matriz es el siguiente:

$$\frac{1}{\gamma_0} \Sigma = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}$$

Notar que exigir que la matriz Σ sea definida positiva es equivalente a imponer que la matriz $\frac{1}{\gamma_0} \Sigma$ también lo sea:

$$\phi^\top \Sigma \phi > 0, \quad \forall \phi \neq \mathbf{0}$$

$$\Leftrightarrow \begin{bmatrix} \phi_0 & \phi_1 & \phi_2 \end{bmatrix} \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \end{bmatrix} > 0$$

Para que se cumpla la restricción, de acuerdo a Soto (2024), “se requiere que el determinante y los subdeterminantes sean positivos. Esto último requiere que $|\phi_1| \leq 1$ y $|\phi_2| \leq 1$, en tanto que el primero exige

$$1 + 2\rho_1^2\rho_2 - 2\rho_1^2 - \rho_2^2 \geq 0 \Rightarrow (\rho_2 - (2\rho_1^2 - 1))(\rho_2 - 1) \leq 0$$

pero como $|\rho_2| \leq 1$, entonces se debe cumplir que:

$$(\rho_2 - (2\rho_1^2 - 1)) \geq 0 \Rightarrow \rho_2 \geq (2\rho_1^2 - 1) \Rightarrow -1 \leq \frac{(\rho_2 - \rho_1^2)}{1 - \rho_1^2} \leq 1$$

Impone restricciones a ρ_2 dado ρ_1 .

Por ejemplo, si $\rho_1 = 0,9$, entonces $0,62 \leq \rho_2 \leq 1$ ⁹.

Así, la secuencia $\{1, 0,9, 0,6, \dots\}$ no es un autocorrelograma de un progreso regresivo $AR(p)$ con $p = 2$, por lo que queda demostrado.

⁹Slides 31 y 32, “Econometría de Series de Tiempo (Parte 2, 2024)”, Soto (2024).

Pregunta 1.1.d

Sea un proceso $MA(\infty)$ descrito por:

$$x_t = \mu + \sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim w.n.$$

Para mostrar que un proceso es estacionario (débilmente) se requiere:

1. $\mathbb{E}[x_t] = \mu_t = \mu, \forall t.$
2. $\gamma_{\tau t} = \mathbb{E}[(x_t - \mu_t)(x_{t-\tau} - \mu_{t-\tau})] = \mathbb{E}[(x_{t+s} - \mu_{t+s})(x_{t+s-\tau} - \mu_{t+s-\tau})], \forall t, s, \tau$

Si se cumple (2), entonces la función de autocovarianza está dada por:

$$\gamma_j = \mathbb{E}[y_t - \mu][y_{t-j} - \mu], \quad \forall t, j$$

Por lo cual, mostrar (2) es equivalente a mostrar que la función γ_j , existe, está bien definida y que es la misma $\forall t$.

Momentos Incondicionales.

Media incondicional.

Tomando esperanza:

$$\begin{aligned} \mathbb{E}[x_t] &= \mu + \mathbb{E}\left[\sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i}\right] + \mathbb{E}[\varepsilon_t] \\ &= \mu \end{aligned} \quad \forall t$$

Por lo tanto, se cumple (1).

Autocovarianza.

Luego de mostrar (1), la función de autocovarianza γ_{jt} está dada por:

$$\begin{aligned}
\gamma_{jt} &= \mathbb{E}[x_t - \mathbb{E}[x_t]][x_{t-j} - \mathbb{E}[x_t]] \\
&= \mathbb{E}[x_t - \mu][x_{t-j} - \mu] \\
&= \mathbb{E}\left[\mu + \sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i} + \varepsilon_t - \mu\right] \left[\mu + \sum_{i=1}^{\infty} \theta_i \varepsilon_{t-j-i} + \varepsilon_{t-j} - \mu\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i} + \varepsilon_t\right] \left[\sum_{i=1}^{\infty} \theta_i \varepsilon_{t-j-i} + \varepsilon_{t-j}\right]
\end{aligned}$$

Notar que por propiedades de ruido blanco, ε_t posee los mismos momentos $\forall t$. Por lo tanto, el resultado de la ecuación anterior es el mismo para todo momento del tiempo. Específicamente, gracias a que $\varepsilon_t \cdot \varepsilon_{t+s} = 0$, $\forall t, s$.

Por simplicidad, usando $q \rightarrow \infty$, entonces el correlograma total está dado por¹⁰:

$$\gamma_{jt} = \gamma_j = \begin{cases} \sigma_\varepsilon^2 \left(\sum_{i=j}^q \theta_i \theta_{i-j} \right), & \text{si } 0 \leq j \leq q \\ 0, & \text{si } j > q \end{cases}$$

Luego, debido a que $q \rightarrow \infty$:

$$\gamma_j = \sigma_\varepsilon^2 \left(\sum_{i=j}^{\infty} \theta_i \theta_{i-j} \right), \quad \forall j$$

Sumabilidad Cuadrática.

Una vez mostrada que la función de autocovarianza no depende del tiempo, sólo queda demostrar que γ_j existe y está bien definida, lo cual sólo requiere que $\sum_{i=j}^{\infty} \theta_i \theta_{i-j}$ sea finita.

Usando propiedades de la función de autocovarianzas cuando no depende de t :

$$\gamma_0 > |\gamma_j|, \quad \text{para } j > 0.$$

Por otro lado, la sumabilidad cuadrática se define como la propiedad de presentar una varianza finita.

$$\sum_{i=0}^{\infty} \theta_i < \infty$$

¹⁰Slide 11, “Econometría de Series de Tiempo (Parte 2, 2024)”, Soto (2024).

Por lo que de cumplirse sumabilidad cuadrática, usando $\gamma_0 > |\gamma_j|$, entonces la función de autocovarianza está bien definida para cualquier orden de j . Con esto, la estacionariedad queda demostrada.

Pregunta 1.1.e

Sea un proceso $MA(\infty)$ descrito por:

$$x_t = \mu + \sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i} + \varepsilon_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

con $\varepsilon_t \sim w.n..$ Suponiendo $\mu = 0$ y usando operador de rezagos:

$$\begin{aligned} x_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} = L^0 \varepsilon_t + \theta_1 L \varepsilon_t + \theta_2 L^2 \varepsilon_t + \dots + \theta_q L^q \varepsilon_t \\ &= (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t \\ &\implies x_t (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)^{-1} = \varepsilon_t \\ &\implies x_t \frac{1}{1 - (-\theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)} = \varepsilon_t \\ &\implies \sum_{i=0}^{\infty} (-\theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)^i x_t = \varepsilon_t \end{aligned}$$

Si bien cada polinomio $(-\theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)^i$ posee soluciones difícilmente expresables en forma analítica, aún así podemos deducir que luego de factorizar, cada término L^i tendrá asociado un coeficiente $-\phi_i$. Notar que $(-\theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)^0 = 1$. Luego:

$$\begin{aligned} \sum_{i=0}^{\infty} (-\theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)^i &= 1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \dots \\ \implies x_t (1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \dots) &= \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + \dots + \varepsilon_t = \varepsilon_t \\ \implies x_t &= \sum_{i=1}^{\infty} \phi_i x_{t-i} + \varepsilon_t \end{aligned}$$

Lo anterior sólo es válido si se cumplen las condiciones de estacionariedad del $AR(\infty)$. Es decir, si se cumple que las raíces del polinomio de rezagos $\Phi(L)$ caen fuera del círculo unitario en el plano complejo, lo que se conoce como “condición de invertibilidad”. Por lo tanto, todo proceso $MA(q)$ tiene representación $AR(\infty)$ si y sólo si, el polinomio asociado cumple la condición de invertibilidad.

Pregunta 1.2.a

Determinación de estacionariedad de proceso (*i*).

El proceso (1) está dado por:

$$y_t = 2y_{t-1} - \frac{3}{2}y_{t-2} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2) \quad (i)$$

Reexpresando con operadores de rezago:

$$\begin{aligned} y_t &= 2Ly_t - \frac{3}{2}L^2y_t + \varepsilon_t \\ \left(1 - 2L + \frac{3}{2}L^2\right)y_t &= \varepsilon_t \end{aligned}$$

Luego, para determinar si el comportamiento del proceso autorregresivo será divergente o convergente, y por ende, estacionario o no, resolvemos las raíces del polinomio característico para verificar que estas caigan dentro del círculo unitario en el plano complejo (Hamilton, 1994, p. 32):

$$\begin{aligned} \left(1 - 2L + \frac{3}{2}L^2\right) &= 0 \\ (c + bL + aL^2) &= 0 \\ \Leftrightarrow (c + bz + az^2) &= 0 \end{aligned}$$

Usando:

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

obtenemos:

$$z = \frac{2 \pm \sqrt{4 - 4(\frac{3}{2})(1)}}{2(\frac{3}{2})} = \frac{2 \pm \sqrt{4 - 6}}{3} = \frac{2 \pm \sqrt{-2}}{3} = \frac{2 \pm i\sqrt{2}}{3}$$

Luego, obteniendo el módulo de z ¹¹:

¹¹El módulo $|z|$ de un número complejo $z = x + iy$ se obtiene mediante la ecuación: $|z| = \sqrt{x^2 + y^2}$.

$$|z| = \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{\sqrt{2}}{3}\right)^2} = \sqrt{\frac{4}{9} + \frac{2}{9}} = 0,816496581 < 1$$

Por lo tanto, como las raíces del polinomio caen dentro del círculo unitario, el proceso es estacionario.

Determinación de estacionariedad de proceso (ii).

El proceso (ii) está dado por:

$$y_t = \frac{1}{2}y_{t-1} + \frac{1}{3}y_{t-2} + 2\varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2) \quad (ii)$$

Reexpresando con operadores de rezago:

$$\left(1 - \frac{1}{2}L - \frac{1}{3}L^2\right)y_t = (1 + 2L)\varepsilon_t$$

Como el lado derecho está conformado por ruidos blancos que por definición son estacionarios, para determinar si el comportamiento del proceso será estacionario, necesitamos asegurarnos que las raíces del polinomio característico del lado izquierdo de la ecuación caigan dentro del círculo unitario en el plano complejo:

$$\begin{aligned} \left(1 - \frac{1}{2}L - \frac{1}{3}L^2\right) &= 0 \\ \Leftrightarrow (c + bz + az^2) &= 0 \end{aligned}$$

$$\Rightarrow z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{\frac{1}{2} \pm \sqrt{\frac{1}{4} - 4(-\frac{1}{3})(1)}}{2(-\frac{1}{3})} = \frac{\frac{1}{2} \pm \sqrt{\frac{3+16}{12}}}{3} = \frac{\frac{1}{2} \pm \sqrt{\frac{19}{12}}}{3}$$

Obteniendo el módulo:

$$|z| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{19}{12}\right)^2} = \sqrt{\frac{4}{9} + \frac{2}{9}} = 1,660404904 > 1$$

Por lo tanto el proceso es no estacionario.

Determinación de estacionariedad de proceso (iii).

El proceso (iii) está dado por:

$$y_t = \frac{1}{2} + \frac{1}{2}y_{t-1} + \frac{1}{2}y_{t-2} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2) \quad (iii)$$

Reexpresando con operadores de rezago:

$$\left(1 - \frac{1}{2}L - \frac{1}{2}L^2\right)y_t = \frac{1}{2} + \varepsilon_t$$

Como el lado derecho está conformado por un ruido blanco, para determinar si el comportamiento del proceso será estacionario necesitamos asegurarnos que las raíces del polinomio característico del lado izquierdo de la ecuación caigan dentro del círculo unitario en el plano complejo:

$$\begin{aligned} \left(1 - \frac{1}{2}L - \frac{1}{2}L^2\right) &= 0 \\ \Leftrightarrow (c + bz + az^2) &= 0 \end{aligned}$$

$$\Rightarrow z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{\frac{1}{2} \pm \sqrt{\frac{1}{4} - 4(-\frac{1}{2})(1)}}{2(-\frac{1}{2})} = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + 2} = -\frac{1}{2} \pm \sqrt{\frac{9}{4}} = -\frac{1}{2} \pm \frac{3}{2}$$

$$z_1 = 1, z_2 = -2$$

Por lo tanto el proceso es no estacionario.

Determinación de estacionariedad de proceso (iv).

El proceso (iv) está dado por:

$$y_t = -\frac{1}{3}y_{t-1} + 2\varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2) \quad (iv)$$

Reexpresando con operadores de rezago:

$$\left(1 + \frac{1}{3}L\right)y_t = (1 + 2L)\varepsilon_t$$

Como el lado derecho está conformado por un ruido blanco, para determinar si el comportamiento del proceso será estacionario necesitamos asegurarnos que las raíces del polinomio característico del lado izquierdo de la ecuación caigan dentro del círculo unitario en el plano complejo:

$$\begin{aligned} \left(1 + \frac{1}{3}L\right) &= 0 \\ \Leftrightarrow (c + bz) &= 0 \end{aligned}$$

Es fácil ver que la raíz del polinomio es -3, por lo que el proceso no será estacionario.

Pregunta 1.2.b

Esperanza incondicional.

Sea el proceso $AR(2)$:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2)$$

tomando operador de esperanzas, y sabiendo que es estacionario:

$$\begin{aligned} \mathbb{E}[y_t] &= \phi_1 \mathbb{E}[y_{t-1}] + \phi_2 \mathbb{E}[y_{t-2}] + \varepsilon_t \approx \\ \mathbb{E}[y_t] &= \phi_1 \mathbb{E}[y_t] + \phi_2 \mathbb{E}[y_t] \\ \mathbb{E}[y_t](1 - \phi_1 - \phi_2) &= 0 \\ \implies \mathbb{E}[y_t] &= 0 \end{aligned}$$

Varianza incondicional.

Siguiendo el desarrollo previamente establecido en la **Pregunta 1.1.b**, por la ecuación (1), es fácil ver que el autocorrelograma parcial de un AR(2) es:

$$\phi_{k,k} = \begin{cases} \phi_1, & \text{si } k = 1, \\ \phi_2, & \text{si } k = 2, \\ 0, & \text{si } k \geq 3, \end{cases}$$

Autocovarianza.

De la misma forma, por el desarrollo de la **Pregunta 1.1.b** sabemos que:

$$\begin{aligned} \gamma_0 &= \frac{2\phi_1\phi_2\gamma_1 + \sigma_\varepsilon^2}{1 - \phi_1^2 - \phi_2^2} \\ \gamma_1 &= \frac{\phi_1}{1 - \phi_2}\gamma_0 \\ \implies &= \frac{\sigma_\varepsilon^2(1 - \phi_2)}{(1 - \phi_2)(1 - \phi_1 - \phi_2) - 2\phi_1^2\phi_2} \end{aligned}$$

Reemplazando y desarrollando para la expresión de γ_1 :

$$\begin{aligned} \gamma_1 &= \frac{\phi_1}{1 - \phi_2}\gamma_0 \\ &= \frac{\phi_1}{1 - \phi_2} \cdot \frac{\sigma_\varepsilon^2(1 - \phi_2)}{(1 - \phi_2)(1 - \phi_1 - \phi_2) - 2\phi_1^2\phi_2} \\ &= \frac{\phi_1}{1} \cdot \frac{\sigma_\varepsilon^2 \cdot 1}{(1 - \phi_2)(1 - \phi_1 - \phi_2) - 2\phi_1^2\phi_2} \\ \gamma_1 &= \frac{\phi_1\sigma_\varepsilon^2}{(1 - \phi_2)(1 - \phi_1 - \phi_2) - 2\phi_1^2\phi_2} \end{aligned}$$

Pregunta 1.2.c

Sean los procesos:

$$(i) : \quad y_t = \varphi \varepsilon_{t-1} + \varepsilon_t,$$

$$(ii) : \quad y_t = -\frac{1}{\varphi} \varepsilon_{t-1} + \varepsilon_t,$$

con $\varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2)$. Ambos procesos son procesos $MA(q)$, con $q = 1$.

Se puede mostrar que la ACF de un proceso $MA(q)$ está dada por¹²:

$$\gamma_j = \begin{cases} \sigma_\varepsilon^2 \left(\sum_{i=j}^q \theta_i \theta_{i-j} \right), & \text{si } 0 \leq j \leq q \\ 0, & \text{si } j > q \end{cases}$$

con $\theta_0 = 1$.

Así, la autocovarianza de un proceso $MA(q)$ es:

$$\gamma_j = \begin{cases} \sigma_\varepsilon^2 \left(\sum_{i=0}^1 \theta_i \theta_{i-j} \right) = \sigma_\varepsilon^2 (\theta_0 \theta_0 + \theta_1 \theta_1) = \sigma_\varepsilon^2 (1 + \theta_1^2), & \text{si } j = 0 \\ \sigma_\varepsilon^2 \left(\sum_{i=1}^1 \theta_i \theta_{i-j} \right) = \sigma_\varepsilon^2 (\theta_1 \theta_0) = \sigma_\varepsilon^2 \theta_1, & \text{si } j = 1 \\ 0, & \text{si } j > 1 \end{cases}$$

El correlograma del proceso $MA(1)$ estará dado por:

$$\rho_j = \begin{cases} \frac{\sigma_\varepsilon^2 (1 + \theta_1^2)}{\sigma_\varepsilon^2 (1 + \theta_1^2)} = 1, & \text{si } j = 0 \\ \frac{\sigma_\varepsilon^2 \theta_1}{\sigma_\varepsilon^2 (1 + \theta_1^2)} = \frac{\theta_1}{(1 + \theta_1^2)}, & \text{si } j = 1 \\ 0, & \text{si } j > 1 \end{cases}$$

Por ende, la autocovarianza de (i) es la siguiente:

$$\gamma_j = \begin{cases} \sigma_\varepsilon^2 (1 + \theta^2), & \text{si } j = 0 \\ \sigma_\varepsilon^2 \theta, & \text{si } j = 1 \\ 0, & \text{si } j > 1 \end{cases}$$

¹²Slide 11, “Econometría de Series de Tiempo (Parte 2, 2024)”, Soto (2024).

Luego, su autocorrelograma es el siguiente:

$$\rho_j = \begin{cases} 1, & \text{si } j = 0 \\ \frac{\theta}{(1+\theta^2)}, & \text{si } j = 1 \\ 0, & \text{si } j > 1 \end{cases}$$

Por otro lado, la autocovarianza de (ii) la siguiente:

$$\gamma_j = \begin{cases} \sigma_\varepsilon^2 \left(1 + \left(\frac{1}{\theta}\right)^2\right), & \text{si } j = 0 \\ \sigma_\varepsilon^2 \left(\frac{1}{\theta}\right), & \text{si } j = 1 \\ 0, & \text{si } j > 1 \end{cases}$$

Notar que $\frac{\frac{1}{\theta}}{1 + \left(\frac{1}{\theta}\right)^2} = \frac{\frac{1}{\theta}}{\frac{\theta^2+1}{\theta^2}} = \frac{\theta^2}{\theta(\theta^2+1)} = \frac{\theta}{(1+\theta^2)}$.

Por ende, el autocorrelograma de (ii) es el siguiente:

$$\rho_j = \begin{cases} 1, & \text{si } j = 0 \\ \frac{\theta_1}{(1+\theta_1^2)} = \frac{\frac{1}{\theta}}{1 + \left(\frac{1}{\theta}\right)^2} = \frac{\theta}{(1+\theta^2)}, & \text{si } j = 1 \\ 0, & \text{si } j > 1 \end{cases}$$

Así, queda demostrado que (i) y (ii) son los mismos.

Pregunta 1.3

Pregunta 1.3.a

Para visualizar los autocorrelogramas de ambos procesos se usó el código “1_3_b_correlogramas_parciales.py” adjunto en esta entrega.

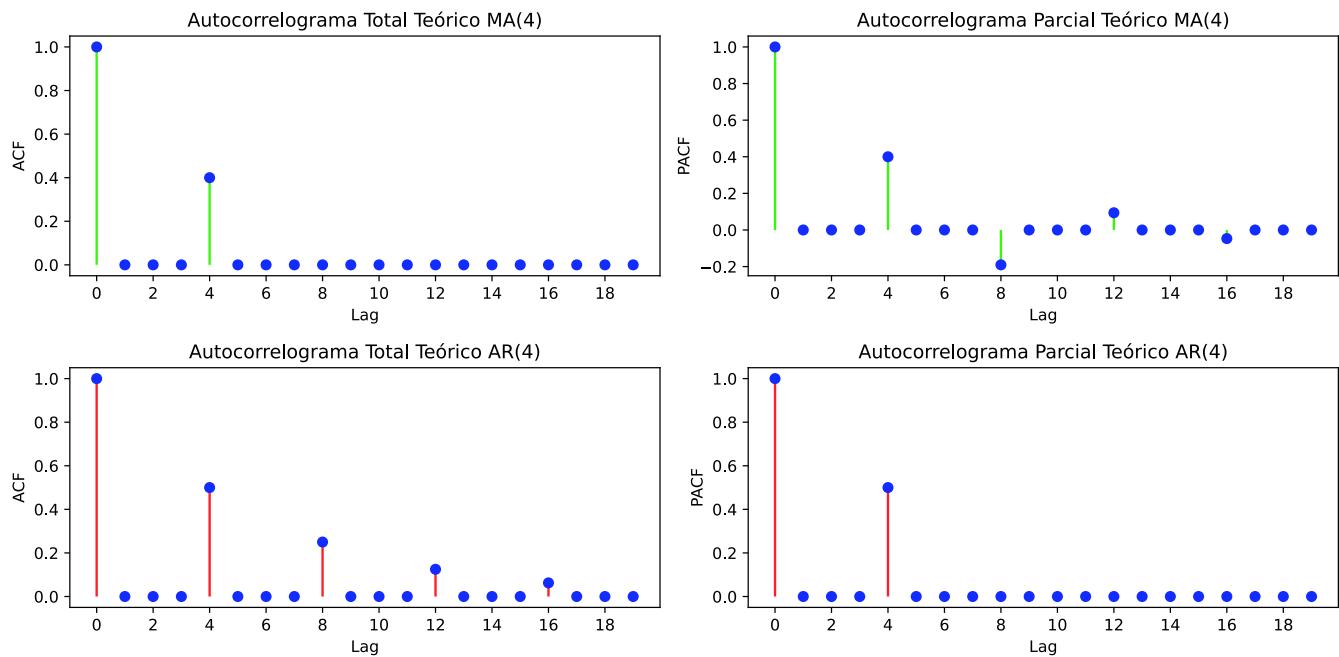
Sean los siguientes procesos:

$$(i) : y_t = \varepsilon_t + \theta_4 \varepsilon_{t-4},$$

$$(ii) : y_t = \phi_4 + \phi_0 y_{t-4} + \varepsilon_t,$$

con $\varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2)$ y $0 < |\{\theta_4, \phi_0, \phi_4\}| < 1$.

Asignando valores a $\theta_4, \phi_0, \phi_4 = 0,5$, el autocorrelograma de los procesos son los siguientes:



Se puede apreciar que el comportamiento de ambas series sugiere la presencia de estacionalidad. A pesar de que ambos autocorrelogramas muestran un comportamiento intermitente, lo que viola el criterio de muerte súbita de un correlograma total *MA* y la muerte súbita de un correlograma parcial *AR* según Box Jenkins, la intermitencia consistente del autocorrelograma parcial *MA* y la del total *AR* permite deducir que el valor del primer al tercer rezago es cero, lo que de por si sólo no se puede inferir con el autocorrelograma total *MA* y el parcial *AR*.

Pregunta 1.3.b

Sea el error de pronóstico $e_t(h) = y_{t+h} - \hat{y}_{t+h}$ con media igual a cero y $\mathbb{E}[y_{t+h} - \hat{y}_{t+h} | H_t]^2$ el *mínimo error cuadrático medio*, donde H_t es la historia del proceso hasta el momento t . A continuación, se demostrará insesgamiento.

Expandiendo el cuadrado del término dentro de la expectativa:

$$\begin{aligned}
\mathbb{E}[y_{t+h} - \hat{y}_{t+h} | H_t]^2 &= \mathbb{E}[y_{t+h}^2 - 2y_{t+h}\hat{y}_{t+h} + \hat{y}_{t+h}^2 | H_t] \\
&= \mathbb{E}[y_{t+h}^2 | H_t] - 2\mathbb{E}[y_{t+h} \underbrace{\hat{y}_{t+h}}_{\mathbb{E}[y_{t+h} | H_t]} | H_t] + \mathbb{E}[\hat{y}_{t+h}^2 | H_t] \\
&= \mathbb{E}[y_{t+h}^2 | H_t] - 2\mathbb{E}[y_{t+h} \cdot \mathbb{E}[y_{t+h} | H_t]] + \mathbb{E}[\hat{y}_{t+h}^2 | H_t] \\
&= \mathbb{E}[y_{t+h}^2 | H_t] - 2\mathbb{E}[y_{t+h} | H_t] \cdot \mathbb{E}[y_{t+h} | H_t] + \mathbb{E}[\underbrace{\hat{y}_{t+h}^2}_{\mathbb{E}[y_{t+h} | H_t]^2} | H_t] \\
&= \mathbb{E}[y_{t+h}^2 | H_t] - \mathbb{E}[y_{t+h} | H_t]^2 \\
&= \text{Var}[y_{t+h} | H_t]
\end{aligned}$$

Como el error cuadrático medio es igual a la varianza condicional, entonces el error de pronóstico sólo varía por incertidumbre asociada a nueva información del presente en la variable original. Es decir, por la incertidumbre de las innovaciones.

Toda la información pasada en $t+h$ no genera incertidumbre en el pronóstico, por lo que el error cuadrático medio es justamente el momento condicional deseado: la incertidumbre del componente aleatorio. Como no hay errores por componentes sistemáticos, el error cuadrático medio es insesgado.

Pregunta 1.3.c

Sea el proceso $z_t = x_t + \varepsilon_t$ con $\varepsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2)$ y x_t una variable aleatoria que no cambia a través del tiempo, tiene media cero y una varianza σ_x^2 , y $\mathbb{C}[x_t, \varepsilon_t] = 0$. Entonces:

$$\mathbb{E}[z_t] = \mathbb{E}[x_t] + \mathbb{E}[\varepsilon_t] = 0$$

$$\mathbb{V}[z_t] = \mathbb{V}[x_t] + \mathbb{V}[\varepsilon_t] + 2\mathbb{C}[x_t, \varepsilon_t] = 2\sigma_\varepsilon^2$$

$$\mathbb{C}[z_t, z_{t-h}] = \mathbb{C}[x_t + \varepsilon_t, x_{t-h} + \varepsilon_{t-h}] \tag{1}$$

$$= \mathbb{C}[x_t, x_{t-h}] + \mathbb{C}[x_t, \varepsilon_{t-h}] + \mathbb{C}[\varepsilon_t, x_{t-h}] + \mathbb{C}[\varepsilon_t, \varepsilon_{t-h}] \tag{2}$$

$$= 0 \tag{3}$$

Lo anterior es equivalente a la suma de dos ruidos blancos, donde al menos uno de ellos distribuye normal. Por lo tanto, podemos concluir que la suma de procesos ruido blanco también es estacionaria y la incertidumbre se amplifica en forma aritmética.

Pregunta 1.3.d

Sea $\{y_t\}$ un proceso $AR(p)$:

$$\Phi(L)y_t = \varepsilon_t$$

Sea $\{z_t\}$ un proceso $MA(q)$:

$$z_t = \theta(L)\varepsilon_t$$

A continuación, se procede a mostrar el proceso que seguirá la serie $x_t = y_t + z_t$. Notar que y_t posee una representación $MA(\infty)$:

$$y_t = \Phi(L)^{-1}\varepsilon_t$$

Entonces, x_t es un proceso $MA(\infty)$ lo que es claro de dilucidar en la siguiente ecuación:

$$x_t = y_t + z_t = \Phi(L)^{-1}\varepsilon_t + \theta(L)\varepsilon_t \equiv \Theta(L)\varepsilon_t = \sum_{i=0}^{\infty} \Theta_i L^i \varepsilon_t$$

Como ambos procesos están compuestos de componentes estacionarios, entonces x_t es estacionario. Sin embargo, se debe cumplir la restricción de sumabilidad cuadrática de las varianzas. Esto es de especial relevancia, ya que como vimos en el inciso anterior, las varianzas se amplifican aritméticamente. Dado que

$$\gamma_0 > |\gamma_j|, \quad \text{para } j > 0.$$

Entonces la sumabilidad cuadrática (varianza finita) impone la siguiente restricción¹³:

$$\sum_{i=0}^{\infty} \Theta_i < \infty$$

¹³Inciso 1.1.d.

Pregunta 2

Pregunta 2.a

En un contexto de modelos autorregresivos, usando la Probabilidad Clásica¹⁴ de *overfitting*¹⁵ como una medida del riesgo de sobreajuste (y de estar alejado del verdadero modelo), se encuentra que al determinar el número de rezagos por Criterio de Información AIC, el riesgo de sobreajuste es considerablemente mayor que al de la determinación de rezagos por BIC.

Calibrando simulaciones de Montecarlo con series macroeconómicas chilenas (Inflación, IMACEC y Tipo de Cambio)¹⁶, se encuentra que la probabilidad de *overfitting* es de alrededor de 30 % con AIC, mientras que la del BIC es de alrededor del 2 %, lo que representa una magnitud de riesgo considerable en el caso del AIC.

Adicionalmente, se encuentra evidencia empírica de que el sobreajuste luego se traspasa a problemas de *forecasting*. Cuantitativamente, se encuentra que el rendimiento predictivo es menor con sobreajuste que con subajuste, y que con datos mensuales año contra año (frecuencia de 12 meses) cuando se estima con más de 12 coeficientes (coeficientes mayor a la frecuencia), el error de proyección aumenta en 20 % aproximadamente cuando se proyecta a 24 meses en adelante. Lo anterior puede ser grave, mientras que el caso del subajuste muestra no ser grave.

Pregunta 2.b

Se pueden encontrar dos motivos principales. En primer lugar, porque al usar simulaciones de Montecarlo conocemos el Proceso Generador de Datos de la serie y podemos analizar con certidumbre si los criterios determinan un número de rezagos erróneo o no. Luego, en segundo lugar, al usar varias series tenemos la posibilidad de hacer inferencia de forma no paramétrica. Al usar un número grande de series sabemos que asintóticamente las distribuciones convergerán.

Luego, al usar una inferencia no paramétrica, esta simplifica bastante. El test pasa a ser la probabilidad clásica ya mencionada y nos permite sólo enfocarnos en si el valor computado de la probabilidad clásica se encuentra por debajo del valor de tolerancia.

Pregunta 2.c

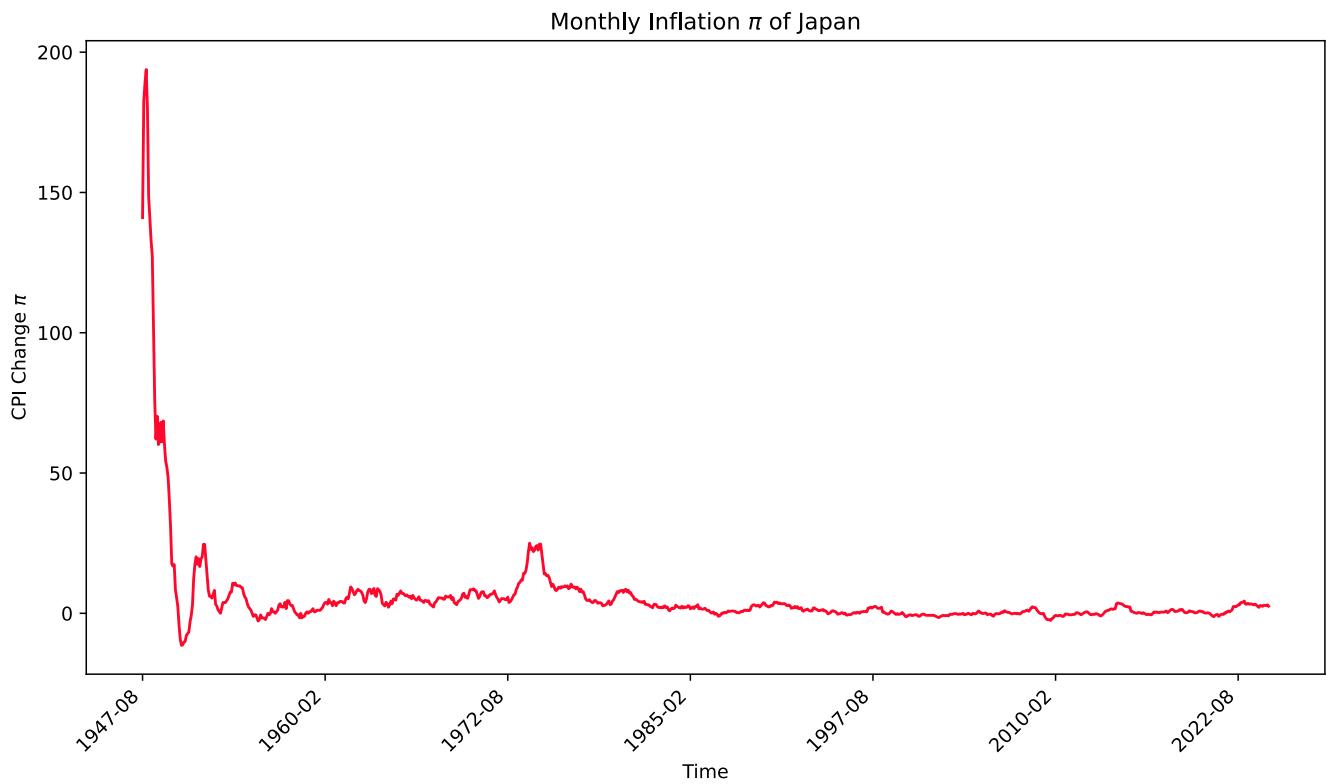
Los códigos de python usados se encuentran adjuntos en esta entrega.

Para esta pregunta se ha escogido Japón como la economía cuyas series serán sometidas al análisis del Paper. A continuación, se visualizan los datos de inflación descargados sin manipular desde el sitio web “*Data Portal*” de *Bank for International Settlements*.

¹⁴Frecuencia de casos en que se encuentra un *overfitting* relativo al total de veces que se estima un modelo simulado.

¹⁵Entendiendo este como el caso donde se determina un número de rezagos mayor al del verdadero modelo (establecido por Montecarlo).

¹⁶Variables representativas de la economía chilena, que concitan atención del público y que presentan diversas características estadísticas. Por lo tanto, el problema es uno económicamente relevante.



A pesar del quiebre estructural presentado al inicio de la muestra, el cual difícilmente podrá ser capturado estadísticamente, la serie parece mostrar un comportamiento estacionario, debido a que la serie presenta comportamiento de reversión a la media. Lo anterior es sometido a prueba mediante el *Augmented Dickey-Fuller test* implementado vía comando de python:

```
# (A Little hardcoded):
Serie = "Japan's Monthly Inflation"

print(f"Augmented Dickey-Fuller test for \'{Serie}\':")
result = adfuller(df_inflation_monthly['OBS_VALUE'].dropna(), autolag='AIC')
print(f"ADF Statistic: {result[0]}")
print(f"p-value: {format(result[1], 'f')}")
# print("Critical Values:")
# for key, value in result[4].items():
#     print(f"    {key}: {value}")
if result[1] < 0.05:
    print(f"\'{Serie}\'" is likely stationary (reject the null hypothesis).")
else:
    print(f"\'{Serie}\'" is likely non-stationary (fail to reject the null hypothesis).")
```

El diseño de las hipótesis del test está dado por:

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

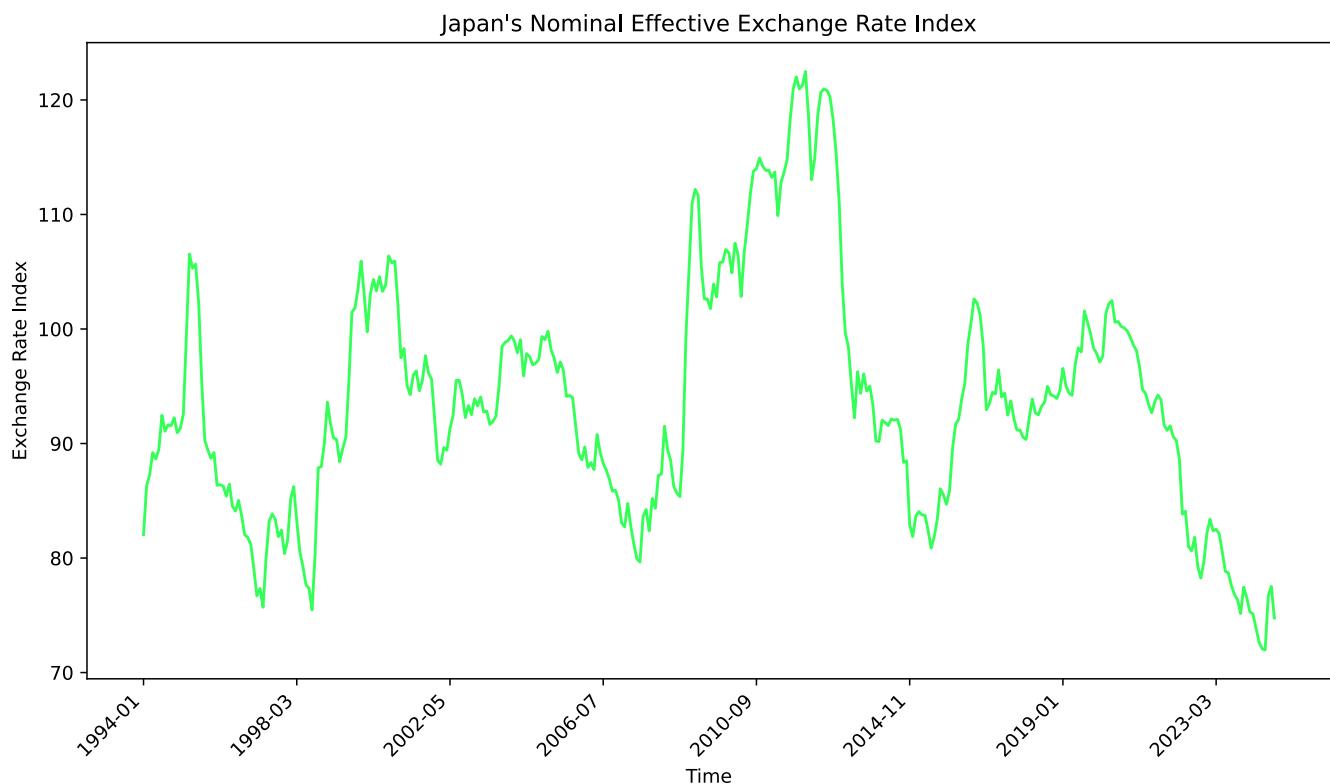
H_1 : La serie es estacionaria.

Augmented Dickey-Fuller Test for Japan's Monthly Inflation

ADF Statistic	-7.778716
p-value	0.000000

Dado que el valor-p se encuentra bajo el nivel de tolerancia estandar $\alpha = 5\%$, rechazamos la hipótesis nula H_0 y establecemos que existe evidencia estadística de que la serie de Inflación de Japón variación mensual año contra año es estacionario y no requiere transformación.

A continuación, se visualizan los datos de Tipo de Cambio de Japón descargados sin manipular desde el sitio web “Data Portal” de *Bank for International Settlements*.

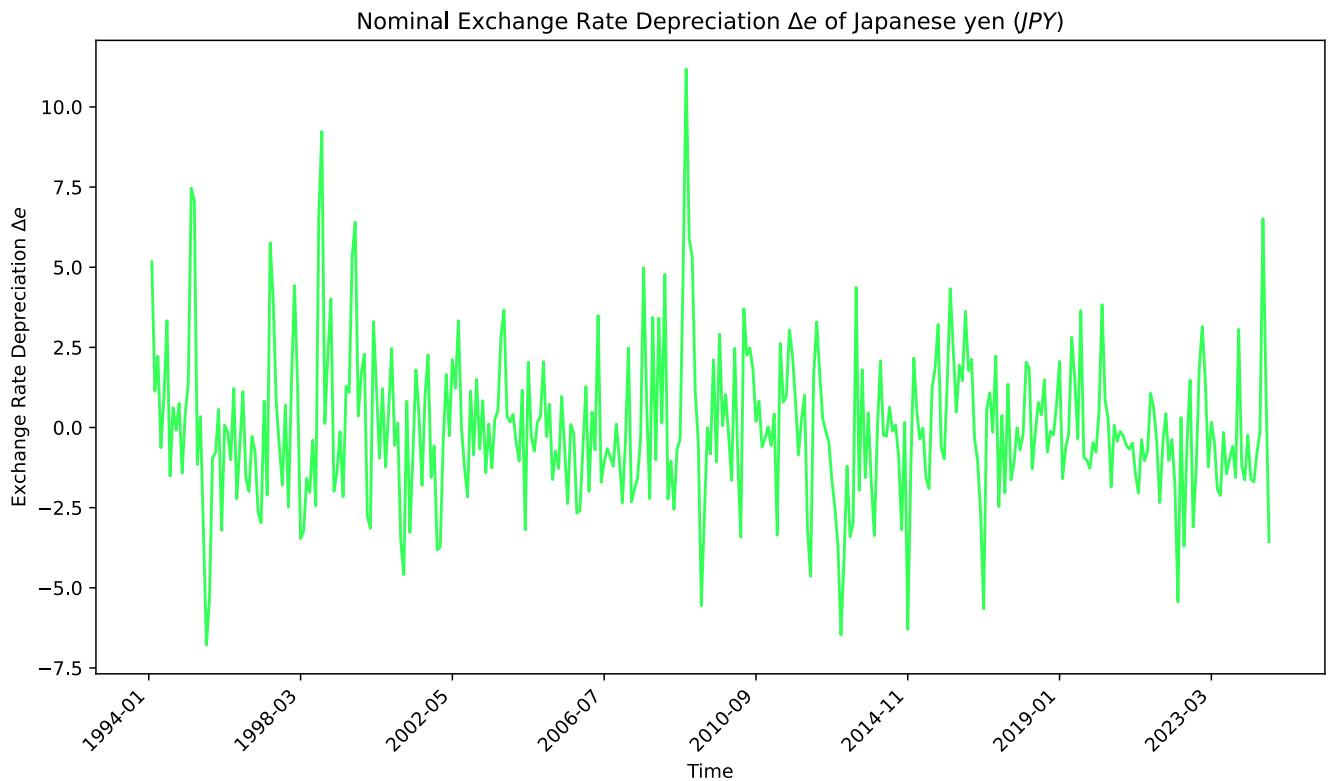


Replicando la aplicación del *Augmented Dickey-Fuller test*, encontramos evidencia de no estacionariedad dado que $p\text{-value} > \alpha$.

Augmented Dickey-Fuller Test for Japan's Nominal Exchange Rate Index

ADF Statistic	-2.689276
p-value	0.075912

Ante esto, se ha aplicado primera diferencia logarítmica:



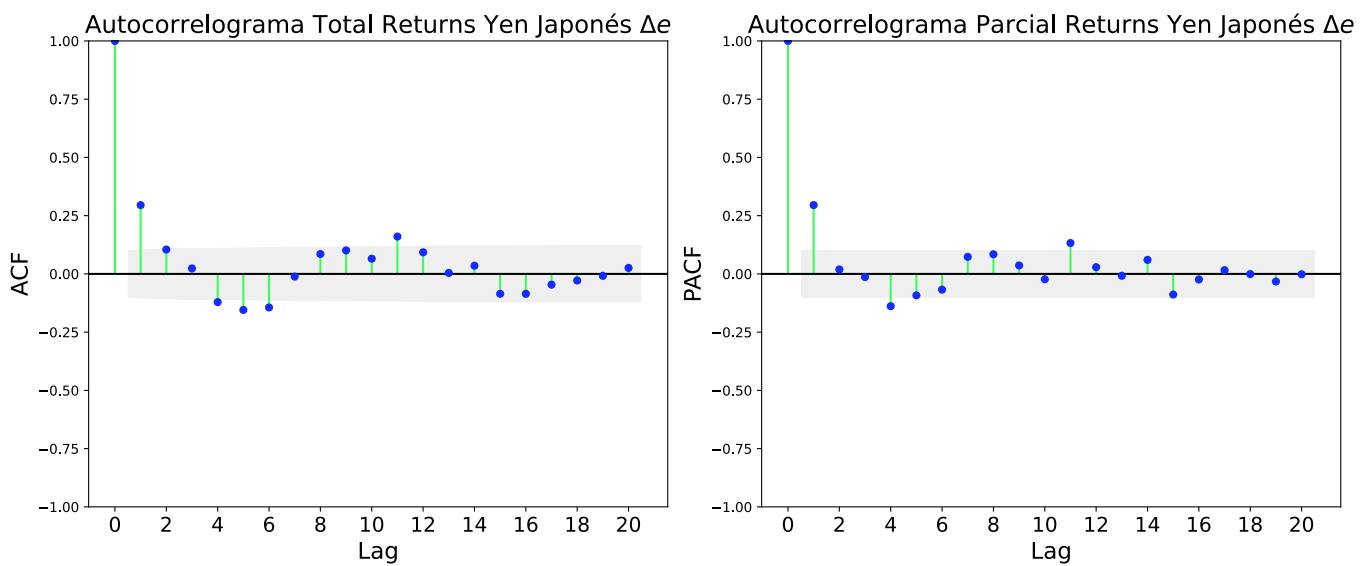
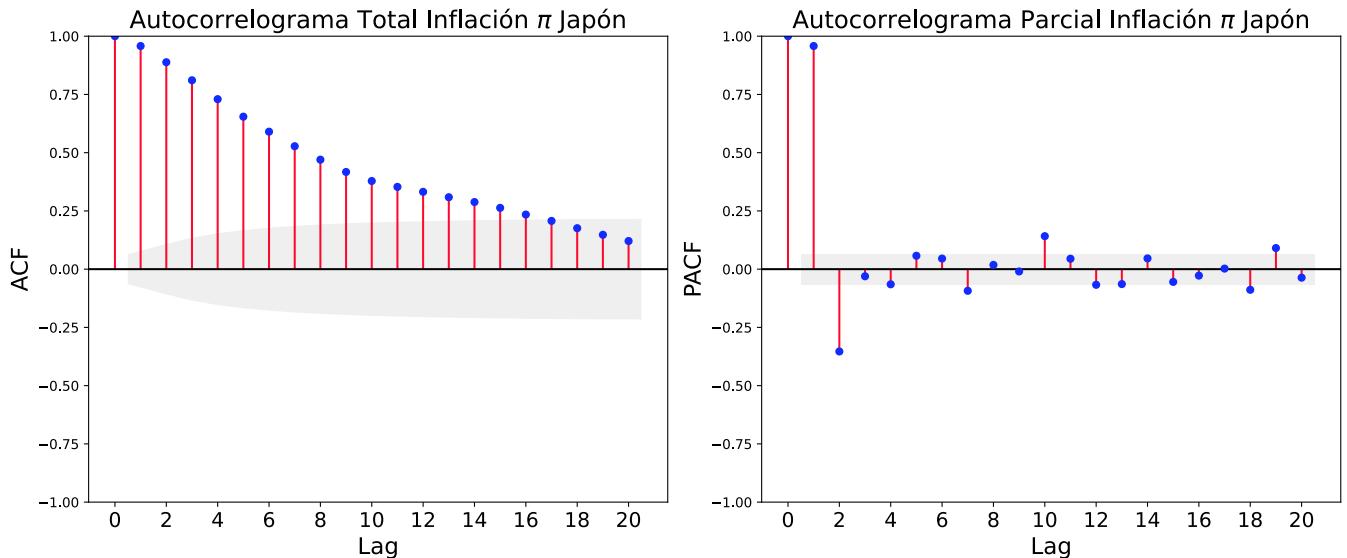
Luego de realizar la transformación, se testea estacionariedad al retorno del Yen Japonés:

Augmented Dickey-Fuller Test for Japan's Nominal Exchange Rate Returns

ADF Statistic	-6.510359
p-value	0.000000

Así, podemos establecer que la última serie ya es estacionaria: p-value α .

A continuación, en línea con Box-Jenkins, se computan los autocorrelogramas empíricos de las series:



Se observa que el autocorrelograma total (ACF) de la inflación muestra un progresiva disipación, mientras que el autocorrelograma parcial (PACF) muestra una súbita disipación en el segundo rezago. Por lo tanto, el modelo para esta serie según el Criterio de Box-Jenkins sería un $AR(2)$.

Se observa que el ACF muere súbitamente en el primer rezago, lo que es propio de un componente $MA(1)$ dominante, con solo algunos rezagos anecdóticamente distintos de cero estadísticamente. Por otro lado, el comportamiento se repite en el PACF, lo que es propio de un $AR(1)$ dominante.

Un modelo $AR(p)$ posee la característica principal de que su comportamiento en el ACF es decayente suavemente, por lo que de determinar un proceso $AR(1)$ en vez de un $ARMA(1,1)$ podría producir problemas de acuerdo a Box-Jenkins, y una calibración problemática. Sin embargo, el ejercicio replicado se da sólo en un

contexto AR . Ante esto, se tomó la opción consciente de sólo establecer el componrnyr $AR(\cdot)$ para el Retorno del Tipo de Cambio de Japón según el PACF. Así, el modelo “verdadero” a considerar en Montecarlo será un $AR(1)$.

A continuación, se estiman los coeficientes de los modelos escogidos para obtener los valores numéricos de los coeficientes, en línea con Box-Jenkins. Los resultados para la Inflación son los siguientes:

Variable	coef.	std. err.	p-value
ϕ_0	0.1670	0.0600	0.0054
ϕ_1	1.2631	0.0228	0.0000
ϕ_2	-0.3215	0.0221	0.0000

Cuadro 2: Resultados Estimación $AR(2)$ Inflación Japón.

Debido a que la estimación corresponde a un $AR(2)$ se ponen a prueba las raíces del modelo, obteniendo los siguientes resultados:

Root	Real	Imaginary	Modulus
AR.1	1.0994	0.0000	1.0994
AR.2	2.8293	0.0000	2.8293

Cuadro 3: Roots Modelo $AR(2)$ Inflación Japón.

```
# Coeficientes Lag Polinomio
ar_params = modelo_estimado.params.filter(like='L')

# Characteristic equation polynomial: 1 - phi_1*L - phi_2*L^2 = 0
ar_poly = np.r_[1, -ar_params.values]

# Compute the roots
roots = np.roots(ar_poly)
roots = (1 / roots)

# Roots consistentes con la documentación oficial:
# https://www.statsmodels.org/v0.12.2/generated/statsmodels.tsa.ar_model.ARResults.roots.html
```

De acuerdo a la documentación oficial del test usado¹⁷: ARResults.roots, “*Stability requires that the roots in modulus lie outside the unit circle.*”. Al encontrar que todas las raíces del polinomio se encuentran fuera del círculo unitario en la columna **Modulus**, el proceso es estacionario y podemos continuar.

A continuación, se estiman los coeficientes del Tipo de Cambio:

¹⁷Disponible en https://www.statsmodels.org/v0.12.2/generated/statsmodels.tsa.ar_model.ARResults.roots.html.

Variable	coef.	std. err.	p-value
ϕ_0	-0.0154	0.1162	0.8943
ϕ_1	0.2972	0.0496	0.0000

Cuadro 4: Resultados Estimación AR(1) Retornos Tipo de Cambio Yen.

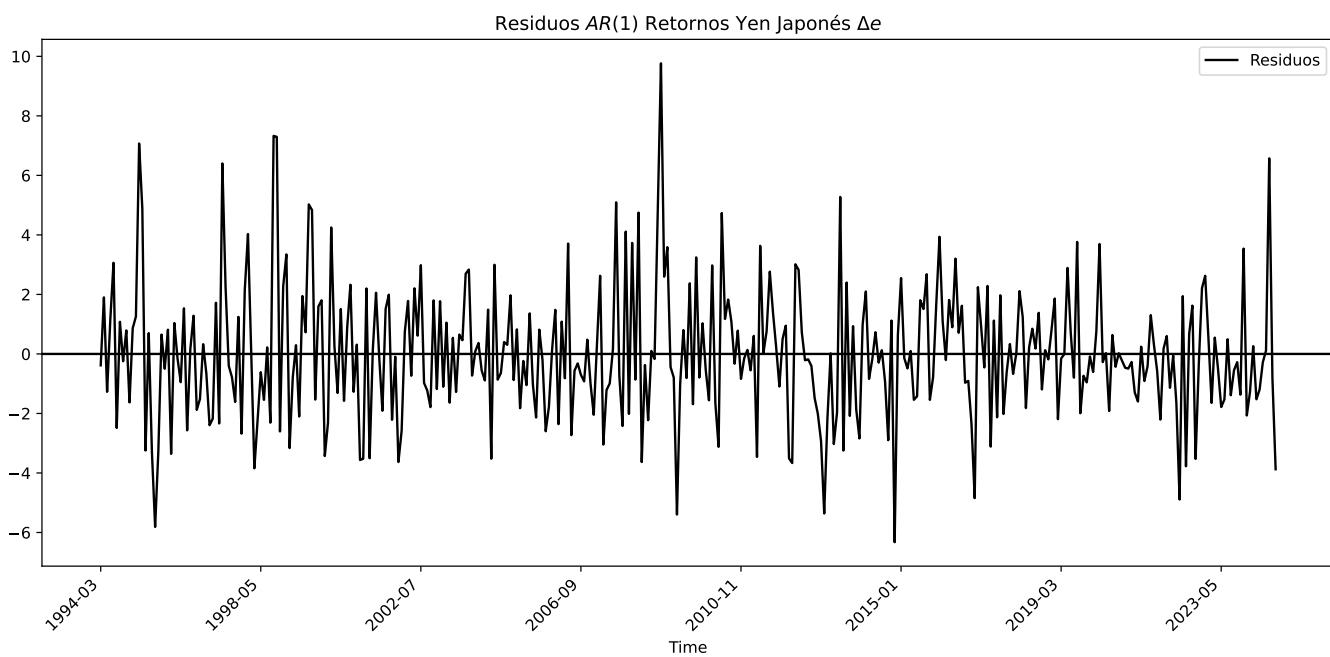
Debido a que la estimación corresponde a un $AR(1)$ se ponen a prueba las raíces del modelo, obteniendo los siguientes resultados:

Root	Real	Imaginary	Modulus
AR.1	3.3649	0.0000	3.3649

Cuadro 5: Roots Modelo $AR(1)$ Tipo de Cambio Japón.

De acuerdo a la documentación oficial del test usado, la raíz del polinomio se encuentra fuera del círculo unitario en la columna **Modulus**, por lo tanto el proceso es estacionario y podemos continuar.

Luego, se estiman los residuos proyectados por el modelo.



Debido a la inclinación por haber escogido un modelo $ARMA(1,1)$ de no ser por el marco restrictivo del contexto AR , se realiza test de ruido blanco para validar la estimación ante riesgo de *misspecification*, encontrando evidencia para la validación:

El diseño de las hipótesis del test está dado por:

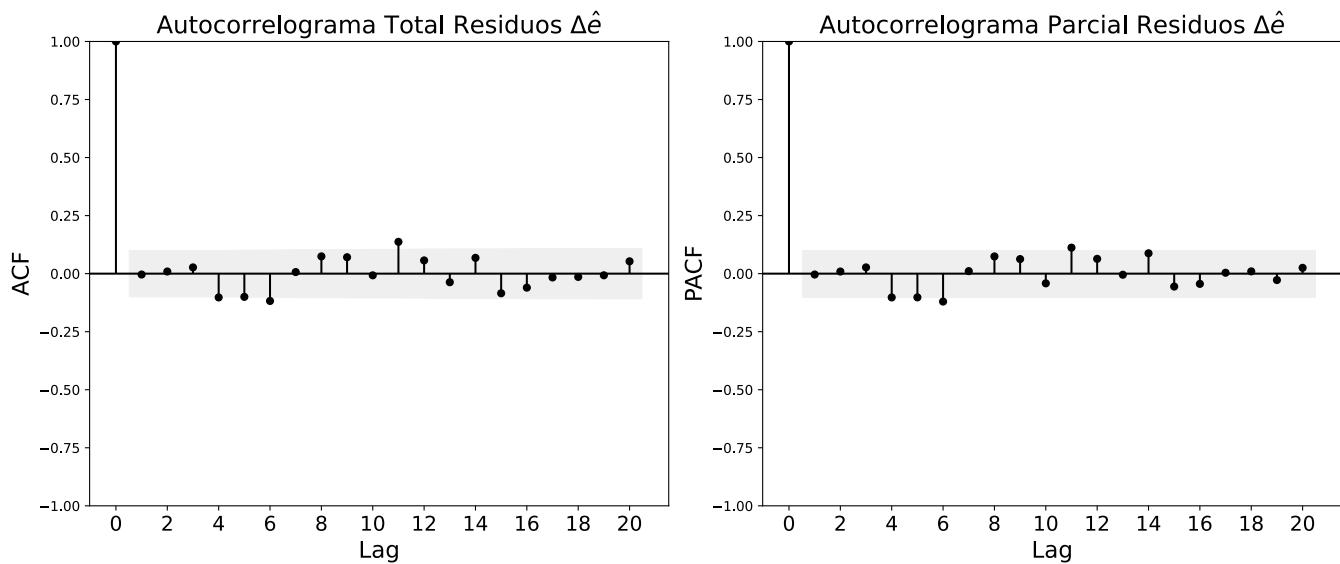
H_0 : La serie sigue un modelo de ruido blanco (i.e., no hay estructura adicional).

H_1 : La serie sigue un modelo ARMA(1,1) con una estructura significativa.

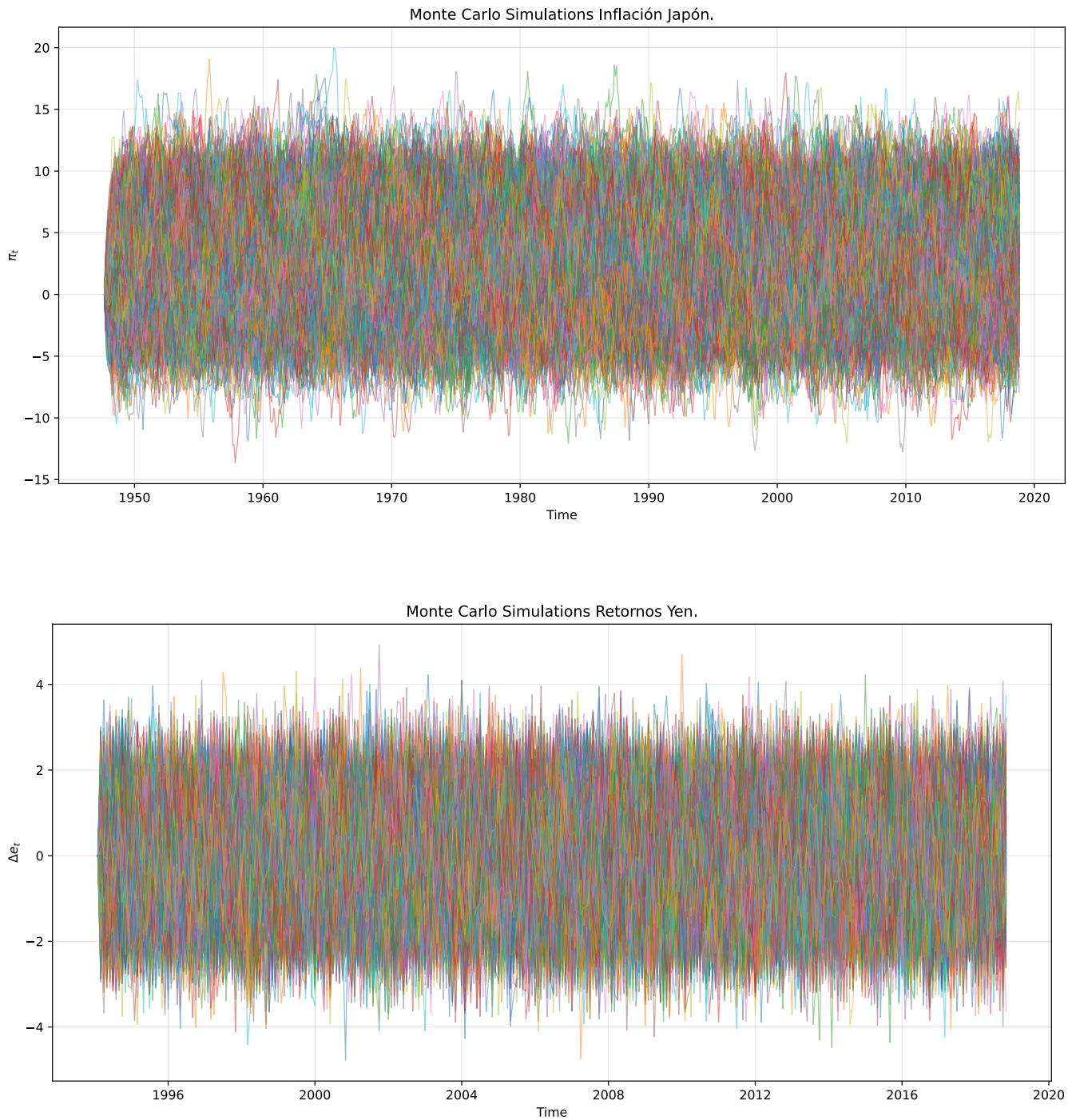
Likelihood Ratio Test (LRT) para ARMA(1,1)

Estadístico LRT	6318.267559
p-valor	0.000000

Se encuentra evidencia de que los residuos son ruido blanco, en línea con los autocorrelogramas de los residuos:

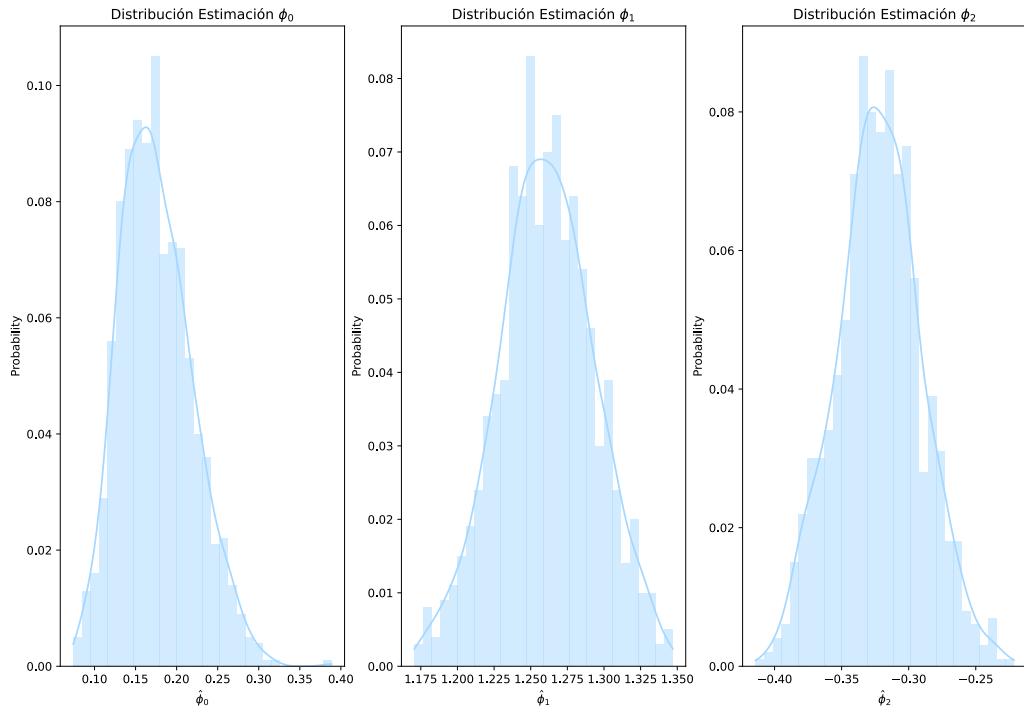


Luego, teniendo los modelos a establecer como DGP de ambas series, se simulan 1000 replicaciones de Montecarlo para Inflación y Retorno del Tipo de Cambio:

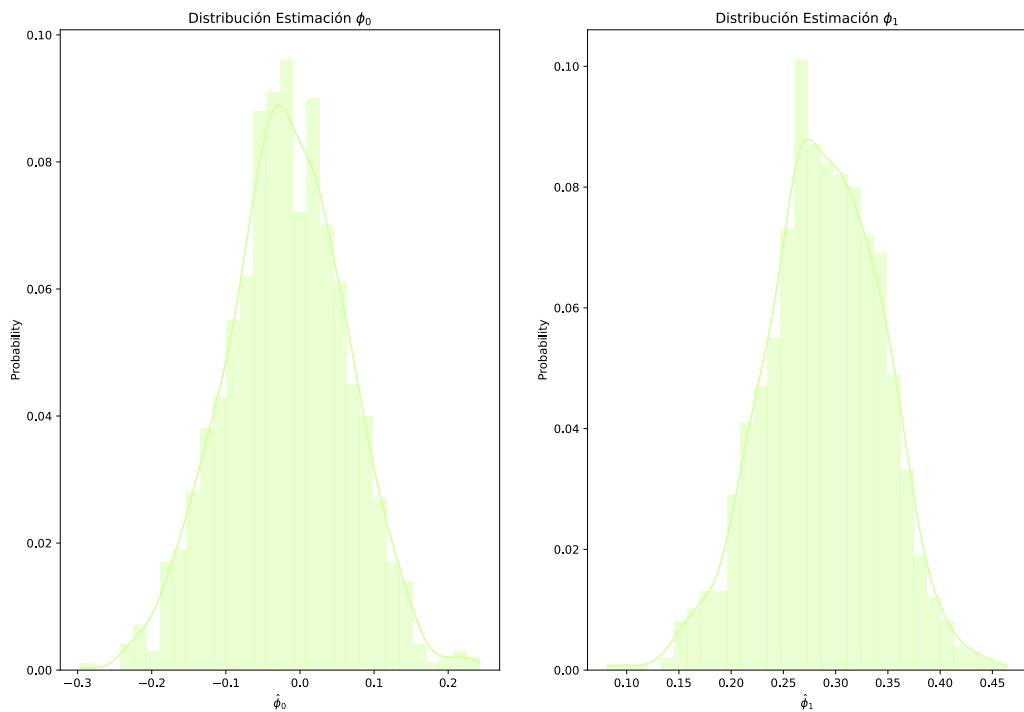


A continuación, se da paso a estimar el modelo verdadero para cada una de las simulaciones. Las distribuciones empíricas de las estimaciones son las siguientes:

Distribuciones de $\hat{\phi}_0$, $\hat{\phi}_1$, y $\hat{\phi}_2$, Inflación Japón



Distribuciones de $\hat{\phi}_0$ y $\hat{\phi}_1$, Retornos Yen



Se puede observar, tal como se esperaba, que los coeficientes están bastante concentrados alrededor de los verdaderos parámetros (notar cotas de eje horizontal). La estimaciones con Montecarlo son consistentes.

Pregunta 2.d

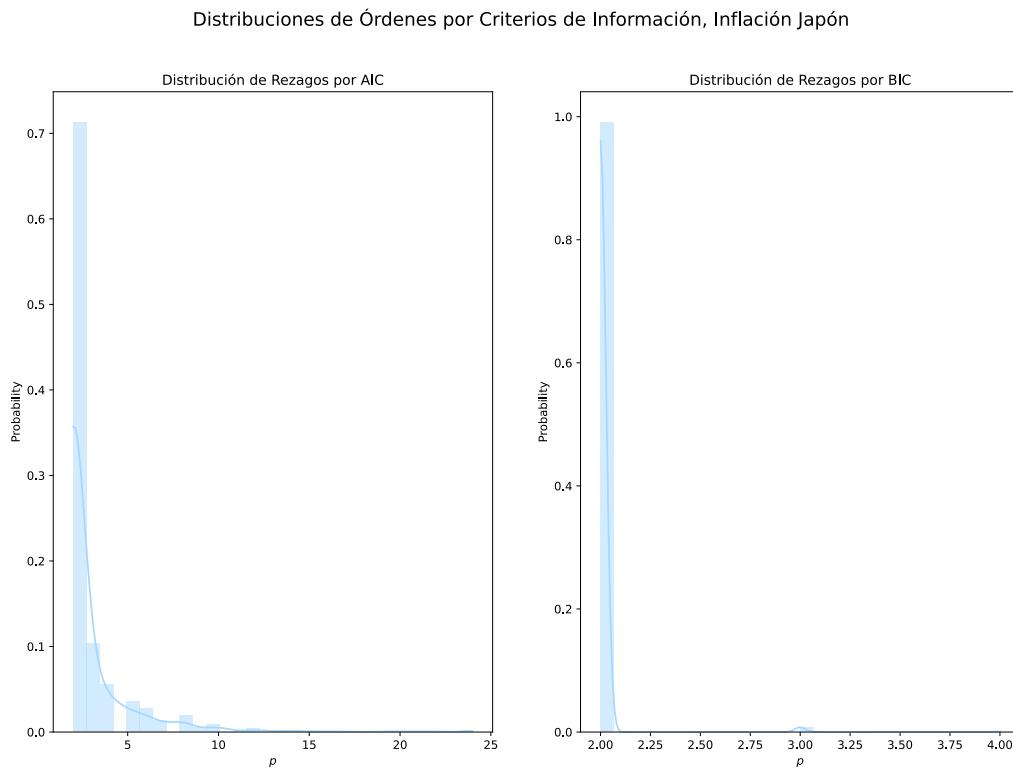
Es importante señalar que este procedimiento presenta una naturaleza en la cual su **complejidad algorítmica y computacional es alta**, traduciéndose en un gran tiempo de ejecución, y en casos, no convergencia. Esto obligó a probar distintos métodos, algoritmos y algoritmos (python y MATLAB) hasta encontrar solución.

Finalmente, se encontró solución en una función creada por Willi Mutschler¹⁸, logrando reducir la complejidad a sólo unos segundos gracias a su función `LagOrderSelectionARp(y, pmax, crit)`. El código final que ejecuta esta pregunta es `pregunta_2.m`, el cual hace uso de la función en `LagOrderSelectionARp.m`.

A continuación, se despliega el tiempo de ejecución de la funcionalidad de selección de rezagos:

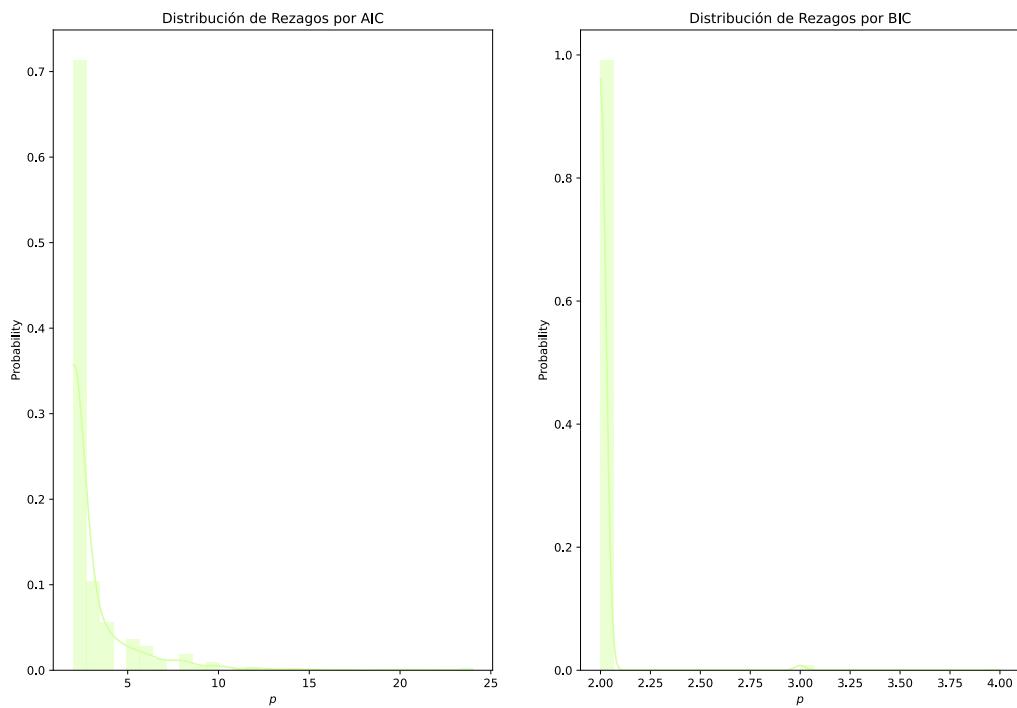
- ```
1 Duration execution Loops AIC/BIC for AR(2): 6.60 seconds
2 Duration execution Loops AIC/BIC for AR(1): 4.09 seconds
```

A continuación, se despliega la distribución del orden seleccionado por cada uno de los criterios:



<sup>18</sup> <https://www.youtube.com/watch?v=9IRfJ63fxpc>

### Distribuciones de Órdenes por Criterios de Información, Retornos Yen



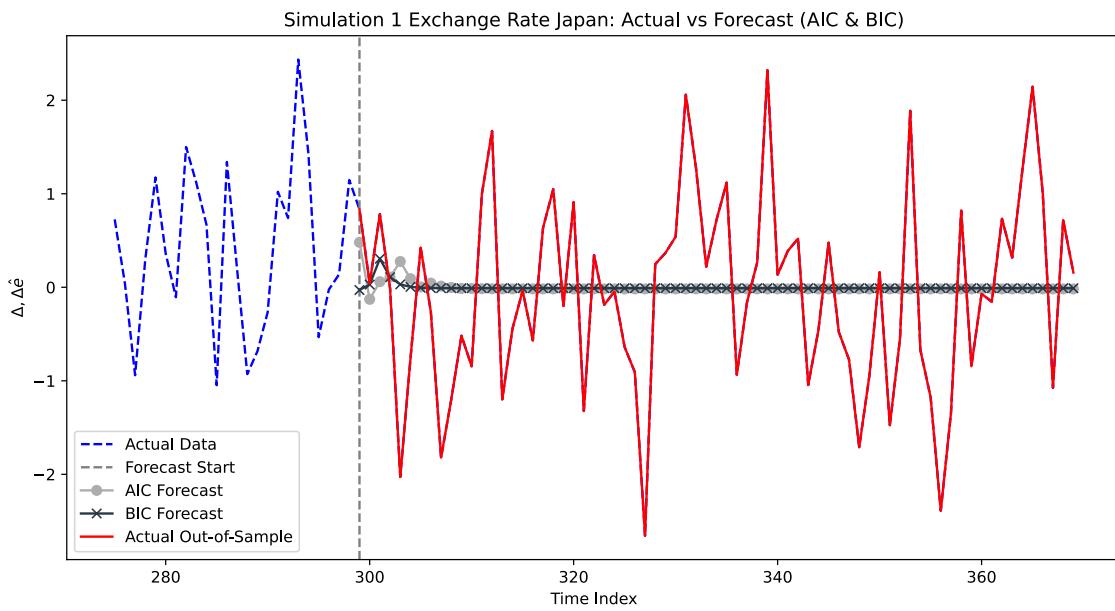
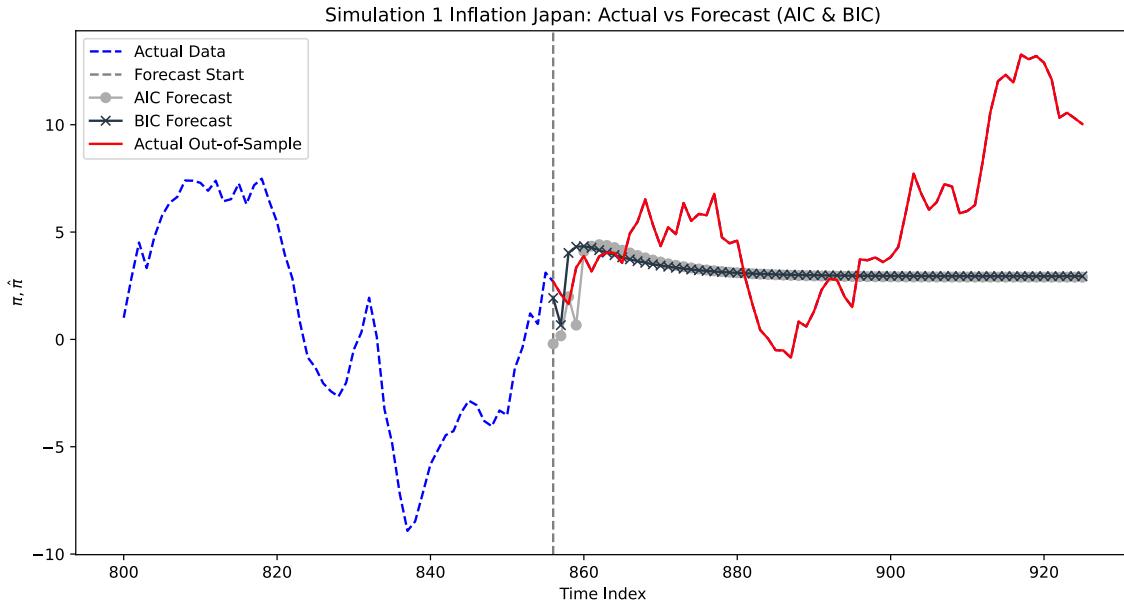
Podemos observar que ambos criterios se acercan al número de parámetros verdadero, 2 y 1 respectivamente, pero se observa tendencia a la sobreestimación por parte del AIC, en línea con los resultados de Medel, 2016.

### Pregunta 2.e

A continuación, se muestra el *dynamic out of sample forecast*<sup>19</sup> para la primera simulación.

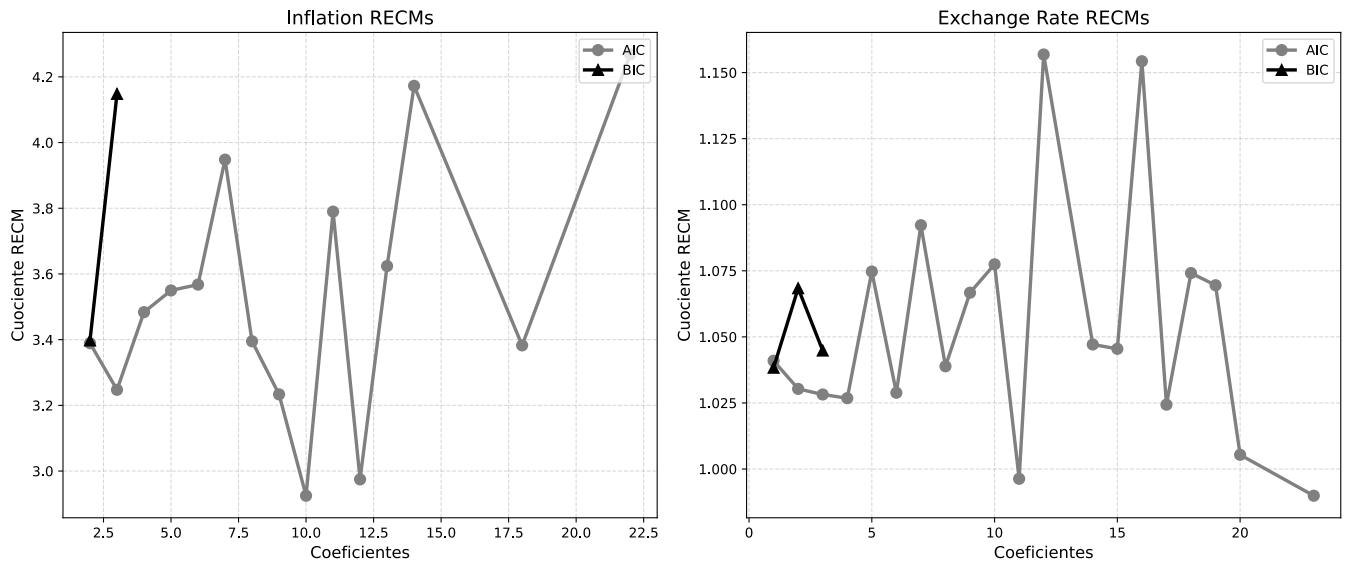
---

<sup>19</sup>Es decir, *forecast que sólo hace uso de la información disponible al momento de hacer la proyección*



Con la validez provista por la visualización anterior, se procede a graficar el raíz del error cuadrático medio de proyección (*RECM*) en función de los rezagos sobre los cuales se obtuvo cada realización  $p = \{0, \dots, 24\}$ .

A continuación, se despliegan los resultados.



## Pregunta 2.f

Al igual que lo observado en Medel (2016), se observan que los errores son crecientes a medida que se incrementa el número de rezagos. A diferencia de Medel (2016), el BIC tiende a determinar muy pocos parámetros (superconsistencia). En ambos series BIC sólo estimó un parámetro adicional en la muestra, y para el caso de la inflación el BIC nunca subestimó.

En cambio, el AIC presenta claros problemas a la hora de hacer pronósticos con un número de rezagos sobreestimado. Por lo tanto, la recomendación sería evitar el AIC a la hora de hacer *forecasting*.

## Referencias

Bank for International Settlements (2024), Consumer prices, BIS WS\_LONG\_CPI 1.0 (data set),  
[https://data.bis.org/topics/CPI/BIS\\_WS\\_LONG\\_CPI\\_1.0/M.JP.771](https://data.bis.org/topics/CPI/BIS_WS_LONG_CPI_1.0/M.JP.771)

Bank for International Settlements (2024), Effective exchange rates, BIS WS\_EER 1.0 (data set),  
[https://data.bis.org/topics/EER/BIS\\_WS\\_EER\\_1.0/M.N.B.JP](https://data.bis.org/topics/EER/BIS_WS_EER_1.0/M.N.B.JP)

Hamilton, J. (1994), “Time Series Analysis”. Princeton University Press.

Medel, C. (2015). “Probabilidad Clásica de Sobreajuste con Criterios de Información: Estimaciones con Series Macroeconómicas Chilenas”. Revista de Análisis Económico – Economic Analysis Review, Universidad Alberto Hurtado/School of Economics and Business, vol. 30(1), pages 57-72.

Munster, W. (2021), “Exercise: Information Criteria for AR(p) model”. YouTube,  
<https://www.youtube.com/watch?v=9IRfJ63fxpc>.

Shumway, R. & Stoffer, D. (2017), “Time Series Analysis and Its Applications: With R Examples”. Springer Texts in Statistics.

Soto, R. (2024), “Econometría de Series de Tiempo (Parte 2, 2024)” [Slides Teoría Econométrica III]. Pontificia Universidad Católica de Chile.

Statsmodels (2020), “ARResults.roots — Statsmodels v0.12.2 Documentation”. Statsmodels.org,  
[www.statsmodels.org/v0.12.2/generated/statsmodels.tsa.ar\\_model.ARResults.roots.html](http://www.statsmodels.org/v0.12.2/generated/statsmodels.tsa.ar_model.ARResults.roots.html)

Turén, J. (2023), “Introduction”[Slides Macroeconomía Aplicada]. Pontificia Universidad Católica de Chile.

Turén, J. (2023), “Univariate Processes”[Slides Macroeconomía Aplicada]. Pontificia Universidad Católica de Chile.