



UNIVERSIDAD COMPLUTENSE
MADRID

ECONOMETRÍA APLICADA



Facultad de Ciencias
Económicas y Empresariales

2

ANÁLISIS UNIVARIANTE DE SERIES TEMPORALES

José Alberto Mauricio

Departamento de Análisis Económico y Economía Cuantitativa



EctrAp-JAM-TR2.pdf

Copyright © 2012 - 2021 J.A.M.
ucm.randomshock.com/ectrap

Versión 8.0 - Septiembre 2021

ANÁLISIS UNIVARIANTE DE SERIES TEMPORALES

BIBLIOGRAFÍA



Box-Jenkins-Reinsel-Ljung (2016): Capítulos 1-9 y 13 – Wei (2006): Capítulos 1-10.

Abraham-Ledolter (1983): Capítulos 5-6 – Vandaele (1983): Capítulos 1-9 y 14.

Introducción al Uso de EViews 4.1 (2021 J.A.M.): Secciones 17 y 19. Apéndices B-D.

En adelante IEV41 ...

CONTENIDO

INTRODUCCIÓN

1

2.1 Objetivos

1

2.2 Series Univariantes

1

2.3 Series Estacionarias y Series No Estacionarias

4

2.4 Modelos Univariantes

8

IDENTIFICACIÓN

13

2.5 Transformaciones de Series No Estacionarias

14

2.6 Propiedades Muestrales de Series Estacionarias

27

2.7 Modelos Estacionarios

43

2.8 Modelos Multiplicativos

69

2.9 Modelos No Estacionarios

73

2.10 Identificación - Resumen

74

ESTIMACIÓN	77
2.11 Estimación - Resumen	77
DIAGNOSIS	84
2.12 Diagnóstico - Resumen	84
[*] PREVISIÓN	89
[*] ANÁLISIS DE INTERVENCIÓN	89

INTRODUCCIÓN

2.1 OBJETIVOS

Describir la evolución temporal observada de una única característica de interés de cierta entidad observable, y **prever** la evolución futura (desconocida) a corto plazo de dicha característica, utilizando tan sólo información (conocida) sobre su pasado. Esta información constituye lo que se conoce como una **serie temporal univariante**.

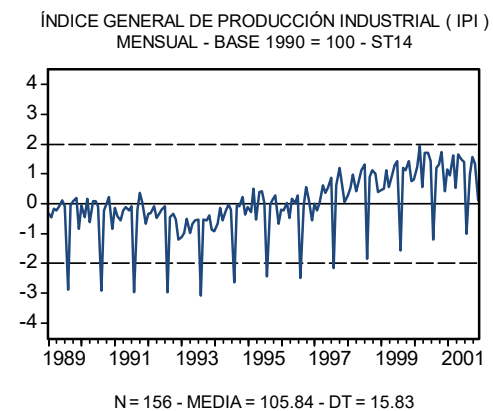
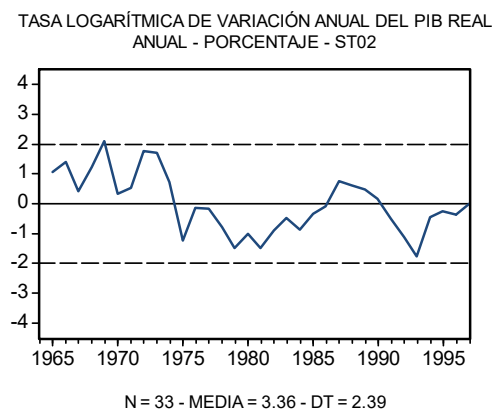
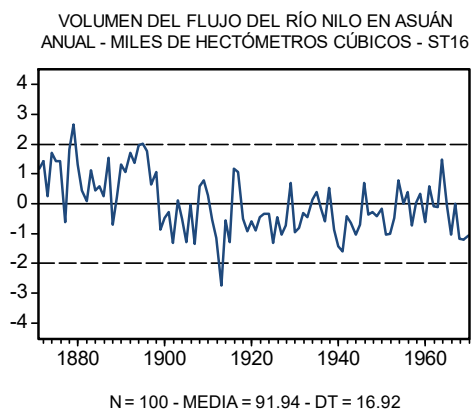
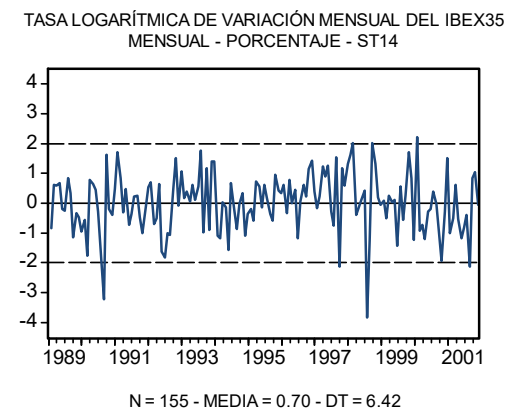
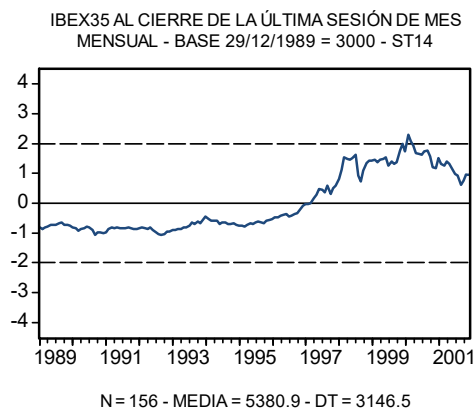
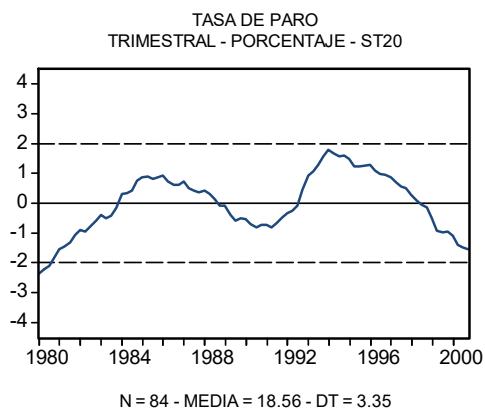
2.2 SERIES UNIVARIANTES

Una **serie temporal univariante** es una secuencia de N observaciones (datos), ordenadas y equidistantes cronológicamente, sobre una única característica de cierta entidad observable en diferentes momentos consecutivos.

Una serie temporal univariante se representa como y_1, y_2, \dots, y_N , $\mathbf{y} = [y_1, y_2, \dots, y_N]'$, $(y_t)_{t=1}^N$, ó $y_t (t = 1, \dots, N)$, donde y_t es la observación número t ($1 \leq t \leq N$) de la serie, y N es el número de observaciones de que consta la serie completa (su tamaño o longitud). En muchas ocasiones, una serie temporal univariante se representa simplemente como y_t .

FIGURA 2.1

Series Temporales Univariantes - Gráficos Estandarizados



SERIES Y GRÁFICOS ESTANDARIZADOS

IEV41: pp. 66-70.

Serie Original:

y_t ($t = 1, 2, \dots, N$): y_1, y_2, \dots, y_N , con

$$\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t, \quad s_y = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})^2}.$$

Serie Estandarizada:

$y_t^* = \frac{y_t - \bar{y}}{s_y}$ ($t = 1, 2, \dots, N$): $y_1^*, y_2^*, \dots, y_N^*$, con

$$\bar{y}^* = 0, \quad s_y^* = 1.$$

① Una serie estandarizada carece de unidades de medida. Por lo tanto, dos o más series estandarizadas están expresadas siempre en la misma escala, lo que las hace perfectamente comparables entre sí (en especial, gráficamente) con independencia de las unidades de medida de las series originales.

② En el eje vertical de un gráfico estandarizado, los valores $\bar{y}^* \pm v \times s_y^* = \pm v$ para cualquier $v \geq 0$ indican valores de la serie original y_t iguales a $\bar{y} \pm v \times s_y$. En particular, en cada gráfico de la Figura 2.1, la línea horizontal continua en 0 indica valores originales de y_t iguales a \bar{y} , y las líneas horizontales discontinuas en 2 y -2 indican valores originales de y_t iguales a $\bar{y} \pm 2 \times s_y$. Los valores calculados de la media y la desviación típica muestrales de y_t suelen incluirse en la parte inferior del gráfico estandarizado correspondiente [como en la Figura 2.1 con las etiquetas MEDIA (\bar{y}) y DT (s_y)]. ■

2.3 SERIES ESTACIONARIAS Y SERIES NO ESTACIONARIAS

Una **serie temporal** es **estacionaria en media** cuando su evolución presenta un nivel medio general que es constante.

Una **serie temporal** es **estacionaria en varianza** cuando es estacionaria en media y, además, su evolución presenta una dispersión constante alrededor de su nivel medio.

Observación 1: Suele decirse que una serie es estacionaria (sin más calificativos) cuando lo es en varianza. Algunos términos asociados con la **estacionariedad** son **estabilidad**, **homogeneidad**, **equilibrio** y **ausencia de tendencia**.

Una **serie temporal** es **no estacionaria en media** cuando su nivel medio no es constante.

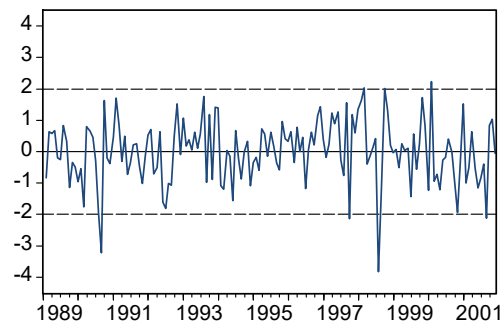
Una **serie temporal** es **no estacionaria en varianza** [i] cuando es no estacionaria en media, o bien [ii] cuando su dispersión alrededor de su nivel medio no es constante, o bien [iii] cuando se dan ambas circunstancias al mismo tiempo.

Observación 2: Los tipos de no estacionariedad más frecuentes son los que tienen que ver con el caso [i] (incluyendo la **estacionalidad**) y con el caso [iii] de la no estacionariedad en varianza. Suele decirse que una serie es no estacionaria (sin más calificativos) cuando no es estacionaria por un sólo motivo o por varios al mismo tiempo. Algunos términos asociados con la **no estacionariedad** son **inestabilidad**, **heterogeneidad**, **desequilibrio** y **tendencia**. Muchas series que son no estacionarias pueden convertirse en series estacionarias mediante la aplicación de transformaciones sencillas.

FIGURA 2.2

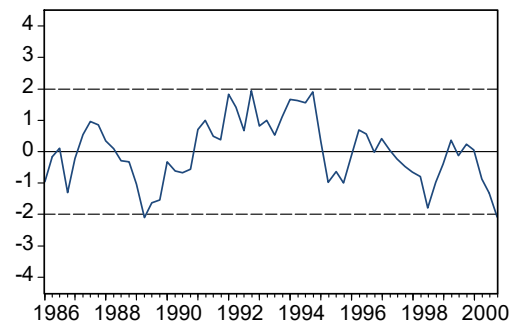
Series Temporales Estacionarias

TASA LOGARÍTMICA DE VARIACIÓN MENSUAL DEL IBEX35
MENSUAL - PORCENTAJE - ST14



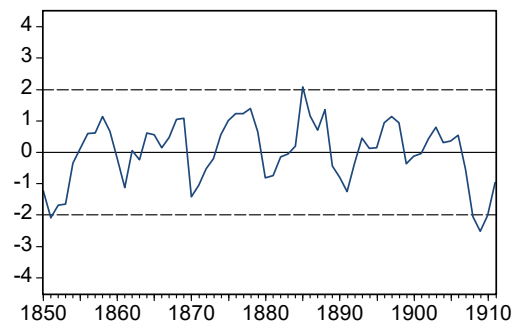
N = 155 - MEDIA = 0.70 - DT = 6.42

DIFERENCIAL DE TIPOS DE INTERÉS B-F EN EEUU
TRIMESTRAL - PUNTOS PORCENTUALES - ST13



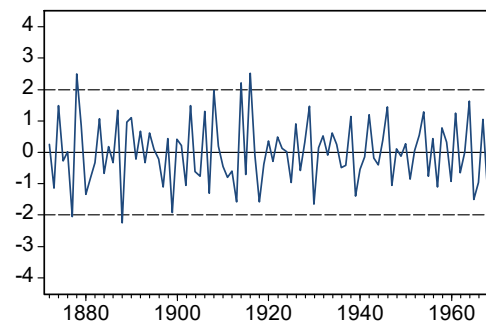
N = 60 - MEDIA = 0.65 - DT = 0.85

VENTAS DE PIELES DE VISÓN - HUDSON'S BAY COMPANY
ANUAL - LOGARITMOS - ST08



N = 62 - MEDIA = 10.79 - DT = 0.40

VARIACIÓN ANUAL DEL FLUJO DEL RÍO NILO EN ASUÁN
MILES DE HECTÓMETROS CÚBICOS - ST16



N = 99 - MEDIA = -0.38 - DT = 16.81

FIGURA 2.3
Series Temporales No Estacionarias

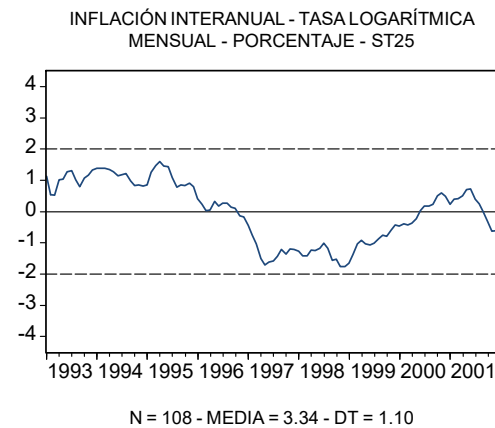
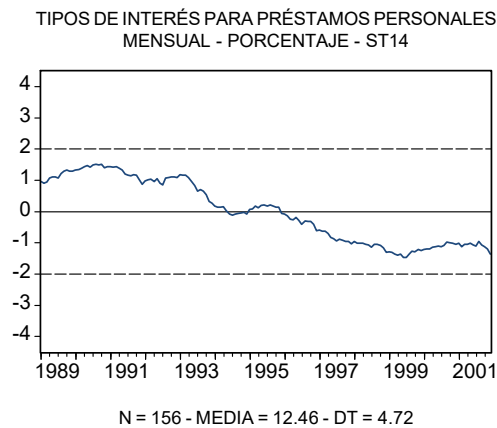
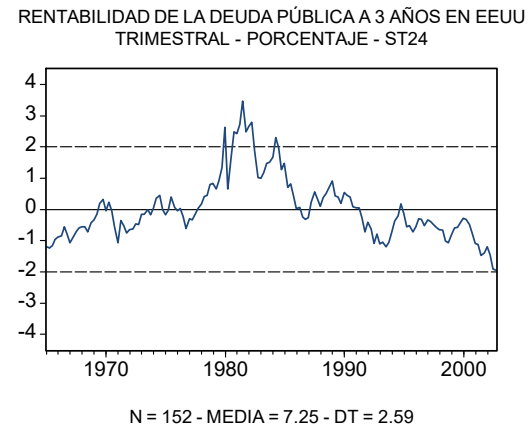
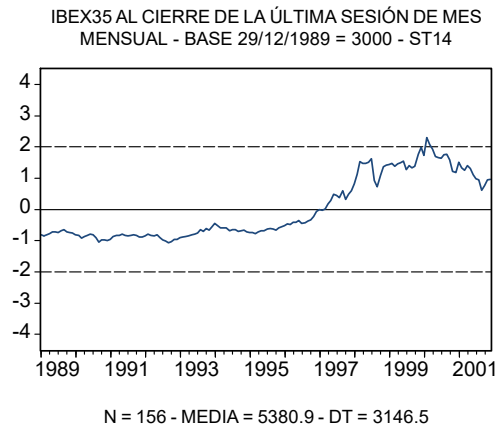
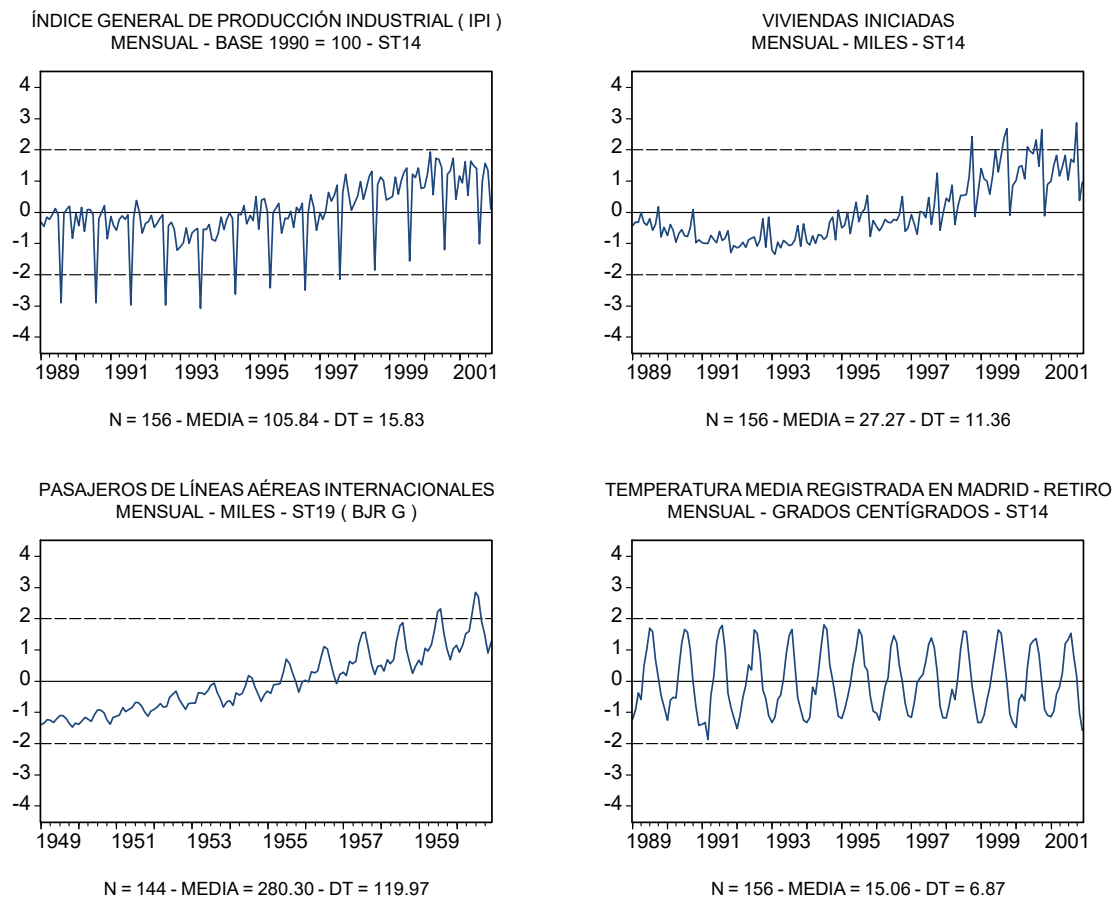


FIGURA 2.4
Series Temporales Estacionales



2.4 MODELOS UNIVARIANTES

Modelo de Ruido Blanco - *White Noise*

En este modelo se supone que una serie estacionaria $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias tales que $Y_t = A_t$, con $(A_t)_{t=1}^N \sim \text{IID}(0, \sigma_A^2)$ (idéntica e independientemente distribuidas). En un modelo de ruido blanco Normal o Gaussiano se supone, además, que cada A_t sigue una distribución Normal.

Observación 3: Una secuencia genérica $(A_t) = (\dots, A_{-1}, A_0, A_1, \dots)$ de variables aleatorias IID con media 0 y varianza $\sigma_A^2 > 0$ se denomina un **proceso** de ruido blanco y se representa como $(A_t) \sim \text{IID}(0, \sigma_A^2)$. Cuando, además, cada A_t sigue una distribución Normal, la secuencia (A_t) es un **proceso** de ruido blanco Normal o Gaussiano, que se representa como $(A_t) \sim \text{NIID}(0, \sigma_A^2)$.

Modelo AR(1) - Autorregresivo (*AutoRegressive*) de Orden 1

En este modelo se supone que una serie estacionaria $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias tales que

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t \quad (t = 1, 2, \dots, N), \quad [1]$$

donde μ y ϕ_1 son parámetros, $|\phi_1| < 1$ (**estacionariedad**), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Modelo AR(2) - Autorregresivo de Orden 2

En este modelo se supone que una serie estacionaria $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias tales que

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + A_t \quad (t = 1, 2, \dots, N), \quad [2]$$

donde μ , ϕ_1 y ϕ_2 son parámetros, las dos raíces de la ecuación $1 - \phi_1 x - \phi_2 x^2 = 0$ están fuera del círculo unitario (**estacionariedad**), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Modelo MA(1) - Media Móvil (*Moving Average*) de Orden 1

En este modelo se supone que una serie estacionaria $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias tales que

$$Y_t = \mu + A_t - \theta_1 A_{t-1} \quad (t = 1, 2, \dots, N), \quad [3]$$

donde μ y θ_1 son parámetros, $|\theta_1| < 1$ (**invertibilidad**), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Modelo de Paseo Aleatorio - *Random Walk*

En este modelo se supone que una serie **no** estacionaria $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una

realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias tales que

$$Y_t = \mu + Y_{t-1} + A_t \quad (t = 1, 2, \dots, N), \quad [4]$$

donde μ es un parámetro (que puede ser igual a cero), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Observación 4: Un paseo aleatorio es un modelo AR(1) con $\phi_1 = 1$, es decir, un modelo AR(1) **no** estacionario. [Nótese que cuando $\phi_1 = 1$, la **ecuación característica** $1 - \phi_1 x = 0$ asociada con un modelo AR(1) tiene una **raíz unitaria** ($x^* = 1$).] Por otro lado, un paseo aleatorio puede escribirse como $Y_t - Y_{t-1} = \mu + A_t$, de manera que una serie temporal y_t ($t = 1, \dots, N$) **no** estacionaria sigue un paseo aleatorio cuando su **diferencia regular de orden 1** $w_t = y_t - y_{t-1}$ ($t = 2, \dots, N$) es una realización particular de un vector de variables aleatorias $\text{IID}(\mu, \sigma_A^2)$ [donde el parámetro μ puede valer cero, en cuyo caso la serie w_t ($t = 2, \dots, N$) seguiría un modelo de ruido blanco].

Observación 5: Un paseo aleatorio es el modelo más sencillo para una **serie integrada de orden 1**, ó **I(1)** (es decir, una serie no estacionaria cuya diferencia regular de orden 1 es una serie estacionaria); por extensión de esta terminología, una serie estacionaria suele denominarse una serie **serie integrada de orden 0**, ó **I(0)**. La notación $y_t \sim \text{I}(1)$ significa que $(y_t)_{t=1}^N$ es una serie I(1) (integrada de orden 1), mientras que la notación $y_t \sim \text{I}(0)$ quiere decir que $(y_t)_{t=1}^N$ es una serie estacionaria (integrada de orden 0).

Observación 6: Las seis series de la Figura 2.5 han sido **simuladas** (generadas artificialmente) a partir de la ecuación que define el modelo correspondiente, utilizando como punto de partida una secuencia de 150 números aleatorios independientes generados con EViews a partir de una distribución $N(0,1)$ (una serie simulada a partir de un modelo de ruido blanco con $\sigma_A^2 = 1$). Todas estas series se encuentran en el archivo ST26-SIM-SERIES.WF1, donde se describe explícitamente el procedimiento seguido para generar cada una de ellas.

FIGURA 2.5
Series Temporales Simuladas

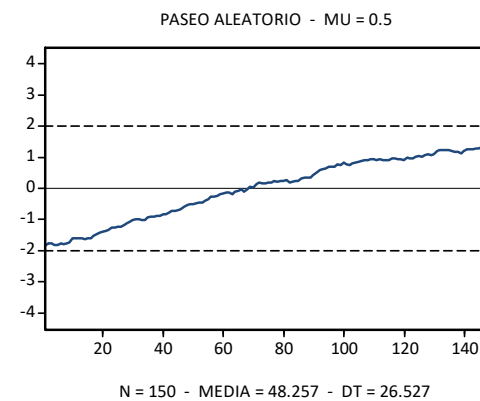
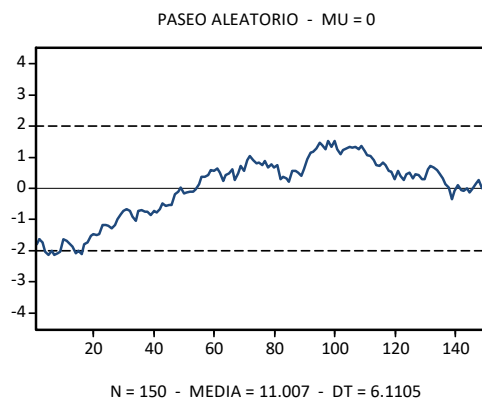
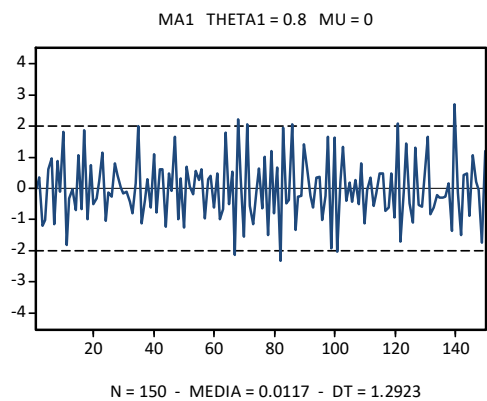
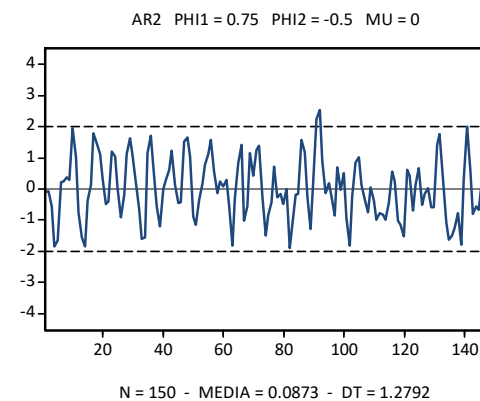
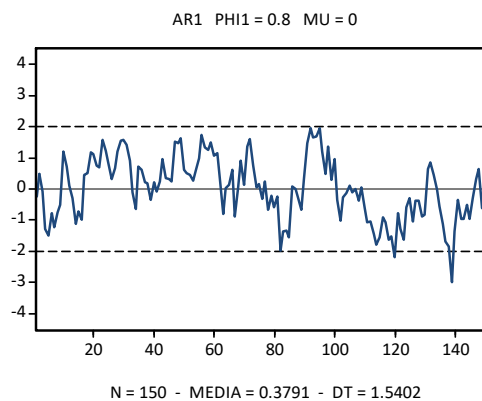
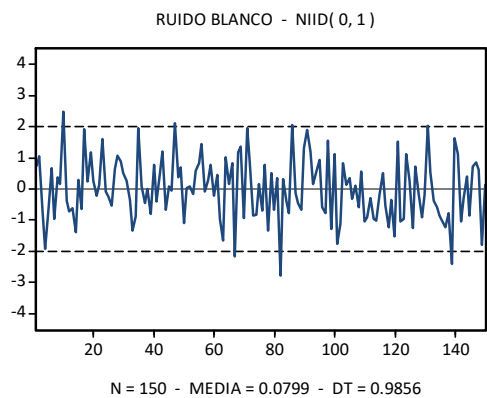
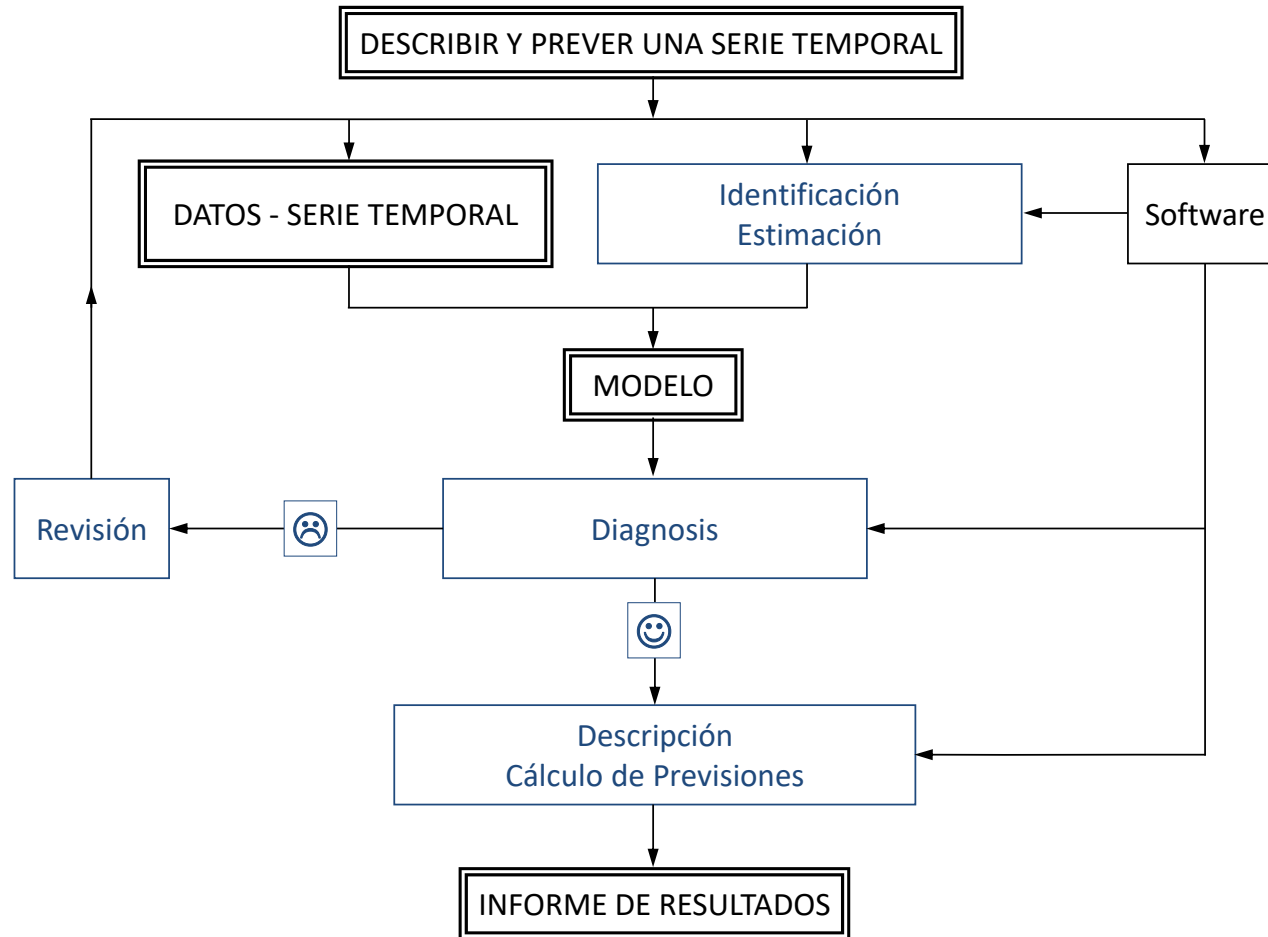


FIGURA 2.6

La Metodología del Análisis Univariante de Series Temporales



IDENTIFICACIÓN

La **identificación** de un modelo inicial (provisional o tentativo) para una serie temporal consiste en el planteamiento de ciertas hipótesis teóricas (estadísticas y matemáticas) que se espera resuman adecuadamente las **propiedades muestrales** más importantes de dicha serie.

La propiedad fundamental de una serie es la que tiene que ver con su carácter estacionario o no estacionario, según el caso.

Cuando una **serie** es **estacionaria**, sus propiedades muestrales más importantes son [i] su **nivel medio**, [ii] su **dispersión** y [iii] su **dinámica** (**inercia** o **memoria**), que pueden resumirse conjuntamente a través de un **modelo univariante estacionario**.

Cuando una **serie** es **no estacionaria**, un enfoque útil consiste en aplicarle primero ciertas **transformaciones** para convertirla en una serie estacionaria y proceder después como en el caso de series estacionarias. Un modelo que se obtiene al juntar dichas transformaciones con un modelo univariante estacionario para la serie transformada (estacionaria) suele denominarse un **modelo univariante no estacionario** para la serie original.

2.5 TRANSFORMACIONES DE SERIES NO ESTACIONARIAS

Aunque existen otras posibilidades, los tipos de no estacionariedad que se observan con más frecuencia en la práctica son los que se indican en la Tabla 2.1.

TABLA 2.1 Tipos Frecuentes de No Estacionariedad			
Nivel Medio General	Estacionalidad	Dispersión No Constante	Dispersión Constante
No Constante	No	A	C
	Si	B	D
Constante	No	F - Poco Frecuente	E - Estacionariedad
	Si	G - Poco Frecuente	H

Observación 7: **Ejemplo A:** Figura 2.1 IBEX 35. **Ejemplo B:** Figura 2.4 Viviendas Iniciadas. **Ejemplo C:** Figura 2.3 Inflación Interanual. **Ejemplo D:** Figura 2.1 Tasa de Paro. **Ejemplo H:** Figura 2.4 Temperaturas Medias.

2.5.1 Estabilización de la Dispersión - Transformación Logarítmica

Cuando la dispersión de una serie temporal no estacionaria y_t ($t = 1, \dots, N$) no es constante porque su **dispersión local** es aproximadamente **proporcional** a su **nivel local**, la dispersión de la **serie transformada logarítmicamente** $y'_t = \ln y_t$ ($t = 1, \dots, N$) suele ser razonablemente estable.

Observación 8: La justificación teórica de esta posibilidad es algo laboriosa, pero se puede ilustrar con ejemplos sobre el efecto de la transformación logarítmica en algunas series reales (Figura 2.7).

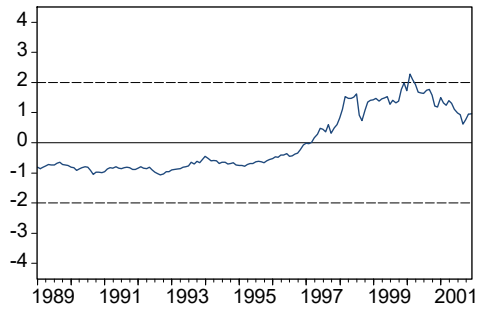
Observación 9: La transformación logarítmica es un caso particular de la llamada **Transformación de Box-Cox**, cuya fórmula general suele escribirse como

$$y'_t = \frac{y_t^m - 1}{m}, \text{ con } |m| \leq 2.$$

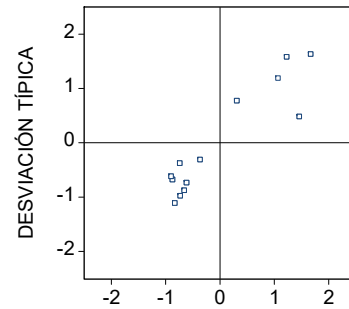
Dado que $\lim_{m \rightarrow 0} [(y_t^m - 1) / m] = \ln y_t$, la elección $m = 0$ proporciona la transformación $y'_t = \ln y_t$. Para otros valores del parámetro m , la fórmula general suele escribirse simplemente como $y'_t = y_t^m$ ("*power transformation*").

Observación 10: Aunque en la transformación general de Box-Cox caben muchas otras posibilidades, con frecuencia la dispersión no constante de series de los tipos A y B en la Tabla 2.1 se puede estabilizar aplicándole a dichas series simplemente una transformación logarítmica. Por lo tanto, la cuestión de cómo transformar una serie para estabilizar su dispersión se reduce, en muchas ocasiones, a decidir si la serie requiere o no una transformación logarítmica, en función de la información que ofrecen su gráfico temporal y su **gráfico desviación típica-media** (Figura 2.7).

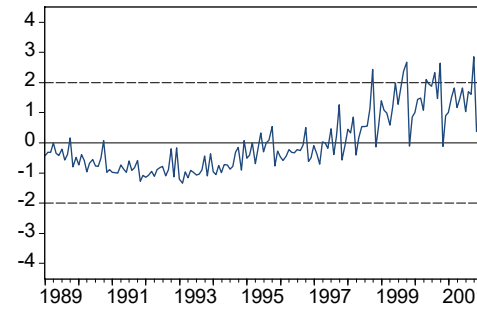
FIGURA 2.7
Transformación Logarítmica - Gráficos Desviación Típica-Media



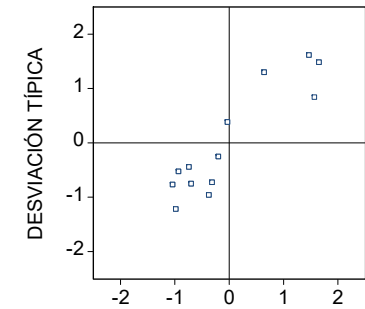
IBEX35 - TIPO A



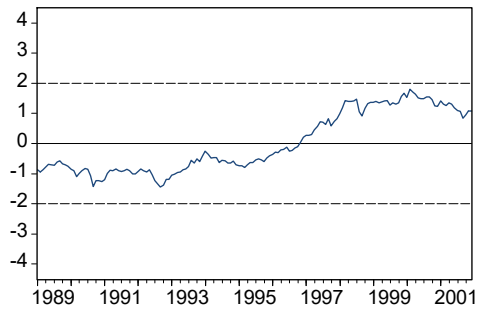
MEDIA



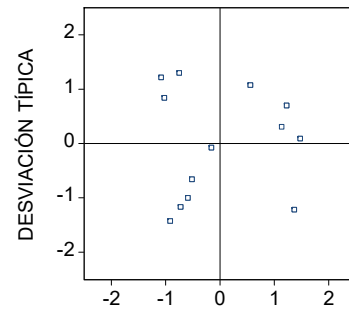
VIVIENDAS INICIADAS - TIPO B



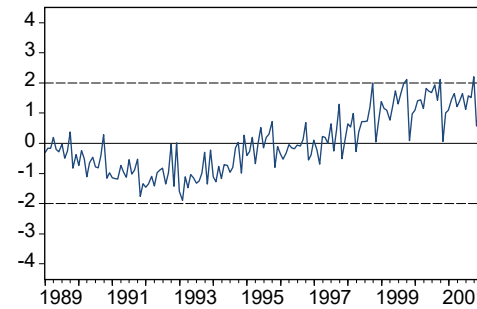
MEDIA



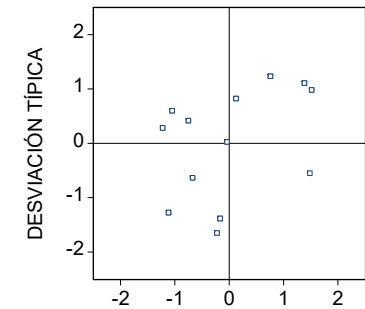
LOG(IBEX35) - TIPO C



MEDIA



LOG(VIVIENDAS INICIADAS) - TIPO D



MEDIA

GRÁFICOS DESVIACIÓN TÍPICA-MEDIA

IEV41: pp. 70-71.

Un **gráfico desviación típica-media** de una serie y_t es la nube de puntos formada por las desviaciones típicas y las medias muestrales locales de y_t en H tramos consecutivos de su historia de tamaños N_1, N_2, \dots, N_H :

$$N \rightarrow \left[\begin{array}{l} N_1 : \left[\begin{array}{c} y_1 \\ \vdots \end{array} \right] \Rightarrow (s_1, \bar{y}_1), \\ N_2 : \left[\begin{array}{c} \vdots \\ \vdots \end{array} \right] \Rightarrow (s_2, \bar{y}_2), \\ \vdots \\ N_H : \left[\begin{array}{c} \vdots \\ y_N \end{array} \right] \Rightarrow (s_H, \bar{y}_H). \end{array} \right].$$

Observación: Por los mismos motivos que los mencionados en relación con el gráfico temporal estandarizado de y_t (unidades de medida, escala), es conveniente estandarizar las series de desviaciones típicas (s_i) y medias (\bar{y}_i) antes de representarlas en el gráfico (nube de puntos) correspondiente.

La utilidad del gráfico desviación típica-media de y_t consiste en ayudar a identificar algún tipo de relación entre la dispersión local y el nivel local de y_t (Figura 2.7), con el fin de transformar y_t para estabilizar su dispersión de acuerdo con las posibilidades teóricas que ofrece la transformación general de Box-Cox. ■

LA TRANSFORMACIÓN DE BOX-COX

Box, G.E.P. y Cox, D.R. (1964), *An Analysis of Transformations*, Journal of the Royal Statistical Society (Series B), Vol. 26, No. 2, pp. 211-252.

Consideramos una serie temporal **no** estacionaria y_t como una realización particular de una secuencia de variables aleatorias Y_t tales que

$$E[Y_t] = \mu_t \text{ y } \text{Var}[Y_t] = \sigma_t^2 \quad (\text{i})$$

existen y **dependen** de t (por lo que **no** son constantes). Si

$$\sigma_t^2 = \sigma^2 \times h(\mu_t)^2 \quad (\text{ii})$$

donde $\sigma^2 > 0$ es una constante y $h(\cdot)$ es una función real tal que $h(\cdot) \neq 0$ para cualquier valor de μ_t , entonces una **transformación estabilizadora de la varianza** de Y_t es cualquier función real $g(\cdot)$ tal que

$$g'(\cdot) = \frac{1}{h(\cdot)} \text{ para cualquier valor de } \mu_t. \quad (\text{iii})$$

Una aproximación de Taylor de primer orden (lineal) a $g(Y_t)$ alrededor de μ_t puede escribirse como $g(Y_t) \cong g(\mu_t) + g'(\mu_t)(Y_t - \mu_t) = [g(\mu_t) - g'(\mu_t)\mu_t] + g'(\mu_t)Y_t$. Por lo tanto, de acuerdo con (i)-(ii),

$$\text{Var}[g(Y_t)] \cong g'(\mu_t)^2 \times \text{Var}[Y_t] = g'(\mu_t)^2 \times [\sigma^2 \times h(\mu_t)^2],$$

que es igual a σ^2 (constante) cuando $g(\cdot)$ satisface (iii). En consecuencia, el propósito de una transformación que satisface (iii) es estabilizar (aproximadamente) la varianza de una secuencia de variables aleatorias Y_t

cuya **dispersión local** depende de su **nivel local** de acuerdo con (ii). Por ejemplo:

① Cuando σ_t es proporcional a μ_t ($\Rightarrow \sigma_t = \sigma\mu_t$), de manera que $\sigma_t^2 = \sigma^2\mu_t^2$, entonces $h(\mu_t) = \mu_t$ en (ii) y $g'(\mu_t) = 1/\mu_t$ en (iii), por lo que $g(\cdot) = \ln(\cdot)$. En este caso, la varianza de la secuencia de variables aleatorias transformadas $Y_t' = \ln Y_t$ es (aproximadamente) constante.

② Cuando σ_t es proporcional a $\sqrt{\mu_t}$ ($\Rightarrow \sigma_t = \sigma\sqrt{\mu_t}$), de manera que $\sigma_t^2 = \sigma^2\mu_t$, entonces $h(\mu_t) = \mu_t^{1/2}$ en (ii) y $g'(\mu_t) = \mu_t^{-1/2}$ en (iii), por lo que $g(\cdot) = 2(\cdot)^{1/2}$. En este caso, la varianza de la secuencia de variables aleatorias transformadas $Y_t' = 2\sqrt{Y_t}$ es (aproximadamente) constante.

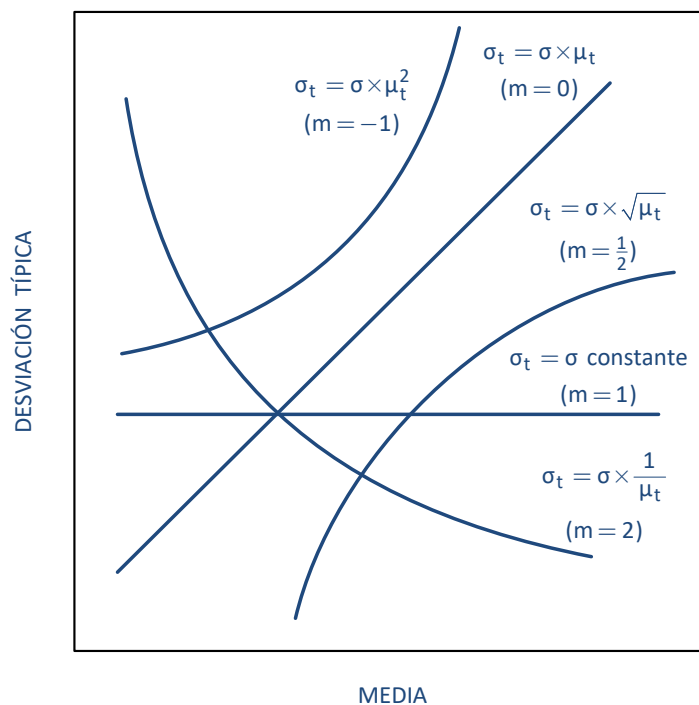
Las dos transformaciones en ① y ② pueden escribirse como

$$Y_t' = g(Y_t) = \frac{Y_t^m - 1}{m}, \quad (\text{iv})$$

que se denomina **transformación de Box-Cox** de parámetro m , donde $|m| \leq 2$ y el -1 del numerador se emplea sólo cuando $m = 0$ porque $\lim_{m \rightarrow 0} [(Y_t^m - 1)/m] = \ln Y_t$. En particular, $m = 0$ proporciona la transformación $Y_t' = \ln Y_t$, mientras que $m = 1/2$ proporciona la transformación $Y_t' = 2\sqrt{Y_t}$.

Observación: El hecho de multiplicar una secuencia de variables aleatorias por una constante no afecta a que su varianza sea o no constante; por lo tanto, la presencia de m en el denominador de la transformación (iv) puede ignorarse en general (lo mismo que el -1 del numerador), de manera que cuando $m \neq 0$, la transformación (iii) suele expresarse simplemente como $Y_t' = Y_t^m$ ("power transformation").

El **gráfico desviación típica-media teórico** de la página siguiente representa las posibilidades de los dos ejemplos anteriores, junto con otras también contempladas en la transformación de Box-Cox. ■



2.5.2 Operador de Retardo

El **operador de retardo** se representa con el símbolo B (a veces L , del inglés *Backshift* o *Lag operator*) y se define de acuerdo con que

$$BX_t = X_{t-1}, B^2X_t = X_{t-2}, \dots, B^dX_t = X_{t-d} \quad (d \geq 1 \text{ entero}), \quad [5]$$

donde X_t es algún objeto referido a un momento t determinado.

2.5.3 Estabilización del Nivel Medio General - Diferenciación Regular

El operador **diferencia regular de orden d** ($d \geq 1$ entero) se representa con el símbolo ∇^d (a veces Δ^d) y se define como $\nabla^d = (1 - B)^d$ (donde B es el operador de retardo), de manera que, en particular,

$$\begin{aligned}\nabla X_t &= (1 - B)X_t = X_t - X_{t-1}, \\ \nabla^2 X_t &= (1 - B)^2 X_t = (1 - 2B + B^2)y_t = X_t - 2X_{t-1} + X_{t-2},\end{aligned}$$

donde X_t es algún objeto referido a un momento t determinado.

Observación 11: En muchas ocasiones, el nivel medio general no constante (tendencia) de series de los tipos C y D en la Tabla 2.1, se puede transformar en un nivel medio general razonablemente estable aplicándole a la serie correspondiente $d \geq 1$ diferencias regulares (en general, no más de 2). Una serie del tipo $\nabla^d y_t$ contiene d observaciones menos (al principio) que la serie y_t de la que procede (Figura 2.8).

2.5.4 Desestacionalización - Diferenciación Estacional

El operador **diferencia estacional de periodo S y orden D** ($D \geq 1$ entero) se representa con el símbolo ∇_S^D (a veces Δ_S^D) y se define como $\nabla_S^D = (1 - B^S)^D$ (donde B es el operador de retardo), de manera que, en particular (cuando $D = 1$),

$$\nabla_S X_t = X_t - X_{t-S},$$

donde X_t es algún objeto referido a un momento t determinado.

Observación 12: En los casos de series estacionales mensuales ($S = 12$ meses) y trimestrales ($S = 4$ trimestres), ∇_{12} y ∇_4 suelen denominarse **diferencias anuales**:

$$\text{Periodo } S = 12 \text{ (mensual):} \quad \nabla_{12} y_t = (1 - B^{12})y_t = y_t - y_{t-12}.$$

$$\text{Periodo } S = 4 \text{ (trimestral):} \quad \nabla_4 y_t = (1 - B^4)y_t = y_t - y_{t-4}.$$

Nótese que $\nabla_S = 1 - B^S \neq (1 - B)^S = \nabla^S$ para todo $S > 1$.

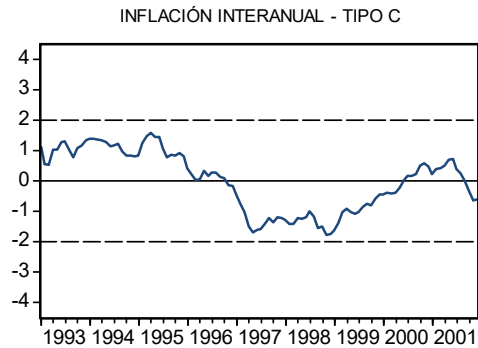
Observación 13: Muchas series estacionales de los tipos D y H en la Tabla 2.1, se pueden transformar en series sin estacionalidad aplicándoles una única diferencia estacional. Una serie del tipo $\nabla_S y_t = y_t - y_{t-S}$ contiene S observaciones menos (al principio) que la serie y_t de la que procede (Figura 2.9).

Observación 14: En conclusión, si una serie temporal y_t ($t = 1, \dots, N$) es no estacionaria por diversos motivos, entonces una serie transformada del tipo

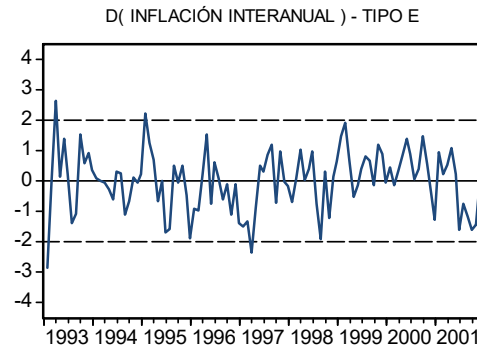
$$w_t = \nabla^d \nabla_S^D y'_t \quad (\text{que consta de } n = N - d - DS \text{ observaciones})$$

puede ser razonablemente estacionaria si la transformación de Box-Cox (con frecuencia $y'_t = y_t$, o bien $y'_t = \ln y_t$), el número d de diferencias regulares (en muchos casos 0, 1 ó 2), y el número D de diferencias estacionales o anuales (en muchos casos 0 ó 1), se escogen adecuadamente (Figura 2.10).

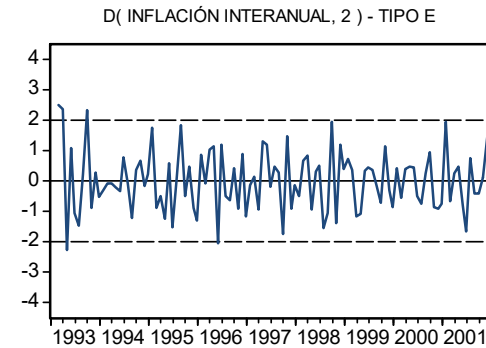
FIGURA 2.8
Diferenciación Regular de Series con Tendencia



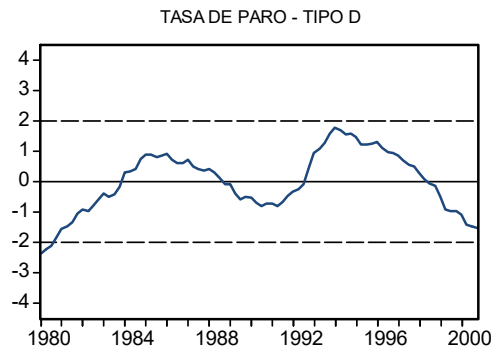
N = 108 - MEDIA = 3.34 - DT = 1.10



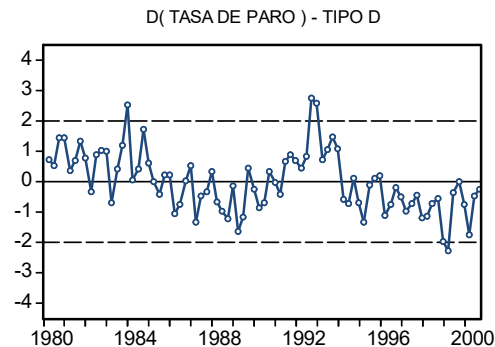
N = 107 - MEDIA = -0.02 - DT = 0.21



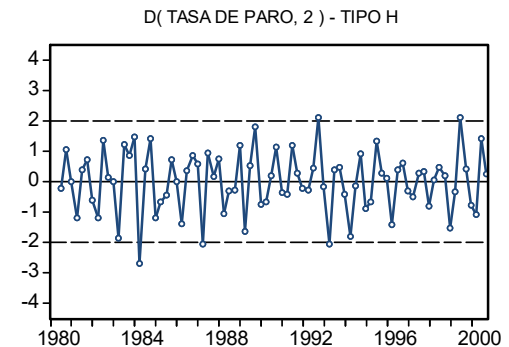
N = 106 - MEDIA = 0.01 - DT = 0.24



N = 84 - MEDIA = 18.56 - DT = 3.35



N = 83 - MEDIA = 0.03 - DT = 0.63



N = 82 - MEDIA = -0.01 - DT = 0.57

FIGURA 2.9
Diferenciación Anual de Series Estacionales

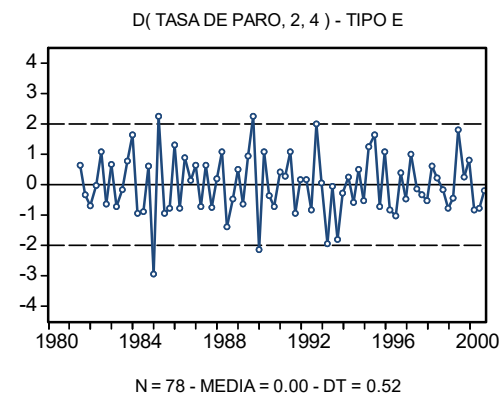
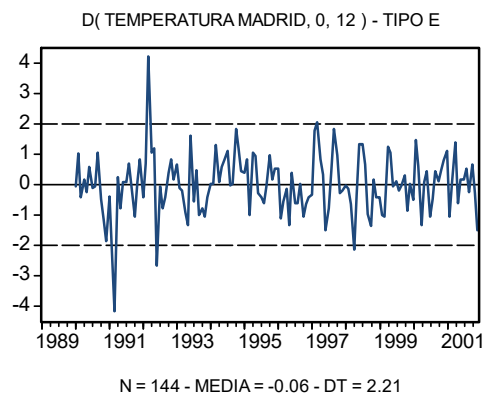
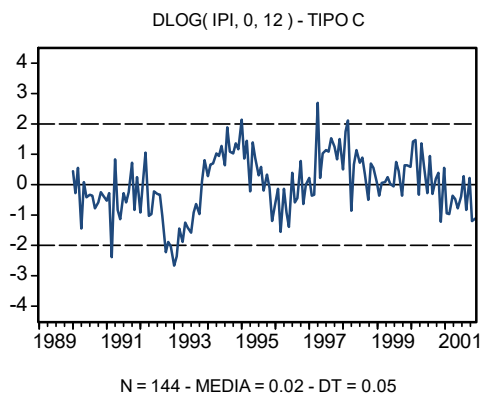
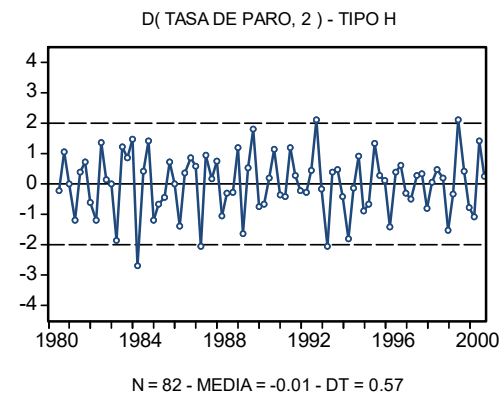
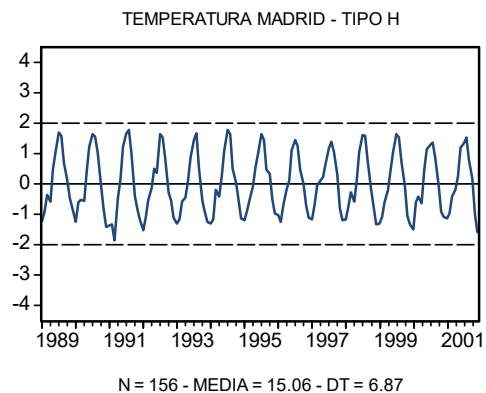
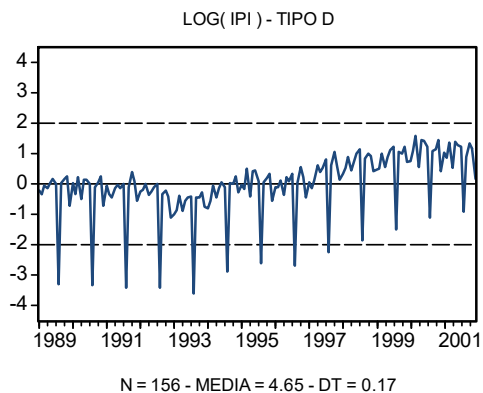
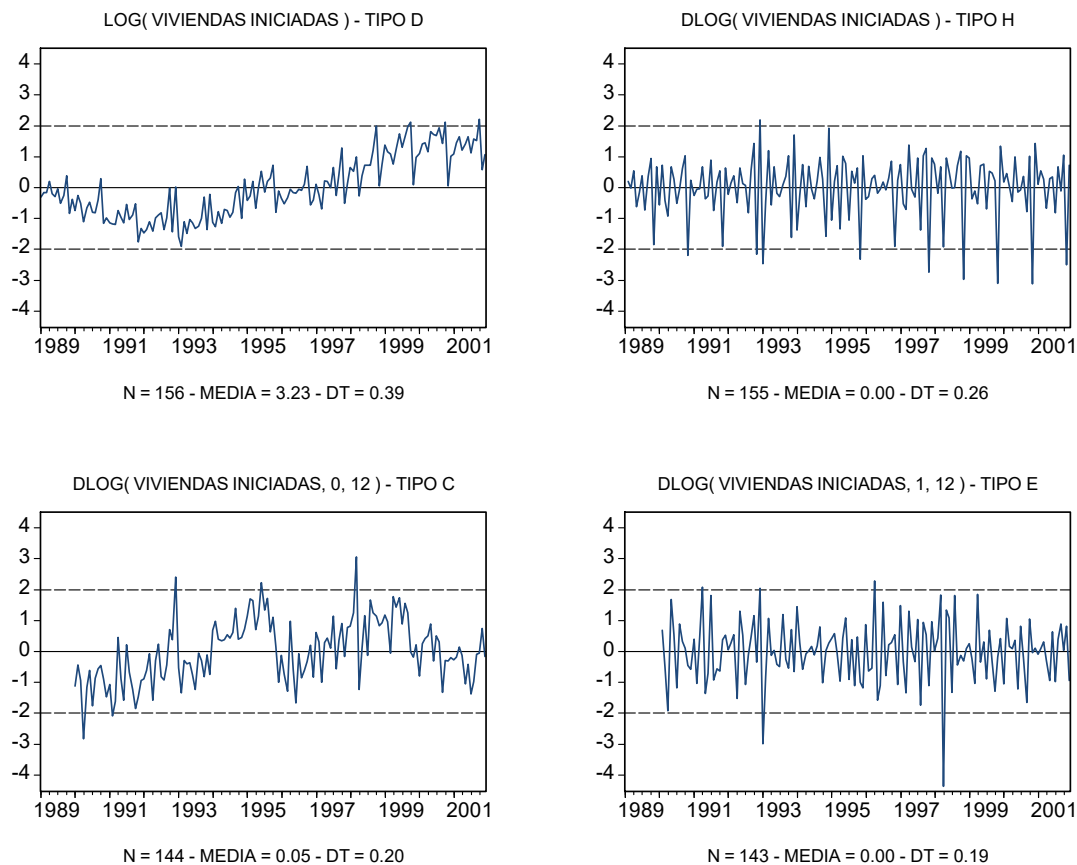


FIGURA 2.10

Efectos de Diferentes Transformaciones sobre una Serie Estacional Mensual



TRANSFORMACIONES

IEV41: pp. 71-74.

La diferenciación regular y la diferenciación estacional son operaciones conmutativas en el sentido de que el orden en el que se aplican sobre una serie es irrelevante (porque son transformaciones lineales). Sin embargo, cuando es necesaria, cualquier transformación de Box-Cox (en concreto, la transformación logarítmica) debe aplicarse siempre en primer lugar (antes de cualquier diferenciación), directamente sobre la serie original (porque cualquier transformación de Box-Cox, incluyendo la transformación logarítmica, es una transformación no lineal).

Las transformaciones requeridas por algunas series no estacionarias de los tipos mencionados en la Tabla 2.1 son las siguientes:

$$\text{Serie ST14 : IBEX35 (Tipo A): } \nabla \ln y_t = \ln y_t - \ln y_{t-1} = \ln \frac{y_t}{y_{t-1}} \cong \frac{y_t - y_{t-1}}{y_{t-1}}.$$

$$\text{Serie ST14 : VIVIN (Tipo B): } \nabla \nabla_{12} \ln y_t = \nabla (\ln y_t - \ln y_{t-12}) = \nabla \left(\ln \frac{y_t}{y_{t-12}} \right) \cong \nabla \left(\frac{y_t - y_{t-12}}{y_{t-12}} \right).$$

$$\text{Serie ST25 : EINF (Tipo C): } \nabla y_t = y_t - y_{t-1}.$$

$$\text{Serie ST20 : TPARO (Tipo D): } \nabla^2 \nabla_4 y_t = \nabla^2 (y_t - y_{t-4}).$$

$$\text{Serie ST14 : TEMP (Tipo H): } \nabla_{12} y_t = y_t - y_{t-12}.$$

Como ocurre en los dos primeros casos, la combinación del ln con alguna/s diferencia/s proporciona series transformadas (estacionarias) que pueden interpretarse en términos de alguna tasa de variación. ■

2.6 PROPIEDADES MUESTRALES DE SERIES ESTACIONARIAS

2.6.1 Nivel y Dispersión

La **media muestral** y la **varianza muestral** de una serie temporal $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ son

$$\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t \quad \text{y} \quad s_y^2 = \frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})^2. \quad [6]$$

La **desviación típica muestral** (s_y) de la serie es la raíz cuadrada de la varianza muestral.

Observación 15: Cuando en un modelo para una serie temporal estacionaria (Apartados 2.4 y 2.7) se supone que $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias con **media teórica** $\mu_Y = E[Y_t]$ y **varianza teórica** $\sigma_Y^2 = \text{Var}[Y_t]$ **constantes**, se puede contrastar la significación de μ_Y con un estadístico t cuyo valor calculado es

$$t = \frac{\bar{y}}{\sqrt{s_y^2/N}},$$

y cuya distribución (aproximada), bajo la hipótesis nula de que $\mu_Y = 0$, es una $t(N-1)$.

2.6.2 Dinámica I - Autocovarianza

La **autocovarianza muestral de orden k** ($k \geq 0$) de una serie temporal $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es la covarianza muestral entre las series y_t e y_{t+k} ($t = 1, \dots, N-k$), o, equivalentemente,

entre las series y_{t-k} e y_t ($t = k + 1, \dots, N$):

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (y_t - \bar{y})(y_{t+k} - \bar{y}) = \frac{1}{N} \sum_{t=k+1}^N (y_{t-k} - \bar{y})(y_t - \bar{y}) \quad (k = 0, 1, 2, \dots). \quad [7]$$

Observación 16: La autocovarianza muestral de orden cero (c_0) es exactamente la varianza muestral (s_y^2) de la serie considerada. Para $k > 0$, la autocovarianza muestral de orden k mide el grado medio de asociación lineal entre cualquier par de componentes de una serie separados entre sí por k períodos (es decir, el grado medio de influencia de cualquier observación y_t sobre y_{t+k} , o, equivalentemente, el grado medio de dependencia de cualquier observación y_t con respecto a y_{t-k}). La secuencia c_0, c_1, c_2, \dots se denomina la **Función de Autocovarianza Muestral** de la serie temporal y_1, y_2, \dots, y_N .

2.6.3 Dinámica II - Autocorrelación Simple

La **autocorrelación simple muestral de orden k** ($k > 0$) de una serie $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una medida **total** o **bruta** del grado de correlación lineal entre cualquier par de componentes de dicha serie separados entre sí por k períodos. Habitualmente (Observación 16),

$$r_k = \text{C}\hat{\text{o}}\text{r}\text{r}[y_{t-k}, y_t] = \frac{c_k}{c_0} \quad (k = 0, 1, 2, \dots) \quad (\Rightarrow r_0 = 1). \quad [8]$$

Observación 17: A diferencia de las autocovarianzas, las autocorrelaciones carecen de unidades de medida y sus valores están comprendidos entre +1 y -1. La secuencia r_1, r_2, \dots se denomina la **Función de Autocorrelación Simple (ACF)**

Muestral de la serie temporal y_1, y_2, \dots, y_N . La representación gráfica de la ACF muestral de una serie se denomina el **correlograma** de dicha serie. La ACF muestral es el instrumento básico para describir en la práctica la **duración** y la **intensidad** de la **dinámica** (**inercia** o **memoria**) de una serie.

Observación 18: Cuando en un modelo para una serie temporal estacionaria (Apartados 2.4 y 2.7) se supone que $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias con **media teórica** $\mu_Y = E[Y_t]$ y **autocovarianzas teóricas** $\gamma_k = \text{Cov}[Y_t, Y_{t+k}] = \text{Cov}[Y_{t-k}, Y_t]$ ($k = 0, 1, 2, \dots$) **constantes**, se puede contrastar la significación individual de cualquier **autocorrelación simple teórica** $\rho_k = \gamma_k / \gamma_0$ utilizando un estadístico cuyo valor calculado es $r_k / \sqrt{1/N}$ y cuya distribución (aproximada), bajo la hipótesis nula de que $\rho_k = 0$, es una $N(0, 1)$ [por lo que suele considerarse que cualquier ρ_k ($k \geq 1$) es *individualmente* significativa al 5% cuando $|r_k| > 1.96 / \sqrt{N}$]. Por otro lado, para decidir si las K primeras autocorrelaciones simples teóricas son *conjuntamente* significativas, suele emplearse el valor calculado del **estadístico de Ljung-Box**

$$Q_{LB} = N(N + 2) \sum_{k=1}^K \frac{r_k^2}{N - k},$$

que, bajo la hipótesis nula de que $\rho_1 = \rho_2 = \dots = \rho_K = 0$, sigue aproximadamente una distribución $\chi^2(K)$.

2.6.4 Dinámica III - Autocorrelación Parcial

La **autocorrelación parcial muestral de orden k** ($k > 0$) de una serie $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una medida **parcial** o **neta** del grado de correlación lineal entre cualquier par de componentes de dicha serie separados entre sí por k períodos, que **no** es debida a la posible correlación entre cada uno de ellos y todos los que se encuentran entre ambos.

Observación 19: Formalmente,

$$r_{11} = r_1, \quad r_{kk} = \text{C\hat{o}rr}[\hat{u}_{t-k}, \hat{u}_t] = \frac{\text{C\hat{o}v}[\hat{u}_{t-k}, \hat{u}_t]}{\text{V\hat{a}r}[\hat{u}_{t-k}]} = \frac{\sum_{t=k+1}^N \hat{u}_{t-k} \hat{u}_t}{\sum_{t=k+1}^N \hat{u}_{t-k}^2} \quad (k \geq 2),$$

donde \hat{u}_t y \hat{u}_{t-k} son las series de residuos MCO de las regresiones de y_t e y_{t-k} ($t = k+1, \dots, N$), respectivamente, sobre las variables intermedias $y_{t-1}, \dots, y_{t-k+1}$. (Nótese que r_{kk} es la estimación MCO de la pendiente en la regresión de \hat{u}_t sobre \hat{u}_{t-k} .) De manera equivalente, r_{kk} ($k \geq 2$) es la estimación MCO del parámetro asociado con y_{t-k} en la regresión de y_t sobre y_{t-1}, \dots, y_{t-k} . Alternativamente, cada r_{kk} ($k \geq 2$) puede calcularse de manera aproximada como el último elemento en el vector del lado izquierdo de la siguiente expresión:

$$\underbrace{\begin{bmatrix} r_{k1} \\ r_{k2} \\ \vdots \\ r_{kk} \end{bmatrix}}_{k \times 1} = \underbrace{\begin{bmatrix} 1 & r_1 & \cdots & r_{k-1} \\ r_1 & 1 & \cdots & r_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & r_{k-2} & \cdots & 1 \end{bmatrix}}_{k \times k}^{-1} \underbrace{\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix}}_{k \times 1},$$

lo que implica, por ejemplo, con $k = 2, 3$, que

$$r_{22} = \frac{r_2 - r_1^2}{1 - r_1^2}, \quad r_{33} = \frac{r_1(r_1^2 + r_2^2 - 2r_2) + r_3(1 - r_1^2)}{(1 - r_2)(1 - 2r_1^2 + r_2)}.$$

Con frecuencia, este cálculo aproximado de la PACF muestral a partir de la ACF muestral correspondiente, se lleva a cabo en la práctica mediante el siguiente procedimiento recursivo:

1. $k = 1$, $r_{11} = r_1$.
2. $k = k + 1$. Calcular la autocorrelación parcial muestral de orden k como

$$r_{kk} = \frac{r_k - \sum_{i=1}^{k-1} r_{k-1,i} r_{k-i}}{1 - \sum_{i=1}^{k-1} r_{k-1,i} r_i}.$$

3. Si $k = K$ (retardo máximo que se desea calcular), parar. En caso contrario, calcular las cantidades auxiliares

$$r_{ki} = r_{k-1,i} - r_{kk} r_{k-1,k-i} \quad (i = 1, 2, \dots, k-1).$$

4. Volver al paso 2.

De esta manera, cada coeficiente de autocorrelación parcial muestral r_{kk} se calcula como una función determinada de los coeficientes de autocorrelación simple muestrales r_1, \dots, r_k ($k = 1, 2, 3, \dots$).

Observación 20: La secuencia r_{11}, r_{22}, \dots se denomina la **Función de Autocorrelación Parcial (PACF) Muestral** de la serie temporal y_1, y_2, \dots, y_N . La ACF y la PACF muestrales constituyen en conjunto el instrumento fundamental para describir en la práctica la **duración** y la **intensidad** de la **dinámica (inercia o memoria)** de una serie temporal.

Observación 21: Cuando en un modelo para una serie temporal estacionaria (Apartados 2.4 y 2.7) se supone que $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ de variables aleatorias con **media teórica** y **autocovarianzas teóricas constantes**, suele considerarse que cualquier **autocorrelación parcial teórica** (Observación 22) es *individualmente* significativa al 5% si $|r_{kk}| > 1.96/\sqrt{N}$ (igual que en la Observación 18 para el caso de cualquier autocorrelación simple teórica).

IEV41: pp. 80-85

FIGURA 2.11
Autocorrelaciones Muestrales de Series Estacionarias Simuladas

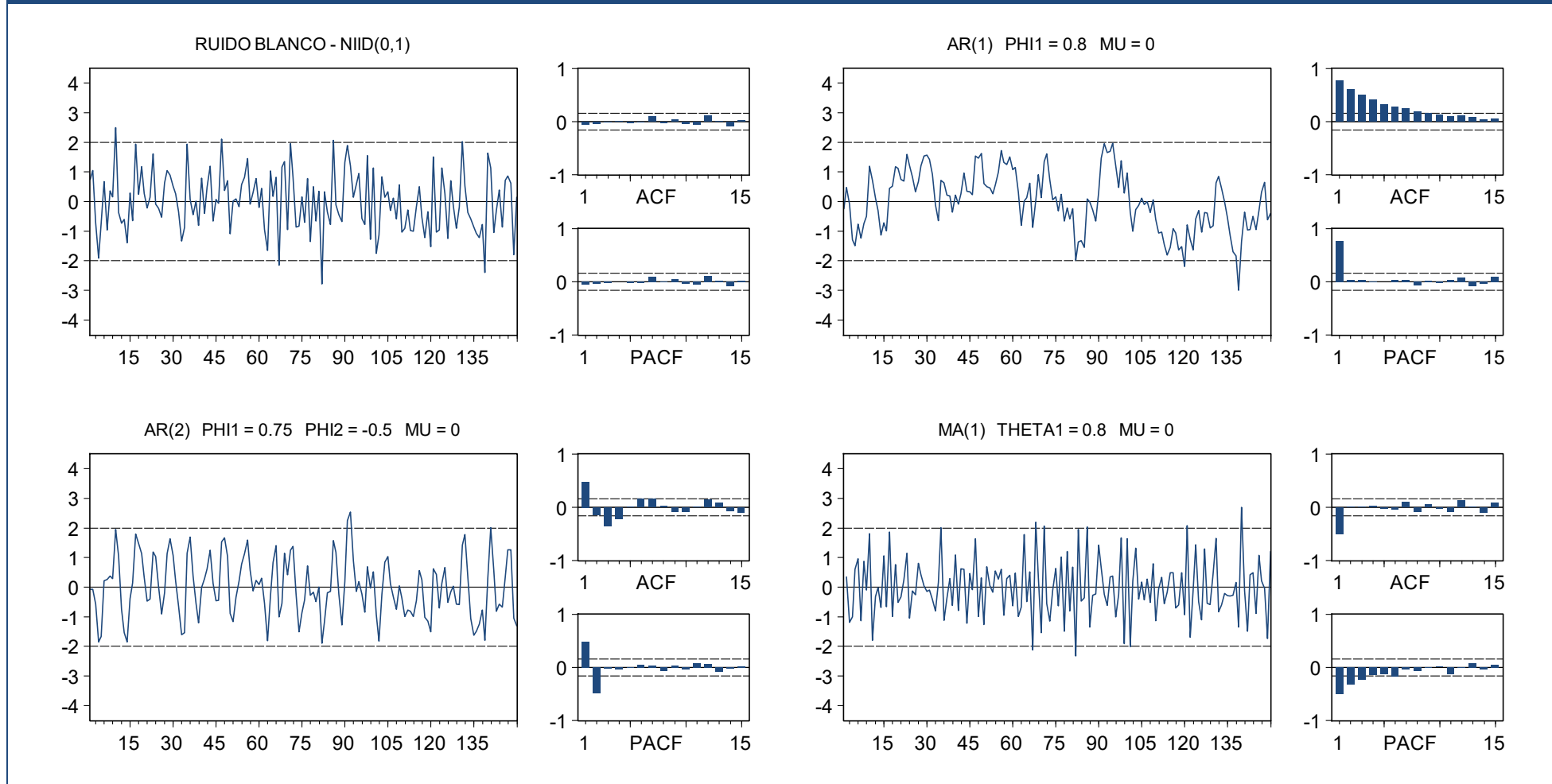


FIGURA 2.12
Autocorrelaciones Muestrales de Series Estacionarias Reales

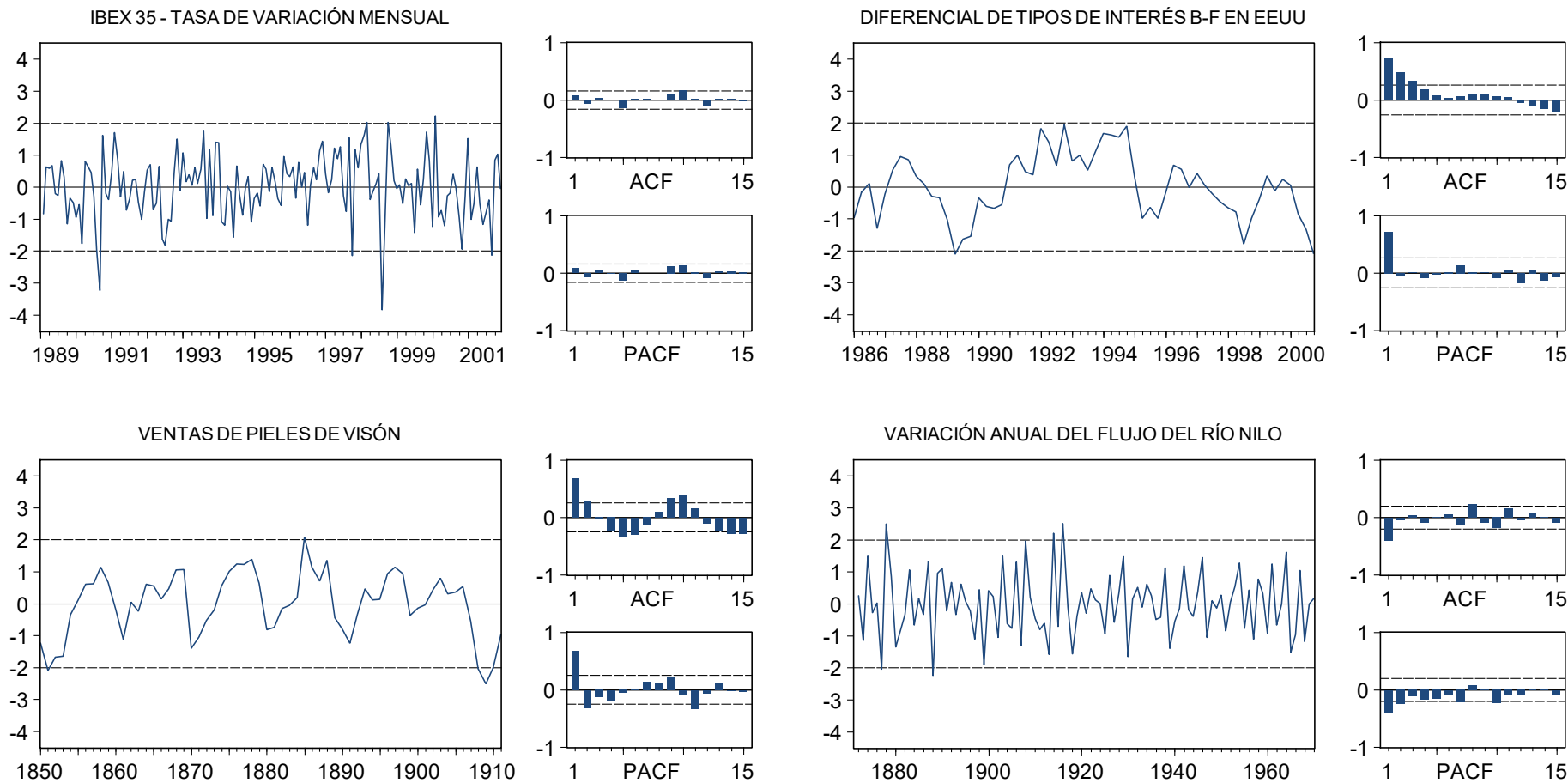


FIGURA 2.13
Autocorrelaciones Muestrales de Series con Tendencia

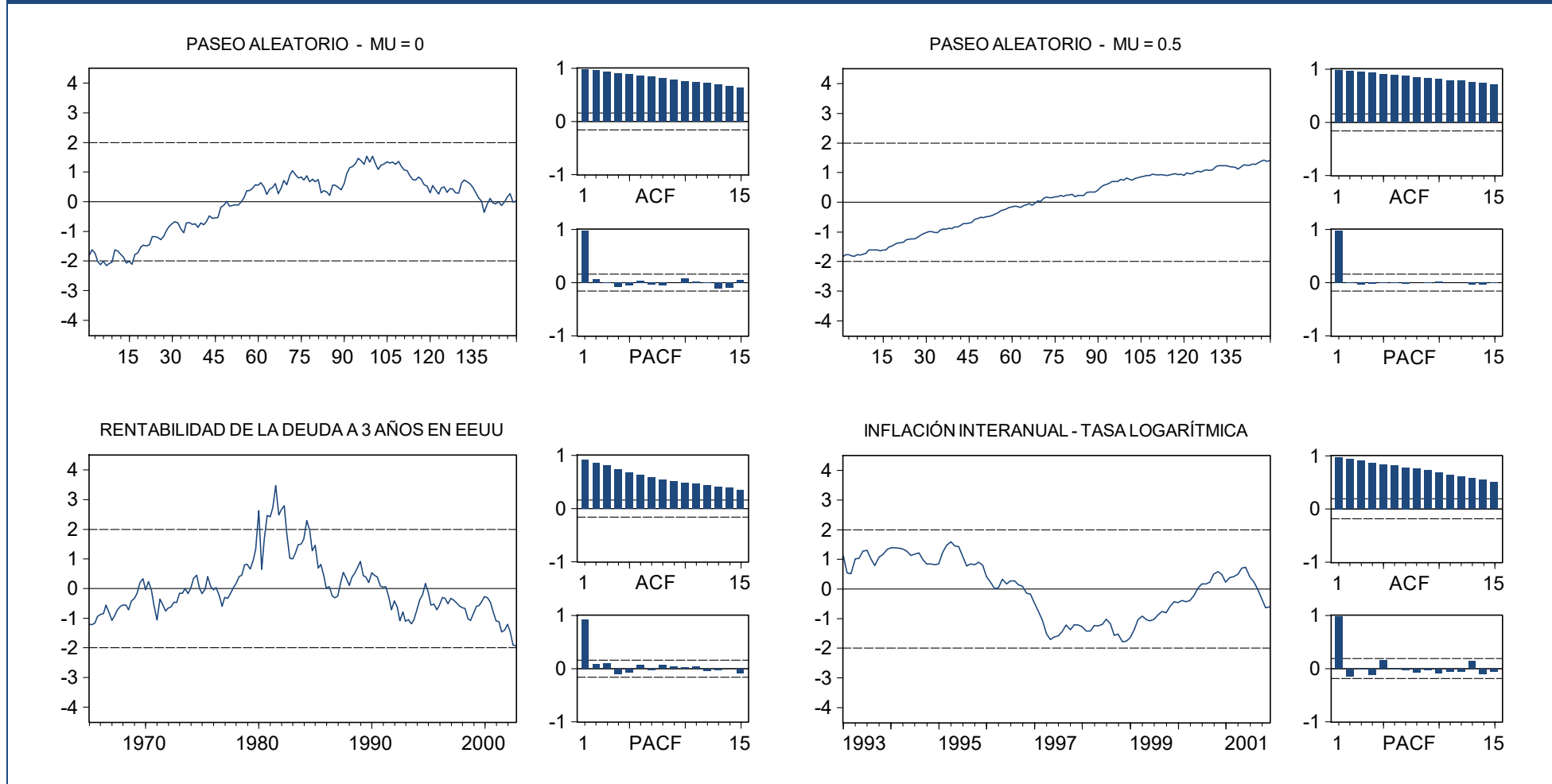
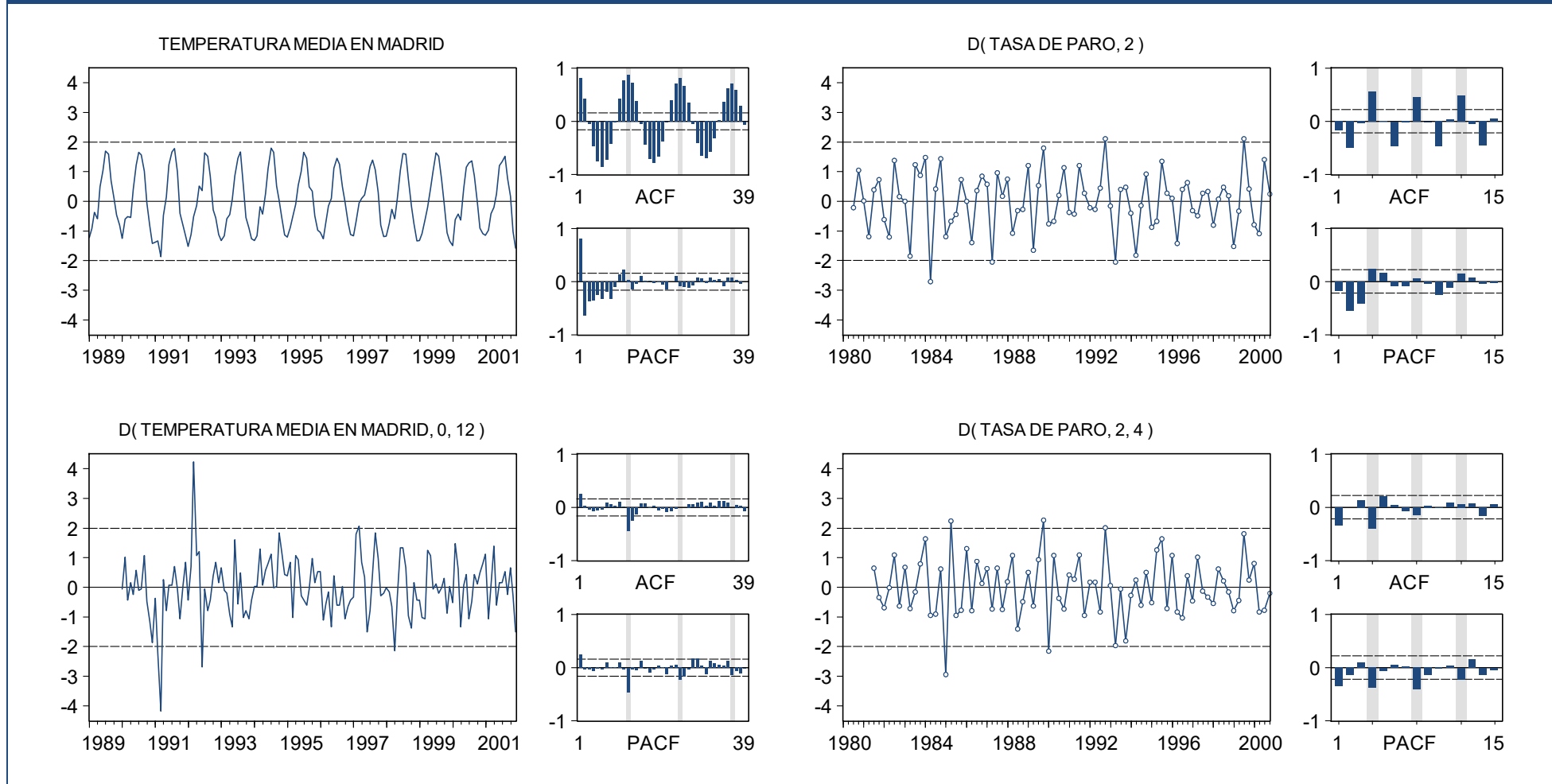


FIGURA 2.14
Autocorrelaciones Muestrales de Series Estacionales



2.6.5 Ejemplo

En este ejemplo se ilustran todas las operaciones mencionadas en esta sección para resumir numéricamente las propiedades muestrales de la serie temporal (NUM06-SERIES.WF1)

$$\mathbf{y} = [2.82, 0.09, -0.97, -1.13, -1.21, -0.81, -0.37, 2.02, 1.52, 2.44]'$$

considerada como una realización particular de un vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_{10}]'$ de variables aleatorias con media μ_Y , varianza σ_Y^2 , y autocovarianzas $\gamma_1, \gamma_2, \dots$ ($\gamma_0 = \sigma_Y^2$) constantes.

Observación: Algunos de los resultados numéricos que figuran en este ejemplo se han obtenido con EViews. Si los cálculos se hacen a mano (con una calculadora) empleando sólo uno o dos decimales, los resultados pueden variar.

Nivel y Dispersión

$$\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t = \frac{1}{10} (2.82 + \dots + 2.44) = 0.440.$$

$$s_y^2 = \frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})^2 = \frac{1}{10} [(2.82 - 0.440)^2 + \dots + (2.44 - 0.440)^2] = 2.284.$$

El valor calculado del estadístico t para el contraste al que se refiere la Observación 15 es

$$t = \bar{y} \times [s_y^2 / N]^{-1/2} = 0.440 \times [2.284 / 10]^{-1/2} = 0.921,$$

con un p -value (bilateral) calculado a partir de una distribución $t(N - 1)$ igual a

$$\alpha^* = 2 \times (1 - \Pr[t(9) \leq 0.921]) = 0.381,$$

por lo que no se puede rechazar la hipótesis nula de que la media teórica μ_Y sea cero.

Observación: Algunos programas (como EViews) emplean como medida de la dispersión de una serie la *cuasivarianza* muestral de dicha serie. Para la serie de este ejemplo, la cuasivarianza muestral es

$$cs_y^2 = \frac{1}{N-1} \sum_{t=1}^N (y_t - \bar{y})^2 = \frac{N}{N-1} s_y^2 = 2.538.$$

Utilizando cs_y^2 (cuasivarianza) como estimación de σ_Y^2 en lugar de s_y^2 ,

$$t = \bar{y} \times [cs_y^2 / N]^{-1/2} = 0.440 \times [2.538 / 10]^{-1/2} = 0.873 \Rightarrow \alpha^* = 2 \times (1 - \Pr[t(9) \leq 0.873]) = 0.405.$$

Por lo tanto, la conclusión del contraste de significación de μ_Y es la misma que al utilizar s_y^2 .

Dinámica I - Autocovarianza

Las autocovarianzas muestrales de órdenes 1, 2, ..., 6 calculadas para la serie y son:

$$\begin{aligned} c_1 &= \frac{1}{N} \sum_{t=1}^{N-1} (y_t - \bar{y})(y_{t+1} - \bar{y}) = \frac{1}{10} [(y_1 - \bar{y})(y_2 - \bar{y}) + \dots + (y_9 - \bar{y})(y_{10} - \bar{y})] \\ &= \frac{1}{10} [(2.82 - 0.440)(0.09 - 0.440) + \dots + (1.52 - 0.440)(2.44 - 0.440)] = 1.013. \end{aligned}$$

$$\begin{aligned}
 c_2 &= \frac{1}{N} \sum_{t=1}^{N-2} (y_t - \bar{y})(y_{t+2} - \bar{y}) = \frac{1}{10} [(y_1 - \bar{y})(y_3 - \bar{y}) + \dots + (y_8 - \bar{y})(y_{10} - \bar{y})] \\
 &= \frac{1}{10} [(2.82 - 0.440)(-0.97 - 0.440) + \dots + (2.02 - 0.440)(2.44 - 0.440)] = 0.313.
 \end{aligned}$$

$$\begin{aligned}
 c_3 &= \frac{1}{N} \sum_{t=1}^{N-3} (y_t - \bar{y})(y_{t+3} - \bar{y}) = \frac{1}{10} [(y_1 - \bar{y})(y_4 - \bar{y}) + \dots + (y_7 - \bar{y})(y_{10} - \bar{y})] \\
 &= \frac{1}{10} [(2.82 - 0.440)(-1.13 - 0.440) + \dots + (-0.37 - 0.440)(2.44 - 0.440)] = -0.570.
 \end{aligned}$$

$$\begin{aligned}
 c_4 &= \frac{1}{N} \sum_{t=1}^{N-4} (y_t - \bar{y})(y_{t+4} - \bar{y}) = \frac{1}{10} [(y_1 - \bar{y})(y_5 - \bar{y}) + \dots + (y_6 - \bar{y})(y_{10} - \bar{y})] \\
 &= \frac{1}{10} [(2.82 - 0.440)(-1.21 - 0.440) + \dots + (-0.81 - 0.440)(2.44 - 0.440)] = -0.911.
 \end{aligned}$$

$$\begin{aligned}
 c_5 &= \frac{1}{N} \sum_{t=1}^{N-5} (y_t - \bar{y})(y_{t+5} - \bar{y}) = \frac{1}{10} [(y_1 - \bar{y})(y_6 - \bar{y}) + \dots + (y_5 - \bar{y})(y_{10} - \bar{y})] \\
 &= \frac{1}{10} [(2.82 - 0.440)(-0.81 - 0.440) + \dots + (-1.21 - 0.440)(2.44 - 0.440)] = -0.991.
 \end{aligned}$$

$$\begin{aligned}
 c_6 &= \frac{1}{N} \sum_{t=1}^{N-6} (y_t - \bar{y})(y_{t+6} - \bar{y}) = \frac{1}{10} [(y_1 - \bar{y})(y_7 - \bar{y}) + \dots + (y_4 - \bar{y})(y_{10} - \bar{y})] \\
 &= \frac{1}{10} [(2.82 - 0.440)(-0.37 - 0.440) + \dots + (-1.13 - 0.440)(2.44 - 0.440)] = -0.714.
 \end{aligned}$$

Observación: La autocovarianza muestral de orden cero c_0 es idéntica a la varianza muestral $s_y^2 = 2.284$.

Dinámica II - Autocorrelación Simple

Teniendo en cuenta que $c_0 = s_y^2 = 2.284$, la ACF muestral de la serie y es:

$$r_1 = \frac{c_1}{c_0} = \frac{1.013}{2.284} = 0.443, \quad r_2 = \frac{c_2}{c_0} = \frac{0.313}{2.284} = 0.137, \quad r_3 = \frac{c_3}{c_0} = -\frac{0.570}{2.284} = -0.250,$$

$$r_4 = \frac{c_4}{c_0} = -\frac{0.911}{2.284} = -0.399, \quad r_5 = \frac{c_5}{c_0} = -\frac{0.991}{2.284} = -0.434, \quad r_6 = \frac{c_6}{c_0} = -\frac{0.714}{2.284} = -0.313.$$

El valor del estadístico Q_{LB} para el contraste al que se refiere la Observación 18 es

$$Q_{LB} = N(N + 2) \sum_{k=1}^K \frac{r_k^2}{N-k} = 10 \times 12 \times \sum_{k=1}^6 \frac{r_k^2}{10-k}$$

$$= 120 \times \left[\frac{0.443^2}{10-1} + \frac{0.137^2}{10-2} + \frac{(-0.250)^2}{10-3} + \frac{(-0.399)^2}{10-4} + \frac{(-0.434)^2}{10-5} + \frac{(-0.313)^2}{10-6} \right] = 14.6,$$

con un *p-value* calculado a partir de una distribución $\chi^2(K)$ igual a

$$1 - \Pr[\chi^2(6) \leq 14.6] = 0.024,$$

por lo que no puede rechazarse al 1% la hipótesis de que las seis primeras autocorrelaciones simples teóricas sean cero (aunque dicha hipótesis debe rechazarse, por ejemplo, al 5%).

Dinámica III - Autocorrelación Parcial

Los tres primeros valores de la PACF muestral pueden calcularse a partir de la relación aproximada entre la PACF y la ACF muestrales que se menciona en la Observación 19:

$$r_{11} = r_1 = 0.443,$$

$$r_{22} = \frac{r_2 - r_1^2}{1 - r_1^2} = \frac{0.137 - 0.443^2}{1 - 0.443^2} = -0.074,$$

$$r_{33} = \frac{r_1(r_1^2 + r_2^2 - 2r_2) + r_3(1 - r_1^2)}{(1 - r_2)(1 - 2r_1^2 + r_2)} = \frac{0.443(0.443^2 + 0.137^2 - 2 \times 0.137) - 0.250(1 - 0.443^2)}{(1 - 0.137)(1 - 2 \times 0.443^2 + 0.137)} = -0.353.$$

Para $k \geq 4$, las expresiones de r_{kk} análogas a éstas son muy complicadas. El procedimiento recursivo de la Observación 19 (utilizado por EViews) proporciona lo siguiente:

⇨ Para $k = 2$:

$$r_{22} = \frac{r_2 - \sum_{i=1}^1 r_{1i}r_{2-i}}{1 - \sum_{i=1}^1 r_{1i}r_i} = \frac{r_2 - r_{11}r_1}{1 - r_{11}r_1} = \frac{0.137 - 0.443 \times 0.443}{1 - 0.443 \times 0.443} = -0.074,$$

$$r_{21} = r_{11} - r_{22}r_{11} = 0.443 + 0.074 \times 0.443 = 0.476.$$

⇒ Para $k = 3$:

$$r_{33} = \frac{r_3 - \sum_{i=1}^2 r_{2i} r_{3-i}}{1 - \sum_{i=1}^2 r_{2i} r_i} = \frac{r_3 - r_{21} r_2 - r_{22} r_1}{1 - r_{21} r_1 - r_{22} r_2} = \frac{-0.250 - 0.476 \times 0.137 + 0.074 \times 0.443}{1 - 0.476 \times 0.443 + 0.074 \times 0.137} = -0.353,$$

$$r_{31} = r_{21} - r_{33} r_{22} = 0.476 - 0.353 \times 0.074 = 0.450,$$

$$r_{32} = r_{22} - r_{33} r_{21} = -0.074 + 0.353 \times 0.476 = 0.094.$$

⇒ Para $k = 4$:

$$\begin{aligned} r_{44} &= \frac{r_4 - \sum_{i=1}^3 r_{3i} r_{4-i}}{1 - \sum_{i=1}^3 r_{3i} r_i} = \frac{r_4 - r_{31} r_3 - r_{32} r_2 - r_{33} r_1}{1 - r_{31} r_1 - r_{32} r_2 - r_{33} r_3} \\ &= \frac{-0.399 + 0.450 \times 0.250 - 0.094 \times 0.137 + 0.353 \times 0.443}{1 - 0.450 \times 0.443 - 0.094 \times 0.137 - 0.353 \times 0.250} = -0.204, \end{aligned}$$

$$r_{41} = r_{31} - r_{44} r_{33} = 0.450 - 0.204 \times 0.353 = 0.378,$$

$$r_{42} = r_{32} - r_{44} r_{32} = 0.094 + 0.204 \times 0.094 = 0.113,$$

$$r_{43} = r_{33} - r_{44} r_{31} = -0.353 + 0.204 \times 0.450 = -0.261.$$

⇒ Para $k = 5$:

$$r_{55} = \frac{r_5 - \sum_{i=1}^4 r_{4i}r_{5-i}}{1 - \sum_{i=1}^4 r_{4i}r_i} = \frac{r_5 - r_{41}r_4 - r_{42}r_3 - r_{43}r_2 - r_{44}r_1}{1 - r_{41}r_1 - r_{42}r_2 - r_{43}r_3 - r_{44}r_4}$$

$$= \frac{-0.434 + 0.378 \times 0.399 + 0.113 \times 0.250 + 0.261 \times 0.137 + 0.204 \times 0.443}{1 - 0.378 \times 0.443 - 0.113 \times 0.137 - 0.261 \times 0.250 - 0.204 \times 0.399} = -0.192,$$

$$r_{51} = r_{41} - r_{55}r_{44} = 0.378 - 0.192 \times 0.204 = 0.339,$$

$$r_{52} = r_{42} - r_{55}r_{43} = 0.113 - 0.192 \times 0.261 = 0.063,$$

$$r_{53} = r_{43} - r_{55}r_{42} = -0.261 + 0.192 \times 0.113 = -0.239,$$

$$r_{54} = r_{44} - r_{55}r_{41} = -0.204 + 0.192 \times 0.378 = -0.131.$$

⇒ Para $k = 6$:

$$r_{66} = \frac{r_6 - \sum_{i=1}^5 r_{5i}r_{6-i}}{1 - \sum_{i=1}^5 r_{5i}r_i} = \frac{r_6 - r_{51}r_5 - r_{52}r_4 - r_{53}r_3 - r_{54}r_2 - r_{55}r_1}{1 - r_{51}r_1 - r_{52}r_2 - r_{53}r_3 - r_{54}r_4 - r_{55}r_5}$$

$$= \frac{-0.313 + 0.339 \times 0.434 + 0.063 \times 0.399 - 0.239 \times 0.250 + 0.131 \times 0.137 + 0.192 \times 0.443}{1 - 0.339 \times 0.443 - 0.063 \times 0.137 - 0.239 \times 0.250 - 0.131 \times 0.399 - 0.192 \times 0.434} = -0.150.$$

La ACF y la PACF muestrales que proporciona EViews para la serie y de este ejemplo son las que se muestran a continuación:

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
			1	0.443	0.443	2.6200	0.106
			2	0.137	-0.074	2.9015	0.234
			3	-0.250	-0.353	3.9695	0.265
			4	-0.399	-0.204	7.1503	0.128
			5	-0.434	-0.192	11.671	0.040
			6	-0.313	-0.150	14.605	0.024

Las bandas de significación individual en esta figura son $\pm \frac{2.0}{\sqrt{N}} = \pm \frac{2.0}{\sqrt{10}} = \pm 0.63$, ya que EViews emplea para dibujar dichas bandas un 2.0 en lugar del valor crítico bilateral al 5% de una $N(0,1)$, que es igual a 1.96 (Observaciones 18 y 21).

2.7 MODELOS ESTACIONARIOS

La hipótesis inicial en cualquier modelo estacionario consiste en suponer que una serie temporal estacionaria $\mathbf{y} = [y_1, y_2, \dots, y_N]'$ es una realización particular de un vector de variables aleatorias $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ con **media teórica** $\mu_Y = E[Y_t]$, **varianza teórica** $\sigma_Y^2 = \text{Var}[Y_t]$, y **autocovarianzas teóricas** $\gamma_k = \text{Cov}[Y_t, Y_{t+k}] = \text{Cov}[Y_{t-k}, Y_t]$ ($k \geq 1$) **constantes**, de manera que (i) μ_Y y σ_Y^2 ($= \gamma_0$) no dependen de t , y (ii) cada γ_k ($k \geq 1$) puede depender del retardo k pero tampoco depende de t .

Cada tipo de modelo tiene asociada una expresión matemática concreta de la que se derivan una media, una varianza, unas autocovarianzas, y, en especial, unas **autocorrelaciones simples** (ACF) y **parciales** (PACF) teóricas determinadas.

Observación 22: Una **ACF teórica** es la secuencia ρ_1, ρ_2, \dots que puede derivarse a partir de la expresión matemática concreta asociada con un modelo determinado, donde $\rho_k = \frac{\gamma_k}{\gamma_0}$ (y, en su caso, $\rho_0 = 1$) mide el grado de correlación lineal **total** o **bruta** entre dos componentes cualesquiera del vector $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$ separados entre sí por un retardo $k \geq 1$ dado (como Y_t e Y_{t-k}). Una **PACF teórica** es la secuencia $\rho_{11}, \rho_{22}, \dots$ que puede derivarse a partir de la expresión matemática concreta asociada con un modelo determinado, donde ρ_{kk} mide el grado de correlación lineal **directa** o **neto** entre dos componentes cualesquiera de \mathbf{Y} separados entre sí por un retardo $k \geq 1$ dado (como Y_t e Y_{t-k}), que **no** es debida a la presencia de los componentes de \mathbf{Y} que se encuentran entre ambos ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$). Análogamente a la PACF muestral de una serie temporal (Observación 19), $\rho_{11} = \rho_1$ y ρ_{kk} ($k \geq 2$) es el parámetro asociado con Y_{t-k} en $Y_t = \rho_{k0} + \rho_{k1}Y_{t-1} + \rho_{k2}Y_{t-2} + \dots + \rho_{kk}Y_{t-k} + U_{tk}$, con U_{tk} independiente de Y_{t-i} para todo $i \geq 1$ y $E[U_{tk}] = 0$. La regresión anterior puede escribirse como

$$\tilde{Y}_t = \rho_{k1}\tilde{Y}_{t-1} + \rho_{k2}\tilde{Y}_{t-2} + \dots + \rho_{kk}\tilde{Y}_{t-k} + U_{tk}, \text{ con } \tilde{Y}_{t-i} = Y_{t-i} - E[Y_t] \text{ (} i = 0, 1, \dots, k \text{),}$$

de manera que multiplicando por \tilde{Y}_{t-i} ($i \geq 1$) y tomando valores esperados:

$$E[\tilde{Y}_{t-i}\tilde{Y}_t] = \rho_{k1}E[\tilde{Y}_{t-i}\tilde{Y}_{t-1}] + \rho_{k2}E[\tilde{Y}_{t-i}\tilde{Y}_{t-2}] + \dots + \rho_{kk}E[\tilde{Y}_{t-i}\tilde{Y}_{t-k}] + E[\tilde{Y}_{t-i}U_{tk}],$$

es decir,

$$\gamma_i = \rho_{k1}\gamma_{i-1} + \rho_{k2}\gamma_{i-2} + \dots + \rho_{kk}\gamma_{i-k}.$$

Dividiendo la expresión anterior por γ_0 :

$$\rho_i = \rho_{k1}\rho_{i-1} + \rho_{k2}\rho_{i-2} + \dots + \rho_{kk}\rho_{i-k} \quad (i \geq 1).$$

Escribiendo esta expresión para $i = 1, \dots, k$, con ($k \geq 2$), resulta finalmente que ρ_{kk} ($k \geq 2$) es el último elemento en el vector del lado izquierdo de la siguiente expresión:

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \rho_{k1} \\ \rho_{k2} \\ \vdots \\ \rho_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix},$$

lo que implica, por ejemplo, con $k = 2, 3$, que

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \quad \rho_{33} = \frac{\rho_1(\rho_1^2 + \rho_2^2 - 2\rho_2) + \rho_3(1 - \rho_1^2)}{(1 - \rho_2)(1 - 2\rho_1^2 + \rho_2)}.$$

De esta manera, cada coeficiente de autocorrelación parcial ρ_{kk} en un modelo estacionario puede expresarse como una función determinada de los coeficientes de autocorrelación simple ρ_1, \dots, ρ_k ($k = 1, 2, 3, \dots$) correspondientes.

En última instancia, la **identificación** de un modelo univariante para una serie estacionaria consiste en la **elección** (tentativa) de un **modelo** cuyas **propiedades teóricas** resuman adecuadamente las **propiedades muestrales** de la serie considerada.

TABLA 2.2 Series Estacionarias y Modelos Estacionarios	
Series Temporales Estacionarias	Modelos ARMA Estacionarios
Observaciones (Datos): y_1, y_2, \dots, y_N	Variables Aleatorias: Y_1, Y_2, \dots, Y_N
Propiedades Muestrales	Propiedades Teóricas
Nivel (Localización)	
$\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t$	$\mu_Y = E[Y_t]$
Dispersión (Variabilidad)	
$s_y^2 = \frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})^2 = c_0$	$\sigma_Y^2 = \text{Var}[Y_t] = E[(Y_t - \mu_Y)^2] = \gamma_0$
Dinámica I - Autocovarianzas	
$c_k = \frac{1}{N} \sum_{t=k+1}^N (y_{t-k} - \bar{y})(y_t - \bar{y})$ $= \frac{1}{N} \sum_{t=1}^{N-k} (y_t - \bar{y})(y_{t+k} - \bar{y})$	$\gamma_k = \text{Cov}[Y_{t-k}, Y_t] = \text{Cov}[Y_t, Y_{t+k}]$ $= E[(Y_{t-k} - \mu_Y)(Y_t - \mu_Y)]$ $= E[(Y_t - \mu_Y)(Y_{t+k} - \mu_Y)]$
Dinámica II - Autocorrelaciones Simples (ACF)	
$r_k = \frac{c_k}{c_0} \quad (\Rightarrow r_0 = 1)$	$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (\Rightarrow \rho_0 = 1)$
Dinámica III - Autocorrelaciones Parciales (PACF)	
$r_{11} = r_1, \quad r_{22} = \frac{r_2 - r_1^2}{1 - r_1^2},$ $r_{33} = \frac{r_1(r_1^2 + r_2^2 - 2r_2) + r_3(1 - r_1^2)}{(1 - r_2)(1 - 2r_1^2 + r_2)}, \dots$	$\rho_{11} = \rho_1, \quad \rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2},$ $\rho_{33} = \frac{\rho_1(\rho_1^2 + \rho_2^2 - 2\rho_2) + \rho_3(1 - \rho_1^2)}{(1 - \rho_2)(1 - 2\rho_1^2 + \rho_2)}, \dots$

2.7.1 Modelo AR(1) - Autorregresivo de Orden 1

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t, \text{ o bien} \quad [9]$$

$$(1 - \phi_1 B)Y_t = \mu + A_t, \quad [10]$$

donde μ y ϕ_1 son parámetros, $|\phi_1| < 1$ (**estacionariedad**), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Observación 23: La condición de estacionariedad de un modelo AR(1) puede enunciarse así: "La **raíz** de la **ecuación característica** $1 - \phi_1 x = 0$ ($x^* = \phi_1^{-1}$) es en valor absoluto mayor que 1 (o bien tiene **módulo** superior a la unidad, o bien se encuentra **fuera del círculo unitario**)".

Un modelo AR(1) estacionario tiene las propiedades teóricas siguientes:

Media
$$\mu_Y = \frac{\mu}{1 - \phi_1}. \quad [11]$$

Varianza
$$\sigma_Y^2 (= \gamma_0) = \frac{\sigma_A^2}{1 - \phi_1^2}. \quad [12]$$

ACF
$$\rho_k = \phi_1 \rho_{k-1} \quad (k \geq 1). \quad [13]$$

Observación 24: Esta ecuación para ρ_k puede resolverse recursivamente a partir de la condición inicial $\rho_0 = 1$ (Observación 22), lo que proporciona $\rho_k = \phi_1^k$ para todo $k \geq 0$ (una **exponencial amortiguada**). Por otro lado, dado

que $\rho_1 = \phi_1$, la varianza [12] de un AR(1) puede escribirse como

$$\sigma_Y^2 = \frac{\sigma_A^2}{1 - \phi_1 \rho_1}.$$

PACF

$$\rho_{kk} = \begin{cases} \rho_1 = \phi_1 & \text{si } k = 1, \\ 0 & \text{para todo } k > 1. \end{cases} \quad [14]$$

Observación 25: La demostración de las propiedades [11]-[14] se basa en las expresiones siguientes:

$$\begin{aligned} Y_t &= \mu + \phi_1 \overbrace{(\mu + \phi_1 Y_{t-2} + A_{t-1})}^{Y_{t-1}} + A_t \\ &= (1 + \phi_1)\mu + (A_t + \phi_1 A_{t-1}) + [\phi_1^2 Y_{t-2}] \\ &= (1 + \phi_1)\mu + A_t + \phi_1 A_{t-1} + \phi_1^2 \overbrace{(\mu + \phi_1 Y_{t-3} + A_{t-2})}^{Y_{t-2}} \\ &= (1 + \phi_1 + \phi_1^2)\mu + (A_t + \phi_1 A_{t-1} + \phi_1^2 A_{t-2}) + [\phi_1^3 Y_{t-3}] \\ &= \vdots \\ &= \left(\sum_{i=0}^{\infty} \phi_1^i \right) \mu + \sum_{i=0}^{\infty} \phi_1^i A_{t-i} \\ &= \frac{\mu}{1 - \phi_1} + \sum_{i=0}^{\infty} \phi_1^i A_{t-i} = \frac{\mu}{1 - \phi_1} + A_t + \phi_1 A_{t-1} + \phi_1^2 A_{t-2} + \phi_1^3 A_{t-3} + \dots \end{aligned} \quad (A)$$

$$\mu_Y = E[Y_t] = E\left[\frac{\mu}{1-\phi_1} + \sum_{i=0}^{\infty} \phi_1^i A_{t-i}\right] = \frac{\mu}{1-\phi_1}. \quad (\text{B})$$

$$\begin{aligned} \sigma_Y^2 (= \gamma_0) &= \text{Var}[Y_t] = \text{Var}\left[\frac{\mu}{1-\phi_1} + \sum_{i=0}^{\infty} \phi_1^i A_{t-i}\right] = \sum_{i=0}^{\infty} \text{Var}[\phi_1^i A_{t-i}] = \\ &= \sum_{i=0}^{\infty} \phi_1^{2i} \sigma_A^2 = \left(\sum_{i=0}^{\infty} \phi_1^{2i}\right) \sigma_A^2 = \frac{\sigma_A^2}{1-\phi_1^2}. \end{aligned} \quad (\text{C})$$

$$\begin{aligned} Y_t &= \mu + \phi_1 Y_{t-1} + A_t \Leftrightarrow Y_t - \frac{\mu}{1-\phi_1} = \mu - \frac{\mu}{1-\phi_1} + \phi_1 Y_{t-1} + A_t = \\ &= \frac{\mu - \phi_1 \mu - \mu}{1-\phi_1} + \phi_1 Y_{t-1} + A_t = \phi_1 Y_{t-1} - \phi_1 \frac{\mu}{1-\phi_1} + A_t = \phi_1 \left(Y_{t-1} - \frac{\mu}{1-\phi_1}\right) + A_t \Leftrightarrow \\ &\Leftrightarrow \tilde{Y}_t = \phi_1 \tilde{Y}_{t-1} + A_t \Leftrightarrow (1 - \phi_1 B) \tilde{Y}_t = A_t, \text{ con } \tilde{Y}_t = Y_t - \mu_Y, \mu_Y = \frac{\mu}{1-\phi_1}. \end{aligned} \quad (\text{D})$$

$$\begin{aligned} \gamma_k &= \text{Cov}[Y_{t-k}, Y_t] = E[\tilde{Y}_{t-k} \tilde{Y}_t] = E[\tilde{Y}_{t-k} (\phi_1 \tilde{Y}_{t-1} + A_t)] = \\ &= \phi_1 \underbrace{E[\tilde{Y}_{t-k} \tilde{Y}_{t-1}]}_{\text{Cov}[Y_{t-k}, Y_{t-1}]} + \underbrace{E[\tilde{Y}_{t-k} A_t]}_{\text{Cov}[Y_{t-k}, A_t]} = \phi_1 \gamma_{k-1} + \underbrace{\text{Cov}[Y_{t-k}, A_t]}_{=0 \text{ para } k \geq 1} = \phi_1 \gamma_{k-1} \quad (k \geq 1). \end{aligned} \quad (\text{E})$$

$$\begin{aligned} \rho_k &= \frac{\gamma_k}{\gamma_0} = \frac{\phi_1 \gamma_{k-1}}{\gamma_0} = \phi_1 \rho_{k-1} \quad (k \geq 1) \Rightarrow \rho_1 = \phi_1 \rho_0 = \phi_1, \\ \rho_2 &= \phi_1 \rho_1 = \phi_1^2, \dots \Rightarrow \rho_k = \phi_1^k \quad (k \geq 0). \end{aligned} \quad (\text{F})$$

$$\sigma_Y^2 (= \gamma_0) = \frac{\sigma_A^2}{1-\phi_1^2} = \frac{\sigma_A^2}{1-\phi_1 \rho_1}. \quad (\text{G})$$

La PACF [14] puede calcularse a partir de la ACF [13] de acuerdo con la Observación 22.

2.7.2 Modelo AR(2) - Autorregresivo de Orden 2

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + A_t, \text{ o bien} \quad [15]$$

$$(1 - \phi_1 B - \phi_2 B^2) Y_t = \mu + A_t, \quad [16]$$

donde μ , ϕ_1 y ϕ_2 son parámetros, las raíces de la ecuación $1 - \phi_1 x - \phi_2 x^2 = 0$ están fuera del círculo unitario (**estacionariedad**), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Observación 26: La condición de estacionariedad de un modelo AR(2) implica que $\phi_2 + \phi_1 < 1$, $\phi_2 - \phi_1 < 1$, y $|\phi_2| < 1$ (aunque, en general, la implicación en sentido inverso no es cierta). Ver Figura 2.15.

Un modelo AR(2) estacionario tiene las propiedades teóricas siguientes:

Media
$$\mu_Y = \frac{\mu}{1 - \phi_1 - \phi_2}. \quad [17]$$

Varianza
$$\sigma_Y^2 = \frac{\sigma_A^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2}. \quad [18]$$

ACF
$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad (k \geq 1). \quad [19]$$

Observación 27: Esta ecuación para ρ_k puede resolverse recursivamente a partir de las condiciones iniciales $\rho_0 = 1$ (Observación 22) y $\rho_1 = \phi_1 / (1 - \phi_2)$ (que se obtiene de [19] con $k = 1$; por otro lado, las dos ecuaciones que resultan

de [19] con $k = 1$ y $k = 2$ se denominan **ecuaciones de Yule-Walker**). En particular, si las dos **raíces** de la ecuación característica son **complejas** (Figura 2.15), la solución de [19] se puede expresar como (IEV41: Apéndice D)

$$\rho_k = M \times D^k \sin(2\pi f k + F) \text{ para todo } k \geq 0,$$

que es una **oscilación armónica amortiguada** de **amplitud** $M \times D^k$, **periodo** $1/f$, **frecuencia** f , **fase inicial** $-F/(2\pi f)$, y **factor de amortiguamiento** D . Estas características de ρ_k son funciones de los parámetros ϕ_1 y ϕ_2 de acuerdo con que $M = 1/\sin(F)$ [$\Leftrightarrow \rho_0 = 1$], $D = \sqrt{-\phi_2}$, $\cos(2\pi f) = \phi_1/(2\sqrt{-\phi_2})$, y $\tan(F) = [(1 - \phi_2)/(1 + \phi_2)] \tan(2\pi f)$.

Observación 28: Con $\rho_1 = \phi_1 / (1 - \phi_2)$ y $\rho_2 = \phi_2 + \phi_1^2 / (1 - \phi_2)$ (ver [19]), la varianza [18] de un AR(2) queda

$$\sigma_Y^2 = \left[\frac{1 - \phi_2}{1 + \phi_2} \right] \left[\frac{\sigma_A^2}{(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \right].$$

PACF

$$\rho_{kk} = \begin{cases} \rho_1 = \phi_1 / (1 - \phi_2) & \text{si } k = 1, \\ \phi_2 & \text{si } k = 2, \\ 0 & \text{para todo } k > 2. \end{cases} \quad [20]$$

Observación 29: Las propiedades [17]-[20] son extensiones de las propiedades [11]-[14] del AR(1). Por ejemplo, la PACF [20] puede calcularse a partir de la ACF [19] [con $\rho_1 = \phi_1 / (1 - \phi_2)$, $\rho_2 = \phi_2 + \phi_1^2 / (1 - \phi_2)$, ...] de acuerdo con la Observación 22. En el mismo sentido, la representación inicial [15]-[16] del modelo AR(2) puede escribirse como

$$\tilde{Y}_t = \phi_1 \tilde{Y}_{t-1} + \phi_2 \tilde{Y}_{t-2} + A_t \Leftrightarrow (1 - \phi_1 B - \phi_2 B^2) \tilde{Y}_t = A_t, \text{ con } \tilde{Y}_t = Y_t - \mu_Y, \mu_Y = \frac{\mu}{1 - \phi_1 - \phi_2}.$$

FIGURA 2.15

Estacionariedad y Autocorrelaciones en Modelos AR(2)

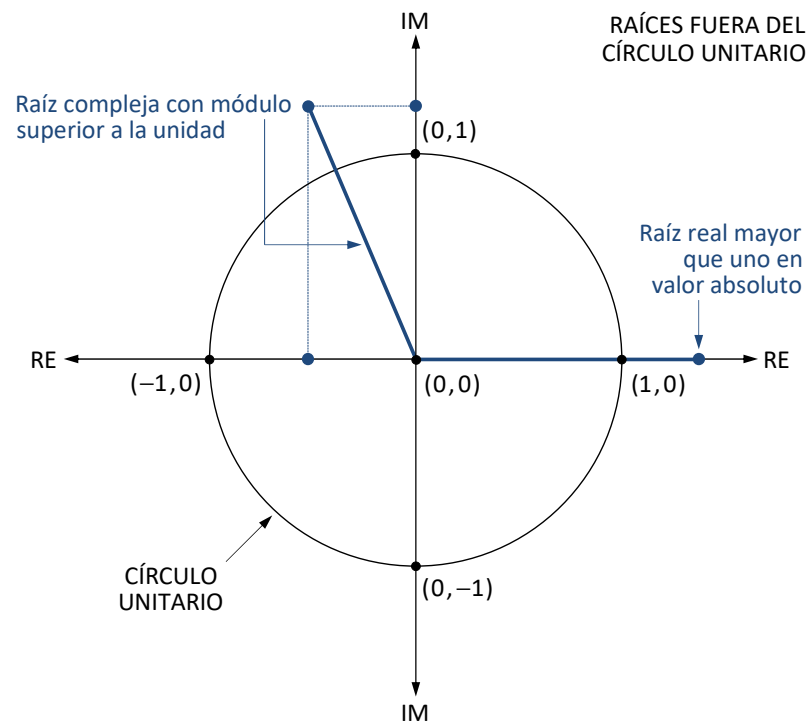
$$1 - \phi_1 x - \phi_2 x^2 = 0 \rightarrow \underbrace{(-\phi_2)}_a x^2 + \underbrace{(-\phi_1)}_b x + \underbrace{1}_c = 0,$$

$$x_1^*, x_2^* = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2} = \frac{\phi_1}{-2\phi_2} \pm \frac{\sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}.$$

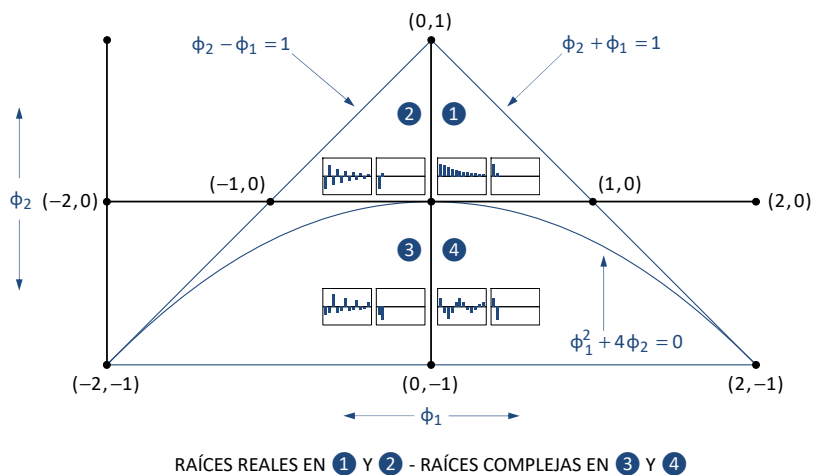
$\phi_1^2 + 4\phi_2 > 0 \rightarrow x_1^* \neq x_2^*$ (reales distintas),

$\phi_1^2 + 4\phi_2 = 0 \rightarrow x_1^* = x_2^* = \frac{\phi_1}{-2\phi_2}$ (reales iguales),

$\phi_1^2 + 4\phi_2 < 0 \rightarrow x_1^*, x_2^* = \frac{\phi_1}{-2\phi_2} \pm \frac{\sqrt{-(\phi_1^2 + 4\phi_2)}}{-2\phi_2} \times i$ (complejas).



AUTOCORRELACIONES EN MODELOS AR(2) ESTACIONARIOS



$$\text{MÓDULO} = \left[\text{PARTE REAL}^2 + \text{PARTE IMAGINARIA}^2 \right]^{\frac{1}{2}}$$

2.7.3 Modelo AR(p) - Autorregresivo de Orden p

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + A_t, \text{ o bien} \quad [21]$$

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = \mu + A_t, \quad [22]$$

donde $\mu, \phi_1, \phi_2, \dots, \phi_p$ son parámetros, las raíces de $1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0$ están fuera del círculo unitario (**estacionariedad**), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Media
$$\mu_Y = \mu \left(1 - \sum_{i=1}^p \phi_i \right)^{-1}. \quad [23]$$

Varianza
$$\sigma_Y^2 = \sigma_A^2 \left(1 - \sum_{i=1}^p \phi_i \rho_i \right)^{-1}. \quad [24]$$

ACF
$$\rho_k = \sum_{i=1}^p \phi_i \rho_{k-i} \quad (k \geq 1). \quad [25]$$

PACF
$$\rho_{kk} = \begin{cases} \rho_1 & \text{si } k = 1, \\ \vdots & \\ \phi_p & \text{si } k = p, \\ 0 & \text{para todo } k > p. \end{cases} \quad [26]$$

Observación 30: La solución general de [25] consiste en una mezcla de exponenciales amortiguadas (asociadas con las raíces reales de la ecuación característica) y oscilaciones armónicas amortiguadas (asociadas con las raíces complejas).

2.7.4 Modelo MA(1) - Media Móvil de Orden 1

$$Y_t = \mu + A_t - \theta_1 A_{t-1}, \text{ o bien} \quad [27]$$

$$Y_t = \mu + (1 - \theta_1 B)A_t, \quad [28]$$

donde μ y θ_1 son parámetros, y $(A_t) \sim \text{IID}(0, \sigma_A^2)$. Cualquier modelo MA(1) tiene las propiedades teóricas siguientes:

Media $\mu_Y = \mu.$ [29]

Varianza $\sigma_Y^2 = (1 + \theta_1^2)\sigma_A^2.$ [30]

ACF $\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & \text{si } k = 1, \\ 0 & \text{para todo } k > 1. \end{cases}$ [31]

PACF $\rho_{kk} = - \left[\frac{1}{\sum_{i=0}^k \theta_1^{2i}} \right] \theta_1^k = - \left[\frac{1 - \theta_1^2}{1 - \theta_1^{2(k+1)}} \right] \theta_1^k$ para todo $k \geq 1.$ [32]

Observación 31: Nótese la **simetría** o **dualidad** entre la formulación y las propiedades de un modelo MA(1) y las de un

modelo AR(1). En este mismo sentido, dado que [29]-[32] son cantidades constantes, un modelo MA(1) es estacionario independientemente de lo que valgan sus parámetros. La demostración de las propiedades [29]-[32] se basa en las expresiones siguientes:

$$\mu_Y = E[Y_t] = E[\mu + A_t - \theta_1 A_{t-1}] = \mu. \quad (\text{A})$$

$$\begin{aligned} \sigma_Y^2 (= \gamma_0) &= \text{Var}[Y_t] = \text{Var}[\mu + A_t - \theta_1 A_{t-1}] = \\ &\text{Var}[A_t] + \theta_1^2 \text{Var}[A_{t-1}] - 2\theta_1 \text{Cov}[A_t, A_{t-1}] = (1 + \theta_1^2) \sigma_A^2. \end{aligned} \quad (\text{B})$$

$$\begin{aligned} Y_t &= \mu + A_t - \theta_1 A_{t-1} \Leftrightarrow Y_t - \mu = A_t - \theta_1 A_{t-1} \Leftrightarrow \\ &\Leftrightarrow \tilde{Y}_t = A_t - \theta_1 A_{t-1} \Leftrightarrow \tilde{Y}_t = (1 - \theta_1 B)A_t, \text{ con } \tilde{Y}_t = Y_t - \mu_Y, \mu_Y = \mu. \end{aligned} \quad (\text{C})$$

$$\begin{aligned} \gamma_1 &= \text{Cov}[Y_{t-1}, Y_t] = E[\tilde{Y}_{t-1} \tilde{Y}_t] = E[(A_{t-1} - \theta_1 A_{t-2})(A_t - \theta_1 A_{t-1})] = \\ &= \underbrace{E[A_{t-1} A_t]}_{\text{Cov}[A_{t-1}, A_t]=0} - \theta_1 \underbrace{E[A_{t-1}^2]}_{\text{Var}[A_t]} - \theta_1 \underbrace{E[A_{t-2} A_t]}_{\text{Cov}[A_{t-2}, A_t]=0} + \theta_1^2 \underbrace{E[A_{t-2} A_{t-1}]}_{\text{Cov}[A_{t-2}, A_{t-1}]=0} = -\theta_1 \sigma_A^2. \end{aligned} \quad (\text{D})$$

$$\gamma_k = \text{Cov}[Y_{t-k}, Y_t] = E[\tilde{Y}_{t-k} \tilde{Y}_t] = E[(A_{t-k} - \theta_1 A_{t-k-1})(A_t - \theta_1 A_{t-1})] = 0 \quad (k \geq 2). \quad (\text{E})$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1 \sigma_A^2}{(1 + \theta_1^2) \sigma_A^2} = -\frac{\theta_1}{1 + \theta_1^2}, \quad \rho_k = \frac{\gamma_k}{\gamma_0} = 0 \quad (k \geq 2). \quad (\text{F})$$

Por su parte, la PACF de un MA(1) se puede calcular a partir de la Observación 22, con la ACF dada en [31]. El hecho de que un modelo MA(1) sea estacionario independientemente de lo que valgan sus parámetros, sugiere la conclusión general de que la estacionariedad es un asunto que requiere atención sólo cuando un modelo contiene

términos de tipo AR. No obstante, en un modelo MA(1) debe ocurrir que $|\theta_1| < 1$ para que el modelo tenga ciertas propiedades adicionales importantes (análogas a algunas de las que tiene automáticamente un modelo de tipo AR).

Cuando $|\theta_1| < 1$ (**invertibilidad**) en un modelo MA(1), puede comprobarse lo siguiente:

⇒ El modelo admite una representación inversa (AR) bien definida. [33]

⇒ Dado un valor para ρ_1 , existe un único valor de θ_1 tal que $\rho_1 = -\theta_1/(1 + \theta_1^2)$. [34]

Observación 32: En relación con [33], cuando $|\theta_1| < 1$ en un MA(1) [ver (C) en la Observación 31],

$$\begin{aligned}
 A_t &= \tilde{Y}_t + \theta_1 A_{t-1} \\
 &= \tilde{Y}_t + \theta_1 \overbrace{(\tilde{Y}_{t-1} + \theta_1 A_{t-2})}^{A_{t-1}} = \tilde{Y}_t + \theta_1 \tilde{Y}_{t-1} + [\theta_1^2 A_{t-2}] \\
 &= \tilde{Y}_t + \theta_1 \tilde{Y}_{t-1} + \theta_1^2 \overbrace{(\tilde{Y}_{t-2} + \theta_1 A_{t-3})}^{A_{t-2}} = \tilde{Y}_t + \theta_1 \tilde{Y}_{t-1} + \theta_1^2 \tilde{Y}_{t-2} + [\theta_1^3 A_{t-3}] \\
 &= \quad \vdots \\
 &= \tilde{Y}_t + \theta_1 \tilde{Y}_{t-1} + \theta_1^2 \tilde{Y}_{t-2} + \dots = \sum_{i=0}^{\infty} \theta_1^i \tilde{Y}_{t-i}, \text{ por lo que} \\
 \tilde{Y}_t &= -\theta_1 \tilde{Y}_{t-1} - \theta_1^2 \tilde{Y}_{t-2} - \dots + A_t.
 \end{aligned}$$

Observación 33: En relación con [34], por ejemplo, $\theta_1 = 0.5$ y $\theta_1 = 2$ dan el mismo $\rho_1 = -0.4$, pero sólo $\theta_1 = 0.5$ es invertible. La condición de invertibilidad $|\theta_1| < 1$ de un modelo MA(1) puede enunciarse así: "La raíz de la **ecuación característica** $1 - \theta_1 x = 0$ ($x^* = \theta_1^{-1}$) tiene módulo superior a la unidad (se encuentra fuera del círculo unitario)".

Consideraciones semejantes a las relacionadas con las propiedades [33]-[34] son también aplicables a modelos MA de cualquier orden. Dado que un modelo AR estacionario satisface automáticamente ese tipo de propiedades, la invertibilidad es un asunto que requiere atención sólo cuando un modelo contiene términos MA.

2.7.5 Modelo MA(2) - Media Móvil de Orden 2

$$Y_t = \mu + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2}, \text{ o bien} \quad [35]$$

$$Y_t = \mu + (1 - \theta_1 B - \theta_2 B^2) A_t, \quad [36]$$

donde μ , θ_1 y θ_2 son parámetros, y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Observación 34: Un modelo MA(2) es invertible cuando las dos raíces de su ecuación característica $1 - \theta_1 x - \theta_2 x^2 = 0$ están fuera del círculo unitario, para lo cual es necesario que $\theta_2 + \theta_1 < 1$, $\theta_2 - \theta_1 < 1$, y $|\theta_2| < 1$.

Cualquier modelo MA(2) tiene las propiedades teóricas siguientes:

Media $\mu_Y = \mu.$ [37]

Varianza $\sigma_Y^2 = (1 + \theta_1^2 + \theta_2^2) \sigma_A^2.$ [38]

ACF $\rho_k = \begin{cases} -[\theta_1(1 - \theta_2)] / (1 + \theta_1^2 + \theta_2^2) & \text{si } k = 1, \\ -\theta_2 / (1 + \theta_1^2 + \theta_2^2) & \text{si } k = 2, \\ 0 & \text{para todo } k > 2. \end{cases}$ [39]

Observación 35: La expresión analítica de la PACF en un modelo MA(2) depende de la naturaleza de las dos raíces de su ecuación característica. Dicha expresión es muy complicada. No obstante, la PACF de un MA(2) sigue las mismas pautas que la ACF (**dualidad**) de un modelo AR(2) [lo mismo que la ACF de un MA(2) y la PACF de un AR(2)]. Las propiedades [37]-[39] son extensiones de las propiedades [29]-[31] del MA(1). En el mismo sentido, la representación inicial [35]-[36] del MA(2) puede escribirse como

$$\tilde{Y}_t = A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} \Leftrightarrow \tilde{Y}_t = (1 - \theta_1 B - \theta_2 B^2) A_t, \text{ con } \tilde{Y}_t = Y_t - \mu_Y, \mu_Y = \mu.$$

2.7.6 Modelo MA(q) - Media Móvil de Orden q

$$Y_t = \mu + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \dots - \theta_q A_{t-q}, \text{ o bien} \quad [40]$$

$$Y_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) A_t, \quad [41]$$

donde $\mu, \theta_1, \theta_2, \dots, \theta_q$ son parámetros, y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Observación 36: Un modelo MA(q) es invertible cuando las raíces de la ecuación $1 - \theta_1 x - \theta_2 x^2 - \dots - \theta_q x^q = 0$ están fuera del círculo unitario.

Cualquier modelo MA(q) tiene las propiedades teóricas siguientes:

Media
$$\mu_Y = \mu. \quad [42]$$

Varianza
$$\sigma_Y^2 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_A^2. \quad [43]$$

Observación 37: Las expresiones analíticas de la ACF y la PACF en un modelo MA general dependen de la naturaleza de las raíces de su ecuación característica. En particular, la ACF y la PACF de un modelo MA general siguen las mismas pautas que la PACF y la ACF, respectivamente (**dualidad**), de un modelo AR general (ver [25]-[26]).

2.7.7 Modelo ARMA(1,1) - Autorregresivo-Media Móvil de Orden 1,1

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t - \theta_1 A_{t-1}, \text{ o bien} \quad [44]$$

$$(1 - \phi_1 B)Y_t = \mu + (1 - \theta_1 B)A_t, \quad [45]$$

donde μ , ϕ_1 y θ_1 son parámetros (con $\phi_1 \neq \theta_1$), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Un modelo ARMA(1,1) estacionario ($|\phi_1| < 1$) tiene las propiedades teóricas siguientes:

Media
$$\mu_Y = \frac{\mu}{1 - \phi_1}. \quad [46]$$

Varianza
$$\sigma_Y^2 = \frac{1 - 2\phi_1\theta_1 + \theta_1^2}{1 - \phi_1^2} \sigma_A^2 = \left[1 + \frac{(\phi_1 - \theta_1)^2}{1 - \phi_1^2} \right] \sigma_A^2. \quad [47]$$

ACF
$$\rho_k = \begin{cases} [(\phi_1 - \theta_1)(1 - \phi_1\theta_1)] / (1 - 2\phi_1\theta_1 + \theta_1^2) & \text{si } k = 1, \\ \rho_1 \phi_1^{k-1} & \text{para todo } k > 1. \end{cases} \quad [48]$$

Observación 38: La PACF de un modelo ARMA(1,1) consiste en un valor inicial $\rho_{11} = \rho_1$ seguido de una secuencia de valores que siguen exactamente la pauta de la PACF de un MA(1). Esto mismo ocurre con la ACF [48], que consiste en un valor inicial ρ_1 seguido de una secuencia de valores que siguen exactamente la pauta de la ACF en un AR(1). En general, la ACF (PACF) de un modelo ARMA sigue las pautas de la ACF (PACF) de su parte AR (MA). Por otro lado, la representación inicial [45]-[45] del ARMA(1,1) puede escribirse como

$$\tilde{Y}_t = \phi_1 \tilde{Y}_{t-1} + A_t - \theta_1 A_{t-1} \Leftrightarrow (1 - \phi_1 B) \tilde{Y}_t = (1 - \theta_1 B) A_t, \text{ con } \tilde{Y}_t = Y_t - \mu_Y, \mu_Y = \frac{\mu}{1 - \phi_1}.$$

2.7.8 Modelo ARMA(p,q) - Autorregresivo-Media Móvil de Orden p,q

$$\begin{aligned} Y_t &= \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} \\ &+ A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \dots - \theta_q A_{t-q}, \text{ o bien} \end{aligned} \quad [49]$$

$$\phi(B) Y_t = \mu + \theta(B) A_t, \quad [50]$$

donde

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad [51]$$

es el **operador** o **polinomio autorregresivo** (AR) del modelo,

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad [52]$$

es el **operador** o **polinomio media móvil** (MA), $(A_t) \sim \text{IID}(0, \sigma_A^2)$, y $\mu, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ son parámetros tales que todas las raíces de la ecuación característica AR

$$\phi(x) = 1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0 \quad [53]$$

están fuera del círculo unitario (**condición de estacionariedad**). Un modelo ARMA(p, q) es **invertible** si todas las raíces de la ecuación característica MA

$$\theta(x) = 1 - \theta_1 x - \theta_2 x^2 - \dots - \theta_q x^q = 0 \quad [54]$$

están fuera del círculo unitario (**condición de invertibilidad**).

Algunas propiedades teóricas de cualquier modelo ARMA(p, q) estacionario e invertible son las siguientes:

⇒ La media del modelo es

$$\mu_Y = E[Y_t] = \frac{\mu}{1 - \phi_1 - \phi_2 - \dots - \phi_p} = \frac{\mu}{\phi(1)}, \quad [55]$$

donde $\phi(1)$ es el valor del operador AR [51] evaluado en $B = 1$.

⇒ El modelo puede escribirse como en [49]-[50], o bien como

$$\phi(B)\tilde{Y}_t = \theta(B)A_t \quad (\text{con } \tilde{Y}_t = Y_t - \mu_Y). \quad [56]$$

⇒ [*] La condición de **estacionariedad** de un modelo ARMA garantiza que los parámetros o **pesos** $\psi_0, \psi_1, \psi_2, \dots$ del polinomio de orden infinito

$$\psi(B) = \frac{\theta(B)}{\phi(B)} = 1 + \psi_1 B + \psi_2 B^2 + \dots = \sum_{i=0}^{\infty} \psi_i B^i \quad (\psi_0 = 1) \quad [57]$$

convergen a cero, por lo que la representación $\tilde{Y}_t = \psi(B)A_t$ (PSI o MA) del modelo implica que $\partial \tilde{Y}_t / \partial A_{t-i} \rightarrow 0$ cuando $i \rightarrow \infty$.

⇒ [*] La condición de **invertibilidad** de un modelo ARMA garantiza que los parámetros o **pesos** $\pi_0, \pi_1, \pi_2, \dots$ del polinomio de orden infinito

$$\pi(B) = \frac{\phi(B)}{\theta(B)} = 1 - \pi_1 B - \pi_2 B^2 - \dots = - \sum_{i=0}^{\infty} \pi_i B^i \quad (\pi_0 = -1) \quad [58]$$

convergen a cero, por lo que la representación $\pi(B)\tilde{Y}_t = A_t$ (PI o AR) del modelo implica que $\partial \tilde{Y}_t / \partial \tilde{Y}_{t-i} \rightarrow 0$ cuando $i \rightarrow \infty$.

Observación 39: [*] En modelos $AR(p)$, $\psi(B)$ es un polinomio infinito (como la ACF) con cada ψ_i ($i \geq 1$) expresable en función de ϕ_1, \dots, ϕ_p , mientras que $\pi(B)$ es un polinomio finito (como la PACF) de grado p con $\pi_i = \phi_i$ ($i = 1, \dots, p$). En modelos $MA(q)$, $\psi(B)$ es un polinomio finito (como la ACF) de grado q con $\psi_i = -\theta_i$ ($i = 1, \dots, q$), mientras que $\pi(B)$ es un polinomio infinito (como la PACF) con cada π_i ($i \geq 1$) expresable en función de $\theta_1, \dots, \theta_q$. En modelos $ARMA(p, q)$, tanto $\psi(B)$ como $\pi(B)$ son polinomios infinitos (como la ACF y la PACF), con cada ψ_i y cada π_i ($i \geq 1$) expresables en función de $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$. En general, los pesos (coeficientes) ψ_i y π_i de [57]-[58] pueden calcularse recursivamente teniendo en cuenta que $\phi(B)\psi(B) = \theta(B)$ y que $\theta(B)\pi(B) = \phi(B)$, de manera que, por un lado,

$$\psi_i = \phi_1\psi_{i-1} + \phi_2\psi_{i-2} + \dots + \phi_p\psi_{i-p} - \theta_i$$

para todo $i > 0$ (donde $\psi_0 = 1$, $\psi_i = 0$ si $i < 0$ y $\theta_i = 0$ si $i > q$), y, por otro lado,

$$\pi_i = \theta_1\pi_{i-1} + \theta_2\pi_{i-2} + \dots + \theta_q\pi_{i-q} + \phi_i$$

para todo $i > 0$ (donde $\pi_0 = -1$, $\pi_i = 0$ si $i < 0$ y $\phi_i = 0$ si $i > p$). Por último, de la relación $\psi(B)\pi(B) = 1$ implícita en [57]-[58] se deduce que

$$\pi_i = -\sum_{j=0}^{i-1} \pi_j\psi_{i-j} \quad (\pi_0 = -1), \quad \psi_i = \sum_{j=0}^{i-1} \psi_j\pi_{i-j} \quad (\psi_0 = 1) \quad \text{para todo } i \geq 1.$$

Las secuencias ψ_1, ψ_2, \dots y π_1, π_2, \dots calculadas a partir de un modelo ARMA estacionario siguen pautas semejantes a las de la ACF y la PACF, respectivamente, del modelo. El cálculo de la ACF y la PACF teóricas de un modelo estacionario (incluyendo el modelo multiplicativo general [62] de la Sección 2.8) puede llevarse a cabo en EViews con el programa PRG07-TACF-TPACF.PRG. El cálculo de los pesos ψ_i y π_i ($i = 1, 2, \dots$) en cualquier modelo univariante (incluyendo el modelo ARIMA general [63] de la Sección 2.9) está programado en PRG08-PSI-PI.PRG. IEV41: pp. 85-94

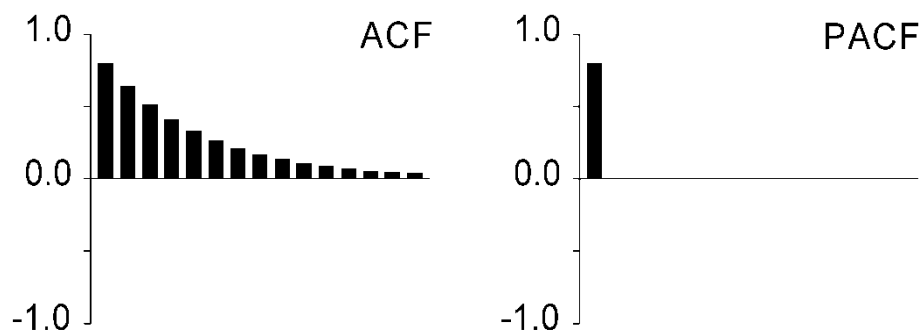


FIGURA 2.16

AR(1): $Y_t = 0.8Y_{t-1} + A_t$.
Parámetro positivo.

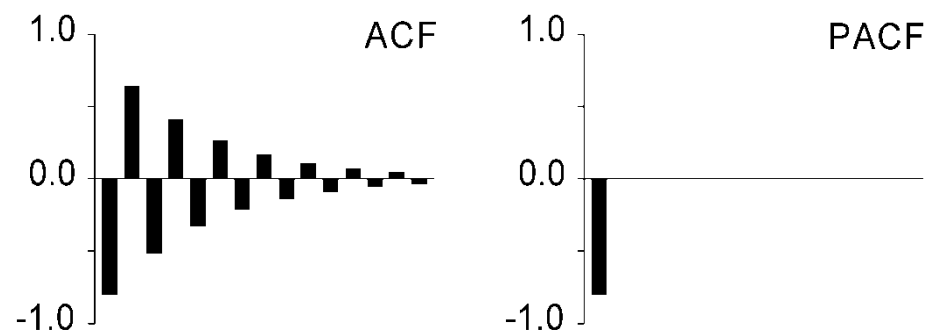


FIGURA 2.17

AR(1): $Y_t = -0.8Y_{t-1} + A_t$.
Parámetro negativo.

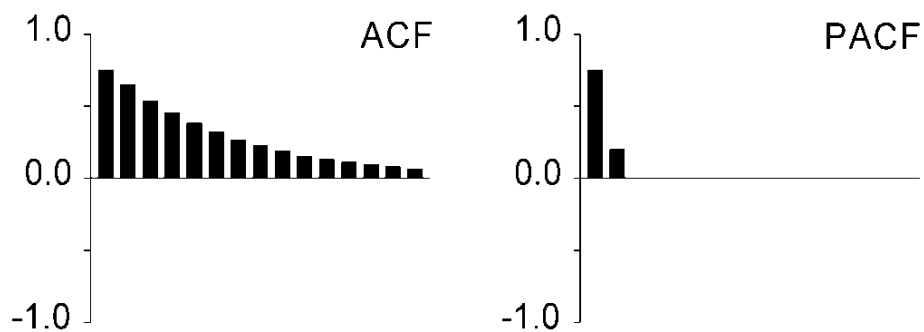


FIGURA 2.18

AR(2): $Y_t = 0.6Y_{t-1} + 0.2Y_{t-2} + A_t$.
Raíces de [53] reales: -4.19, +1.19 (dominante).

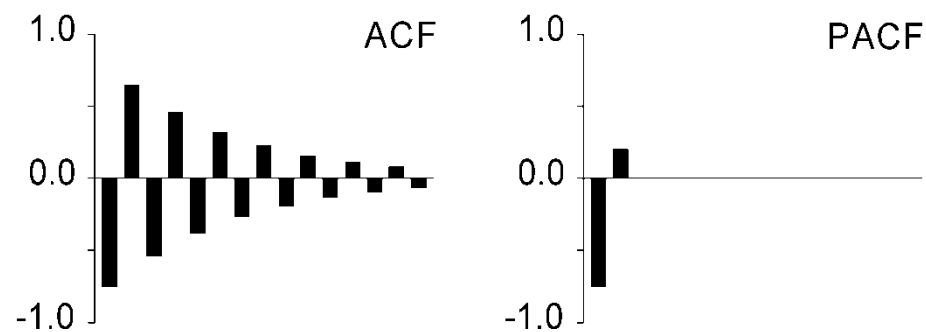


FIGURA 2.19

AR(2): $Y_t = -0.6Y_{t-1} + 0.2Y_{t-2} + A_t$.
Raíces de [53] reales: +4.19, -1.19 (dominante).

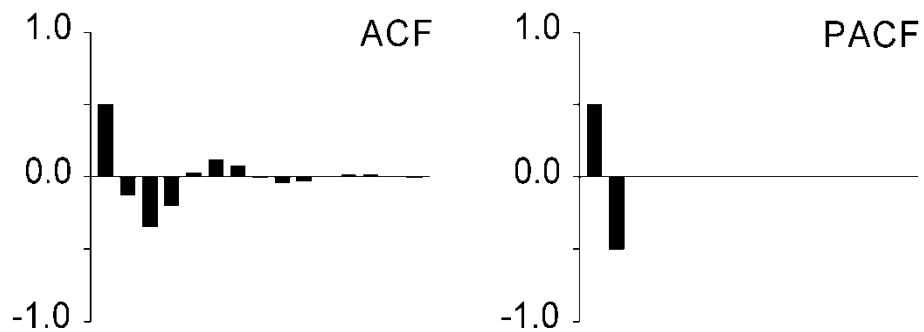


FIGURA 2.20

AR(2): $Y_t = 0.75Y_{t-1} - 0.5Y_{t-2} + A_t$.
 Raíces de [53] complejas: $+0.75 \pm 1.20i$.

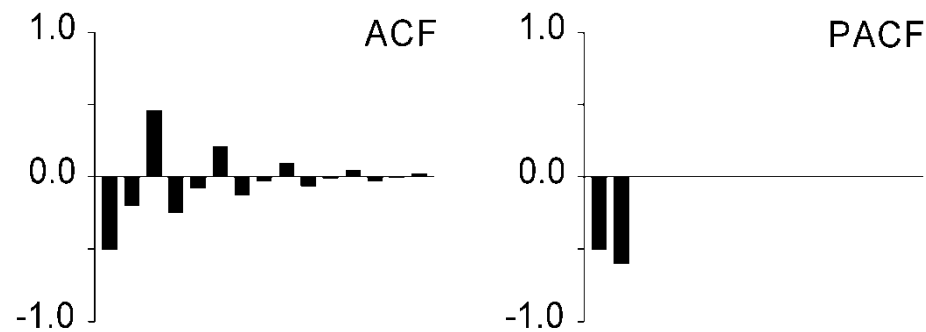


FIGURA 2.21

AR(2): $Y_t = -0.8Y_{t-1} - 0.6Y_{t-2} + A_t$.
 Raíces de [53] complejas: $-0.67 \pm 1.11i$.

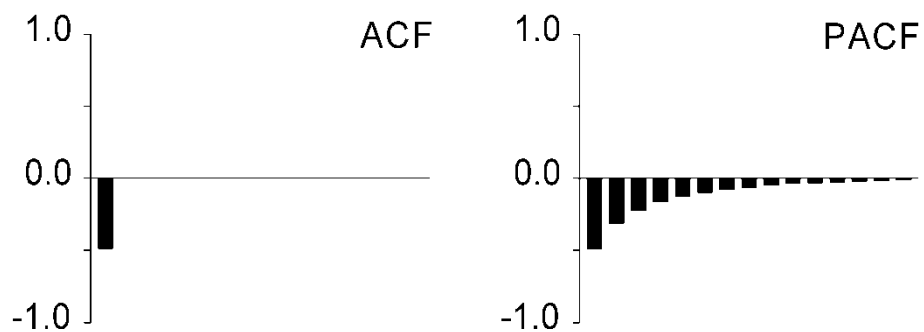


FIGURA 2.22

MA(1): $Y_t = A_t - 0.8A_{t-1}$.
 Parámetro positivo.

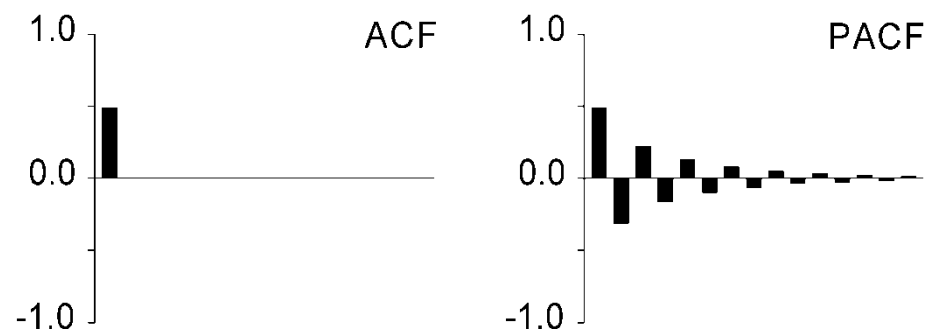


FIGURA 2.23

MA(1): $Y_t = A_t + 0.8A_{t-1}$.
 Parámetro negativo.

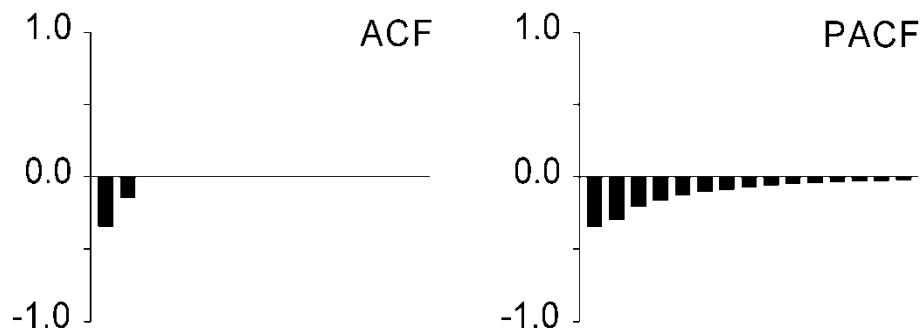


FIGURA 2.24

MA(2): $Y_t = A_t - 0.6A_{t-1} - 0.2A_{t-2}$.
 Raíces de [54] reales: $-4.19, +1.19$ (dominante).

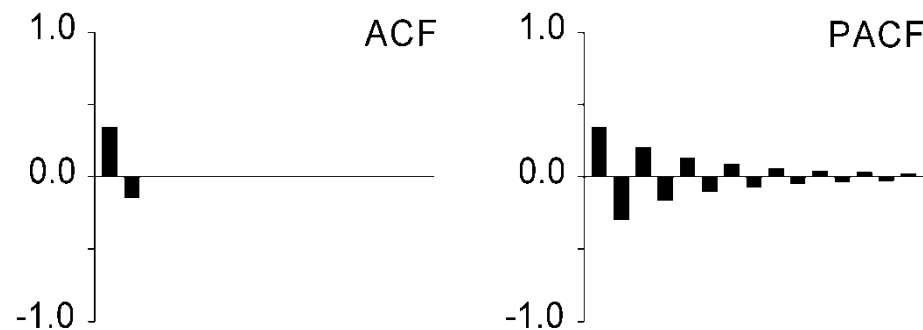


FIGURA 2.25

MA(2): $Y_t = A_t + 0.6A_{t-1} - 0.2A_{t-2}$.
 Raíces de [54] reales: $+4.19, -1.19$ (dominante).

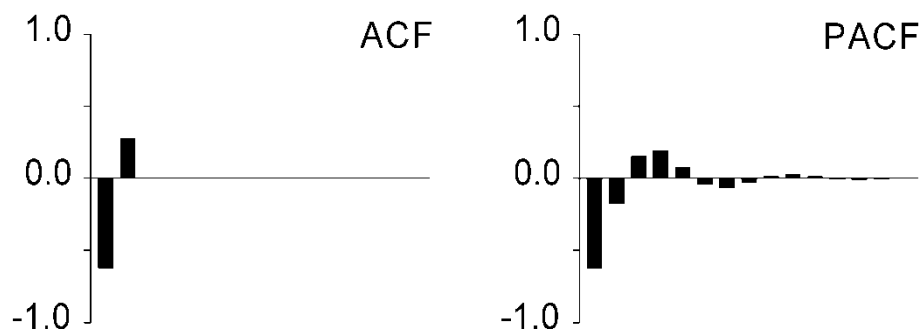


FIGURA 2.26

MA(2): $Y_t = A_t - 0.75A_{t-1} + 0.5A_{t-2}$.
 Raíces de [54] complejas: $+0.75 \pm 1.20i$.

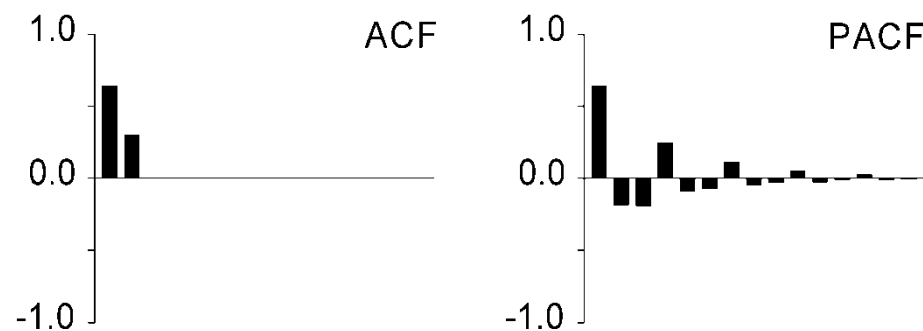


FIGURA 2.27

MA(2): $Y_t = A_t + 0.8A_{t-1} + 0.6A_{t-2}$.
 Raíces de [54] complejas: $-0.67 \pm 1.11i$.

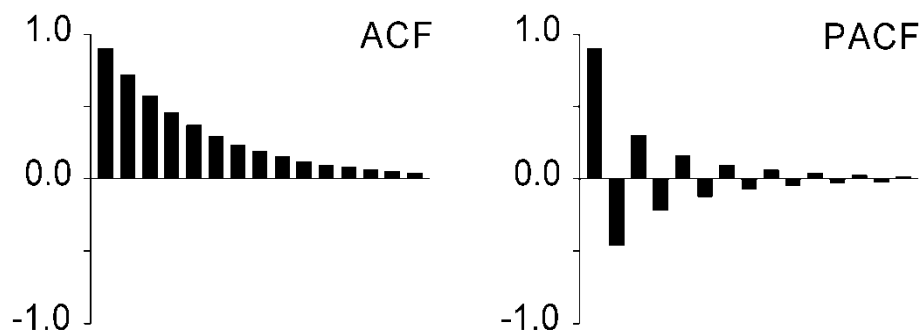


FIGURA 2.28

ARMA(1,1): $Y_t = 0.8Y_{t-1} + A_t + 0.8A_{t-1}$.

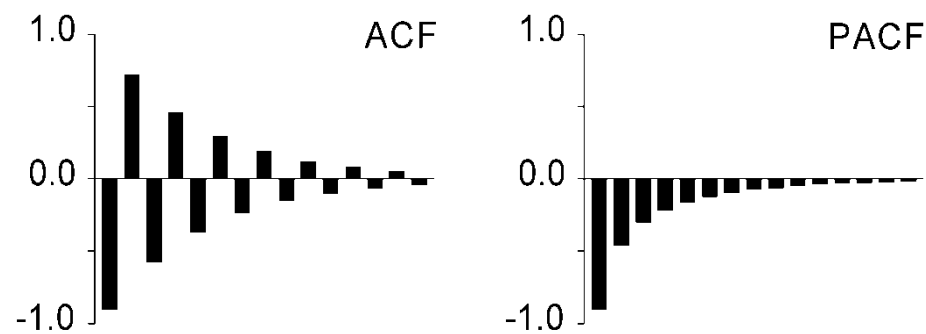


FIGURA 2.29

ARMA(1,1): $Y_t = -0.8Y_{t-1} + A_t - 0.8A_{t-1}$.

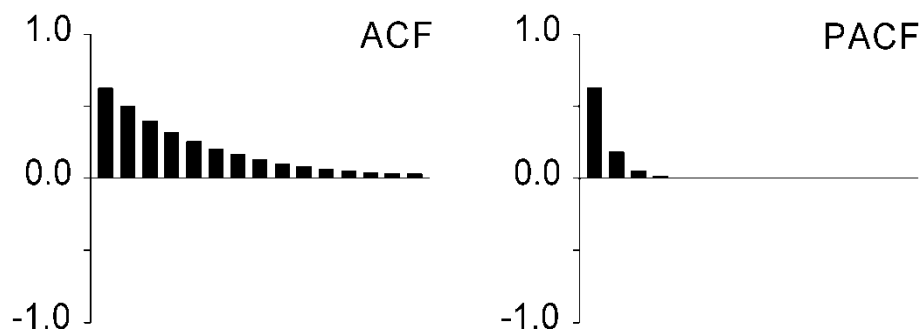


FIGURA 2.30

ARMA(1,1): $Y_t = 0.8Y_{t-1} + A_t - 0.3A_{t-1}$.

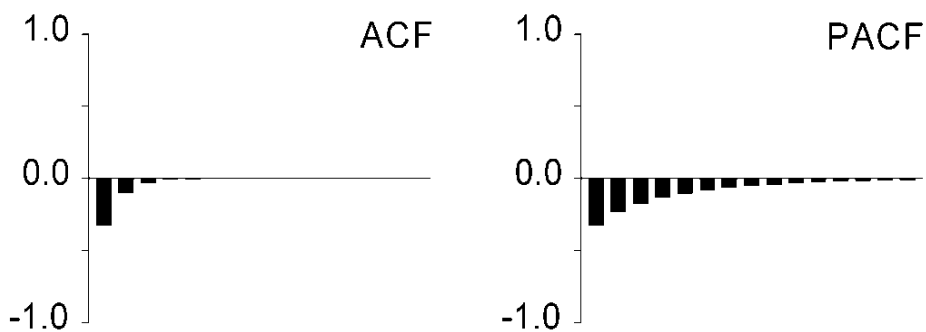


FIGURA 2.31

ARMA(1,1): $Y_t = 0.3Y_{t-1} + A_t - 0.8A_{t-1}$.

TABLA 2.3

Algunas Propiedades Teóricas de Modelos ARMA

Propiedades	Modelo AR(p)	Modelo MA(q)	Modelo ARMA(p,q)
Condición de estacionariedad	Raíces de $\phi(x) = 0$ fuera del círculo unitario	Siempre estacionario	Raíces de $\phi(x) = 0$ fuera del círculo unitario
Condición de invertibilidad	Siempre invertible	Raíces de $\theta(x) = 0$ fuera del círculo unitario	Raíces de $\theta(x) = 0$ fuera del círculo unitario
ACF	Infinita: Mezcla de exponenciales y oscilaciones amortiguadas	Finita: Igual a 0 después de $k = q$	Infinita: Mezcla de exponenciales y oscilaciones amortiguadas después de $k = q - p$
PACF	Finita: Igual a 0 después de $k = p$	Infinita: Dominada por mezcla de exponenciales y oscilaciones amortiguadas	Infinita: Dominada por mezcla de exponenciales y oscilaciones amortiguadas después de $k = p - q$
\tilde{Y}_t en términos de A_{t-i}	$\tilde{Y}_t = \phi^{-1}(B)A_t$	$\tilde{Y}_t = \theta(B)A_t$	$\tilde{Y}_t = \phi^{-1}(B)\theta(B)A_t$
Operador $\psi(B)$	$\psi(B) = \phi^{-1}(B)$ (infinito)	$\psi(B) = \theta(B)$ (finito)	$\psi(B) = \phi^{-1}(B)\theta(B)$ (infinito)
\tilde{Y}_t en términos de \tilde{Y}_{t-i}	$\phi(B)\tilde{Y}_t = A_t$	$\theta^{-1}(B)\tilde{Y}_t = A_t$	$\theta^{-1}(B)\phi(B)\tilde{Y}_t = A_t$
Operador $\pi(B)$	$\pi(B) = \phi(B)$ (finito)	$\pi(B) = \theta^{-1}(B)$ (infinito)	$\pi(B) = \theta^{-1}(B)\phi(B)$ (infinito)

2.8 MODELOS MULTIPLICATIVOS

Con mucha frecuencia, la ACF y la PACF muestrales de series estacionarias obtenidas a partir de series estacionales, presentan pautas significativas en sus retardos múltiplos del período estacional (**retardos estacionales**, como 12, 24, 36, ... para series mensuales, ó 4, 8, 12, ... para series trimestrales) y alrededor de dichos retardos.

Estas pautas representan la **dinámica estacional** (un mes con respecto al mismo mes en años anteriores) de la serie original que aún está presente en la serie transformada (estacionaria), y que acompaña a (e interactúa con) la **dinámica regular** (un mes con respecto a los meses inmediatamente anteriores en el mismo año) de dicha serie.

2.8.1 Ejemplo - Modelo $AR(1) \times AR(1)_{12}$ para Series Mensuales

$$(1 - \phi_1 B)(1 - \Phi_1 B^{12})\tilde{Y}_t = A_t, \quad [59]$$

donde ϕ_1 y Φ_1 son parámetros (con $|\phi_1| < 1, |\Phi_1| < 1$), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Los términos $AR(1)$ **regular** $(1 - \phi_1 B)$ y $AR(1)$ **anual** $(1 - \Phi_1 B^{12})$ describen la dinámica regular y anual, respectivamente, de una serie estacionaria.

2.8.2 Ejemplo - Modelo MA(1) × MA(1)₁₂ para Series Mensuales

$$\tilde{Y}_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})A_t, \quad [60]$$

donde θ_1 y Θ_1 son parámetros (con $|\theta_1| < 1, |\Theta_1| < 1$), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Los términos MA(1) **regular** $(1 - \theta_1 B)$ y MA(1) **anual** $(1 - \Theta_1 B^{12})$ describen la dinámica regular y anual, respectivamente, de una serie estacionaria.

2.8.3 Ejemplo - Modelo MA(1) × AR(2)₁₂ para Series Mensuales

$$(1 - \Phi_1 B^{12} - \Phi_2 B^{24})\tilde{Y}_t = (1 - \theta_1 B)A_t, \quad [61]$$

donde Φ_1, Φ_2 y θ_1 son parámetros (con las raíces de $1 - \Phi_1 x - \Phi_2 x^2 = 0$ fuera del círculo unitario, y $|\theta_1| < 1$), y $(A_t) \sim \text{IID}(0, \sigma_A^2)$.

Los términos MA(1) **regular** $(1 - \theta_1 B)$ y AR(2) **anual** $(1 - \Phi_1 B^{12} - \Phi_2 B^{24})$ describen la dinámica regular y anual, respectivamente, de una serie estacionaria.

2.8.4 Modelo Multiplicativo General ARMA(p,q) × ARMA(P,Q)_s

$$\phi_p(B)\Phi_P(B^S)\tilde{Y}_t = \theta_q(B)\Theta_Q(B^S)A_t. \quad [62]$$

- ⇒ Comparar con [56] y [49]-[50]. En particular, $\phi_p(B)$ y $\theta_q(B)$ son los operadores AR (estacionario) y MA (invertible) habituales (que se denominan **operadores regulares**).
- ⇒ $\Phi_P(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS} = 1 - \sum_{i=1}^P \Phi_i B^{iS}$ es el **operador AR estacional o anual**, con las raíces de $\Phi_P(x) = 0$ fuera del círculo unitario.
- ⇒ $\Theta_Q(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS} = 1 - \sum_{i=1}^Q \Theta_i B^{iS}$ es el **operador MA estacional o anual**, con las raíces de $\Theta_Q(x) = 0$ fuera del círculo unitario.

Observación 40: La **ACF teórica** de un modelo $\text{ARMA}(p, q) \times \text{ARMA}(P, Q)_S$ presenta las pautas siguientes: [i] en los primeros retardos (1, 2, 3, ...) aparece la ACF de la parte regular $\text{ARMA}(p, q)$, [ii] en los retardos anuales ($S, 2S, 3S, \dots$) aparece la ACF de la parte anual $\text{ARMA}(P, Q)_S$, [iii] la pauta de la ACF de la parte regular se repite a la izquierda (en el sentido \leftarrow) y a la derecha (en el sentido \rightarrow) de cada coeficiente anual, con signo igual al producto de los signos del coeficiente anual y de los coeficientes regulares correspondientes.

Por su parte, la **PACF teórica** de un modelo $\text{ARMA}(p, q) \times \text{ARMA}(P, Q)_S$ presenta las pautas siguientes: [i] en los primeros retardos (1, 2, 3, ...) aparece la PACF de la parte regular $\text{ARMA}(p, q)$, [ii] en los retardos anuales ($S, 2S, 3S, \dots$) aparece la PACF de la parte anual $\text{ARMA}(P, Q)_S$, [iii] a la derecha de cada coeficiente anual se repite (en el sentido \rightarrow) la pauta de la PACF de la parte regular, con el signo cambiado si el coeficiente anual es positivo, o con su signo si dicho coeficiente es negativo, [iv] a la izquierda de cada coeficiente anual se repite (en el sentido \leftarrow) la pauta de la ACF de la parte regular, con signo igual al producto de los signos del coeficiente anual y de los coeficientes regulares correspondientes.

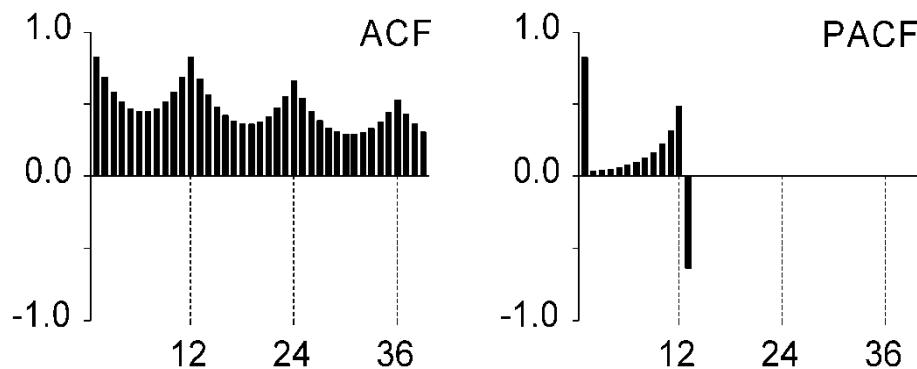


FIGURA 2.32

$AR(1) \times AR(1)_{12}$, con $\phi_1 = 0.8$, $\Phi_1 = 0.8$.

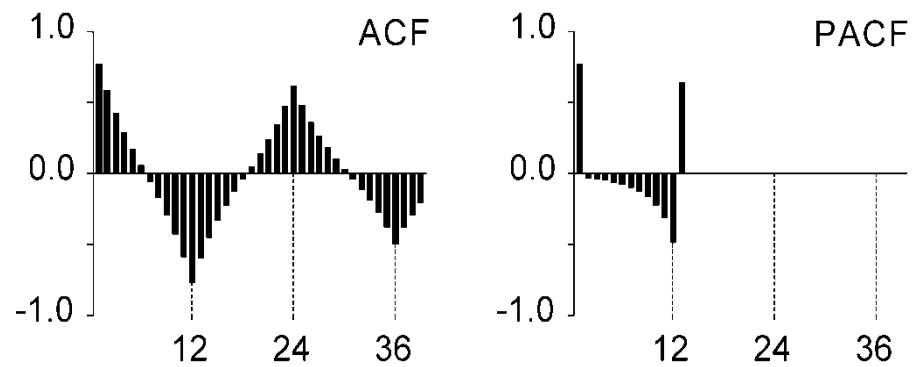


FIGURA 2.33

$AR(1) \times AR(1)_{12}$, con $\phi_1 = 0.8$, $\Phi_1 = -0.8$.

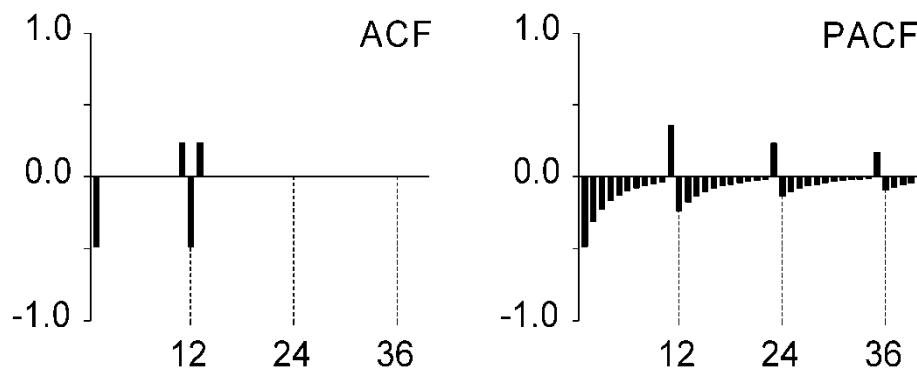


FIGURA 2.34

$MA(1) \times MA(1)_{12}$, con $\theta_1 = 0.8$, $\Theta_1 = 0.8$.

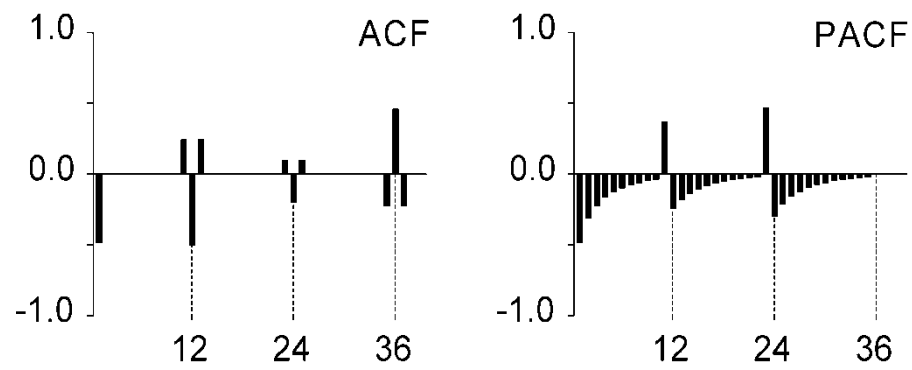


FIGURA 2.35

$MA(1) \times AR(2)_{12}$, con $\theta_1 = 0.8$, $\Phi_1 = -0.8$, $\Phi_2 = -0.6$.

2.9 MODELOS NO ESTACIONARIOS

Un modelo univariante no estacionario para una serie no estacionaria es el modelo que resulta de considerar conjuntamente [i] las transformaciones que requiere dicha serie para hacerla estacionaria y [ii] un modelo univariante estacionario para la serie transformada.

2.9.1 Modelo Multiplicativo General $ARIMA(p,d,q) \times ARIMA(P,D,Q)_S$

$$\phi_p(B)\Phi_P(B^S)[\nabla^d \nabla_S^D Y'_t - \mu_W] = \theta_q(B)\Theta_Q(B^S)A_t, \quad [63]$$

donde $\mu_W = E[W_t]$ con $W_t = \nabla^d \nabla_S^D Y'_t$ (comparar con [62]).

2.9.2 Dos Interpretaciones Equivalentes

El modelo general [63] puede interpretarse como un **modelo no estacionario** para una serie (original) no estacionaria que requiere [i] una transformación de Box-Cox, [ii] d diferencias regulares y [iii] D diferencias estacionales de periodo S para hacerla estacionaria, o bien como un **modelo estacionario** para una serie estacionaria que se obtiene aplicándole la secuencia de transformaciones [i]-[iii] a una serie (original) no estacionaria.

2.10 IDENTIFICACIÓN - RESUMEN

ENTRADA: $\mathbf{y} = [y_1, \dots, y_N]'$ (serie temporal original); S (período estacional).

1. Elegir m en la transformación de Box-Cox – Gráfico desviación típica-media. Se obtiene y'_t ($t = 1, \dots, N$), con $y'_t = \ln y_t$ si $m = 0$, $y'_t = y_t^m$ si $m \neq 0$.
2. Elegir d y D – Gráfico temporal (tendencia, estacionalidad) y ACF (amortiguamiento) de y'_t ($t = 1, \dots, N$). Se obtiene una **serie transformada** (tentativamente **estacionaria**) con $n = N - d - DS$ observaciones del tipo $w_t = \nabla^d \nabla_S^D y'_t$.
3. Incluir o no μ_W – Contraste de significación a partir de la serie transformada w_t .
4. Elegir p, P, q, Q – ACF y PACF muestrales de la serie transformada w_t comparadas con ACF y PACF teóricas de modelos $\text{ARMA}(p, q) \times \text{ARMA}(P, Q)_S$ estacionarios:
 - ⇒ Considerar siempre en primer lugar lo más obvio, empezando con la parte regular.
 - ⇒ Un valor grande en la ACF o en la PACF muestrales en un retardo k dado, puede **no** ser relevante si se encuentra aislado y k es grande.
 - ⇒ Una configuración reconocible de valores pequeños **sí** puede ser relevante.

SALIDA: Modelo tentativo de tipo $ARIMA(p, d, q) \times ARIMA(P, D, Q)_S$ para la serie original y_t [o, equivalentemente, modelo tentativo de tipo $ARMA(p, q) \times ARMA(P, Q)_S$ para la serie transformada w_t].

Observación 41 - Contrastes Formales de Raíces Unitarias: A veces pueden ayudar a elegir d (el número de diferencias regulares que requiere una serie para estabilizar su nivel medio general) en el paso 2 anterior.

IEV41: p. 85

Observación 42 - Tasas de Variación: La **tasa convencional de variación** de y_t es (en proporción, o en tanto por 1)

$$tcv_t(k) = \frac{y_t}{y_{t-k}} - 1 = \frac{y_t - y_{t-k}}{y_{t-k}} = \frac{(1 - B^k)y_t}{y_{t-k}},$$

mientras que la **tasa logarítmica de variación** de y_t es (en proporción, o en tanto por 1)

$$tlv_t(k) = \ln[1 + tcv_t(k)] = \ln \frac{y_t}{y_{t-k}} = \ln y_t - \ln y_{t-k} = (1 - B^k) \ln y_t,$$

que es aproximadamente igual a $tcv_t(k)$ para valores de $tcv_t(k)$ "pequeños":

$$\ln(1 + x) \cong \ln(1 + x_0) + \frac{1}{1 + x_0} \times [(1 + x) - (1 + x_0)] = \ln(1 + x_0) + \frac{x - x_0}{1 + x_0} = x \text{ cuando } x_0 = 0.$$

Dado que muchas series temporales requieren en la práctica una transformación logarítmica y algún tipo de diferenciación para hacerlas estacionarias, las series resultantes pueden interpretarse en términos de diferentes tasas logarítmicas de variación. Por ejemplo, si y_1, \dots, y_N es una serie mensual, entonces:

- ⇒ $w_{t1} = \nabla \ln y_t$ es una tasa logarítmica de variación mensual.
- ⇒ $w_{t2} = \nabla_{12} \ln y_t$ es una tasa logarítmica de variación interanual.
- ⇒ $w_{t3} = \nabla^2 \ln y_t$ es una variación (absoluta) mensual de una tasa logarítmica de variación mensual.
- ⇒ $w_{t4} = \nabla \nabla_{12} \ln y_t$ es una variación (absoluta) mensual de una tasa logarítmica de variación interanual, o bien una variación (absoluta) interanual de una tasa logarítmica de variación mensual.

EJEMPLOS DE IDENTIFICACIÓN

1. ST02: Serie TVPIB.
2. ST14: Serie VIVIN.
3. ST19: Serie AIRLINE.
4. ST20: Serie TPARO.

ESTIMACIÓN

2.11 ESTIMACIÓN - RESUMEN

ENTRADA: $\mathbf{w} = [w_1, \dots, w_n]'$ (serie estacionara con $n = N - d - DS$ observaciones del tipo $w_t = \nabla^d \nabla_S^D y_t'$); modelo $\text{ARMA}(p, q) \times \text{ARMA}(P, Q)_S$ para la serie \mathbf{w} .

1. Estimación de los **parámetros** del modelo (máxima verosimilitud, mínimos cuadrados).
2. Cálculo de los **residuos** del modelo estimado.

SALIDA: Estimaciones de los parámetros ($\hat{\mu}_W$, $\hat{\phi} = [\hat{\phi}_1, \dots, \hat{\phi}_p]'$, $\hat{\Phi} = [\hat{\Phi}_1, \dots, \hat{\Phi}_P]'$, $\hat{\theta} = [\hat{\theta}_1, \dots, \hat{\theta}_q]'$, $\hat{\Theta} = [\hat{\Theta}_1, \dots, \hat{\Theta}_Q]'$, $\hat{\sigma}_A^2$); estimación de la matriz de varianzas-covarianzas del estimador utilizado (\mathbf{V}); serie de **residuos** ($\hat{a}_1, \dots, \hat{a}_n$).

Observación 43 - Estimación de modelos univariantes: La **función de verosimilitud exacta** (FVE) asociada con el modelo univariante [63] es la función de densidad de un vector $\mathbf{W} = [W_1, \dots, W_n]'$ (con $W_t = \nabla^d \nabla_S^D Y_t'$) del que se supone procede la serie (estacionaria) $\mathbf{w} = [w_1, \dots, w_n]'$, considerada como función de los parámetros que figuran en [63]. Si \mathbf{W} sigue una distribución Normal multivariante [es decir, bajo la hipótesis de que $(A_t) \sim \text{NIID}(0, \sigma_A^2)$], entonces la FVE asociada con el modelo univariante [63] es

$$L(\boldsymbol{\beta}, \sigma_A^2) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Gamma}_n|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \tilde{\mathbf{w}}' \boldsymbol{\Gamma}_n^{-1} \tilde{\mathbf{w}}\right), \quad (1)$$

donde $\boldsymbol{\beta}$ contiene todos los parámetros del modelo [63] excepto σ_A^2 , $\boldsymbol{\Gamma}_n = \text{Var}[\mathbf{W}]$, y $\tilde{\mathbf{w}} = \mathbf{w} - \text{E}[\mathbf{W}]$, con $\text{E}[\mathbf{W}] = [\mu_W, \dots, \mu_W]'$. Dado que $\boldsymbol{\Gamma}_n$ siempre puede escribirse como $\boldsymbol{\Gamma}_n = \sigma_A^2 \boldsymbol{\Omega}$ (definiendo $\boldsymbol{\Omega} = \sigma_A^{-2} \boldsymbol{\Gamma}_n$), el logaritmo (ln) de la FVE (1) puede escribirse como

$$\ln L(\boldsymbol{\beta}, \sigma_A^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma_A^2 - \frac{1}{2} \ln |\boldsymbol{\Omega}| - \frac{\sigma_A^{-2}}{2} \tilde{\mathbf{w}}' \boldsymbol{\Omega}^{-1} \tilde{\mathbf{w}}. \quad (2)$$

Por ejemplo, para modelos AR(1) (ver 2.7.1) y MA(1) (ver 2.7.4),

$$\boldsymbol{\Omega} = \frac{1}{1-\phi_1^2} \begin{bmatrix} 1 & \phi_1 & \phi_1^2 & \dots & \phi_1^{n-2} & \phi_1^{n-1} \\ \phi_1 & 1 & \phi_1 & \dots & \phi_1^{n-3} & \phi_1^{n-2} \\ \phi_1^2 & \phi_1 & 1 & \dots & \phi_1^{n-4} & \phi_1^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_1^{n-2} & \phi_1^{n-3} & \phi_1^{n-4} & \dots & 1 & \phi_1 \\ \phi_1^{n-1} & \phi_1^{n-2} & \phi_1^{n-3} & \dots & \phi_1 & 1 \end{bmatrix}, \quad \boldsymbol{\Omega} = \begin{bmatrix} 1 + \theta_1^2 & -\theta_1 & 0 & \dots & 0 & 0 \\ -\theta_1 & 1 + \theta_1^2 & -\theta_1 & \dots & 0 & 0 \\ 0 & -\theta_1 & 1 + \theta_1^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 + \theta_1^2 & -\theta_1 \\ 0 & 0 & 0 & \dots & -\theta_1 & 1 + \theta_1^2 \end{bmatrix}. \quad (3)$$

En todos los casos, los elementos de $\boldsymbol{\Omega} = \sigma_A^{-2} \boldsymbol{\Gamma}_n$ en (2) sólo dependen [como se ilustra en (3)] de los parámetros AR y MA del modelo [63] (recogidos en $\boldsymbol{\phi} = [\phi_1, \dots, \phi_p]'$, $\boldsymbol{\Phi} = [\Phi_1, \dots, \Phi_P]'$, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_q]'$ y $\boldsymbol{\Theta} = [\Theta_1, \dots, \Theta_Q]'$); el cálculo de dichos elementos es semejante al de las ACFs teóricas de las Secciones 2.7-2.8.

Para obtener estimaciones por **Máxima Verosimilitud** (MV) de $\boldsymbol{\beta}$ y σ_A^2 , se requiere maximizar numéricamente (1) ó

(2) con respecto a β y a σ_A^2 . Si $\hat{\beta}$ y $\hat{\sigma}_A^2$ son las estimaciones numéricas de β y σ_A^2 que maximizan (localmente) (2), entonces ocurre necesariamente que

$$\frac{\partial \ln L(\hat{\beta}, \hat{\sigma}_A^2)}{\partial \sigma_A^2} = 0 \Leftrightarrow \hat{\sigma}_A^2 = \frac{S(\hat{\beta})}{n}, \quad (4)$$

donde $S(\hat{\beta})$ es el valor de la suma de cuadrados $S(\beta) = \tilde{\mathbf{w}}' \Omega^{-1} \tilde{\mathbf{w}}$ evaluada en $\beta = \hat{\beta}$. Sustituyendo σ_A^2 por $\frac{S(\beta)}{n}$ en (2), se obtiene el ln de la FVE **concentrada** en β :

$$l(\beta) = -\frac{n}{2} [\ln(2\pi) + 1 - \ln n] - \frac{n}{2} \ln \left(\tilde{\mathbf{w}}' \Omega^{-1} \tilde{\mathbf{w}} \times |\Omega|^{\frac{1}{n}} \right),$$

de manera que

$$\text{MAX } L(\beta, \sigma_A^2) \Leftrightarrow \text{MAX } \ln L(\beta, \sigma_A^2) \Leftrightarrow \text{MAX } l(\beta) \Leftrightarrow \text{MIN } \tilde{\mathbf{w}}' \Omega^{-1} \tilde{\mathbf{w}} \times |\Omega|^{\frac{1}{n}}.$$

En resumen, para estimar β y σ_A^2 por MV se requiere (i) minimizar numéricamente con respecto a β la función

$$F(\beta) = \tilde{\mathbf{w}}' \Omega^{-1} \tilde{\mathbf{w}} \times |\Omega|^{\frac{1}{n}} \quad (5)$$

(lo que proporciona una estimación puntual $\hat{\beta}$ de β), y (ii) estimar posteriormente σ_A^2 como, por ejemplo,

$$\hat{\sigma}_A^2 = \frac{S(\hat{\beta})}{n-m}, \quad (6)$$

donde $m = 1 + p + P + q + Q$ es el número de parámetros contenidos en β . [Nótese que la estimación (6), que no coincide con la que figura en (4), es análoga a la estimación de mínimos cuadrados de σ^2 en modelos de regresión.]

El problema fundamental para calcular una estimación MV de β consiste en que (5) es una función no lineal complicada de β . Por ejemplo, para un modelo AR(1), a partir de (3) puede comprobarse que

$$\mathbf{\Omega}^{-1} = \begin{bmatrix} 1 & -\phi_1 & 0 & \cdots & 0 & 0 \\ -\phi_1 & 1 + \phi_1^2 & -\phi_1 & \cdots & 0 & 0 \\ 0 & -\phi_1 & 1 + \phi_1^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 + \phi_1^2 & -\phi_1 \\ 0 & 0 & 0 & \cdots & -\phi_1 & 1 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} (1 - \phi_1^2)^{\frac{1}{2}} & 0 & 0 & \cdots & 0 & 0 \\ -\phi_1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\phi_1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -\phi_1 & 1 \end{bmatrix},$$

donde \mathbf{P} es tal que $\mathbf{P}'\mathbf{P} = \mathbf{\Omega}^{-1}$; por lo tanto, en este caso el primer factor de $F(\boldsymbol{\beta})$ en (5) es

$$\tilde{\mathbf{w}}'\mathbf{\Omega}^{-1}\tilde{\mathbf{w}} = (\mathbf{P}\tilde{\mathbf{w}})'(\mathbf{P}\tilde{\mathbf{w}}) = (1 - \phi_1^2)(w_1 - \mu_W)^2 + \sum_{t=2}^n [(w_t - \mu_W) - \phi_1(w_{t-1} - \mu_W)]^2,$$

mientras que el segundo factor de $F(\boldsymbol{\beta})$ en (5) es

$$|\mathbf{\Omega}|^{\frac{1}{n}} = |\mathbf{\Omega}^{-1}|^{-\frac{1}{n}} = |\mathbf{P}'\mathbf{P}|^{-\frac{1}{n}} = |\mathbf{P}|^{-\frac{2}{n}} = \left[(1 - \phi_1^2)^{\frac{1}{2}} \right]^{-\frac{2}{n}} = (1 - \phi_1^2)^{-\frac{1}{n}}.$$

En general, no posible obtener analíticamente una expresión para $\boldsymbol{\beta}$ que minimice (5), por lo que en la práctica, para obtener una estimación MV de $\boldsymbol{\beta}$, es necesario minimizar (5) mediante algún **método de cálculo iterativo** para minimizar numéricamente funciones de varias variables (como el método de *Gauss-Newton*, el de *Newton-Raphson*, el de *Levenberg-Marquardt*, o algún método *cuasi-Newton*). Las características de cualquier método de este tipo son:

⇒ Partiendo de una **estimación inicial** $\hat{\boldsymbol{\beta}}_0$ proporcionada por el usuario, un método iterativo genera una secuencia de estimaciones $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \dots$ que, bajo ciertas condiciones, converge a un **mínimo local** de (5). En general, no puede asegurarse que dicho mínimo sea un **mínimo global**.

- ⇒ Si (5) tiene varios mínimos locales, el mínimo al que converge la secuencia anterior depende crucialmente de la estimación inicial $\hat{\beta}_0$.
- ⇒ Si (5) es una función relativamente **plana**, puede resultar difícil localizar numéricamente un mínimo. Esta situación es análoga a la **multicolinealidad aproximada** en modelos de regresión lineal, e indica la posible presencia de términos redundantes (sobrantes) en el modelo [63].

Los detalles sobre el funcionamiento de este tipo de métodos y muchos otros aplicables a la resolución de diversos problemas numéricos, pueden consultarse, por ejemplo, en [i] Monahan, J.F. (2011), *Numerical Methods of Statistics (Second Edition)*, Cambridge University Press, y [ii] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. (2002), *Numerical Recipes in C - The Art of Scientific Computing (Second Edition)*, Cambridge University Press.

Aunque en la actualidad no existen razones objetivas para hacerlo, en la práctica a veces se estima un modelo como [63] por **Mínimos Cuadrados** (MC) (lo que se justifica suponiendo que $|\Omega|^{1/n}$ en (5) tiende a 1 para n suficientemente grande), incluso considerando alguna aproximación al término $\tilde{\mathbf{w}}' \Omega^{-1} \tilde{\mathbf{w}}$ en (5) en lugar de su expresión exacta, como la aproximación basada en el cálculo de *retroprevisiones* (Box-Jenkins-Reinsel-Ljung (2016), Capítulo 7), o la aproximación empleada por EViews (IEV41: p. 99). En teoría, todos los estimadores de β definidos implícitamente a través de las estimaciones que minimizan numéricamente $F(\beta)$ en (5) o alguna aproximación a $F(\beta)$, tienen las mismas propiedades asintóticas (consistencia, Normalidad y eficiencia asintóticas); sin embargo, cuando n no es grande, o cuando los verdaderos valores de los parámetros AR y MA están próximos a la no estacionariedad o a la no invertibilidad, pueden darse diferencias importantes entre las propiedades de unos y otros estimadores, que son en general favorables al estimador MV.

Por último, la estimación \mathbf{V} de la matriz de covarianzas del estimador de β , se puede calcular a partir de la matriz de derivadas segundas (matriz *hessiana*) de la función que se haya minimizado para calcular $\hat{\beta}$. Por ejemplo, si $\hat{\beta}$ es una estimación obtenida por MV [es decir, un mínimo local de (5)], entonces \mathbf{V} se calcula a partir de la matriz

$[\nabla^2 F(\hat{\beta})]^{-1}$ [la inversa de la matriz hessiana de $F(\beta)$ evaluada en $\beta = \hat{\beta}$]; teniendo en cuenta que $\nabla^2 F(\hat{\beta})$ representa la **curvatura** de $F(\beta)$ en $\beta = \hat{\beta}$, cuanto mayor sea dicha curvatura, con mayor precisión (menor varianza) estará localizado el verdadero valor de β ; por el contrario, si $F(\beta)$ en $\beta = \hat{\beta}$ tiene poca curvatura, el verdadero valor de β estará localizado con poca precisión (mucho varianza).

Observación 44 - Cálculo de los residuos: Igual que en un modelo de regresión, la serie de residuos $(\hat{a}_t)_{t=1}^n$ asociada con la estimación del modelo [63] contiene las diferencias entre los valores observados y los valores ajustados para la variable dependiente $W_t = \nabla^d \nabla_S^D Y_t'$ de [63]. Una vez estimado (como se describe en la Observación 43 anterior), [63] puede escribirse como $\hat{\Phi}'(B)(w_t - \hat{\mu}_W) = \hat{\Theta}'(B)\hat{a}_t$, o bien como $\hat{\Phi}'(B)w_t = \hat{\mu} + \hat{\Theta}'(B)\hat{a}_t$, donde

$$\begin{aligned}\hat{\Phi}'(B) &= \hat{\phi}_p(B) \times \hat{\Phi}_P(B^S) = 1 - \sum_{i=1}^{p'} \hat{\Phi}'_i B^i \quad (p' = p + PS), \\ \hat{\Theta}'(B) &= \hat{\theta}_q(B) \times \hat{\Theta}_Q(B^S) = 1 - \sum_{i=1}^{q'} \hat{\Theta}'_i B^i \quad (q' = q + QS), \\ \hat{\mu} &= \hat{\Phi}'(1) \times \hat{\mu}_W.\end{aligned}$$

De manera más explícita, el modelo [63] estimado puede escribirse, por lo tanto, como

$$w_t - \sum_{i=1}^{p'} \hat{\Phi}'_i w_{t-i} = \hat{\mu} + \hat{a}_t - \sum_{i=1}^{q'} \hat{\Theta}'_i \hat{a}_{t-i} \quad (t = 1, \dots, n).$$

En consecuencia, cada residuo \hat{a}_t se calcula como

$$\hat{a}_t = w_t - \left[\hat{\mu} + \sum_{i=1}^{p'} \hat{\Phi}'_i w_{t-i} - \sum_{i=1}^{q'} \hat{\Theta}'_i \hat{a}_{t-i} \right] \quad (t = 1, \dots, n), \quad (7)$$

donde el término entre corchetes es el **valor ajustado** o **previsto** para $W_t = \nabla^d \nabla_S^D Y_t'$ correspondiente al valor observado w_t (de manera que \hat{a}_t puede interpretarse como un **error de previsión** intramuestral).

Para evaluar (7), se requieren algunos valores de w_t y de \hat{a}_t para $t < 1$ (es decir, algunos *valores premuestrales*). Por ejemplo, en un modelo con $p = 2$, $q = 1$, $P = 0$ y $Q = 0$, (7) queda

$$\hat{a}_t = w_t - \left[\hat{\mu} + \hat{\phi}_1 w_{t-1} + \hat{\phi}_2 w_{t-2} - \hat{\theta}_1 \hat{a}_{t-1} \right] \quad (t = 1, \dots, n),$$

de manera que para calcular \hat{a}_1 hacen falta w_0 , w_{-1} y \hat{a}_0 (también hace falta w_0 para calcular \hat{a}_2). En la práctica, dichos valores pueden estimarse utilizando la serie estacionaria $(w_t)_{t=1}^n$ y el modelo [63] estimado (lo cual suele hacerse de forma automática cuando el modelo se ha estimado por MV), o bien (aunque no es recomendable, especialmente cuando n no es grande) se puede suponer que $w_t = \hat{\mu}_W$ (la media estimada de $W_t = \nabla^d \nabla_S^D Y_t'$) y que $\hat{a}_t = 0$ (el valor esperado de A_t) para $t < 1$.

EJEMPLOS DE ESTIMACIÓN

1. ST02: Serie TVPIB.
2. ST14: Serie VIVIN.
3. ST19: Serie AIRLINE.
4. ST20: Serie TPARO.

IEV41: pp. 98-100

DIAGNOSIS

2.12 DIAGNOSIS - RESUMEN

ENTRADA: Modelo estimado; residuos.

1. Diagnósis de **parámetros**: significación (individual y conjunta); correlaciones estimadas entre estimadores; estacionariedad e invertibilidad.
2. Diagnósis de **residuos** (reemplazan a las perturbaciones NIID en el modelo estimado): gráfico temporal; media muestral; ACF y PACF residuales (Ljung-Box); Normalidad.
3. Posible revisión del modelo según las conclusiones de los dos pasos anteriores.

SALIDA: Si modelo revisado \Rightarrow Volver a estimación. Si modelo adecuado \Rightarrow Utilizar para describir y prever la serie original.

Observación 45 - Diagnósis de parámetros: Las operaciones siguientes se refieren a los **parámetros** AR y MA incluidos en el modelo identificado y estimado:

\Rightarrow Contrastar la **significación individual** de cada parámetro del modelo y la **significación conjunta** de los parámetros

de cada operador $\phi(B)$, $\Phi(B)$, $\theta(B)$, $\Theta(B)$. A veces, estos contrastes indican que pueden eliminarse algunos parámetros no significativos.

- ⇒ Examinar la matriz \mathbf{R} de **correlaciones** estimadas entre los estimadores de los parámetros, que puede calcularse a partir de \mathbf{V} como $\mathbf{R} = \mathbf{D}^{-1}\mathbf{V}\mathbf{D}^{-1}$, donde \mathbf{D} es una matriz diagonal cuyos elementos son las raíces cuadradas de los elementos en la diagonal principal de \mathbf{V} . La presencia de correlaciones elevadas suele indicar que el modelo está **sobrep parametrizado**.
- ⇒ Comprobar que los operadores AR y MA estimados satisfacen las condiciones de estacionariedad y de invertibilidad, respectivamente. La presencia de una raíz de $\hat{\phi}(x) = 0$ o de $\hat{\Phi}(x) = 0$ próxima o igual a 1 indica **subdiferenciación**, por lo que es conveniente aplicar una diferencia regular o estacional, respectivamente, a la serie w_t . Por su parte, la presencia de una raíz de $\hat{\theta}(x) = 0$ o de $\hat{\Theta}(x) = 0$ próxima o igual a 1 indica **sobrediferenciación**, por lo que es conveniente eliminar una diferencia regular o estacional, respectivamente, de la serie w_t (además, quizás, de reconsiderar la presencia de un término constante en el modelo revisado).
- ⇒ Factorizar los operadores AR y MA estimados y comprobar que no tienen **factores comunes**. Si dichos operadores tienen factores aproximadamente comunes, en general deben cancelarse para simplificar el modelo.

Observación 46 - Diagnóstico de residuos: Las operaciones siguientes se refieren a los **residuos** calculados a partir del modelo identificado y estimado:

- ⇒ Examinar el gráfico temporal de la serie de residuos estandarizada. La presencia de cualquier tipo de tendencia suele indicar la conveniencia de aplicar una diferencia regular o estacional (según sea la tendencia) a la serie w_t . Si se observa que la dispersión de los residuos no es constante, probablemente la transformación de Box-Cox incluida en w_t no es adecuada.
- ⇒ Si el modelo no contiene el parámetro μ_W , contrastar la significación de la media de las perturbaciones utilizando

la media residual; si se rechaza la hipótesis nula del contraste, añadir el parámetro μ_W al modelo.

- ⇒ Examinar la ACF y la PACF residuales. Un amortiguamiento muy lento de la ACF indica que la serie de residuos no es estacionaria, por lo que conviene aplicar una diferencia (regular o estacional, según convenga) adicional a la serie w_t . Si la ACF y la PACF presentan alguna configuración reconocible, añadir al modelo los operadores AR y MA correspondientes. Para determinar si las autocorrelaciones simples entre las perturbaciones del modelo son todas iguales a cero hasta cierto retardo g , puede utilizarse el valor calculado del estadístico de Ljung-Box,

$$Q_{LB}(\hat{a}_t) = n(n+2) \sum_{k=1}^g \frac{r_k(\hat{a}_t)^2}{n-k},$$

que bajo la hipótesis nula de que $\rho_1(A_t) = \rho_2(A_t) = \dots = \rho_g(A_t) = 0$, sigue una distribución $\chi^2(g-b)$ (aproximadamente), donde b es el número total de parámetros AR y MA estimados.

- ⇒ Examinar el histograma de la serie de residuos y utilizar el estadístico de Jarque-Bera para contrastar la hipótesis de Normalidad de las perturbaciones.

Observación 47 - Otras Operaciones: Las operaciones siguientes pueden llevarse a cabo para completar la diagnosis de un modelo identificado y estimado:

- ⇒ Comparar la ACF y la PACF muestrales de la serie estacionaria w_t con (i) la ACF y la PACF teóricas de la parte estacionaria del modelo estimado, y con (ii) la ACF y la PACF residuales.
- ⇒ Si se han identificado y estimado varios modelos alternativos para la misma serie, se pueden comparar entre sí a través de los valores de las varianzas (o de las desviaciones típicas) residuales y de algunos **criterios de información**, y, especialmente, a través de algunos **criterios de evaluación de previsiones**.
- ⇒ Evaluar la posible presencia de observaciones influyentes mediante un **análisis de intervención** adecuado.

RAÍCES UNITARIAS - NO ESTACIONARIEDAD - NO INVERTIBILIDAD

Ejemplo 1: Modelo AR(1) no estacionario:

$$\begin{aligned} \text{M1: } (1 - \phi_1 B)(Y_t - C) &= A_t, \text{ con } \phi_1 = 1 \Rightarrow \text{M1': } \underbrace{(1 - B)}_{\nabla} (Y_t - C) = A_t. \\ \text{M1': } \nabla(Y_t - C) &= \nabla Y_t - \underbrace{\nabla C}_{=0} = A_t \Rightarrow \text{M1': } \nabla Y_t = A_t. \end{aligned}$$

Ejemplo 2: Modelo IMA(1,1) no invertible:

$$\begin{aligned} \text{M2: } \nabla Y_t &= (1 - \theta_1 B)A_t, \text{ con } \theta_1 = 1 \Rightarrow \text{M2': } \nabla Y_t = \underbrace{(1 - B)}_{\nabla} A_t = \nabla A_t. \\ \text{M2': } Y_t - C &= A_t. \\ \text{C en M2' porque } \text{M2'} &\Leftrightarrow \nabla Y_t - \underbrace{\nabla C}_{=0} = \nabla A_t \end{aligned}$$

Aplicación: Inclusión de un factor MA(1) como “testigo de sobrediferenciación”:

$$\begin{aligned} \text{M3: } (1 - \phi_1 B)(Y_t - C) &= A_t, \\ \text{M3': } (1 - \phi_1 B)\nabla Y_t &= (1 - \theta_1 B)A_t. \\ \theta_1 = 1 \text{ en M3'} &\Rightarrow \text{M3' no invertible} \Rightarrow \text{Volver a M3,} \\ |\theta_1| < 1 \text{ en M3'} &\Rightarrow \text{M3' invertible} \Rightarrow \text{M3' adecuado.} \end{aligned}$$

[*]



CRITERIOS DE INFORMACIÓN

Para un modelo que contiene m parámetros y ha sido estimado con n observaciones, el criterio de información de Akaike (AIC) y el criterio de información bayesiano (BIC) o de Schwarz son las cantidades

$$\begin{aligned} \text{AIC} &= -\frac{2\hat{l}}{n} + \frac{2m}{n} \approx \ln(\hat{\sigma}_A^2) + \frac{2m}{n}, \\ \text{BIC} &= -\frac{2\hat{l}}{n} + \frac{m \ln(n)}{n} \approx \ln(\hat{\sigma}_A^2) + \frac{m \ln(n)}{n}, \end{aligned}$$

donde \hat{l} es el logaritmo de la función de verosimilitud asociada con el modelo estimado (Observación 43), evaluada en las estimaciones obtenidas de los parámetros de dicho modelo.

Los criterios de información suelen emplearse para comparar entre sí diferentes modelos alternativos para una transformación (estacionaria) dada de una serie y_t . Dichos criterios **no** deben utilizarse para discriminar ni entre diferentes transformaciones alternativas de y_t , ni entre modelos estimados con diferentes tamaños muestrales efectivos (que pueden variar en función del método de estimación empleado).

Teniendo en cuenta lo anterior, un modelo es tanto más adecuado cuanto más **reducidos** son los valores de sus AIC y BIC. Ambos criterios expresan un compromiso entre la obtención de un buen ajuste (a través del primer sumando de cada uno de ellos) y la obtención de un modelo con un número reducido de parámetros (a través del segundo sumando). En este sentido, BIC suele seleccionar modelos más **escuetos** que AIC. Con muestras grandes, el empleo de BIC suele ser preferible al empleo de AIC, aunque AIC puede ser preferible con algunas muestras cortas. En todo caso, la selección de un modelo entre varios alternativos no debería basarse únicamente en el empleo rutinario de estos dos criterios. ■

EJEMPLOS DE DIAGNOSIS

1. ST02: Serie TVPIB [incluyendo testigo MA(1) de sobrediferenciación].
2. ST14: Serie VIVIN.
3. ST19: Serie AIRLINE.
4. ST20: Serie TPARO.

IEV41: pp. 100-105

PREVISIÓN

[*]

IEV41: pp. 105-107

ANÁLISIS DE INTERVENCIÓN

[*]

IEV41: pp. 108-114