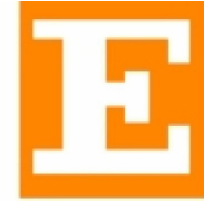




UNIVERSIDAD COMPLUTENSE  
MADRID

## **ECONOMETRÍA II**

<http://www.ucm.es/info/ecocuan/ectr2>



Facultad de Ciencias  
Económicas y Empresariales

### **TEMA 1**

## **ANÁLISIS UNIVARIANTE DE SERIES TEMPORALES**

**José Alberto Mauricio**

**Departamento de Economía Cuantitativa**

<http://www.ucm.es/info/ecocuan>

**COPYRIGHT © 2005 - 2007 José Alberto Mauricio**

<http://www.ucm.es/info/ecocuan/jam>

Este documento puede utilizarse exclusivamente como instrumento para la docencia oficial de la asignatura

ECONOMETRÍA II

que se imparte en la Facultad de Ciencias Económicas de la Universidad Complutense de Madrid. No se permite almacenar, reproducir o distribuir por medio alguno, ni tampoco utilizar este documento en cualquier sentido, fuera de los términos mencionados anteriormente. La obtención de este documento (Ectr2-JAM-Tema1.pdf) en la dirección de Internet


<http://www.ucm.es/info/ecocuan/ectr2>

implica la aceptación de que su uso estará limitado a los términos anteriores.

**Última revisión: 16 de febrero de 2007.**

## CONTENIDO









<b>1.1</b>	<b>Introducción</b>	<b>1</b>
	Series univariantes .....	2
	Series estacionarias.....	8
	Series no estacionarias.....	10
	Series estacionales .....	11
	Series simuladas a partir de modelos ARIMA .....	16
<b>1.2</b>	<b>Procesos Estocásticos Estacionarios</b>	<b>20</b>
➔	<b>Estacionariedad .....</b>	<b>21</b>
➔	<b>ACF – PACF teóricas .....</b>	<b>23</b>
	Autocovarianzas .....	23
	Autocorrelaciones simples .....	24
	Autocorrelaciones parciales.....	27










➔	<b>ACF – PACF muestrales .....</b>	<b>28</b>
	Series estacionarias simuladas .....	32
	Series estacionarias reales.....	33
	Series no estacionarias.....	34
➔	<b>Procesos estacionarios y series no estacionarias.....</b>	<b>35</b>
	Práctica EV-01 .....	36

**1.3 Modelos ARMA 38**

---

➔	<b>Estacionariedad – Invertibilidad.....</b>	<b>40</b>
➔	<b>ACF – PACF teóricas en modelos ARMA .....</b>	<b>43</b>
	AR(1).....	43
	AR(2).....	44
	MA(1) .....	46
	MA(2) .....	48
	ARMA(1,1) .....	49
➔	<b>Tabla – Resumen .....</b>	<b>54</b>

	Práctica EV-02 .....	55
<b>1.4</b>	<b>Procesos Estocásticos No Estacionarios</b> .....	<b>58</b>
	<b>No estacionariedad en varianza</b> .....	<b>59</b>
	Transformación de Box - Cox .....	60
	Práctica EV-03 .....	65
	<b>No estacionariedad en media</b> .....	<b>65</b>
	Tendencias – Diferenciación regular .....	73
	Práctica EV-04 .....	75
<b>1.5</b>	<b>Modelos ARIMA</b> .....	<b>76</b>
	<b>Tendencias en modelos invertibles</b> .....	<b>77</b>
	<b>Tendencias en modelos no invertibles</b> .....	<b>79</b>
<b>1.6</b>	<b>Modelos para Procesos Estacionales</b> .....	<b>80</b>
	<b>Series temporales estacionales</b> .....	<b>80</b>

	Diferenciación estacional .....	81
	Práctica EV-05 .....	84
	<b>Modelos ARIMA multiplicativos .....</b>	<b>84</b>
<b>1.7</b>	<b>El Enfoque Box - Jenkins para la Elaboración de Modelos ARIMA</b>	<b>88</b>
	<b>Identificación .....</b>	<b>89</b>
	Práctica EV-06 .....	91
	<b>Estimación .....</b>	<b>92</b>
	Práctica EV-07 .....	98
	<b>Diagnosis .....</b>	<b>98</b>
	Criterios de información .....	102
	Práctica EV-08 .....	103
<b>1.8</b>	<b>Previsión con Modelos ARIMA</b>	<b>104</b>
	<b>La función de previsión.....</b>	<b>105</b>
	ARMA(1,1) .....	107

	AR(1).....	109
	MA(1) .....	111
	IMA(1,1) .....	112
➔	<b>Previsiones puntuales – Intervalos de confianza.....</b>	<b>114</b>
📁	Práctica EV-09 .....	118
➔	<b>Criterios de evaluación de previsiones .....</b>	<b>119</b>
📁	Práctica EV-10 .....	122

## **1.9 Algunos Casos Prácticos** **X**

---

📁	<b>PIB .....</b>	<b>X</b>
📁	<b>Flujo anual del río Nilo .....</b>	<b>X</b>
📁	<b>Tasa de paro .....</b>	<b>X</b>
📁	<b>Temperatura media en Madrid-Retiro .....</b>	<b>X</b>
📁	<b>Viviendas iniciadas.....</b>	<b>X</b>
📁	<b>...</b>	

## BIBLIOGRAFÍA

**Manual: Peña (1989) capítulo 15.**

**Bibliografía adicional:**

**Box, Jenkins y Reinsel (1994) capítulos 1 - 9.**

**Heij, de Boer, Franses, Kloek, van Dijk (2004) capítulos 7.1 – 7.3.**

**Peña (2005) capítulos 1 - 11.**

**Verbeek (2004) capítulo 8.**

## 1.1 INTRODUCCIÓN

---

### Objetivos del análisis univariante de series temporales:

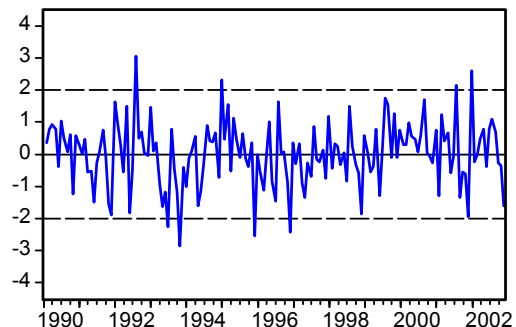
- ⇒ **Describir** la evolución temporal observada de alguna característica de interés de una unidad observable determinada.
- ⇒ **Prever** la evolución futura (desconocida) a corto plazo de dicha característica, utilizando tan sólo información (conocida) sobre su pasado. Esta información constituye lo que se conoce como una **serie temporal univariante**.

### Serie temporal univariante:

Secuencia de  $N$  observaciones (datos), ordenadas y equidistantes cronológicamente, sobre una única característica de una unidad observable en diferentes momentos.

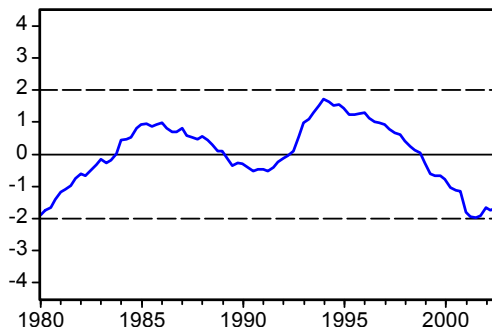
**Representaciones:**  $y_1, y_2, \dots, y_N$ ;  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$ ;  $(y_t)_{t=1}^N$ ;  $(y_t : t = 1, \dots, N)$ , donde  $y_t$  es la observación número  $t$  ( $1 \leq t \leq N$ ) de la serie y  $N$  es el número de observaciones de que consta la serie completa (el tamaño o la longitud de la serie).

VARIACIÓN MENSUAL DE TIPOS DE INTERÉS PARA PRÉSTAMOS PERSONALES - MENSUAL - PUNTOS PORCENTUALES - TS27



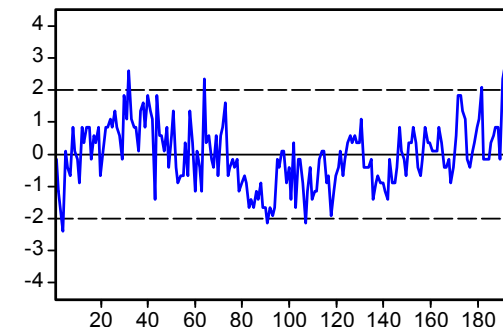
N = 155 - MEDIA = -0.08 - DT = 0.35

TASA DE PARO TRIMESTRAL - PORCENTAJE - TS31



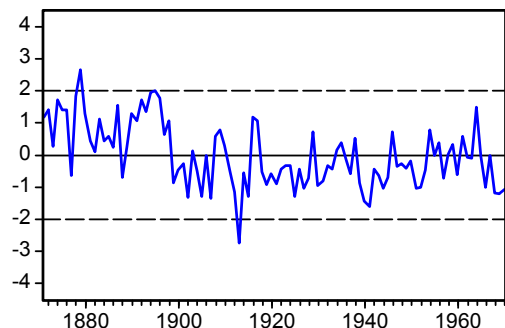
N = 92 - MEDIA = 17.89 - DT = 3.86

PROCESO QUÍMICO NO CONTROLADO - CONCENTRACIÓN LECTURAS CADA DOS HORAS - TS10 ( BJR A )



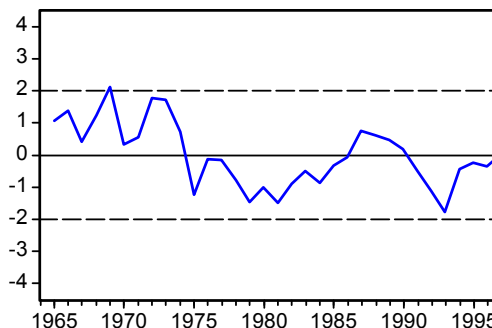
N = 197 - MEDIA = 17.06 - DT = 0.40

VOLUMEN DEL FLUJO DEL RÍO NILO EN ASUÁN ANUAL - MILES DE HECTÓMETROS CÚBICOS - TS01



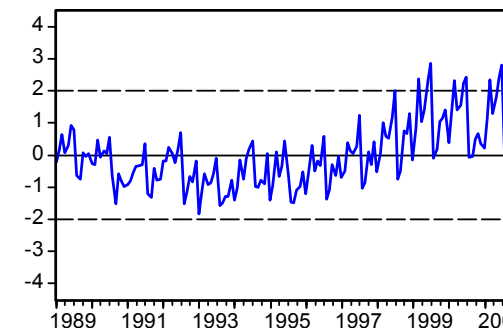
N = 100 - MEDIA = 91.94 - DT = 16.92

TASA LOGARÍTMICA DE VARIACIÓN ANUAL DEL PIB REAL ANUAL - PORCENTAJE - TS29



N = 33 - MEDIA = 3.36 - DT = 2.39

MATRICULACIONES ORDINARIAS DE TURISMOS MENSUAL - UNIDADES - TS30



N = 156 - MEDIA = 92788 - DT = 27006

Figura 1.1 – Ejemplos de series temporales univariantes

## Hipótesis de trabajo:

Una **serie temporal** (univariante) es una **realización particular** de una **muestra** procedente de un **proceso estocástico** (univariante).

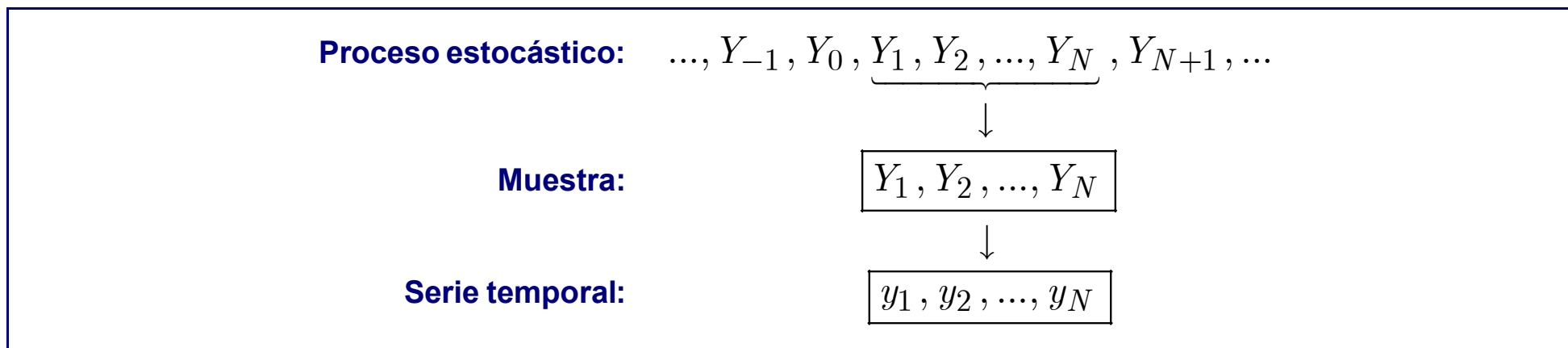
⇒ **Proceso estocástico:** Secuencia de variables aleatorias (v.a.s), ordenadas y equidistantes cronológicamente, referidas a una única característica de una unidad observable en diferentes momentos o fechas.

**Representaciones:**  $(Y_t : t = 0, \pm 1, \pm 2, \dots)$ ;  $(Y_t)$ , donde  $Y_t$  es una variable aleatoria (v.a.) referida a la característica de la unidad observable considerada en el momento  $t$ .

⇒ **Muestra:** Una muestra de tamaño o longitud  $N$  (finito) procedente de un proceso estocástico  $(Y_t)$  es una secuencia de  $N$  componentes **consecutivos** de  $(Y_t)$ .

**Representaciones:**  $Y_1, Y_2, \dots, Y_N$ ;  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$ ;  $(Y_t)_{t=1}^N$ ;  $(Y_t : t = 1, \dots, N)$ , donde  $Y_t$  es el componente o la observación número  $t$  ( $1 \leq t \leq N$ ) de la muestra y  $N$  es el tamaño o la longitud de la muestra completa.

**Observación 1:** El diagrama de la Figura 1.2 de la página siguiente representa la hipótesis de trabajo anterior.



**Figura 1.2 – Procedencia de una serie temporal**

**Observación 2:** Aunque es posible concebir ciertos experimentos controlados (repetibles varias veces bajo idénticas condiciones de partida) que permitan obtener diferentes realizaciones particulares finitas de un proceso estocástico, en muchas ocasiones es *imposible, prohibitivamente costoso o moralmente inaceptable* ejecutar dichos experimentos. La imposibilidad de controlar las condiciones a partir de las que se desarrollan las actividades en las que están implicadas las unidades observables a las que se refieren, es lo que hace que muchos procesos estocásticos sólo puedan observarse una única vez.

### Objetivo 1 del análisis univariante de series temporales:

Elaborar un **modelo estadístico** para una muestra  $\mathbf{Y}$  procedente de un proceso estocástico  $(Y_t)$  a partir de **una única** realización particular de  $\mathbf{Y}$  (una serie temporal  $\mathbf{y}$ ).

## Objetivo 2 del análisis univariante de series temporales:

Utilizar el modelo elaborado para describir en términos probabilísticos algunos componentes futuros  $Y_{N+l}$  ( $l \geq 1$ ) del proceso  $(Y_t)$ , es decir, **prever**  $Y_{N+1}, Y_{N+2}, \dots$

**Observación 3:** En general, una serie temporal se refiere a un **período muestral** que tan sólo es una parte de la historia del proceso estocástico del que procede dicha serie; ver Figura 1.2. No obstante, si las circunstancias sociales o naturales del período muestral al que se refiere la serie considerada se mantienen relativamente estables después de dicho período, entonces se espera que las conclusiones obtenidas del análisis de dicha serie sean aplicables también a momentos posteriores, al menos a corto plazo. Esta idea justifica el empleo de un modelo elaborado con una muestra dada para describir la evolución temporal de un proceso estocástico después del período muestral considerado.

## Otros objetivos del análisis univariante de series temporales:

- ⇒ **Contrastar** (presentar evidencia empírica en favor o en contra de) alguna teoría sobre la característica o variable a la que se refiere la serie considerada.
- ⇒ Servir como **punto de partida** para analizar posibles **relaciones** entre la serie considerada y otras series, mediante la elaboración de un modelo multivariante o vectorial.

### **Método de trabajo:**

Seleccionar un modelo para la muestra  $\mathbf{Y}$  dentro de una clase general de modelos de reconocida utilidad en la práctica (**modelos ARIMA**), que implique ciertas propiedades **teóricas** para el proceso estocástico  $(Y_t)$  del que procede  $\mathbf{Y}$  que resulten compatibles con las propiedades **muestrales** observadas en la serie temporal  $\mathbf{y}$ .

### **Inconveniente 1:**

Para elaborar un modelo para  $\mathbf{Y}$  a partir de una única serie temporal  $\mathbf{y}$ , hay que suponer ciertas simplificaciones sobre la estructura probabilística del proceso  $(Y_t)$  del que procede  $\mathbf{Y}$ , de manera que lo que no se suponga se pueda inferir a partir de una única realización particular (una única serie temporal) de  $\mathbf{Y}$ .

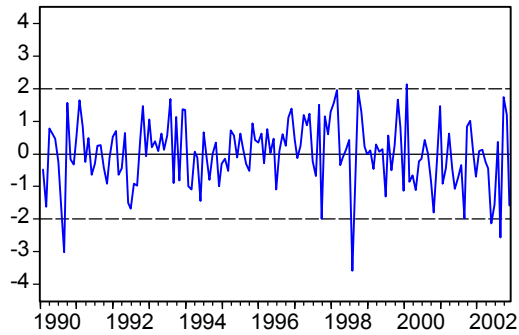
### **Solución – Hipótesis de estacionariedad:**

Algunas propiedades teóricas de un proceso estocástico implicadas por la hipótesis de estacionariedad son las siguientes:

- ⇒ **Estacionariedad:** Cierta estado de **equilibrio** estadístico que caracteriza la evolución temporal de un proceso estocástico que ha generado una serie temporal.
- ⇒ **Media:** Valor **constante** en el tiempo que mide el **nivel** alrededor del cual evoluciona un proceso estocástico estacionario.
- ⇒ **Varianza:** Valor **constante** en el tiempo que mide la **dispersión** o la **variabilidad** de la evolución temporal de un proceso estocástico estacionario alrededor de su media.
- ⇒ **Autocorrelaciones:** Valores **constantes** en el tiempo que miden el grado de **asociación lineal** entre cada par de componentes de un proceso estocástico estacionario separados por diferentes intervalos temporales o **retardos**.

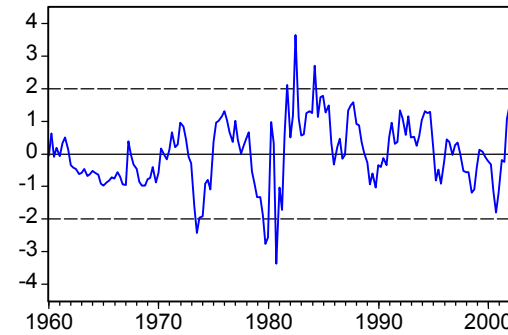
**Observación 4:** Aunque en la Sección 1.2 se dan algunas definiciones formales del concepto de estacionariedad, las consideraciones anteriores ponen de manifiesto que cuando un proceso estocástico es estacionario, sus propiedades estadísticas se simplifican notablemente con respecto a las de un proceso que no sea estacionario, lo cual facilita la descripción de su estructura probabilística completa a partir de una única realización finita del mismo. La Figura 1.3 de la página siguiente contiene los gráficos estandarizados de cuatro series temporales que podrían considerarse **estacionarias**, es decir, series que podrían resultar compatibles con la hipótesis de haber sido generadas por procesos estocásticos estacionarios.

TASA LOGARÍTMICA DE VARIACIÓN MENSUAL DEL IBEX35  
MENSUAL - PORCENTAJE - TS27



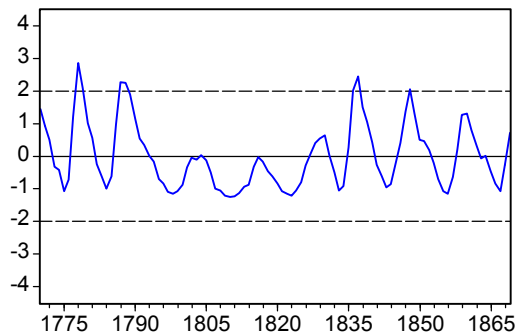
N = 155 - MEDIA = 0.49 - DT = 6.80

DIFERENCIAL DE LA DEUDA EN EEUU A 36 Y A 3 MESES  
TRIMESTRAL - PUNTOS PORCENTUALES - TS37



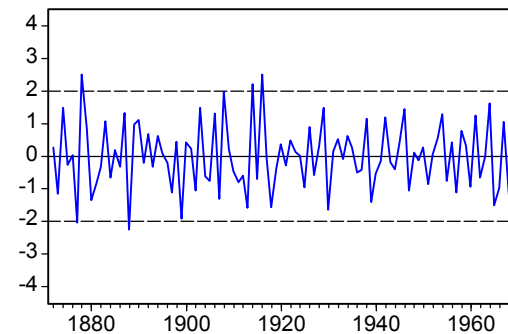
N = 172 - MEDIA = 1.02 - DT = 0.85

MANCHAS SOLARES (WÖLFER SUNSPOT NUMBERS)  
NÚMERO MEDIO ANUAL - TS14 (BJR E)



N = 100 - MEDIA = 46.93 - DT = 37.37

VARIACIÓN ANUAL DEL FLUJO DEL RÍO NILO EN ASUÁN  
ANUAL - HECTÓMETROS CÚBICOS - TS01



N = 99 - MEDIA = -383.83 - DT = 16813.19

**Figura 1.3 – Ejemplos de series temporales estacionarias**

## Inconveniente 2:

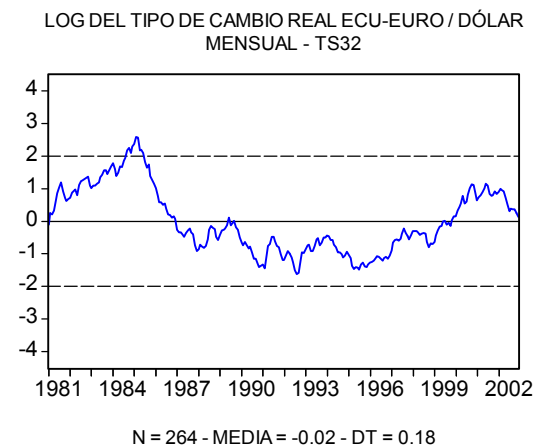
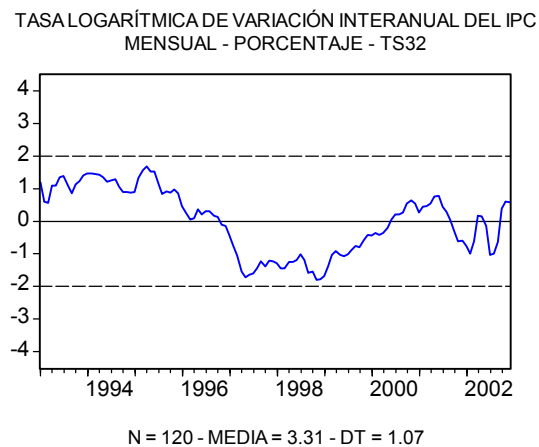
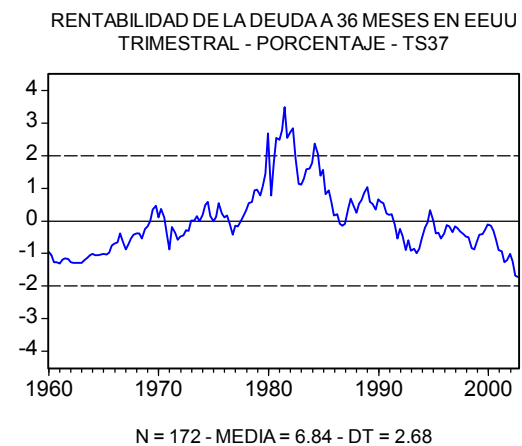
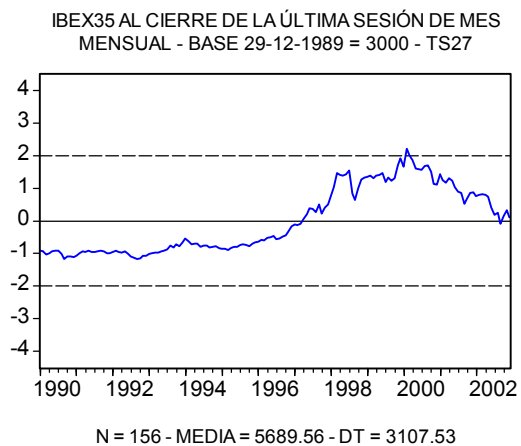
Muchas series temporales no pueden considerarse generadas por procesos estocásticos estacionarios (es decir, son series **no estacionarias**), porque presentan ciertas **tendencias** claras en su evolución temporal (de manera que no presentan afinidad hacia algún valor constante en el tiempo), porque su variabilidad no es constante, o porque son **estacionales**.

**Observación 5:** Las Figuras 1.4 y 1.5 de las dos páginas siguientes contienen los gráficos estandarizados de ocho series temporales que son claramente **no estacionarias**, es decir, series que **no** resultan compatibles con la hipótesis de haber sido generadas por procesos estocásticos estacionarios.

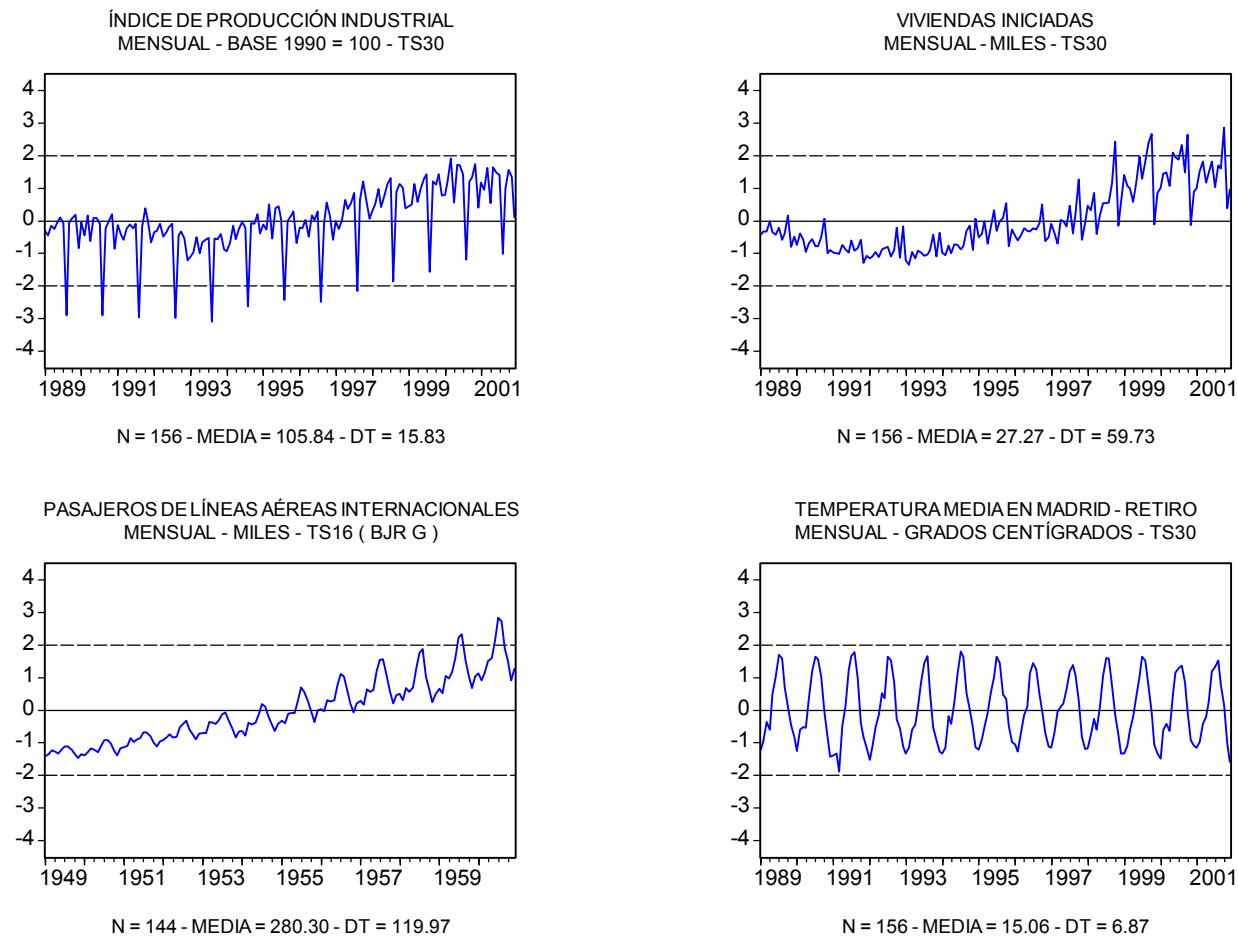
## Solución – Transformaciones de las series originales:

En muchas ocasiones, una serie temporal no estacionaria puede transformarse de forma adecuada para obtener una serie temporal de aspecto estacionario. La serie transformada puede utilizarse entonces para elaborar un modelo estadístico que resulte compatible con la hipótesis de estacionariedad.

**Observación 6:** Las transformaciones más utilizadas en la práctica para obtener series temporales estacionarias son la **transformación logarítmica**, la **diferenciación regular** y la **diferenciación estacional**.



**Figura 1.4 – Ejemplos de series temporales no estacionarias**



**Figura 1.5 – Ejemplos de series temporales estacionales**

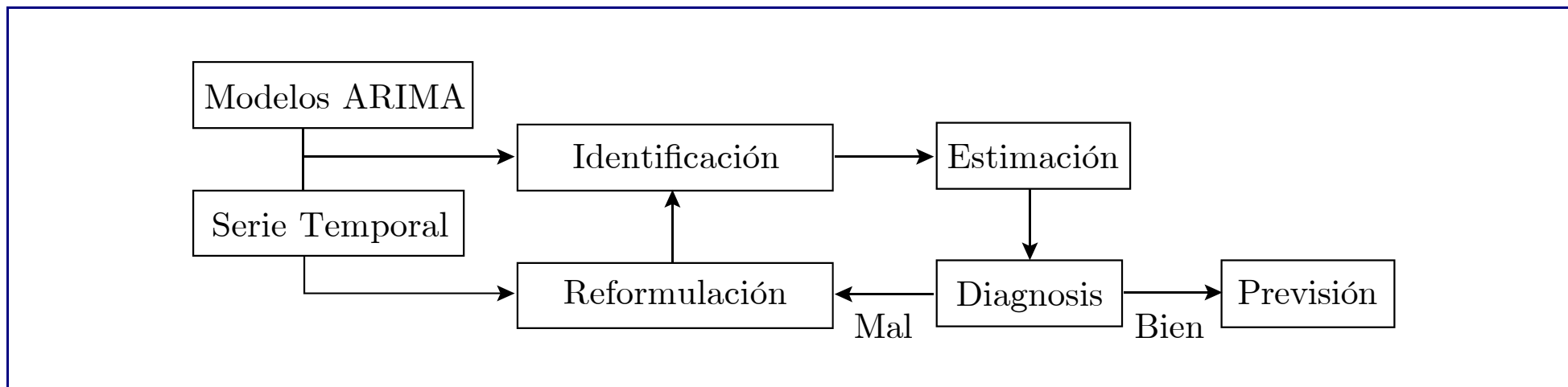
## Etapas en la construcción de un modelo ARIMA:

- ⇒ **Identificación:** Selección de un modelo que implique ciertas propiedades teóricas compatibles con las propiedades muestrales correspondientes observadas en  $\mathbf{y}$ .
- ⇒ **Estimación:** Asignación de valores numéricos a los parámetros del modelo.
- ⇒ **Diagnosis:** Comprobación del ajuste del modelo a los datos utilizados.

**Observación 7:** La construcción en la práctica de un modelo ARIMA a partir de una serie temporal observada, es un proceso **iterativo** que está descrito en el diagrama de la Figura 1.6 de la página siguiente.

## Previsión:

Utilización de un modelo adecuado para estimar la media y la varianza de la distribución de probabilidad de  $Y_{N+l}$  ( $l \geq 1$ ) **condicionada** por toda la información disponible hasta el momento  $N$ . La media de dicha distribución,  $E_N[Y_{N+l}]$  ( $l \geq 1$ ), se denomina **función de previsión en origen  $N$  a horizonte  $l$** , y se utiliza para calcular **previsiones puntuales** de los valores futuros  $Y_{N+1}, Y_{N+2}, \dots$ :  $\hat{y}_N(1) \equiv \hat{E}_N[Y_{N+1}]$ ,  $\hat{y}_N(2) \equiv \hat{E}_N[Y_{N+2}]$ ,  $\dots$ .



**Figura 1.6 – Proceso de construcción de un modelo ARIMA**

**Observación 8:** Un modelo ARIMA para  $(Y_t)$  se plantea mediante una expresión matemática concreta (una **ecuación en diferencias**) para  $Y_t$  (el componente genérico del proceso considerado), que se espera implique unas **propiedades teóricas** sobre los momentos de primer y segundo orden (medias, varianzas y covarianzas) de las distribuciones conjuntas de los componentes de  $(Y_t)$  que sean compatibles con las **propiedades muestrales** correspondientes observadas en una serie temporal  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$ . Cuando dichas distribuciones conjuntas son Normales, sus propiedades de primer y segundo orden caracterizan completamente la estructura probabilística de  $(Y_t)$ .

### Ejemplo 1 – Modelo ARIMA(0,0,0), o Ruido Blanco (*White Noise*)

Un proceso de **ruido blanco** univariante es una secuencia  $(A_t)$  de variables aleatorias

escalares idéntica e independientemente distribuidas con media 0 y varianza  $\sigma_A^2$ , lo cual suele representarse como  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ . Cuando cada  $A_t$  sigue una distribución Normal,  $(A_t)$  se denomina un proceso de ruido blanco **Normal** o **Gaussiano**, lo cual suele representarse como  $(A_t) \sim \text{NIID}(0, \sigma_A^2)$ .

### Ejemplo 2 – Modelo ARIMA(1,0,0), o AR(1)

Un proceso estocástico univariante estacionario  $(Y_t)$  sigue un modelo **AR(1)** (autorregresivo de orden 1, del inglés *AutoRegressive*) cuando

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t \quad \text{para todo } t = 0, \pm 1, \pm 2, \dots, \quad (1.1)$$

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

### Ejemplo 3 – Modelo ARIMA(2,0,0), o AR(2)

Un proceso estocástico univariante estacionario  $(Y_t)$  sigue un modelo **AR(2)** (autorregresivo de orden 2) cuando

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + A_t \quad \text{para todo } t = 0, \pm 1, \pm 2, \dots, \quad (1.2)$$

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

#### Ejemplo 4 – Modelo ARIMA(0,0,1), o MA(1)

Un proceso estocástico univariante estacionario  $(Y_t)$  sigue un modelo **MA(1)** (media móvil de orden 1, del inglés *Moving Average*) cuando

$$Y_t = \mu + A_t - \theta_1 A_{t-1} \quad \text{para todo } t = 0, \pm 1, \pm 2, \dots, \quad (1.3)$$

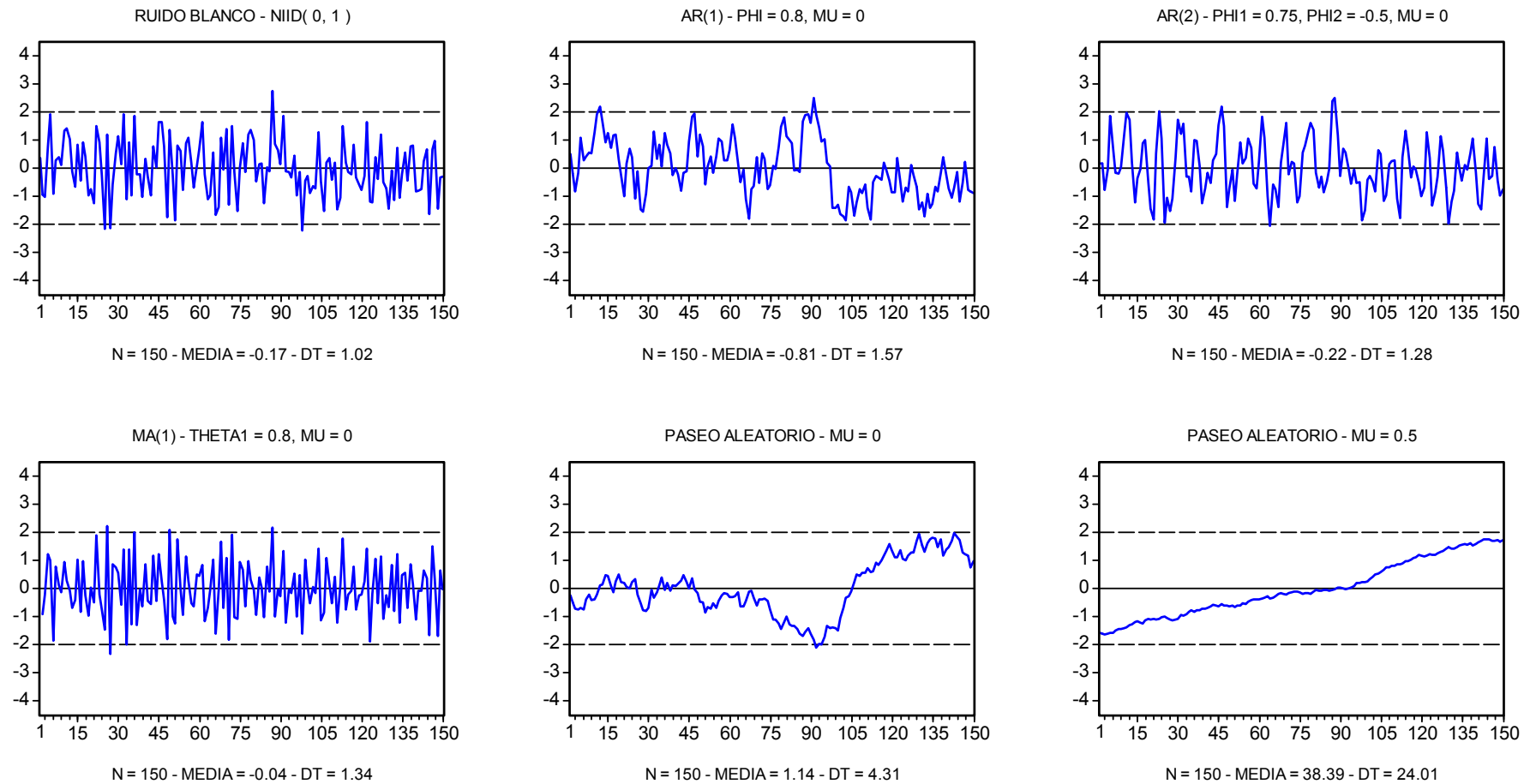
donde  $\mu$  y  $\theta_1$  son parámetros,  $|\theta_1| < 1$  (**condición de invertibilidad**), y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ .

#### Ejemplo 5 – Modelo ARIMA(0,1,0), o Paseo Aleatorio (*Random Walk*)

Un proceso estocástico univariante **no** estacionario  $(Y_t)$  es un **paseo aleatorio** cuando

$$Y_t = \mu + Y_{t-1} + A_t, \quad (1.4)$$

donde  $\mu$  es un parámetro (que en muchas ocasiones vale cero) y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ .



**Figura 1.7 – Series simuladas a partir de varios modelos ARIMA**

**Observación 9:** 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  $1 - \phi_1 x = 0$  asociada con un modelo AR(1) tiene una **raíz unitaria** ( $x^* = 1$ ).] Por otro lado, un paseo aleatorio puede escribirse como  $\nabla Y_t = \mu + A_t$ , donde  $\nabla Y_t \equiv Y_t - Y_{t-1}$ , de manera que un proceso estocástico univariante no estacionario ( $Y_t$ ) es un paseo aleatorio cuando su **diferencia regular de orden 1** ( $\nabla Y_t \equiv (Y_t - Y_{t-1})$ ) es un proceso estacionario IID( $\mu, \sigma_A^2$ ) [donde el parámetro  $\mu$  puede valer cero, en cuyo caso el proceso ( $\nabla Y_t$ ) sería ruido blanco].

**Observación 10:** Un paseo aleatorio constituye el caso más sencillo de lo que suele denominarse un **proceso integrado de orden 1**, ó **I(1)** (es decir, un proceso no estacionario cuya diferencia regular de orden 1 es un proceso estacionario); por extensión de esta terminología, un proceso estacionario suele denominarse un proceso **I(0)**. La notación  $(Y_t) \sim I(1)$  significa que  $(Y_t)$  es un proceso I(1) (integrado de orden 1), mientras que la notación  $(Y_t) \sim I(0)$  significa que  $(Y_t)$  es un proceso estacionario (integrado de orden 0).

**Observación 11:** Las seis series de la Figura 1.7 han sido simuladas (generadas artificialmente) a partir de la ecuación que define el modelo correspondiente, utilizando en cada caso una secuencia de 150 realizaciones independientes de una distribución N(0,1) generadas por el programa EViews. Estas series se encuentran en el archivo TS39-Sim-ARIMA.wf1, donde se describe explícitamente la ecuación empleada para simular cada serie.

## Ejemplo 6 – Previsión con modelos ARIMA

Las **funciones de previsión** en **origen**  $N$  a **horizonte**  $l \geq 1$  asociadas con algunos de los modelos ARIMA considerados en los ejemplos anteriores son las siguientes:

Ruido blanco:  $E_N[A_{N+l}] = E[A_{N+l}] = 0$  para todo  $l \geq 1$ .

AR(1):  $E_N[Y_{N+l}] = \mu + \phi_1 E_N[Y_{N+l-1}]$   
 $= \left( \sum_{i=0}^{l-1} \phi_1^i \right) \mu + \phi_1^l Y_N$  para todo  $l \geq 1$ .

MA(1):  $E_N[Y_{N+l}] = \begin{cases} \mu - \theta_1 E_N[A_N], & l = 1, \\ \mu & , l > 1. \end{cases}$

Paseo aleatorio:  $E_N[Y_{N+l}] = \mu + E_N[Y_{N+l-1}]$  para todo  $l \geq 1$ .  
 $= l\mu + Y_N$

### Ejemplo 7 – El tipo de cambio y la paridad del poder adquisitivo

Sea  $E_t$  el **tipo de cambio nominal** de la moneda de un país A con respecto a la moneda de un país B (por ejemplo,  $E_t$  podría ser el precio en euros de un dólar de EEUU, medido en euros/dólar). Si  $PA_t$  y  $PB_t$  son dos **índices de precios** referidos a los países A y B, respectivamente, la teoría de la **paridad del poder adquisitivo (PPA)** (en inglés *purchasing*

*power parity*) establece la siguiente relación de equilibrio entre  $E_t$  y  $PA_t / PB_t$ :

$$E^* = \gamma \frac{PA^*}{PB^*}, \quad (1.5)$$

donde  $\gamma \neq 0$  puede ser igual a 1 (PPA **absoluta**) o distinto de 1 (PPA **relativa**, que permite ciertas incompatibilidades entre los índices  $P_t$  y  $P_t^*$ ). Nótese que (1.5) implica que

$$\ln E^* = \beta_0 + \ln \frac{PA^*}{PB^*} \quad (\beta_0 \equiv \ln \gamma), \quad (1.6)$$

de manera que, de acuerdo con la PPA, la **elasticidad a largo plazo** del tipo de cambio nominal con respecto a los precios relativos debe ser unitaria. Alternativamente, (1.6) puede escribirse como

$$\ln RE^* = \beta_0, \quad (1.7)$$

donde  $RE^* \equiv E^* (PB^* / PA^*)$  es el **tipo de cambio real** a largo plazo. Por lo tanto, la PPA enunciada en (1.6) puede verificarse empíricamente contrastando si el proceso  $(\ln RE_t)$  es estacionario, donde  $RE_t \equiv E_t \times (PB_t / PA_t)$ ; ver Figura 1.4.

## 1.2 PROCESOS ESTOCÁSTICOS ESTACIONARIOS

---

⇒ Elaborar un modelo estadístico para una muestra  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$  exige al menos estimar el vector de medias y la matriz de varianzas-covarianzas,

$$\boldsymbol{\mu} \equiv E[\mathbf{Y}] = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix}, \quad \boldsymbol{\Sigma} \equiv E[(\mathbf{Y} - \boldsymbol{\mu})(\mathbf{Y} - \boldsymbol{\mu})'] = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{bmatrix},$$

de la distribución de probabilidad de  $\mathbf{Y}$ . En conjunto,  $\boldsymbol{\mu}$  y  $\boldsymbol{\Sigma}$  contienen  $N + N(N + 1) / 2$  parámetros distintos, que **no** pueden estimarse con precisión utilizando tan sólo una única realización particular de  $\mathbf{Y}$  (una única serie temporal de  $N$  observaciones).

⇒ Elaborar un modelo para  $\mathbf{Y}$  exige suponer alguna simplificación sobre su distribución de probabilidad que reduzca el número de parámetros distintos contenidos en  $\boldsymbol{\mu}$  y en  $\boldsymbol{\Sigma}$ .

## Estacionariedad

En algunos casos, las características muestrales de una serie temporal (ver Figura 1.3) sugieren que el proceso estocástico del que procede dicha serie se encuentra en cierto estado de **equilibrio estadístico** llamado **estacionariedad**.

### 1.2.1 Definición

Un proceso estocástico  $(Y_t)$  es **estrictamente estacionario** si y sólo si para cualesquiera  $n \geq 1$  momentos  $t_1 < t_2 < \dots < t_n$  de su historia, la distribución de probabilidad conjunta de  $[Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}]'$  coincide con la de  $[Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_n+h}]'$  para cualquier número entero  $h = \pm 1, \pm 2, \dots$  (distinto de cero).

### 1.2.2 Definición

Un proceso estocástico  $(Y_t)$  con  $E[Y_t] < \infty$  para todo  $t = 0, \pm 1, \pm 2, \dots$  es **estacionario en media** o **débilmente estacionario de primer orden** si y sólo si  $E[Y_t]$  es constante (no depende de  $t$ ) para todo  $t = 0, \pm 1, \pm 2, \dots$

### 1.2.3 Definición

Un proceso estocástico  $(Y_t)$  con  $E[Y_t^2] < \infty$  para todo  $t = 0, \pm 1, \pm 2, \dots$  es **estacionario en autocovarianza** o **débilmente estacionario de segundo orden** si y sólo si:

- $E[Y_t]$  y  $\text{Var}[Y_t]$  son constantes (no dependen de  $t$ ) para todo  $t = 0, \pm 1, \pm 2, \dots$ ,
- $\text{Cov}[Y_t, Y_{t+k}]$  depende a lo sumo de  $k$  (entero) pero no de  $t$  para todo  $t = 0, \pm 1, \pm 2, \dots$

### 1.2.4 Definición

Un proceso estocástico  $(Y_t)$  es **Normal** o **Gaussiano** cuando para cualesquiera  $n \geq 1$  momentos  $t_1 < t_2 < \dots < t_n$  de su historia, la distribución de probabilidad conjunta de  $[Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}]'$  es una distribución Normal  $n$ -variante.

**Estacionariedad en autocovarianza + Normalidad  $\Rightarrow$  Estacionariedad estricta**

**Figura 2.1 – Estacionariedad**

**Observación:** En general, el tipo de estacionariedad definido en 1.2.3 es suficiente para todos los planteamientos que se abordan en esta asignatura. En adelante, cuando se hable de estacionariedad sin calificativos se entenderá que se está hablando de estacionariedad en autocovarianza. Cuando un proceso estocástico  $(Y_t)$  es estacionario, suele decirse que  $(Y_t)$  es un proceso **integrado de orden 0**, ó **I(0)**, lo que se representa como  $(Y_t) \sim I(0)$  (aunque esta terminología no es del todo correcta; ver Definición 1.4.5).

## ACF – PACF teóricas

### 1.2.5 Definición

La **autocovarianza de orden  $k$**  ( $k > 0$ ) de un proceso  $(Y_t)$  estacionario se representa con el símbolo  $\gamma_k$  y se define como  $\gamma_k \equiv \text{Cov}[Y_t, Y_{t+k}]$ .

**Media:**

$$\mu_Y \equiv E[Y_t].$$

**Varianza:**

$$\sigma_Y^2 \equiv \text{Var}[Y_t] \equiv E[(Y_t - \mu_Y)^2].$$

**Autocovarianza de orden  $k$ :**  $\gamma_k \equiv \text{Cov}[Y_t, Y_{t+k}] \equiv E[(Y_t - \mu_Y)(Y_{t+k} - \mu_Y)]$  ( $k = 1, 2, \dots$ ).

**Figura 2.2 – Media, varianza y autocovarianzas de un proceso estacionario**

**Observación 1:** La autocovarianza de orden  $k$  de  $(Y_t)$  es la covarianza entre cualquier par de componentes de  $(Y_t)$  separados entre sí por un intervalo temporal o **retardo**  $k > 0$  dado. Por 1.2.3,  $\gamma_k$  puede depender de  $k$ , pero no depende de los momentos concretos a los que se refieran los componentes de  $(Y_t)$  considerados.

**Observación 2:** Aunque  $\gamma_k$  se ha definido para cualquier  $k > 0$ , también puede definirse la autocovarianza de orden cero como  $\gamma_0 \equiv \text{Cov}[Y_t, Y_t] \equiv \text{Var}[Y_t]$ , que es la varianza del proceso estacionario  $(Y_t)$ . Por otro lado, la autocovarianza de orden  $k$  de  $(Y_t)$  también puede definirse como  $\gamma_k \equiv \text{Cov}[Y_{t-k}, Y_t]$ , ya que para cualquier valor de  $k \geq 0$  dado,  $Y_{t-k}$  e  $Y_t$  están separados entre sí por el mismo retardo que  $Y_t$  e  $Y_{t+k}$ . Considerada como una función del retardo  $k$ , la secuencia  $(\gamma_k : k = 0, 1, 2, \dots)$  se denomina **la función de autocovarianza** del proceso estacionario  $(Y_t)$ .

### 1.2.6 Definición

La **autocorrelación simple de orden  $k$**  ( $k > 0$ ) de un proceso  $(Y_t)$  estacionario se representa con el símbolo  $\rho_k$  y se define como

$$\rho_k \equiv \frac{\text{Cov}[Y_t, Y_{t+k}]}{\text{Var}[Y_t]^{1/2} \text{Var}[Y_{t+k}]^{1/2}} \equiv \frac{\gamma_k}{\gamma_0}. \quad (2.1)$$

**Autocorrelación simple de orden  $k$ :**  $\rho_k \equiv \frac{\gamma_k}{\gamma_0} \quad (k = 1, 2, \dots); \quad \rho_0 = 1.$

**Figura 2.3 – Autocorrelaciones simples de un proceso estacionario**

**Observación 1:** La autocorrelación simple de orden  $k$  de  $(Y_t)$  es el coeficiente de correlación lineal simple entre cualquier par de componentes de  $(Y_t)$  separados entre sí por un retardo  $k > 0$  dado. Por 1.2.3,  $\rho_k$  puede depender de  $k$ , pero no depende de los momentos concretos a los que se refieran los componentes de  $(Y_t)$  considerados.

**Observación 2:** Considerada como una función de  $k$ ,  $(\rho_k : k = 1, 2, \dots)$  se denomina **la función de autocorrelación simple (ACF, del inglés *AutoCorrelation Function*)** del proceso estacionario  $(Y_t)$ . Dado que cada  $\rho_k$  es un coeficiente de correlación, suele decirse que la ACF de  $(Y_t)$  representa la **duración** y la **intensidad** de la **memoria** del proceso  $(Y_t)$ .

**Observación 3:** Si una muestra  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$  procede de un proceso estacionario, entonces su vector de medias y su matriz de varianzas-covarianzas son

$$\boldsymbol{\mu} \equiv E[\mathbf{Y}] = \begin{bmatrix} \mu_Y \\ \mu_Y \\ \vdots \\ \mu_Y \end{bmatrix}, \quad \boldsymbol{\Sigma} \equiv \text{Var}[\mathbf{Y}] = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{N-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{N-1} & \gamma_{N-2} & \cdots & \gamma_0 \end{bmatrix} = \sigma_Y^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{N-1} \\ \rho_1 & 1 & \cdots & \rho_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N-1} & \rho_{N-2} & \cdots & 1 \end{bmatrix},$$

que sólo contienen  $1 + N$  parámetros distintos:  $\mu_Y, \gamma_0, \gamma_1, \dots, \gamma_{N-1}$ , o bien  $\mu_Y, \sigma_Y^2, \rho_1, \dots, \rho_{N-1}$ . Las dos matrices que figuran en la expresión anterior suelen representarse como  $\boldsymbol{\Gamma}_N$  (**matriz de autocovarianzas de  $\mathbf{Y}$** ) y  $\mathbf{P}_N$  (**matriz de autocorrelaciones de  $\mathbf{Y}$** ), de manera que la matriz de varianzas-covarianzas de la muestra  $\mathbf{Y}$  procedente de  $(Y_t)$  es  $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}_N = \sigma_Y^2 \mathbf{P}_N$ . A pesar de que la hipótesis de estacionariedad simplifica notablemente la distribución de probabilidad de  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$ , aún no pueden estimarse con precisión  $1 + N$  parámetros utilizando tan sólo una única realización particular de  $\mathbf{Y}$  (una única serie temporal  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$  de  $N$  observaciones). La solución a este problema consiste en expresar la media y la función de autocovarianza (o la media, la varianza y la ACF) de  $(Y_t)$  en términos de un número reducido de parámetros a través de algún **modelo ARMA** para  $(Y_t)$ ; ver Sección 1.3.

### 1.2.7 Definición

La **autocorrelación parcial de orden  $k$**  ( $k > 0$ ) de un proceso  $(Y_t)$  estacionario se representa con el símbolo  $\phi_{kk}$  y se define como el parámetro  $\phi_{kk}$  en la regresión

$$\tilde{Y}_t = \phi_{k1}\tilde{Y}_{t-1} + \phi_{k2}\tilde{Y}_{t-2} + \dots + \phi_{kk}\tilde{Y}_{t-k} + U_t, \quad (2.2)$$

donde  $\tilde{Y}_{t-i} \equiv Y_{t-i} - \mu_Y$  ( $i = 0, 1, \dots, k$ ) y  $U_t$  es independiente de  $Y_{t-i}$  para todo  $i \geq 1$ .

**Observación 1:** La regresión (2.2) anterior puede escribirse como  $Y_t = \phi_{k0} + \phi_{k1}Y_{t-1} + \phi_{k2}Y_{t-2} + \dots + \phi_{kk}Y_{t-k} + U_t$ , donde  $\phi_{k0} \equiv (1 - \phi_{k1} - \phi_{k2} - \dots - \phi_{kk})\mu_Y$ , lo que implica que  $\Delta Y_t = \phi_{kk} \times \Delta Y_{t-k}$  cuando  $\Delta Y_{t-1} = \Delta Y_{t-2} = \dots = \Delta Y_{t-k+1} = \Delta U_t = 0$ ; es decir,  $\phi_{kk}$  en (2.2) representa el **efecto parcial** o **ceteris paribus** de  $Y_{t-k}$  sobre  $Y_t$ . Por lo tanto,  $\phi_{kk}$  es una medida del grado de asociación lineal entre dos componentes cualesquiera de  $(Y_t)$  separados entre sí por un retardo  $k \geq 1$  dado (como  $Y_t$  e  $Y_{t-k}$ ), que **no** es debida a la posible correlación entre cada uno de ellos y todos los componentes de  $(Y_t)$  que se encuentran entre ambos ( $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$ ).

**Observación 2:** A partir de la regresión (2.2) anterior, puede comprobarse que  $\phi_{kk} = |\mathbf{A}_k| / |\mathbf{B}_k|$ , donde

$$\mathbf{A}_k \equiv \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{k-4} & \rho_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{bmatrix}, \quad \mathbf{B}_k \equiv \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{k-4} & \rho_{k-3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{bmatrix}.$$

La matriz  $\mathbf{B}_k$  es una matriz de autocorrelaciones del mismo tipo que  $\mathbf{P}_N$  en la Observación 3 de 1.2.6. La matriz  $\mathbf{A}_k$  es idéntica a  $\mathbf{B}_k$  excepto porque la última columna de  $\mathbf{B}_k$  está reemplazada por el vector columna  $[\rho_1, \rho_2, \dots, \rho_k]'$ . En particular, para  $k = 1, 2, 3$  resulta que

$$\phi_{11} = \rho_1, \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \quad \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1(\rho_1^2 + \rho_2^2 - 2\rho_2) + \rho_3(1 - \rho_1^2)}{(1 - \rho_2)(1 - 2\rho_1^2 + \rho_2)},$$

lo que ilustra que, en general, cada coeficiente de autocorrelación parcial  $\phi_{kk}$  de un proceso estrictamente estacionario es una función concreta de los coeficientes de autocorrelación simple  $\rho_1, \rho_2, \dots, \rho_k$  ( $k = 1, 2, 3, \dots$ ).

**Observación 3:** Considerada como una función del retardo  $k$ , la secuencia  $(\phi_{kk} : k = 1, 2, \dots)$  se denomina **la función de autocorrelación parcial (PACF, del inglés *Partial AutoCorrelation Function*)** del proceso  $(Y_t)$ .

$$\text{Autocorrelación parcial de orden } k: \quad \phi_{kk} = \frac{|\mathbf{A}_k|}{|\mathbf{B}_k|} \quad (k = 1, 2, \dots); \quad \phi_{11} = \rho_1.$$

**Figura 2.4 – Autocorrelaciones parciales de un proceso estacionario**

## ACF – PACF muestrales

- ⇒ En la práctica, es imposible estimar a partir de una serie temporal  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$  de  $N$  observaciones, la media, la varianza, la ACF y la PACF del proceso estacionario  $(Y_t)$  del que supuestamente procede dicha serie. En particular, sólo en el vector de medias y en la matriz de varianzas-covarianzas de la muestra  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$  asociada con la serie  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$  figuran  $1 + N$  parámetros distintos.
- ⇒ No obstante, en la práctica **sí** pueden estimarse con cierta precisión las  $K$  primeras autocorrelaciones simples y parciales de  $(Y_t)$  a partir de una serie temporal de  $N$  observaciones, siempre que  $K$  sea un número bastante más reducido que  $N$ .
- ⇒ Dichas estimaciones resultan muy útiles para identificar un modelo inicial para  $(Y_t)$  dentro de la clase modelos para procesos estacionarios que se describe en la Sección 1.3.

### 1.2.8 Definición

La **media muestral** y la **varianza muestral** de una muestra  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$  asociada con una serie temporal  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$  son

$$\hat{\mu}_Y \equiv \frac{1}{N} \sum_{t=1}^N Y_t \quad \text{y} \quad \hat{\sigma}_Y^2 \equiv \frac{1}{N} \sum_{t=1}^N (Y_t - \hat{\mu}_Y)^2, \quad (2.3)$$

respectivamente, que suelen emplearse como **estimadores** de la media  $\mu_Y$  y la varianza  $\sigma_Y^2$  del proceso estacionario  $(Y_t)$  del que procede  $\mathbf{Y}$ .

**Observación 1:** Los estimadores  $\hat{\mu}_Y$  y  $\hat{\sigma}_Y^2$  proporcionan **estimaciones** numéricas  $\bar{y}$  y  $s_y^2$ , respectivamente,

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

cuando se reemplaza en (2.3) cada variable aleatoria  $Y_t$  por su valor observado  $y_t$  ( $t = 1, \dots, N$ ).

**Observación 2:** Para contrastar hipótesis sobre  $\mu_Y$  (el verdadero valor de la media de un proceso estacionario) del estilo de  $\mu_Y = m$  frente a alternativas del tipo  $\mu_Y \neq m$ ,  $\mu_Y > m$  ó  $\mu_Y < m$  (donde  $m$  es un número dado, como 0), puede utilizarse un estadístico  $t$  basado en que

$$E[\hat{\mu}_Y] = \mu_Y, \quad \text{Var}[\hat{\mu}_Y] \cong \frac{\sigma_Y^2}{N} \left( 1 + 2 \sum_{k=1}^{\infty} \rho_k \right).$$

Por lo tanto, el estadístico  $t_{\mathbf{Y}} = (\hat{\mu}_Y - m) / (\hat{\sigma}_Y^2 / N)^{1/2}$ , cuyo valor calculado es  $t = (\bar{y} - m) / (s_y^2 / N)^{1/2}$ , sigue aproximadamente una distribución  $t(N - 1)$ .

### 1.2.9 Definición

La **correlación simple muestral de orden  $k$**  ( $k > 0$ ) de una muestra  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$

asociada con una serie temporal  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$  es

$$\hat{\rho}_k \equiv \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (k = 1, 2, \dots) \quad (2.4)$$

(ver Figura 2.3), donde

$$\begin{aligned} \hat{\gamma}_k &\equiv \frac{1}{N} \sum_{t=1}^{N-k} (Y_t - \hat{\mu}_Y)(Y_{t+k} - \hat{\mu}_Y) \\ &\equiv \frac{1}{N} \sum_{t=k+1}^N (Y_{t-k} - \hat{\mu}_Y)(Y_t - \hat{\mu}_Y) \quad (k = 0, 1, 2, \dots). \end{aligned} \quad (2.5)$$

**Observación 1:** El estimador  $\hat{\gamma}_0$  en (2.5) es el mismo estimador que  $\hat{\sigma}_Y^2$  en (2.3). Los **estimadores**  $\hat{\gamma}_k$  y  $\hat{\rho}_k$  proporcionan **estimaciones** numéricas  $c_k$  y  $r_k$ , respectivamente,

$$\begin{aligned} c_k &\equiv \frac{1}{N} \sum_{t=1}^{N-k} (y_t - \bar{y})(y_{t+k} - \bar{y}) \equiv \frac{1}{N} \sum_{t=k+1}^N (y_{t-k} - \bar{y})(y_t - \bar{y}) \quad (k = 0, 1, 2, \dots), \\ r_k &= \frac{c_k}{c_0} \quad (k = 1, 2, \dots), \end{aligned}$$

cuando se reemplaza en (2.5) cada variable aleatoria  $Y_t$  por su valor observado  $y_t$  ( $t = 1, \dots, N$ ).

**Observación 2:** La secuencia de valores numéricos ( $r_k : k = 1, 2, \dots$ ) se denomina la **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.

**Observación 3:** Bajo ciertas condiciones,  $(\hat{\rho}_k) \sim \text{NIID}(0, 1 / N)$ , de manera que cualquier autocorrelación simple  $\rho_k$  ( $k \geq 1$ ) de  $(Y_t)$  puede considerarse *individualmente* significativa al 5% cuando  $|r_k| > 1.96 / \sqrt{N}$ . Por otro lado, para determinar si las  $K$  primeras autocorrelaciones simples de  $(Y_t)$  son *conjuntamente* significativas (donde  $K$  es un entero bastante más reducido que  $N$ ), 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  $\chi^2(K)$ .

### 1.2.10 Definición

La **correlación parcial muestral de orden  $k$**  ( $k > 0$ ) de una muestra  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$  es el estimador MCO o MV  $\hat{\phi}_{kk}$  del parámetro  $\phi_{kk}$  en el modelo de regresión (2.2) considerado para  $t = 1, 2, \dots, N$ .

**Observación 1:** Otro estimador razonable para  $\phi_{kk}$  es el que se obtiene al evaluar el cociente de la Figura 2.4, reemplazando en las matrices  $\mathbf{A}_k$  y  $\mathbf{B}_k$  cada autocorrelación simple teórica  $\rho_k$  por su estimador (2.4). La secuencia de valores numéricos  $(r_{kk} : k = 1, 2, \dots)$  que se obtiene al aplicar los estimadores  $\hat{\phi}_{11}, \hat{\phi}_{22}, \dots$  a una serie  $y_1, y_2, \dots, y_N$ , se denomina la **PACF muestral** de dicha serie.

**Observación 2:** Bajo ciertas condiciones,  $(\hat{\phi}_{kk}) \sim \text{NIID}(0, 1 / N)$ , de manera que cualquier autocorrelación parcial  $\phi_{kk}$  ( $k \geq 1$ ) de  $(Y_t)$  puede considerarse *individualmente* significativa al 5% cuando  $|r_{kk}| > 1.96 / \sqrt{N}$ .

1.2.11 Ejemplo

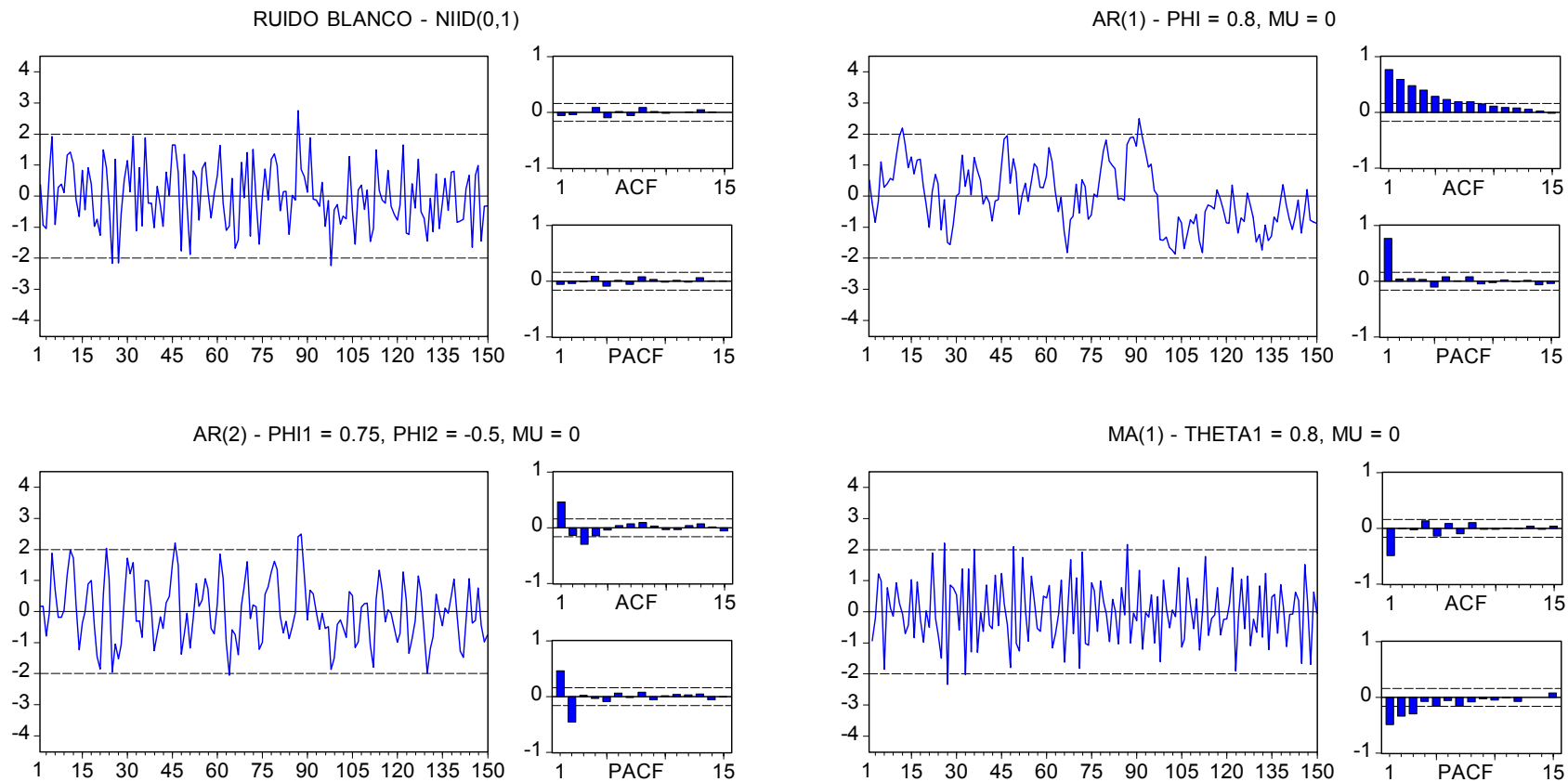
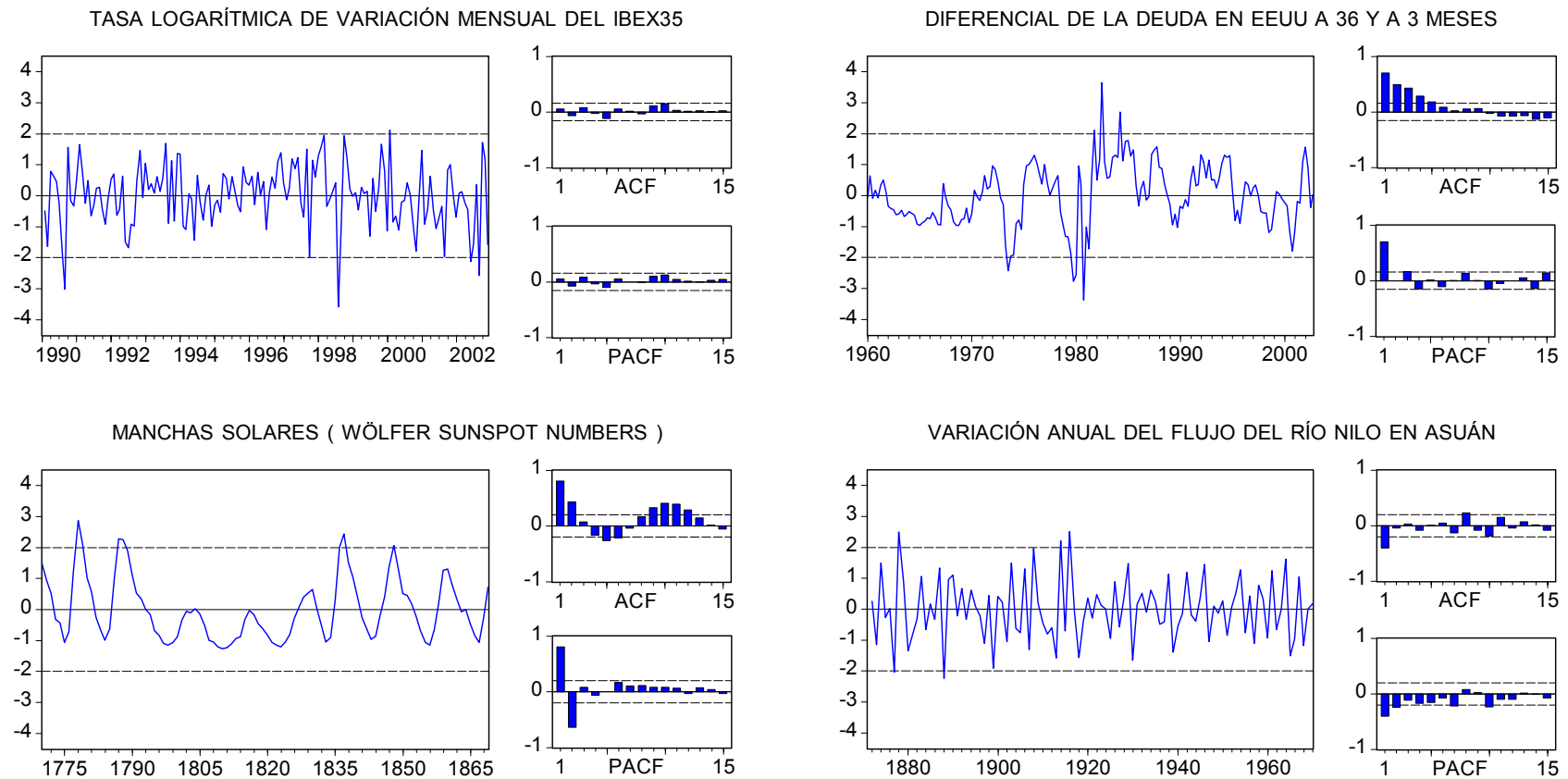


Figura 2.5 – Autocorrelaciones muestrales de varias series estacionarias simuladas

## 1.2.12 Ejemplo



**Figura 2.6 – Autocorrelaciones muestrales de varias series estacionarias reales**

1.2.13 Ejemplo

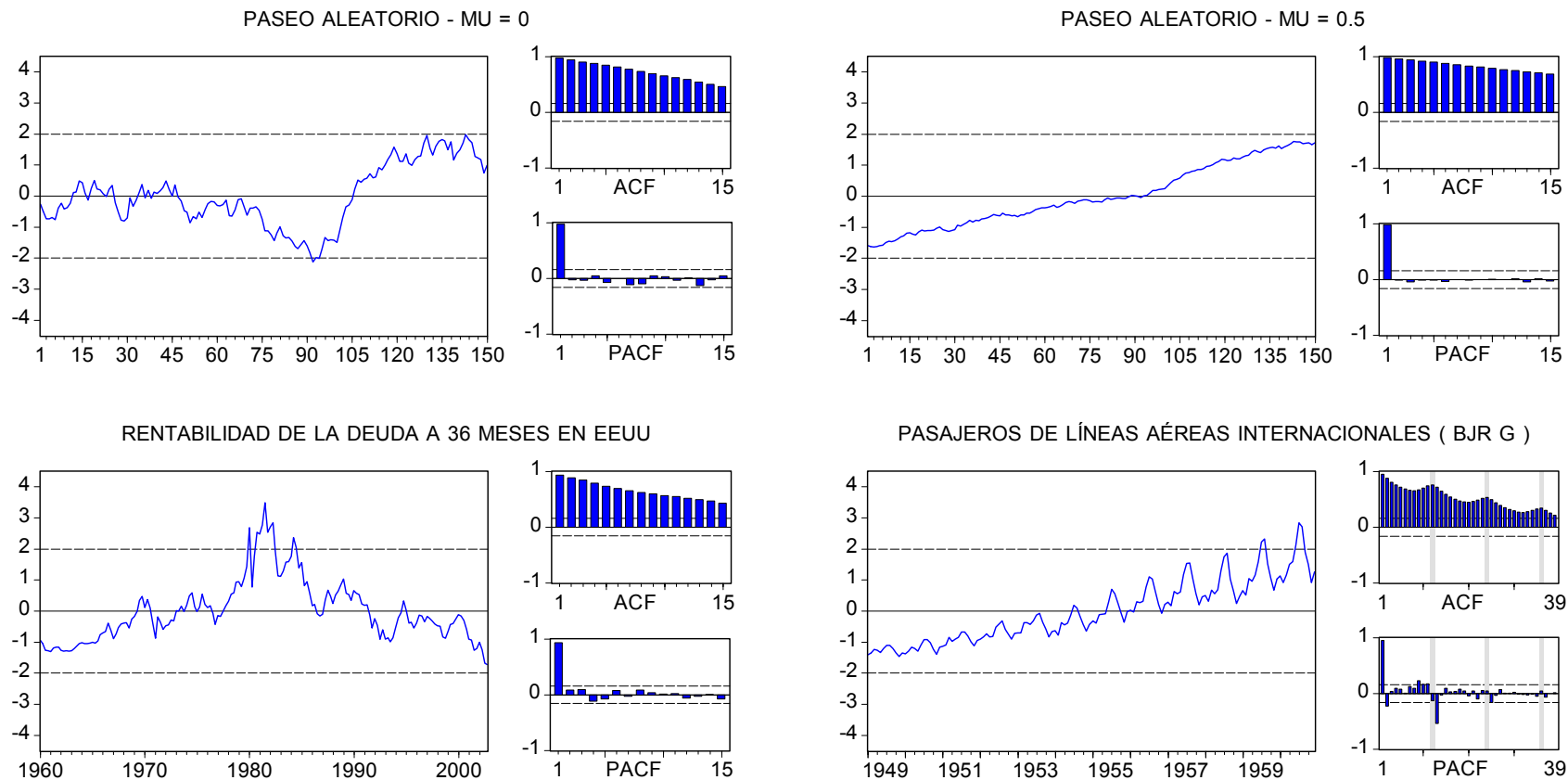


Figura 2.7 – Autocorrelaciones muestrales de varias series no estacionarias

**Observación 1:** Aunque la ACF y la PACF muestrales se han definido en 1.2.8 - 1.2.10 en relación con una serie temporal estacionaria, también pueden calcularse a partir de cualquier secuencia de datos. En particular, la ACF muestral de una **serie no estacionaria** suele decrecer muy lentamente, mientras que la ACF muestral de una **serie estacional** suele presentar el mismo tipo de comportamiento periódico que dicha serie. Por lo tanto, el correlograma de una serie temporal es un instrumento visual muy útil para decidir sobre la estacionariedad de dicha serie.

**Observación 2:** La ACF y la PACF muestrales de una serie temporal estacionaria son dos instrumentos fundamentales para la **identificación** de un modelo ARMA inicial para el proceso estocástico  $(Y_t)$  del que supuestamente procede dicha serie. Por otro lado, la ACF y la PACF muestrales de la **serie de residuos** asociada con un modelo estimado para  $(Y_t)$ , también son dos instrumentos fundamentales en la **diagnos**is de dicho modelo.

## Procesos estacionarios y series no estacionarias

- ⇒ Muchas series temporales (como las del Ejemplo 1.2.13) no pueden considerarse generadas por procesos estocásticos estacionarios (es decir, son series **no estacionarias**), porque presentan ciertas **tendencias** claras en su evolución temporal (de manera que no presentan afinidad hacia algún valor constante en el tiempo), porque su variabilidad no es constante, porque son **estacionales**, o por varias combinaciones de estos motivos.
- ⇒ No obstante, muchas series temporales no estacionarias se pueden **transformar** de forma adecuada para obtener series de aspecto estacionario (como las de los ejemplos 1.2.11 y

1.2.12), que pueden ser utilizadas para elaborar modelos estadísticos compatibles con la hipótesis de estacionariedad. Estos modelos se presentan en la Sección 1.3.

⇒ En la secciones 1.4 y 1.5 se presentan algunas técnicas sencillas para detectar series no estacionarias y para transformar adecuadamente dichas series (dos operaciones que se llevan a cabo al principio de cualquier análisis empírico), así como los modelos estadísticos que resultan al combinar dichas transformaciones con los modelos estacionarios de la sección 1.3.

### **Práctica EV-01**

1. Obtener series temporales para su tratamiento con EViews.
2. Transformar series temporales.
3. Trabajar con submuestras.
4. Elaborar e interpretar gráficos temporales estandarizados.
5. Calcular estadísticos muestrales básicos.

6. Contrastar hipótesis sobre la media de un proceso estocástico.
7. Calcular ACFs y PACFs muestrales, estadísticos de Ljung-Box y *p-values* asociados.
8. Resolver a mano y con EViews el ejercicio numérico propuesto en clase.

## 1.3 MODELOS ARMA

---

Un **modelo ARMA** para un proceso estocástico **estacionario** ( $Y_t$ ) es una representación matemática de  $Y_t$  (el estado de dicho proceso en cualquier momento de su historia) que permite expresar el vector de medias y la matriz de varianzas de cualquier muestra finita  $\mathbf{Y} \equiv [Y_1, Y_2, \dots, Y_N]'$  procedente de  $(Y_t)$  en términos de un **número reducido de parámetros**, todos los cuales pueden ser estimados con precisión a partir de una única realización particular  $\mathbf{y} \equiv [y_1, y_2, \dots, y_N]'$  de  $\mathbf{Y}$ .

### 1.3.1 Definición

Un proceso estocástico **estacionario** ( $Y_t$ ) sigue un **modelo autorregresivo - media móvil** de orden  $(p, q)$ , o ARMA( $p, q$ ) (del inglés *AutoRegressive - Moving Average*), si y sólo si

$$\begin{aligned}
 Y_t = & \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} \\
 & + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \dots - \theta_q A_{t-q}
 \end{aligned}
 \tag{3.1}$$

para todo  $t = 0, \pm 1, \pm 2, \dots$ , donde  $(A_t) \sim \text{IID}(0, \sigma_A^2)$  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 polinomial

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0 \quad (3.2)$$

están fuera del círculo unitario (**condición de estacionariedad**). Un modelo ARMA( $p, q$ ) descrito por la ecuación (3.1) es **invertible** si todas las raíces de la ecuación polinomial

$$1 - \theta_1 x - \theta_2 x^2 - \dots - \theta_q x^q = 0 \quad (3.3)$$

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

### 1.3.2 Definición

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

$$BX_t \equiv X_{t-1}, B^d X_t \equiv X_{t-d} \quad (d \geq 1 \text{ entero}), \quad (3.4)$$

donde  $X_t$  es una variable (real o aleatoria) referida a un momento  $t$  determinado.

### 1.3.3 Definición

La ecuación (3.1) puede escribirse alternativamente como

$$\phi(B)Y_t = \mu + \theta(B)A_t, \quad (3.5)$$

donde

$$\phi(B) \equiv 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (3.6)$$

es el **operador** o **polinomio autorregresivo** (AR) del modelo [comparar con (3.2)] y

$$\theta(B) \equiv 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (3.7)$$

es el **operador** o **polinomio media móvil** (MA) [comparar con (3.3)].

## Estacionariedad – Invertibilidad

⇒ Cuando un proceso estacionario ( $Y_t$ ) sigue un modelo ARMA( $p, q$ ) descrito por (3.1) ó (3.5), la **esperanza incondicional** de ( $Y_t$ ) es

$$\mu_Y \equiv E[Y_t] = \frac{\mu}{1 - \phi_1 - \phi_2 - \dots - \phi_p} = \frac{\mu}{\phi(1)}, \quad (3.8)$$

donde  $\phi(1)$  es el valor del operador AR (3.6) evaluado en  $B = 1$ . Por lo tanto, (3.5) puede escribirse alternativamente como

$$\phi(B)(Y_t - \mu_Y) = \theta(B)A_t, \quad \text{o bien} \quad \phi(B)\tilde{Y}_t = \theta(B)A_t, \quad (3.9)$$

donde  $\tilde{Y}_t \equiv Y_t - E[Y_t] \equiv Y_t - \mu_Y$  para todo  $t = 0, \pm 1, \pm 2, \dots$

⇒ La condición de **estacionariedad** de un modelo ARMA enunciada en 1.3.1 garantiza que los parámetros o **pesos**  $\psi_0, \psi_1, \psi_2, \dots$  del polinomio de orden infinito

$$\psi(B) \equiv \frac{\theta(B)}{\phi(B)} \equiv 1 + \psi_1 B + \psi_2 B^2 + \dots \equiv \sum_{i=0}^{\infty} \psi_i B^i \quad (\psi_0 \equiv 1) \quad (3.10)$$

satisfacen la condición  $\sum_{i=0}^{\infty} |\psi_i| < \infty$ , que es suficiente para que un proceso  $(Y_t)$  tal que  $\tilde{Y}_t = \psi(B)A_t$  sea estacionario (**Teorema de Wold**). Ver Ejemplo 1.3.4.

⇒ La condición de **invertibilidad** de un modelo ARMA enunciada en 1.3.1 garantiza que los parámetros o **pesos**  $\pi_0, \pi_1, \pi_2, \dots$  del polinomio de orden infinito

$$\pi(B) \equiv \frac{\phi(B)}{\theta(B)} \equiv 1 - \pi_1 B - \pi_2 B^2 - \dots \equiv - \sum_{i=0}^{\infty} \pi_i B^i \quad (\pi_0 \equiv -1) \quad (3.11)$$

satisfacen la condición  $\sum_{i=0}^{\infty} |\pi_i| < \infty$ , lo que implica que cuando (3.9) se escribe como  $\pi(B)\tilde{Y}_t = A_t$ ,  $(Y_t)$  es un proceso estacionario tal que el **efecto parcial** de  $Y_{t-k}$  ( $k > 0$ ) sobre  $Y_t$  tiende a cero a medida que el retardo  $k$  aumenta. Ver Ejemplo 1.3.6.

**Observación:** Dados unos valores para los parámetros  $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$  de un modelo ARMA, los pesos  $\psi$  y los pesos  $\pi$  de (3.10) y (3.11) 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 (3.10) y (3.11) 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  $(\psi_i : i = 1, 2, \dots)$  y  $(\pi_i : i = 1, 2, \dots)$  calculadas a partir de un modelo ARMA( $p, q$ ) para un proceso estacionario  $(Y_t)$ , siguen exactamente las mismas pautas que la ACF y la PACF teóricas de  $(Y_t)$ , respectivamente, implicadas por dicho modelo.

## ACF – PACF teóricas en modelos ARMA

⇒ Cuando un proceso estacionario  $(Y_t)$  sigue un modelo ARMA( $p, q$ ) descrito por (3.1), (3.5) ó (3.9), la esperanza, la varianza, la ACF y la PACF de  $(Y_t)$  sólo dependen de los  $p + q + 2$  parámetros  $(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \mu$  y  $\sigma_A^2)$  que figuran en dicho modelo.

### 1.3.4 Ejemplo – Modelo autorregresivo de primer orden o AR(1)

Cuando un proceso estacionario  $(Y_t)$  sigue un modelo ARMA(1,0) ó AR(1)

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t \quad (3.12.1)$$

con  $|\phi_1| < 1$  (ver 1.3.1) y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ , puede comprobarse lo siguiente:

$$Y_t = \phi_1^{t-k} y_k + \sum_{i=0}^{t-k-1} \phi_1^i A_{t-i} + \left( \sum_{i=0}^{t-k-1} \phi_1^i \right) \mu \quad \text{para } t > k, \quad (3.12.2)$$

donde  $k$  es un origen temporal dado e  $y_k$  es un valor inicial dado. Si  $k \rightarrow -\infty$  y  $|\phi_1| < 1$ :

$$Y_t = \frac{\mu}{1 - \phi_1} + \sum_{i=0}^{\infty} \phi_1^i A_{t-i} \quad \text{para todo } t = 0, \pm 1, \pm 2, \dots \quad (3.12.3)$$

Media: 
$$\mu_Y = \frac{\mu}{1 - \phi_1}. \quad (3.12.4)$$

ACF: 
$$\rho_k = \phi_1 \rho_{k-1} \quad (k \geq 1). \quad (3.12.5)$$

**Observación:** La expresión anterior es una ecuación en diferencias de orden 1 que puede resolverse recursivamente a partir de la condición inicial  $\rho_0 = 1$  (ver Figura 2.3), lo que proporciona  $\rho_k = \phi_1^k$  para todo  $k \geq 0$  (una **exponencial amortiguada**).

PACF: 
$$\phi_{kk} = \begin{cases} \phi_1 & \text{si } k = 1, \\ 0 & \text{para todo } k > 1. \end{cases} \quad (3.12.6)$$

Varianza: 
$$\sigma_Y^2 = \frac{\sigma_A^2}{1 - \phi_1 \rho_1} = \frac{\sigma_A^2}{1 - \phi_1^2}. \quad (3.12.7)$$

### 1.3.5 Ejemplo – Modelo autorregresivo de segundo orden o AR(2)

Cuando un proceso estacionario  $(Y_t)$  sigue un modelo ARMA(2,0) ó AR(2)

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + A_t \quad (3.13.1)$$

con  $\phi_2 + \phi_1 < 1$ ,  $\phi_2 - \phi_1 < 1$ ,  $|\phi_2| < 1$  (ver 1.3.1) y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ , se puede

comprobar lo siguiente:

Media: 
$$\mu_Y = \frac{\mu}{1 - \phi_1 - \phi_2}. \quad (3.13.2)$$

ACF: 
$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad (k \geq 1). \quad (3.13.3)$$

**Observación:** La expresión anterior es una ecuación en diferencias que puede resolverse recursivamente a partir de las condiciones iniciales  $\rho_0 = 1$  (ver Figura 2.3) y  $\rho_1 = \phi_1 / (1 - \phi_2)$  [que se obtiene de (3.13.3) con  $k = 1$ ; por otro lado, las dos ecuaciones que resultan de (3.13.3) con  $k = 1$  y  $k = 2$  se denominan **ecuaciones de Yule-Walker**]. En particular, si las dos raíces de (3.2) con  $p = 2$  son complejas, la solución de (3.13.3) es

$$\rho_k = \frac{D^k \text{sen}(2\pi f k + F)}{\text{sen}(F)} \quad \text{para todo } k \geq 0,$$

cuyo numerador representa una **oscilación armónica amortiguada** de **amplitud**  $D^k$ , **período**  $1/f$ , **frecuencia**  $f$ , **fase inicial**  $-F / (2\pi f)$  y **factor de amortiguamiento**  $D$ ; estas características de  $\rho_k$  son funciones de los parámetros  $\phi_1$  y  $\phi_2$  de acuerdo con que  $D = \sqrt{-\phi_2}$ ,  $\cos(2\pi f) = \phi_1 / (2\sqrt{-\phi_2})$  y  $\tan(F) = [(1 - \phi_2) / (1 + \phi_2)] \tan(2\pi f)$ .

PACF: 
$$\phi_{kk} = \begin{cases} \phi_1 / (1 - \phi_2) & \text{si } k = 1, \\ \phi_2 & \text{si } k = 2, \\ 0 & \text{para todo } k > 2. \end{cases} \quad (3.13.4)$$

$$\text{Varianza: } \sigma_Y^2 = \frac{\sigma_A^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2} = \left[ \frac{1 - \phi_2}{1 + \phi_2} \right] \left[ \frac{\sigma_A^2}{(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \right]. \quad (3.13.5)$$

### 1.3.6 Ejemplo – Modelo media móvil de primer orden o MA(1)

Cuando un proceso estacionario  $(Y_t)$  sigue un modelo ARMA(0,1) ó MA(1)

$$Y_t = \mu + A_t - \theta_1 A_{t-1} \quad (3.14.1)$$

con  $|\theta_1| < 1$  (ver 2.3.1) y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ , puede comprobarse lo siguiente:

$$\text{Media: } \mu_Y = \mu. \quad (3.14.2)$$

$$\text{ACF: } \rho_k = \begin{cases} -\theta_1 / (1 + \theta_1^2) & \text{si } k = 1, \\ 0 & \text{para todo } k > 1. \end{cases} \quad (3.14.3)$$

**Observación:** La expresión anterior indica que en un modelo MA(1) invertible ( $|\theta_1| < 1$ ),  $|\rho_1| < \frac{1}{2}$ . No obstante, si de un proceso estacionario  $(Y_t)$  sólo se sabe que  $(Y_t) \sim \text{MA}(1)$  con  $|\rho_1| < \frac{1}{2}$ , entonces tanto  $\theta_1 = \theta$  con  $|\theta| < 1$  (**invertible**) como  $\theta_1 = \theta^{-1}$  con  $|\theta^{-1}| > 1$  (**no invertible**) proporcionan exactamente el mismo valor para  $\rho_1$ , en el sentido que (3.14.3) con  $k = 1$  vale exactamente lo mismo para  $\theta_1 = \theta$  que para  $\theta_1 = \theta^{-1}$ . En consecuencia, una única ACF como (3.14.3) con  $|\rho_1| < \frac{1}{2}$ , puede representarse con dos modelos MA(1), uno invertible y el otro no (lo

mismo ocurre en cualquier modelo que incluye estructura MA de cualquier orden). En este sentido, la condición de invertibilidad puede interpretarse como una condición para la **unicidad** de la representación MA que puede estar asociada con un proceso estacionario dado.

$$\text{PACF:} \quad \phi_{kk} = - \left[ \frac{1}{\sum_{i=0}^k \theta_1^{2i}} \right] \theta_1^k = - \left[ \frac{1 - \theta_1^2}{1 - \theta_1^{2(k+1)}} \right] \theta_1^k \quad \text{para todo } k \geq 1. \quad (3.14.4)$$

$$\text{Varianza:} \quad \sigma_Y^2 = (1 + \theta_1^2) \sigma_A^2. \quad (3.14.5)$$

Por otro lado [ver (3.12.2)]:

$$\begin{aligned} A_t &= \theta_1^{t-k} a_k + \sum_{i=0}^{t-k-1} \theta_1^i Y_{t-i} - \left( \sum_{i=0}^{t-k-1} \theta_1^i \right) \mu \\ \Rightarrow Y_t &= -\theta_1^{t-k} a_k - \sum_{i=1}^{t-k-1} \theta_1^i Y_{t-i} + \left( \sum_{i=0}^{t-k-1} \theta_1^i \right) \mu + A_t \end{aligned} \quad \text{para } t > k, \quad (3.14.6)$$

donde  $k$  es un origen temporal dado y  $a_k$  es un valor inicial dado. Si  $k \rightarrow -\infty$  y  $|\theta_1| < 1$ :

$$Y_t = \frac{\mu}{1 - \theta_1} - \sum_{i=1}^{\infty} \theta_1^i Y_{t-i} + A_t \quad \text{para todo } t = 0, \pm 1, \pm 2, \dots \quad (3.14.7)$$

### 1.3.7 Ejemplo – Modelo media móvil de segundo orden o MA(2)

Si un proceso estacionario  $(Y_t)$  sigue un modelo ARMA(0,2) ó MA(2)

$$Y_t = \mu + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} \quad (3.15.1)$$

con  $\theta_2 + \theta_1 < 1$ ,  $\theta_2 - \theta_1 < 1$ ,  $|\theta_2| < 1$  (ver 2.3.1) y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ , entonces:

Media: 
$$\mu_Y = \mu. \quad (3.15.2)$$

ACF: 
$$\rho_k = \begin{cases} -[\theta_1(1 - \theta_2)] / (1 + \theta_1^2 + \theta_2^2) & \text{si } k = 1, \\ -\theta_2 / (1 + \theta_1^2 + \theta_2^2) & \text{si } k = 2, \\ 0 & \text{para todo } k > 2. \end{cases} \quad (3.15.3)$$

Varianza: 
$$\sigma_Y^2 = (1 + \theta_1^2 + \theta_2^2) \sigma_A^2. \quad (3.15.4)$$

**Observación:** La expresión analítica de la PACF en un modelo MA(2) depende de la naturaleza de las dos raíces de (3.3) con  $q = 2$ ; en general, dicha expresión es muy complicada. No obstante, la PACF sigue exactamente las mismas pautas que la serie  $(\pi_i : i = 1, 2, \dots)$  asociada con (3.11) cuando  $\phi(B) = 1$  y  $\theta(B) = 1 - \theta_1 B - \theta_2 B^2$ . En otros términos, la PACF en un modelo MA(2) es semejante a la ACF en un modelo AR(2); ver Observación en 1.3.5.

### 1.3.8 Ejemplo – Modelo autorregresivo-media móvil de orden (1,1) o ARMA(1,1)

Cuando un proceso estacionario  $(Y_t)$  sigue un modelo ARMA(1,1)

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t - \theta_1 A_{t-1} \quad (3.16.1)$$

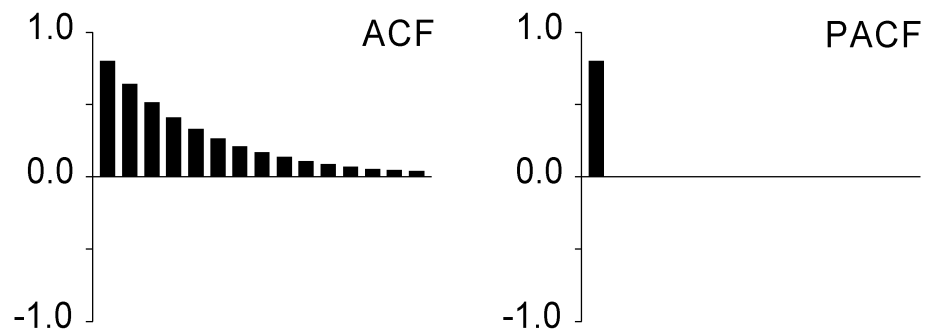
con  $|\phi_1| < 1$ ,  $|\theta_1| < 1$  (ver 2.3.1) y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ , puede comprobarse lo siguiente:

Media: 
$$\mu_Y = \frac{\mu}{1 - \phi_1}. \quad (3.16.2)$$

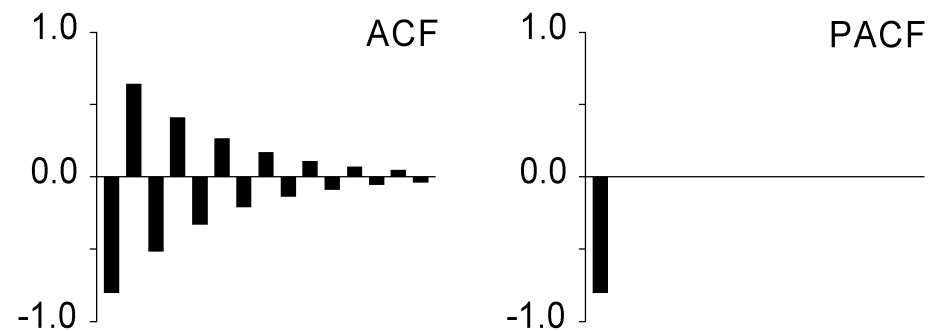
ACF: 
$$\rho_k = \begin{cases} [(\phi_1 - \theta_1)(1 - \phi_1\theta_1)] / (1 - 2\phi_1\theta_1 + \theta_1^2) & \text{si } k = 1, \\ \rho_1 \phi_1^{k-1} & \text{para todo } k > 1. \end{cases} \quad (3.16.3)$$

Varianza: 
$$\sigma_Y^2 = \frac{1 - 2\phi_1\theta_1 + \theta_1^2}{1 - \phi_1^2} \sigma_A^2 = \left[ 1 + \frac{(\phi_1 - \theta_1)^2}{1 - \phi_1^2} \right] \sigma_A^2. \quad (3.16.4)$$

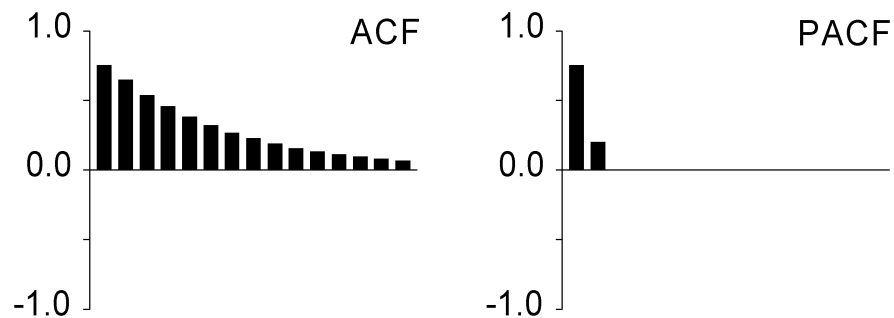
**Observación:** La PACF en un modelo ARMA(1,1) consiste en un valor inicial  $\phi_{11} = \rho_1$  seguido de una secuencia de valores que siguen exactamente la pauta de la PACF en un modelo MA(1); ver Ejemplo 1.3.6. Esto es lo mismo que ocurre con la ACF (3.16.3), que consiste en un valor inicial  $\rho_1$  seguido de una secuencia de valores que siguen exactamente la pauta de la ACF en un modelo AR(1); ver Ejemplo 1.3.4.

**Figura 3.1**

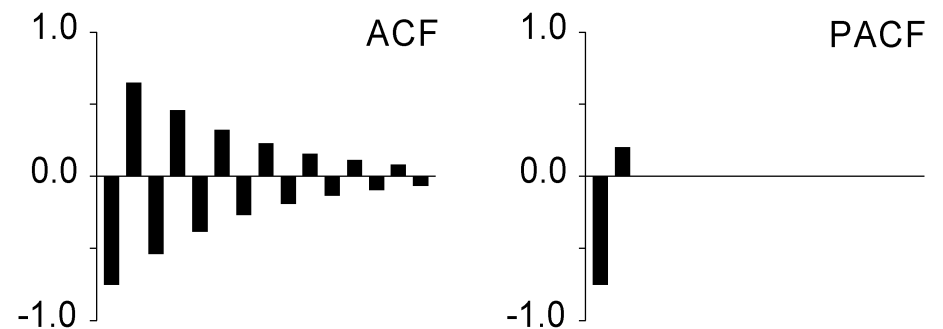
AR(1):  $Y_t = 0.8Y_{t-1} + A_t$ .  
Parámetro positivo.

**Figura 3.2**

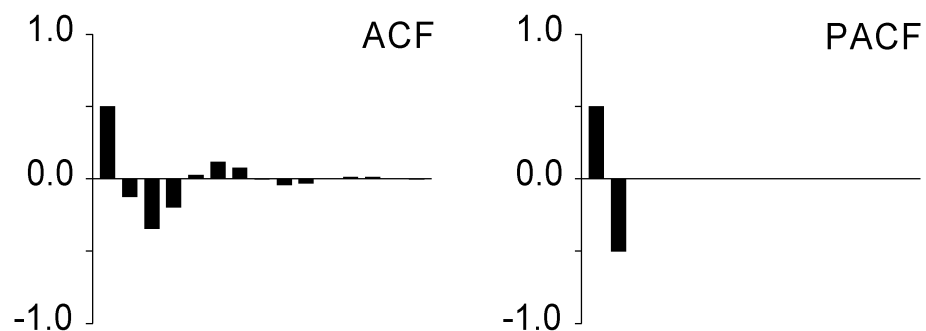
AR(1):  $Y_t = -0.8Y_{t-1} + A_t$ .  
Parámetro negativo.

**Figura 3.3**

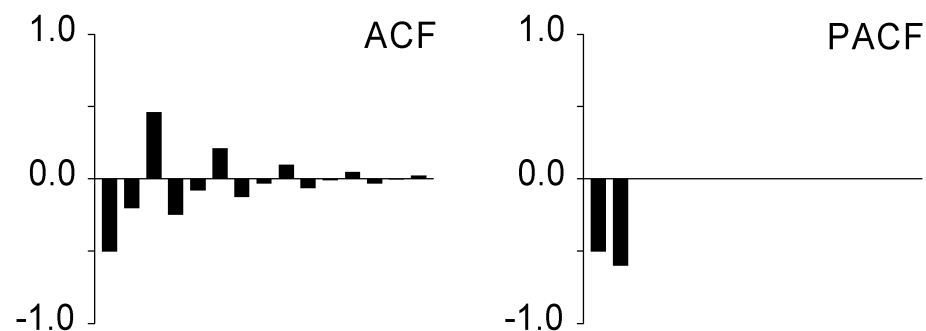
AR(2):  $Y_t = 0.6Y_{t-1} + 0.2Y_{t-2} + A_t$ .  
Raíces de (3.2) reales:  $-4.19, +1.19$  (dominante).

**Figura 3.4**

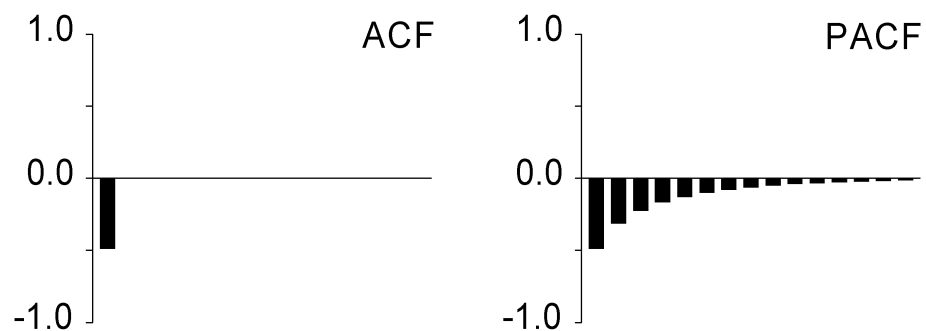
AR(2):  $Y_t = -0.6Y_{t-1} + 0.2Y_{t-2} + A_t$ .  
Raíces de (3.2) reales:  $+4.19, -1.19$  (dominante).

**Figura 3.5**

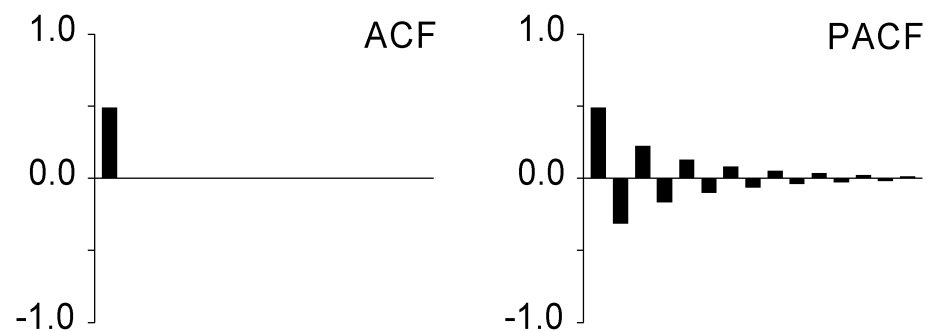
AR(2):  $Y_t = 0.75Y_{t-1} - 0.5Y_{t-2} + A_t$ .  
 Raíces de (3.2) complejas:  $+0.75 \pm 1.20i$ .

**Figura 3.6**

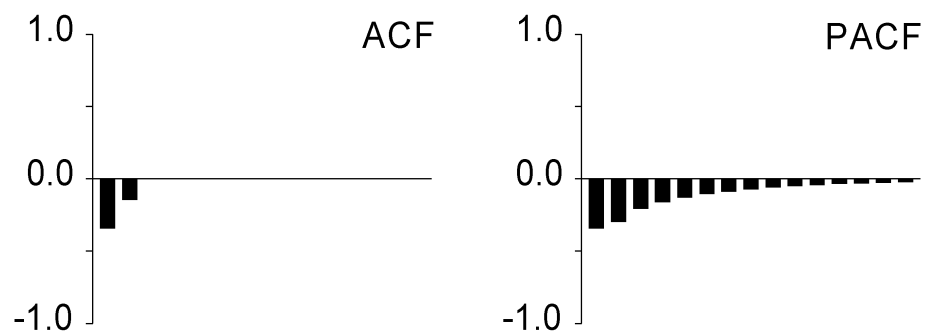
AR(2):  $Y_t = -0.8Y_{t-1} - 0.6Y_{t-2} + A_t$ .  
 Raíces de (3.2) complejas:  $-0.67 \pm 1.11i$ .

**Figura 3.7**

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

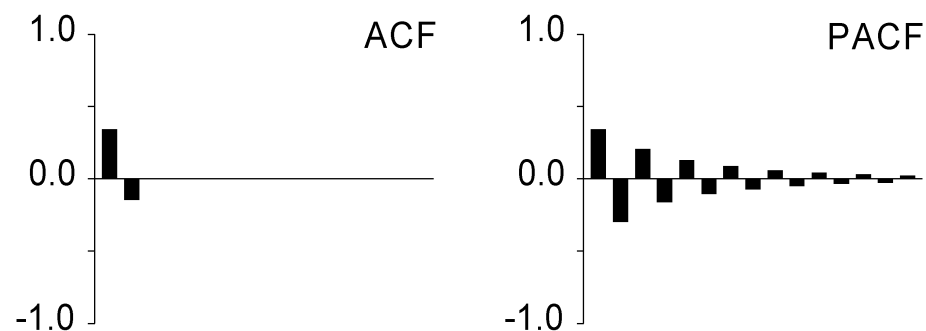
**Figura 3.8**

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

**Figura 3.9**

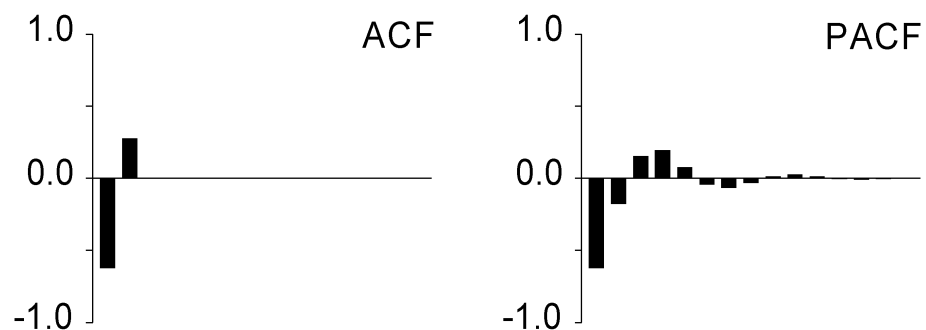
$$\text{MA}(2): Y_t = A_t - 0.6A_{t-1} - 0.2A_{t-2}.$$

Raíces de (3.3) reales:  $-4.19, +1.19$  (dominante).

**Figura 3.10**

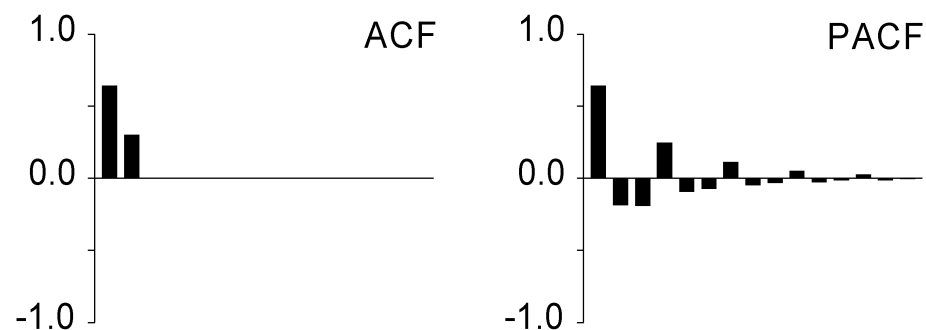
$$\text{MA}(2): Y_t = A_t + 0.6A_{t-1} - 0.2A_{t-2}.$$

Raíces de (3.3) reales:  $+4.19, -1.19$  (dominante).

**Figura 3.11**

$$\text{MA}(2): Y_t = A_t - 0.75A_{t-1} + 0.5A_{t-2}.$$

Raíces de (3.3) complejas:  $+0.75 \pm 1.20i$ .

**Figura 3.12**

$$\text{MA}(2): Y_t = A_t + 0.8A_{t-1} + 0.6A_{t-2}.$$

Raíces de (3.3) complejas:  $-0.67 \pm 1.11i$ .

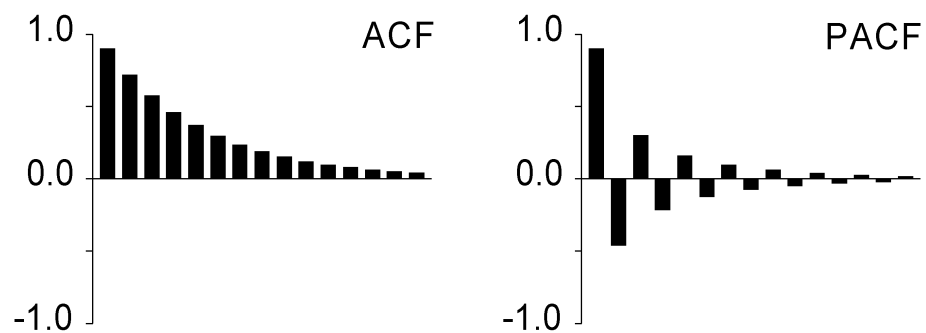


Figura 3.13

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

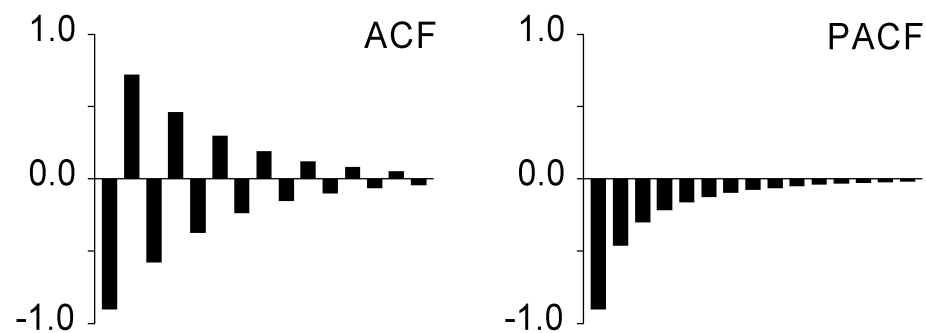


Figura 3.14

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

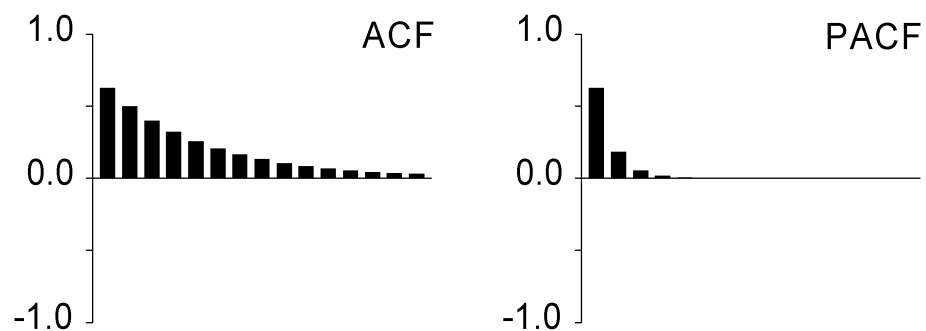


Figura 3.15

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

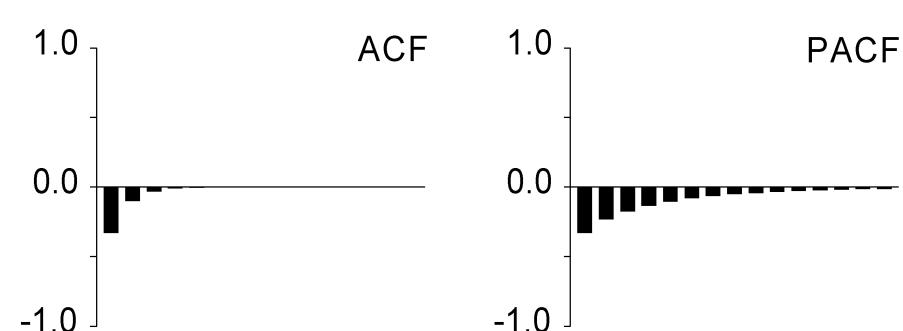


Figura 3.16

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

TABLA – RESUMEN	Modelo AR( $p$ )	Modelo MA( $q$ )	Modelo ARMA( $p, q$ )
Representación de $\tilde{Y}_t$ en términos de $\tilde{Y}_{t-i}$	$\phi(B)\tilde{Y}_t = A_t$	$\theta^{-1}(B)\tilde{Y}_t = A_t$	$\theta^{-1}(B)\phi(B)\tilde{Y}_t = A_t$
Representación de $\tilde{Y}_t$ 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 $\pi(B)$	$\pi(B) = \phi(B)$ (finito)	$\pi(B) = \theta^{-1}(B)$ (infinito)	$\pi(B) = \theta^{-1}(B)\phi(B)$ (infinito)
Operador $\psi(B)$	$\psi(B) = \phi^{-1}(B)$ (infinito)	$\psi(B) = \theta(B)$ (finito)	$\psi(B) = \phi^{-1}(B)\theta(B)$ (infinito)
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 teórica	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 teórica	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 $k = p - q$

**Observación:** El archivo TS39-Sim-ARIMA.wf1 contiene (entre otras) varias series temporales simuladas (generadas) con EViews 3.1 a partir de todos los modelos ARMA cuyas ACF y PACF teóricas aparecen representadas en las figuras 3.1 - 3.13 anteriores. Todas las series se han obtenido a partir de la serie RB, que es una realización particular de tamaño 150 de un proceso de ruido blanco Gaussiano con varianza igual a uno. El modelo utilizado para generar cada serie aparece en la vista por defecto en EViews (SpreadSheet) de cada serie. Las series del archivo TS39-Sim-ARIMA.wf1 pueden utilizarse para observar su aspecto en un gráfico temporal estandarizado, para examinar el aspecto de las ACF y PACF muestrales correspondientes (como en la Figura 2.5), y para comparar dichas funciones de autocorrelación muestrales con las funciones de autocorrelación teóricas de las figuras 3.1 - 3.13.

## Práctica EV-02

1. Realizar algunas de las operaciones a las que se refiere la observación anterior.
2. Repasar las Figuras 2.5 y 2.6 de la Sección 1.2 anterior.

**Demostración de (3.12.2) y (3.14.6)**

Dado un origen temporal  $k$  y un valor inicial  $y_k$ , (3.12.1) implica lo siguiente:

$$Y_{k+1} = \mu + \phi_1 y_k + A_{k+1}. \quad [1]$$

$$\begin{aligned} Y_{k+2} &= \mu + \phi_1 Y_{k+1} + A_{k+2} \\ &= \mu + \left[ \phi_1 \mu + \phi_1^2 y_k + \phi_1 A_{k+1} \right] + A_{k+2} \\ &= (1 + \phi_1) \mu + \phi_1^2 y_k + \left[ A_{k+2} + \phi_1 A_{k+1} \right]. \end{aligned} \quad [2]$$

$$\begin{aligned} Y_{k+3} &= \mu + \phi_1 Y_{k+2} + A_{k+3} \\ &= \mu + \left[ (\phi_1 + \phi_1^2) \mu + \phi_1^3 y_k + \phi_1 A_{k+2} + \phi_1^2 A_{k+1} \right] + A_{k+3} \\ &= (1 + \phi_1 + \phi_1^2) \mu + \phi_1^3 y_k + \left[ A_{k+3} + \phi_1 A_{k+2} + \phi_1^2 A_{k+1} \right], \end{aligned} \quad [3]$$

que puede escribirse como

$$Y_{k+3} = \phi_1^3 y_k + \sum_{i=0}^{3-1} \phi_1^i A_{k+3-i} + \left( \sum_{i=0}^{3-1} \phi_1^i \right) \mu.$$

Por lo tanto, para obtener  $Y_t \equiv Y_{k+(t-k)}$ , sólo hay que reemplazar cada  $3$  en la expresión anterior por  $t - k$ , lo que proporciona directamente (3.12.2).

Por otro lado, el modelo (3.14.1) puede escribirse como  $A_t = -\mu + \theta_1 A_{t-1} + Y_t$ , que tiene una estructura completamente análoga (dual) a la del modelo AR(1) de la demostración anterior. Por lo tanto, reemplazando en (3.12.2)  $Y_t$  por  $A_t$ ,  $\phi_1$  por  $\theta_1$ ,  $y_k$  por  $a_k$ ,  $A_{t-i}$  por  $Y_{t-i}$ , y  $\mu$  por  $-\mu$ , se obtiene directamente (3.14.6) para el modelo MA(1) (3.14.1).

## 1.4 PROCESOS ESTOCÁSTICOS NO ESTACIONARIOS

---

- ⇒ Muchas series temporales (como las de las Figuras 1.4 y 1.5) no pueden considerarse generadas por procesos estocásticos estacionarios (es decir, son series **no estacionarias**), porque presentan ciertas **tendencias** claras en su evolución temporal (de manera que no presentan afinidad hacia algún valor constante en el tiempo), porque su dispersión no es constante, porque son **estacionales**, o por varias combinaciones de estos motivos.
- ⇒ No obstante, muchas series temporales no estacionarias se pueden **transformar** de forma adecuada para obtener series de aspecto estacionario (como las de la Figura 1.3), que pueden ser utilizadas como punto de partida para elaborar en la práctica modelos  $ARMA(p,q)$  como los descritos en la Sección 1.3 anterior.
- ⇒ Para convertir una serie no estacionaria en otra estacionaria, suelen emplearse en la práctica dos tipos de transformaciones: un primer tipo para estabilizar su dispersión (es decir, para inducir **estacionariedad en varianza**) y un segundo tipo para estabilizar su nivel (es decir, para eliminar su tendencia o para inducir **estacionariedad en media**).

## No estacionariedad en varianza

### 1.4.1 Definición

Sea  $(Y_t)$  un proceso estocástico **no** estacionario tal que

$$\mu_t \equiv E[Y_t] \quad \text{y} \quad \sigma_t^2 \equiv \text{Var}[Y_t]$$

existen y **dependen** de  $t$  (por lo que **no** son constantes; ver 1.2.1 - 1.2.3). Si

$$\sigma_t^2 = \sigma^2 \times h(\mu_t)^2, \quad (4.1)$$

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

$$g'(\cdot) = \frac{1}{h(\cdot)} \quad \text{para cualquier valor de } \mu_t. \quad (4.2)$$

**Observación:** Una aproximación de Taylor de primer orden (lineal) a  $Y_t' \equiv g(Y_t)$  alrededor de  $\mu_t$  indica que  $Y_t' \approx g(\mu_t) + g'(\mu_t)(Y_t - \mu_t)$ ; por lo tanto,  $\text{Var}[Y_t'] \approx g'(\mu_t)^2 \sigma_t^2 = \sigma^2 g'(\mu_t)^2 h(\mu_t)^2$ , que es igual  $\sigma^2$  (constante) cuando  $g(\cdot)$  satisface (4.2). En consecuencia, el propósito de una transformación que satisface (4.2) es estabilizar la varianza de un proceso no estacionario  $(Y_t)$  cuya **dispersión local** depende de su **nivel local** de acuerdo con (4.1).

### 1.4.2 Ejemplo

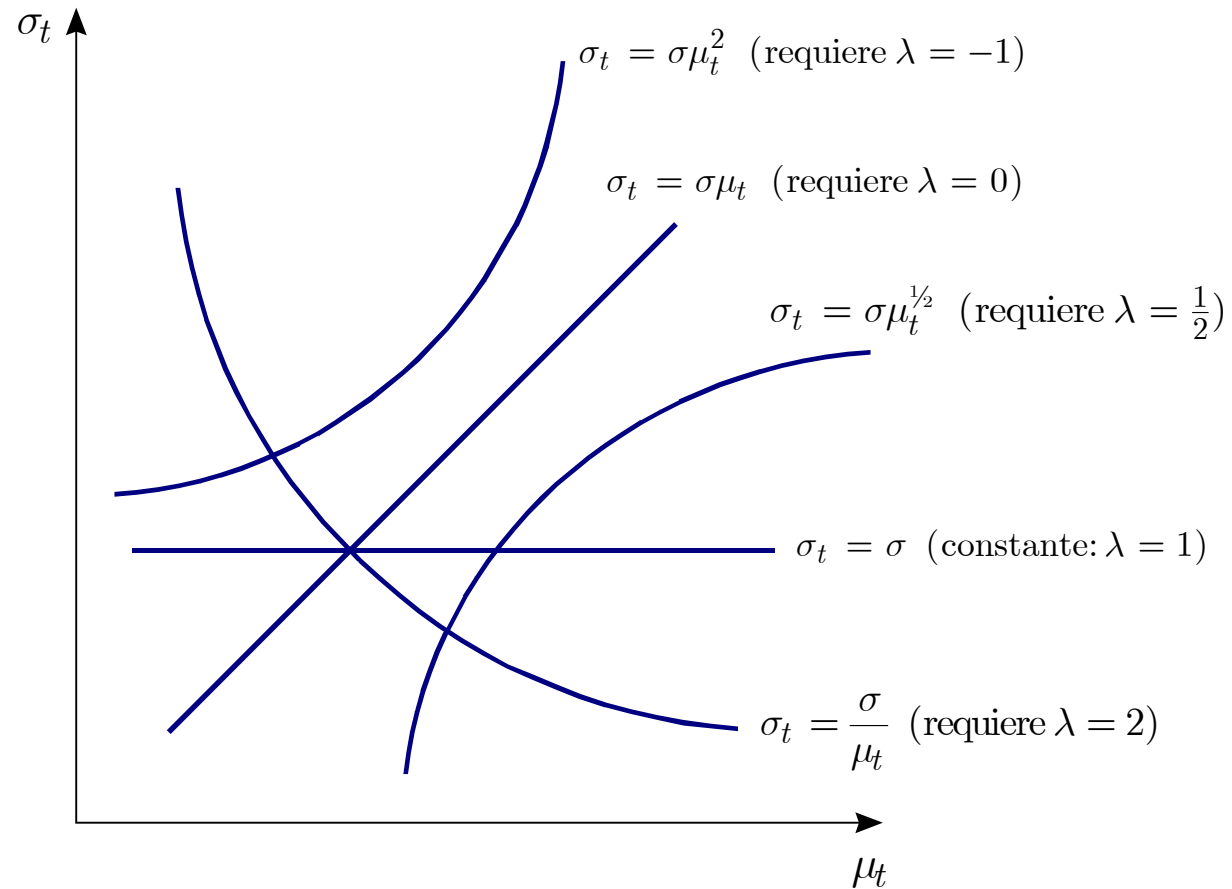
Sea  $(Y_t)$  un proceso estocástico no estacionario tal que  $\mu_t \equiv E[Y_t]$  y  $\sigma_t^2 \equiv \text{Var}[Y_t]$  no son constantes. Si  $\sigma_t$  es proporcional a  $\mu_t$ , de manera que  $\sigma_t^2 = \sigma^2 \mu_t^2$ , entonces  $h(\mu_t) = \mu_t$  en (4.1) y  $g'(\mu_t) = \mu_t^{-1}$  en (4.2), es decir,  $g(\cdot) = \ln(\cdot)$ ; en este caso, la varianza del proceso  $(Y_t') = (\ln Y_t)$  es (aproximadamente) constante. Si  $\sigma_t^2$  es proporcional a  $\mu_t$ , de manera que  $\sigma_t^2 = \sigma^2 \mu_t$ , entonces  $h(\mu_t) = \mu_t^{1/2}$  en (4.1) y  $g'(\mu_t) = \mu_t^{-1/2}$  en (4.2), es decir,  $g(\cdot) = 2(\cdot)^{1/2}$ ; en este caso, la varianza de  $(Y_t') = (2Y_t^{1/2})$  es (aproximadamente) constante.

**Observación 1:** Las dos transformaciones del Ejemplo 1.4.2 pueden escribirse como

$$Y_t' \equiv g(Y_t) = \frac{(Y_t + m)^\lambda - 1}{\lambda},$$

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

**Observación 2:** El hecho de multiplicar un proceso estocástico por una constante no afecta a que su varianza sea o no sea constante; por lo tanto, la presencia de  $\lambda$  en el denominador de la transformación de Box-Cox anterior puede ignorarse en general (lo mismo que el  $-1$  del numerador). La Figura 4.1 representa las dos posibilidades del Ejemplo 1.4.2, junto con otras también contempladas (aunque menos frecuentes) en la transformación de Box - Cox.



**Figura 4.1 – No estacionariedad en varianza: Transformaciones de Box - Cox**

## 1.4.3 Ejemplo - I

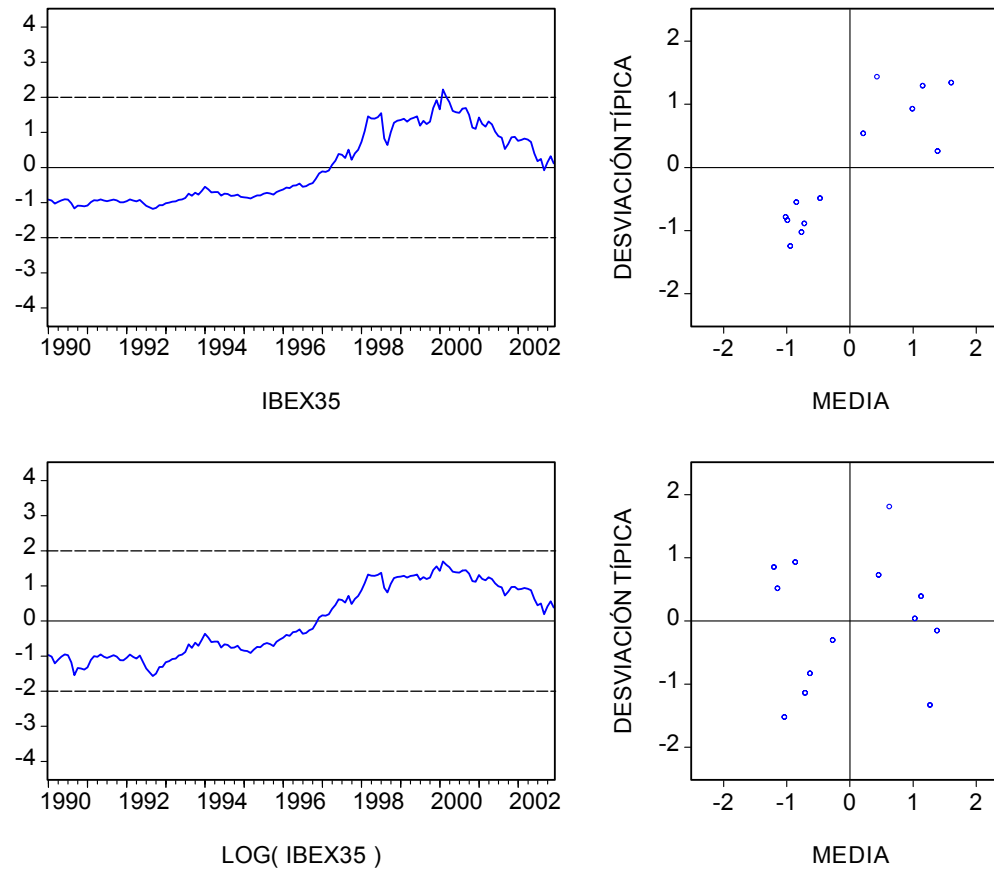


Figura 4.2 – Gráficos desviación típica - media muestrales estandarizados: ver Figura 1.4

1.4.3 Ejemplo - II

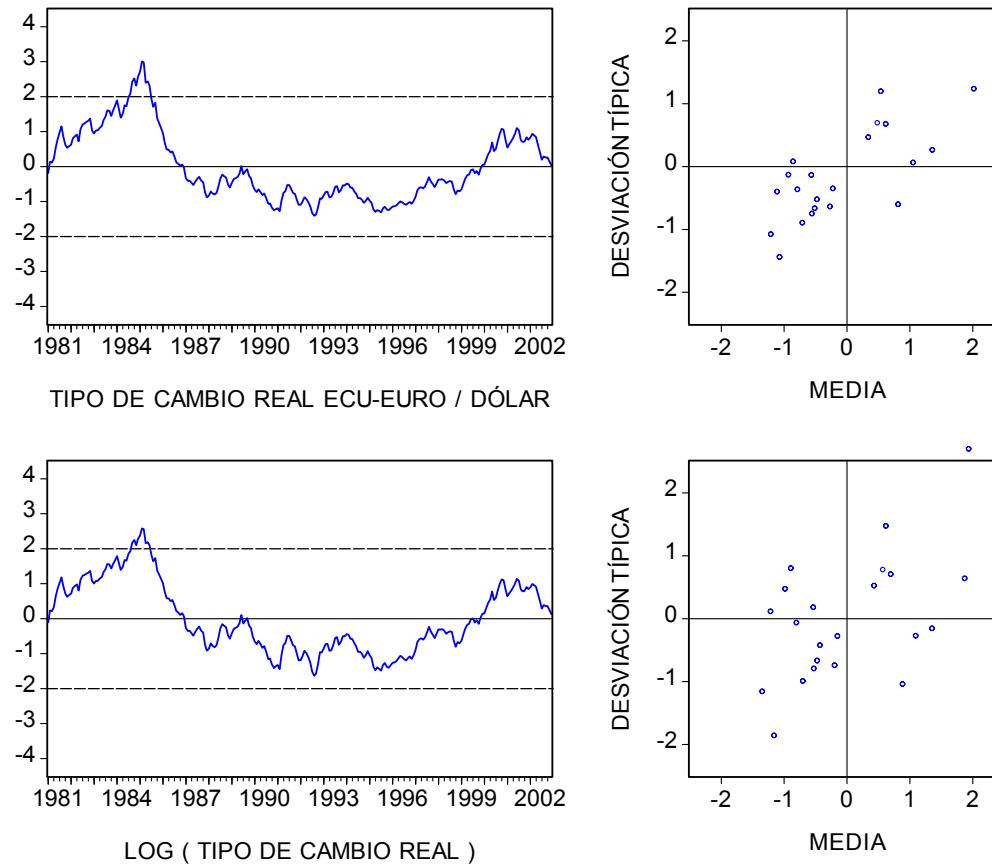


Figura 4.3 – Gráficos desviación típica - media muestrales estandarizados: ver Figura 1.4

## 1.4.3 Ejemplo - III

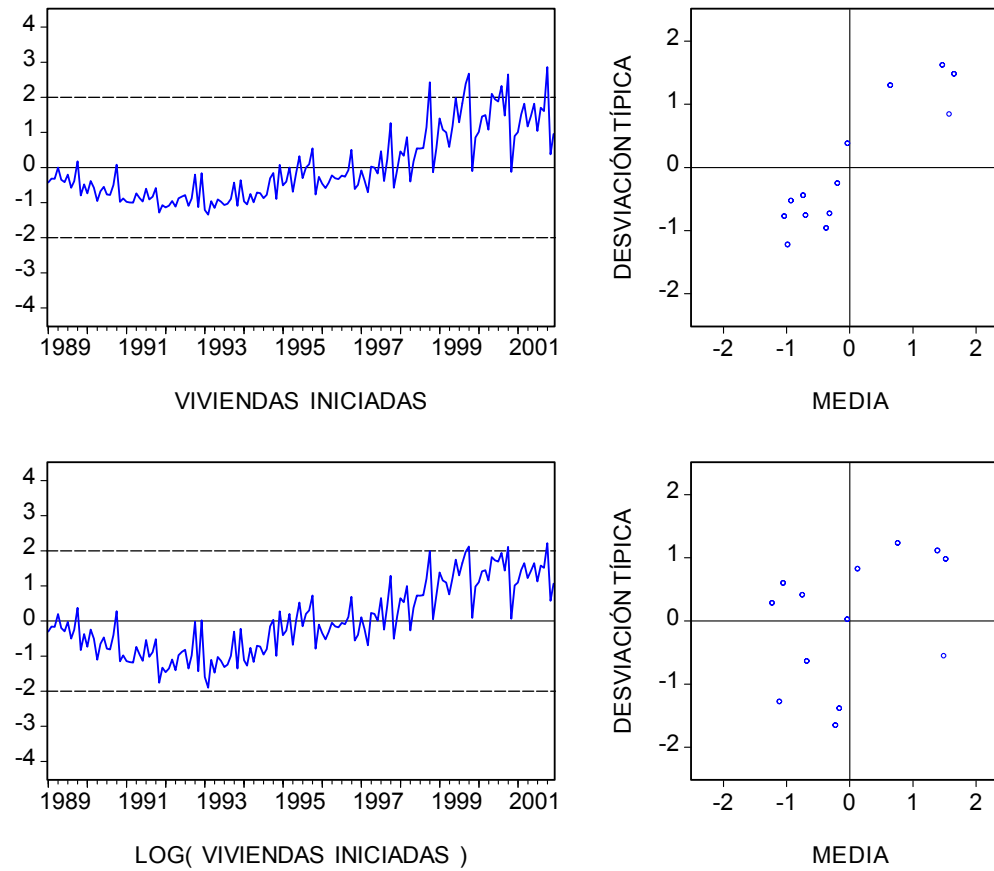


Figura 4.4 – Gráficos desviación típica - media muestrales estandarizados: ver Figura 1.5

**Práctica EV-03**

Elaborar (con el programa DT-Media.prg) e interpretar los gráficos del Ejemplo 1.4.3 anterior.

**No estacionariedad en media****1.4.4 Definición**

El operador **diferencia regular de orden  $d$**  ( $d \geq 1$  entero) se representa con el símbolo  $\nabla^d$  (a veces  $\Delta^d$ ) y se define como  $\nabla^d \equiv (1 - B)^d$  (donde  $B$  es el operador de retardo de 1.3.2), de manera que

$$\nabla X_t \equiv X_t - X_{t-1}, \quad \nabla^2 X_t \equiv X_t - 2X_{t-1} + X_{t-2}, \dots, \quad (4.3)$$

donde  $X_t$  es una variable (real o aleatoria) referida a un momento  $t$  determinado.

**1.4.5 Definición**

Un proceso estocástico  $(Y_t)$  es **integrado de orden  $d$**  ( $d \geq 0$  entero), ó **I( $d$ )**, si y sólo si el

proceso  $(\nabla^d Y_t)$  sigue un modelo ARMA( $p, q$ ) **estacionario** e **invertible** (ver 1.3.1). En tal caso, suele escribirse  $(Y_t) \sim I(d)$ .

**Observación:** En particular: (i) un proceso estocástico  $(Y_t)$  es integrado de orden 0, ó  $I(0)$ , si y sólo si  $(Y_t)$  sigue un modelo ARMA( $p, q$ ) estacionario e invertible (para lo cual es necesario que  $(Y_t)$  sea un proceso estacionario; ver 1.2.1 - 1.2.4), (ii)  $(Y_t) \sim I(1)$  si y sólo si  $(\nabla Y_t) \sim I(0)$ , (iii)  $(Y_t) \sim I(2)$  si y sólo si  $(\nabla^2 Y_t) \sim I(0)$ , ... Como se ilustra en los ejemplos 1.4.6 - 1.4.8, si  $(Y_t) \sim I(d)$  con  $d \geq 1$ , entonces  $(Y_t)$  es un proceso **no** estacionario.

### 1.4.6 Ejemplo

Un **paseo aleatorio** (en inglés *random walk*) es un proceso estocástico  $(Y_t)$  tal que

$$Y_t = \mu + Y_{t-1} + A_t, \quad (4.4)$$

donde  $\mu$  es un parámetro (que puede valer 0) y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ . (4.4) implica que

$$\nabla Y_t \equiv Y_t - Y_{t-1} = \mu + A_t, \quad (4.5)$$

de manera que  $(\nabla Y_t)$  sigue un modelo ARMA(0,0) (con término constante si  $\mu \neq 0$ ). Por lo tanto,  $(\nabla Y_t) \sim I(0)$ , lo que implica que un paseo aleatorio es un proceso  $I(1)$ .

Por otro lado, (4.4) es una **ecuación en diferencias**, cuya **solución general** a partir de un

**origen temporal**  $k < t$  dado es [ver (3.12.2) con  $\phi_1 = 1$ ]

$$Y_t = \left[ (y_k - \mu k) + A_t + \sum_{i=1}^{t-k-1} A_{t-i} \right] + \mu t \quad \text{para } t > k. \quad (4.6)$$

Si, por ejemplo,  $k = 0$  e  $y_0 = 0$ , (4.6) implica que  $E[Y_t] = \mu t$  y  $\text{Var}[Y_t] = \sigma_A^2 t$  dependen de  $t$ , por lo que  $(Y_t)$  es un proceso no estacionario (incluso cuando  $\mu = 0$ ).

**Observación 1:** A partir de (4.6) con  $k = 0$  e  $y_0 = 0$ , puede comprobarse que  $\text{Corr}[Y_t, Y_{t+h}] = [t/(t+h)]^{1/2}$  para todo  $h > 0$ , lo que indica nuevamente que un paseo aleatorio es un proceso no estacionario. Además, para cualquier  $t$  dado, la correlación anterior tiende a cero con  $h$  muy lentamente; de hecho, cuanto mayor es  $t$ , más lentamente tiende  $\text{Corr}[Y_t, Y_{t+h}]$  a cero (por ejemplo, con  $t = 500$ ,  $\text{Corr}[Y_t, Y_{t+h}] > 0.95$  incluso para  $h = 50$ ). Análogamente, puede comprobarse que  $\text{Corr}[Y_{t-h}, Y_t] = [(t-h)/t]^{1/2}$  para todo  $h > 0$ , que no coincide en general con  $\text{Corr}[Y_t, Y_{t+h}]$ . En cualquier caso, el **decrecimiento lento** de la ACF de un paseo aleatorio es característico de cualquier proceso integrado de orden  $d \geq 1$ , lo que contrasta claramente con el **decrecimiento exponencial** de la ACF de un proceso estacionario, ó  $I(0)$ , que sigue cualquiera de los modelos ARMA estacionarios e invertibles de la Sección 1.3.

**Observación 2:** Un proceso integrado de orden  $d \geq 1$  también se denomina un proceso **homogéneamente no estacionario** de orden  $d \geq 1$ , en el sentido que cuando  $(Y_t) \sim I(d)$ ,  $(Y_t)$  presenta una **tendencia** en su evolución temporal, pero sus fluctuaciones alrededor de dicha tendencia constituyen un proceso estocástico estacionario con media cero, por lo que dichas fluctuaciones son homogéneas. En otros términos, cuando  $(Y_t) \sim I(d)$ , su evolución en cada intervalo temporal de su historia (**componente irregular** o **transitorio**) es esencialmente independiente de su tendencia (**componente tendencial** o **permanente**). (El tipo de no estacionariedad que se caracteriza por una clara

dependencia de la evolución temporal local de un proceso con respecto a su tendencia, denominado **no estacionariedad explosiva**, no se considera aquí por no ser representativo de la mayoría de las series temporales económicas).

**Observación 3:** Según lo anterior, si  $(Y_t) \sim I(d)$ ,  $Y_t$  debe admitir una representación del tipo  $Y_t = T_t[Y_t] + I_t[Y_t]$ , donde  $T_t[Y_t]$  es la tendencia de  $(Y_t)$ , e  $I_t[Y_t]$  es la parte irregular de  $(Y_t)$  (que debe ser un proceso estacionario con media cero). La tendencia de un proceso estocástico  $(Y_t)$  puede definirse como el nivel al que tiende  $Y_t$  desde un **origen temporal** suficientemente anterior al momento  $t$ . Por lo tanto, si  $T_t[Y_t]$  no es constante, entonces  $(Y_t)$  es necesariamente un proceso no estacionario (porque **no** converge a largo plazo a algún nivel de equilibrio).

### 1.4.7 Ejemplo

Cuando  $(Y_t)$  es un paseo aleatorio descrito por (4.4), la forma funcional de (4.6) no varía aunque el origen temporal  $k < t$  tienda a  $-\infty$ . En este caso, puede considerarse que

$$T_t[Y_t] = \left[ (y_k - \mu k) + \sum_{i=1}^{t-k-1} A_{t-i} \right] + \mu t, \quad I_t[Y_t] = A_t, \quad (4.7)$$

o bien

$$T_t[Y_t] = \mu + Y_{t-1}, \quad I_t[Y_t] = A_t. \quad (4.8)$$

$T_t[Y_t]$  en (4.7) es la ecuación de una recta de nivel  $\beta_0 \equiv \left[ (y_k - \mu k) + \sum_{i=1}^{t-k-1} A_{t-i} \right]$  (una variable aleatoria) y pendiente  $\mu$  (una constante). Por lo tanto, la tendencia de un paseo

aleatorio tiene dos componentes: uno **determinista** asociado con  $\mu t$  (que no está presente si  $\mu = 0$ ), y otro **estocástico** o **aleatorio** asociado con  $\beta_0$ , cada uno de cuyos sumandos  $A_{t-1}$ ,  $A_{t-2}$ , ...,  $A_{k+1}$  representa una contribución permanente aunque aleatoria al nivel de  $(Y_t)$ .

Por otro lado, (4.5) implica que  $T_t[\nabla Y_t] = \mu$  (constante),  $I_t[\nabla Y_t] = A_t$ , de manera que  $(\nabla Y_t)$  es en este caso un proceso estocástico sin tendencia.

### 1.4.8 Ejemplo

Sea  $(Y_t)$  un proceso estocástico tal que

$$Y_t = Y_{t-1} + A_t - \theta_1 A_{t-1}, \quad (4.9)$$

donde  $\theta_1$  es un parámetro tal que  $|\theta_1| < 1$  y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ . (4.9) implica que

$$\nabla Y_t \equiv Y_t - Y_{t-1} = A_t - \theta_1 A_{t-1}, \quad (4.10)$$

de manera que  $(\nabla Y_t)$  sigue un modelo MA(1) invertible. Por lo tanto,  $(\nabla Y_t) \sim \text{I}(0)$ , lo que implica que un proceso estocástico descrito por (4.9) es un proceso I(1).

Por otro lado, la solución general de (4.9) a partir de un origen temporal  $k < t$  dado es

$$Y_t = (y_k - \theta_1 a_k) + A_t + \sum_{i=1}^{t-k-1} (1 - \theta_1) A_{t-i} \quad \text{para } t > k. \quad (4.11)$$

Si  $k = 0$  e  $y_0 = a_0 = 0$ , (4.11) implica que  $\text{Var}[Y_t] = \sigma_A^2 + (1 - \theta_1)^2 \sigma_A^2 (t - 1)$  depende de  $t$ , por lo que  $(Y_t)$  es un proceso no estacionario.

Dado que la forma funcional de la solución (4.11) no varía aunque el origen temporal  $k < t$  tienda a  $-\infty$ , a partir de (4.9) y (4.11) puede considerarse en este caso que

$$T_t[Y_t] = (y_k - \theta_1 a_k) + A_{t-1} + \sum_{i=2}^{t-k-1} (1 - \theta_1) A_{t-i}, \quad I_t[Y_t] = A_t - \theta_1 A_{t-1}, \quad (4.12)$$

o bien

$$T_t[Y_t] = Y_{t-1}, \quad I_t[Y_t] = A_t - \theta_1 A_{t-1}. \quad (4.13)$$

$T_t[Y_t]$  en (4.12) es puramente estocástica, cada uno de cuyos sumandos  $A_{t-1}$ ,  $(1 - \theta_1)A_{t-2}$ , ...,  $(1 - \theta_1)A_{k+1}$  representa una contribución permanente aunque aleatoria al nivel de  $(Y_t)$ .

Por otro lado, (4.11) implica que  $T_t[\nabla Y_t] = 0$ ,  $I_t[\nabla Y_t] = A_t - \theta_1 A_{t-1}$ , de manera que  $(\nabla Y_t)$  es en este caso un proceso estocástico sin tendencia.

### 1.4.9 Ejemplo

Sea  $(Y_t) \sim I(0)$  un proceso estocástico que sigue un modelo AR(1) estacionario del tipo

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t \quad \text{para todo } t = 0, \pm 1, \pm 2, \dots, \quad (4.14)$$

donde  $\mu$  y  $\phi_1$  son parámetros,  $|\phi_1| < 1$ , y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ . La solución de (4.14) a partir de un origen temporal  $k < t$  dado es [ver (3.12.2)]:

$$Y_t = \phi_1^{t-k} y_k + \sum_{i=0}^{t-k-1} \phi_1^i A_{t-i} + \left( \sum_{i=0}^{t-k-1} \phi_1^i \right) \mu \quad \text{para } t > k, \quad (4.15)$$

que tiende a

$$Y_t = \frac{\mu}{1 - \phi_1} + \sum_{i=0}^{\infty} \phi_1^i A_{t-i} \quad \text{para todo } t = 0, \pm 1, \pm 2, \dots \quad (4.16)$$

cuando  $k \rightarrow -\infty$ . Por lo tanto, en este caso puede considerarse que

$$T_t[Y_t] = \frac{\mu}{1 - \phi_1}, \quad I_t[Y_t] = \sum_{i=0}^{\infty} \phi_1^i A_{t-i} = \phi_1 I_{t-1}[Y_t] + A_t, \quad (4.17)$$

de manera que  $(Y_t)$  es un proceso estocástico sin tendencia (porque es estacionario).

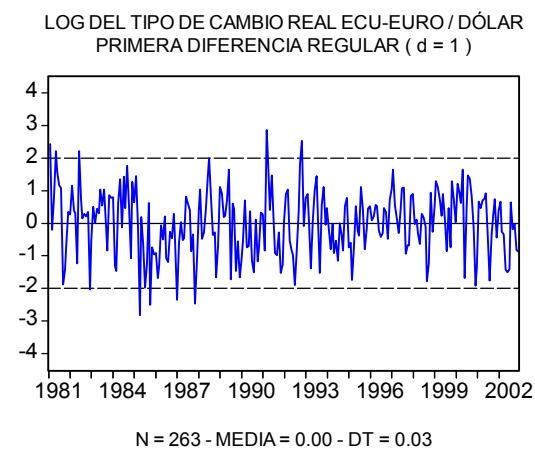
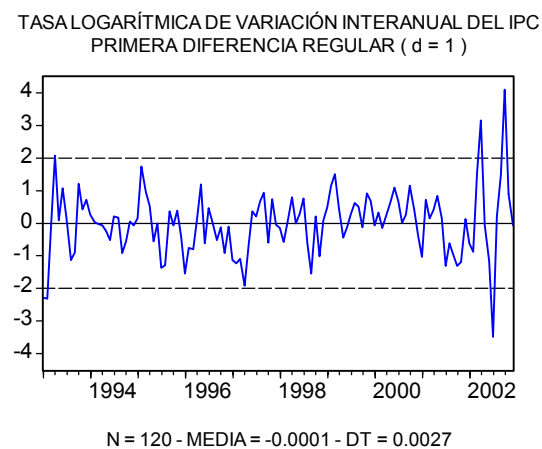
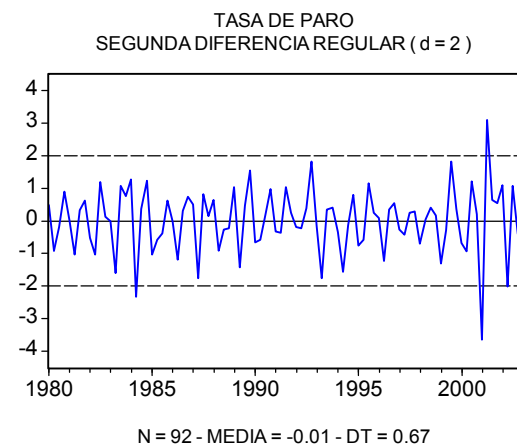
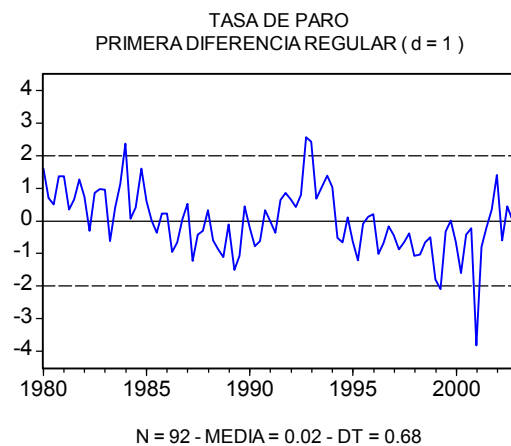
**Observación 1:** Los ejemplos 1.4.6 - 1.4.9 sugieren las siguientes conclusiones generales en relación con un proceso estocástico  $(Y_t) \sim I(d)$  ( $d \geq 1$ ) (que se denomina, a veces, en inglés *difference-stationary*):

⇒  $(Y_t)$  es un proceso **no** estacionario que presenta en su evolución temporal una **tendencia estocástica**, acompañada a veces por alguna tendencia determinista.

⇒  $(W_t) \equiv (\nabla^d Y_t)$  es un proceso  $I(0)$  que, por lo tanto, no presenta tendencia alguna en su evolución temporal.

**Observación 2:** A veces, la tendencia de un proceso  $(Y_t)$  es puramente determinista, como ocurre por ejemplo cuando  $Y_t = \beta_0 + \beta_1 t + A_t$  con  $\beta_1 \neq 0$  y  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ ; en este caso,  $\nabla Y_t = \beta_1 + \nabla A_t$  es un proceso estacionario que sigue un modelo MA(1) **no invertible**, por lo que  $(Y_t)$  **no** es un proceso  $I(1)$  (un proceso de este tipo se denomina en inglés *trend-stationary*). Por otro lado, si  $(Y_t) \sim I(0)$ , como ocurre por ejemplo si  $Y_t = \beta_0 + A_t$ ,  $(Y_t)$  no requiere ninguna diferencia regular para transformarlo en un proceso  $I(0)$ , lo que se observa, por ejemplo, en que  $\nabla Y_t = \nabla A_t$  también sigue un modelo MA(1) **no invertible**. En general, la no invertibilidad indica un uso incorrecto por exceso del operador  $\nabla$  (**sobrediferenciación**), lo que puede llevar a caracterizar erróneamente la tendencia de un proceso.

**Observación 3:** En la práctica, muchas series temporales no estacionarias que presentan tendencias claras en su evolución temporal (ver Figuras 1.1 y 1.4), se pueden transformar en series estacionarias aplicándoles 1 ó 2 diferencias regulares, lo que significa que dichas series han podido ser generadas por procesos integrados de orden 1 ó 2, respectivamente; ver Figura 4.5. Cuando se supone que una serie  $(y_t)$  ha sido generada por un proceso  $(Y_t) \sim I(d)$ , suele decirse que  $(y_t) \sim I(d)$ . En tal caso, la serie (supuestamente) estacionaria  $(w_t) \equiv (\nabla^d y_t)$  puede utilizarse (como se explica en la Sección 1.7) para elaborar un modelo ARMA( $p, q$ ) estacionario e invertible para el proceso  $(W_t)$  del que procede la serie  $(w_t)$ . El modelo correspondiente para el proceso  $(Y_t)$  tal que  $(W_t) \equiv (\nabla^d Y_t)$ , se denomina un modelo ARIMA( $p, d, q$ ) para  $(Y_t)$ ; ver Sección 1.5.



**Figura 4.5 – Diferencias regulares de algunas series no estacionarias**

**Práctica EV-04**

1. Ilustrar la presencia de diferentes tendencias en varias series no estacionarias.
2. Elaborar e interpretar los gráficos de la Figura 4.5 anterior.

## 1.5 MODELOS ARIMA

---

### 1.5.1 Definición

Un proceso estocástico  $(Y_t)$  es integrado de orden  $d$  ( $d \geq 0$  entero) si y sólo si  $(Y_t)$  sigue un **modelo autorregresivo-integrado-media móvil** de orden  $(p, d, q)$ , o  $\text{ARIMA}(p, d, q)$  (del inglés *AutoRegressive-Integrated-Moving Average*), del tipo

$$\phi(B)\nabla^d Y_t = \mu + \theta(B)A_t, \quad (5.1)$$

donde las raíces de las ecuaciones  $\phi(x) = 0$  y  $\theta(x) = 0$  están fuera del círculo unitario.

**Observación 1:** (5.1) puede interpretarse como un modelo  $\text{ARMA}(p, q)$  estacionario para la diferencia regular de orden  $d$  de  $(Y_t)$ , en el sentido que (5.1) es equivalente a  $\phi(B)W_t = \mu + \theta(B)A_t$  con  $(W_t) \equiv (\nabla^d Y_t) \sim \text{I}(0)$ ; ver también (3.5) - (3.9). Por otro lado, (5.1) también puede interpretarse como un modelo  $\text{ARMA}(p+d, q)$  **no** estacionario para  $(Y_t)$  del tipo  $\phi'(B)Y_t = \mu + \theta(B)A_t$ , cuyo operador autorregresivo  $\phi'(B) \equiv \phi(B)\nabla^d \equiv 1 - \sum_{i=1}^{p'} \phi'_i B^i$  ( $p' \equiv p + d$ ) es tal que la ecuación  $\phi'(x) = 0$  tiene exactamente  $d$  **raíces unitarias** [porque  $\phi'(x) \equiv \phi(x) \times (1-x)^d$ ].

**Observación 2:** Como se ilustra con dos ejemplos en la Observación 2 del Ejemplo 1.4.9, la invertibilidad de (5.1) garantiza que el modelo  $\text{ARIMA}$  que sigue  $(Y_t) \sim \text{I}(d)$  describe adecuadamente el tipo de tendencia que presenta  $(Y_t)$  en su evolución temporal (puramente estocástica, o bien con un componente estocástico y otro determinista); para más detalles, ver los dos apartados que siguen al Ejemplo 1.5.2.

### 1.5.2 Ejemplo

El paseo aleatorio de los ejemplos 1.4.6 y 1.4.7 es un modelo ARIMA(0,1,0) para un proceso I(1). El modelo del Ejemplo 1.4.8 es un modelo ARIMA(0,1,1), o IMA(1,1), para un proceso I(1). Un modelo ARIMA(1,1,0), o ARI(1,1), para un proceso I(1) es

$$(1 - \phi_1 B)\nabla Y_t = \mu + A_t, \text{ o bien } Y_t = \mu + (1 + \phi_1)Y_{t-1} - \phi_1 Y_{t-2} + A_t,$$

donde  $|\phi_1| < 1$ , mientras que un modelo ARIMA(1,1,1) para un proceso I(1) es

$$(1 - \phi_1 B)\nabla Y_t = \mu + (1 - \theta_1 B)A_t, \text{ o bien } Y_t = \mu + (1 + \phi_1)Y_{t-1} - \phi_1 Y_{t-2} + A_t - \theta_1 A_{t-1},$$

donde  $|\phi_1| < 1$ ,  $|\theta_1| < 1$  y  $\phi_1 \neq \theta_1$ . Por su parte, todos los modelos de los ejemplos 1.3.4 - 1.3.8 y 1.4.9 son modelos ARIMA( $p,0,q$ ), o ARMA( $p,q$ ), para procesos I(0).

### Tendencias en modelos invertibles

⇨ El modelo ARIMA( $p,d,q$ ) para  $(Y_t) \sim I(d)$  de la Definición 1.5.1 puede escribirse como

$$\phi(B)[\nabla^d Y_t - \mu_W] = \theta(B)A_t, \quad (5.2)$$

donde  $\mu_W \equiv E[\nabla^d Y_t] = \mu / \phi(1)$  [ver (3.5) - (3.9)], o bien como

$$\nabla^d Y_t = \mu_W + \frac{\theta(B)}{\phi(B)} A_t = \mu_W + \psi(B) A_t, \quad (5.3)$$

con  $\psi(B) \equiv \theta(B) / \phi(B)$ ; ver la explicación alrededor de (3.10). Por lo tanto,

$$Y_t = \mu_t + \psi(B) A_t^*, \quad \text{o bien} \quad Y_t = \mu_t + \psi^*(B) A_t, \quad (5.4)$$

donde  $\mu_t$ ,  $A_t^*$  y  $\psi^*(B) \equiv \sum_{i=0}^{\infty} \psi_i^* B^i$  ( $\psi_0^* \equiv 1$ ) se definen de acuerdo con que

$$\nabla^d \mu_t = \mu_W, \quad \nabla^d A_t^* = A_t, \quad \nabla^d \psi^*(B) = \psi(B). \quad (5.5)$$

⇒ En consecuencia, un modelo ARIMA( $p, d, q$ ) para un proceso  $(Y_t) \sim I(d)$  implica que

$$Y_t = \underbrace{\mu_t + \psi(B)(A_t^* - A_t)}_{T_t[Y_t]} + \underbrace{\psi(B)A_t}_{I_t[Y_t]} = \underbrace{\mu_t + [\psi^*(B) - \psi(B)]A_t}_{T_t[Y_t]} + \underbrace{\psi(B)A_t}_{I_t[Y_t]}, \quad (5.6)$$

donde la tendencia  $T_t[Y_t]$  es la suma de un **componente determinista** y un **componente estocástico**, y la parte irregular  $I_t[Y_t]$  es un proceso ARMA estacionario e invertible.

**Observación:** La expresión  $\nabla^d \mu_t = \mu_W$  en (5.5) es una ecuación en diferencias cuya solución general es del tipo  $\mu_t = g(t) + [\mu_W / d!] t^d$ , donde  $g(t)$  es un polinomio de grado igual o inferior a  $d - 1$ , incluyendo la posibilidad de

que  $g(t) = 0$ . Por lo tanto,  $\mu_t$  en (5.6),  $\mu_W$  en (5.2), ó  $\mu$  en (5.1), representan un componente tendencial determinista de grado  $d$  en  $(Y_t)$ , con respecto a cuyos coeficientes un modelo ARIMA tan sólo especifica el número  $d$  de diferencias regulares y (explícita o implícitamente) el valor del parámetro  $\mu_W$  (que puede ser 0, en cuyo caso la tendencia determinista de  $(Y_t)$  sería a lo sumo de grado  $d - 1$ ). En particular, puede comprobarse que (5.5) - (5.6) proporcionan exactamente los resultados de los ejemplos 1.4.7 - 1.4.9.

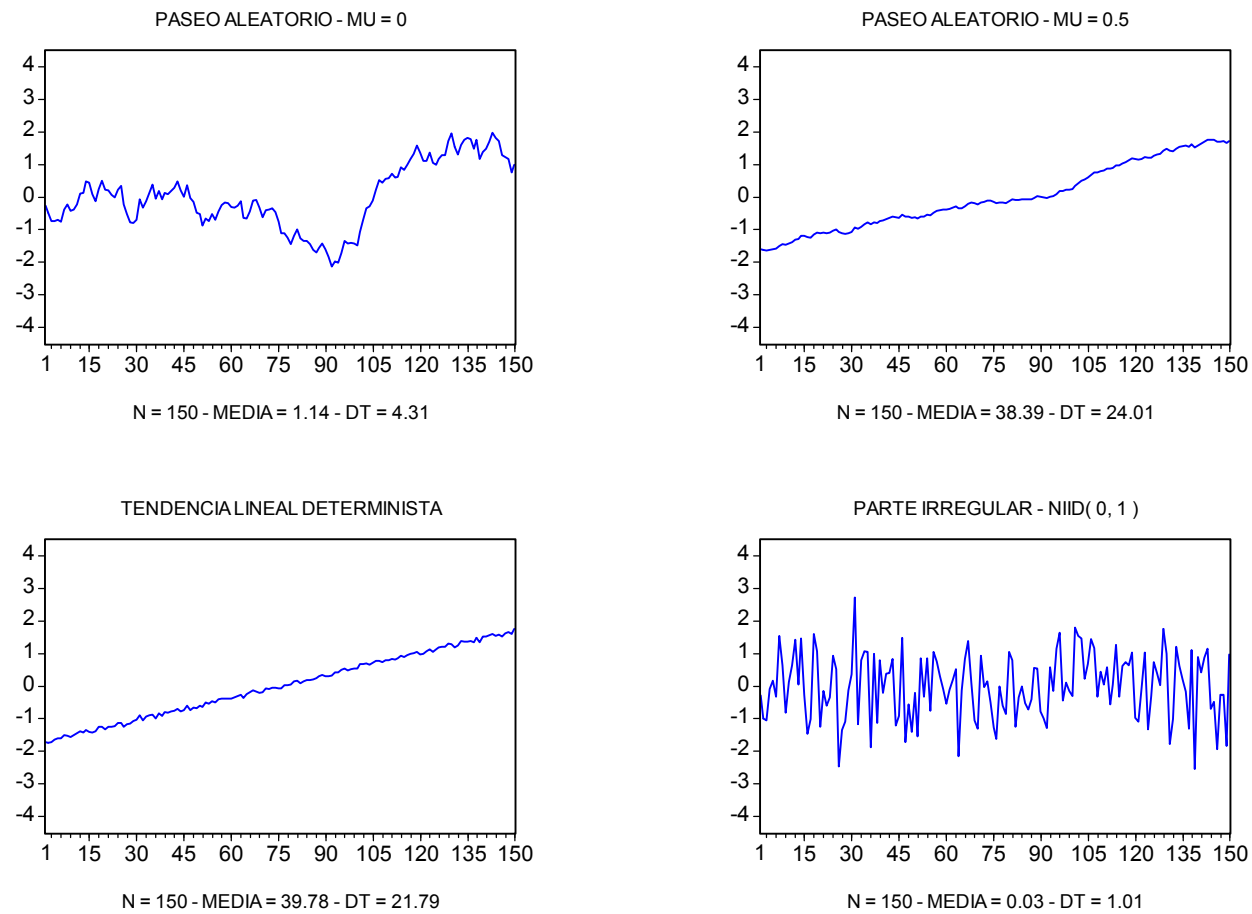
## Tendencias en modelos no invertibles

⇒ Cuando el operador MA en (5.2) es tal que  $\theta(x) = 0$  tiene exactamente  $d$  raíces unitarias,  $\theta(B)$  puede factorizarse como  $\theta(B) = \vartheta(B)\nabla^d$ , donde  $\vartheta(B)$  es un operador invertible. En este caso,  $(Y_t)$  **no** es un proceso  $I(d)$  (ver 1.4.5 ó 1.5.1), y (5.6) daría la falsa impresión de que  $(Y_t)$  tiene un componente tendencial estocástico.

⇒ Cuando  $\theta(B) = \vartheta(B)\nabla^d$ , (5.3) - (5.5) implican que

$$Y_t = \mu_t + \frac{\vartheta(B)}{\phi(B)} A_t,$$

por lo que o bien la tendencia de  $(Y_t)$  es puramente determinista (si  $d \geq 1$  y  $\mu \neq 0$ ), o bien  $(Y_t) \sim I(0)$  y no tiene tendencia (si  $d = 1$  y  $\mu = 0$ ); ver Observación 2 en 1.4.9.



**Figura 5.1 – Series simuladas con distintas tendencias y la misma parte irregular**

## 1.6 MODELOS PARA PROCESOS ESTACIONALES

---

⇒ La **estacionalidad** es una forma de **no estacionariedad** que se observa con frecuencia en muchas series temporales (como las de la Figura 1.5).

### **Series temporales estacionales**

⇒ Se dice que una serie temporal es estacional de periodo  $S$  cuando presenta ciertas pautas en su evolución temporal que se repiten periódicamente cada  $S$  observaciones.

⇒ Muchas series mensuales (trimestrales) son estacionales de periodo  $S = 12$  meses ( $S = 4$  trimestres) porque cada año se repiten ciertas pautas en su evolución temporal.

⇒ El gráfico temporal de una serie y su ACF muestral son instrumentos muy útiles para detectar la posible presencia de estacionalidad en dicha serie.

⇒ Igual que la **diferenciación regular** de la Sección 1.4 se emplea para recoger la tendencia general de una serie no estacionaria, la **diferenciación estacional** se emplea para recoger el comportamiento periódico no estacionario de una serie estacional.

### 1.6.1 Definición

El operador **diferencia estacional de periodo  $S$  y orden  $D$**  ( $D \geq 1$  entero) se representa con el símbolo  $\nabla_S^D$  (a veces  $\Delta_S^D$ ) y se define como  $\nabla_S^D \equiv (1 - B^S)^D$  (donde  $B$  es el operador de retardo de 1.3.2), de manera que, en particular (cuando  $D = 1$ ),

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

donde  $X_t$  es una variable (real o aleatoria) referida a un momento  $t$  determinado.

**Observación 1:** En los casos de series estacionales mensuales ( $S = 12$  meses) y trimestrales ( $S = 4$  trimestres),  $\nabla_{12}$  y  $\nabla_4$  suelen denominarse **diferencias anuales**. Nótese que  $\nabla_S \equiv 1 - B^S \neq (1 - B)^S \equiv \nabla^S$  para todo  $S > 1$ .

**Observación 2:** En la práctica, cuando  $(y_t)$  es una serie estacional de período  $S$ , suele ocurrir que  $(\nabla_S^D y_t)$  no presenta estacionalidad si  $D$  se escoge adecuadamente [aunque  $(\nabla_S^D y_t)$  puede ser no estacionaria por otros motivos]. En este sentido, es muy poco frecuente que el orden de diferenciación estacional  $D$  requerido sea mayor que 1.

**Observación 3:** En conclusión, recapitulando todas las causas posibles de no estacionariedad consideradas a lo largo de este Tema 1, si una serie temporal  $(y_t)$  es no estacionaria por diversos motivos, entonces  $(w_t) \equiv (\nabla^d \nabla_S^D y_t')$  puede ser una serie razonablemente estacionaria si la transformación de Box - Cox (ver 1.4.1 - 1.4.3), 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; ver Figuras 6.1 y 6.2.

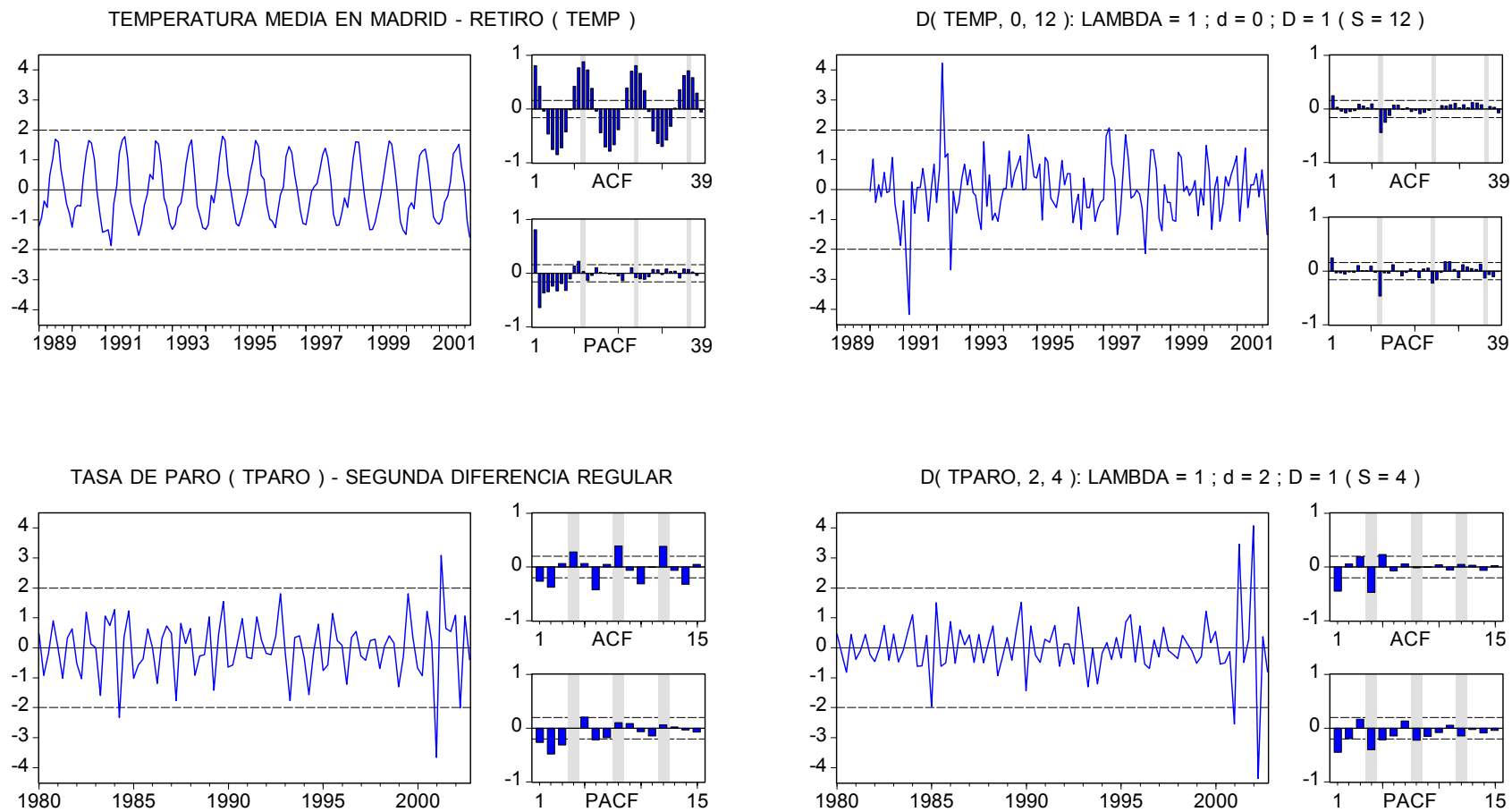
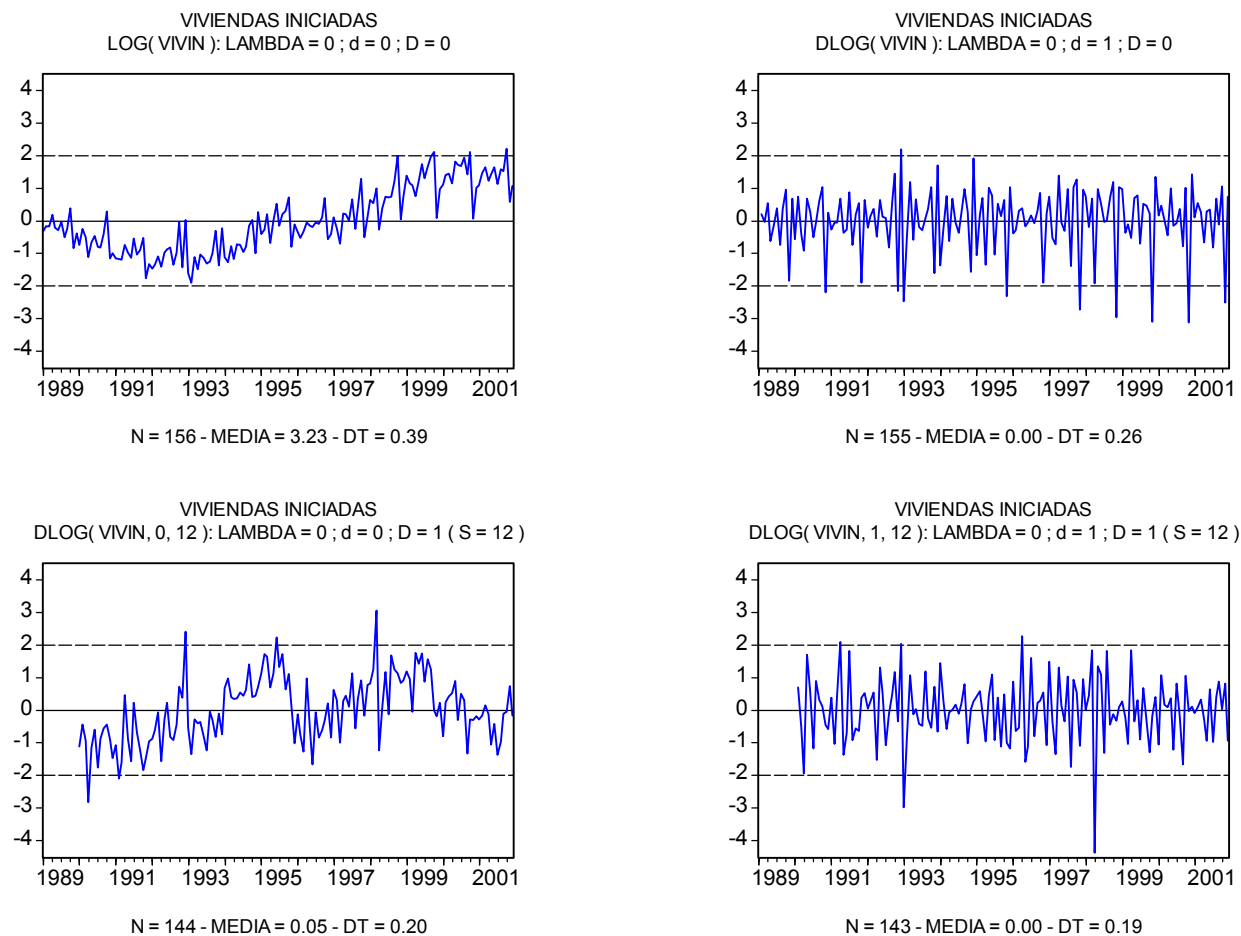


Figura 6.1 – Diferenciación anual de dos series estacionales sin tendencia



**Figura 6.2 – Distintas transformaciones de una serie estacional mensual**

## Práctica EV-05

1. Ilustrar la presencia de pautas periódicas no estacionarias en varias series estacionales.
2. Elaborar e interpretar los gráficos de las Figuras 6.1 y 6.2 anteriores.

## Modelos ARIMA multiplicativos

### 1.6.2 Definición

Un proceso estocástico  $(Y_t)$  sigue un modelo  $ARIMA(p, d, q) \times ARIMA(P, D, Q)_S$  si y sólo si

$$\phi_p(B)\Phi_P(B^S) \left[ \nabla^d \nabla_S^D Y_t - \mu_W \right] = \theta_q(B)\Theta_Q(B^S)A_t, \quad (6.1)$$

donde:

⇨  $(\nabla^d \nabla_S^D Y_t)$  es un proceso estocástico **estacionario** que resulta de aplicar a  $(Y_t)$  una diferencia regular de orden  $d$  [que recoge la tendencia general de  $(Y_t)$ ] y una diferencia

estacional o anual de periodo  $S$  y orden  $D$  [que recoge la estacionalidad de  $(Y_t)$ ] que, en conjunto, estabilizan el nivel de  $(Y_t)$ .

⇒  $\mu_W \equiv E[W_t]$  (constante), con  $(W_t) \equiv (\nabla^d \nabla_S^D Y_t)$  (estacionario).

⇒  $\phi_p(B)$  y  $\theta_q(B)$  son los operadores AR (estacionario) y MA (invertible) habituales (que en un modelo como (6.1) suelen denominarse **operadores regulares**); ver 1.3.3.

⇒  $\Phi_P(B^S) \equiv 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS} \equiv 1 - \sum_{i=1}^P \Phi_i B^{iS}$  es el **operador AR estacional o anual** (con las raíces de  $1 - \sum_{i=1}^P \Phi_i x^i = 0$  fuera del círculo unitario).

⇒  $\Theta_Q(B^S) \equiv 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS} \equiv 1 - \sum_{i=1}^Q \Theta_i B^{iS}$  es el **operador MA estacional o anual**  $P$  (con las raíces de  $1 - \sum_{i=1}^Q \Theta_i x^i = 0$  fuera del círculo unitario).

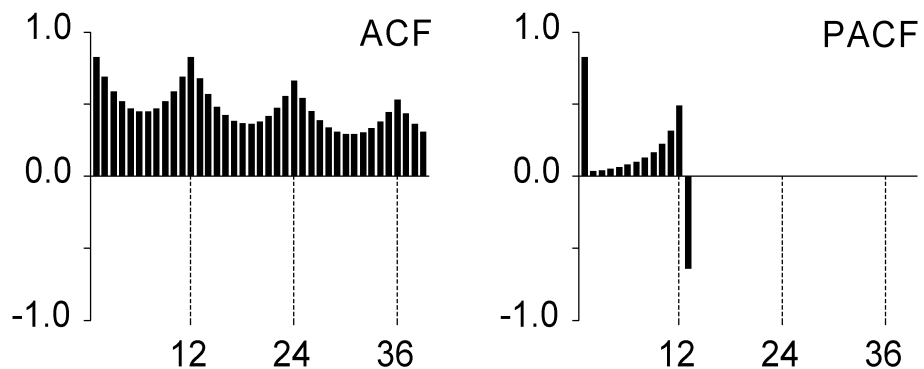
⇒  $(A_t) \sim \text{IID}(0, \sigma_A^2)$ .

**Observación 1:** En (6.1), los operadores regulares  $\phi_p(B)$  y  $\theta_q(B)$  se emplean para modelizar la correlación entre pares de componentes de  $(W_t) \equiv (\nabla^d \nabla_S^D Y_t)$  separados entre sí por  $k = 1, 2, 3, \dots$  períodos básicos (en muchos casos meses o trimestres), mientras que los operadores estacionales  $\Phi_P(B^S)$  y  $\Theta_Q(B^S)$  describen la correlación entre pares de componentes de  $(W_t) \equiv (\nabla^d \nabla_S^D Y_t)$  separados entre sí por  $k = S, 2S, 3S, \dots$  períodos básicos (en muchos casos 1, 2, 3, ... años).

**Observación 2:** Para determinar en la práctica si un modelo como (6.1) representa adecuadamente el proceso que ha podido generar una serie estacional  $(y_t)$  de periodo  $S$ , es necesario conocer la ACF y la PACF teóricas implicadas por (6.1) (con  $d = D = 0$ ) y compararlas con la ACF y la PACF muestrales calculadas a partir de la serie estacionaria  $(w_t) \equiv (\nabla^d \nabla_S^D y_t)$  (es decir, la serie estacional adecuadamente diferenciada). Ver Figuras 6.3 - 6.6.

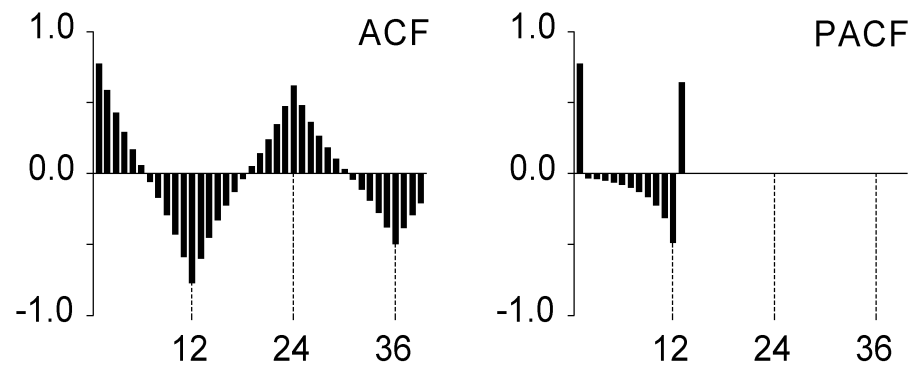
**Observación 3:** En general, la **ACF teórica** de un modelo  $\text{ARMA}(p, q) \times \text{ARMA}(P, Q)_S$  presenta las pautas siguientes: (i) en los primeros retardos  $(1, 2, 3, \dots)$  aparece la ACF de la parte regular  $\text{ARMA}(p, q)$ , (ii) en los retardos anuales  $(S, 2S, 3S, \dots)$  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. En general, 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  $(1, 2, 3, \dots)$  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.

**Observación 4:** En la práctica, con independencia del orden de diferenciación regular  $d$  requerido por una serie estacional para estabilizar su nivel, el orden de diferenciación estacional  $D$  requerido para el mismo fin prácticamente nunca es superior a 1. Por otro lado, especialmente para series mensuales ( $S = 12$ ), los órdenes  $P$  y  $Q$  de (6.1) raramente son mayores que 1 (sobre todo en el caso de  $Q$ ). En conclusión, con valores de  $p, P, d, D, q$  y  $Q$  muy reducidos, suelen obtenerse modelos adecuados en muchos casos prácticos. La manera de elaborar dichos modelos y la forma de utilizarlos para calcular previsiones se describen en las dos secciones siguientes.



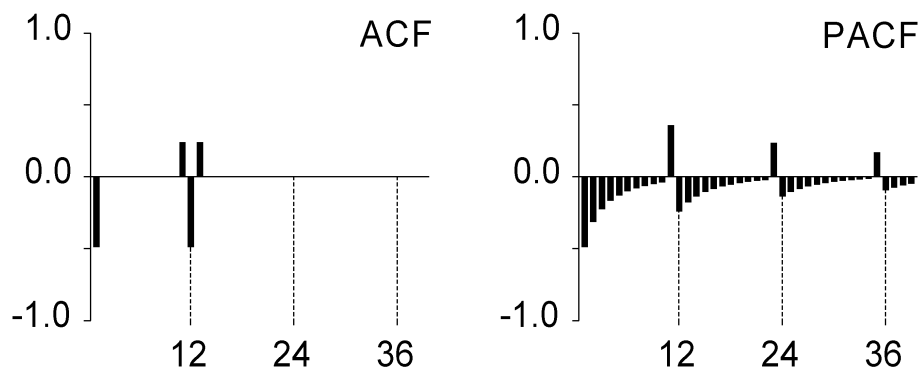
**Figura 6.3**

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



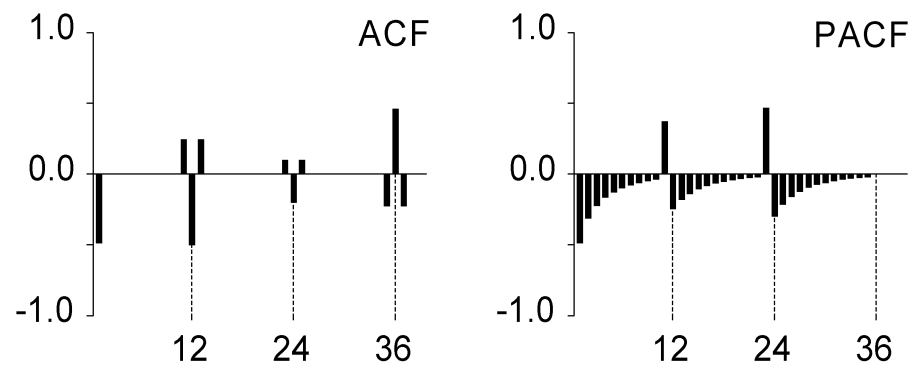
**Figura 6.4**

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



**Figura 6.5**

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



**Figura 6.6**

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

## 1.7 EL ENFOQUE BOX - JENKINS PARA LA ELABORACIÓN DE MODELOS ARIMA

---

El modelo ARIMA más general para  $(Y_t)$  considerado en las secciones anteriores es

$$\phi_p(B)\Phi_P(B^S)[\nabla^d\nabla_S^D Y_t' - \mu_W] = \theta_q(B)\Theta_Q(B^S)A_t, \quad (7.1)$$

donde:

- ⇒  $\mu_W \equiv E[W_t]$  (constante), con  $(W_t) \equiv (\nabla^d\nabla_S^D Y_t')$  **estacionario**, que resulta de aplicar a  $(Y_t)$  una transformación de Box - Cox, una diferencia regular de orden  $d$ , y una diferencia estacional o anual de periodo  $S$  y orden  $D$ .
- ⇒  $\phi_p(B) \equiv 1 - \sum_{i=1}^p \phi_i B^i$ ,  $\theta_q(B) \equiv 1 - \sum_{i=1}^q \theta_i B^i$  (operadores regulares).
- ⇒  $\Phi_P(B^S) \equiv 1 - \sum_{i=1}^P \Phi_i B^{iS}$ ,  $\Theta_Q(B^S) \equiv 1 - \sum_{i=1}^Q \Theta_i B^{iS}$  (operadores anuales).
- ⇒ Todas las raíces de  $\phi_p(x) = 0$ ,  $\theta_q(x) = 0$ ,  $\Phi_P(x) = 0$ , y  $\Theta_Q(x) = 0$  están **fuera del círculo unitario** (estacionariedad e invertibilidad).
- ⇒  $(A_t) \sim \text{IID}(0, \sigma_A^2)$  [**ruido blanco, Normal o Gaussiano** si  $(A_t) \sim \text{NIID}(0, \sigma_A^2)$ ].

## Identificación

**Entrada:**  $(y_t)_{t=1}^N$  (serie temporal original);  $S$  (período estacional).

- 1.1 Elegir  $\lambda$  y  $m$  en la transformación de Box - Cox – Gráfico desviación típica - media. Se obtiene una serie  $(y'_t)$ , con  $y'_t = \ln(y_t + m)$  si  $\lambda = 0$ ,  $y'_t = (y_t + m)^\lambda$  si  $\lambda \neq 0$ .
- 1.2 Elegir  $d$  y  $D$  – Gráfico temporal (tendencias) y ACF muestral (amortiguamiento) de  $(y'_t)$ . Se obtiene una serie  $(w_t)$ , con  $w_t \equiv \nabla^d \nabla_S^D y'_t$  (estacionaria).
- 1.3 Incluir o no  $\mu_W$  – Contraste de significación de  $\mu_W \equiv E[W_t] \equiv E[\nabla^d \nabla_S^D Y'_t]$  a partir de la media muestral de la serie  $(w_t)$ .
- 1.4 Elegir  $p$ ,  $P$ ,  $q$  y  $Q$  – ACF y PACF muestrales de  $(w_t)$  comparadas con ACF y PACF teóricas de modelos  $\text{ARMA}(p, q) \times \text{ARMA}(P, Q)_S$ . El criterio central aquí es el reconocimiento de configuraciones:
  - 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:**  $(w_t)_{t=1}^n$ , con  $w_t \equiv \nabla^d \nabla_S^D y_t'$ ,  $n \equiv N - d - DS$  (serie transformada y diferenciada); órdenes  $p$ ,  $P$ ,  $q$  y  $Q$  [modelo identificado para el proceso (estacionario)  $(W_t) \equiv (\nabla^d \nabla_S^D Y_t')$ ].

**Observación 1 - Contrastes formales de raíces unitarias:** A veces, pueden ayudar a elegir  $d$  (número de diferencias regulares) en el paso 2.1 anterior. Ver, por ejemplo, en EViews, el enlace a **Unit Root Tests** al final de **Help** → **EViews Help Topics** ... → **Statistical Views and Procedures** → **Series Views**; ver también Box, Jenkins y Reinsel (1994), páginas 208 - 211.

**Observación 2 - Tasas de variación:** La **tasa convencional de variación (neta)** de una serie  $(y_t)$  es

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

mientras que la **tasa logarítmica de variación** de una serie  $(y_t)$  se define como

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

que es aproximadamente igual a  $tcv_t(k)$  para valores de  $tcv_t(k)$  pequeños. La tasa convencional es **multiplicativa**, en el sentido que

$$[1 + tcv_t(k)] = \frac{y_t}{y_{t-k}} = \frac{y_t}{y_{t-1}} \frac{y_{t-1}}{y_{t-2}} \dots \frac{y_{t-k+1}}{y_{t-k}} = \prod_{i=0}^{k-1} [1 + tcv_{t-i}(1)],$$

por lo que la tasa logarítmica es **aditiva**:

$$tlv_t(k) = \ln \prod_{i=0}^{k-1} [1 + tcv_{t-i}(1)] = \sum_{i=0}^{k-1} \ln tlv_{t-i}(1).$$

Dado que muchas series temporales requieren una transformación logarítmica y una diferencia regular para darles aspecto estacionario, las series resultantes pueden interpretarse como tasas de variación. Por ejemplo, si  $(y_t)$  es una serie mensual, entonces:

- $(\nabla \ln y_t)$  es la tasa logarítmica de variación mensual de  $(y_t)$ .
- $(\nabla_{12} \ln y_t)$  es la tasa logarítmica de variación interanual de  $(y_t)$ .
- $(\nabla^2 \ln y_t)$  es la variación (absoluta) mensual de  $(\nabla \ln y_t)$ .
- $(\nabla \nabla_{12} \ln y_t)$  es la variación (absoluta) mensual de  $(\nabla_{12} \ln y_t)$ , o bien la variación (absoluta) interanual de  $(\nabla \ln y_t)$ ; análogamente,  $(\nabla^2 \nabla_{12} \ln y_t)$  puede interpretarse de varias formas alternativas.

### Práctica EV-06

1. Identificar un modelo inicial para la serie SUNSPOTS del archivo TS14-BJR-E.wf1.
2. Identificar un modelo inicial para la serie PIB del archivo TS29-PIB-CPI.wf1.
3. Identificar un modelo inicial para la serie TEMP del archivo TS30-Seasonal.wf1.
4. Identificar un modelo inicial para la serie AIRLINE del archivo TS16-BJR-G.wf1.
5. Identificar un modelo inicial para la serie VIVIN del archivo TS30-Seasonal.wf1.

## Estimación

**Entrada:**  $(w_t)_{t=1}^n$  (serie transformada y diferenciada); modelo identificado para el proceso (estacionario)  $(W_t) \equiv (\nabla^d \nabla_S^D Y_t')$ ; estimaciones iniciales de los parámetros del modelo; variables de control para el método de estimación escogido.

2.1 MV  $\Leftrightarrow$  Min  $\tilde{\mathbf{w}}' \mathbf{\Omega}^{-1} \tilde{\mathbf{w}} \times |\mathbf{\Omega}|^{\frac{1}{n}}$ , con  $\mathbf{\Omega} \equiv \sigma_A^{-2} \text{Var}[\mathbf{W}]$  (teórica, implicada por el modelo).

2.2 Problemas: evaluar la forma cuadrática  $\tilde{\mathbf{w}}' \mathbf{\Omega}^{-1} \tilde{\mathbf{w}}$  y el determinante  $|\mathbf{\Omega}|$ ; estimación no lineal (no existen soluciones analíticas).

2.3 Alternativa a MV: MC  $\Leftrightarrow$  Min  $\tilde{\mathbf{w}}' \mathbf{\Omega}^{-1} \tilde{\mathbf{w}}$  (aproximaciones: *retrovisión*, EViews, ...).

**Salida:**  $\hat{\mu}_W$ ,  $\hat{\Phi} \equiv [\hat{\phi}_1, \dots, \hat{\phi}_p]'$ ,  $\hat{\Phi} \equiv [\hat{\Phi}_1, \dots, \hat{\Phi}_P]'$ ,  $\hat{\theta} \equiv [\hat{\theta}_1, \dots, \hat{\theta}_q]'$ ,  $\hat{\Theta} \equiv [\hat{\Theta}_1, \dots, \hat{\Theta}_Q]'$ ,  $\hat{\sigma}_A^2$  (modelo estimado);  $\mathbf{V}$  (matriz de covarianzas estimadas);  $(\hat{a}_t)_{t=1}^n$  (serie de **residuos**).

**Observación 1 - Estimación de modelos univariantes:** La **función de verosimilitud exacta** (FVE) asociada con el modelo univariante (7.1) es la función de densidad de la muestra  $\mathbf{W}$  de la que procede la serie (estacionaria)  $\mathbf{w} \equiv [w_1, \dots, w_n]'$ , considerada como función de los parámetros que figuran en (7.1). Si el proceso  $(W_t) \equiv (\nabla^d \nabla_S^D Y_t')$  del que procede  $\mathbf{W}$  es Normal o Gaussiano [es decir, bajo la hipótesis de que  $(A_t) \sim \text{NIID}(0, \sigma_A^2)$ ], entonces la FVE asociada con el modelo univariante (7.1) 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 de (7.1) excepto  $\sigma_A^2$ ,  $\boldsymbol{\Gamma}_n \equiv \text{Var}[\mathbf{W}]$  (ver la Observación 3 de 1.2.6), y  $\tilde{\mathbf{w}} \equiv \mathbf{w} - \text{E}[\mathbf{W}]$ , 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  $\boldsymbol{\Omega} \equiv \sigma_A^{-2} \boldsymbol{\Gamma}_n$ ), el logaritmo neperiano ( $\ln$ ) 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 1.3.4) y MA(1) (ver 1.3.6),

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

En todos los casos, los elementos de  $\boldsymbol{\Omega} \equiv \sigma_A^{-2} \boldsymbol{\Gamma}_n$  en (2) sólo dependen [como se ilustra en (3)] de los parámetros AR y MA del modelo (7.1) (recogidos en  $\boldsymbol{\phi} \equiv [\phi_1, \dots, \phi_p]'$ ,  $\boldsymbol{\Phi} \equiv [\Phi_1, \dots, \Phi_P]'$ ,  $\boldsymbol{\theta} \equiv [\theta_1, \dots, \theta_q]'$  y  $\boldsymbol{\Theta} \equiv [\Theta_1, \dots, \Theta_Q]'$ ); el cálculo de dichos elementos es semejante al de las ACFs teóricas de los Ejemplos (1.3.4) - (1.3.8).

Para obtener estimaciones por **Máxima Verosimilitud** (MV) de  $\beta$  y  $\sigma_A^2$ , se requiere maximizar numéricamente (1) ó (2) con respecto a  $\beta$  y a  $\sigma_A^2$ . Si  $\hat{\beta}$  y  $\hat{\sigma}_A^2$  son las estimaciones numéricas de  $\beta$  y  $\sigma_A^2$  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  $S(\beta) \equiv \tilde{\mathbf{w}}' \Omega^{-1} \tilde{\mathbf{w}}$  evaluada en  $\beta = \hat{\beta}$ . Sustituyendo  $\sigma_A^2$  por  $\frac{S(\beta)}{n}$  en (2), se obtiene el ln de la FVE **concentrada** en  $\beta$ :

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

de manera que

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

En resumen, para estimar  $\beta$  y  $\sigma_A^2$  por MV se requiere (i) minimizar numéricamente con respecto a  $\beta$  la función

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

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

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

donde  $m \equiv 1 + p + P + q + Q$  es el número de parámetros contenidos en  $\beta$ . [Nótese que la estimación (6), que no coincide con la que figura en (4), es análoga a la estimación MCO de  $\sigma^2$  en el Modelo Lineal General.]

El problema fundamental para calcular una estimación MV de  $\beta$  consiste en que (5) es una función no lineal

complicada de  $\beta$ . 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  $F(\beta)$  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  $F(\beta)$  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 es posible obtener analíticamente una expresión para  $\beta$  que minimice (5), por lo que en la práctica, para obtener una estimación MV de  $\beta$  se requiere minimizar (5) mediante algún **método de cálculo iterativo** para minimizar funciones de varias variables (como el método de *Newton - Raphson*, el de *Gauss - Newton*, el de *Levenberg - Marquardt*, o algún método *cuasi - Newton*). Las características esenciales de cualquier método de este tipo son:

- Partiendo de una **estimación inicial**  $\hat{\beta}_0$  proporcionada por el usuario, se obtiene una secuencia de estimaciones  $\hat{\beta}_1, \hat{\beta}_2, \dots$  que, bajo ciertas condiciones, converge a un **mínimo local** de (5). En general, no puede asegurarse que dicho mínimo sea un **mínimo global** de (5).
- 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 el Modelo Lineal General, e indica la posible presencia de términos redundantes (sobrantes) en el modelo (7.1).

Aunque en la actualidad no existen razones objetivas para hacerlo, en la práctica a veces se estima un modelo como (7.1) 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* de Box, Jenkins y Reinsel (1994), Capítulo 7, o la aproximación empleada por EViews, descrita en [Help](#) → [EViews Help Topics](#) ... → [Estimation Methods](#) → [Time Series Models](#)]. En teoría, todos los estimadores de  $\beta$  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, se han observado 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  $\beta$ , 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  $\beta$ ; por el contrario, si  $F(\beta)$  en  $\beta = \hat{\beta}$  tiene poca curvatura, el verdadero valor de  $\beta$  estará localizado con poca precisión (muchas varianzas).

**Observación 2 - Cálculo de residuos:** Como en el Modelo Lineal General, la serie de residuos  $(\hat{a}_t)_{t=1}^n$  asociada con la estimación del modelo (7.1) contiene las diferencias entre los valores observados y los valores ajustados para la variable dependiente  $W_t \equiv \nabla^d \nabla_S^D Y_t'$  ( $t = 1, \dots, n$ ) de (7.1). Una vez estimado (como se describe en la Observación 1 anterior), (7.1) 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

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

De manera más explícita, el modelo (7.1) 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  $W_t \equiv \nabla^d \nabla_S^D Y_t'$ , 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 (7.1) 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 del proceso estacionario ( $W_t$ )] y que  $\hat{a}_t = 0$  (la media del proceso de ruido blanco ( $A_t$ )) para  $t < 1$ .

### Práctica EV-07

Estimar algunos de los modelos identificados en la Práctica EV-06 (incluyendo, al menos, el de la serie PIB del archivo TS29-PIB-CPI.wf1).

## Diagnosis

**Entrada:** Modelo estimado; residuos.

- 3.1 Parámetros: significación (individual y conjunta); correlaciones estimadas entre estimadores; estacionariedad e invertibilidad.
- 3.2 Residuos (reemplazan a un proceso de ruido blanco en el modelo estimado): gráfico temporal; ACF y PACF residuales (Ljung - Box); media muestral; Normalidad.
- 3.3 Posible reformulación según los resultados de 3.1 y 3.2.

**Salida:** Si modelo reformulado  $\Rightarrow$  Volver a estimación. Si modelo adecuado  $\Rightarrow$  Utilizar para prever (otros fines: análisis de intervención, análisis multivariante, ...).

**Observación 1 - Diagnósis de parámetros:** Las operaciones siguientes referidas a los **parámetros** del modelo identificado y estimado, son especialmente útiles en la etapa de diagnóstico:

- Contrastar la **significación individual** de cada parámetro, o la **significación conjunta** de varios parámetros, según interés. A veces, estos contrastes indican que pueden eliminarse algunos parámetros no significativos.
- Examinar la matriz **R** de **correlaciones** estimadas entre los estimadores de los parámetros, que puede calcularse a partir de **V** como  $\mathbf{R} = \mathbf{D}^{-1}\mathbf{V}\mathbf{D}^{-1}$ , donde **D** es una matriz diagonal cuyos elementos son las raíces cuadradas de los elementos en la diagonal principal de **V**. La presencia de correlaciones elevadas suele indicar que el modelo está **sobrep parametrizado**.
- 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.

- 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 naturaleza de la tendencia implícita en el modelo reformulado).
- Cuando en la etapa de identificación no está claro si  $d$  (el número de diferencias regulares) debe ser igual a uno o igual a dos (también cuando se duda entre  $d = 1$  y  $d = 0$ ), una buena práctica consiste en escoger  $d = 2$  y añadir un factor MA(1) regular (**testigo de sobrediferenciación**) al modelo identificado con  $d = 2$ ; si el modelo estimado resulta tal que  $\hat{\theta}(x) = 0$  tiene una raíz unitaria, se reformula el modelo con  $d = 1$  y sin el factor MA(1) no invertible (reconsiderando, además, la naturaleza de la tendencia implícita en el modelo reformulado).
- Aunque en muchos casos es suficiente llevar a cabo las operaciones descritas en los dos puntos anteriores de manera informal, en ocasiones se requiere **contrastar formalmente** la presencia de **raíces unitarias** en los operadores AR y MA del modelo considerado. Una dificultad asociada con esta cuestión es que los estadísticos habituales para contrastar hipótesis sobre los parámetros no siguen distribuciones estándar (Normal o  $t$  de Student) cuando se emplean para contrastar la presencia de raíces unitarias. Ver Brockwell y Davis (2002), páginas 193 - 198; Shin, D.W., y Fuller, W.A. (1998), "Unit root tests based on unconditional maximum likelihood estimation for the autoregressive moving average," *Journal of Time Series Analysis*, 19, 591 - 599; y Davis, R.A., Chen, M., y Dunsmuir, W.T.M. (1995), "Inference for MA(1) processes with a root on or near the unit circle," *Probability and Mathematical Statistics*, 15, 227 - 242.
- Cuando en la etapa de identificación no están claros los órdenes  $p$ ,  $P$ ,  $q$  y  $Q$ , se puede llevar a cabo un ejercicio de **sobreajuste**, que consiste en (i) incluir (sin un respaldo muestral claro) de uno en uno factores AR y MA de orden

1 ó 2, (ii) estimar el modelo resultante, y (iii) comprobar si cada factor añadido es significativo.

**Observación 2 - Diagnósis de residuos:** Las operaciones siguientes referidas a los **residuos** del modelo identificado y estimado, son especialmente útiles en la etapa de diagnóstico:

- 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.
- Examinar el histograma de la serie de residuos y utilizar el estadístico de Jarque-Bera para contrastar la hipótesis de Normalidad de las perturbaciones.
- Si el modelo no contiene el parámetro  $\mu_W$ , contrastar la significación de la media de las perturbaciones utilizando la media residual; si se rechaza la hipótesis nula del contraste, añadir el parámetro  $\mu_W$  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 aproximadamente una distribución  $\chi^2(g-m)$ , donde  $m$  es el número total de parámetros estimados (excluyendo la varianza de las perturbaciones).

### 1.7.1 Definición

Para un modelo que contiene  $m$  parámetros y ha sido estimado con  $n$  observaciones, el **criterio de información de Akaike** es la cantidad

$$\text{AIC} \equiv -\frac{2 \ln L^*}{n} + \frac{2m}{n},$$

donde  $L^*$  es el valor de la función de verosimilitud asociada con el modelo estimado, evaluada en las estimaciones obtenidas de los parámetros de dicho modelo.

### 1.7.2 Definición

Para un modelo que contiene  $m$  parámetros y ha sido estimado con  $n$  observaciones, el **criterio de información bayesiano de Schwarz** es la cantidad

$$\text{BIC} \equiv -\frac{2 \ln L^*}{n} + \frac{m \ln(n)}{n},$$

donde  $L^*$  es el valor de la función de verosimilitud asociada con el modelo estimado, evaluada en las estimaciones obtenidas de los parámetros de dicho modelo.

**Observación 1:** En el contexto de la elaboración de un modelo ARIMA a partir de una serie temporal  $(y_t)$ , los criterios

de información AIC y BIC (que a veces se denomina SC) suelen emplearse para decidir sobre los órdenes  $p$ ,  $P$ ,  $q$  y  $Q$  asociados con una transformación (estacionaria) dada de  $(y_t)$ . 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).

**Observación 2:** 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, puede comprobarse que el BIC siempre selecciona un modelo más **escueto** que el AIC. De hecho, el BIC es siempre preferible al AIC con muestras grandes, aunque el AIC puede ser preferible con algunas muestras cortas. En todo caso, la selección de un modelo no debería basarse únicamente en el empleo de estos criterios (aunque pueden incorporarse al procedimiento de identificación descrito al principio de esta sección).

## Práctica EV-08

Diagnosticar y, en su caso, reformular, algunos de los modelos estimados en la Práctica EV-07 (incluyendo, al menos, el de la serie PIB del archivo TS29-PIB-CPI.wf1).

## 1.8 PREVISIÓN CON MODELOS ARIMA

---

### 1.8.1 Definición

La **previsión puntual** en **origen**  $k$  (dado) a **horizonte**  $l \geq 1$  (variable) de un proceso estocástico  $(Y_t)$ , se representa con el símbolo  $Y_k(l)$  y se define como el **valor esperado** de  $Y_{k+l}$  **condicionado** por toda la información disponible hasta un momento  $k$  dado de la historia de  $(Y_t)$ ; es decir,

$$Y_k(l) \equiv \mathbf{E}_k[Y_{k+l}] \text{ para } l = 1, 2, \dots \quad (8.1)$$

**Observación:** Considerada como una función del horizonte de previsión  $l \geq 1$ ,  $Y_k(l)$  se denomina la **función de previsión** en origen  $k$  (dado) del proceso  $(Y_t)$ . En la práctica, el origen de previsión  $k$  suele ser igual al momento  $N$  de la última observación de una serie temporal  $y_1, y_2, \dots, y_N$  empleada en la elaboración de un modelo para  $(Y_t)$ .

### 1.8.2 Definición

El **error de previsión** en origen  $k$  (dado) a horizonte  $l \geq 1$  (variable) de un proceso estocástico  $(Y_t)$ , es la variable aleatoria

$$E_k(l) \equiv Y_{k+l} - Y_k(l) \text{ para } l = 1, 2, \dots, \quad (8.2)$$

es decir, la diferencia entre el **valor observado**  $Y_{k+l}$  de  $(Y_t)$  en un momento  $k + l$  de su historia y el **valor previsto**  $Y_k(l)$  en origen  $k$  a horizonte  $l$  de  $(Y_t)$ .

## La función de previsión

### 1.8.3 Definición

La **función de previsión** en origen  $k$  a horizonte  $l \geq 1$  de un proceso  $(Y_t) \sim I(d)$  que sigue un **modelo ARIMA** del tipo  $\phi'(B)Y_t = \mu + \theta(B)A_t$  (ver Observación 1 de 1.5.1), donde  $\phi'(B) \equiv \phi(B)\nabla^d \equiv 1 - \sum_{i=1}^{p'} \phi'_i B^i$  ( $p' \equiv p + d$ ), es

$$Y_k(l) \equiv E_k[Y_{k+l}] = \mu + \sum_{i=1}^{p'} \phi'_i E_k[Y_{k+l-i}] + E_k[A_{k+l}] - \sum_{i=1}^q \theta_i E_k[A_{k+l-i}], \quad (8.3)$$

o bien [ver (5.4)]

$$Y_k(l) \equiv E_k[Y_{k+l}] = \mu_{k+l} + E_k[\psi^*(B)A_{k+l}], \quad (8.4)$$

donde  $\psi^*(B) \equiv \sum_{i=0}^{\infty} \psi_i^* B^i$  ( $\psi_0^* \equiv 1$ ) se define de acuerdo con que  $\nabla^d \psi^*(B) = \psi(B)$ .

**Observación 1:** La función de previsión (8.3) de un proceso  $(Y_t) \sim \text{ARIMA}(p, d, q)$  se utiliza en la práctica para

calcular previsiones puntuales  $\hat{y}_k(1), \hat{y}_k(2), \dots$  a partir de un modelo estimado. Para ello, en (8.3) se reemplazan los parámetros  $\mu, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$  por sus valores estimados, se reemplaza cada  $E_k[Y_j]$  por  $y_j$  (un dato) si  $j \leq k$ , o bien por  $\hat{y}_k(j-k)$  (una previsión) si  $j > k$ , y se reemplaza cada  $E_k[A_j]$  por  $\hat{a}_j$  (un residuo) si  $j \leq k$ , o bien por 0 (la esperanza incondicional de  $A_j$ ) si  $j > k$ . Los errores de previsión calculados según (8.2), asociados con las previsiones  $\hat{y}_k(1), \hat{y}_k(2), \dots$ , suelen representarse como  $\hat{e}_k(l) \equiv y_{k+l} - \hat{y}_k(l)$  ( $l = 1, 2, \dots$ ) (que sólo pueden calcularse cuando  $y_{k+1}, y_{k+2}, \dots$  son datos conocidos); por su parte, la función de previsión (8.4) se utiliza para describir las propiedades estadísticas de los errores de previsión (ver Definición 1.8.4).

**Observación 2:** Un modelo ARIMA( $p, d, q$ ) estimado puede escribirse como  $\hat{\phi}(B)\nabla^d y_t = \hat{\mu} + \hat{\theta}(B)\hat{a}_t$ , o bien como  $\hat{\phi}'(B)y_t = \hat{\mu} + \hat{\theta}(B)\hat{a}_t$ , con  $\hat{\phi}'(B) \equiv \hat{\phi}(B)\nabla^d \equiv (1 - \sum_{i=1}^p \hat{\phi}_i B^i)(1 - B)^d \equiv 1 - \sum_{i=1}^{p'} \hat{\phi}'_i B^i$ ,  $\hat{\theta}(B) \equiv 1 - \sum_{i=1}^q \hat{\theta}_i B^i$ ; en consecuencia, cada residuo  $\hat{a}_t$  de un modelo estimado puede escribirse como

$$\hat{a}_t = y_t - \left[ \hat{\mu} + \sum_{i=1}^{p'} \hat{\phi}'_i y_{t-i} - \sum_{i=1}^q \hat{\theta}_i \hat{a}_{t-i} \right],$$

donde el término entre corchetes es un **valor ajustado** que puede interpretarse como una previsión en origen  $t-1$  a horizonte 1. Por lo tanto,  $\hat{a}_t$  puede interpretarse como un error de previsión en origen  $t-1$  a horizonte 1.

### 1.8.4 Definición

El **error de previsión** en origen  $k$  a horizonte  $l \geq 1$  de un proceso  $(Y_t) \sim I(d)$  que sigue un **modelo ARIMA** del tipo  $\phi'(B)Y_t = \mu + \theta(B)A_t$  (ver 1.8.3), puede escribirse, de acuerdo con (5.4) y (8.4), como

$$E_k(l) \equiv Y_{k+l} - Y_k(l) = \psi^*(B)A_{k+l} - E_k[\psi^*(B)A_{k+l}],$$

donde  $\psi^*(B)A_{k+l} = \sum_{i=0}^{\infty} \psi_i^* A_{k+l-i} = A_{k+l} + \psi_1^* A_{k+l-1} + \psi_2^* A_{k+l-2} + \dots$ . Por lo tanto,

$$E_k(l) = A_{k+l} + \psi_1^* A_{k+l-1} + \dots + \psi_{l-1}^* A_{k+1} = \sum_{i=0}^{l-1} \psi_i^* A_{k+l-i}, \quad (8.5)$$

donde los parámetros  $\psi_i^*$  del polinomio  $\psi^*(B) \equiv \sum_{i=0}^{\infty} \psi_i^* B^i$  ( $\psi_0^* \equiv 1$ ) pueden calcularse teniendo en cuenta que  $\nabla^d \psi^*(B) = \psi(B)$ , con  $\psi(B) \equiv \theta(B) / \phi(B)$ ; ver la explicación alrededor de (3.10).

**Observación:** (8.5) implica que  $E[E_k(l)] = 0$  y que  $\text{Var}[E_k(l)] = \sigma_A^2 \sum_{i=0}^{l-1} (\psi_i^*)^2$  en cualquier modelo ARIMA con cualquier valor para  $d \geq 0$ ; en particular, si  $d = 0$ ,  $\psi^*(B) = \psi(B)$ , por lo que  $\text{Var}[E_k(l)] = \sigma_A^2 \sum_{i=0}^{l-1} \psi_i^2$  en este caso. Como  $\text{Var}[E_k(l)]$  depende en cualquier caso del horizonte  $l \geq 1$  pero **no** del origen  $k$  de previsión,  $\text{Var}[E_k(l)]$  suele escribirse simplemente como  $V(l)$ ,  $l = 1, 2, \dots$

### 1.8.5 Ejemplo – Previsión con un modelo ARMA(1,1)

El modelo ARMA(1,1), o ARIMA(1,0,1), para un proceso  $(Y_t) \sim I(0)$  es (Ejemplo 1.3.8)

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t - \theta_1 A_{t-1}.$$

Escribiendo esta expresión en  $t = N + l$  ( $l \geq 1$ ) y aplicando al resultado el operador  $E_N[\cdot]$ ,

se obtiene que la función de previsión en origen  $N$  a horizonte  $l \geq 1$  es

$$Y_N(l) = \mu + \phi_1 E_N[Y_{N+l-1}] + E_N[A_{N+l}] - \theta_1 E_N[A_{N+l-1}]. \quad (8.6)$$

Para  $l = 1$ , (8.6) queda

$$Y_N(1) = \mu + \phi_1 Y_N - \theta_1 A_N = \mu + \phi_1 \left( Y_N - \frac{\theta_1}{\phi_1} A_N \right), \quad (8.7)$$

mientras que para  $l \geq 2$ , (8.6) queda

$$Y_N(l) = \mu + \phi_1 Y_N(l-1). \quad (8.8)$$

(8.8) es una ecuación en diferencias, cuya solución a partir de la condición inicial (8.7) es

$$Y_N(l) = \left( \sum_{i=0}^{l-1} \phi_1^i \right) \mu + \phi_1^l \left( Y_N - \frac{\theta_1}{\phi_1} A_N \right) \text{ para todo } l \geq 1. \quad (8.9)$$

(8.9) se denomina la **función final de previsión** de  $(Y_t) \sim \text{ARMA}(1,1)$ , que converge al valor esperado de  $(Y_t)$ ,  $E[Y_t] = \mu / (1 - \phi_1)$ , cuando  $|\phi_1| < 1$  y  $l \rightarrow \infty$ .

**Observación:** Cada error de previsión puede calcularse (junto con su varianza) a partir de la diferencia entre cada  $Y_{N+l} = \mu + \phi_1 Y_{N+l-1} + A_{N+l} - \theta_1 A_{N+l-1}$  y la previsión (8.7) - (8.8) correspondiente, para  $l = 1, 2, 3, \dots$ :

$$E_N(1) = A_{N+1},$$

$$\begin{aligned}
 E_N(2) &= \phi_1 E_N(1) + A_{N+2} - \theta_1 A_{N+1} = A_{N+2} + (\phi_1 - \theta_1) A_{N+1}, \\
 E_N(3) &= \phi_1 E_N(2) + A_{N+3} - \theta_1 A_{N+2} = A_{N+3} + (\phi_1 - \theta_1) A_{N+2} + \phi_1 (\phi_1 - \theta_1) A_{N+1}, \\
 &\dots, \\
 E_N(l) &= \sum_{i=0}^{l-1} \psi_i A_{N+l-i}, \text{ donde } \psi_0 \equiv 1, \psi_i = \phi_1^{i-1} (\phi_1 - \theta_1) \text{ (} i \geq 1 \text{)}.
 \end{aligned}$$

Por lo tanto,

$$V(l) = \sigma_A^2 \sum_{i=0}^{l-1} \psi_i^2 = \left[ 1 + (\phi_1 - \theta_1)^2 \sum_{i=1}^{l-1} \phi_1^{2(i-1)} \right] \sigma_A^2,$$

que converge a la varianza de  $(Y_t)$ ,  $\text{Var}(Y_t) = [1 + (\phi_1 - \theta_1)^2 / (1 - \phi_1^2)] \sigma_A^2$ , cuando  $|\phi_1| < 1$  y  $l \rightarrow \infty$ .

### 1.8.6 Ejemplo – Previsión con un modelo AR(1)

El modelo AR(1), o ARIMA(1,0,0), para un proceso  $(Y_t) \sim I(0)$  es (Ejemplo 1.3.4)

$$Y_t = \mu + \phi_1 Y_{t-1} + A_t.$$

Escribiendo esta expresión en  $t = N + l$  ( $l \geq 1$ ), y aplicando al resultado el operador  $E_N[\cdot]$ , se obtiene que la función de previsión en origen  $N$  a horizonte  $l \geq 1$  es

$$Y_N(l) = \mu + \phi_1 E_N[Y_{N+l-1}] + E_N[A_{N+l}]. \quad (8.10)$$

Para  $l = 1$ , (8.10) queda

$$Y_N(1) = \mu + \phi_1 Y_N, \quad (8.11)$$

mientras que para  $l \geq 2$ , (8.10) queda

$$Y_N(l) = \mu + \phi_1 Y_N(l-1). \quad (8.12)$$

(8.12) es una ecuación en diferencias, cuya solución a partir de la condición inicial (8.11) es

$$Y_N(l) = \left( \sum_{i=0}^{l-1} \phi_1^i \right) \mu + \phi_1^l Y_N \quad \text{para todo } l \geq 1. \quad (8.13)$$

(8.13) es la **función final de previsión** de  $(Y_t) \sim \text{AR}(1)$ , que converge al valor esperado de  $(Y_t)$ ,  $E[Y_t] = \mu / (1 - \phi_1)$ , cuando  $|\phi_1| < 1$  y  $l \rightarrow \infty$ .

**Observación:** Cada error de previsión puede calcularse (junto con su varianza) a partir de la diferencia entre cada  $Y_{N+l} = \mu + \phi_1 Y_{N+l-1} + A_{N+l}$  y la previsión (8.11) - (8.12) correspondiente, para  $l = 1, 2, 3, \dots$ :

$$\begin{aligned} E_N(1) &= A_{N+1}, \\ E_N(2) &= \phi_1 E_N(1) + A_{N+2} = A_{N+2} + \phi_1 A_{N+1}, \\ E_N(3) &= \phi_1 E_N(2) + A_{N+3} = A_{N+3} + \phi_1 A_{N+2} + \phi_1^2 A_{N+1}, \\ &\dots, \\ E_N(l) &= \sum_{i=0}^{l-1} \phi_1^i A_{N+l-i}, \quad \text{donde } \psi_i = \phi_1^i \quad (i \geq 0). \end{aligned}$$

Por lo tanto,

$$V(l) = \sigma_A^2 \sum_{i=0}^{l-1} \psi_i^2 = \left( \sum_{i=0}^{l-1} \phi_1^{2i} \right) \sigma_A^2,$$

que converge a la varianza de  $(Y_t)$ ,  $\text{Var}(Y_t) = [1/(1 - \phi_1^2)]\sigma_A^2$ , cuando  $|\phi_1| < 1$  y  $l \rightarrow \infty$ .

### 1.8.7 Ejemplo – Previsión con un modelo MA(1)

El modelo MA(1), o ARIMA(0,0,1), para un proceso  $(Y_t) \sim I(0)$  es (Ejemplo 1.3.6)

$$Y_t = \mu + A_t - \theta_1 A_{t-1}.$$

Escribiendo esta expresión en  $t = N + l$  ( $l \geq 1$ ), y aplicando al resultado el operador  $E_N[\cdot]$ , se obtiene que la función de previsión en origen  $N$  a horizonte  $l \geq 1$  es

$$Y_N(l) = \mu + E_N[A_{N+l}] - \theta_1 E_N[A_{N+l-1}]. \quad (8.14)$$

Para  $l = 1$ , (8.14) queda

$$Y_N(1) = \mu - \theta_1 A_N, \quad (8.15)$$

mientras que para  $l \geq 2$ , (8.14) queda

$$Y_N(l) = \mu. \quad (8.16)$$

(8.16) es la **función final de previsión** de  $(Y_t) \sim \text{MA}(1)$ , que coincide con  $E[Y_t] = \mu$ .

**Observación 1:** Cada error de previsión puede calcularse (junto con su varianza) a partir de la diferencia entre cada  $Y_{N+l} = \mu + A_{N+l} - \theta_1 A_{N+l-1}$  y la previsión (8.15) - (8.16) correspondiente, para  $l = 1, 2, 3, \dots$ :

$$\begin{aligned} E_N(1) &= A_{N+1}, \\ E_N(2) &= A_{N+2} - \theta_1 A_{N+1}, \\ E_N(3) &= A_{N+3} - \theta_1 A_{N+2}, \\ &\dots, \\ E_N(l) &= \sum_{i=0}^{l-1} \psi_i A_{N+l-i}, \text{ donde } \psi_0 \equiv 1, \psi_1 = -\theta_1, \psi_i = 0 \text{ (} i \geq 2 \text{)}. \end{aligned}$$

Por lo tanto,

$$V(1) = \sigma_A^2, \quad V(l) = (1 + \theta_1^2) \sigma_A^2 \quad (l \geq 2),$$

que a partir de  $l = 2$  es constante y coincide con  $\text{Var}(Y_t) = (1 + \theta_1^2) \sigma_A^2$ .

**Observación 2:** En general, puede comprobarse si  $(Y_t)$  es un proceso  $I(0)$ , de manera que  $(Y_t) \sim \text{ARMA}(p, q)$  estacionario e invertible,  $Y_N(l) \rightarrow E(Y_t)$  y  $V(l) \rightarrow \text{Var}(Y_t)$  cuando el horizonte de previsión  $l$  tiende a infinito.

### 1.8.8 Ejemplo – Previsión con un modelo IMA(1,1)

El modelo IMA(1,1), o ARIMA(0,1,1), para un proceso  $(Y_t) \sim I(1)$ , puede escribirse como

$$Y_t = \mu + Y_{t-1} + A_t - \theta_1 A_{t-1}.$$

Escribiendo esta expresión en  $t = N + l$  ( $l \geq 1$ ) y aplicando al resultado el operador  $E_N[\cdot]$ ,

se obtiene que la función de previsión en origen  $N$  a horizonte  $l \geq 1$  es

$$Y_N(l) = \mu + E_N[Y_{N+l-1}] + E_N[A_{N+l}] - \theta_1 E_N[A_{N+l-1}], \quad (8.17)$$

Para  $l = 1$ , (8.17) queda

$$Y_N(1) = \mu + Y_N - \theta_1 A_N, \quad (8.18)$$

mientras que para  $l \geq 2$ , (8.17) queda

$$Y_N(l) = \mu + Y_N(l-1). \quad (8.19)$$

(8.19) es una ecuación en diferencias, cuya solución a partir de la condición inicial (8.18) es

$$Y_N(l) = l\mu + Y_N - \theta_1 A_N \quad \text{para todo } l \geq 1. \quad (8.20)$$

(8.20) es la **función final de previsión** de  $(Y_t) \sim \text{IMA}(1, 1)$ , que no depende de  $l$  si  $\mu = 0$ , pero diverge cuando  $l \rightarrow \infty$  si  $\mu \neq 0$ .

**Observación 1:** Cada error de previsión puede calcularse (junto con su varianza) a partir de la diferencia entre cada  $Y_{N+l} = \mu + Y_{N+l-1} + A_{N+l} - \theta_1 A_{N+l-1}$  y la previsión (8.18) - (8.19) correspondiente, para  $l = 1, 2, 3, \dots$ :

$$E_N(1) = A_{N+1},$$

$$\begin{aligned}
 E_N(2) &= E_N(1) + A_{N+2} - \theta_1 A_{N+1} = A_{N+2} + (1 - \theta_1)A_{N+1}, \\
 E_N(3) &= E_N(2) + A_{N+3} - \theta_1 A_{N+2} = A_{N+3} + (1 - \theta_1)A_{N+2} + (1 - \theta_1)A_{N+1}, \\
 &\dots, \\
 E_N(l) &= \sum_{i=0}^{l-1} \psi_i^* A_{N+l-i}, \text{ donde } \psi_0^* \equiv 1, \psi_i^* = (1 - \theta_1) \text{ (} i \geq 1 \text{)}.
 \end{aligned}$$

Por lo tanto,

$$V(l) = \sigma_A^2 \sum_{i=0}^{l-1} (\psi_i^*)^2 = [1 + (1 - \theta_1)^2 (l - 1)] \sigma_A^2,$$

de manera que  $V(l) \rightarrow \infty$  cuando  $l \rightarrow \infty$ .

**Observación 2:** Cuando  $(Y_t)$  es un paseo aleatorio,  $Y_t = \mu + Y_{t-1} + A_t$ , los resultados del Ejemplo (1.8.8) con  $\theta_1 = 0$  indican que  $Y_N(l) = l\mu + Y_N$  para todo  $l \geq 1$ , con  $V(l) = l\sigma_A^2$ . En general, puede comprobarse que cuando  $(Y_t)$  es un proceso  $I(d)$  con  $d \geq 1$ , de manera que  $(Y_t) \sim \text{ARIMA}(p, d, q)$  con  $d \geq 1$ , tanto  $Y_N(l)$  como  $V(l)$  son funciones que divergen cuando el horizonte de previsión  $l$  tiende a infinito; en particular,  $V(l) \rightarrow \infty$  cuando  $l \rightarrow \infty$ , lo cual contrasta claramente con los resultados de los ejemplos 1.8.5 - 1.8.7 para procesos  $I(0)$ .

## Previsiones puntuales – Intervalos de confianza

El modelo ARIMA más general considerado en la Sección 1.7,

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

puede escribirse como

$$\Phi^*(B)Y_t' = \mu + \Theta^*(B)A_t, \quad (8.21)$$

donde

$$\begin{aligned} \Phi^*(B) &\equiv \phi_p(B)\Phi_P(B^S)\nabla^d\nabla_S^D \equiv 1 - \sum_{i=1}^{p^*} \Phi_i^*B^i, \\ \Theta^*(B) &\equiv \theta_q(B)\Theta_Q(B^S) \equiv 1 - \sum_{i=1}^{q^*} \Theta_i^*B^i, \end{aligned}$$

$\mu \equiv [\phi_p(1)\Phi_P(1)]\mu_W$ ,  $p^* \equiv (p+d) + (P+D)S$ ,  $q^* \equiv q + QS$ . Por lo tanto, (8.21) queda

$$Y_t' = \mu + \sum_{i=1}^{p^*} \Phi_i Y_{t-i}' + A_t - \sum_{i=1}^{q^*} \Theta_i^* A_{t-i}. \quad (8.22)$$

**Cálculo de previsiones puntuales para la serie transformada:**

De acuerdo con las Definiciones 1.8.1 y 1.8.3, el modelo (8.22) para  $(Y_t')$  implica que

$$\hat{y}'_N(l) = \hat{\mu} + \sum_{i=1}^{p^*} \hat{\Phi}_i [y'_{N+l-i}] - \sum_{i=1}^{q^*} \hat{\Theta}_i^* [a_{N+l-i}] \quad (l \geq 1), \quad (8.23)$$

donde

$$[y'_j] \equiv \begin{cases} y'_j & , \text{ si } j \leq N, \\ \hat{y}'_N(j-N) & , \text{ si } j > N. \end{cases} \quad [a_j] \equiv \begin{cases} \hat{a}_j \equiv y'_j - \hat{y}'_{j-1}(1) & , \text{ si } j \leq N, \\ 0 & , \text{ si } j > N. \end{cases}$$

Cálculo de intervalos de confianza para la serie transformada:

$$IC'_{1-\alpha}(l) = [\hat{y}'_{N+l}(\mp)] = [\hat{y}'_N(l) \mp u \times \sqrt{\hat{v}(l)}], \quad (8.24)$$

donde  $u = 1.65, 1.96$  ó  $2.58$ , según si  $\alpha = 0.10, 0.05$  ó  $0.01$ ,  $\hat{v}(l) \equiv \hat{\sigma}_A^2 \sum_{i=0}^{l-1} (\hat{\psi}_i^*)^2$  (ver la Definición 1.8.4), y

$$\hat{\psi}_i^* = -\hat{\Theta}_i^* + \sum_{j=1}^{p^*} \hat{\Phi}_j^* \hat{\psi}_{i-j}^* \quad (i \geq 1), \quad (8.25)$$

con  $\hat{\psi}_0^* \equiv 1$ ,  $\hat{\psi}_i^* = 0$  si  $i < 0$ , y  $\hat{\Theta}_i^* = 0$  si  $i > q^*$ .

Cálculo de previsiones puntuales e intervalos de confianza para la serie original:

$$\hat{y}_N(l) \equiv \begin{cases} \exp[\hat{y}'_N(l)], & \text{si } \lambda = 0, \\ [\hat{y}'_N(l)]^{\frac{1}{\lambda}}, & \text{si } \lambda \neq 0, \end{cases} \quad \hat{y}_{N+l}(\mp) \equiv \begin{cases} \exp[\hat{y}'_{N+l}(\mp)], & \text{si } \lambda = 0, \\ [\hat{y}'_{N+l}(\mp)]^{\frac{1}{\lambda}}, & \text{si } \lambda \neq 0, \end{cases} \quad (8.26)$$

donde  $\lambda$  es el parámetro de la transformación Box - Cox empleada (si el parámetro  $m$  en dicha transformación es distinto de cero, se resta  $m$  en las cuatro expresiones anteriores).

### 1.8.9 Ejemplo

Modelo estimado con EViews 3.1 (Jul 11 2001 build) para el proceso **PIB real en España**,  $(Y_t)$ , con datos anuales (serie PIB del archivo TS29-PIB-CPI.wf1) desde 1964 hasta 1997 (34 observaciones):

$$\begin{aligned} (1 - 0.6747B) [\nabla \ln y_t - 0.0311] &= \hat{a}_t, \\ (0.1302) \qquad \qquad \qquad (0.0096) & \\ n = 32, \hat{\sigma}_A &= 0.0176. \end{aligned} \tag{8.27}$$

(8.27) también puede escribirse como

$$w_t = 0.0101 + 0.6747w_{t-1} + \hat{a}_t, \tag{8.28}$$

donde  $w_t \equiv \nabla \ln y_t$ , o bien como

$$y'_t = 0.0101 + 1.6747y'_{t-1} - 0.6747y'_{t-2} + \hat{a}_t, \tag{8.29}$$

donde  $y'_t \equiv \ln y_t$ , dependiendo de si se quiere prever la tasa logarítmica de variación anual del PIB real,  $(W_t) \equiv (\nabla \ln Y_t)$ , o bien el PIB real directamente,  $(Y_t)$ .

(8.28) es un modelo AR(1) estimado para  $(W_t) \equiv (\nabla \ln Y_t)$ ; por lo tanto (ver Ejemplo 1.8.6),  $\hat{w}_{32}(1) = 0.0101 + 0.6747 w_{32} = 0.0329$ ,  $\hat{w}_{32}(2) = 0.0101 + 0.6747 \hat{w}_{32}(1) = 0.0323$ , ..., que converge a la media estimada de  $(W_t) \equiv (\nabla \ln Y_t)$ ,  $\hat{\mu}_W = 0.0311$  (un 3.11%).

Por su parte, (8.29) es un modelo ARI(1,1) estimado para  $(Y'_t) \equiv (\ln Y_t)$ ; por lo tanto:

$$\hat{y}_{32}(1) = \exp [ 0.0101 + 1.6747 y'_{32} - 0.6747 y'_{31} ],$$

$$\hat{y}_{32}(2) = \exp [ 0.0101 + 1.6747 \hat{y}'_{32}(1) - 0.6747 y'_{32} ],$$

$$\hat{y}_{32}(3) = \exp [ 0.0101 + 1.6747 \hat{y}'_{32}(2) - 0.6747 \hat{y}'_{32}(1) ],$$

...

que es una secuencia creciente (no convergente) con el horizonte de previsión.

### Práctica EV-09

Utilizar algunos de los modelos elaborados en la Sección 1.7 para calcular previsiones.

## Criterios de evaluación de previsiones

### 1.8.10 Definición

Sea  $\hat{y}_N(1), \hat{y}_N(2), \dots, \hat{y}_N(L)$  una secuencia de previsiones calculadas en origen  $N$  a horizontes  $1, 2, \dots, L$ . La **raíz del error cuadrático medio** (**RMSE**, del inglés *Root Mean Squared Error*) asociada con dicha secuencia es

$$RMSE \equiv \sqrt{\frac{1}{L} \sum_{l=1}^L [\hat{y}_N(l) - y_{N+l}]^2} \equiv \sqrt{\frac{1}{L} \sum_{l=1}^L \hat{e}_N(l)^2},$$

es decir, la desviación típica muestral de los errores de previsión  $\hat{e}_N(1), \hat{e}_N(2), \dots, \hat{e}_N(L)$ .

### 1.8.11 Definición

Sea  $\hat{y}_N(1), \hat{y}_N(2), \dots, \hat{y}_N(L)$  una secuencia de previsiones calculadas en origen  $N$  a horizontes  $1, 2, \dots, L$ . El **error absoluto medio** (**MAE**, del inglés *Mean Absolute Error*) asociado con dicha secuencia es

$$MAE \equiv \frac{1}{L} \sum_{l=1}^L |\hat{y}_N(l) - y_{N+l}| \equiv \frac{1}{L} \sum_{l=1}^L |\hat{e}_N(l)|,$$

es decir, la media muestral de los valores absolutos de los errores de previsión.



$$\nabla\nabla_{12} \ln y_t = \underset{(0.0696)}{(1 - 0.6771B)} \underset{(0.0002)}{(1 - 0.8857B^{12})} \hat{a}_t, \quad (8.31)$$

$$n = 119, \hat{\sigma}_A = 0.1371, \text{AIC} = -1.1190, \text{BIC} = -1.0723.$$

Por ejemplo, teniendo en cuenta que  $\nabla\nabla_{12} \equiv (1 - B)(1 - B^{12}) = 1 - B - B^{12} + B^{13}$ , el modelo estimado (8.31) queda

$$y'_t = y'_{t-1} - y'_{t-12} + y'_{t-13} + \hat{a}_t - 0.6771\hat{a}_{t-1} - 0.8857\hat{a}_{t-12} + 0.5997\hat{a}_{t-13}, \quad (8.32)$$

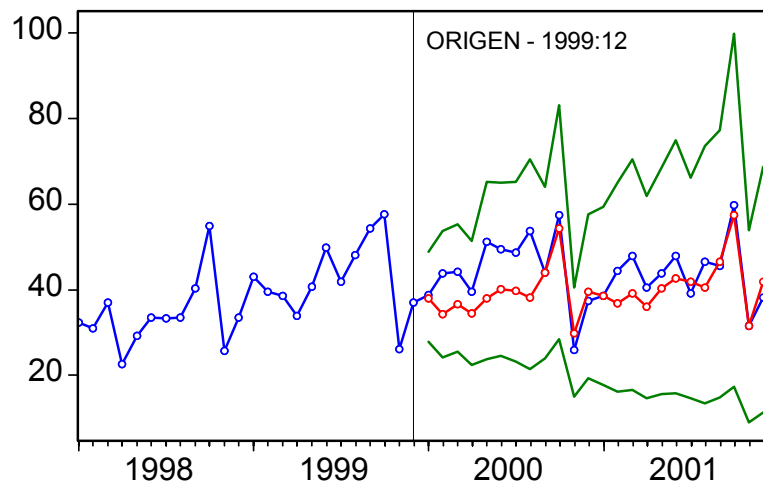
donde  $y'_t \equiv \ln y_t$ . A partir de (8.32), las previsiones puntuales del proceso original ( $Y_t$ ), calculadas en origen  $n = 119$  (1999:12) a horizontes 1, 2, ..., 24, son:

$$\hat{y}_{119}(1) = \exp [ y'_{119} - y'_{108} + y'_{107} - 0.6771\hat{a}_{119} - 0.8857\hat{a}_{108} + 0.5997\hat{a}_{107} ] = 38.524,$$

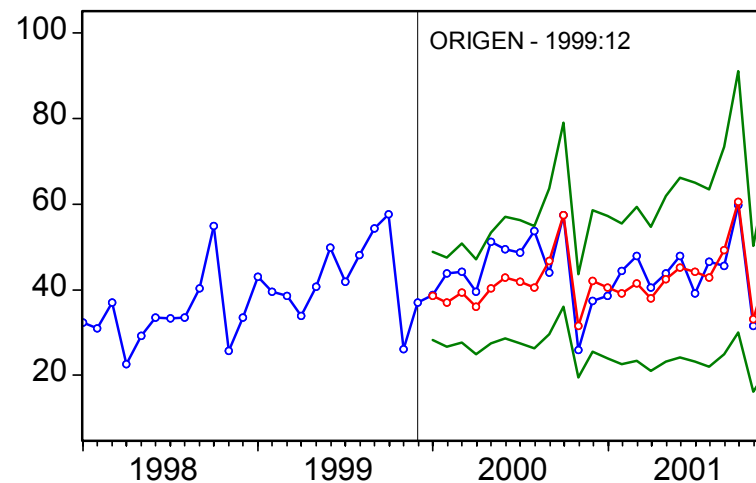
$$\hat{y}_{119}(2) = \exp [ \hat{y}'_{119}(1) - y'_{109} + y'_{108} - 0.8857\hat{a}_{109} + 0.5997\hat{a}_{108} ] = 37.083,$$

...

$$\hat{y}_{119}(24) = \exp [ \hat{y}'_{119}(23) - \hat{y}'_{119}(12) + \hat{y}'_{119}(11) ] = 44.286.$$



PREVISIONES CALCULADAS CON EL MODELO (8.30)  
 RMSE = 6.586 - MAE = 5.206 - MAPE = 11.44%



PREVISIONES CALCULADAS CON EL MODELO (8.31)  
 RMSE = 5.405 - MAE = 4.461 - MAPE = 10.27%

**Figura 8.1 – Previsiones para la serie VIVIN con dos modelos alternativos**

## Práctica EV-10

Realizar las operaciones a las que se refiere el Ejemplo 1.8.13 anterior.