

Caja 19

EZEQUIEL URIEL

Catedrático de Econometría
y Métodos Estadísticos
Universidad Literaria de Valencia

ANÁLISIS DE
SERIES TEMPORALES
modelos ARIMA

1985



COLECCIÓN
ábaco

PARANINFO

SA

MADRID



Prefacio

Este libro constituye una introducción al análisis de series temporales utilizando el enfoque de los modelos ARIMA. Dicho enfoque se ha popularizado a partir de 1970 con el libro de Box y Jenkins «Time Series Analysis: Forecasting y Control», de forma que actualmente se consideran prácticamente sinónimas las expresiones *modelos ARIMA* y *enfoque Box-Jenkins*. La finalidad última de la construcción de los modelos ARIMA es la predicción y, más concretamente, la predicción a corto plazo, habiéndose obtenido resultados satisfactorios en distintas áreas tales como la macroeconomía, la economía empresarial, la sociología, los procesos tecnológicos, etc.

En mi opinión, el estudio teórico de los modelos ARIMA debe complementarse con su aplicación práctica a series reales, ya que esta aplicación es la que permite profundizar en el proceso de inferencia con este tipo de modelos. La elaboración de modelos ARIMA es impensable sin la ayuda de un ordenador dada la laboriosidad de los cálculos. Hace muy pocos años tan sólo, el acceso a un ordenador constituyía una barrera importante para muchas personas. En la actualidad el panorama ha cambiado sustancialmente con la introducción en el mercado de los ordenadores personales, para muchos de los cuales se han elaborado programas que permiten instrumentar la metodología ARIMA. Se encuentra disponible un paquete de programas en lenguaje BASIC diseñado especialmente por el autor para microordenadores.

No obstante, es importante señalar que la aplicación mecánica de la metodología ARIMA puede conducir a la elaboración de modelos inadecuados que, en muchos casos, carecerán de sentido. Por otra parte, la comprensión de los modelos ARIMA a un nivel suficiente para su aplicación práctica no requiere grandes conocimientos de carácter estadístico o matemático. En este sentido «Análisis de series temporales: modelos ARIMA» se sitúa a un nivel introductorio, aunque al mismo tiempo se profundiza en los mecanismos de generación de las series temporales. Para obtener provecho de su lectura únicamente se requiere tener conocimientos básicos de cálculo, álgebra y estadística.

Se expone a continuación la estructura del libro. El primer capítulo proporciona una panorámica de los distintos métodos de previsión, situando el enfoque ARIMA dentro de este contexto. Los capítulos 2 y 3, que están dedicados a los procesos estocásticos y a los modelos lineales respectivamente, suministran la base estadística imprescindible para acometer la elaboración de los modelos ARIMA. En capítulos 4 a 7 se examinan las diferentes fases de la elaboración de dichos modelos. Se ha dedicado el capítulo 8 a los modelos estacionales, ya que revisten gran importancia en la prá-

© EZEQUIEL URIEL JIMENEZ
Valencia, España, 1985

Reservados los derechos de
edición, reproducción o adaptación

IMPRESO EN ESPAÑA
PRINTED IN SPAIN

ISBN: 84-283-1398-9
Depósito legal: M. 22.844.—1985

PARANINFO S.A. Magallanes, 25 - 28015 MADRID

(3-118)

ARTES GRÁFICAS BENZAL, S. A. Virtudes, 7. 28010 MADRID

PREFACIO

tica, pues la mayor parte de las series que se utilizan en economía tienen un claro componente estacional.

A continuación de los capítulos 3, 4, 5, 6, 7 y 8 se proponen algunos ejercicios cuyas soluciones aparecen al final del libro. En series reales o en series simuladas largas no se ha indicado ninguna solución precisamente porque en la elaboración práctica de modelos ARIMA existe casi siempre una dosis de ambigüedad que impide proponer una única solución. Ello no es, por supuesto, un obstáculo para construir un modelo razonable que sea coherente con los datos.

El libro se completa con cuatro apéndices. En el primer apéndice se expone el concepto del operador de retardos, que es un instrumento matemático que permite simplificar desde el punto de vista algebraico la manipulación de los procesos estocásticos. En el segundo apéndice se hace una exposición de las ecuaciones en diferencias finitas, ya que a nuestro juicio el conocimiento de estas ecuaciones es esencial para la comprensión de las características y propiedades de los modelos ARIMA. El tercer apéndice está dedicado a la simulación de modelos por el método de Montecarlo que resulta muy útil para comprender cómo se generan los datos con un modelo estocástico. Por último, en el cuarto apéndice aparecen series de datos reales y series simuladas que han sido utilizadas para proponer ejercicios y que pueden ser de gran utilidad para realizar prácticas en un ordenador.

En la elaboración de este libro he contado con la colaboración de muchas personas. Así, he recibido valiosas sugerencias de numerosos colegas que han permitido mejorar sustancialmente su contenido. Particularmente, quiero expresar mi agradecimiento por sus aportaciones a los profesores José Luis Raymond, Juan Muro y María Luisa Moltó y al matemático Joaquín Asiaín.

En la mecanografía de los originales deseo agradecer el esfuerzo realizado por Reyes Gomar, Teresa Asiaín y Javier Ferrer, dado el tipo de materia y mi pésima caligrafía. En la elaboración de los gráficos han prestado una importante colaboración Miguel Angel García, José Gutiérrez y Francisco Hontoria. Finalmente, en la preparación de la presente edición, por parte de editorial Paraninfo, he contado con los consejos y eficiente ayuda de Felisa Cuevas y Miguel Mangada.

En cualquier caso, los errores que pueda haber a lo largo de este libro son de mi absoluta responsabilidad.

Valencia, abril de 1985

Índice

1. LA PREVISIÓN EN EL ÁREA ECONÓMICA	11
1.1. Necesidad de previsión en economía	11
1.2. Clasificación de los métodos de previsión	12
2. PROCESOS ESTOCÁSTICOS	20
2.1. Introducción	20
2.2. Concepto de proceso estocástico	24
2.3. Procesos estacionarios y ergódicos	27
2.4. Procesos lineales	29
3. MODELOS LINEALES	32
3.1. Modelos autorregresivos (AR)	32
3.2. Modelos de medias móviles (MA)	44
3.3. Modelos mixtos autorregresivos—medias móviles (ARMA)	53
3.4. Procesos no estacionarios	60
Ejercicios	64
4. ELABORACIÓN DE MODELOS ARIMA: FASE DE IDENTIFICACIÓN	69
4.1. Metodología de elaboración de modelos ARIMA	69
4.2. Identificación de modelos estacionarios	73
4.3. Identificación de un modelo ARMA(p, q)	88
4.4. Análisis de estacionariedad	93
4.5. Identificación del término independiente	98
Ejercicios	99
5. ESTIMACIÓN DE UN MODELO ARMA	109
5.1. Planteamiento	109
5.2. Enfoque condicional	111
5.3. Enfoque no condicional	120
Ejercicios	123
6. VALIDACIÓN	126
6.1. Introducción	126
6.2. Análisis de los residuos	127
6.3. Análisis de los coeficientes estimados	132
6.4. Introducción de parámetros adicionales en el modelo	133

ÍNDICE

6.5. Bondad del ajuste	134
6.6. Análisis de estabilidad	137
6.7. Reformulación del modelo	137
Ejercicios	139
7. PREDICCIÓN	143
7.1. Introducción	143
7.2. Formas alternativas de presentar un modelo ARIMA	143
7.3. El predictor óptimo	145
7.4. Cálculo del predictor	147
7.5. Contraste de la estabilidad estructural en la predicción	150
Ejercicios	151
8. MODELOS ESTACIONALES	153
8.1. Introducción	153
8.2. Modelos estacionales puros y estacionarios	153
8.3. Modelos estacionales multiplicativos estacionarios	159
8.4. Modelos estacionales no estacionarios	165
8.5. Elaboración de un modelo ARIMA estacional	167
Ejercicios	169

A mi hija Mónica

APÉNDICES

A 1. Operador de diferencias y de retardos	177
A 1.1. Operador de diferencias Δ	177
A 1.2. Operador de retardos L	178
A 2. Ecuaciones ordinarias en diferencias finitas: lineales y con coeficientes constantes	180
A 2.1. Solución general de la ecuación homogénea	182
A 2.2. Solución particular	191
A 3. Simulación con modelos ARIMA por el método de Montecarlo	197
A 3.1. Generación de números aleatorios uniformes	198
A 3.2. Generación de números aleatorios de distribuciones no uniformes	201
A 3.3. Generación de números aleatorios de la distribución normal	202
A 3.4. Generación de series con modelos ARIMA	203
A 4. Datos	206
SOLUCIONES A EJERCICIOS	215
REFERENCIAS BIBLIOGRÁFICAS	259
ÍNDICE ALFABÉTICO DE MATERIAS	266

CAPÍTULO 1

La previsión en el área económica

1.1. Necesidad de previsión en economía

Tanto en economía de la empresa como en el campo macroeconómico, se plantea el problema de la toma de decisiones, es decir, la elección de una opción entre diversas alternativas. Cada opción dará lugar a un resultado distinto que puede ser medido en términos de utilidad, coste, beneficio, o cualquier otra magnitud, dependiendo del problema que se esté considerando. Pero en definitiva y en el caso general se podría medir en términos de utilidad o desutilidad. Ahora bien, el resultado concreto que se obtenga dependerá de situaciones que se puedan producir fuera del ámbito de influencia del decisor.

En el cuadro 1.1. queda reflejada esta problemática, con n opciones posibles para el decisor (d_i), p sucesos excluyentes (S_j) y $n \times p$ resultados distintos, a los que hemos designado por U_{ij} .

Cuadro 1.1.

sucesos decisiones \	S_1	S_2	...	S_p
d_1	U_{11}	U_{12}		U_{1p}
d_2	U_{21}	U_{22}		U_{2p}
:				
d_n	U_{n1}	U_{n2}		U_{np}

Cuando se toman decisiones, el decisor se encuentra generalmente en ambiente de incertidumbre respecto a los sucesos que se pueden producir en el futuro. Cada decisor tiene asociada una ganancia, o pérdida, que se debe cuantificar de acuerdo con la utilidad o desutilidad del resultado. Así, si se produce el suceso S_1 y se ha tomado la decisión d_1 , se tendrá una ganancia o pérdida a la que se ha designado por U_{11} .

Así pues, el problema con el que se enfrenta el decisor es elegir entre decisiones alternativas, teniendo en cuenta la utilidad de sus decisiones ante cada uno de los su-

cesos posibles. Estos sucesos son hechos, situados generalmente en el futuro, o que el decisor desconoce. En cualquier caso, el decisor podrá lograr unos mejores resultados si en alguna medida logra reducir la incertidumbre sobre los sucesos situados en el futuro¹. Al reducir la incertidumbre sobre el futuro, van dirigidas precisamente las técnicas de previsión.

Con las técnicas de previsión se trata, pues, de hacer pronósticos lo más acertadamente posible sobre sucesos que todavía no han tenido lugar. Para hacer estas previsiones una pitonisa utiliza como herramienta una bola de cristal, aunque implícitamente puede tener en cuenta la historia pasada. Por el contrario, las previsiones que realiza el económetra o el estadístico están basadas en un análisis explícito de la información proporcionada por los sucesos ocurridos en un pasado más o menos inmediato.

En el epígrafe siguiente se hace una exposición de los diferentes métodos de previsión, lo que permitirá situar en este contexto la elaboración de los modelos ARIMA cuyo tratamiento en profundidad se realizará en los capítulos siguientes.

1.2. Clasificación de los métodos de previsión

Desde un punto de vista metodológico, los métodos de previsión se pueden agrupar, según puede verse en la figura 1.1., en dos grandes bloques: métodos cualitativos o tecnológicos y métodos cuantitativos.

Básicamente los métodos cualitativos se utilizan en aquellos casos en los que el pasado no proporciona una información directa sobre el fenómeno considerado, como ocurre en el caso de la aparición de nuevos productos en el mercado. Así, si se pretende efectuar la previsión de ventas de televisores tridimensionales (sin gasas), no se puede recurrir a la información sobre el volumen de ventas en el pasado, ya que este producto no existe todavía en el mercado. También los métodos cualitativos pueden ser aplicados a investigaciones de tipo político o sociológico.

En la predicción de carácter cualitativo o tecnológico, los métodos estadísticos juegan un papel relativamente secundario. En estos casos lo más importante es contar con un grupo de expertos, dotados de intuición y sagacidad, que tengan un buen conocimiento tecnológico o especializado acerca del fenómeno cuya proyección en el futuro se trata de analizar. En este contexto, los métodos estadísticos serían utilizables en la organización y sistematización de las opiniones de los expertos.

Entre los múltiples métodos de carácter cualitativo, se van a exponer brevemente los tres siguientes: «brainstorming», Delphi y cross-impact. Los métodos anteriores están ordenados de menor a mayor exigencia de formalismo matemático.

En el método «brainstorming» —literalmente tormenta de cerebros— la previsión se efectúa a partir de la discusión entre un grupo de expertos donde se crea un determinado ambiente para facilitar que afloren nuevas ideas. La aplicación de este método no requiere prácticamente la utilización de instrumentos estadísticos.

En el método Delphi, cuya denominación procede del famoso oráculo griego, la

¹ Para una introducción al tema de la teoría de la decisión puede verse Lindley (1971).

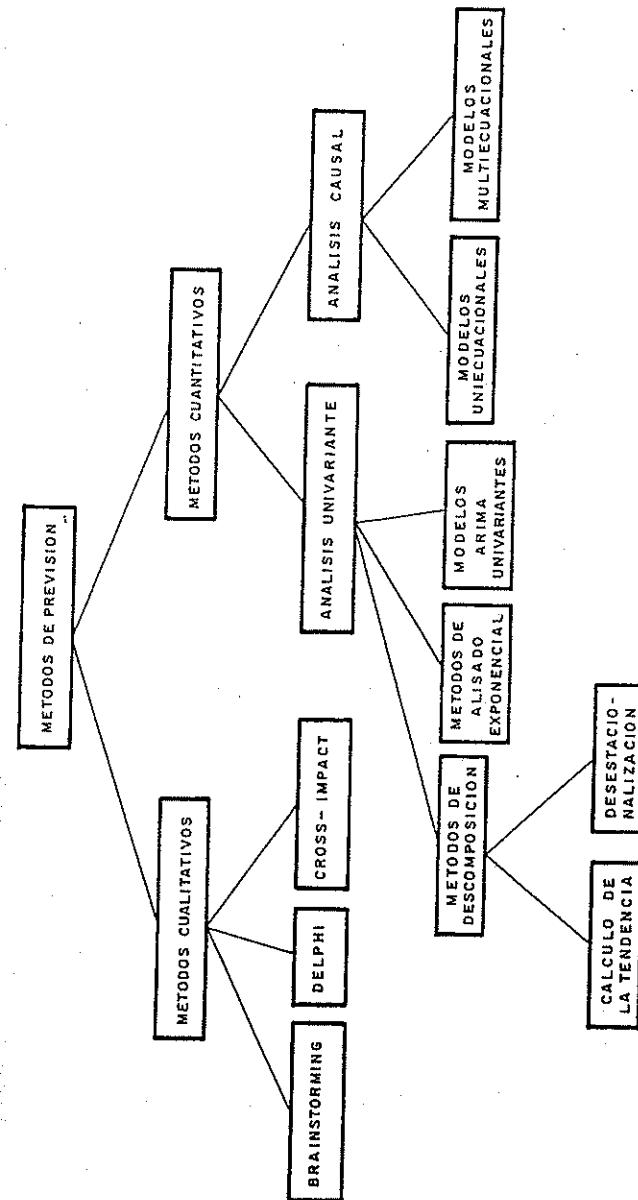


Fig. 1.1.

previsión está basada en la utilización sistemática e iterativa de juicios de opinión de un grupo de expertos hasta llegar a un acuerdo. En este proceso se trata de evitar las influencias de individuos o grupos dominantes, y al mismo tiempo que exista una realimentación de forma que se facilite el acuerdo final.

El método Delphi fue desarrollado originalmente en la Rand Corporation, debiendo destacarse las aportaciones de Helmer y Rescher (1959), Helmer (1966 a) y Helmer (1966 b). Una evaluación de los resultados que se han obtenido con su aplicación puede verse en Ament (1970), Sackman (1975) y Lindstone y Turoff (1975).

El cross-impact, o impacto cruzado, es un método que tiene una mayor exigencia de procedimientos estadísticos. Este método es utilizado para evaluar la función de distribución de escenarios de anticipación, en los que se tienen en cuenta simultáneamente diversos sucesos. En el método Delphi se evalúan de forma individual uno o varios sucesos; en cambio la característica del método cross-impact es el tratamiento conjunto, y por tanto el impacto cruzado de los sucesos que configuran un escenario determinado exigiendo evaluaciones subjetivas de probabilidades de ocurrencia, tanto simples como condicionales.

Una de las primeras aplicaciones del método cross-impact fue realizada por Gordon y Hayward (1968), pudiendo verse una descripción del mismo en Rochberg, Gordon y Helmer (1970). Sobre los problemas que plantea su aplicación práctica es interesante la crítica de Amara (1972).

En las previsiones de carácter cuantitativo, se parte del supuesto de que se tiene registrada información sobre el pasado acerca del fenómeno que se quiere estudiar. Generalmente la información sobre el pasado aparece en forma de series temporales. De forma descriptiva, puede decirse que una serie temporal² consiste en un conjunto de observaciones acerca de una variable n , observada a intervalos regulares de tiempo. Ejemplos de series temporales son las disponibilidades líquidas observadas día a día, o también los índices de precios de consumo que elabora mensualmente el Instituto Nacional de Estadística.

En los métodos cuantitativos la misión del estadístico consiste en extraer toda la información posible contenida en los datos, y en base al patrón de conducta seguida en el pasado, realizar conjeturas sobre el futuro. Un punto muy importante es la cantidad de información contenida en una serie temporal, según puede verse en dos ejemplos contrapuestos.

En primer lugar se considerará una serie temporal constituida por las observaciones sobre la hora de salida del sol de todos los días durante 50 años. El problema sería predecir a qué hora va a salir el sol mañana. Con una serie de este tipo es muy fácil realizar esta predicción. Bien es cierto que, en lugar de utilizar las observaciones del pasado, se puede realizar también la predicción en base al conocimiento de las leyes sobre el movimiento de los astros. En cualquier caso, y es lo que interesa recalcar aquí, en un fenómeno de este tipo a partir de la observación del pasado se puede hacer una buena predicción de cuál va a ser el futuro. ¿Por qué sucede esto? Sencillamente porque la serie contiene mucha información siendo los valores pasados de gran utilidad para predecir el futuro.

² En el capítulo 2 se analizará con mayor profundidad el tema de las series temporales.

Como segundo ejemplo, puede considerarse una serie temporal en la que se ha registrado el primer premio de la lotería de los sorteos que se han celebrado en los últimos 50 años. Con esta información disponible se presenta el problema de predecir cuál va a ser el primer premio en el próximo sorteo de lotería. Aun siendo del mismo tamaño —supóngase que los sorteos son diarios— y habiéndose registrado formalmente de la misma manera, está claro que la serie del primer premio de lotería es una serie que contiene poca información. En realidad, esta información no sirve absolutamente de nada para hacer la predicción del futuro, suponiendo que se cumplen todas las normas y principios que debe seguir un sorteo aleatorio.

En definitiva, se han puesto dos casos extremos de series temporales. En el primer caso, la variable observada tiene un esquema o patrón de comportamiento fijo, mientras que en segundo caso no existe ningún patrón de comportamiento.

En términos estadísticos se dice que la primera serie es determinista, mientras que la segunda es puramente aleatoria.

En general, las series económicas son series que contienen componentes deterministas y componentes aleatorios. Si designamos por D_t al componente determinista y por N_t al componente aleatorio, una serie temporal Y_t podría expresarse de la siguiente forma:

$$Y_t = D_t + N_t \quad (I-1)$$

En la expresión anterior los dos componentes están relacionados de forma aditiva, aunque pueden concebirse otras formas distintas en la integración de ambos componentes.

Una vez planteado el problema, con los métodos de previsión cuantitativos se pretende conocer los componentes subyacentes de una serie y su forma de integración, con objeto de realizar previsiones del futuro. Respecto a estos métodos, se pueden considerar, según puede verse en la figura I.1., dos grandes enfoques alternativos a los que denominaremos análisis univariante y análisis causal.

En el análisis univariante, a diferencia del análisis causal, se trata de hacer previsiones de valores futuros de una variable, utilizando como información únicamente la contenida en los valores pasados de la serie temporal que mide la evolución de la variable objeto de estudio.

En el análisis univariante, en una primera aproximación, se pueden considerar tres grandes grupos: métodos de descomposición, métodos de alisado exponencial y modelos ARIMA univariantes.

En los métodos de descomposición se parte de que el patrón o esquema de generación de una serie temporal se puede descomponer en varios subesquemas. Generalmente se distinguen los siguientes componentes: tendencia, factor cíclico, movimiento estacional y movimiento irregular³.

La tendencia refleja las variaciones a largo plazo. El factor cíclico consiste en variaciones superiores al año que no son estrictamente periódicas. Por ejemplo, en eco-

³ Sobre una introducción a los métodos de descomposición puede consultarse Uriel y Casado (1981).

nomía se manejan unos ciclos cortos con una duración aproximada de 30 meses, aunque unas veces pueden ser de 36 meses y otras de 28.

Los movimientos estacionales se caracterizan por tener una periodicidad de naturaleza fija, aunque la amplitud puede ser variable. Así, por ejemplo, el índice de producción industrial en España experimenta un descenso considerable todos los años en el mes de agosto.

Finalmente, el movimiento irregular sería el componente de la serie no sujeto a ninguna periodicidad en el tiempo. No debe confundirse con la parte aleatoria, ya que no coincide conceptualmente con la misma. El movimiento irregular no está sujeto a ninguna periodicidad, mientras que el componente aleatorio se caracteriza por no ser previsible. Puede pensarse, por ejemplo, que las disponibilidades líquidas sufren en un momento dado una alteración importante como consecuencia de una medida tomada por el Banco de España. Éste sería un movimiento irregular, pero no aleatorio, puesto que en alguna forma se puede predecir de antemano el resultado de una actuación de este tipo; en definitiva, sería considerada aleatoria aquella parte de los movimientos irregulares que no es predecible.

La serie Y_t —que es la observada por el estadístico—, en el contexto de los métodos de descomposición, se puede obtener de la siguiente forma:

$$Y_t = T_t \cdot C_t \cdot E_t \cdot I_t \quad (1-2)$$

donde:

C_t = Tendencia

C_t = Factor cíclico

E_t = Movimiento estacional

I_t = Movimiento irregular.

El esquema de integración de los cuatro componentes para formar la serie que se refleja en (1-2) es multiplicativo.

Otra alternativa sería obtener la serie mediante suma de los cuatro componentes citados, es decir:

$$Y_t = T_t + C_t + E_t + I_t \quad (1-3)$$

En los métodos de descomposición, se parte de un esquema previo y se trata de aislar cada uno de los componentes. Particularmente, la tendencia y la estacionalidad son los componentes que presentan menos dificultades para su aislamiento. En lo que se refiere al cálculo de la tendencia, un procedimiento consiste en ajustar una curva a los datos mediante una ecuación matemática y después hacer proyecciones en el futuro mediante esta ecuación. Las funciones matemáticas más empleadas son la recta, parábola, función exponencial, curva logística, curva de Gompertz, etc.

Los métodos de desestacionalización tienen por objeto aislar el factor estacional: E_t . Una vez aislado, si en la identidad (1-2) se dividen ambos miembros por E_t , se obtiene el siguiente resultado:

$$Y_{ot} = \frac{Y_t}{E_t} = T_t \times C_t \times I_t \quad \forall t \quad (1-4)$$

A la serie formada por los valores sucesivos de Y_{ot} se le denomina serie desestacionalizada.

Entre los métodos de desestacionalización debe destacarse el X-11 desarrollado en el Bureau of Census del Departamento de Comercio de los Estados Unidos, por Shiskin y otros (1967). La aplicación del método X-11 requiere la utilización de un ordenador, dada la complejidad de los cálculos. Este método es el más extendido en la actualidad y, entre otras instituciones, está siendo utilizado por el banco de España. Posteriormente Estela Dagum (1975, 1978) ha introducido ciertos refinamientos en el método, basados precisamente en el empleo del enfoque ARIMA.

Cuando se ajusta una función matemática para el cálculo de la tendencia se dice que se ha estimado una tendencia *global*. Con este término se indica que la función matemática empleada es válida para todos los momentos del tiempo y que los parámetros permanecen constantes. Así, por ejemplo, si en un caso una recta refleja adecuadamente la tendencia y el esquema es aditivo, la serie temporal se puede expresar de la siguiente forma:

$$Y_t = \beta_0 + \beta_1 t + u_t \quad (1-5)$$

donde u_t recogería el resto de los componentes.

Si se ajusta el modelo anterior por mínimos cuadrados se obtendría la siguiente función estimada:

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 t \quad (1-6)$$

que constituirá una representación de la tendencia.

La ecuación (1-6) proporciona una tendencia global porque $\hat{\beta}_0$ y $\hat{\beta}_1$ tienen la consideración de constantes y porque se supone en principio que la función es válida para todo t .

Los métodos de alisado exponencial permiten también calcular los valores de la tendencia, pero a diferencia de la tendencia de carácter *global*, en cada punto se hace una revisión en función de las observaciones más recientes. Por esta razón se dice que los métodos de alisado exponencial proporcionan una tendencia de carácter local.

Algunos métodos de alisado exponencial se pueden contemplar también como casos particulares del método de mínimos cuadrados ponderados: weighted (discounted) least squares, en terminología anglosajona. En este método, considerando el modelo siguiente, en el que $f(t)$ es una función de tiempo,

$$Y_t = f(t) + u_t$$

se trata de minimizar la expresión

$$\sum w_t [y_t - f(t)]^2 \quad 0 < w_t < 1 \quad (1-7)$$

seleccionando los pesos w_t de forma que tengan un decrecimiento exponencial. Así, las observaciones más recientes tendrían un mayor peso en la determinación de la tendencia correspondiente al momento actual.

Entre los autores que han desarrollado los métodos de alisado en sus distintas variantes pueden destacarse las aportaciones de Brown (1956, 1959, 1963), Holt (1957), Brown y Meyer (1961), Muth (1960), Winters (1960), Trigg (1964) y Trigg y Leach (1967). Una exposición de estos métodos puede verse en Makridakis y Wheelwright (1978), o en la obra más reciente de Abraham y Ledolter (1983).

Tanto en los métodos de descomposición como de alisado exponencial, el analista establece un esquema «a priori» y después procede a los cálculos estadísticos correspondientes. En los modelos ARIMA univariantes se hace un planteamiento inicial de carácter general. Se considera que la serie temporal objeto de estudio ha sido generada por un proceso estocástico. Las técnicas de elaboración de los modelos ARIMA van dirigidas precisamente a identificar el modelo generador de las observaciones, para después, en un proceso iterativo, estimar y verificar el modelo, que una vez aceptado se utiliza para predecir valores futuros de la serie temporal. Los modelos ARIMA se popularizaron a partir de 1970 merced a la obra de los estadísticos Box y Jenkins (1976). Conviene señalar que los métodos de alisado exponencial son casos particulares de los modelos ARIMA. Para referirse a la elaboración de los modelos ARIMA se utiliza también la expresión de análisis de series temporales, aunque lógicamente esta última expresión tiene una mayor amplitud.

Antes de concluir esta breve panorámica del análisis univariante conviene señalar que en ciertos desarrollos realizados recientemente se han integrado en un solo enfoque los métodos de descomposición y los modelos ARIMA. En este sentido Harvey (1984) ha formulado los modelos estructurales que pueden considerarse como la suma de tres modelos ARIMA —para la tendencia, estacionalidad y movimiento irregular— ninguno de los cuales se observa directamente. Estos modelos pertenecen a la clase de los modelos UCARIMA (Unobserved Components ARIMA), y han sido estudiados por Engle (1978) y Nerlove y otros (1978).

Dentro de los métodos cuantitativos el otro gran bloque está integrado por el análisis causal, denominado así porque en la explicación de la variable o variables objeto de estudio intervienen factores externos. Quizás el término de causalidad, por las connotaciones de todo tipo que puede tener, sea demasiado fuerte para describir la forma en que inciden los factores externos en la explicación de una variable. Sin embargo, este término ha adquirido su carta de naturaleza especialmente en el análisis estadístico de variables relacionadas temporalmente.

Hasta hace poco tiempo los manuales de econometría han estado dedicados casi exclusivamente al estudio de los modelos causales estocásticos, bien uniecuacionales o bien multiecuacionales. En los últimos años existe la tendencia de incorporar en estos manuales el estudio de series temporales siguiendo básicamente el enfoque de Box y Jenkins. Véase, por ejemplo, Pulido (1983). No tiene sentido sin embargo decir que los métodos económétricos convencionales —refiriéndonos con ello a los modelos causales— y los métodos de series temporales son contrapuestos, ya que son de carácter complementario.

Por otra parte, se han propuesto métodos para elaborar modelos causales desde la perspectiva de las series temporales. En esta línea pueden citarse los trabajos de Box y Jenkins (1976) sobre la función de transferencia, de Granger y Newbold (1977) sobre modelos bivariantes, de Box y Tiao (1975) sobre análisis de intervención, y de Jenkins (1979), Jenkins y Alavi (1981) y Tiao y Box (1981), sobre modelos ARIMA

multivariantes. Una visión panorámica, anterior a estas últimas aportaciones, de modelización mediante el enfoque Box-Jenkins puede verse en Treadway y otros (1978).

Simplificando un poco, la diferencia entre el enfoque económico tradicional y el enfoque de series temporales, acerca de la elaboración de modelos causales, sería la siguiente: en el primero, el énfasis se sitúa en el modelo de partida, efectuándose contrastes sobre la adecuación entre modelo y datos; por el contrario, en el enfoque de series temporales el punto de partida son los datos y quizás alguna idea general sobre el fenómeno que se trata de modelizar, siendo el modelo el resultado final de la investigación.

Volviendo a la distinción entre métodos univariantes y métodos causales puede surgir el siguiente interrogante: en una investigación concreta, ¿qué línea es conveniente seguir? Cuando se dispone de un modelo teórico adecuado, de información estadística suficiente y de una previsión precisa de los factores externos o variables explicativas es preferible utilizar un modelo causal como instrumento para la previsión. No obstante en muchas ocasiones no se dan simultáneamente esas circunstancias, siendo necesario recurrir a análisis de carácter univariante. En cualquier caso, siempre es útil, como una primera aproximación a un fenómeno, efectuar un análisis univariante de las variables involucradas en el modelo econométrico. Para la utilización de la metodología ARIMA univariante, como en los demás métodos de análisis univariante, solamente se requiere información, en forma de serie temporal, de la variable objeto de estudio. Ahora bien, cuando el número de observaciones es pequeño⁴ no sería recomendable la utilización de la metodología ARIMA, debido a que algunos de los instrumentos que la configuraron no funcionan adecuadamente con muestras pequeñas.

⁴ Para tener una referencia, aunque un tanto arbitraria, puede considerarse que es pequeña si el número de observaciones es inferior a 60.

CAPÍTULO 2

Procesos estocásticos

2.1. Introducción

En el análisis económico se pueden utilizar datos de corte transversal y datos de serie temporal.

Los datos de corte transversal se caracterizan porque van todos referidos a un mismo momento del tiempo. Su denominación procede precisamente de que se obtienen mediante un *corte* en el eje del tiempo. Estos datos suelen consistir en información acerca de familias, empresas, etc. En general es un conjunto de datos de carácter microeconómico. A este conjunto se le denomina información censal cuando contiene información exhaustiva de todos los elementos del colectivo que se está investigando; por ejemplo, el censo de población de España de 1981. Si solamente se investiga una parte del colectivo, la información obtenida es de carácter muestral. Con objeto de que se puedan hacer inferencias acerca de todo el colectivo es necesario que la muestra investigada se haya obtenido mediante un procedimiento aleatorio. Así, entre julio de 1973 y julio de 1974 el Instituto Nacional de Estadística realizó una encuesta de presupuestos familiares en la que se investigó la estructura del gasto de cerca de 25.000 familias, a partir de las cuales se pudo hacer una inferencia acerca del gasto de *todas* las familias españolas, puesto que se trataba de una muestra obtenida por un procedimiento aleatorio. En la figura 2.1. se ha representado el consumo total de cuatro hipotéticas familias, pertenecientes a esta muestra, a las que se han designado por las letras *i*, *j*, *k* y *l*. Se ha considerado, para simplificar, que el período al que va referido el consumo de esas familias es 1973. Al realizar el proceso de inferencia se obtiene el consumo per cápita español para 1973.

Con los datos de corte transversal, no se plantean en principio problemas de interrelación entre los diferentes elementos que integran la muestra o, en su caso, el censo obtenido. Así, es razonable considerar que el consumo de la familia *i*-ésima es independiente del correspondiente a la familia *j*-ésima. Como hipótesis inicial se puede suponer que el consumo de estas cuatro familias y de las restantes, están generados por una distribución normal de media μ , tal como se refleja en la figura 2.2.a). En abscisas, se representa el consumo, mientras que la ordenada representa la densidad probabilística de cada punto.

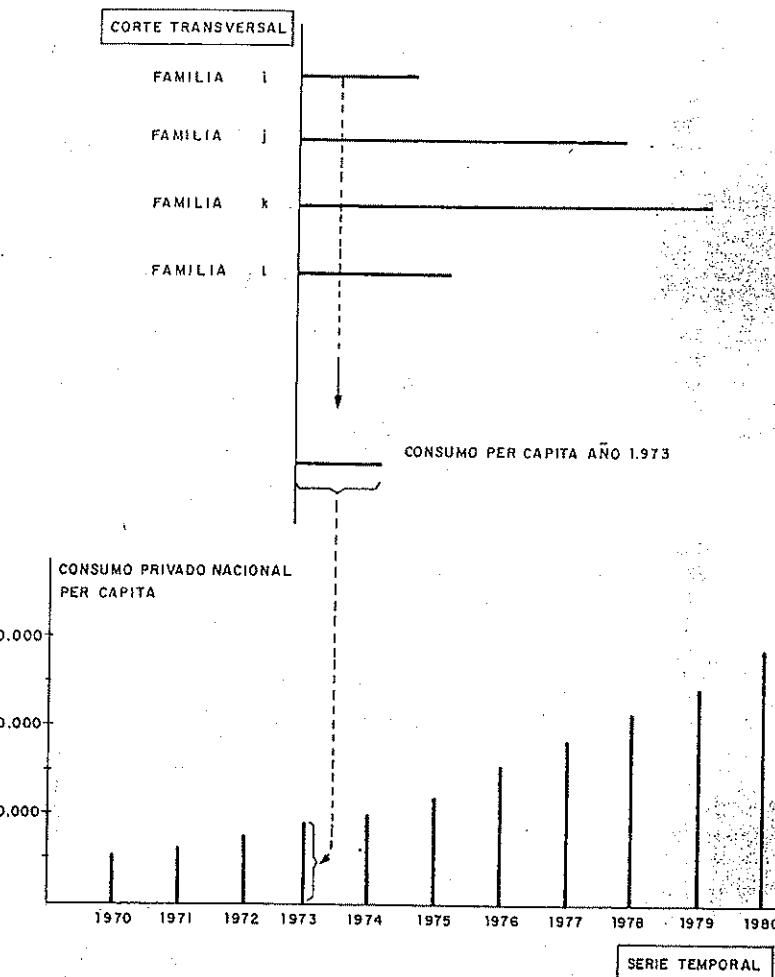
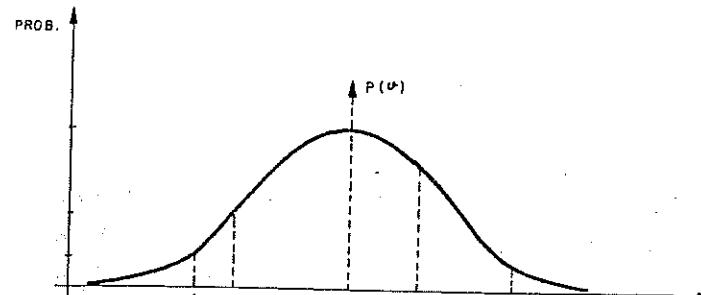


Fig. 2.1.

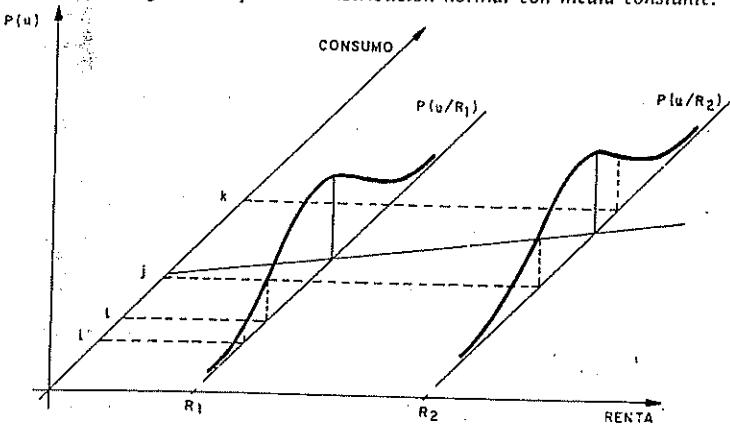
El modelo generador puede expresarse así:

$$C_i = \mu + v_i \quad (2-1)$$

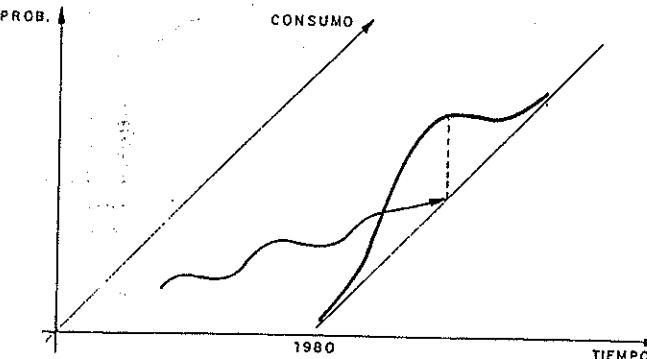
donde v_i es una variable aleatoria en la que $E(v_i) = 0$.



a) El consumo generado por una distribución normal con media constante.



b) El consumo generado por distribuciones normales con medias en función de la renta.



c) El consumo generado por un proceso estocástico.

Puede objetarse que el supuesto adoptado es excesivamente simplificador, ya que el consumo viene determinado en buena medida por la renta. Para recoger esta influencia en lugar de una media fija, se considerará que ésta es variable y dependiente de la renta:

$$\mu_i = \alpha + \beta R_i \quad (2-2)$$

En cualquier caso la renta no determina completamente el consumo, ya que seguirá existiendo un componente aleatorio, al que se designará por u_i . Entonces el consumo vendrá dado por:

$$C_i = \alpha + \beta R_i + u_i \quad (2-3)$$

donde u_i es también una variable aleatoria en que $E(u_i) = 0, \forall i$.

La ecuación (2-3) es justamente un modelo de regresión lineal simple. Como una parte del consumo se explica mediante la renta, la dispersión de la variable u_i será menor que la correspondiente a la variable v_i .

Aunque será muy difícil que coincidan exactamente las rentas de dos familias, a efecto ilustrativo, se ha considerado el supuesto de que las familias i y j disponen de la misma renta R_i , siendo R_j la renta de las familias j y k . En la parte b) de la figura 2.2 se ha representado la distribución probabilística del consumo para las rentas R_i y R_j . Dada la forma en que se han generado estos datos de corte transversal, parece razonable adoptar el supuesto de que las distribuciones probabilísticas de cada nivel de renta son independientes entre sí.

Con los datos procedentes de serie temporal el planteamiento es distinto. Una serie temporal está constituida por unas observaciones de una variable a intervalos regulares de tiempo. La información que se maneja en una serie temporal es generalmente de carácter agregado, aunque el grado de agregación dependerá del sistema que se esté tratando. Así, en una empresa se obtiene la serie temporal de ventas mediante agregación de las ventas efectuadas a cada uno de los clientes. A nivel macroeconómico la agregación suele hacerse por métodos estadísticos, es decir, mediante inferencia al conjunto de la población de lo observado en un subconjunto de la misma. Por ejemplo, el consumo privado nacional per cápita que publica la Contabilidad Nacional de España. En la figura 2.1 se ha reflejado esta serie para el período 1970-1981 (los datos están expresados en pesetas corrientes).

En cualquier caso, cuando se utiliza una serie temporal no parece razonable adoptar el supuesto de que el valor que toma la variable estudiada en un período de tiempo es independiente de los valores que ha tomado dicha variable en los períodos que la han precedido. Así, en la parte c) de la figura 2.2, se ha representado —prescindiendo en este caso de la influencia de la renta— la distribución probabilística del consumo en el año 1980. Se han diferenciado dos componentes. Por una parte, la evolución pasada del consumo, que para 1980 sitúa a éste en el punto marcado por la flecha. El segundo componente es una variable aleatoria; la suma de los dos componentes dará lugar al consumo observado en dicho año. Para analizar la problemática de fenómenos de este tipo se utiliza la teoría de los procesos estocásticos. Las nociones básicas de los procesos estocásticos serán examinadas en los epígrafes siguientes.

2.2. Concepto de proceso estocástico

Un proceso estocástico se define como una familia de variables aleatorias que corresponden a momentos sucesivos del tiempo. Será designado por $Y(t, u)$, donde t es el tiempo y u es la variable aleatoria.

Si se fija t en un valor t_0 , $Y(t_0, u)$ será una variable aleatoria similar a las que se han utilizado en el epígrafe anterior.

Por el contrario, si u se fija en u_0 , entonces para cada momento de tiempo, el proceso tomaría un solo valor $Y(t, u_0)$. Si se consideran los valores que ha tomado en los diversos momentos del tiempo, se tendría una función del tiempo como la que aparece en la parte c) de la figura 2.2.

La determinación de las características de un proceso estocástico puede hacerse mediante dos formas alternativas, bien a partir de funciones de distribución conjunta o bien a partir de los momentos.

Antes de exponer el primer procedimiento se va a examinar brevemente el significado de las funciones de distribución probabilísticas en un proceso estocástico. Cuando se fija un valor en el tiempo, el proceso estocástico pasa a ser una variable aleatoria que tendrá su correspondiente distribución de probabilidad. Así, para $t = t_i$, la distribución de probabilidad será expresada por $F[Y(t_i)]$. Si en lugar de un valor fijamos dos valores del tiempo, se obtendría una variable bidimensional con una función de distribución bivariante. Así para $t = t_i$ y $t = t_j$, la distribución de probabilidad sería $F[Y(t_i), Y(t_j)]$.

En general, para un conjunto finito de valores del tiempo, se obtiene una función de distribución conjunta. Así para t_1, t_2, \dots, t_n , la función de distribución conjunta sería:

$$F[Y(t_1), Y(t_2), \dots, Y(t_n)] \quad (2-4)$$

Teniendo presente lo que antecede, se dice que un proceso estocástico está perfectamente caracterizado cuando se pueden determinar las funciones de distribución conjunta para cada conjunto finito de variables del proceso, es decir, para cada valor finito de n en (2-4).

La determinación de las características del proceso a partir de las funciones de distribución es en general un procedimiento complicado, por lo que se acostumbra a utilizar preferentemente el método de los momentos. En una distribución de probabilidad se pueden calcular momentos de diverso orden, aunque son los momentos de primero y segundo orden los más utilizados.

En un proceso estocástico al que se designará, para simplificar la notación, por Y_t , la media o momento de primer orden se define de la siguiente forma:

$$\mu_t = E(Y_t) \quad (2-5)$$

El subíndice t señala que la media será en general distinta para cada período de tiempo.

Como momentos de segundo orden respecto a la media es preciso considerar,

además de la varianza, las covarianzas entre variables referidas a distintos momentos de tiempo o autocovarianzas a las que se designarán por

$$\gamma_{t,s} = \text{cov}(Y_t, Y_s) = E(Y_t - \mu_t)(Y_s - \mu_s) \quad (2-6)$$

Cuando $s = t$ se tiene definida la varianza:

$$\gamma_{t,t} = \text{var}(Y_t) = E(Y_t - \mu_t)^2 \quad (2-7)$$

Como forma alternativa de caracterización de un proceso estocástico se utilizan los coeficientes de autocorrelación.

$$\rho_{t,s} = \frac{\text{cov}(Y_t, Y_s)}{\sqrt{\text{var}(Y_t) \text{var}(Y_s)}} \quad (2-8)$$

Las autocorrelaciones conjuntamente con las varianzas proporcionan idéntica información que las autocovarianzas. Sin embargo, es preferible utilizar las autocorrelaciones ya que éstas proporcionan unas medidas relativas a diferencia de lo que ocurre con las autocovarianzas que vienen afectadas por la escala que se utiliza.

La caracterización de un proceso estocástico mediante los momentos de primero y segundo orden es en principio más incompleta que cuando se hace mediante funciones de distribución. Ahora bien, si el proceso es *normal* éste queda perfectamente caracterizado, a través de los dos primeros momentos.

En el contexto de los procesos estocásticos, ¿cómo se conceptualizaría a una serie temporal? Aunque en una serie temporal se dispone de una observación para cada período de tiempo, no se obtiene en general de forma determinista como sería el caso en una función exacta del tiempo. Una serie temporal en general tendrá un carácter aleatorio¹, y se puede interpretar como una muestra de tamaño 1 tomada en períodos sucesivos de tiempo en un proceso estocástico. En este sentido, a una serie temporal se le considera como una realización de un proceso estocástico. A diferencia del muestreo aleatorio simple donde cada extracción es independiente de las demás (por ejemplo, el comportamiento de una familia es independiente del correspondiente a la familia extraída en el lugar anterior), en una serie temporal el dato extraído para un período concreto no será, en general, independiente de los datos extraídos para períodos anteriores (por ejemplo, como ya se ha indicado, el consumo privado nacional del año 1980 vendrá condicionado por los valores que haya tomado esta variable en años anteriores).

La información que se maneja en el área económica adopta en muchos casos la forma de serie temporal. Así, las magnitudes de la Contabilidad Nacional, las series monetarias del Banco de España, la serie de ventas de una empresa, etc. En economía, y en general en las ciencias sociales, la información de una serie se obtiene mediante *observación pasiva*, es decir, sin control sobre los factores que influyen sobre la

¹ Si una serie no tuviera un carácter aleatorio, no tendría sentido el registrar los valores correspondientes a distintos períodos, ya que se podrían obtener simplemente por aplicación de una fórmula. Así, si $Y_t = 10 + 4t$, el valor de Y_t para $t = 10$ se obtiene de forma inmediata.

variable objeto de investigación. Por ello aunque se disponga de una serie muy larga, debe considerarse toda ella como una sola realización de un proceso estocástico. Para ilustrar esto, pensemos imaginariamente que disponemos de un índice mensual de precios de consumo desde 1800 a la actualidad; no sería razonable suponer que a partir de una fecha determinada, por ejemplo, a partir de 1900, comienza una segunda realización del proceso estocástico, ya que el contexto en que se produce la observación de los precios a principio de siglo XIX difiere del correspondiente a principios del siglo XX. Debido al carácter pasivo de la toma de información, en economía, en general, sólo se dispone de una sola realización para cada proceso estocástico. En las ciencias experimentales el panorama es muy diferente. Aquí el investigador juega el papel activo, ya que controla los factores que influyen sobre un fenómeno, pudiendo repetir un experimento en condiciones prácticamente idénticas a las que se ha realizado el anterior. Por ejemplo, si se considera la emisión de una determinada partícula en un proceso productivo a lo largo de una jornada de 8 horas, el experimentador puede controlar el proceso de forma que cada día dicho proceso se inicie en unas condiciones idénticas. Si se midiera diariamente el peso de las partículas emitidas cada hora, se tendría un cuadro como el adjunto, donde Y_{it} designa el peso de las partículas en el día i y durante la hora t :

hora \ dia	1	2	3	4	5	6	7	8
1	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}	Y_{16}	Y_{17}	Y_{18}
2	Y_{21}	Y_{22}	Y_{23}	Y_{24}	Y_{25}	Y_{26}	Y_{27}	Y_{28}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
N	Y_{N1}	Y_{N2}	Y_{N3}	Y_{N4}	Y_{N5}	Y_{N6}	Y_{N7}	Y_{N8}

Cada línea del cuadro anterior se puede contemplar como una realización de un mismo proceso estocástico. Si se dispone de un número amplio de días en que se han efectuado las mediciones del peso de las partículas emitidas, se estaría en condiciones de poder realizar inferencias acerca del proceso.

Así, una estimación de la emisión durante la primera hora se realizaría mediante el siguiente estadístico:

$$\hat{\mu}_1 = \frac{1}{N} \sum_{i=1}^N Y_{i1} \quad (2-9)$$

Lo mismo se haría con las demás medidas: varianza y autocovarianzas, utilizadas en la caracterización del proceso.

En las ciencias sociales, la problemática que se plantea es mucho más complicada, puesto que con el número de datos obtenidos para una sola realización se debe estimar un número superior de parámetros. Así, si se dispone de n datos, con ellos se deberían estimar n medias y n varianzas, sin contar con las autocovarianzas que también se necesitan para poder caracterizar el proceso. En otras palabras, tal como

se ha planteado, el problema no tiene ninguna solución. Para poder, a partir de una sola realización, efectuar inferencias sobre un proceso estocástico es preciso imponer restricciones a este último. Las restricciones impuestas habitualmente son que sea estacionario y ergódico.

2.3. Procesos estacionarios y ergódicos

Para definir la estacionariedad de un proceso se puede utilizar, como se hizo anteriormente para su caracterización, bien las funciones distribución o alternativamente los momentos.

Se dice que un proceso estocástico es estacionario en sentido estricto cuando al realizar un mismo desplazamiento en el tiempo de todas las variables de cualquier distribución conjunta finita, resulta que esta distribución no varía.

Considerando la función de distribución conjunta

$$F(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}) \quad (2-10)$$

si se adopta el supuesto de que a todos los elementos de la anterior distribución se desplazan m períodos, la nueva función de distribución conjunta será

$$F(Y_{t_1+m}, Y_{t_2+m}, \dots, Y_{t_k+m}) \quad (2-11)$$

Si el proceso es estacionario en sentido estricto se deberá verificar que

$$F(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}) = F(Y_{t_1+m}, Y_{t_2+m}, \dots, Y_{t_k+m}) \quad (2-12)$$

e igualmente, se deberá obtener un resultado análogo para cualquier otra distribución conjunta que tenga carácter finito.

También en este caso es más complejo el análisis de la estacionariedad a partir de las funciones de distribución que si se efectúa a partir de los momentos. En contrapartida, el concepto de estacionariedad será más limitado. Se dice que un proceso es estacionario de primer orden, o en media, si se verifica que:

$$E(Y_t) = \mu \quad \forall t \quad (2-13)$$

Por tanto, en un proceso estacionario en media, la esperanza matemática, o media teórica, permanece constante a lo largo del tiempo.

Se dice que un proceso es estacionario de segundo orden (o en sentido amplio) cuando se verifican las dos condiciones siguientes:

1.) La varianza es finita y permanece constante a lo largo del tiempo, es decir:

$$E(Y_t - \mu)^2 = \sigma^2 < \infty \quad \forall t \quad (2-14)$$

2.") La autocovarianza entre dos períodos distintos de tiempo únicamente viene afectada por el lapso de tiempo transcurrido entre esos dos períodos. Así:

$$E(Y_{t+k} - \mu)(Y_t - \mu) = \gamma_k \quad \forall t \quad (2-15)$$

que sería una autocovarianza de orden k , por ser éste el lapso de tiempo que separa a Y_t de Y_{t+k} . Su valor, γ_k es independiente de cuál sea el período t que se considere. Como puede verse, la varianza del proceso es simplemente la autocovarianza de orden 0.

Al definir un proceso estacionario en sentido amplio, se tiene en cuenta implícitamente que el proceso es también estacionario en media, ya que tanto en la varianza como en las autocovarianzas, el símbolo μ no viene afectado por ningún subíndice.

Si un proceso es estacionario en sentido estricto también será estacionario en sentido amplio, pero no ocurrirá necesariamente el caso recíproco, pues el proceso puede ser no estacionario para momentos superiores al segundo. También aquí, si el proceso es estacionario en sentido amplio, y además es normal, se verifica que el proceso es estacionario en sentido estricto.

En un proceso estacionario las autocorrelaciones quedan definidas de la siguiente forma:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad k \geq 0 \quad (2-16)$$

De acuerdo con (2-15), se verificará que $\gamma_k = \gamma_{-k}$, y por consiguiente $\rho_k = \rho_{-k}$. La representación gráfica de ρ_k para $k = 0, 1, 2, 3, \dots$, recibe el nombre de correlograma.

Cuando el proceso es estacionario, en principio pueden estimarse los parámetros $\mu, \gamma_0, \gamma_1, \gamma_2, \dots$ a partir de una sola realización. Disponiendo de la muestra Y_1, Y_2, \dots, Y_T se pueden utilizar los siguientes estimadores:

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T Y_t$$

$$\hat{\gamma}_0 = \frac{1}{T} \sum_{t=1}^T (Y_t - \hat{\mu})^2 \quad (2-17)$$

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (Y_{t+k} - \hat{\mu})(Y_t - \hat{\mu})$$

Naturalmente, al crecer k se dispone de menos observaciones en el cálculo de $\hat{\gamma}_k$. Así, para $\hat{\gamma}_{T-1}$ solamente se dispondría de una sola observación.

Además de la estacionariedad, es necesario que el proceso estocástico goce de la propiedad de ergodicidad, con objeto de que el proceso de inferencia pueda realizarse de una forma adecuada. El concepto de ergodicidad se examinará de forma intuitiva.

Cuando valores de la serie temporal alejados en el tiempo están muy correlacionados, es decir, cuando ρ_k se mantiene en unas cotas elevadas para un k grande, suce-

derá que al aumentar el tamaño de la muestra se añade poca información nueva. La consecuencia de este hecho en el plano estadístico será que los estimadores obtenidos no son consistentes, ya que el aumento del tamaño de la muestra no tendrá una especial utilidad, puesto que se tendrá que calcular un mayor número de autocovarianzas para caracterizar adecuadamente el proceso.

Cuando se verifica la propiedad de ergodicidad las fórmulas (2-17) permiten obtener estimadores consistentes. Una condición necesaria, aunque no suficiente, de la ergodicidad es que

$$\lim_{T \rightarrow \infty} \rho_k = 0 \quad (2-18)$$

En lo sucesivo, se supondrá siempre que el proceso es ergódico.

Cuando un proceso es estacionario —y también ergódico— todo el problema de inferencia se simplifica de forma considerable. Ahora bien, ¿es razonable suponer que las series económicas que usualmente manejamos los economistas están generadas por procesos estocásticos estacionarios? En general, puede afirmarse que las series económicas no son estacionarias. Puede pensarse, por ejemplo, en el PIB, el índice de precios de consumo, la oferta monetaria, etc. A la vista de esto, una conclusión inmediata podría ser que para un economista tiene poco interés la utilización de procesos estocásticos estacionarios. Pero en realidad no es así, pues mediante sencillas transformaciones en la mayor parte de los casos las series económicas se pueden convertir en series aproximadamente estacionarias, siendo entonces aplicable el proceso de inferencia correspondiente a procesos de este tipo.

2.4. Procesos lineales

El foco de atención lo vamos a centrar en una clase especial de los procesos estacionarios y ergódicos: los procesos lineales. Estos procesos se caracterizan porque se pueden representar como una combinación lineal de variables aleatorias. Se van a examinar los siguientes tipos de procesos lineales: procesos puramente aleatorios, procesos autorregresivos, procesos de medias móviles y los obtenidos como combinación de estos dos últimos. En primer lugar se verá la definición de estos procesos, y después se justificará el por qué se ha restringido el campo de estudio a estos tipos exclusivamente.

El proceso puramente aleatorio es el más simple de todos. Podría expresarse de la siguiente forma:

$$Y_t = \epsilon_t \quad (2-19)$$

donde ϵ_t satisface las siguientes propiedades:

$$E[\epsilon_t] = 0 \quad \forall t$$

$$E[\epsilon_t]^2 = \sigma^2 \quad \forall t$$

$$E[\epsilon_t \epsilon_{t'}] = 0 \quad t \neq t'$$
(2-20)

Así pues, ϵ_t se caracteriza porque su media —igual a 0— y su varianza son constantes a lo largo del tiempo y porque no existe relación entre valores referidos a momentos distintos del tiempo. Las características de ϵ_t son idénticas a las que tiene la perturbación aleatoria en el modelo de regresión lineal múltiple bajo las hipótesis básicas.

En el tratamiento de series temporales, se suele designar a un proceso puramente aleatorio con la denominación de «ruido blanco» (o «white noise» en terminología anglosajona).

De ahora en adelante, por ϵ_t designaremos *únicamente* a una variable aleatoria que goza de las propiedades enunciadas en (2-20).

Naturalmente, al no existir relación entre valores referidos a momentos distintos del tiempo, el «ruido blanco» está alejado de la concepción intuitiva que hemos dado de un proceso estocástico. O, invirtiendo los términos, para manejar variables aleatorias del tipo ϵ_t no haría falta la teoría de los procesos estocásticos.

Esta última afirmación es cierta. Sin embargo el «ruido blanco» es una pieza clave en la construcción de procesos estocásticos más complicados como pueden ser los procesos autorregresivos y los procesos de medias móviles, que se examinarán a continuación.

El proceso autorregresivo de orden p , o utilizando la notación usual, un proceso $AR(p)$ se expresa de la siguiente forma:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad (2-21)$$

Como puede verse en un proceso $AR(p)$, aparece un «ruido blanco» referido al momento actual y la variable desfasada para distintos períodos, siendo p el retardo máximo que aparece en el proceso. La denominación de autorregresivo procede de que Y_t se obtiene mediante regresión sobre valores desfasados de la propia variable. Los procesos autorregresivos fueron introducidos por primera vez por Yule (1927).

Un proceso de medias móviles de orden q , o un proceso $MA(q)$ ² viene dado por:

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (2-22)$$

La expresión de medias móviles hace referencia a que la variable Y_t se obtiene como un promedio de variables de «ruido blanco» (en este caso $q+1$), siendo las θ_i los coeficientes de ponderación; como las variables que forman parte de este promedio varían a lo largo del tiempo, reciben el apelativo de móviles. La introducción de estos procesos también se debe a Yule (1921, 1926).

Finalmente, mediante una combinación de un proceso autorregresivo y un proceso de medias móviles se obtiene un proceso $ARMA(p, q)$, donde p indica el retardo máximo de la parte autorregresiva y q señala el correspondiente a la parte de las medias móviles. La divulgación y popularización de este tipo de procesos se debe funda-

²Las letras MA son iniciales de «moving average», expresión inglesa que se corresponde a la expresión castellana de medias móviles.

mentalmente a la obra de Box y Jenkins (1976), aunque fueron estudiados anteriormente por Wold (1938) y Barlett (1946). La expresión de un proceso $ARMA(p, q)$ es la siguiente:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (2-23)$$

Procede ahora justificar el por qué se ha centrado el estudio en ciertos tipos de procesos.

En el año 1938 el profesor Wold enunció y demostró el siguiente teorema:

Cualquier proceso estacionario Y_t puede representarse únicamente como la suma de dos procesos mutuamente incorrelacionados.

$$Y_t = D_t + X_t \quad (2-24)$$

donde D_t es linealmente determinista y X_t es un proceso $MA(\infty)$. En contraposición al proceso determinista, a la parte de medias móviles se le denomina puramente no determinista.

La parte determinista D_t puede ser una función exacta de tiempo. Por ejemplo,

$$D_t = A \cos(\omega t) \quad (2-25)$$

describe una oscilación sinusoidal a lo largo del tiempo.

El caso más simple de D_t consistiría en hacerlo igual a una constante:

$$D_t = \mu$$

Volviendo al Teorema de Wold, conviene recalcar que la representación (2-24) se aplica a *cualquier* proceso estacionario, sea o no lineal. Si dejamos a un lado la parte determinista, la parte no determinista se representa mediante un $MA(\infty)$. A efectos prácticos, no se puede manipular un MA con infinitos retardos. Sin embargo es razonable esperar que muchos procesos estacionarios se puedan aproximar mediante un MA con un número no muy elevado de retardos. En la aproximación de un $MA(\infty)$ también se puede utilizar un AR o un ARMA finito, según se verá en el capítulo siguiente dedicado a los modelos lineales, donde se utiliza el término modelo como equivalente al de proceso. El significado dado a ambos términos es el mismo. Quizás el término de modelo tiene la connotación de mayor formalización, por lo que se utiliza con mayor frecuencia en el estudio de procesos concretos.

CAPÍTULO 3

Modelos lineales

En este capítulo se examinarán sucesivamente los siguientes modelos lineales estacionarios: autorregresivos (AR), de medias móviles (MA) y mixtos (ARMA) en particular. Finalmente, dentro de los modelos no estacionarios se estudiarán los modelos ARIMA.

En este capítulo el método utilizado para exponer los modelos lineales es el método deductivo. Así, a partir de unas hipótesis iniciales se obtendrán las propiedades de los diferentes tipos de modelos. El conocimiento de estas propiedades facilitará la realización del proceso de inferencia cuyo tratamiento se verá en los capítulos siguientes.

3.1. Modelos autorregresivos (AR)

En el capítulo anterior, un modelo autorregresivo de orden p , o abreviadamente un modelo AR(p), se definió de la siguiente forma:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (3-1)$$

donde ε_t es una variable «ruido blanco».

Utilizando el operador polinomial de retardos¹:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

el modelo puede expresarse de forma compacta:

$$\phi(L) Y_t = \varepsilon_t \quad (3-2)$$

A continuación se van a analizar las características de los modelos AR(1) y AR(2). Los resultados obtenidos se generalizarán a un modelo AR(p).

Modelo AR(1).

Un modelo AR(1) viene definido por

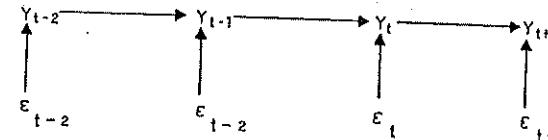
$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t \quad (3-3)$$

¹ Véase en el apéndice A.1. el significado de operador polinomial de retardos.

o, utilizando el operador de retardos, por

$$(1 - \phi_1 L) Y_t = \varepsilon_t$$

Mediante un diagrama de flechas se puede ver claramente cómo se relacionan la variable del proceso Y_t y el ruido blanco ε_t .



Cada variable ruido blanco influye, pues, sobre los valores de Y correspondientes al mismo período; o a períodos posteriores; pero nunca ejerce influencia sobre los valores de Y correspondientes a períodos anteriores. Una consecuencia importante es que

$$E[\varepsilon_t Y_{t-\tau}] = 0 \quad \forall \tau > 0 \quad (3-4)$$

Para que el proceso definido en (3-3) sea estacionario, la raíz del polinomio característico

$$1 - \phi_1 L = 0$$

debe caer fuera del círculo unidad². Es decir

$$|L| = ||/|\phi_1| > 1$$

lo que equivale a que

$$|\phi_1| < 1$$

Para analizar el comportamiento del proceso (3-3) se va a suponer en principio que el proceso se ha iniciado en el período $-N$. Por sustituciones sucesivas se tendrá que

$$\begin{aligned} Y_t &= \phi_1 [\phi_1 Y_{t-2} + \varepsilon_{t-1}] + \varepsilon_t = \\ &= \dots = \\ &= \sum_{j=0}^{t-N-1} \phi_1^j \varepsilon_{t-j} + \phi_1^{t+N} Y_{-N} \end{aligned} \quad (3-5)$$

Si se toman esperanzas en (3-5), se deduce que:

$$E[Y_t] = \phi_1^{t+N} Y_{-N} \quad (3-6)$$

donde Y_{-N} , que representa las condiciones iniciales, se ha tratado como variable no estocástica. En general, se considera que un proceso estocástico se inicia en un período infinitamente lejano.

² Véase en el apéndice A.2. una exposición de las ecuaciones en diferencias finitas.

Si $|\phi_1| < 1$, entonces la esperanza (3-6) decrecerá en valor absoluto a medida que tome valores más elevados. Todo lo contrario ocurrirá cuando $|\phi_1| > 1$, siempre que Y_{-N} no sea exactamente igual a 0. Cuando $\phi_1 = 1$ entonces se verifica que $E(Y_t) = Y_{-N}$; si fuera $\phi_1 = -1$ se daría una alternancia de signos en el valor de la esperanza. En consecuencia, para un proceso que se inicia en un punto finito del tiempo, si el valor inicial no es nulo, la media de Y_t únicamente se mantiene constante cuando $\phi_1 = 1$.

Ahora bien, si el proceso se considera iniciado en $-\infty$, entonces para $|\phi_1| < 1$, y cualquiera que sea el valor inicial, se verificará que

$$E[Y_t] = 0 \quad (3-7)$$

Si en el modelo (3-3) se incluye una constante, se tendrá que

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t \quad (3-8)$$

Entonces, si el proceso se inicia en $-\infty$ y es estacionario, se verificará que la media del proceso será constante para cualquier valor de t .

$$E(Y_t) = E(Y_{t-1}) = \mu \quad (3-9)$$

Tomando esperanzas en (3-8) y teniendo en cuenta lo anterior, se verificará que

$$\mu = \frac{\delta}{1 - \phi_1} \quad (3-10)$$

Sin pérdida de generalidad se supondrá en lo sucesivo que $\delta = 0$.

Un proceso iniciado en $-N$ tendrá la siguiente varianza:

$$\begin{aligned} \gamma_0 &= E[Y_t - E(Y_t)]^2 = E[Y_t - \phi_1^{t-N} Y_{-N}]^2 = E\left[\sum_{j=0}^{t-N-1} \phi_1^j \varepsilon_{t-j}\right]^2 = \sigma_\varepsilon^2 \sum_{j=0}^{N-1} \phi_1^{2j} \\ &\quad [\text{teniendo en cuenta }^3] \\ &= \sigma_\varepsilon^2 \frac{1 - \phi_1^{2(N-1)}}{1 - \phi_1^2} \end{aligned} \quad (3-11)$$

Puede decirse finalmente que tanto para $\phi_1 = 1$ como para $|\phi_1| > 1$, la varianza (3-11) se hace infinita cuanto el proceso arranca desde $-\infty$. Así pues, procesos de estas características tienen una naturaleza explosiva.

³ Si a_1 y a_n son respectivamente el primer y último término de una progresión geométrica y r es la razón, entonces se verificará que la suma de los términos de la progresión será:

$$S = \frac{a_1 - a_n}{1 - r}$$

Si el proceso estacionario, es decir, si $|\phi_1| < 1$ y se ha iniciado $-\infty$, se verifica que:

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2} \quad (3-12)$$

En lo sucesivo, y mientras no se diga lo contrario, se adoptará el supuesto de que el proceso arranca desde $-\infty$, y que cumple la condición de estacionariedad.

De todas formas, a efectos prácticos aunque el proceso se haya iniciado en un período finito, la varianza se estabiliza muy pronto en torno a valores próximos a (3-12).

Si multiplicamos ambos miembros de (3-3) por $Y_{t-\tau}$ se verificará que

$$\gamma_\tau = E[Y_t Y_{t-\tau}] = \phi_1 E[Y_{t-1} Y_{t-\tau}] + E[\varepsilon_t Y_{t-\tau}] \quad (3-13)$$

Para $\tau > 0$, de acuerdo con el diagrama de flechas expuesto anteriormente, se verifica que

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} \quad \tau > 0 \quad (3-14)$$

La ecuación (3-14) es una ecuación en diferencias *homogénea* de primer orden. También es una ecuación en diferencias de primer orden y con el mismo parámetro, el proceso generador (3-3); en cambio esta última no es homogénea debido a la presencia del ruido blanco ε_t .

La solución de la ecuación (3-14) será:

$$\gamma_\tau = A \lambda^\tau = \gamma_0 \phi_1^\tau \quad (3-15)$$

Si en (3-14) se dividen ambos miembros por γ_0 resulta que

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \phi_1 \rho_{\tau-1} \quad \tau > 0 \quad (3-16)$$

La solución de la ecuación en diferencias (3-16) será

$$\rho_\tau = \phi_1^\tau \rho_0 = \phi_1^\tau \quad (3-17)$$

En las autocovarianzas la condición inicial viene determinada por el valor de la varianza; en cambio en las autocorrelaciones como son medidas relativas, la condición inicial es siempre $\rho_0 = 1$. En las figuras 3.1.a) y 3.1.b) se han representado valores sucesivos de ρ_τ , es decir, el correlograma, para $\phi_1 = 0.8$ y $\phi_1 = -0.8$ respectivamente. En el primer caso, el correlograma muestra un decrecimiento exponencial puro. En cambio, cuando $\phi_1 = -0.8$ el correlograma muestra que ρ sigue un decrecimiento exponencial, pero con alternancia de signos. En la figura 3.2. partes a) y b) puede verse la representación de dos realizaciones que se han generado artificialmente

por el método Montecarlo⁴ a partir de los procesos $Y_t = 0.8 Y_{t-1} + \varepsilon_t$ e $Y_t = -0.8 Y_{t-1} + \varepsilon_t$. En ambos casos se ha tomado como valor inicial $Y_0 = 0$.

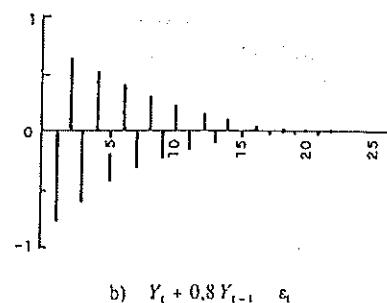
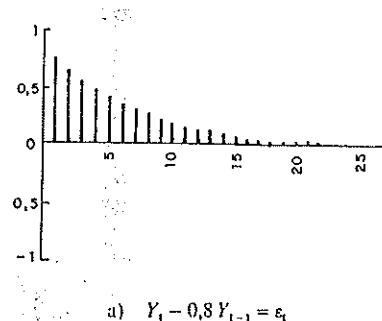


Fig. 3.1. Funciones de autocorrelación de modelos AR(1).

En un proceso AR(1) estacionario se puede despejar Y_t , utilizando el inverso del operador polinomial de retardos. Así,

$$Y_t = \phi^{-1}(L)\varepsilon_t = \frac{1}{1 - \phi_1 L}\varepsilon_t \quad (3-18)$$

Naturalmente, considerando que $|\phi_1 L| < 1$, la fracción del segundo miembro puede contemplarse como la suma de los términos de una progresión geométrica infinita convergente de razón $\phi_1 L$. Es decir:

$$Y_t = \sum_{j=0}^{\infty} (\phi_1 L)^j \varepsilon_t = \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{t-j} \quad (3-19)$$

A idéntica conclusión se hubiera llegado si en (3-5) se hubiera continuado de forma indefinida la sustitución.

De acuerdo con la terminología utilizada en el capítulo anterior, el modelo (3-19) es un modelo de medias móviles con infinitos retardos. En consecuencia, se observa por tanto que partiendo de un modelo AR(1) inicial se ha pasado a un modelo MA(∞).

⁴Véase en el apéndice A.3 una exposición de la generación de modelos mediante el método de Monte-Carlo.

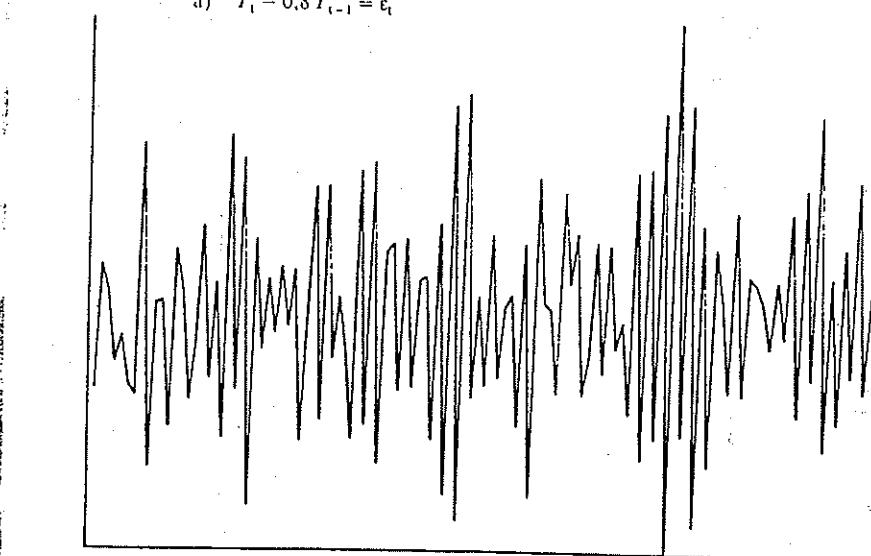
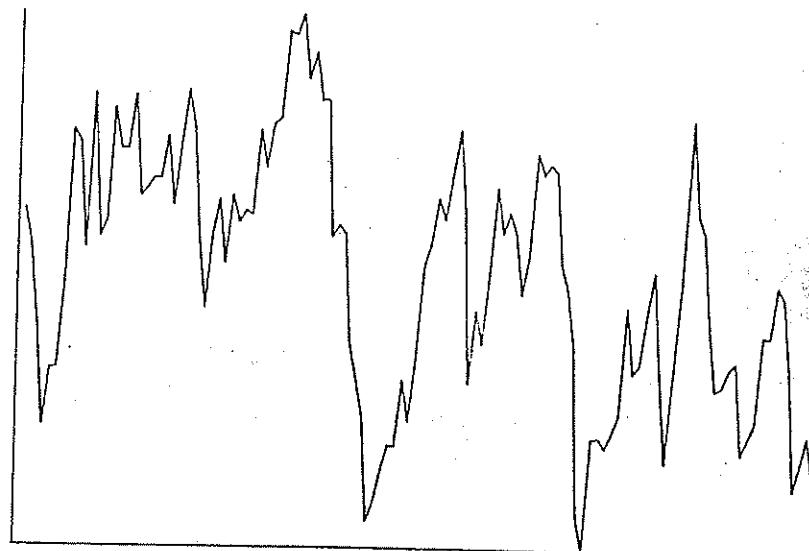


Fig. 3.2. Realizaciones de modelos AR(1).

Modelo AR(2)

Un modelo AR(2) viene definido por

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t \quad (3-20)$$

o, utilizando el operador de retardos,

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = \epsilon_t$$

Para que el proceso anterior sea estacionario se requiere que las raíces de la ecuación

$$1 - \phi_1 L - \phi_2 L^2 = 0 \quad (3-21)$$

estén situadas fuera del círculo unidad.

Si se cumplen las condiciones de estacionariedad —supuesto que se adoptará de aquí en adelante— se verifica que $E(Y_t) = 0$.

Si se multiplican ambos miembros de (3-20) por $Y_{t-\tau}$ y se toman esperanzas, se tiene que

$$E[Y_t Y_{t-\tau}] = \phi_1 E[Y_{t-1} Y_{t-\tau}] + \phi_2 E[Y_{t-2} Y_{t-\tau}] + E[\epsilon_t Y_{t-\tau}] \quad (3-22)$$

Teniendo en cuenta que

$$E[Y_t Y_{t-\tau}] = E[\epsilon_t (\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t)] = \sigma_\epsilon^2 \quad (3-23)$$

la expresión (3-22) para $\tau = 0$ dará el siguiente resultado:

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_\epsilon^2 \quad (3-24)$$

Para valores de $\tau > 0$ se tendrá que

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} \quad \tau > 0 \quad (3-25)$$

La ecuación de autocovarianzas (3-25) es una ecuación en diferencias homogénea de segundo orden. Dividiendo ambos miembros por γ_0 se obtiene la ecuación en diferencias relativa a las autocorrelaciones.

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} \quad \tau > 0 \quad (3-26)$$

La solución de la ecuación en diferencias (3-26) será en general⁵ del tipo:

$$\rho_\tau = A_1 \lambda_1^\tau + A_2 \lambda_2^\tau \quad (3-27)$$

donde $1/\lambda_1$ y $1/\lambda_2$ serán las raíces del polinomio característico (3-21).

⁵ Según puede verse en el apéndice A.2., cuando las raíces λ_1 y λ_2 son iguales, la solución de la ecuación en diferencias presenta otra forma.

Las constantes arbitrarias se determinan a partir de condiciones iniciales: $\rho_0 = 1$ y $\rho_1 = \phi_1/(1 - \phi_2)$. Este último valor se deduce haciendo $\tau = 1$ en (3-26). Por tanto,

$$\begin{aligned} \rho_0 &= A_1 \lambda_1^0 + A_2 \lambda_2^0 = 1 \\ \rho_1 &= A_1 \lambda_1 + A_2 \lambda_2 = \phi_1/(1 - \phi_2) \end{aligned} \quad (3-28)$$

Resolviendo este sistema para A_1 y A_2 , y sustituyendo los valores obtenidos en (3-27), se puede aplicar esta fórmula para determinar el valor de ρ_τ para cualquier valor de $\tau > 0$. Debe tenerse en cuenta que A_1 , A_2 , λ_1 y λ_2 se han calculado a partir de ϕ_1 y ϕ_2 . Ahora bien, puede plantearse el problema inverso, es decir, determinar ϕ_1 y ϕ_2 a partir del correlograma. En este último caso, haciendo $\tau = 1$ y $\tau = 2$ en (3-26) se obtiene el siguiente sistema de ecuaciones:

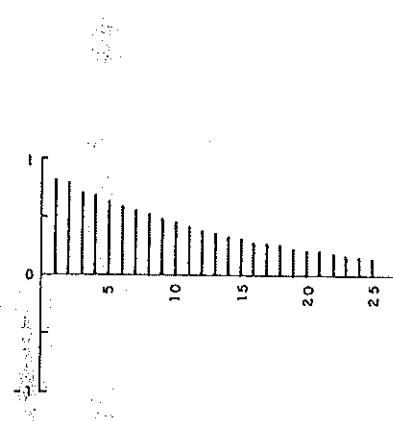
$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \quad (3-29)$$

El sistema anterior recibe la denominación de ecuaciones de Yule-Walker, como tributo a las aportaciones de Yule (1927) y Walker (1931). Resolviéndolo para ϕ_1 y ϕ_2 se tiene que,

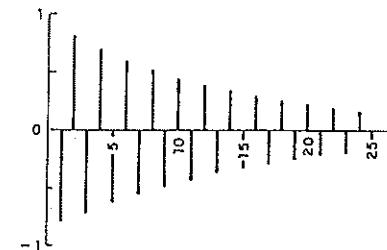
$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \quad (3-30)$$

En la figura 3.3. se han representado los correlogramas correspondientes a cuatro modelos AR(2). Las raíces de los modelos a) y b) son reales; en ambos casos se observa que los coeficientes de autocorrelación decuen exponencialmente, si bien en el modelo b) el decrecimiento exponencial va acompañado de una alternancia de signos. Las raíces de los modelos c) y d) son complejas, dando lugar a oscilaciones sinusoidales que decuen también de forma geométrica; en el modelo d) es más difícil apreciar el movimiento sinusoidal debido a que va acompañado de una alternancia de signos. Para cada uno de los 4 modelos AR(2) se ha simulado una realización de 120 observaciones mediante el método de Montecarlo. Estas realizaciones se han representado en la figura 3.4.

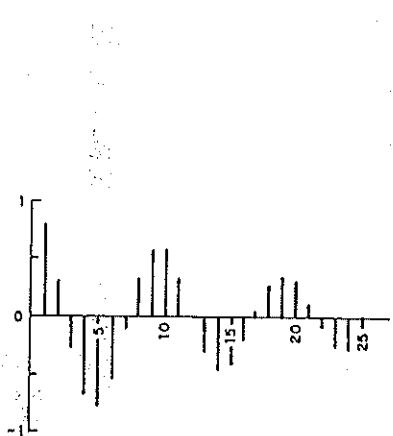
NUMERO DE OBSERVACIONES: 120



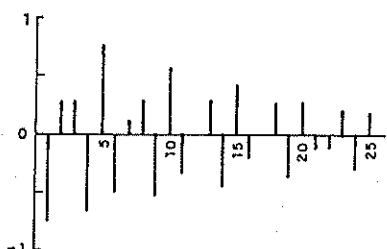
$$a) Y_t = 0,4 Y_{t-1} - 0,5 Y_{t-2} + \epsilon_t$$



$$b) Y_t + 0,4 Y_{t-1} - 0,5 Y_{t-2} = \epsilon_t$$

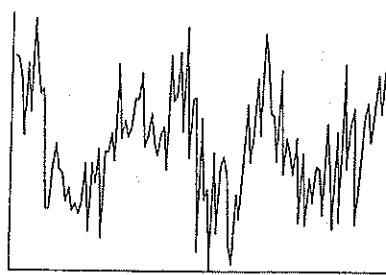


$$c) Y_t - 1,5 Y_{t-1} - 0,9 Y_{t-2} = \epsilon_t$$

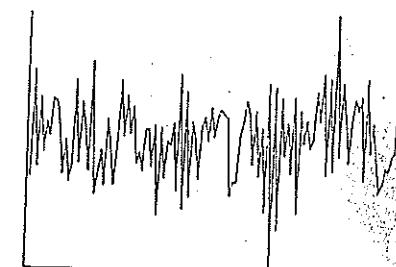


$$d) Y_t + 1,5 Y_{t-1} + 0,9 Y_{t-2} = \epsilon_t$$

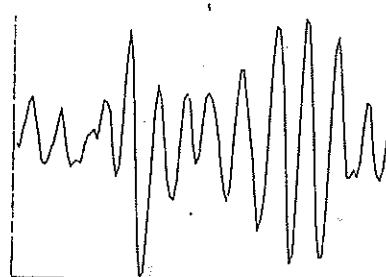
Fig. 3.3. Funciones de autocorrelación de modelos AR(2).



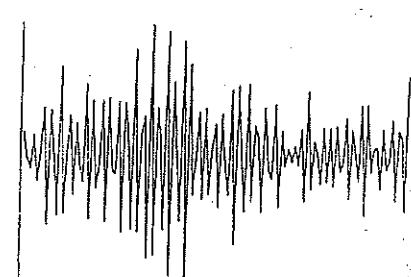
$$a) Y_t = 0,4 Y_{t-1} - 0,5 Y_{t-2} + \epsilon_t$$



$$b) Y_t + 0,4 Y_{t-1} - 0,5 Y_{t-2} = \epsilon_t$$



$$c) Y_t - 1,5 Y_{t-1} - 0,9 Y_{t-2} = \epsilon_t$$



$$d) Y_t + 1,5 Y_{t-1} + 0,9 Y_{t-2} = \epsilon_t$$

Fig. 3.4. Realizaciones de modelos AR(2).

En un proceso AR(2), y mediante utilización del inverso del operador polinomial e retardos, se puede despejar Y_t :

$$Y_t = \phi^{-1}(L)\epsilon_t = \frac{1}{1 - \phi_1 L - \phi_2 L^2} \epsilon_t$$

Si $1/\lambda_1$ y $1/\lambda_2$ son las raíces del polinomio característico (3-21), se puede proce-

der a la factorización del denominador que aparece en la expresión anterior, de la siguiente forma⁶:

$$Y_t = \frac{1}{(1 - \lambda_1 L)(1 - \lambda_2 L)} \varepsilon_t \quad (3-31)$$

Desarrollando la fracción anterior, en forma de fracciones parciales, se tendrá:

$$\frac{1}{(1 - \lambda_1 L)(1 - \lambda_2 L)} = \frac{A}{1 - \lambda_1 L} + \frac{B}{1 - \lambda_2 L} \quad (3-32)$$

A y B son constantes que se determinan de la siguiente forma. En primer lugar se eliminan los denominadores en (3-32):

$$1 = (1 - \lambda_2 L)A + (1 - \lambda_1 L)B = (A + B) + (-\lambda_2 A - \lambda_1 B)L \quad (3-33)$$

Igualando miembro a miembro los coeficientes de las potencias de L de idéntico grado se tiene que

$$\begin{aligned} A + B &= 1 \\ -\lambda_2 A - \lambda_1 B &= 0 \end{aligned}$$

La solución del anterior sistema permite obtener que

$$A = \lambda_1 / (\lambda_1 - \lambda_2) \quad y \quad B = -\lambda_2 / (\lambda_1 - \lambda_2)$$

En consecuencia, (3-31) se puede expresar de la siguiente forma:

$$\begin{aligned} Y_t &= \left[\frac{\lambda_1}{\lambda_1 - \lambda_2} \frac{1}{1 - \lambda_1 L} - \frac{\lambda_2}{\lambda_1 - \lambda_2} \frac{1}{1 - \lambda_2 L} \right] \varepsilon_t = \\ &= \left[\frac{\lambda_1}{\lambda_1 - \lambda_2} \sum_{j=0}^{\infty} (\lambda_1 L)^j - \frac{\lambda_2}{\lambda_1 - \lambda_2} \sum_{j=0}^{\infty} (\lambda_2 L)^j \right] \varepsilon_t \end{aligned} \quad (3-34)$$

Así pues, (3-34) es la suma de dos procesos MA(∞) que también será otro proceso de medias móviles de infinitos retardos. Los coeficientes del proceso resultante se pueden obtener a partir de (3-34) sumando los coeficientes para cada nivel de retardos. En el epígrafe 3.3 se examinará un procedimiento alternativo, que es más sencillo de aplicar, para calcular los coeficientes del proceso MA(∞) a que equivale el modelo AR(2).

⁶ La forma de factorización más usual de (3-21) hubiera sido la siguiente:

$$\left[L - \frac{1}{\lambda_1} \right] \left[L - \frac{1}{\lambda_2} \right] = 0$$

Para pasar a la factorización (3-31) que es la que interesa en nuestro caso se ha multiplicado por $\lambda_1 \lambda_2$.

Modelo AR(p).

Un modelo AR(p) viene definido por:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (3-35)$$

o, alternativamente, por

$$\phi(L) Y_t = \varepsilon_t$$

donde

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

Para que el proceso sea estacionario se requiere que las raíces de la ecuación polinomial

$$1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p = 0$$

estén fuera del círculo unidad.

Si en el modelo (3-35) se incluye un término independiente

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \delta + \varepsilon_t \quad (3-36)$$

entonces bajo el supuesto de estacionariedad, tomando esperanzas en (3-36) y denominado $\mu = E(Y_t) \forall t$, se tiene que

$$\mu = \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu$$

por tanto

$$\mu = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p} \quad (3-37)$$

En lo sucesivo se hará el supuesto, sin pérdida de generalidad, de que $\delta = 0$.

Multiplicando ambos miembros de (3-35) por $Y_{t-\tau}$ y tomando esperanzas se tiene que,

$$\gamma_t = \phi_1 \gamma_{t-1} + \phi_2 \gamma_{t-2} + \dots + \phi_p \gamma_{t-p} + E[\varepsilon_t Y_{t-\tau}] \quad (3-38)$$

Para $\tau = 0$, se obtiene

$$\gamma_0 = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_\varepsilon^2 \quad (3-39)$$

Para valores de $\tau > 0$ el resultado obtenido es el siguiente:

$$\gamma_t = \phi_1 \gamma_{t-1} + \phi_2 \gamma_{t-2} + \dots + \phi_p \gamma_{t-p} \quad \tau > 0 \quad (3-40)$$

Dividiendo ambos miembros de (3-40) por γ_0 se obtiene la ecuación en diferencias de orden p relativa a las autocorrelaciones,

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} + \dots + \phi_p \rho_{\tau-p} \quad (3-41)$$

Tomando $\rho_0, \rho_1, \rho_2, \dots, \rho_{p-1}$ como condiciones iniciales determinadas a partir de los coeficientes $\phi_1, \phi_2, \dots, \phi_p$, la solución de la ecuación (3-41) permite calcular los valores de ρ_τ para $\tau \geq p$.

Particularizando (3-41) para $\tau = 1, 2, \dots, p$ se obtiene el sistema de ecuaciones de Yule-Walker:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\dots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \quad (3-42)$$

Resolviendo el anterior sistema resulta que

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} \quad (3-43)$$

Si en el proceso (3-35) multiplicamos ambos miembros por $\phi^{-1}(L)$ se obtiene

$$Y_t = \phi^{-1}(L) \varepsilon_t = \frac{1}{\phi(L)} \varepsilon_t \quad (3-44)$$

De esta forma se ha pasado de un modelo AR(p) a un modelo MA(∞) según se puede ver al desarrollar el segundo miembro de (3-44) teóricamente de análoga forma a como se hizo en el caso del modelo AR(2).

3.2. Modelos de medias móviles (MA)

Un modelo de medias móviles de orden q , o brevemente un modelo MA(q), se define de la siguiente forma:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3-45)$$

Conviene señalar que los coeficientes $\theta_1, \dots, \theta_q$ van precedidos por el signo negativo, por cuestión meramente de conveniencia en la notación.

Utilizando el operador polinomial de retardos,

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q \quad (3-46)$$

el modelo de medias móviles se puede expresar de forma compacta

$$Y_t = \theta(L) \varepsilon_t \quad (3-47)$$

En el modelo (3-47) la media es cero, cualesquiera que sean los valores de θ_i . En efecto

$$E[Y_t] = \theta(L) E[\varepsilon_t] = 0 \quad (3-48)$$

Si en el modelo (3-45) se incluye un término constante

$$Y_t = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3-49)$$

entonces al tomar esperanzas matemáticas en la expresión anterior resulta

$$E[Y_t] = \delta \quad (3-50)$$

Así pues, en los modelos de medias móviles, la media del proceso coincide con el término independiente, que aparece en el segundo miembro. Sin pérdida de generalidad se supondrá en lo sucesivo que $\delta = 0$.

A continuación se examinan las propiedades de un MA(1) y de un MA(2), para generalizarlas posteriormente a un MA(q).

Modelo MA(1)

Un modelo MA(1) viene definido por

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} = (1 - \theta_1 L) \varepsilon_t \quad (3-51)$$

Si se multiplican ambos miembros de (3-51) por $Y_{t-\tau}$, y se toman esperanzas matemáticas se tiene que

$$E[Y_t Y_{t-\tau}] = E[\varepsilon_t - \theta_1 \varepsilon_{t-1}] [\varepsilon_{t-\tau} - \theta_1 \varepsilon_{t-\tau-1}] \quad (3-52)$$

Puede verse que para $\tau = 0$, teniendo en cuenta que $E[\varepsilon_t \varepsilon_{t'}] = 0$ para $t \neq t'$, resulta

$$\gamma_0 = E[Y_t]^2 = E[\varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 - 2\theta_1 \varepsilon_t \varepsilon_{t-1}] = [1 + \theta_1^2] \sigma_\varepsilon^2 \quad (3-53)$$

Para $\tau = 1$

$$\gamma_1 = E[\varepsilon_t - \theta_1 \varepsilon_{t-1}] [\varepsilon_{t-1} - \theta_1 \varepsilon_{t-2}] = -\theta_1 \sigma_\varepsilon^2 \quad (3-54)$$

Para valores de $\tau > 1$, se deduce que

$$\gamma_\tau = 0 \quad \tau > 1 \quad (3-55)$$

Dividiendo ambos miembros de (3-54) y (3-55) por γ_0 se obtienen los coeficientes de autocorrelación

$$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2} \quad \tau = 1 \quad (3-56)$$

$$\rho_\tau = 0 \quad \tau > 1 \quad (3-57)$$

En la figura 3.5 se ha representado el correlograma de un modelo MA(1) para valores del parámetro $\theta_1 = 0,8$ y $\theta_1 = -0,8$.

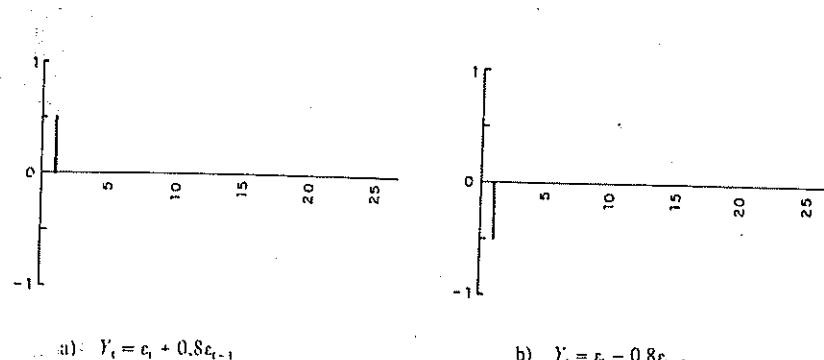
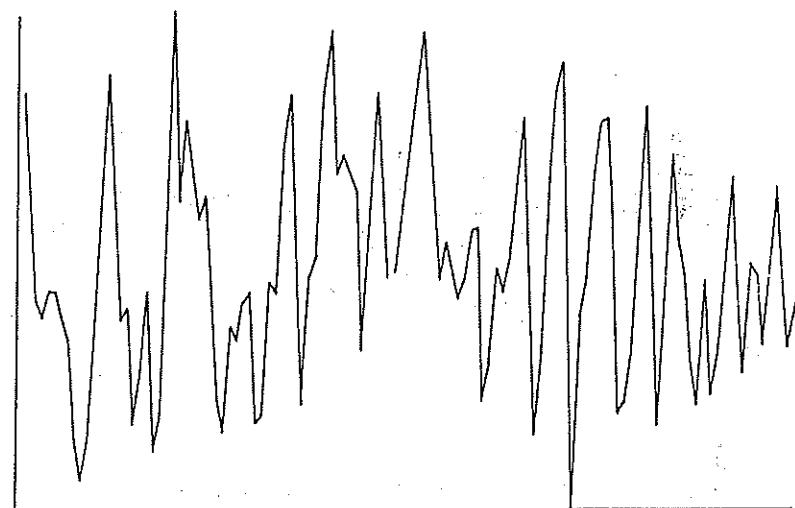
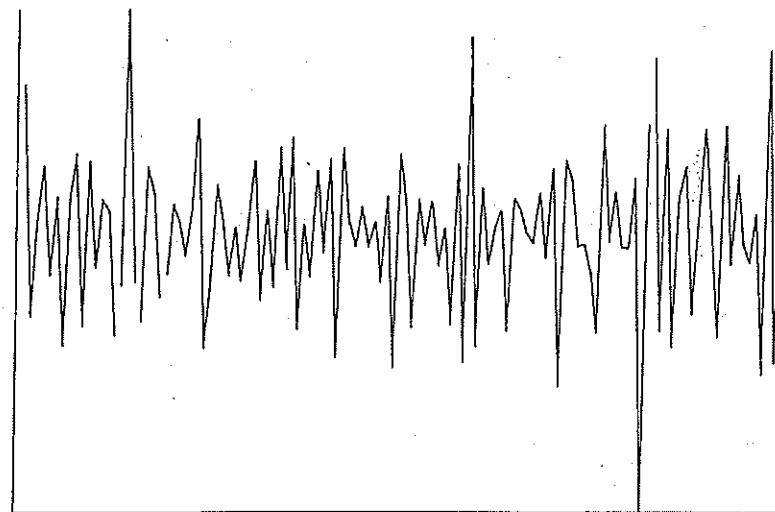


Fig. 3.5. Funciones de autocorrelación de modelos MA(1).

Asimismo, en la figura 3.6 se han representado las realizaciones correspondientes a estos procesos.



a) $Y_t = \varepsilon_t + 0,8\varepsilon_{t-1}$



b) $Y_t = \varepsilon_t - 0,8\varepsilon_{t-1}$

Fig. 3.6. Realizaciones de modelos MA(1).

A la vista de la expresión (3-53) puede verse que un modelo MA(1) es siempre estacionario con independencia del valor que tenga el parámetro θ_1 .

Si el modelo (3-51) se expresa así:

$$\varepsilon_t = Y_t + \theta_1 \varepsilon_{t-1}$$

y se efectúan sustituciones sucesivas, resulta que

$$\begin{aligned} \varepsilon_t &= Y_t + \theta_1 [Y_{t-1} + \theta_1 \varepsilon_{t-2}] = \\ &\dots \\ &= Y_t + \theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots + \theta_1^N Y_{t-N} + \theta_1^{N+1} \varepsilon_{t-N-1} \end{aligned} \quad (3-58)$$

Si $|\theta_1| < 1$, el último término de (3-58) tiene menos peso a medida que sea mayor N , y además el peso de cada Y_t retardada decrece a medida que crece el número de retardos. En consecuencia, si se cumple siempre esta condición podrá expresarse así el modelo (3-51):

$$\varepsilon_t = Y_t + \theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots \quad (3-59)$$

De esta forma, se ha pasado de un MA(1) a un AR(∞). A la condición que ha permitido pasar de un modelo a otro, es decir, que $|\theta_1| < 1$ se le denomina condición de invertibilidad.

También, cuando $|\theta_1| < 1$ puede hacerse

$$\varepsilon_t = \frac{1}{1 - \theta_1 L} Y_t = \sum_{j=0}^{\infty} (\theta_1 L)^j = Y_t + \theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots$$

La ecuación (3-51) puede contemplarse como una ecuación en diferencias de ε_t (Y_t constituiría en este caso la parte no homogénea). Para que esta ecuación sea estable se requiere que la raíz del polinomio característico

$$1 - \theta_1 L = 0$$

caiga fuera del círculo unidad, es decir

$$|L| = \left| \frac{1}{\theta_1} \right| > 1$$

o, de forma equivalente, que

$$|\theta_1| < 1$$

Como puede verse, la condición de invertibilidad de un modelo MA(1) es equivalente en sentido formal a la condición de estacionariedad de un modelo AR(1). Ahora bien, conviene no olvidar que un modelo MA(1) es siempre estacionario y que la condición de invertibilidad se establece para poder pasar a un modelo AR(∞).

El problema de la invertibilidad tiene otras implicaciones. Para ver una de ellas, se

va a calcular ρ_1 para un valor θ_1^* tal que $\theta_1^* > 1$. Evidentemente de acuerdo con (3-56)

$$\rho_1^* = \frac{-\theta_1^*}{1 + \theta_1^{*2}} \quad (3-60)$$

La función de autocorrelación anterior corresponde a un proceso no invertible. En cambio, si se toma como parámetro $\theta^{**} = 1/\theta^*$ el proceso será necesariamente invertible. El coeficiente de autocorrelación de primer orden de este nuevo proceso será

$$\rho_1^{**} = \frac{-(1/\theta^*)}{1 + (1/\theta^*)^2} = \frac{-1/\theta^*}{\theta^{*2} + 1} = \frac{-\theta_1^*}{1 + \theta_1^{*2}} \quad (3-61)$$

En general, puede concluirse que a cada proceso no invertible le corresponde otro proceso invertible que tiene exactamente el mismo coeficiente de autocorrelación. Esto plantea problemas de identificación, ya que, invirtiendo los términos, cuando se tiene un coeficiente de autocorrelación no se sabe si pertenece a un proceso invertible o a uno no invertible. La única excepción a esta ambigüedad se presenta cuando $|\theta_1| = 1$.

Con objeto de evitar esta ambigüedad, y a fin de lograr unas predicciones eficientes, se supondrá siempre que los procesos considerados son invertibles.

Al comparar un AR(1) estacionario con un MA(1) se ve que en el primero la función de autocorrelación tiene infinitos términos no nulos mientras que en este último existe un sólo término no nulo. En lo que respecta al coeficiente de autocorrelación de primer orden, existe una diferencia entre ambos tipos de proceso. Mientras que en un AR(1): ρ_1 únicamente está limitado por la condición de que $|\rho_1| < 1$, en un MA(1), el máximo valor que puede tomar en términos absolutos es $|\rho_1| = 1/2$.

En efecto, derivando (3-56) con respecto a θ_1 , e igualando a 0, resulta

$$\frac{d\rho_1}{d\theta_1} = \frac{-(1 + \theta_1^2) + 2\theta_1^2}{(1 + \theta_1^2)^2} = 0 \quad (3-62)$$

Resolviendo

$$\theta_1 = \pm 1$$

el máximo de ρ_1 en valor absoluto será, de acuerdo con (3-56),

$$\rho_1 = |1/2|$$

Modelo MA(2)

Un modelo MA(2) viene definido por

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} = (1 - \theta_1 L - \theta_2 L^2) \varepsilon_t \quad (3-63)$$

Multiplicando ambos miembros de (3-63) por $Y_{t-\tau}$ y tomando esperanzas, resulta

$$E[Y_t Y_{t-\tau}] = E[\epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}] [\epsilon_{t-\tau} - \theta_1 \epsilon_{t-\tau-1} - \theta_2 \epsilon_{t-\tau-2}] \quad (3-64)$$

Para distintos valores de τ se obtienen los siguientes resultados

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2) \sigma_\epsilon^2 \quad \tau = 0 \quad (3-65)$$

$$\gamma_1 = (-\theta_1 + \theta_1 \theta_2) \sigma_\epsilon^2 \quad \tau = 1 \quad (3-66)$$

$$\gamma_2 = (-\theta_2) \sigma_\epsilon^2 \quad \tau = 2 \quad (3-67)$$

$$\gamma_\tau = 0 \quad \tau > 2 \quad (3-68)$$

A partir de las expresiones anteriores se obtienen fácilmente los coeficientes de autocorrelación

$$\rho_1 = \frac{-\theta_1 + \theta_1 \theta_2}{1 + \theta_1^2 + \theta_2^2} \quad (3-69)$$

$$\rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} \quad (3-70)$$

$$\rho_\tau = 0 \quad \tau > 2 \quad (3-71)$$

Para que un proceso MA(2) sea invertible se requiere que las raíces del polinomio característico

$$1 - \theta_1 L - \theta_2 L^2 = 0 \quad (3-72)$$

caigan fuera del círculo unidad.

En las figuras 3.7 y 3.8 se han representado respectivamente los correlogramas y realizaciones correspondientes a cuatro modelos MA(2) invertibles.

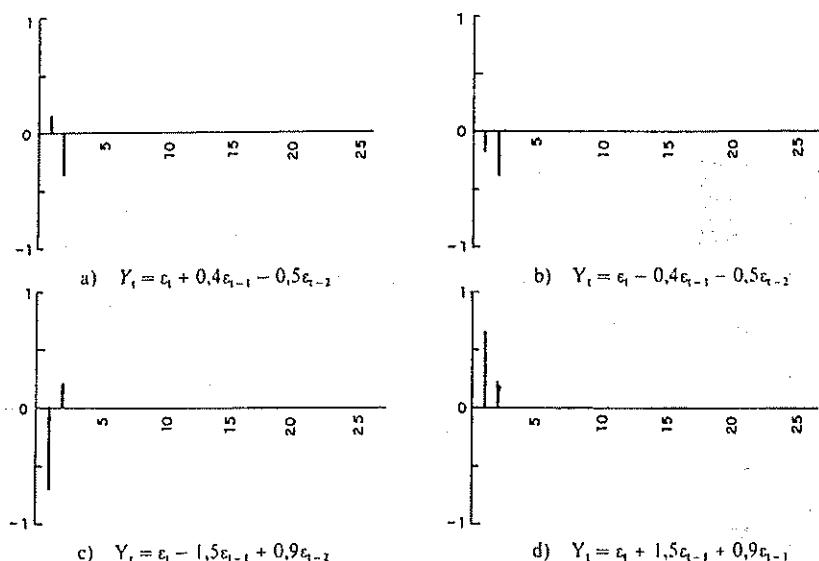


Fig. 3.7. Funciones de autocorrelación de modelos MA(2).

NUMERO DE OBSERVACIONES: 120

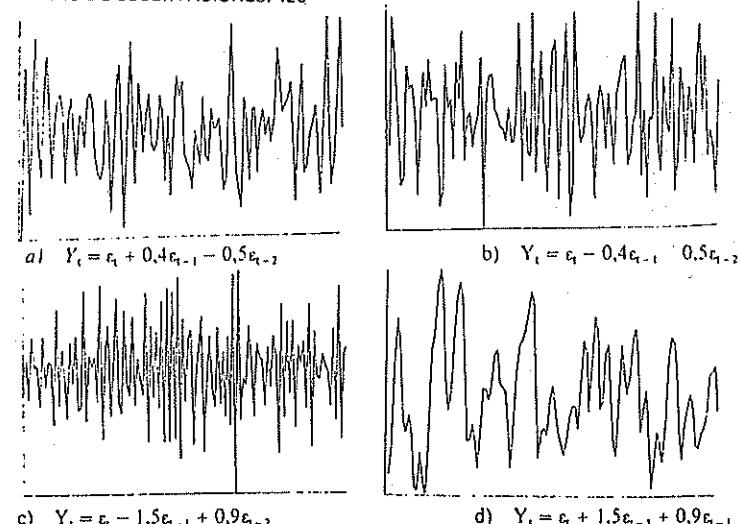


Fig. 3.8. Realizaciones de modelos MA(2).

Modelo MA(q)

La generalización de las propiedades de un modelo MA(q) no presenta ningún problema. Si en (3-45) se multiplican ambos miembros por $Y_{t-\tau}$ y se toman esperanzas, se obtienen los siguientes resultados

$$\gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2)\sigma_\epsilon^2 \quad (3-73)$$

$$\gamma_\tau = \begin{cases} (-\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_{q-\tau}\theta_q)\sigma_\epsilon^2 & \tau = 1, 2, \dots, q \\ 0 & \tau > q \end{cases} \quad (3-74)$$

Por tanto, en las autocovarianzas existe un corte brusco en q , ya que después de ese retardo su valor es igual a 0. El mismo fenómeno se presenta con los coeficientes de autocorrelación que vendrán dados por

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_{q-\tau}\theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & \tau = 1, 2, \dots, q \\ 0 & \tau > q \end{cases} \quad (3-75)$$

Cuando se analizaron las propiedades de un MA(1) se vio que $|\rho_1| \leq 1/2$. En general, los coeficientes de autocorrelación ρ_τ de un MA(q) están acotados. Davies y otros (1974) demostraron que

$$\text{Max } \rho_\tau = \begin{cases} \cos \pi/(H+1) & \text{si } q+1 \text{ es divisible por } \tau \\ \cos \pi/(H+1) & \text{si } q+1 \text{ no es divisible por } \tau \end{cases}$$

donde H es el número entero mayor que no excede de $(q+1)/\tau$. Aplicando la fórmula anterior para ver el máximo de ρ_1 en un MA(1) y teniendo en cuenta que $H=2$ ya que $(q+1)/\tau=2$, se obtiene el siguiente resultado

$$\max \rho_1 = \cos(\pi/3) = 0,5$$

En el cuadro 3.1 se han recogido los valores máximos de los coeficientes de autocorrelación en procesos de medias móviles de distinto orden.

Para que un modelo MA(q) sea invertible se requiere que las raíces de la ecuación polinomial

$$1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q = 0$$

caigan fuera del círculo unidad.

Cuadro 3.1. Valores máximos de los coeficientes de autocorrelación en procesos de medias móviles de distinto orden

Orden del retardo (τ)	Orden del proceso de medias móviles (q)												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0,50	0,71	0,81	0,87	0,90	0,92	0,94	0,95	0,96	0,97	0,97	0,97	0,98
2		0,50	0,50	0,71	0,71	0,81	0,81	0,87	0,87	0,90	0,90	0,92	0,92
3			0,50	0,50	0,50	0,71	0,71	0,71	0,81	0,81	0,81	0,87	0,87
4				0,50	0,50	0,50	0,50	0,71	0,71	0,71	0,71	0,81	0,81
5					0,50	0,50	0,50	0,50	0,50	0,71	0,71	0,71	0,71
6						0,50	0,50	0,50	0,50	0,50	0,50	0,71	0,71
7							0,50	0,50	0,50	0,50	0,50	0,50	0,50
8								0,50	0,50	0,50	0,50	0,50	0,50
9									0,50	0,50	0,50	0,50	0,50
10										0,50	0,50	0,50	0,50
11											0,50	0,50	0,50
12												0,50	0,50
13													0,50

3.3. Modelos mixtos autorregresivos-medias móviles (ARMA)

Un modelo ARMA (p, q) viene definido de la siguiente forma:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (3-76)$$

Utilizando los operadores polinomiales de retardo, el modelo queda expresado en forma compacta del siguiente modo:

$$\phi(L)Y_t = \theta(L)\epsilon_t \quad (3-77)$$

Para que el modelo sea estacionario se requiere que las raíces de la ecuación polinomial

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p = 0 \quad (3-78)$$

caigan fuera del círculo unidad.

Si se cumplen las condiciones de estacionariedad, el modelo ARMA (p, q) se puede expresar como un MA(∞), pudiendo representarse de la siguiente forma:

$$Y_t = \frac{\theta(L)}{\phi(L)} \epsilon_t = \psi(L)\epsilon_t \quad (3-79)$$

MODELOS LINEALES

Por tanto los coeficientes del operador polinomial $\psi(L)$, que tiene infinitos elementos, deben cumplir la siguiente identidad

$$\phi(L)\psi(L) \equiv \theta(L) \quad (3-80)$$

o, en notación más detallada,

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1 + \psi_1 L + \psi_2 L^2 + \dots) \equiv (1 - \theta_1 L - \dots - \theta_q L^q) \quad (3-81)$$

A partir de la identidad (3-81), se pueden deducir un conjunto de ecuaciones que nos permiten obtener los ψ_i , en función de los coeficientes ϕ_i y θ_i . Así en un modelo ARMA (1,1), la identidad anterior sería

$$(1 - \phi_1 L)(1 + \psi_1 L + \psi_2 L^2 \dots) = (1 - \theta_1 L) \quad (3-82)$$

Realizando la multiplicación del primer miembro y agrupando según el grado del operador de retardos, se tiene que:

$$[1 + (\psi_1 - \phi_1)L + (\psi_2 - \phi_1\psi_1)L^2 + (\psi_3 - \phi_1\psi_2)L^3 + \dots] = (1 - \theta_1 L) \quad (3-83)$$

Igualando término a término los coeficientes de las distintas potencias de L en ambos miembros, se obtiene

$$\psi_1 - \phi_1 = -\theta_1 \quad \text{de donde} \quad \psi_1 = \phi_1 - \theta_1$$

$$\psi_2 - \phi_1\psi_1 = 0 \quad \text{de donde} \quad \psi_2 = \phi_1\psi_1$$

Para $\tau > 1$ los coeficientes ψ_τ se pueden obtener de forma recursiva a partir de la ecuación en diferencias

$$\psi_\tau = \phi_1\psi_{\tau-1} \quad \text{para } \tau > 1 \quad (3-84)$$

De forma análoga en un modelo ARMA (p,q), los coeficientes $\psi_1, \psi_2, \dots, \psi_q$ se pueden obtener de ecuaciones derivadas del sistema (3-80). Los valores de ψ_τ para $\tau > q$ se pueden deducir de la siguiente ecuación en diferencias

$$\psi_\tau = \phi_1\psi_{\tau-1} + \dots + \phi_p\psi_{\tau-p} \quad \tau > q \quad (3-85)$$

Para que el modelo (3-76) sea invertible, se requiere que las raíces de la ecuación polinomial

$$\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q = 0 \quad (3-86)$$

caigan fuera del círculo unidad.

MODELOS MIXTOS AUTORREGRESIVOS-MEDIAS MÓVILES (ARMA)

Si se cumplen las condiciones de invertibilidad, el modelo ARMA(p, q) se puede expresar mediante un AR(∞):

$$\varepsilon_t = \frac{\phi(L)}{\theta(L)} Y_t = \pi(L)Y_t \quad (3-87)$$

donde

$$\pi(L) = 1 - \pi_1 L - \pi_2 L^2 - \dots$$

Los coeficientes del operador polinomial $\pi(L)$ deben cumplir la siguiente igualdad

$$\pi(L)\theta(L) \equiv \phi(L) \quad (3-88)$$

Siguiendo un idéntico razonamiento al que se hizo para pasar a un MA(∞), se obtiene que para $\tau > p$ los coeficientes π_τ siguen la siguiente ecuación en diferencias

$$\pi_\tau = \theta_1\pi_{\tau-1} + \dots + \theta_q\pi_{\tau-q} \quad \tau > p \quad (3-89)$$

En el modelo (3-76) la media es 0. Si se añade al segundo miembro un término constante δ , la media del proceso se deduce a partir de la siguiente expresión:

$$E[\phi(L)Y_t] = \delta + E[\theta(L)\varepsilon_t] \quad (3-90)$$

Por tanto

$$\phi(L)E[Y_t] = \delta \quad (3-91)$$

Si el proceso es estacionario entonces $E[Y_t] = \mu \forall t$, con lo que

$$\mu = \frac{\delta}{1 - \phi_1 L - \phi_p L^p} = \frac{\delta}{1 - \phi_1 - \dots - \phi_p} \quad (3-92)$$

ya que al aplicar el operador L a una constante —en este caso δ — se obtiene el valor de la constante.

A continuación se examinan las propiedades de un modelo ARMA (1,1), para después generalizarlas a un proceso ARMA (p, q).

Modelo ARMA (1, 1)

Un modelo ARMA (1,1) viene dado por la expresión

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (3-93)$$

Multiplicando ambos miembros de (3-93) por $Y_{t-\tau}$ y tomando esperanzas tenemos que

$$\gamma_\tau = E[Y_t Y_{t-\tau}] = \phi_1 \gamma_{\tau-1} + E[\varepsilon_t Y_{t-\tau}] - \theta_1 E[\varepsilon_{t-1} Y_{t-\tau}]$$

Teniendo en cuenta que

$$E[\varepsilon_t Y_t] = \sigma_\varepsilon^2 \quad (3-94)$$

$$E[\varepsilon_{t-1} Y_t] = E[\varepsilon_{t-1} (\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1})] = (\phi_1 - \theta_1) \sigma_\varepsilon^2 \quad (3-95)$$

Para $\tau = 0$ se obtiene la siguiente expresión

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_\varepsilon^2 - \theta_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2 \quad (3-96)$$

Para $\tau = 1$, fácilmente se comprueba que:

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2 \quad (3-97)$$

Para $\tau > 1$ resulta que

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} \quad \tau > 1 \quad (3-98)$$

Sustituyendo el valor de γ_τ de (3-97) en (3-96) se obtiene

$$\gamma_0 = \phi_1 [\phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2] + \sigma_\varepsilon^2 - \theta_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2$$

Operando

$$\gamma_0 = \frac{1 - 2\theta_1 \phi_1 + \theta_1^2}{1 - \phi_1^2} \sigma_\varepsilon^2 \quad (3-99)$$

Sustituyendo a su vez este valor obtenido en (3-97) y operando se tiene que

$$\gamma_1 = \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_\varepsilon^2 \quad (3-100)$$

Dividiendo ambos miembros de (3-100) y de (3-98) por γ_0 se obtiene la siguiente secuencia de coeficientes de autocorrelación

$$\rho_1 = \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 - 2\theta_1 \phi_1 + \theta_1^2} \quad (3-101)$$

$$\rho_\tau = \phi_1 \rho_{\tau-1} \quad \tau > 1 \quad (3-102)$$

En la figura 3.9 se ha representado la función de autocorrelación correspondiente a cuatro modelos ARMA (1, 1). Los coeficientes de autocorrelación decrecen siempre en valor absoluto tomando como condición inicial ρ_1 . Recuérdese que en los modelos AR(1) existía también un decrecimiento exponencial pero la condición inicial era ρ_0 . En algunos modelos, como son los casos a) y b), la parte de medias móviles atenúa considerablemente la condición inicial a partir de la cual se inicia el decaimiento exponencial. En la figura 3.10 se ha representado una realización correspondiente a cada uno de los modelos.

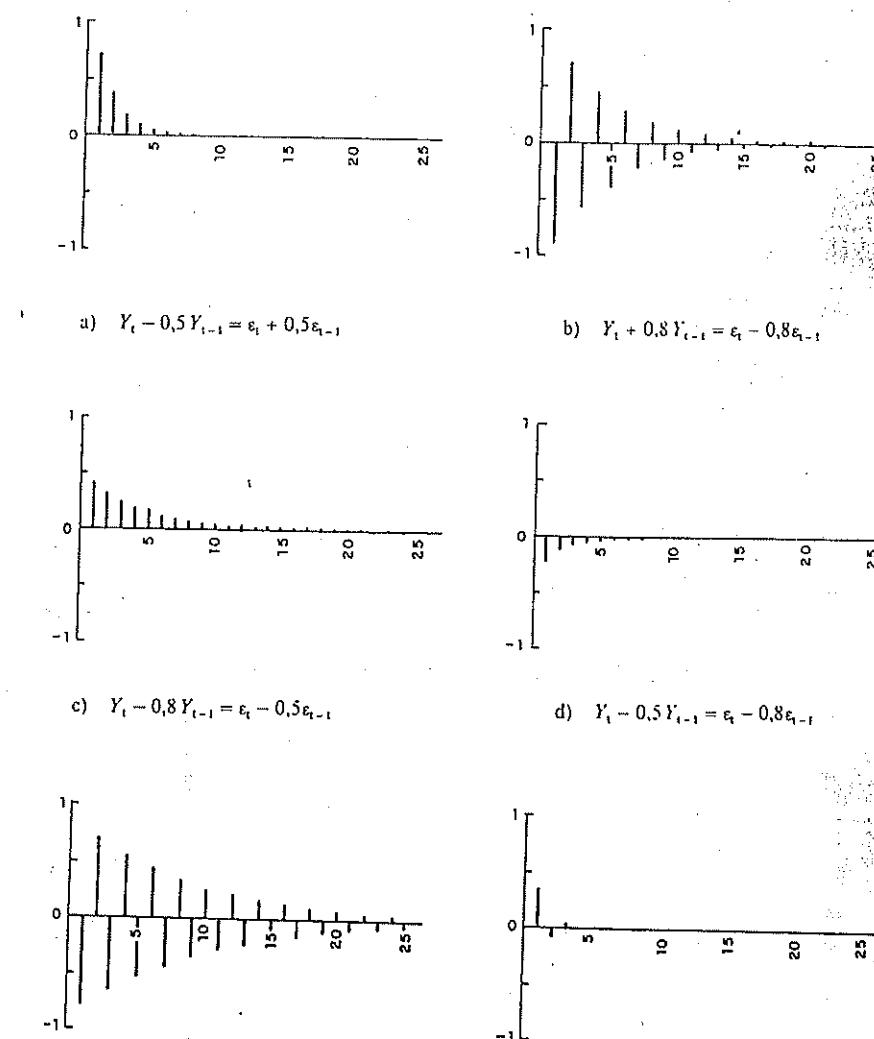


Fig. 3.9. Funciones de autocorrelación de modelos ARMA (1,1).

NUMERO DE OBSERVACIONES: 120

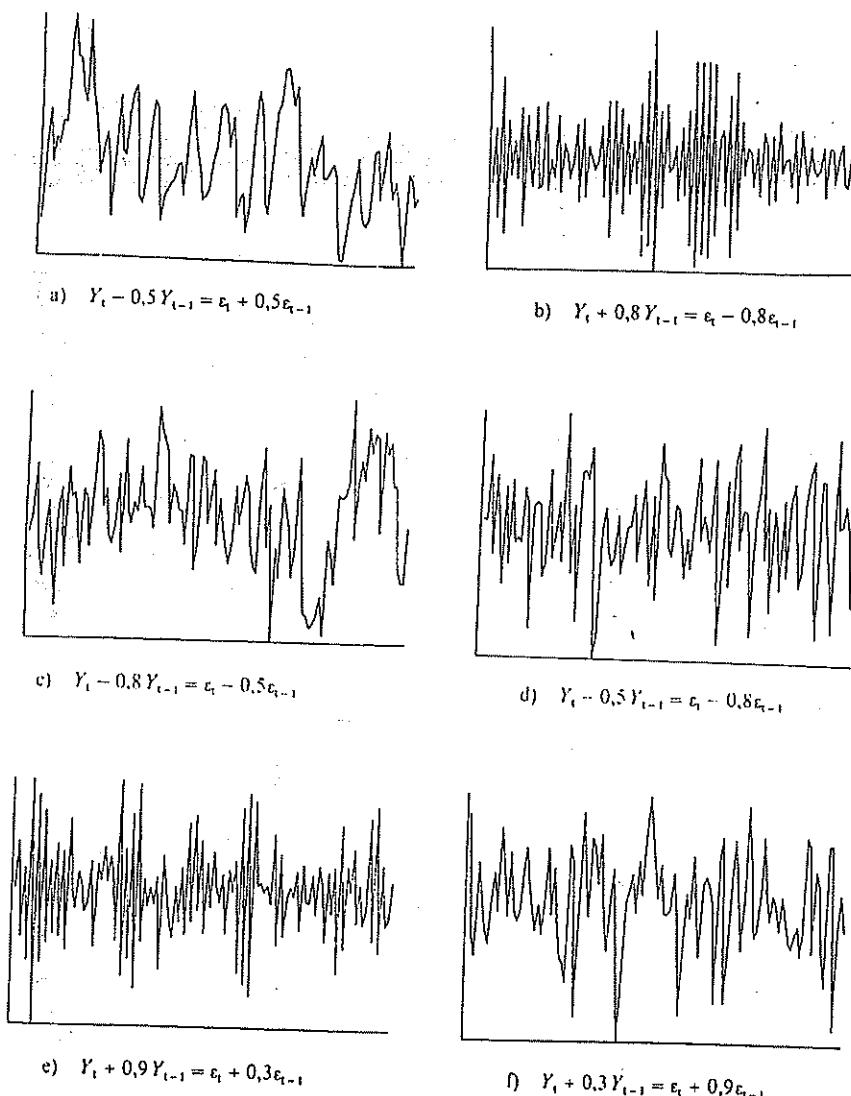


Fig. 3.10. Realizaciones de modelos ARMA (1,1).

Modelo ARMA (p, q)

Un modelo ARMA (p, q) viene definido según (3-76). Si en esta expresión se multiplican ambos miembros por $Y_{t-\tau}$ y se toman esperanzas, se obtiene:

$$\gamma_\tau - \phi_1 \gamma_{\tau-1} - \dots - \phi_p \gamma_{\tau-p} = E[\epsilon_t Y_{t-\tau}] - \theta_1 E[\epsilon_{t-1} Y_{t-\tau}] - \dots - \theta_q E[\epsilon_{t-q} Y_{t-\tau}] \quad (3-103)$$

En la evaluación de las esperanzas del segundo miembro es preciso tener en cuenta que:

$$E[\epsilon_t Y_{t'}] = 0 \quad \text{para } t' < t \quad (3-104)$$

Por tanto, se verifica que:

$$\gamma_\tau - \phi_1 \gamma_{\tau-1} - \dots - \phi_p \gamma_{\tau-p} = 0 \quad \text{para } \tau > q \quad (3-105)$$

Esta ecuación en diferencias permite obtener las autocovarianzas a partir de $\tau > q$.

Análogamente, dividiendo en (3-105) por γ_0 se obtiene:

$$\rho_\tau - \phi_1 \rho_{\tau-1} - \dots - \phi_p \rho_{\tau-p} = 0 \quad \text{para } \tau > q \quad (3-106)$$

que es la ecuación en diferencias que permite generar los coeficientes de autocorrelación para $\tau > q$. En la determinación de los q primeros valores de ρ_τ interviene la parte de medias móviles del modelo.

En los modelos ARMA (p, q) es conveniente factorizar la parte autorregresiva y la parte de medias móviles, con objeto de analizar si existen raíces comunes.

Sean $\lambda_1, \lambda_2, \dots, \lambda_p$ las raíces de la ecuación en diferencias

$$\lambda^p - \phi_1 \lambda^{p-1} - \dots - \phi_p = 0$$

y sean $\delta_1, \delta_2, \dots, \delta_q$ las raíces de la ecuación en diferencias

$$\delta^q - \theta_1 \delta^{q-1} - \dots - \theta_q = 0$$

Una forma alternativa, por tanto, de expresar (3-76) sería la siguiente:

$$(1 - \lambda_1 L) \dots (1 - \lambda_p L) Y_t = (1 - \delta_1 L) \dots (1 - \delta_q L) \epsilon_t \quad (3-107)$$

Si existiera alguna raíz idéntica en los dos miembros (por ejemplo, si $\lambda_1 = \delta_q$) entonces el modelo (3-76) estaría sobreparametrizado innecesariamente, ya que un modelo con las mismas propiedades sería un ARMA ($p-1, q-1$). En este caso, se verificaría que

$$(1 - \lambda_1 L) \dots (1 - \lambda_p L) Y_t = (1 - \delta_1 L) \dots (1 - \delta_{q-1} L) \epsilon_t \quad (3-108)$$

3.4. Procesos no estacionarios

En los procesos tratados en los tres primeros epígrafes se ha impuesto las condiciones de estacionariedad y/o invertibilidad. Como ya se ha indicado, en buena parte las series económicas se deben considerar generadas por procesos no estacionarios. Así pues, si se desea obtener un tratamiento de las series basado en la teoría de los procesos estocásticos es necesario ampliar nuestro campo para incluir los procesos, al menos algunos tipos, no estacionarios.

En principio cabe imaginar infinitas formas por las cuales se puede introducir la no estacionariedad en un proceso. Sin embargo interesa considerar solamente ciertos tipos de procesos no estacionarios que sean adecuados para describir el comportamiento de series económicas, y, al mismo tiempo, sean fácilmente transformables en procesos estacionarios con objeto de utilizar las ventajas que ofrecen estos últimos. Desde esta perspectiva, dentro de los modelos no estacionarios se considerará en primer lugar y a modo de introducción el modelo:

$$Y_t = Y_{t-1} + \epsilon_t \quad (3-109)$$

El modelo anterior es un AR(1) con el coeficiente $\phi_1 = 1$. A este modelo se le denomina también «paseo aleatorio», o «random walk» en terminología anglosajona. Cuando el proceso se inicia en un pasado remoto, mediante sustituciones sucesivas se puede expresar así:

$$Y_t = \sum_{j=0}^t \epsilon_{t-j} \quad (3-110)$$

La varianza del proceso es infinita, como ya se indicó al analizar (3-11), y el proceso, en consecuencia, es no estacionario.

Haciendo de nuevo el supuesto de que el proceso se inicia en $-N$, la varianza para el período t viene dada por la siguiente expresión:

$$\gamma_{tt} = (t + N)\sigma_\epsilon^2 \quad (3-111)$$

El proceso tampoco es estacionario, ya que toma un valor distinto para cada t .

De manera análoga, se obtiene la autocovarianza de orden τ para el momento t

$$\gamma_{t,t-\tau} = (t + N - \tau)\sigma_\epsilon^2 \quad (3-112)$$

El coeficiente de autocorrelación de orden τ para el momento t será

$$\rho_{t,\tau} = \frac{Y_{t-\tau}}{\sqrt{Y_{tt}} \sqrt{Y_{t,t-\tau}}} = \frac{t + N - \tau}{\sqrt{t + N} \sqrt{t + N - \tau}} = \sqrt{\frac{t + N - \tau}{t + N}} \quad (3-113)$$

El valor que toma ρ_t depende claramente del período t de referencia. Sin embargo para t suficientemente grande su valor se situaría muy cerca de 1. Cuando t tiende hacia infinito —o bien el período inicial se remonta hacia $-\infty$ — entonces el coeficiente de autocorrelación no está definido.

Sin embargo, el proceso (3-109) puede transformarse fácilmente en estacionario. En efecto, el proceso

$$w_t = Y_t - Y_{t-1} = \Delta Y_t \quad (3-114)$$

de acuerdo con la ecuación (3-109), se tiene que

$$w_t = \Delta Y_t = \epsilon_t \quad (3-115)$$

es decir, el proceso transformado —obtenido al tomar primeras diferencias en el proceso original— es ruido blanco.

Tomando diferencias de primer orden se pasa de Y_t al proceso w_t . Consideremos ahora el caso inverso, es decir, cómo obtener Y_t a partir del proceso w_t . Por sustituciones sucesivas se tiene que

$$\begin{aligned} Y_t &= w_t + Y_{t-1} = w_t + w_{t-1} + Y_{t-2} \\ &\dots \\ &= w_t + w_{t-1} + w_{t-2} + w_{t-3} + w_{t-4} + \dots \end{aligned}$$

Por tanto, el proceso Y_t se obtiene sumando, o lo que es lo mismo, integrando el proceso w_t .

Por esta razón, se dice que el «paseo aleatorio» pertenece a la clase de modelos integrados. Esta clase está constituida por todos aquellos modelos que se pueden transformar en estacionarios mediante la toma de diferencias de un determinado orden, o dicho de otra forma, modelos integrados son aquellos que se pueden obtener mediante suma o integración de un proceso estacionario. A estos modelos se les denomina también modelos no estacionarios homogéneos y han sido estudiados por Tintner (1940), Tintner y Rao (1963), Yaglom (1955) y, más recientemente, por Box y Jenkins (1976), siendo estos últimos autores los que más han contribuido a su divulgación.

A un proceso integrado Y_t se le denomina proceso ARIMA(p, d, q) si tomando diferencias de orden d se obtiene un proceso estacionario w_t del tipo ARMA (p, q). La I central del término ARIMA indica integrado.

Así pues, se tendría que

$$\begin{aligned} w_t &= \Delta^d Y_t = (1 - L)^d Y_t \\ (1 - \phi_1 L - \dots - \phi_p L^p) w_t &= (1 - \theta_1 L - \dots - \theta_q L^q) \epsilon_t \end{aligned} \quad (3-116)$$

o sustituyendo w_t en la segunda expresión:

$$(1 - \phi_1 L - \dots - \phi_p L^p) (1 - L)^d Y_t = (1 - \theta_1 L - \dots - \theta_q L^q) \epsilon_t$$

En forma más compacta, un proceso ARIMA (p, d, q) se define así

$$\phi(L)(1 - L)^d Y_t = \theta(L) \epsilon_t \quad (3-117)$$

Los modelos ARIMA constituyen una clase particular de procesos no estacionarios. Sin embargo, en muchos casos son suficientes para representar el comportamiento de las series económicas. Sobre todo si al mismo tiempo se tienen en cuenta otro tipo de transformaciones que se exponen a continuación.

Cuando una serie económica se observa a lo largo de un período dilatado de tiempo ocurre con frecuencia que también la varianza viene afectada por una tendencia, y esta tendencia no desaparece al tomar diferencias. Cuando se presenta esta circunstancia la transformación adecuada puede consistir en tomar logaritmos. Los estadísticos Box y Cox (1964) definieron una transformación instantánea —en el sentido de que no están involucrados simultáneamente varios períodos de tiempo— de carácter más general. Esta transformación se define por:

$$Y_t^{(0)} = \begin{cases} (Y_t^\lambda - 1)/\lambda & \lambda \neq 0 \\ \ln Y_t & \lambda = 0 \end{cases} \quad (3-118)$$

Cuando el parámetro es $\lambda = 1$, la transformación Box-Cox consiste prácticamente en tomar valores originales.

Cuando el parámetro es $\lambda = 0$, sustituyendo en la primera igualdad se ve que da lugar a una indeterminación del tipo 0/0. Sin embargo aplicando la regla de l'Hôpital se tiene que

$$\lim_{\lambda \rightarrow 0} Y_t^{(0)} = \lim_{\lambda \rightarrow 0} \frac{Y_t^\lambda - 1}{\lambda} = \lim_{\lambda \rightarrow 0} \frac{Y_t^\lambda \ln Y_t}{1} = \ln Y_t \quad (\text{para } Y_t > 0)$$

que es justamente el valor propuesto para $Y_t^{(0)}$ en (3-118) cuando $\lambda = 0$.

Combinando una diferencia con la toma de logaritmos se obtiene la transformación $\Delta \ln Y_t$, que reviste un gran interés, ya que es adecuada para el tratamiento de la mayoría de las series económicas, si no se tiene en cuenta el problema de la estacionalidad. Por otra parte a la expresión $\Delta \ln Y_t$ se le denomina tasa de variación natural, siendo una aproximación a la tasa de variación relativa. Veamos a continuación en qué consiste esta aproximación.

Si se efectúa un desarrollo de Taylor⁷ de $\ln Y_t = \ln(Y_t/Y_{t-1})$ en torno al valor 1 se obtiene que

$$\ln Y_t = \ln Y_{t-1} + \frac{1}{Y_{t-1}} (Y_t - Y_{t-1}) + \dots$$

⁷ Recuérdese que el desarrollo de Taylor de $f(x)$ en torno al punto a viene dado por:

$$f(x) = f(a) + (x - a)f'(x)|_{x=a} + \frac{(x - a)^2}{2!} f''(x)|_{x=a} + \frac{(x - a)^3}{3!} f'''(x)|_{x=a} + \dots$$

$$\begin{aligned} \ln \frac{Y_t}{Y_{t-1}} &= \left[\frac{Y_t}{Y_{t-1}} - 1 \right] + \left[\frac{1}{Y_t/Y_{t-1}} \right] Y_t/Y_{t-1} - 1 + \\ &+ \frac{1}{2} \left[\frac{Y_t}{Y_{t-1}} - 1 \right]^2 \left[-\frac{1}{(Y_t/Y_{t-1})^2} \right] Y_t/Y_{t-1} - 1 + \\ &+ \frac{1}{3} \left[\frac{Y_t}{Y_{t-1}} - 1 \right]^3 \left[\frac{1}{(Y_t/Y_{t-1})^3} \right] Y_t/Y_{t-1} - 1 + \dots \\ &= \left[\frac{Y_t}{Y_{t-1}} - 1 \right] - \frac{1}{2} \left[\frac{Y_t}{Y_{t-1}} - 1 \right]^2 + \frac{1}{3} \left[\frac{Y_t}{Y_{t-1}} - 1 \right]^3 + \dots \end{aligned} \quad (3-119)$$

Designando por \dot{Y}_t a la tasa de variación relativa (Y_t/Y_{t-1}) - 1 se tiene que

$$\Delta \ln Y_t = \dot{Y}_t - \frac{1}{2} \dot{Y}_t^2 + \frac{1}{3} \dot{Y}_t^3 + \dots \quad (3-120)$$

Tomando la aproximación lineal, es decir, del primer término del segundo miembro de (3-120) se puede establecer que

$$\Delta \ln Y_t \approx \dot{Y}_t \quad (3-121)$$

Esta aproximación sería tanto mejor cuanto menor sea la tasa de variación. Así para un incremento relativo del 5% —es decir, del 0,05— se obtiene que $\ln 1,05 = 0,0487901$. Puede comprobarse que utilizando tres términos se obtiene una aproximación de 5 cifras decimales

$$0,0487901 \approx 0,05 - \frac{1}{2}(0,05)^2 + \frac{1}{3}(0,05)^3 + \dots$$

La tasa de variación natural tiene la ventaja de que goza de la propiedad de aditividad. De esta forma, la tasa de variación natural se puede obtener como suma de las tasas de variación natural de los intervalos en que está dividido el período. Por ejemplo:

$$\begin{aligned} \Delta \ln Y_1 + \Delta \ln Y_2 + \Delta \ln Y_3 &= \\ &= \ln Y_1 - \ln Y_0 + \ln Y_2 - \ln Y_1 + \ln Y_3 - \ln Y_2 = \\ &= \ln Y_3 - \ln Y_0 \\ &= \ln \frac{Y_3}{Y_0} \end{aligned}$$

Ejercicios

[1] Sea el modelo AR(1) siguiente:

$$Y_t = 0,8 Y_{t-1} + \varepsilon_t$$

$$\sigma_\varepsilon^2 = 2$$

Se pide:

- a) ¿Es estacionario?
- b) ¿Es invertible?
- c) Calcular la secuencia y_0, y_1, \dots, y_5
- d) Calcular la secuencia p_1, p_2, \dots, p_5
- e) Calcular los coeficientes $\psi_1, \psi_2, \dots, \psi_5$ del modelo MA(∞) en que, si es posible, puede transformarse el AR(1) dado.

[2] Sea el modelo MA(1) siguiente:

$$Y_t = \varepsilon_t - 0,9 \varepsilon_{t-1}$$

$$\sigma_\varepsilon^2 = 4$$

Se pide:

- a) ¿Es estacionario?
- b) ¿Es invertible?
- c) Calcular la secuencia y_0, y_1, \dots, y_5
- d) Calcular la secuencia p_1, p_2, \dots, p_5
- e) Calcular los coeficientes $\pi_1, \pi_2, \dots, \pi_5$ correspondientes al modelo invertido AR(∞).

[3] Sea el modelo AR(2) siguiente:

$$Y_t = 0,6 Y_{t-1} + 0,3 Y_{t-2} + \varepsilon_t$$

$$\sigma_\varepsilon^2 = 3$$

Se pide:

- a) ¿Es estacionario?
- b) ¿Es invertible?
- c) Calcular la secuencia y_0, y_1, \dots, y_5
- d) Calcular la secuencia p_1, p_2, \dots, p_5

[4] Sea el modelo MA(2) siguiente:

$$Y_t = \varepsilon_t - 0,4 \varepsilon_{t-1} + 1,2 \varepsilon_{t-2}$$

$$\sigma_\varepsilon^2 = 2$$

Se pide:

- a) ¿Es estacionario?
- b) ¿Es invertible?
- c) Calcular la secuencia y_0, y_1, \dots, y_5
- d) Calcular la secuencia p_1, p_2, \dots, p_5

[5] Sea el modelo ARMA(1, 1) siguiente:

$$Y_t = 0,9 Y_{t-1} + \varepsilon_t - 0,8 \varepsilon_{t-1}$$

$$\sigma_\varepsilon^2 = 5$$

Se pide:

- a) ¿Es estacionario?
- b) ¿Es invertible?
- c) Calcular la secuencia y_0, y_1, \dots, y_5
- d) Calcular la secuencia p_1, p_2, \dots, p_5

[6] Sea el proceso AR(1)

$$Y_t = \delta + 0,9 Y_{t-1} + \varepsilon_t$$

donde δ es una constante.

Pruebe que los coeficientes de autocorrelación no dependen de los valores que se asignen a δ .

[7] Se dispone de la siguiente secuencia de coeficientes de autocorrelación:

$$\rho_1 = 0,70 \quad \rho_2 = 0,49 \quad \rho_3 = 0,34 \quad \rho_4 = 0,24 \quad \rho_5 = 0,17$$

¿De qué proceso ARMA(p, q) pueden provenir estos coeficientes? Considere únicamente valores bajos de p y q .

[8] Se dispone de la siguiente secuencia de coeficientes de autocorrelación:

$$\rho_1 = 0,80 \quad \rho_2 = 0,82 \quad \rho_3 = 0,74 \quad \rho_4 = 0,71 \quad \rho_5 = 0,66$$

¿De qué proceso ARMA(p, q) pueden provenir estos coeficientes? Considere únicamente valores bajos de p y q .

[9] En un modelo AR(2)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

Se conoce que $\phi_1 = 0,7$ y $\rho_1 = 0,9$

¿Podría deducir el valor de ϕ_2 ?

[10] En un modelo AR(2)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

¿Podría deducir el valor ϕ_2 si se da la información $\phi_1 = -0,4$ y $\rho_1 = 0,7$?

[11] En un modelo AR(2)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

se conoce que $\phi_1 = 0,8$ y $\phi_2 = 0,6$

¿Podría deducir el valor de ϕ_2 ?

[12] Se dispone de la siguiente secuencia de coeficientes de autocorrelación:

$$\rho_1 = 0,95 \quad \rho_2 = 0,90 \quad \rho_3 = 0,84 \quad \rho_4 = 0,80 \quad \rho_5 = 0,75$$

¿Es verosímil que estos coeficientes puedan proceder de un MA(15)?

NOTA. — Sobre los coeficientes $\rho_6 \rho_7 \dots \rho_{15}$, haga los supuestos que crea más convenientes.

[13] Se sabe que en un país el índice mensual de producción industrial está generado por el siguiente modelo:

$$Y_t = 0,8 Y_{t-1} - 0,17 Y_{t-2} + 40 + \epsilon_t$$

Se pide:

- ¿Es estacionario el proceso?
- En el caso de que las raíces sean imaginarias, exprese la solución de la parte homogénea en forma de función de seno.
- En el país a que va referido el índice se considera la hipótesis de que los ciclos cortos oscilan entre 2 y 3 años. ¿Es consistente el modelo con esta hipótesis?
- El índice tiene como base diciembre 1970 = 100 y en enero de 1971 toma el valor de 102. Tomando estos valores como condiciones iniciales, determine el valor de Y_t en enero de 1984 sin considerar la parte estocástica.

[14] En el modelo MA(2) invertible.

$$Y_t = \epsilon_t - 1,2\epsilon_{t-1} + 0,32\epsilon_{t-2}$$

Se desea calcular los coeficientes π_1 , π_2 , π_3 , π_4 y π_5 del correspondiente AR(∞)

- Utilizando el procedimiento de las fracciones parciales.
- Utilizando la igualdad $\pi(L)\theta(L) = \phi(L)$.

[15] En el modelo AR(2) estacionario

$$Y_t = Y_{t-1} - 0,21 Y_{t-2} + \epsilon_t$$

se desea calcular los coeficientes ψ_1 , ψ_2 , ψ_3 , ψ_4 y ψ_5 del correspondiente MA(∞)

- Utilizando el procedimiento de las fracciones parciales.
- Utilizando la igualdad $\psi(L)\phi(L) = \theta(L)$.

[16] En el modelo ARMA (1, 1)

$$Y_t = 0,8 Y_{t-1} + \epsilon_t + 0,7\epsilon_{t-1}$$

Se pide:

- Calcular los coeficientes ψ_1 , ψ_2 , ψ_3 , ψ_4 y ψ_5 del correspondiente MA(∞)
- Calcular los coeficientes π_1 , π_2 , π_3 , π_4 y π_5 del correspondiente AR(∞)

[17] Sea el modelo MA(2)

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$$

Se sabe que las raíces de la ecuación característica

$$\lambda^2 - \theta_1 \lambda - \theta_2 = 0$$

son $\lambda_1 = 0,6$ y $\lambda_2 = 0,3$

¿Podría determinar cuáles son los valores de θ_1 y θ_2 ?

[18] Sea el modelo ARMA (1, 1)

$$Y_t = 0,6 Y_{t-1} + \epsilon_t - 0,6\epsilon_{t-1}$$

Se pide

- Calcular la función de autocorrelación.
- ¿Puede simplificarse este modelo?

[19] Sea el modelo ARMA (2, 2)

$$Y_t = 0,9 Y_{t-1} - 0,2 Y_{t-2} + \epsilon_t - 1,3\epsilon_{t-1} + 0,4\epsilon_{t-2}$$

¿Puede introducirse alguna simplificación en este modelo?

[20] Sea el modelo

$$Y_t = 1,8 Y_{t-1} - 0,81 Y_{t-2} + \epsilon_t$$

Se pide hallar la solución de la ecuación en diferencias que genera los coeficientes de autocorrelación.

[21] Sea el modelo

$$Y_t = 0,4 Y_{t-1} + 0,9 Y_{t-2} + \epsilon_t$$

Se pide:

- Analizar las condiciones de estacionariedad.
- En el caso de que no sea estacionario, ¿puede hacerse alguna transformación para convertirlo en estacionario?

22. Sea el modelo

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

Considera la transformación

$$w_t = \Delta Y_t$$

¿Es estacionario el proceso w_t ?

23. Sea el modelo

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t$$

Se pide:

- a) ¿Es el modelo estacionario?
- b) Si se hace la transformación $w_t = \Delta^2 Y_t$, el proceso correspondiente a w_t es estacionario?
- c) Discutir la invertibilidad del proceso resultante.

24. Sea el proceso ARIMA (0, 1, 1)

$$Y_t = Y_{t-1} + \varepsilon_t - 0.9\varepsilon_{t-1}$$

Se pide:

- a) Expresar la ecuación anterior como un modelo AR(∞).
- b) Denominando π_i a los coeficientes del AR(∞), demostrar que $\sum_{j=1}^{\infty} \pi_j = 1$. ¿Es válida la demostración para todos los modelos ARIMA (0, 1, 1)?

25. Sea el modelo

$$Y_t = (\beta_0 + \beta_1 t)\varepsilon_t$$

Se pide:

- a) El modelo propuesto ¿es estacionario en media y en varianza?
- b) Si se hace la transformación $w_t = \Delta Y_t$, el modelo transformado ¿es estacionario en media y en varianza?
- c) En el caso de que en el punto b) no se haya obtenido un modelo estacionario, ¿cómo transformaría el modelo original a fin de que cumpla las condiciones de estacionariedad?

CAPÍTULO 4

Elaboración de modelos ARIMA: fase de identificación

4.1. Metodología de elaboración de modelos ARIMA

Como ya se ha visto en el capítulo 2, una serie temporal se puede contemplar como una realización de un proceso estocástico.

En el capítulo 3 se han examinado las propiedades de los procesos ARIMA (p, d, q) que son una clase particular de procesos estocásticos, pero que sin embargo pueden ser utilizados para describir el comportamiento de la mayor parte de las series económicas. Por esta razón se considerará que una serie temporal dada es una realización de un proceso estocástico específico ARIMA (p, d, q). Se dice que se trata de un proceso específico —o estructura— ARIMA (p, d, q) cuando están determinados numéricamente los parámetros $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, d, \lambda, \mu$ y σ_e^2 . En la figura 4.1. se ha reflejado el mecanismo de generación de una serie temporal.

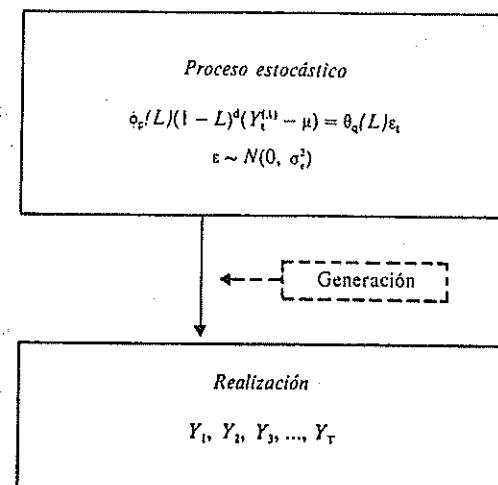


Fig. 4.1. Mecanismo de generación de una serie temporal.

Con este mecanismo, a partir de una muestra concreta de «ruidos blancos» $\epsilon_1, \epsilon_2, \dots, \epsilon_T$ y de valores iniciales de Y_0, Y_1, \dots, Y_{p-1} y de $(\epsilon_0, \epsilon_1, \dots, \epsilon_{q+1})$ se produce una realización del proceso, o dicho de otra forma, se generan los valores Y_1, Y_2, \dots, Y_T .

Naturalmente con una estructura ARIMA (p, d, q) dada se pueden obtener infinitas realizaciones. En cada realización la muestra de «ruidos blancos» será diferente y por consiguiente también será diferente la serie temporal generada. Ahora bien, todas las series generadas que procedan de una misma estructura conservarán entre sí un parentesco, una similitud.

Hasta ahora, se ha adoptado la perspectiva de un proceso conocido previamente; a partir del cual se han generado una o más realizaciones. A partir de aquí se adoptará otra perspectiva: se considerará el supuesto de que se conocen los valores de una serie temporal, y el objetivo será llegar a un conocimiento del mecanismo que verosimilmente haya podido generar la serie. Esta nueva perspectiva queda reflejada en la figura 4.2.

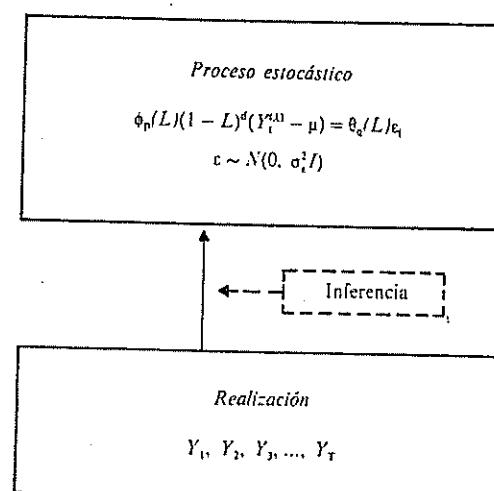


Fig. 4.2. Inferencia sobre un proceso estocástico.

Para efectuar inferencias se puede disponer de una o más realizaciones, es decir, de una o más series generadas de un mismo proceso. En economía, y en general en el ámbito de las ciencias sociales, la inferencia, como ya se indicó en el capítulo 2, hay que efectuarla casi siempre en base a una sola realización.

A la vista de lo que antecede, la elaboración de un modelo ARIMA consiste en la búsqueda de un proceso ARIMA (p, d, q) que verosimilmente haya podido generar la serie temporal objeto de estudio. Las fases que se pueden considerar en la elaboración de un modelo ARIMA se recogen en la figura 4.3. La elaboración de un modelo ARIMA requiere la utilización de un ordenador; en la figura 4.3 se distingue entre tareas que ejecuta el ordenador —cálculos— y las tareas que corresponden al

“0

analista, que básicamente consisten en la toma de decisiones a la vista de los resultados obtenidos. De todas formas, y utilizando la teoría de reconocimiento de patrones o técnicas similares, pueden diseñarse programas que además de la ejecución de los cálculos adopten decisiones. Véase Vallis y Prat (1983).

En la fase de identificación¹, en una primera etapa se procede a efectuar un análisis de estacionariedad de la serie. En el caso de que se trate de una serie no estacionaria se aplican las transformaciones adecuadas al objeto de convertirla en estacionaria. Así, si se dispone de una serie Y_t , se elegirá unos valores de d y λ de forma que la serie:

$$w_t = (1 - L)^d Y_t^{(\lambda)}$$

goce de la propiedad de estacionariedad. Tanto a Y_t como a w_t se le puede restar el valor de una media si se considera necesario.

En la segunda etapa de la fase de identificación se procede a determinar el orden de la parte autorregresiva (es decir, p) y el orden de la parte de medias móviles (es decir, q) del proceso ARMA que se considere haya podido generar a la serie estacionaria w_t .

En los epígrafes 2 y 3 de este capítulo se examinarán estas etapas de la fase de identificación, pero, por razones didácticas, en un orden inverso a la forma en que se procede en la práctica. Así en el epígrafe 4.2 se examinará la identificación de procesos estacionarios, es decir, la determinación de p y q de un proceso ARIMA. En el epígrafe 4.3, que lleva el título de análisis de estacionariedad, se procederá a la determinación de d (número de diferencias a tomar) y λ (transformación de Box-Cox adecuada) para poder obtener una serie estacionaria. Como último punto, en el epígrafe 4.4 se examinará si la media del proceso w_t se puede considerar o no distinta de 0.

En la fase de estimación, se obtienen unos valores estimados para los parámetros $\phi_1, \dots, \phi_q, \theta_1, \dots, \theta_q, \sigma_w^2$ y, en su caso, para μ_w (media del proceso) del proceso ARMA (p, q) . En general, en el proceso de estimación será necesario utilizar procedimientos no lineales, según se verá en el capítulo 5.

Una vez concluida la fase de estimación, se tiene conocimiento de un proceso que, hipotéticamente, ha podido generar la serie temporal transformada w_t , y a partir de la cual se puede obtener la serie original Y_t . En consecuencia, la serie w_t podría expresarse así:

$$\phi(L)(w_t - \mu_w) = \theta(L)\hat{\epsilon}_t$$

A partir de la expresión anterior, conociendo los parámetros de la serie w_t , se puede determinar la serie $\hat{\epsilon}_t$, que si el modelo fuera adecuado debería ser una serie

¹ El término de identificación en el análisis de series temporales tiene un sentido distinto al que se le suele dar en los manuales de econometría. Así, mientras que en el análisis de series temporales identificación se utiliza como un término equivalente a especificación del modelo, en econometría la identificación es un problema que se plantea en el contexto de modelos de ecuaciones simultáneas. En este último sentido se dice que un modelo es identificable cuando se puede encontrar un solo modelo de forma estructural compatible con la función de densidad conjunta de la información muestral. Véase a este respecto Goldberger (1964, págs. 306 y siguientes).

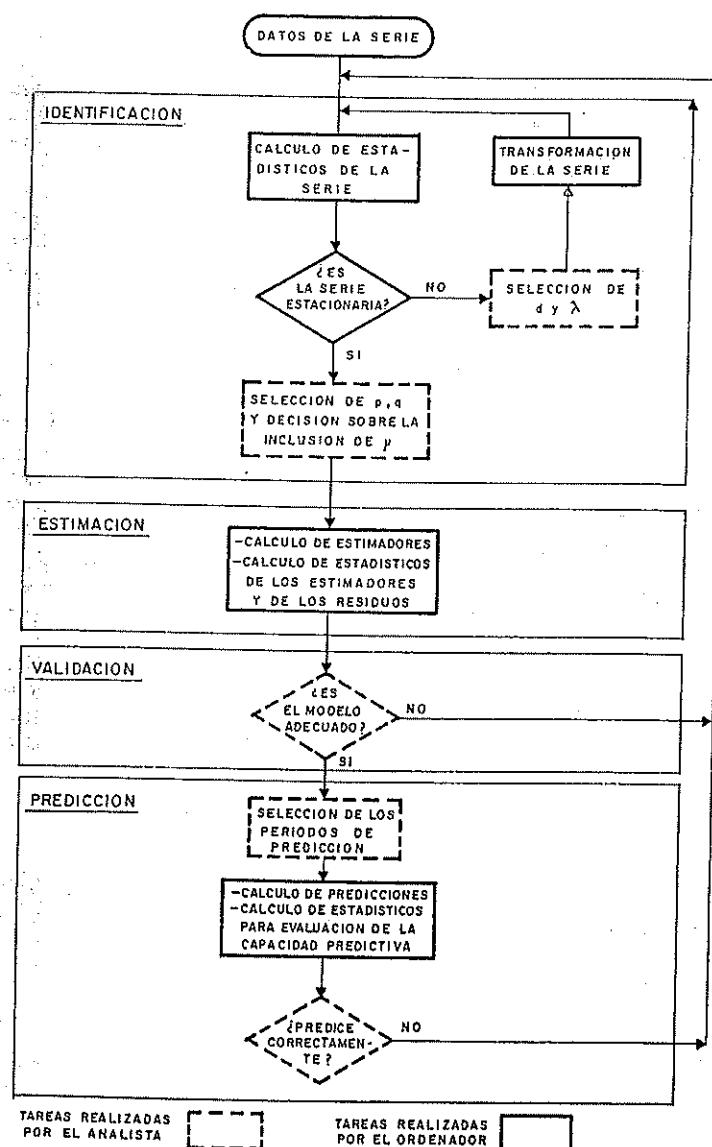


Fig. 4.3. Fases de elaboración de un modelo ARIMA.

que se aproxima al «ruido blanco» e. La fase de validación va dirigida precisamente a establecer si se produce o no esa adecuación entre datos y modelo.

Finalmente, en la fase de predicción, se realizan pronósticos en términos probabilísticos de valores futuros de la variable. La predicción es el verdadero banco de prueba de un modelo: en el caso de existir discrepancias de carácter sistemático entre los valores pronosticados y los que se observen realmente, se deberá cuestionar la validez del modelo identificado. Las fases de validación y predicción se analizarán en los capítulos 6 y 7.

En la práctica los datos que se utilizan para la predicción tienen en muchas ocasiones un carácter estacional. El capítulo 8 está dedicado al tratamiento de la estacionalidad en el contexto de los modelos ARIMA.

En la elaboración de un modelo ARIMA se suelen tener en cuenta dos principios: a) selección del modelo de forma iterativa, b) principio de parametrización escueta.

En la figura 4.3 queda reflejado que la identificación del modelo se hace de forma iterativa, mediante la línea que conduce desde la validación y la predicción hasta la identificación. Con esto quiere indicarse que las fases de validación y de predicción pueden poner de manifiesto que existe una inadecuación entre los datos y el modelo identificado; en este caso se debe proceder a reidentificar el modelo generador.

El principio de la parametrización escueta que proponen los autores Box y Jenkins consiste en postular un modelo generador con el mínimo número de parámetros posibles, y únicamente proceder a su ampliación en caso de que sea estrictamente necesario para explicar el comportamiento de la serie objeto de estudio.

4.2. Identificación de modelos estacionarios

Instrumentos estadísticos de identificación

Como instrumentos básicos de identificación de los modelos estacionarios se utilizarán la función de autocorrelación estimada (FACE) y la función de autocorrelación parcial (FACPE). En los capítulos 2 y 3, en la caracterización de los procesos estacionarios, y de los procesos ARMA en particular, se utilizaron los coeficientes de autocorrelación teóricos; ahora se examinará en qué medida los coeficientes de autocorrelación estimados a partir de una muestra, y cuya secuencia constituye la FACE, son útiles para identificar el proceso generador.

Por otra parte, la función de autocorrelación parcial estimada es un instrumento utilizado especialmente para la identificación de modelos ARMA. A este estadístico le corresponde también un concepto teórico, que será también examinado dentro de este epígrafe.

Según cuál sea el tamaño de la muestra, la identificación del proceso generador puede revestir distinto grado de dificultad. Por ello será analizada también la incidencia del tamaño de la muestra sobre algunos procedimientos de identificación.

Función de autocorrelación estimada (FACE)

A partir de una muestra de tamaño N de valores² de una serie estacionaria w_t se puede calcular un coeficiente de autocorrelación muestral de orden τ , mediante el siguiente estimador:

$$r_\tau = \frac{\sum_{t=\tau+1}^N (w_t - \bar{w})(w_{t-\tau} - \bar{w})}{\sum_{t=1}^N (w_t - \bar{w})^2} \quad (4-1)$$

Puede pensarse en otros estimadores alternativos. Por ejemplo, el que se obtendría a partir de (4-1) dividiendo numerador y denominador por $N - \tau$ y N respectivamente, ya que así existiría una correspondencia con el número de sumandos. Sin embargo, entre las distintas alternativas, el estimador (4-1) se considera que es el que tiene menor sesgo. La secuencia de valores de r_τ para $\tau = 1, 2, 3, \dots$ constituye el correlograma estimado o FACE. Como fácilmente puede verse, a medida que crece el orden del retardo disminuye el número de sumandos del numerador de la fórmula (4-1). Así, para r_{N-1} solamente se dispondrá de un solo sumando en el numerador, y r_N ya no se podría calcular. Naturalmente al disminuir el número de sumandos en el numerador se pierde eficiencia en la estimación; por ello no se recomienda calcular coeficientes de autocorrelación para tamaños superiores a un tercio o a un cuarto del tamaño de la muestra.

Como sucede con cualquier tipo de estimador, interesa conocer la distribución de r_τ al objeto de poder efectuar inferencias de tipo probabilístico. La esperanza de (4-1) bajo el supuesto de que $p_\tau = 0$ sigue asintóticamente una distribución normal con media 0, según demostró Anderson (1942):

$$E(r_\tau) \approx 0 \quad (4-2)$$

Barlett (1946) calculó una aproximación de la varianza y covarianzas de las autocorrelaciones estimadas que vienen dadas por las expresiones:

$$\text{var}(r_\tau) \approx \frac{1}{N} \sum_{k=-\tau}^{\tau} [p_k^2 + p_{k+\tau} p_{k-\tau} - 4p_\tau p_k p_{k-\tau} + 2p_k^2 p_\tau^2] \quad (4-3)$$

$$\text{cov}(r_\tau, r_{\tau+s}) \approx \frac{1}{N} \sum_{k=-\infty}^{\infty} [p_k p_{k+s} + p_{k+\tau-s} p_{k-\tau} - 2p_{\tau+s} p_k p_{k-\tau-s} - 2p_{\tau+s} p_k p_{k-\tau} + 2p_\tau p_{\tau+s} p_k^2] \quad (4-4)$$

Cuando el proceso es un MA(q), los coeficientes de autocorrelación teóricos son

² En cada diferencia que se toma se pierde una observación al comienzo de la serie de acuerdo con la definición que se ha dado del operador diferencia. Por ello, si la muestra de Y_t es de tamaño T , para w_t quedará reducida a $T - d = N$. Para simplificar la notación, a los N valores de la serie estacionaria w_t se les ha designado por w_1, \dots, w_N .

iguales a 0 para $\tau > q$, con lo cual el sumatorio de la fórmula (4-3) queda reducido a límites finitos. Además, tanto en la fórmula (4-3) como en la (4-4) se puede prescindir de todos los sumandos excepto del primero, ya que su valor será en general muy pequeño. En consecuencia, para un proceso MA(q) —en que $p_\tau = 0$ para $\tau > q$ — como una segunda aproximación se puede utilizar las expresiones siguientes:

$$\text{var}(r_\tau) \approx \frac{1}{N} \sum_{k=-q}^q p_k^2 \quad (4-5)$$

$$\text{cov}(r_\tau, r_{\tau+s}) \approx \frac{1}{N} \sum_{k=-q}^q p_k p_{k+s} \quad (4-6)$$

Como una aproximación empírica para calcular la varianza se sustituyen los p_k , que son desconocidos, por los correspondientes coeficientes de autocorrelación estimados. Además, como se desconoce de antemano el tipo y el orden del proceso, se adopta el supuesto de que $p_k = 0$ para $k \geq \tau$. Es decir, en el cálculo de la varianza para cada retardo τ se está adoptando el supuesto de que se trata de un modelo MA de orden $\tau - 1$. En consecuencia, la fórmula empírica utilizada es la siguiente:

$$\text{var}(r_\tau) \approx \frac{1}{N} (1 + 2 \sum_{k=1}^{\tau-1} r_k^2) \quad (4-7)$$

Como r_τ tiene una distribución aproximadamente normal, en base a (4-2) y (4-7) se pueden construir intervalos del 95 % que vendrán dados por:

$$(-1,96 \sqrt{\text{var}(r_\tau)}, +1,96 \sqrt{\text{var}(r_\tau)}) \quad (4-8)$$

Conviene tener también en cuenta que los coeficientes de autocorrelación están correlacionados entre sí como lo muestran las expresiones (4-4) ó (4-6). Como consecuencia de este tipo de relaciones pueden aparecer en la FACE o correlograma estimado ciertas ondulaciones que pueden no corresponderse con el comportamiento teórico del modelo. En este sentido, y especialmente cuando los valores de los coeficientes de autocorrelación son elevados, se puede esperar que la FACE converja a cero más lentamente que la FACT, según han analizado Jenkins y Watts (1968).

La función de autocorrelación parcial teórica y estimada

A diferencia de lo que ocurre con un proceso de medias móviles, en un AR, cualquiera que sea su orden, la función de autocorrelación teórica empieza a decrecer a partir de un determinado retardo, pero nunca se hace 0. Por esta circunstancia, es difícil discriminar a la vista del correlograma el orden de un proceso autorregresivo salvo que se trate de un AR(1). Precisamente, los coeficientes de autocorrelación parcial constituyen un instrumento diseñado para que sirva de ayuda de forma especial en la identificación de modelos autorregresivos. A continuación se examinará el significado de estos coeficientes.

Por medio de las ecuaciones de Yule-Walker se pueden calcular los coeficientes de un proceso AR(p) en función de los coeficientes de autocorrelación, ya que

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} \quad (4-9)$$

Ahora bien, cuando se trabaja con series reales, el problema radica en que se desconoce p .

Si el sistema (4-9) se resuelve sucesivamente para un AR(1), AR(2), AR(3), se obtendrán los siguientes coeficientes:

$$\begin{array}{ll} \text{AR(1)} & \phi_{11} \\ \text{AR(2)} & \phi_{11} \phi_{22} \\ \text{AR(3)} & \phi_{11} \phi_{22} \phi_{33} \\ \cdots & \cdots \\ \text{AR}(p) & \phi_{11} \phi_{22} \phi_{33} \dots \phi_{pp} \\ \text{AR}(p+1) & \phi_{p+1,1} \phi_{p+1,2} \phi_{p+1,3} \dots \phi_{p+1,p}, \phi_{p+1,p+1} \end{array}$$

Pues bien, si se toma el último coeficiente de cada uno de los procesos se obtiene la función de autocorrelación parcial teórica (FACPT). Así pues

$$\phi_{11}, \phi_{22}, \phi_{33}, \dots, \phi_{pp}, \phi_{p+1,p+1}, \dots$$

constituyen la FACPT.

¿Cuál es el comportamiento de estos coeficientes? Bajo el supuesto de que p es el orden del proceso autorregresivo, serán distintos de 0 para retardos iguales o inferiores a p , ya que por hipótesis los coeficientes ϕ_1, ϕ_2, \dots y ϕ_p son distintos de 0. En cambio, para retardos superiores a p serán iguales a 0.

Si se considera un proceso MA(q), aplicando sustituciones sucesivas se llegaría a un proceso AR(∞). Los coeficientes del proceso transformado decaen a partir del retardo q según una ecuación en diferencias. Por tanto, en un proceso MA(q) los coeficientes de autocorrelación parcial no se igualarán nunca a 0, aunque a partir de un retardo q decaerán de forma rápida. Así pues, la FACPT de un proceso MA se comporta de manera análoga a como lo hace la FACT en un AR y recíprocamente.

Lógicamente, en un proceso ARMA(p, q) tanto la FACPT como la FACT tienen infinitos elementos distintos de 0, ya que darían lugar, bien a un MA(∞) o bien a un AR(∞).

En las figuras 4.4, 4.5 y 4.6 se han representado los coeficientes de autocorrelación parcial teóricos correspondientes a procesos AR, MA y ARMA respectivamente.

La función de autocorrelación parcial estimada (FACPE), que es el instrumento empírico utilizado en la identificación, puede obtenerse de dos formas alternativas pero prácticamente equivalentes.

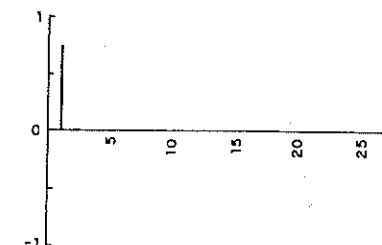
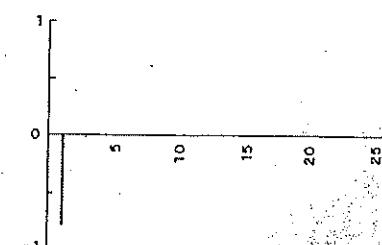
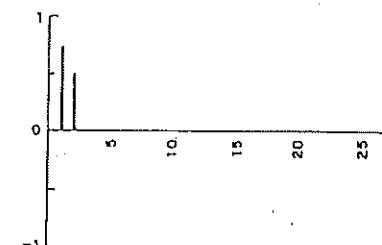
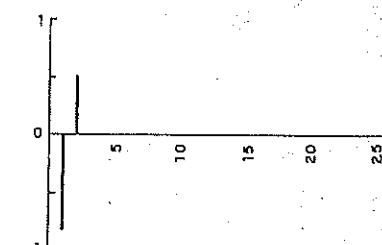
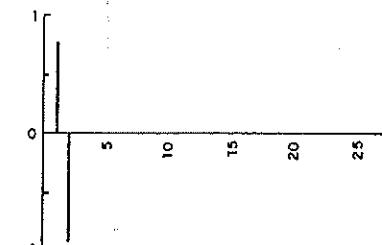
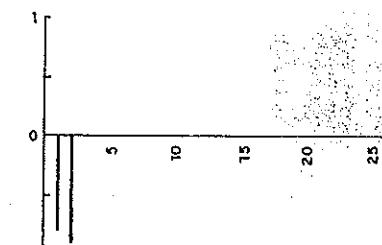
a) $Y_t - 0.8 Y_{t-1} = \epsilon_t$ b) $Y_t + 0.8 Y_{t-1} = \epsilon_t$ c) $Y_t - 0.4 Y_{t-1} - 0.5 Y_{t-2} = \epsilon_t$ d) $Y_t + 0.4 Y_{t-1} - 0.5 Y_{t-2} = \epsilon_t$ e) $Y_t - 1.5 Y_{t-1} + 0.9 Y_{t-2} = \epsilon_t$ f) $Y_t + 1.5 Y_{t-1} + 0.9 Y_{t-2} = \epsilon_t$

Fig. 4.4. Funciones de autocorrelación parcial de procesos AR(1) y AR(2).

ELABORACIÓN DE MODELOS ARIMA: FASE DE IDENTIFICACIÓN

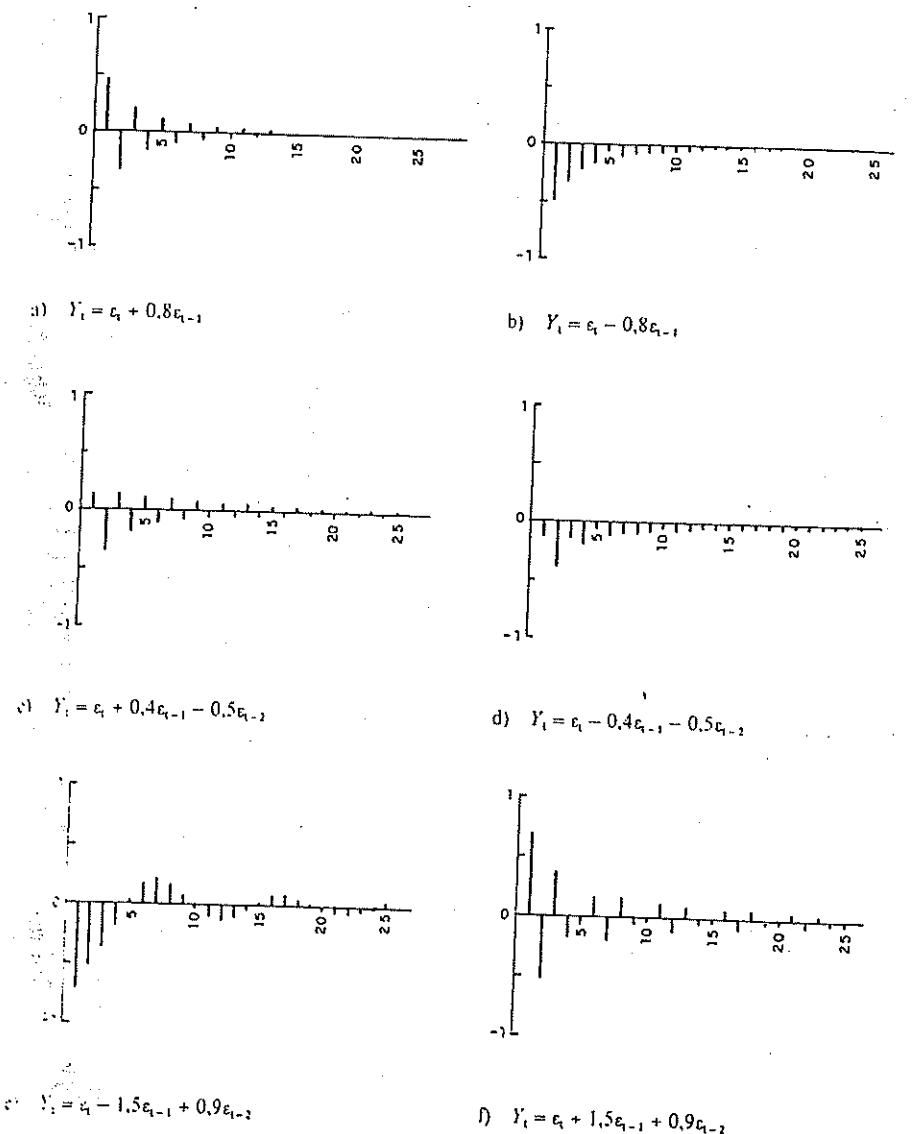


Fig. 4.5. Funciones de autocorrelación parcial de procesos MA(1) y MA(2).

IDENTIFICACIÓN DE MODELOS ESTACIONARIOS

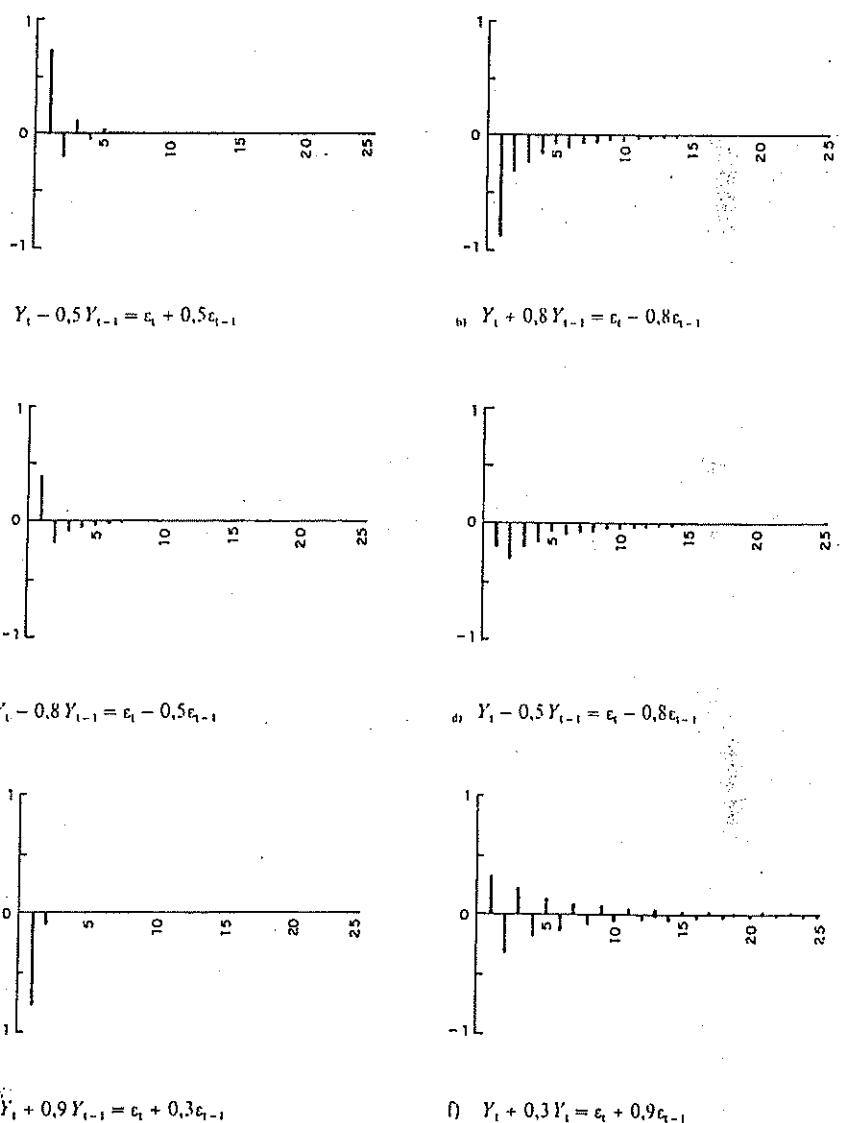


Fig. 4.6. Funciones de autocorrelación parcial de procesos ARMA (1, 1).

En la primera alternativa, se utiliza el sistema de Yule-Walker a partir de los coeficientes de autocorrelación estimados. La secuencia sería la siguiente:

$$\begin{aligned}\hat{\phi}_{11} &= r_1 \\ \begin{bmatrix} \hat{\phi}_{21} \\ \hat{\phi}_{12} \end{bmatrix} &= \begin{bmatrix} 1 & r_1 \\ r_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \\ \begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{22} \\ \hat{\phi}_{32} \end{bmatrix} &= \begin{bmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}\end{aligned}\quad (4-10)$$

.....
.....

y así sucesivamente.

De los cálculos anteriores únicamente se retienen $\hat{\phi}_{11}, \hat{\phi}_{22}, \hat{\phi}_{33}, \dots$ que son los coeficientes que configuran la FACPE. Durbin (1960) ideó un procedimiento que permite obtener estos coeficientes de forma recursiva, sin necesidad de invertir matrices que obviamente aumentan de tamaño a medida que crece el número de retardos.

El procedimiento de Durbin viene dado por las siguientes expresiones:

$$\begin{aligned}\hat{\phi}_{kk} &= \frac{r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j} \\ \hat{\phi}_{kj} &= \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \hat{\phi}_{k-1,k-j} \quad j = 1, 2, \dots, k-1\end{aligned}\quad (4-11)$$

Como se trata de un cálculo recursivo, este procedimiento tiene el inconveniente de que se van acumulando los errores de redondeo. Por esta razón, en los cálculos por ordenador se suelen utilizar algoritmos de doble precisión.

La segunda alternativa consiste en obtener los coeficientes mediante regresión. Así, para sucesivos retardos las ecuaciones ajustadas serán

$$\begin{aligned}\hat{Y}_t &= \hat{\phi}_{11} Y_{t-1} \\ \hat{Y}_t &= \hat{\phi}_{21} Y_{t-1} + \hat{\phi}_{22} Y_{t-2} \\ &\vdots \\ \hat{Y}_t &= \hat{\phi}_{k1} Y_{t-1} + \hat{\phi}_{k2} Y_{t-2} + \dots + \hat{\phi}_{kk} Y_{t-k}\end{aligned}\quad (4-12)$$

Desde esta perspectiva, un coeficiente de autocorrelación parcial estimado (por ejemplo, $\hat{\phi}_{kk}$) puede contemplarse como la correlación parcial existente entre Y_{t-k} e Y_t cuando se han eliminado de ambas variables el efecto de $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$. Si se trata de un proceso AR(p), el coeficiente $\hat{\phi}_{kk}$ estará próximo a 0 cuando $k > p$ ya que el

correspondiente coeficiente teórico $\phi_{kk} = 0$. Para $k \leq p$, $\hat{\phi}_{kk}$ tendrá un valor significativamente distinto de 0 si el tamaño de la muestra es suficientemente grande.

En un proceso AR(p), los coeficientes de autocorrelación parcial aproximadamente tienen distribución normal con media 0 y varianza:

$$\text{var}(\hat{\phi}_{kk}) \approx \frac{1}{N} \quad k > p \quad (4-13)$$

según el resultado obtenido por Quenouille (1949). En la práctica, se utiliza empíricamente esta distribución para calcular intervalos de confianza para todos los coeficientes de autocorrelación parcial obtenidos, con independencia de cuál sea el tipo de proceso, que naturalmente se desconoce de antemano.

El tamaño de la muestra

La identificación del modelo generador será tanto más fácil cuanto mayor sea el tamaño de la muestra. Para analizar la incidencia de este factor, a partir de la estructura

$$Y_t = 0.7 Y_{t-1} + \varepsilon_t \quad (Y_0 = 0)$$

se han generado muestras de distinto tamaño, desde $T = 30$ hasta $T = 1000$. En las figuras 4.7 a 4.12 se ha representado los valores de la serie, los coeficientes de autocorrelación y autocorrelación parcial estimados, junto con una banda de confianza del 95 %. También, y con objeto de que sirva de referencia en cada figura, aparecen los coeficientes de autocorrelación y autocorrelación parcial teóricos derivados directamente de la estructura adoptada.

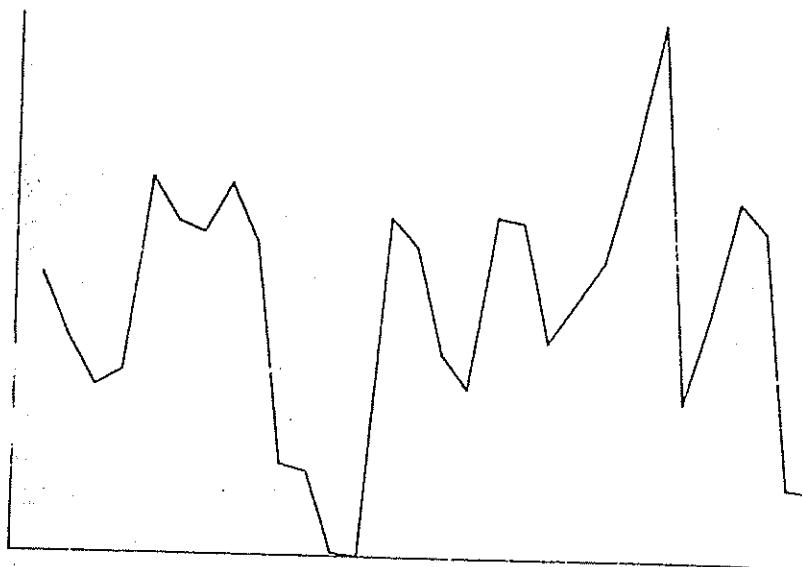
Cuando el tamaño de la muestra es 30 (figura 4.7), el comportamiento de la FACE y de la FACPE no guarda ninguna similitud con el comportamiento de las correspondientes funciones teóricas. Además, prácticamente ninguno de los coeficientes es significativamente distinto de 0 para un nivel de confianza del 95 %, que es el nivel a que están construidas las bandas. Por esta razón tanto la FACE como la FACPE pueden provenir hipotéticamente de un proceso ruido blanco. El coeficiente que supera en mayor medida las bandas de confianza es el $\hat{\phi}_{11,11}$, lo que confirma la inutilidad de los instrumentos de identificación cuando el tamaño de la muestra es tan pequeño.

Para un tamaño igual a 50 solamente son significativos estadísticamente los coeficientes $\hat{\rho}_{11}$, $\hat{\phi}_{11}$ y $\hat{\phi}_{16,16}$. Se esboza sin embargo el parentesco de las funciones estimadas con las funciones teóricas. Así, en la FACE se advierte un cierto decrecimiento exponencial mientras que en la FACPE, se produce un corte brusco después del primer coeficiente.

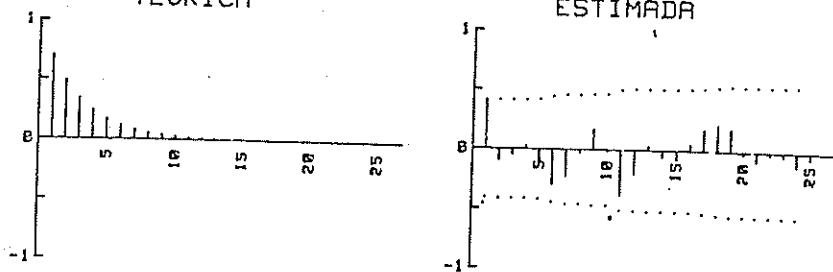
A partir de muestras de tamaño 100, aparece claro el decrecimiento exponencial en la FACE y el corte brusco en la FACPE después del primer coeficiente. Sin embargo, el decrecimiento exponencial de la FACE va acompañado de una especie de movimiento ondulatorio que seguramente es una consecuencia de la correlación exis-

Número de observaciones: 30

A) SERIE SIMULADA



B) FUNCION DE AUTOCORRELACION
TEORICA ESTIMADA



C) FUNCION DE AUTOCORRELACION PARCIAL
TEORICA ESTIMADA

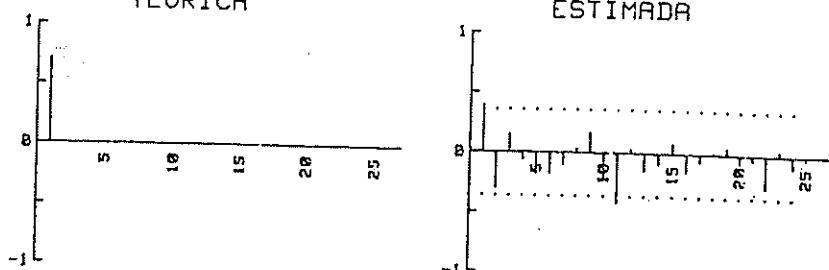
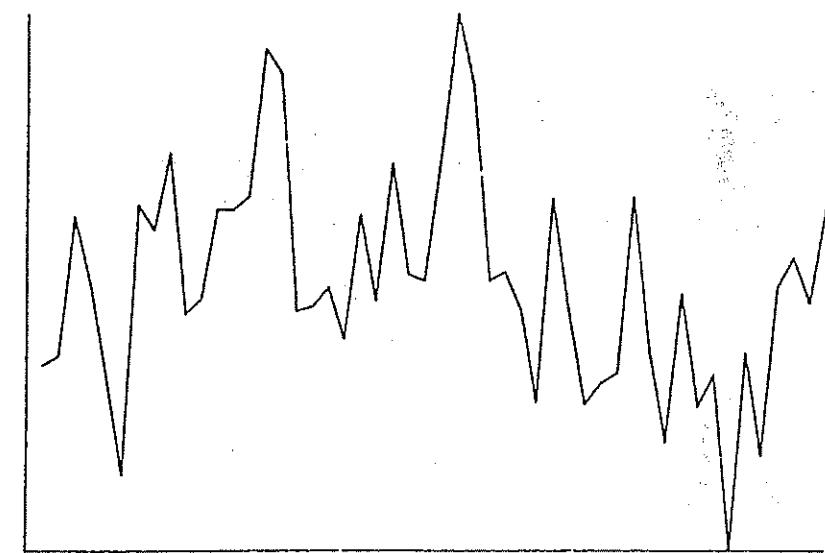


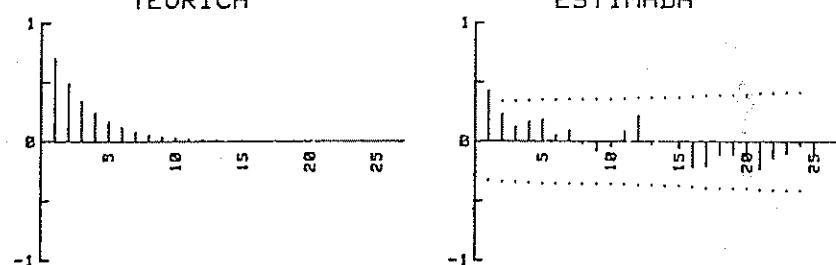
Fig. 4.7

Número de observaciones: 50

A) SERIE SIMULADA



B) FUNCION DE AUTOCORRELACION
TEORICA ESTIMADA



C) FUNCION DE AUTOCORRELACION PARCIAL
TEORICA ESTIMADA

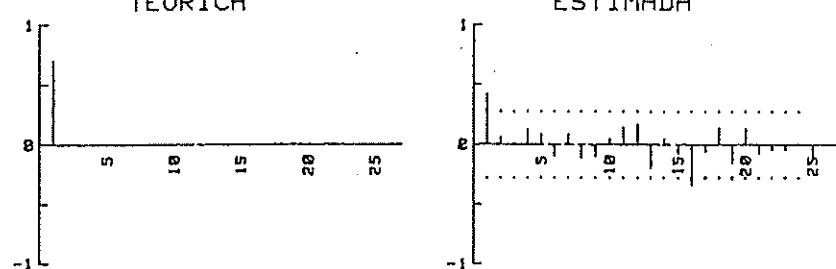
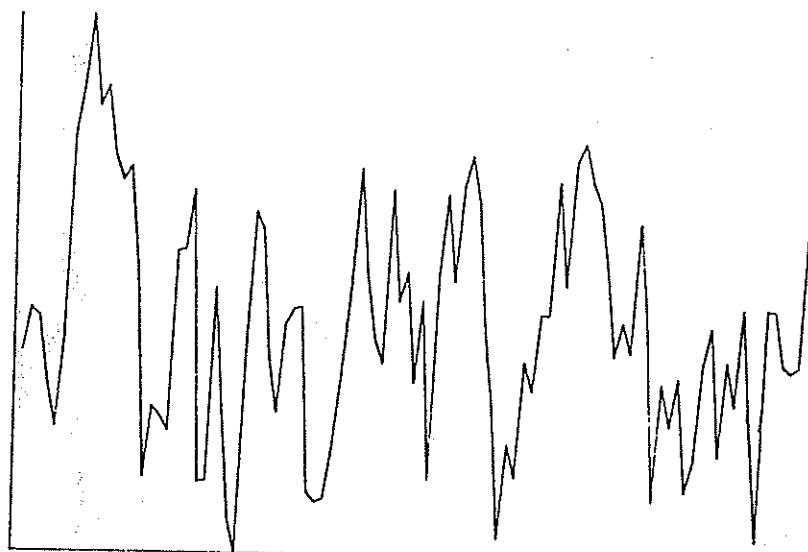


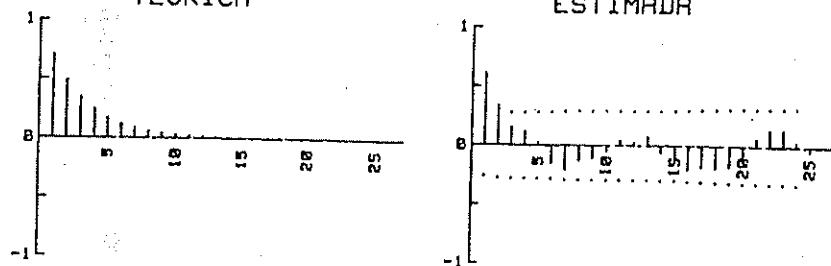
Fig. 4.8

Número de observaciones: 100

A) SERIE SIMULADA



B) FUNCION DE AUTOCORRELACION
TEORICA ESTIMADA



C) FUNCION DE AUTOCORRELACION PARCIAL
TEORICA ESTIMADA

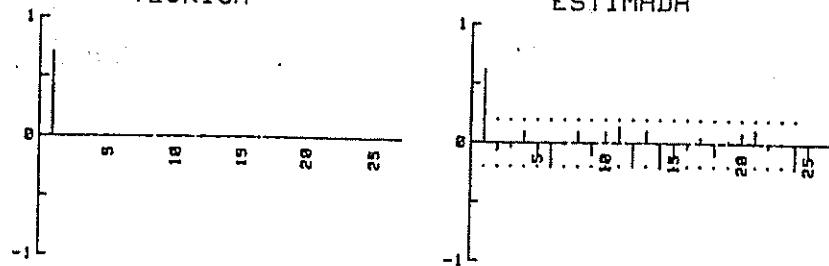
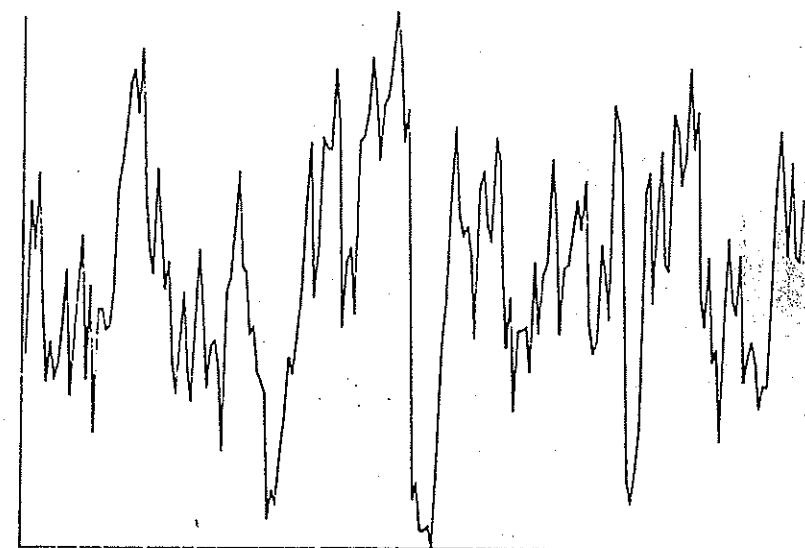


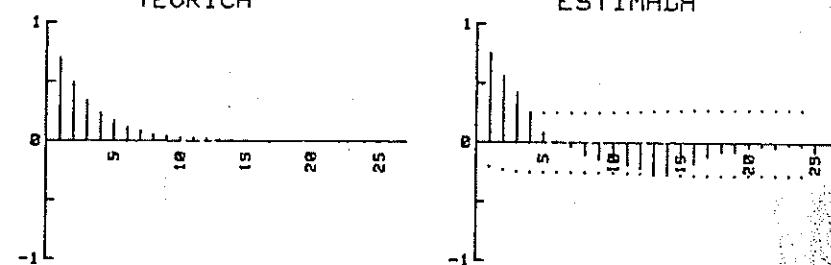
Fig. 4.9

Número de observaciones: 200

A) SERIE SIMULADA



B) FUNCION DE AUTOCORRELACION
TEORICA ESTIMADA



C) FUNCION DE AUTOCORRELACION PARCIAL
TEORICA ESTIMADA

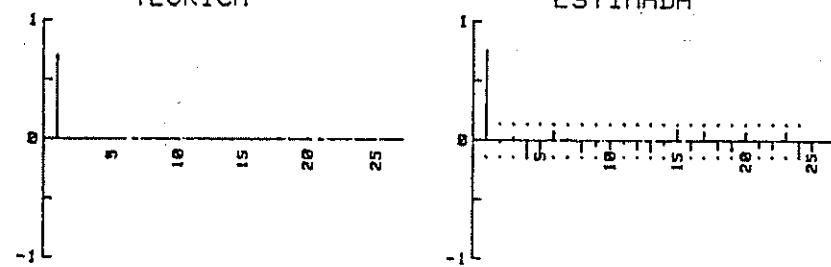
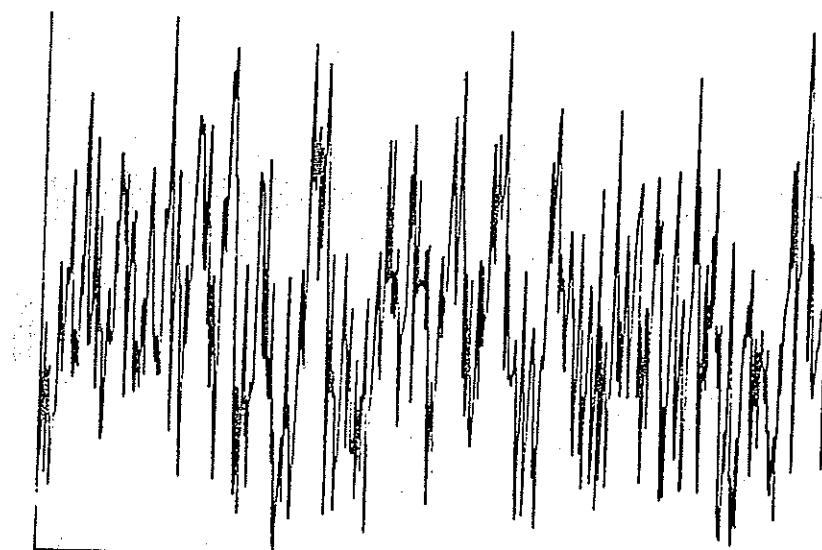


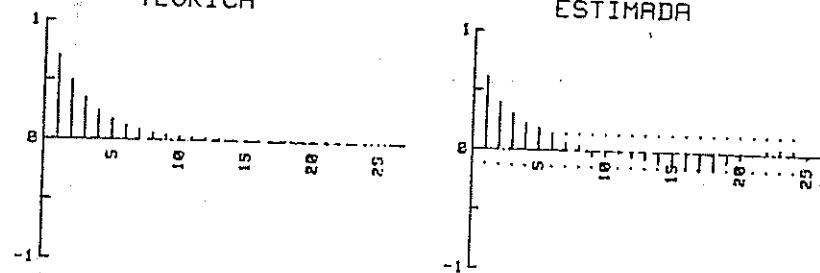
Fig. 4.10

Número de observaciones: 500

A) SERIE SIMULADA



B) FUNCION DE AUTOCORRELACION
TEORICA
ESTIMADA



C) FUNCION DE AUTOCORRELACION PARCIAL
TEORICA
ESTIMADA

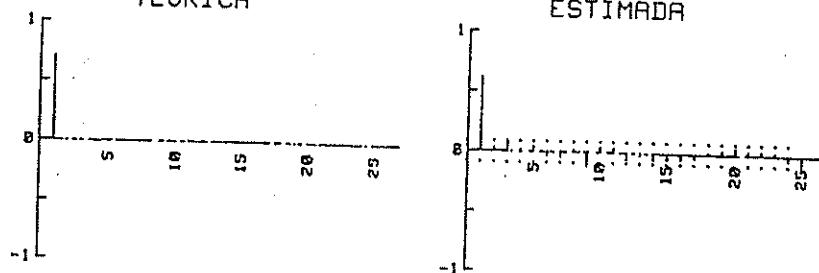
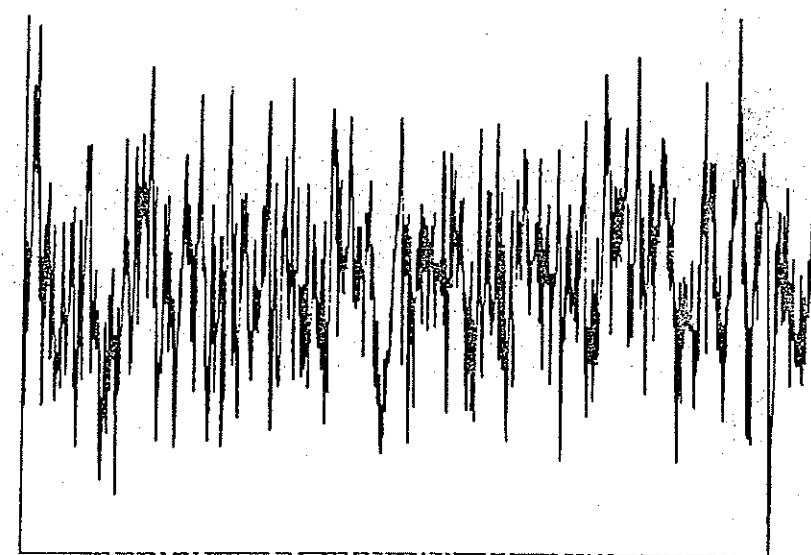


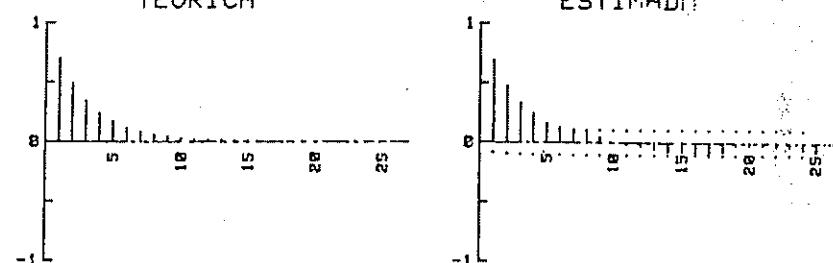
Fig. 4.11

Número de observaciones: 1000

A) SERIE SIMULADA



B) FUNCION DE AUTOCORRELACION
TEORICA
ESTIMADA



C) FUNCION DE AUTOCORRELACION PARCIAL
TEORICA
ESTIMADA

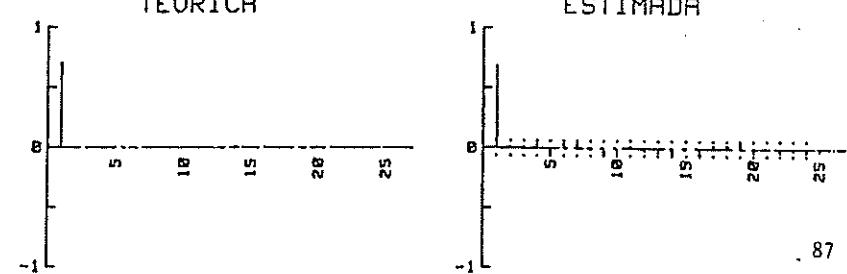


Fig. 4.12

tente entre los coeficientes estimados. Este movimiento ondulatorio no llega a desaparecer para una muestra de tamaño 1000.

Al evaluar los resultados obtenidos conviene advertir que el comportamiento de las series será distinto en cada experimento que se efectúe. Rosa E. Fernández (1984) ha realizado un estudio sistemático mediante simulación con métodos de Montecarlo para distintos tamaños de muestra. No obstante, como una conclusión general se puede establecer que en las series cuyo número de observaciones sea pequeño, la identificación del proceso generador a partir de los instrumentos propuestos es muy difícil. Algunos autores consideran pequeño un tamaño de muestra inferior a 50 ó 60 observaciones, aunque siempre hay que tener en cuenta las características particulares de la serie que se está analizando.

4.3. Identificación de un modelo ARMA (p, q)

En la identificación del modelo generador ARMA(p, q) los instrumentos básicos son, como ya se ha dicho, la FACE y la FACPE. El proceso de identificación consiste en comparar el comportamiento de estos estadísticos con funciones de autocorrelación y autocorrelación parcial teóricas correspondientes a distintos modelos teóricos con los que puedan guardar similitud, teniendo en cuenta que nunca cabe esperar una similitud perfecta debido a errores de muestreo. A continuación, y con objeto de que sirva de referencia, se va a recapitular el comportamiento de la FACT y de la FACPT en modelos AR, MA y ARMA.

Modelos AR(p)

En un modelo AR(p), los coeficientes de autocorrelación presentan un decrecimiento rápido que puede ser geométrico puro, geométrico con alternancia de signos, sinusoidal o una mezcla de las anteriores. El tipo de decrecimiento concreto dependerá del signo y de las características de las raíces de la ecuación en diferencias del proceso generador. Por otra parte, los coeficientes de autocorrelación parcial son igual a 0, para retardos superiores a p .

En la figura 4.13 se ha representado la FACT y la FACPT para distintos procesos AR(1) y AR(2). Los puntos situados en el interior del triángulo representan distintas combinaciones de modelos AR(2) estacionarios. (Un modelo AR(1) es un caso particular de un AR(2) en que $\phi_2 = 0$.) Las rectas que delimitan el triángulo son

$$\phi_2 = 1 - \phi_1$$

$$\phi_2 = 1 + \phi_1$$

$$\phi_2 = -1$$

Por otra parte, la región que aparece debajo del arco de parábola corresponde a modelos cuya ecuación en diferencias tiene raíces imaginarias.

El comportamiento de la FACT y de la FACPT de los procesos AR(2) ha sido estudiado en detalle por Stralkowski, Wu y Devor (1970).

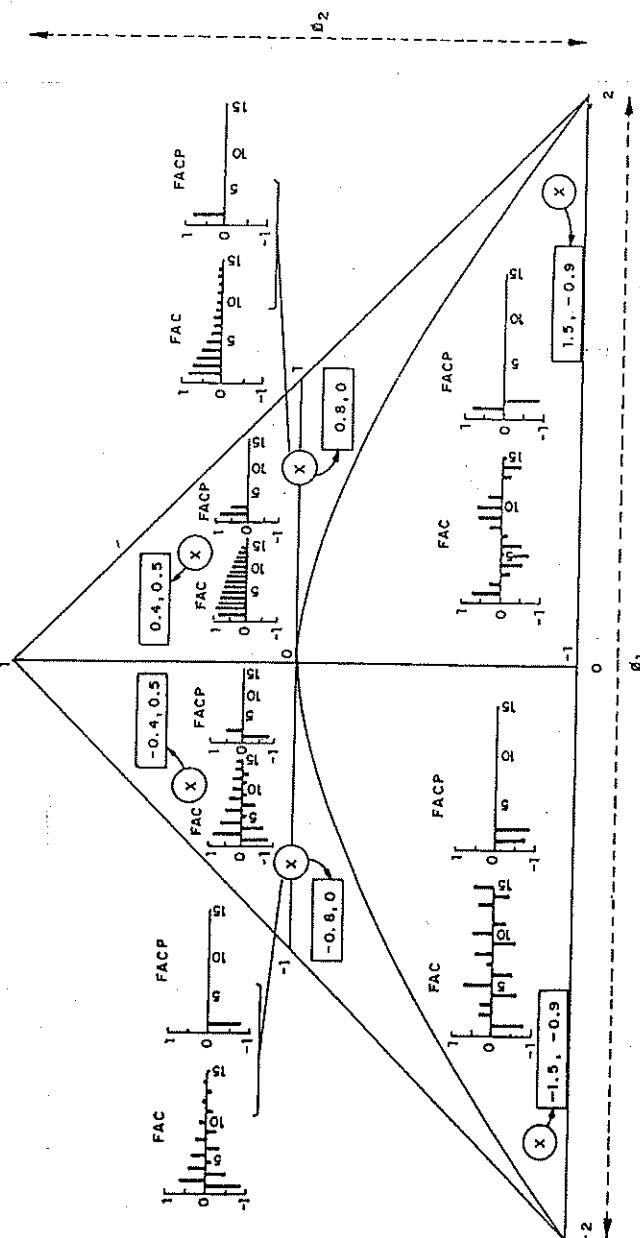


Fig. 4.13

Modelos MA(q)

En los modelos MA el comportamiento es el inverso del presentado en los modelos AR. Así, en un modelo MA(q) la FACT se hace 0 para retardos superiores a q ; por el contrario, la FACPT decrece de forma rápida pero sin llegar a anularse nunca.

En la figura 4.14 se refleja la FACT y la FACPT de 6 estructuras de modelos MA(1) y MA(2). Todas las estructuras están dentro de la zona de invertibilidad, definida por los puntos interiores del triángulo formado por las rectas

$$\theta_2 = 1 - \theta_1$$

$$\theta_2 = 1 + \theta_1$$

$$\theta_2 = -1$$

También aquí, la región que aparece debajo del arco de parábola corresponde a modelos con raíces imaginarias.

El comportamiento teórico de la FACT y de la FACPT de un MA(2) puede verse con mayor detalle en Stralkowski, Wu y Devor (1974).

Modelos ARMA(p, q)

Los modelos ARMA(p, q) son los que presentan mayores dificultades en lo que se refiere a su identificación. En un modelo ARMA(p, q), tanto la FACT como la FACPT contienen teóricamente un número infinito de coeficientes distintos de 0.

En la FACT los coeficientes p_1, \dots, p_q no tienen un patrón fijo de comportamiento, ya que dependen de los valores de los coeficientes de medias móviles del modelo. Para $\tau > q$, los coeficientes p_τ se determinan según una ecuación en diferencias y decrecen según una mezcla de oscilaciones sinusoidales y/o exponenciales amortiguadas. La FACPT también presenta p valores iniciales sin un patrón fijo, seguidos de una mezcla de oscilaciones sinusoidales y/o exponenciales amortiguadas.

En la figura 4.15 se ha representado la FACT y la FACPT de 6 modelos ARMA(1, 1) estacionarios e invertibles. La zona conjunta de invertibilidad y estacionariedad está definida por todas las combinaciones situadas en el interior del cuadrado de la figura. Para un mayor detalle, véase Stralkowski, Wu y Devor (1974).

Conviene tener presente que la diagonal trazada en la figura 4.15 corresponde a aquellas combinaciones en que $\phi_1 = \theta_1$. Cuando se da esta circunstancia el modelo ARMA queda reducido a un modelo de ruido blanco.

En efecto,

$$(1 - \phi_1 L) Y_t = (1 - \theta_1 L) \epsilon_t \quad (4-14)$$

se convierte en

$$Y_t = \epsilon_t \quad \text{para} \quad \phi_1 = \theta_1 \quad (4-15)$$

También, las combinaciones cercanas a esa diagonal presentan dificultades de identificación, puesto que no serán fácilmente distinguibles de un ruido blanco.

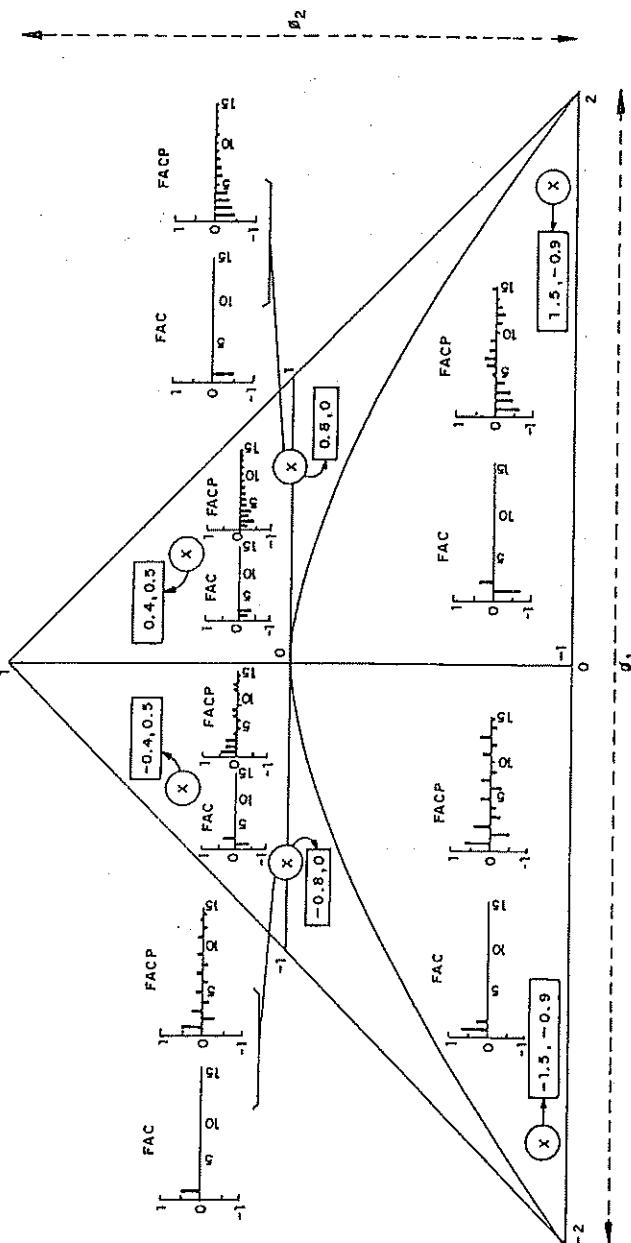


Fig. 4.14

En el cuadro 4.1. se hace un resumen del comportamiento de la FACT y de la FACPT en modelos AR(p), MA(q) y ARMA(p, q).

CUADRO 4.1. La FACT y la FACPT en modelos AR(p), MA(q) y ARMA(p, q).

	FACT	FACPT
AR(p)	Decrecimiento rápido de tipo geométrico puro, y geométrico con alteración de signos, sinusoidal o mezcla de varios tipos.	Se anula para retardos superiores a p .
MA(q)	Se anula para retardos superiores a q .	Decrecimiento rápido de tipo exponencial y/o sinusoidal.
ARMA(p, q)	Los primeros valores iniciales no tienen patrón fijo y van seguidos de una mezcla de oscilaciones sinusoidales y/o exponenciales amortiguadas.	Los primeros valores iniciales no tienen patrón fijo y van seguidos de una mezcla de oscilaciones sinusoidales y/o exponenciales amortiguadas.

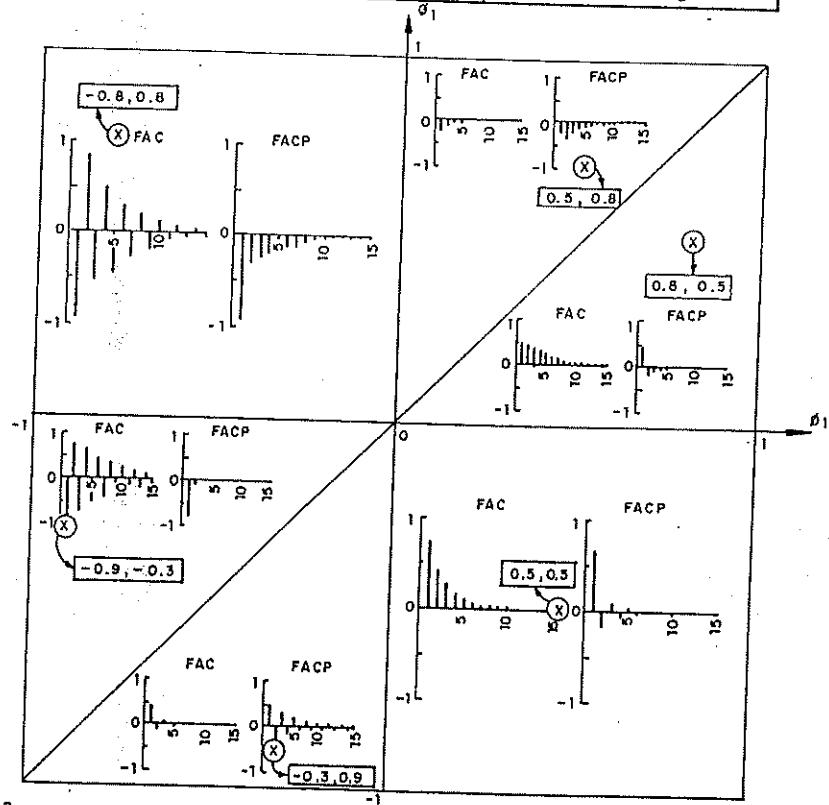


Fig. 4.15

En la práctica los procesos mixtos ARMA son difíciles de identificar mediante estos instrumentos, ya que teóricamente no se produce un corte ni en la FACT ni en la FACPT.

Con objeto de facilitar la identificación de los modelos ARIMA, en los últimos años se han desarrollado nuevos instrumentos, entre los que se pueden destacar los siguientes: la función de autocorrelación inversa en el dominio de las frecuencias de Cleveland (1972), la función de autocorrelación inversa en el dominio temporal desarrollada por Hipel, McLeod y Lennox (1977), el criterio S-array y R-array de Gray, Kelly y McIntire (1979), el método de la esquina de Beguin, Gourieroux y Monfort (1980) y la función de autocorrelación extendida de Tsay y Tiao (1981). Sobre este último método, que teóricamente al menos parece reunir ventajas sobre los anteriores, puede verse una exposición, así como un análisis mediante simulación en Gea y Uriel (1984).

4.4. Análisis de estacionariedad

Como ya se ha indicado en múltiples ocasiones, las series económicas que se manejan son en general de carácter no estacionario. Con objeto de poder aplicar la metodología de los modelos ARIMA es preciso realizar las transformaciones adecuadas. En concreto, en consonancia con el análisis efectuado en 3.4 se contemplarán dos tipos de transformaciones: a) la toma de diferencias de orden d y b) la transformación instantánea del tipo Box-Cox, mediante el parámetro λ . A continuación se exponen los instrumentos utilizados y su aplicación en la identificación de los parámetros d y λ .

Tratamiento de la no estacionariedad en media: Identificación de d

Tomando diferencias de orden d se conseguirá en general que la serie sea estacionaria en media. El primer problema que se plantea es ¿qué instrumentos estadísticos se pueden utilizar? Como dos instrumentos iniciales —más adelante se añadirá un tercero— se utilizarán el gráfico de la serie y la función de autocorrelación estimada (FACE).

El examen visual de la trayectoria de la serie a lo largo del tiempo puede dar una idea de si es o no estacionaria en media. Si existe algún valor en torno al cual la serie va oscilando pero sin alejarse de forma permanente de dicho valor, entonces se puede considerar que la serie es estacionaria en media. Naturalmente si no se cumple este requisito, la serie debería someterse a diferenciaciones con objeto de convertirla en estacionaria.

En cualquier caso, es conveniente realizar esta inspección ocular conjuntamente con un examen de la FACE. Desde un punto de vista teórico, la FACT no está definida para un proceso no estacionario que arranque en un pasado remoto. No obstante para un proceso «random walk»³ que arranque en un período finito, los coeficientes

³ Un proceso «random walk» es un proceso en que la parte autorregresiva $[(1 - L)Y_t]$ tiene una sola raíz que es unitaria, mientras que la parte de medias móviles se reduce a un simple ruido blanco según se vio en 3.4.

ρ_{it} se mantienen muy próximos a 1 a medida que aumenta t de acuerdo con la expresión (3-112). En general, puede considerarse que éste será el comportamiento de los procesos con raíces unitarias. Tomando como pauta este comportamiento, si los coeficientes de la FACE no decaen rápidamente sería un indicio claro de que la serie es no estacionaria. Si así ocurriera, se tomaría una diferencia, y se volvería a analizar la FACE de la serie diferenciada. De esta forma se continuaría hasta obtener una serie diferenciada de orden d en la que los coeficientes de la FACE decaigan rápidamente a partir de un determinado retardo. En la práctica es suficiente con tomar $d = 1$ o $d = 2$ para obtener una serie estacionaria en media.

Un problema que se puede plantear es el siguiente: ¿qué sucedería si se toman más diferencias de las necesarias? Para analizar esta cuestión, se va a considerar de nuevo un modelo «random walk» cuya expresión es la siguiente:

$$Y_t = Y_{t-1} + \varepsilon_t \quad (4-16)$$

Si se toma $d = 1$, es decir, si se toma $w_t = \Delta Y_t$, se obtiene el proceso

$$w_t = \Delta Y_t = \varepsilon_t \quad (4-17)$$

que, como es evidente, ya se vio, es estacionario e invertible.

Si se toma $d = 2$, entonces $w_t = \Delta^2 Y_t$ lo que daría lugar a:

$$w_t = \Delta^2 Y_t = \Delta(\Delta Y_t) = \Delta(\varepsilon_t) = \varepsilon_t - \varepsilon_{t-1} \quad (4-18)$$

Este segundo proceso sería también estacionario, pero no invertible.

Por otra parte, la varianza en el proceso (4-17) es

$$\text{var}(\Delta Y_t) = \text{var}(\varepsilon_t) = \sigma_\varepsilon^2 \quad (4-19)$$

mientras que la varianza correspondiente a (4-18) es la siguiente

$$\text{var}(\Delta^2 Y_t) = \text{var}(\varepsilon_t - \varepsilon_{t-1}) = 2\sigma_\varepsilon^2 \quad (4-20)$$

Las consecuencias de aplicar en este modelo un número de diferencias mayor que el necesario han sido que el proceso final sea no invertible y con una varianza mayor que si se hubiera aplicado un número de diferencias correcto (en este caso $d = 1$). Otra consecuencia es la de sobreparametrización ya que el proceso (4-16) se ha representado mediante un ARIMA (0, 2, 1), cuando en realidad se podría representar de una forma más escueta mediante un ARIMA (0, 1, 0).

En cualquier caso, y ésta sería una consecuencia de carácter general, si se toman más diferencias de las requeridas, el proceso final será siempre estacionario. Si al estimar un proceso, en la parte de medias móviles se obtienen alguna o algunas raíces cercanas a 1 sería indicativo de que ha existido sobrediferenciación de la serie.

En el modelo considerado, cuya varianza para $d = 0$ es teóricamente infinita, al pasar de $d = 1$ (diferencia correcta) a $d = 2$, la varianza se ha duplicado. Este hecho se ha tomado como base para sugerir un nuevo instrumento de identificación de d .

Este instrumento consiste en calcular la varianza de la serie original y de la serie sometida a diferentes diferenciaciones, tomándose como diferenciación adecuada aquella para la cual la varianza es mínima.

Este puede ser un instrumento a utilizar, conjuntamente con los otros señalados anteriormente, pero conviene advertir que lo que sucede con el modelo (4-16) no tiene validez general. No obstante Prothero y Wallis (1976), que proponen también un método complementario al método de la varianza; consideran que este último método funciona adecuadamente cuando son elevadas las diferencias entre las varianzas correspondientes a distintas transformaciones de la serie.

Finalmente, se va a plantear otro problema que puede surgir en la identificación de d . Si en un modelo alguna de las raíces de la parte autorregresiva está próxima a la unidad será difícil distinguirla de una raíz unitaria. Así el comportamiento del modelo ARIMA (0, 1, 1)

$$Y_t - Y_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (4-21)$$

no sería fácilmente distinguible del comportamiento del modelo ARIMA (1, 0, 1)

$$Y_t - \phi_1 Y_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (4-22)$$

cuando ϕ_1 está cercana a la unidad.

Cuando se presentan dudas de este tipo, se recomienda optar por un modelo del primer tipo, ya que presenta una estructura más flexible en lo que respecta a la predicción.

La inspección del gráfico de la serie es también útil en la determinación del orden de diferenciación. Así, según indican Box y Jenkins (1976), cuando la serie muestra una tendencia persistente o bien deambula hacia arriba o hacia abajo sin mostrar afinidad hacia un nivel determinado, será necesario en general tomar alguna diferenciación. Si la serie es no estacionaria en nivel, como ocurre con la representada en la parte a) de la figura 4.16, puede transformarse tomando $d = 1$ en una serie aproximadamente estacionaria. Si la serie no es estacionaria en nivel, ni en pendiente, como la serie representada en la parte b), el número de diferencias a tomar deberá ser al menos 2.

Tratamiento de la no estacionariedad en varianza: Identificación de λ

La toma de diferencias de un determinado orden es suficiente en muchos casos para obtener series estacionarias en media y en varianza. No obstante, en series económicas que se extienden a lo largo de un período dilatado de tiempo y que están afectadas por una fuerte tendencia, suele ser necesario efectuar además alguna transformación instantánea del tipo de Box-Cox, para obtener una serie estacionaria en varianza y que al mismo tiempo tenga distribución normal. Como se recordará, este tipo de transformación quedaba definida mediante el parámetro λ . Cuando $\lambda = 1$, equivale a no transformar la serie original; cuando $\lambda = 0$, la transformación de Box-Cox consiste en tomar logaritmos neperianos; en los demás casos se eleva la serie a

un exponente fraccionario. En la práctica, las alternativas que se utilizan son $\lambda = 0$ y $\lambda = 1$.

A continuación se van a examinar dos instrumentos que se utilizan para detectar si la serie es estacionaria o no en varianza.

El primer instrumento consiste simplemente en el gráfico de la serie. Una inspección ocular de este gráfico puede revelar si se mantiene o no la dispersión de la serie en torno al nivel existente en cada momento. Es relativamente frecuente que a medida que aumenta el nivel, aumente también la dispersión, siendo aconsejable en ese caso tomar logaritmos neperianos, es decir, la transformación $\lambda = 0$.

El segundo instrumento a utilizar es el gráfico rango-media. Antes de examinar la forma de aplicar este instrumento se procederá a dar una justificación teórica.

Considérese el proceso

$$Y_t = e^{\beta_0 + \beta_1 t + \epsilon_t} \quad (4-23)$$

Si se toman logaritmos neperianos se tiene que:

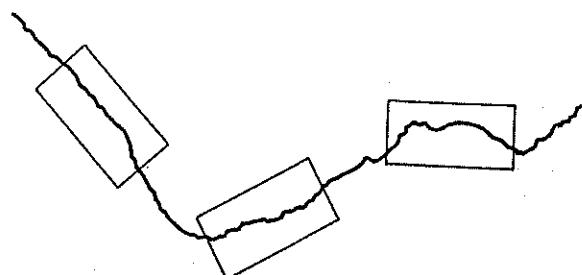
$$\ln Y_t = \beta_0 + \beta_1 t + \epsilon_t \quad (4-24)$$

cuya distribución, teniendo en cuenta las hipótesis sobre ϵ_t , es la siguiente:

$$\ln Y_t \sim N[\beta_0 + \beta_1 t, \sigma^2_{\epsilon}] \quad (4-25)$$



a) Serie no estacionaria en nivel ($d = 1$)



b) Serie no estacionaria en nivel ni en pendiente ($d = 2$)

Fig. 4.16. Gráficos de series no estacionarias [Fuente: Box y Jenkins (1976)].

Así pues, el proceso (4-25) es estacionario en varianza aunque no en media.

Como $\ln Y_t$ tiene una distribución normal, se verifica que Y_t tiene una distribución log-normal con los parámetros siguientes:

$$E[Y_t] = e^{\beta_0 + \beta_1 t + \frac{1}{2}\sigma^2_{\epsilon}} \quad (4-26)$$

$$\sigma_{Y_t} = e^{\beta_0 + \beta_1 t + \sigma^2_{\epsilon}} - e^{\beta_0 + \beta_1 t + \frac{1}{2}\sigma^2_{\epsilon}} \quad (4-27)$$

Tanto la esperanza como la desviación típica de Y_t dependen de t . En consecuencia Y_t no es estacionaria, ni en media ni en varianza.

La relación entre la desviación típica y la esperanza de Y_t es la siguiente:

$$\frac{\sigma_{Y_t}}{E[Y_t]} = e^{\frac{1}{2}\sigma^2_{\epsilon}} - 1 = k \quad (4-28)$$

o lo que es lo mismo

$$\sigma_{Y_t} = k E[Y_t] \quad (4-29)$$

Si β_1 es positivo, la relación entre σ_{Y_t} y $E[Y_t]$ se podrá representar mediante una recta con pendiente positiva, de acuerdo con (4-29).

En cambio, si representamos σ_{rav_t} (en ordenadas) y $E(\ln Y_t)$ (en abcisas) se obtendrá una línea recta paralela al eje de abcisas, ya que σ^2_{ϵ} no varía al variar t , y por tanto al variar $E(\ln Y_t)$.

Las anteriores consideraciones justifican la utilización de un gráfico dispersión media como instrumento para detectar si una determinada serie es estacionaria en varianza. Se suele utilizar el rango o recorrido como medida de la dispersión dada la sencillez de cálculo. Ahora bien, a partir de una sola serie temporal no se puede estimar la media y el rango para cada valor de t . Para poder obviar este problema lo que se hace es dividir la serie en varios intervalos, calculándose para cada uno de ellos la media (como medida de nivel) y el rango (como medida de dispersión). Si los datos son de carácter estacional, se suele tomar la longitud del intervalo igual a la del período estacional; así para datos mensuales cada uno de los intervalos tendrá una longitud de 12 datos. En el gráfico se representan cada par de valores de rango y media; si los correspondientes puntos están más o menos alineados en torno a una línea recta con pendiente ascendente sería indicativo de que los datos no son estacionarios en varianza y de que la toma de logaritmos sería la transformación apropiada para conseguir la estacionariedad. Cuando el gráfico no muéstre ningún esquema claro, o bien los puntos estén alineados en torno a una línea paralela al eje de abscisas, entonces no será necesario realizar ningún tipo de transformación instantánea. Caben otras posibilidades como las señaladas en la figura 4.17.

Puede ocurrir que los datos no sean sensibles a la elección de λ ; en este caso el gráfico de rango/media de los datos originales y el correspondiente de los datos transformados serían muy similares. Naturalmente si se da esta circunstancia no sería aconsejable realizar ninguna transformación sobre los datos originales. Conviene tener en cuenta por otra parte que, como han probado Nelson y Granger (1979), la identifi-

cación del modelo ARIMA puede ser muy sensible al λ seleccionado. Algunos autores como Box y Jenkins (1976) proponen estimar λ junto con los demás parámetros del modelo por el método de máxima verosimilitud.

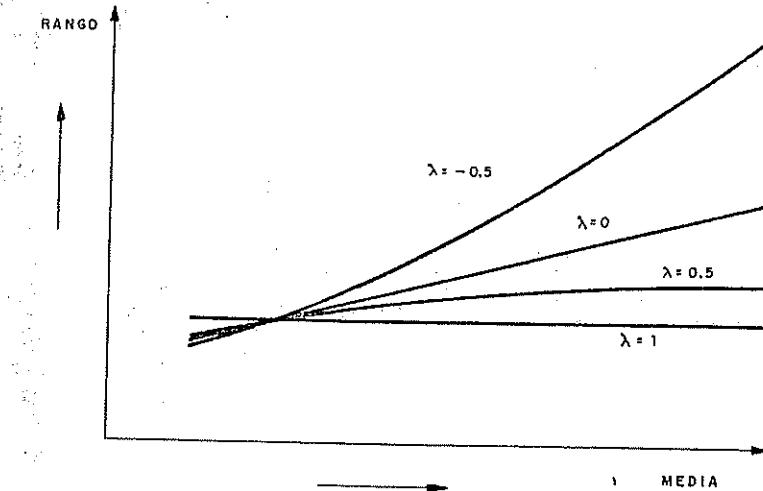


Fig. 4.17. Relación teórica entre rango y media para diferentes valores de λ de la transformación de Box-Cox.

4.5. Identificación del término independiente

Sin pérdida de generalidad, en el tratamiento teórico del capítulo 3 se ha considerado el supuesto de que la media del proceso es nula. Ahora en la etapa de identificación se trata de determinar si se incluye la media del proceso $w_i - (\mu_w)$ —o alternativamente un término independiente (δ)—en la especificación del proceso.

Si los datos de una serie $\{w_i\}$ proceden de un muestreo aleatorio simple sobre una población normal—, en otras palabras, si los datos constituyen una realización de

un proceso de ruido blanco con media distinta de 0—se utilizaría, para el contraste de la hipótesis nula de que $\mu = 0$, el siguiente estadístico

$$t_{N-1} = \frac{\bar{w}}{\sqrt{S_w^2/(N-1)}} \quad (4-30)$$

donde

\bar{w} y S_w^2 son la media y la varianza de la serie $\{w_i\}$.

En la fórmula anterior, se utiliza el resultado de que bajo las hipótesis establecidas $\hat{\sigma}_w^2 = S_w^2/(N-1)$ es un estimador insesgado y consistente de la varianza de la media muestral. Sin embargo, cuando los datos están autocorrelacionados, como ocurre en un proceso ARIMA, dicho estimador no resulta válido por ser inconsistente. En este caso, el estimador de la varianza muestral puede aproximarse mediante la siguiente expresión:

$$\hat{\sigma}_w^2 = \frac{S_w^2}{N} (1 + 2r_1 + 2r_2 + \dots + 2r_k) \quad (4-31)$$

seleccionando K de forma que se incluyan los coeficientes que sean significativos. Conviene tener en cuenta que al calcular la expresión (4-31), se puede obtener un resultado negativo en cuyo caso no sería aplicable.

Ejercicios

- [1] De la serie Y_t se dispone de las siguientes observaciones

t	1	2	3	4	5	6	7	8	9	10
Y_t	10	-5	8	-7	6	-5	9	1	6	-3

Se pide:

- Calcular r_1 , r_2 , r_3
- Analizar la significatividad estadística de r_1 , r_2 y r_3 , estableciendo las hipótesis que considere convenientes.

- [2] En el proceso estocástico generador

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

se conoce el valor del coeficiente de autocorrelación parcial teórico $\phi_{11} = 0.9$. Sabiendo que $\rho_2 = 0.8$, ¿puede calcular el valor de ϕ_{12} y ϕ_{22} ?

3) La serie temporal Y_t está integrada por las siguientes observaciones:

t	1	2	3	4	5	6	7	8	9	10
Y_t	7	9	10	6	5	3	1	2	3	4

Se pide:

- a) Calcular r_1 , r_2 y r_3 .
- b) Calcular una banda de confianza del 95 % para p_1 , p_2 y p_3 .
- c) Calcular $\hat{\phi}_{11}$, $\hat{\phi}_{12}$ y $\hat{\phi}_{33}$ a partir de las ecuaciones de Yule-Walker.
- d) Calcular una banda de confianza del 95 % para $\hat{\phi}_{11}$, $\hat{\phi}_{12}$ y $\hat{\phi}_{33}$.

4) La serie temporal Y_t está integrada por las siguientes desviaciones:

t	1	2	3	4	5	6	7	8	9	10
Y_t	2	3	1	-1	-4	-2	0	2	1	-2

Se pide:

- a) Calcular r_1 , r_2 y r_3 .
- b) Calcular $\hat{\phi}_{11}$, $\hat{\phi}_{12}$ y $\hat{\phi}_{33}$ por Yule-Walker.
- c) Calcular $\hat{\phi}_{11}$, $\hat{\phi}_{12}$ y $\hat{\phi}_{33}$ mediante regresión.
- d) ¿Cuál es el origen de la diferencia entre los resultados obtenidos en b) y c)?

5) A partir de una serie temporal de 100 elementos se han obtenido los siguientes resultados:

Retardo	1	2	3	4	5
r_t	-0,902	0,816	-0,714	0,671	-0,584

Se pide:

- a) Contrastar la significatividad de r_1 , r_2 y r_3 , r_4 y r_5 para un nivel de confianza del 95 %.
- b) Calcular $\hat{\phi}_{11}$, $\hat{\phi}_{12}$ y $\hat{\phi}_{33}$ mediante las ecuaciones de Yule-Walker.
- c) Contrastar la significatividad de $\hat{\phi}_{11}$, $\hat{\phi}_{12}$ y $\hat{\phi}_{33}$ al nivel del 95 %.
- d) Identifique p y q del proceso ARMA(p, q) que hipotéticamente ha podido generar los datos suministrados.

6) A partir de una serie temporal integrada por 144 observaciones se han obtenido los siguientes resultados:

Retardo	1	2	3	4	5
r_t	-0,490	0,040	-0,026	0,009	0,013
$\hat{\phi}_{rr}$	-0,490	-0,318	-0,227	-0,159	-0,113

Se pide:

- a) Analizar la significatividad de los coeficientes de autocorrelación parcial.
- b) Identificar el posible proceso generador de los datos propuestos.

7) A partir de 200 observaciones de la serie temporal Y_t se han obtenido los siguientes resultados:

Retardo	1	2	3	4	5
r_t	-0,909	0,727	-0,545	0,454	-0,364
$\hat{\phi}_{rr}$	-0,909	-0,317	-0,225	-0,136	-0,041

Se pide identificar el posible proceso generador.

8) A partir de la muestra de 200 observaciones se han obtenido los siguientes resultados:

Retardo	1	2	3	4	5
r_t	0,777	0,277	-0,268	-0,468	-0,701
$\hat{\phi}_{rr}$	0,777	-0,859	-0,082	0,041	-0,057

Se pide identificar el posible proceso generador.

9) A partir de la muestra de 100 observaciones se han obtenido los siguientes resultados:

Retardo	1	2	3	4	5
r_t	0,736	0,232	0,035	-0,115	0,013
$\hat{\phi}_{rr}$	0,736	-0,541	0,359	-0,191	0,047

Se pide identificar el hipotético proceso generador.

10) A partir de 240 observaciones de un índice de producción se han obtenido los siguientes resultados:

Serie	Retardo	1	2	3	4	5	6	7	8
Y_t	r_t	0,881	0,843	0,829	0,844	0,826	0,830	0,803	0,801
	$\hat{\phi}_{rr}$	0,881	0,295	0,224	0,282	0,072	0,165	-0,035	0,064
ΔY_t	r_t	-0,510	0,042	-0,069	0,016	0,082	0,054	0,014	0,007
	$\hat{\phi}_{rr}$	-0,510	-0,172	-0,108	-0,070	-0,149	-0,089	-0,005	0,019

Se pide:

- a) ¿Le parece correcto transformar la serie tomando diferencias de primer orden?
- b) Identificar el posible proceso generador.

11) Sea el modelo

$$Y_t = \phi_1 Y_{t-1} + \epsilon_t$$

A una serie temporal de Y_t generada por este proceso se le aplica una diferencia de primer orden obteniéndose el proceso $w_t = \Delta Y_t$.

¿En cuál de las dos series Y_t y w_t esperaría que fuera mayor la varianza?

ELABORACIÓN DE MODELOS ARIMA: FASE DE IDENTIFICACIÓN

[12] Con datos de la serie A del apéndice 4 (Tasa de nacimientos por mil habitantes en España) a la que se designaría por Y_t se ha calculado los siguientes estadísticos utilizando el

EJERCICIOS

Serie Y_t

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
r_t	0,93	0,88	0,84	0,78	0,74	0,69	0,66	0,63	0,58	0,55	0,51	0,49	0,45	0,42	0,40
\hat{a}_t	0,11	0,18	0,23	0,27	0,29	0,31	0,33	0,35	0,36	0,37	0,38	0,39	0,40	0,41	0,41
$\hat{\sigma}_e$	0,93	0,15	0,01	-0,11	0,06	-0,08	0,11	0,02	-0,09	-0,01	0,02	0,06	-0,12	0,06	-0,02

$$\bar{Y} = 24,98 \quad S_y = 5,57$$

Serie ΔY_t

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
r_t	-0,38	0,09	0,15	-0,09	0,11	-0,22	-0,04	0,17	-0,14	-0,01	-0,07	0,21	-0,10	-0,01	0,04
\hat{a}_t	0,11	0,13	0,13	0,13	0,13	0,14	0,14	0,14	0,14	0,14	0,14	0,14	0,15	0,15	0,15
$\hat{\sigma}_e$	-0,38	-0,17	0,10	0,02	0,11	-0,20	-0,24	0,04	0,03	-0,02	-0,17	0,12	-0,01	0,06	-0,01

$$\bar{W} = -0,24 \quad S_w = 1,44$$

Serie $W_t = \Delta^2 Y_t$

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
r_t	-0,64	0,08	0,15	-0,16	0,19	-0,19	-0,01	0,19	-0,16	0,07	-0,13	0,22	0,14	0,01	0,07
\hat{a}_t	0,11	0,15	0,15	0,15	0,16	0,16	0,16	0,16	0,17	0,17	0,17	0,17	0,17	0,17	0,17
$\hat{\sigma}_e$	-0,64	-0,55	-0,30	-0,29	0,06	0,05	-0,23	-0,18	0,10	-0,07	0,20	0,03	0,08	0,00	0,12

$$\bar{W} = 0,03 \quad S_w = 2,41$$

En la figura 4.18 se ha representado la evolución de la serie en el tiempo, la FACE y la FACPE de las tres transformaciones anteriores.

A la vista de la información anterior

- ¿Qué valor d seleccionaría con objeto de obtener una serie estacionaria?
- Una vez seleccionado d , identifíquese p y q del proceso resultante.
- Incluiría un término independiente en el modelo seleccionado?

Tanto en el punto a) como en el b) puede considerar varias alternativas.

TASA DE NACIMIENTOS POR 1.000 HABITANTES
NÚMERO DE OBSERVACIONES: 81

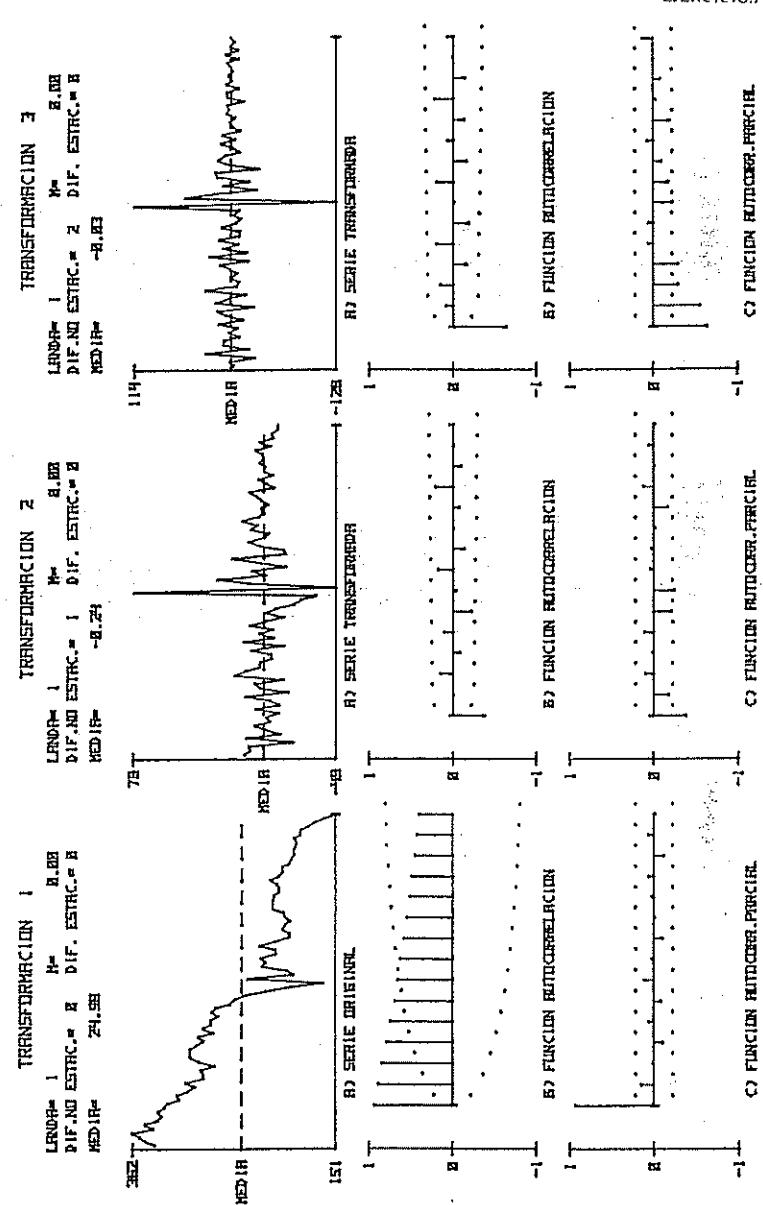


Fig. 4.18.

- 13 Designando por Y_t a la serie B del apéndice 4 —Renta nacional neta al coste de los factores deflactada, en millones de pesetas— se ha calculado los siguientes estadísticos utilizando el período 1940-1980:

Serie Y_t

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	0,94	0,87	0,80	0,72	0,64	0,56	0,48	0,39	0,32	0,25
\hat{a}_{rt}	0,16	0,26	0,32	0,37	0,40	0,43	0,44	0,46	0,46	0,47
$\hat{\delta}_{tt}$	0,94	-0,08	-0,09	-0,06	-0,06	-0,07	-0,08	-0,03	0,01	0,01

Rango, desviación típica y media por subperiodos.

Período	Rango	Desviación típica	Media
1941-1950	120	45	583
1951-1960	383	129	883
1961-1970	1.031	333	1.526
1971-1980	986	353	2.758
1941-1980	2.643	910	1.479

Serie $w_t = \Delta Y_t$

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	0,62	0,41	0,32	0,39	0,32	0,16	0,14	0,13	0,08	0,11
\hat{a}_{rt}	0,16	0,21	0,23	0,24	0,26	0,27	0,27	0,27	0,27	0,27
$\hat{\delta}_{tt}$	0,62	0,05	0,09	0,25	-0,05	-0,14	-0,08	-0,03	-0,09	0,19

$$\bar{W} = 66 \quad S_w = 61$$

Serie $w_t = \Delta^2 Y_t$

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	-0,23	-0,17	-0,20	-0,18	0,14	-0,17	-0,03	0,07	-0,13	-0,03
\hat{a}_{rt}	0,16	0,17	0,17	0,18	0,18	0,19	0,19	0,19	0,19	0,19
$\hat{\delta}_{tt}$	-0,23	-0,23	-0,34	-0,04	0,09	-0,13	-0,02	0,07	-0,24	-0,06

$$\bar{W} = 0,7 \quad S_w = 52$$

Serie $\ln Y_t$

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	0,94	0,88	0,82	0,75	0,68	0,60	0,53	0,45	0,38	0,30
\hat{a}_{rt}	0,16	0,26	0,32	0,37	0,41	0,43	0,45	0,47	0,48	0,49
$\hat{\delta}_{tt}$	0,94	-0,07	-0,05	-0,07	-0,06	-0,09	-0,03	-0,05	-0,05	-0,04

Rango, desviación típica y media por subperiodos.

Período	Rango	Desviación típica	Media
1941-1950	0,207	0,078	13,273
1951-1960	0,453	0,150	13,681
1961-1970	0,700	0,226	14,216
1971-1980	0,379	0,135	14,822
1941-1980	1,808	0,619	14,022

Serie $w_t = \Delta \ln Y_t$

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	0,21	0,05	-0,14	0,13	0,12	-0,06	-0,09	-0,08	-0,07	0,16
\hat{a}_{rt}	0,16	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,18	0,18
$\hat{\delta}_{tt}$	0,21	0,01	-0,16	0,20	0,07	-0,16	0,00	-0,03	-0,13	0,24

$$\bar{W} = 0,045 \quad S_w = 0,038$$

Serie $w_t = \Delta^2 \ln Y_t$

Coeficientes de autocorrelación y autocorrelación parcial.

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	-0,39	0,00	-0,28	0,16	0,15	-0,10	-0,03	0,01	-0,15	0,20
\hat{a}_{rt}	0,16	0,18	0,18	0,19	0,20	0,20	0,20	0,20	0,20	0,20
$\hat{\delta}_{tt}$	-0,39	-0,18	-0,43	-0,24	0,04	-0,12	-0,09	0,06	-0,30	-0,06

$$\bar{W} = 0,000 \quad S_w = 0,047$$

En las figuras 4.19 y 4.20 se han representado la evolución de la serie en el tiempo, la FACE y la FACPE de las seis transformaciones anteriores.

A la vista de la información anterior:

- Representar los gráficos rango-media y desviación típica-media para Y_t y $\ln Y_t$.
- Limitando la elección a $\lambda = 0$ y $\lambda = 1$, ¿qué transformación de Box-Cox considera más adecuada?
- ¿Qué valores d y λ seleccionaría para obtener una serie estacionaria?
- Una vez determinados d y λ , identifique p y q del proceso resultante. En este punto, como en el anterior, puede considerar diversas alternativas.
- ¿Incluiría un término independiente en el modelo seleccionado?

RENTA NACIONAL NETA DEFLACTADA (C.F.)
NUMERO DE OBSERVACIONES UTILIZADAS 41

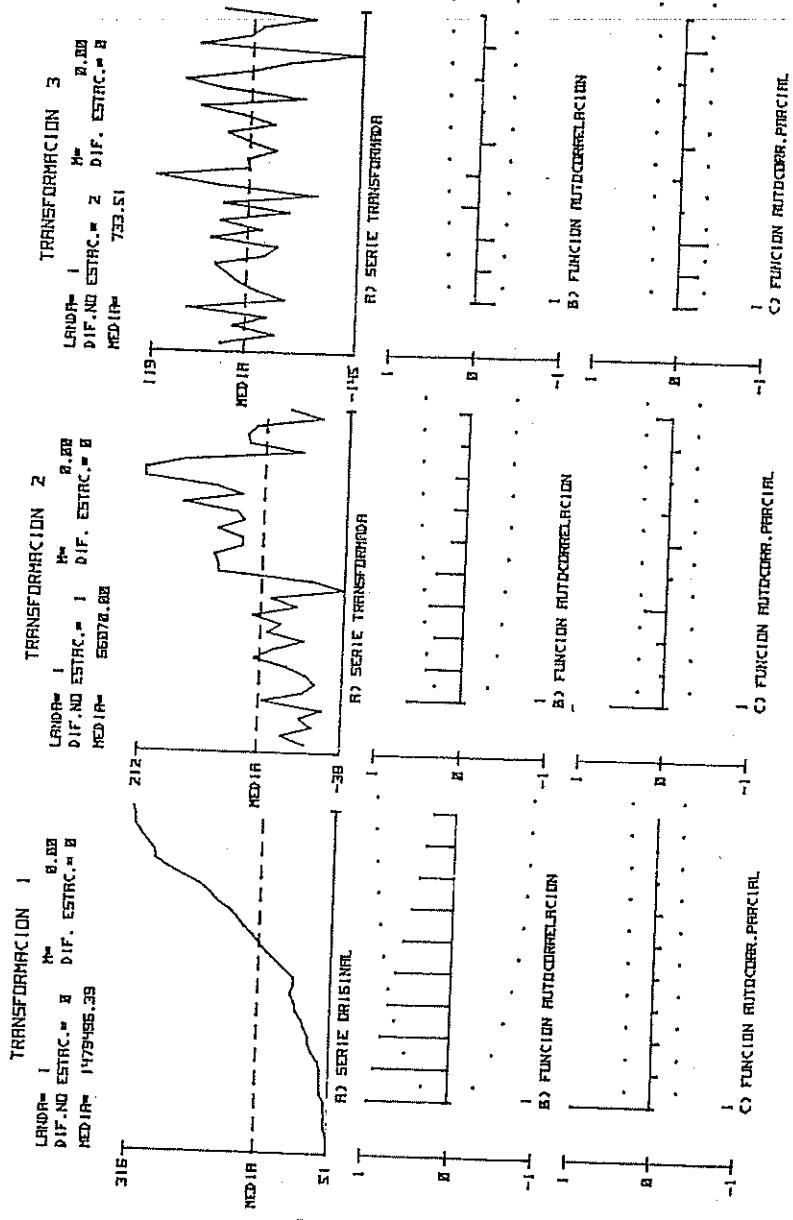


Fig. 4.19.

RENTA NACIONAL NETA DEFLACTADA (C.F.)
NUMERO DE OBSERVACIONES UTILIZADAS 41

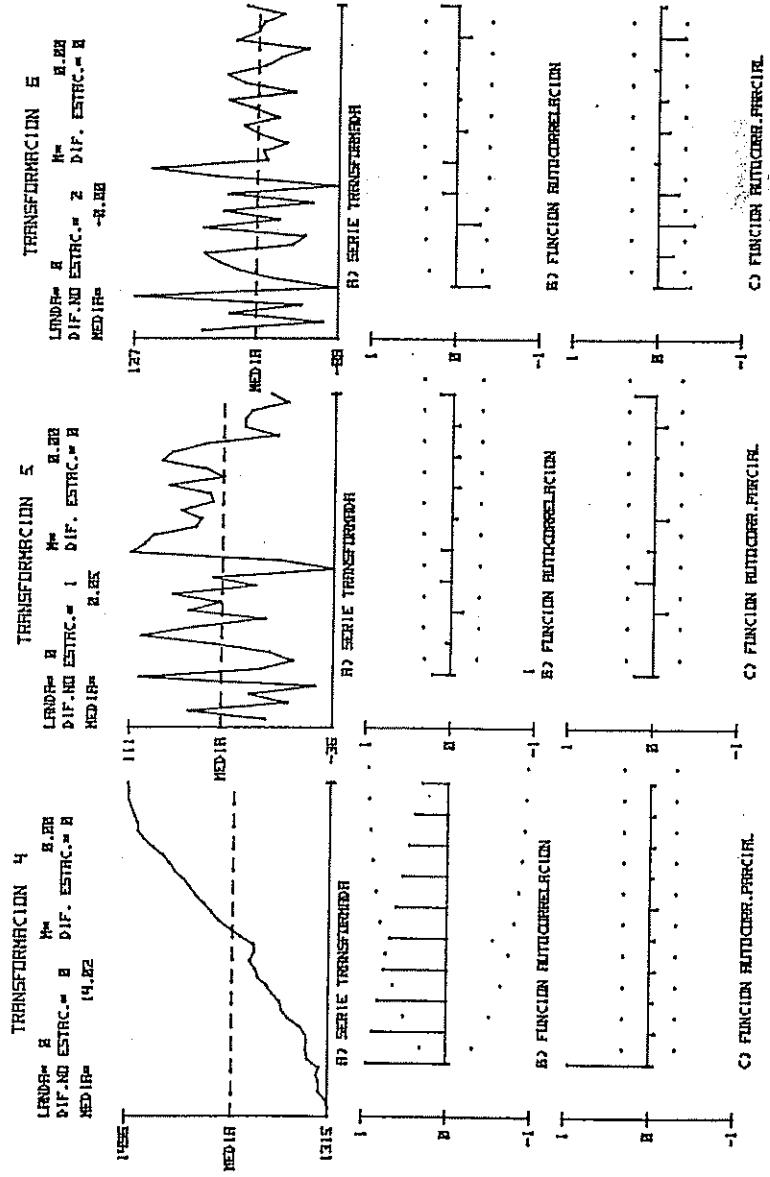


Fig. 4.20.

- [14] Denominando Y_t a la serie C del apéndice 4 —serie simulada— y utilizando los 100 primeros datos:
- Calcular la FACE y la FACPE de Y_t , ΔY_t y $\Delta^2 Y_t$.
 - ¿Qué orden de diferencias sería el adecuado para obtener una serie estacionaria?
 - Una vez realizada la transformación que considere adecuada, identifique p y q del proceso resultante.
- [15] Denominando Y_t a la serie D del apéndice 4 —serie simulada— y utilizando los 200 primeros datos:
- Calcular la FACE y la FACPE de Y_t , ΔY_t y $\Delta^2 Y_t$.
 - ¿Qué orden de diferencias sería el adecuado para obtener una serie estacionaria?
 - Una vez realizada la transformación que considere adecuada, identifique p y q del proceso resultante.
- [16] Denominando Y_t a la serie E del apéndice 4 —serie simulada— y utilizando los 100 primeros datos:
- Calcular la FACE y la FACPE de Y_t , ΔY_t y $\Delta^2 Y_t$, $\ln Y_t$, $\Delta \ln Y_t$ y $\Delta^2 \ln Y_t$.
 - Calcular el estadístico rango-media para cada una de las transformaciones del apartado a), utilizando tramos de 10 observaciones.
 - ¿Qué transformaciones efectuaría para obtener una serie estacionaria?
 - Una vez realizada la transformación que considere adecuada, identifique p y q del proceso resultante.

CAPÍTULO 5

Estimación de un modelo ARMA

5.1. Planteamiento

En la fase de identificación, en una primera etapa a la serie Y_t se le aplica una transformación de forma que la serie

$$w_t = (1 - L)^d Y_t^{(0)} \quad (5-1)$$

sea estacionaria.

En la segunda etapa de la fase de identificación se procede a identificar los valores de la parte autorregresiva (p) y de la parte de medias móviles (q) del hipotético proceso generador ARMA (p, q). Así se llega a la expresión:

$$w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p} = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} + \delta_w \quad (5-2)$$

Se ha incluido un término constante δ_w , que se relaciona con la media del proceso $w_t (\mu_w)$ de la siguiente forma:

$$\mu_w = \frac{\delta_w}{1 - \phi_1 - \dots - \phi_p} \quad (5-3)$$

Alternativamente el proceso (5-2) se puede expresar así:

$$(w_t - \mu_w) - \phi_1 (w_{t-1} - \mu_w) - \dots - \phi_p (w_{t-q} - \mu_w) = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (5-4)$$

En la fase de estimación, como es lógico, el objetivo es la obtención de estimadores de los parámetros $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \mu_w$ y σ_ϵ^2 . Para designar a todos los parámetros ϕ_i y θ_i se utilizará el vector fila

$$\beta' = [\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q]$$

- Los supuestos que se adoptarán en el proceso de estimación son los siguientes:
- $\varepsilon \sim N(0, \sigma^2, I)$.
 - El proceso w_t es estacionario.
 - El proceso w_t es invertible.
 - δ_w y, en consecuencia, μ_w son nulos.

Para realizar la estimación se dispone de una muestra de tamaño T de la variable Y . Al tomar diferencias de orden d , la variable w_t constará de $T-d$ elementos. Con objeto de simplificar la notación, a la diferencia $T-d$ se le designará por N , y a los elementos de la muestra de w_t se les designará por el vector fila:

$$w' = (w_1, w_2, \dots, w_N)$$

Si se adopta el criterio minimocuadrático, los estimadores se obtienen minimizando la expresión:

$$S = \sum \varepsilon_t^2 \quad (5-5)$$

Si se adopta el criterio maximoverosímil, entonces los estimadores se obtendrán mediante maximización de la función de verosimilitud

$$L = f(\varepsilon_1, \dots, \varepsilon_N) \quad (5-6)$$

Cualquiera que sea el criterio adoptado, a la estimación de un modelo ARMA se plantea por una parte el problema de determinación de los valores iniciales y por otra parte el problema de que, en general, los modelos de este tipo son no lineales. A continuación se examina el alcance y el tipo de solución que se puede dar a cada uno de estos problemas.

Determinación de valores iniciales

Si en el modelo (5-2) se dan valores a t desde $t=1$ hasta $t=N$, se observa que el conocimiento de los parámetros β , y de una muestra w' no es suficiente para determinar los valores de ε_t . Así, para $t=1$, se tiene que

$$\varepsilon_1 = w_1 - \phi_1 w_0 - \dots - \phi_p w_{-p+1} + \theta_1 \varepsilon_0 - \dots + \theta_q \varepsilon_{-q+1} \quad (5-7)$$

En la expresión anterior, bajo los supuestos considerados, se desconocen $w_0, \dots, w_{-p+1}, \varepsilon_0, \dots, \varepsilon_{-q+1}$. Estos valores constituyen precisamente las condiciones iniciales; en adelante se les designará mediante los vectores fila

$$\begin{aligned} w^{0r} &= [w_0, w_{-1}, \dots, w_{-p+1}] \\ \varepsilon^{0r} &= [\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-q+1}] \end{aligned}$$

Tanto en la minimización de (5-5), como en la maximización de (5-6) hay que tener en cuenta el papel que juegan estos valores iniciales. A este respecto en el proce-

so de estimación se puede adoptar un enfoque condicional o alternativamente un enfoque al que se le puede designar con la denominación de no condicional. En el cuadro 5.1 se señala para cada enfoque qué valores se determinan en el proceso de estimación y qué valores se consideran como condiciones iniciales.

Cuadro 5.1. Enfoque condicional y no condicional en la estimación de un modelo ARMA (p, q).

Enfoque	Valores a estimar	Valores dados
Condicional	$\beta' = [\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q]$ $\sigma_{\varepsilon}^2 (\mu_w \text{ en su caso})$	$w' = [w_1, \dots, w_N]$ $\varepsilon^{0r} = [w_0, w_{-1}, \dots, w_{-p+1}]$ $\varepsilon^{0r} = [\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-q+1}]$
No condicional	$\beta' = [\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q]$ $w^{0r} = [w_0, w_{-1}, \dots, w_{-p+1}]$ $\varepsilon^{0r} = [\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-q+1}]$ $\sigma_{\varepsilon}^2 (\mu_w \text{ en su caso})$	$w' = [w_1, \dots, w_N]$

Como es lógico, en los dos enfoques aparece la muestra observada w' como valores dados. En el enfoque condicional también aparecen como valores dados los valores iniciales w^{0r} y ε^{0r} que son en principio desconocidos; en este enfoque y antes de iniciar la estimación, lo que se hace precisamente es asignar unos valores apropiados a w^{0r} y a ε^{0r} basados en las hipótesis estadísticas del modelo. En el enfoque no condicional, los valores iniciales se determinan en el proceso de estimación, conjuntamente con los parámetros del modelo. En los epígrafes 5.2 y 5.3, se examinan los dos enfoques, comenzando con el enfoque condicional que es más fácil de aplicar, pero menos eficiente estadísticamente que el enfoque no condicional.

No linealidad

Si se verifica que $q \geq 1$, el modelo será no lineal en los coeficientes. En efecto, si se utilizan operadores polinomiales de retardo, el modelo (5-2) se puede expresar así:

$$\varepsilon_t = \frac{\phi(L)}{\theta(L)} w_t \quad (5-8)$$

La presencia en el denominador del segundo miembro de (5-8) del operador $\theta(L)$ distinto de 1 hace que la expresión a minimizar sea no lineal en los parámetros. En un modelo de este tipo no es posible obtener soluciones analíticas; en su lugar habrán de utilizarse métodos iterativos según se verá en el epígrafe 5.2.

5.2. Enfoque condicional

Con el enfoque condicional se examinará en primer lugar el método de mínimos cuadrados que se aplicará sucesivamente a modelos AR, a modelos MA y a modelos

ARMA. Como segundo método de estimación se examinará el método de máxima verosimilitud aplicado a modelos ARMA.

Estimación por mínimos cuadrados de un modelo AR.

Sea un modelo AR(p)

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \varepsilon_t \quad (5-9)$$

Si el proceso es estacionario, entonces

$$E[w_t] = E[w_{t-1}] = \dots = E[w_{t-p}] = \dots = \mu_w$$

que por hipótesis es igual a 0.

Una solución al problema de los valores iniciales es suponer que son igual a su media, es decir

$$w_0 = w_{-1} = \dots = w_{-p+1} = 0$$

Esta solución puede ser inadecuada en muchos casos debido a que por influencia del transiente —o parte sistemática de la ecuación en diferencias—, los valores inicia-

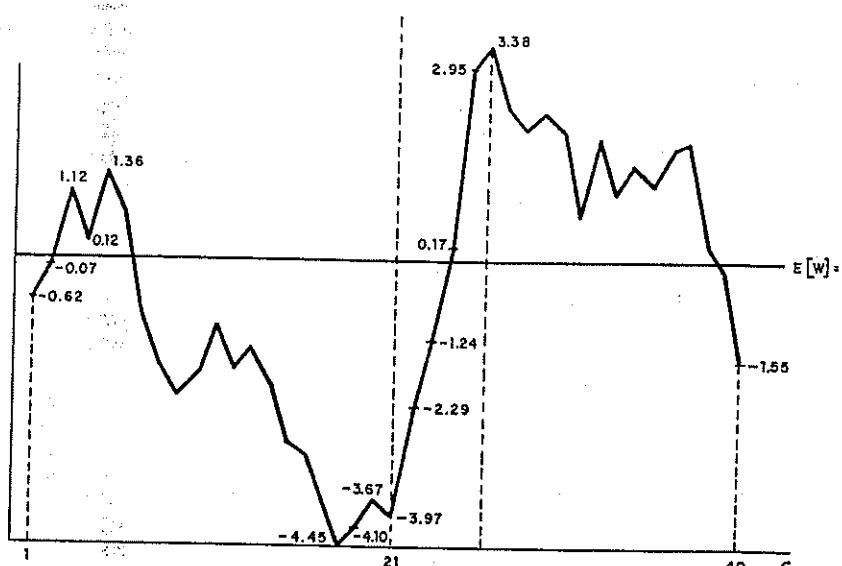


Fig. 5.1. Muestra generada a partir del modelo $w_t = 0.95 w_{t-1} + \varepsilon_t$.

les pueden estar muy alejados del valor medio teórico de la serie. El siguiente ejemplo es ilustrativo de este problema.

En base al modelo

$$w_t = 0.95 w_{t-1} + \varepsilon_t$$

y considerando $w_0 = 0$ se ha generado una muestra de 40 observaciones que se ha representado en la figura 5.1.

Se adopta el supuesto de que al investigador se le han facilitado únicamente los últimos 20 datos, en base a los cuales ha identificado que el proceso generador de la serie es un AR(1) y que $E(w_t) = 0$. La muestra de que dispone es:

$$w' = (w_1 \dots w_N) = (-3.97 \dots -2.29 \dots -1.55)$$

Como no se conoce w_0 el investigador le asigna su valor medio teórico, es decir, lo hace igual a 0. Este valor está en este caso muy alejado del valor w_0 obtenido en la serie simulada ($w_0 = 3.67$), lo que puede provocar distorsiones importantes en el proceso de estimación.

Para evitar este problema se puede adoptar otra solución más conveniente. Consiste en dividir la muestra en dos partes, la primera de las cuales se utiliza para determinación de los valores iniciales, siendo la segunda parte la que se utiliza en el proceso de estimación.

En el ejemplo propuesto, solamente se requiere un valor inicial, ya que se trata de un modelo AR(1). Por tanto:

$$w' = (w_1 \dots w_N) = (-2.29 \dots -1.55)$$

$$w^{(0)} = (w_1) = (-3.97)$$

Teniendo en cuenta ambas posibilidades para la determinación de valores iniciales, en un modelo AR(p) la expresión a minimizar sería la siguiente:

$$S = \sum_{t=m}^N \varepsilon_t^2 = \sum_{t=m}^N (w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p})^2 \quad (5-10)$$

donde

$$\begin{aligned} m &= 1 && \text{si se hace } w_0 = w_{-1} = \dots = w_{-p+1} = E(w_t) = 0 \\ m &= p+1 && \text{si se toman } w_1, \dots, w_p \text{ como valores iniciales.} \end{aligned}$$

La minimización de S con respecto a los parámetros no plantea ningún problema, ya que al ser las ecuaciones resultantes lineales en los parámetros, se obtiene una solución analítica mediante las fórmulas usuales de regresión mínima cuadrática.

Con objeto de conseguir una aproximación a los problemas que se plantean en la estimación de un modelo de medias móviles se va a contemplar previamente la minimización de S de un modelo AR(1) desde la perspectiva de un proceso iterativo.

En dicho modelo se verifica que

$$\varepsilon_t = w_t - \phi_1 w_{t-1} \quad (5-11)$$

En el proceso de estimación, las w_t aparecen como constantes y ϕ_1 será la variable independiente, ya que se trata de encontrar un valor de ϕ_1 tal que resulte maximizada S . Así pues, (5-11) se puede expresar de la siguiente forma:

$$\varepsilon_t = f(\phi_1) \quad (5-12)$$

Desarrollando en serie ε_t en el entorno de ϕ_1^0 mediante la aproximación de Taylor y tomando los dos primeros términos, se tiene que:

$$\varepsilon_t \approx \varepsilon_t^0 + (\phi_1 - \phi_1^0) \left[\frac{\partial \varepsilon_t}{\partial \phi_1} \right]_{\phi_1 = \phi_1^0} \quad (5-13)$$

Ahora bien

$$\frac{\partial \varepsilon_t}{\partial \phi_1} = -w_{t-1} \quad (5-14)$$

es independiente de ϕ_1 , por lo que

$$\frac{\partial^h \varepsilon_t}{\partial \phi_1^h} = 0 \quad \text{para } h > 1$$

Por esta razón, en la expresión (5-13) se puede sustituir el signo \approx por el signo igual.

Así pues

$$\varepsilon_t = \varepsilon_t^0 - (\phi_1 - \phi_1^0) (w_{t-1}) \quad (5-15)$$

o, alternativamente,

$$\varepsilon_t^0 = (\phi_1 - \phi_1^0) (w_{t-1}) + \varepsilon_t \quad (5-16)$$

Por tanto, la minimización de $\sum \varepsilon_t^2$ sería equivalente a realizar la regresión de ε_t^0 —generada a partir de un valor arbitrario ϕ_1^0 — sobre w_{t-1} . El parámetro a estimar sería $\phi - \phi_1^0$.

En consecuencia

$$(\phi_1 \hat{\wedge} \phi_1^0) = \frac{\sum \varepsilon_t^0 w_{t-1}}{\sum w_{t-1}^2} \quad (5-17)$$

La regresión de ε_t sobre w_{t-1} puede contemplarse como un primer paso de un

proceso iterativo en que, partiendo de un valor arbitrario ϕ_1^0 , las estimaciones se van actualizando según el mecanismo siguiente:

$$\phi_1^h = \phi_1^{h-1} + (\phi_1 \hat{\wedge} \phi_1^0) \quad (5-18)$$

Así, una vez obtenido (5-17), aplicando la fórmula anterior, se tiene para la primera iteración que

$$\begin{aligned} \phi_1^1 &= \phi_1^0 + (\phi_1 \hat{\wedge} \phi_1^0) \\ &= \phi_1^0 + \frac{\sum \varepsilon_{t0} w_{t-1}}{\sum w_{t-1}^2} \\ &= \phi_1^0 + \frac{\sum (\nu_t - \phi_1^0 w_{t-1}) w_{t-1}}{\sum w_{t-1}^2} \\ &= \phi_1^0 + \frac{\sum w_t w_{t-1}}{\sum w_{t-1}^2} - \phi_1^0 \\ &= \frac{\sum w_t w_{t-1}}{\sum w_{t-1}^2} \end{aligned} \quad (5-19)$$

De la deducción anterior se pueden obtener las siguientes conclusiones:

- a) El estimador ϕ_1^1 coincide con el estimador $\hat{\phi}_1$ en la regresión de w_t sobre w_{t-1} .
- b) El estimador ϕ_1^1 no viene afectado por el valor inicial de ϕ_1^0 . Así cualquiera que fuera el valor inicial de ϕ_1^0 , el estimador ϕ_1^1 hubiera sido el mismo.
- c) El proceso iterativo se detiene en la primera iteración al haber llegado directamente al estimador que minimiza a S . Si se continúa el proceso iterativo se obtendría para ϕ_1^2, ϕ_1^3 , etc., idénticas expresiones que para (5-19).

Las tres conclusiones anteriores se derivan del hecho de que el modelo AR(1) es un modelo lineal en los parámetros al que se puede aplicar una solución analítica. Por ello, cuando se enfoca la estimación como un proceso iterativo, es suficiente con una sola iteración para llegar a la solución óptima.

Estimación por mínimos cuadrados de un modelo MA

El examen de toda la problemática que se plantea en la estimación de los modelos de medias móviles se puede realizar a partir del modelo más sencillo, es decir, a partir de un MA(1). Un modelo de este tipo se expresa de la siguiente forma:

$$\varepsilon_t = w_t + \theta_1 \varepsilon_{t-1} \quad (5-20)$$

En los modelos de medias móviles no se plantean problemas especiales en la fijación de los valores iniciales, ya que por hipótesis las ε_t están incorrelacionadas entre sí. Por tanto, en el modelo (5-20) parece razonable hacer $\varepsilon_0 = 0$, ya que este valor

coincide con su media teórica. Fijado este valor inicial se pretende expresar ε_t solamente en función de la variable observable w_t , del parámetro θ_i y de ε_0 . Así, mediante sustituciones sucesivas se tiene que:

$$\begin{aligned} \varepsilon_t &= w_t + \theta_i \varepsilon_{t-1} \\ &= w_t + \theta_i w_{t-1} + \theta_i^2 \varepsilon_{t-2} \\ &\dots \\ &= w_t + \theta_i w_{t-2} + \theta_i^2 w_{t-1} + \dots + \theta_i^{i-1} w_t + \theta_i^i \varepsilon_0 \end{aligned} \quad (5-21)$$

De la expresión anterior se deduce que ε_t no es función lineal de θ_i , a diferencia de lo que ocurre en los modelos autorregresivos. Por ello, si ε_t se expresa mediante la aproximación lineal del desarrollo de Taylor, es preciso mantener el signo \simeq según puede verse a continuación:

$$\varepsilon_t \simeq \varepsilon_t^0 - (\theta_i - \theta_i^0) \left[\frac{\partial \varepsilon_t}{\partial \theta_i} \right]_{\theta_i = \theta_i^0} \quad (5-22)$$

Las derivadas $\left[\frac{\partial \varepsilon_t}{\partial \theta_i} \right]$ evaluadas para $\theta_i = \theta_i^0$ se pueden calcular de dos formas:

a) De forma analítica.

Derivando ambos miembros de (5-20) con respecto a θ_i se tendría que

$$\frac{\partial \varepsilon_t}{\partial \theta_i} = \theta_i \frac{\partial \varepsilon_{t-1}}{\partial \theta_i} + \varepsilon_{t-1} \quad (5-23)$$

En la obtención de (5-23) se ha tenido en cuenta que en este proceso w_t juega el papel de constante y, por tanto, su derivada respecto a θ_i es nula.

El cálculo de (5-23) se haría de forma recursiva, sustituyendo θ_i por θ_i^0 y teniendo en cuenta que $\partial \varepsilon_t / \partial \theta_i = 0$, ya que por hipótesis $\varepsilon_0 = 0$. A diferencia de lo que ocurre en los modelos autorregresivos, esta derivada no es independiente del valor θ_i^0 que se toma para evaluarla.

b) De forma numérica.

En el cálculo numérico se aplica el concepto de derivada pero tomando incrementos finitos. Así

$$\frac{\partial \varepsilon_t}{\partial \theta_i} = \frac{w_t + (\theta_i^0 + h) \varepsilon_{t-1} - (w_t + \theta_i^0 \varepsilon_{t-1})}{h} \quad (5-24)$$

donde h es un número arbitrariamente pequeño (p. ej., $h = 0,01$).

Si se designa

$$x_t^0 = \left[\frac{\partial \varepsilon_t}{\partial \theta_i} \right]_{\theta_i = \theta_i^0} \quad (5-25)$$

y se considera válida la aproximación lineal (5-22) se tendría que:

$$\varepsilon_t^0 = (\theta_i - \theta_i^0) x_t^0 + \varepsilon_t \quad (5-26)$$

Calculando la regresión ε_t^0 sobre x_t^0 se tendría que:

$$(\theta_i \hat{\Delta} \theta_i^0) = \frac{\sum \varepsilon_t^0 x_t^0}{\sum x_t^{02}} \quad (5-27)$$

La expresión anterior constituye la primera iteración del proceso; la fórmula de actualización de las estimaciones sucesivas sería la siguiente:

$$\theta_i^h = \theta_i^{h-1} + (\theta_i \hat{\Delta} \theta_i^{h-1}) \quad (5-28)$$

Los cálculos se continúan hasta que se obtenga una convergencia satisfactoria, por ejemplo, hasta que

$$|\theta_i^h - \theta_i^{h-1}| < r$$

donde r es un número prefijado tan pequeño como se quiera (un número r usual es 0,001).

A diferencia de un modelo autorregresivo en que solamente se efectúa una iteración para llegar a la solución final, en un modelo MA —o análogamente en un modelo ARMA— son necesarias en general varias iteraciones, pudiendo ocurrir además que el proceso iterativo no conduzca al mínimo absoluto de la función que se quiere minimizar, sino a un mínimo local. Este último fenómeno queda ilustrado en la figura 5.2 para un modelo MA(1).

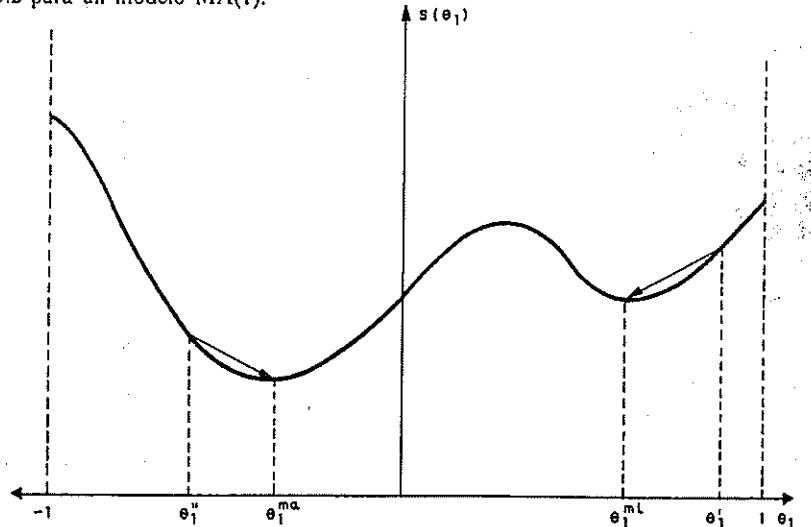


Fig. 5.2. Proceso de minimización en un MA(1).

En ordenadas se ha tomado la suma de los cuadrados de los residuos S , y se ha añadido entre paréntesis θ_i para resaltar que el valor de S depende del valor que tome θ_i . En abcisas se ha tomado θ_i con un recorrido correspondiente a un modelo invertido. Si se tomara como valor inicial $\theta_i^0 = \theta_i^1$, el proceso iterativo conduciría al valor fijo θ_i^m , que es un mínimo local, pero no un mínimo absoluto. En cambio si se parte de $\theta_i^0 = \theta_i^1$ se llegaría al mínimo absoluto de la función $S(\theta_i)$.

Estimación por mínimos cuadrados de un modelo ARMA

El procedimiento de estimación que se ha visto para un MA(1) se puede generalizar a un modelo ARMA(p, q). Este modelo se puede expresar de la siguiente forma:

$$\epsilon_t = w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (5-29)$$

Haciendo el vector $\beta' = [\beta_1, \beta_2 \dots \beta_{p+q}] = [\phi_1 \dots \phi_p, \theta_1 \dots \theta_q]$ y tomando como válida la aproximación lineal del desarrollo de Taylor se tendría que:

$$\epsilon_t = \epsilon_t^0 - \sum_{i=1}^{p+q} [\beta_i - \beta_i^0] X_i^0 \quad (5-30)$$

donde

$$X_i^0 = \begin{bmatrix} \frac{\partial \epsilon_t}{\partial \beta_i} \end{bmatrix} \Big|_{\beta = \beta_i^0} \quad (5-31)$$

Si se designa

$$X^0 = \begin{bmatrix} X_{11}^0 & X_{12}^0 & \dots & X_{1p+q}^0 \\ X_{21}^0 & X_{22}^0 & \dots & X_{2p+q}^0 \\ \dots & \dots & \dots & \dots \\ X_{N1}^0 & X_{N2}^0 & \dots & X_{Np+q}^0 \end{bmatrix}$$

Se puede expresar (5-30) para las N observaciones de forma compacta de la siguiente forma:

$$\epsilon^0 = X^0[\beta - \beta^0] + \epsilon \quad (5-32)$$

ϵ^0 y ϵ son vectores de N elementos.

La estimación de (5-32) por mínimos cuadrados sería la siguiente:

$$\hat{\beta} = [X^0 X^0]^T X^0 \epsilon^0 \quad (5-33)$$

y el mecanismo de actualización vendría dado por

$$\beta^h = \beta^{h-1} + [\hat{\beta} - \hat{\beta}^{h-1}] \quad (5-34)$$

El proceso iterativo continuaría hasta que se obtuviera una convergencia satisfactoria para cada uno de los parámetros.

Con objeto de poder efectuar contrastes estadísticos, para la iteración final (f) se calcula la estimación de la matriz de covarianzas de los estimadores mediante la fórmula:

$$V = s^2 (X^0 X^0)^{-1} \quad (5-35)$$

donde

$$s^2 = \frac{\epsilon^0 \epsilon^0}{N - p - q} \quad (5-36)$$

Debe tenerse en cuenta que todas las fórmulas anteriores solamente son aproximadas, ya que se han efectuado los cálculos bajo el supuesto de una aproximación lineal.

Con objeto de obtener un proceso de convergencia más rápido, se introducen ciertos refinamientos en el procedimiento que se ha explicado. En este sentido, el algoritmo de Marquardt (1963) es muy utilizado en los paquetes de ordenador.

Estimación por máxima verosimilitud de un modelo ARMA

Para efectuar la estimación maximoverosímil es preciso utilizar las hipótesis acerca de la distribución probabilística de ϵ_t .

La función de densidad de ϵ_t viene dada por:

$$f(\epsilon_t / \sigma_t^2, \beta, y^0, \epsilon^0) = (2\pi)^{-1/2} (\sigma_t^2)^{-1/2} e^{-\frac{1}{2\sigma_t^2} \epsilon_t^2} \quad (5-37)$$

$$= (2\pi)^{-1/2} (\sigma_t^2)^{-1/2} e^{-\frac{1}{2\sigma_t^2} (w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q})^2}$$

La función de densidad conjunta para $\epsilon_1, \dots, \epsilon_N$ será igual al producto de las funciones de densidad marginales, ya que estas variables por hipótesis están incorrelacionadas entre sí. En consecuencia

$$f(\epsilon_1 \dots \epsilon_N / \sigma_t^2, \beta, y^0, \epsilon^0) = (2\pi)^{-N/2} (\sigma_t^2)^{-N/2} e^{-\frac{1}{2\sigma_t^2} \sum \epsilon_t^2} \quad (5-38)$$

$$= (2\pi)^{-N/2} (\sigma_t^2)^{-N/2} e^{-\frac{1}{2\sigma_t^2} \sum (w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q})^2}$$

La maximización de la anterior función respecto a β conduce a la obtención de unos estimadores maximoverosímiles. Ahora bien, maximizar la anterior función equivale a minimizar el sumatorio que aparece en el exponencial, y por tanto coincide

con los estimadores mínimos cuadráticos. Obsérvese también que (5-38) puede contemplarse también como la función de densidad conjunta del vector w .

5.3. Enfoque no condicional

El enfoque no condicional se caracteriza porque los valores iniciales no se consideran como datos.

En lugar de minimizar $\sum \varepsilon_t^2$, en este enfoque se minimizan la suma de los cuadrados de las esperanzas condicionales de los ruidos dada la muestra $w_1 \dots w_N$. Considerando que el proceso se inicia en un pasado remoto la suma a minimizar sería la siguiente:

$$S = \sum_{t=-\infty}^N E_c[(\varepsilon_t)^2] \quad (5-39)$$

En un proceso estacionario, que es el caso que se está considerando aquí, las esperanzas condicionadas a la muestra de w' se hacen pequeñas a medida que nos alejamos del período muestral. En un proceso MA(q) se verifica $E_c(\varepsilon_t) = 0$ para $t < -q+1$, ya que, al estar incorrelacionada entre sí las ε_i , únicamente $\varepsilon_{-q+1} \dots \varepsilon_0$, $\varepsilon_1 \dots \varepsilon_N$ han intervenido en la generación de la muestra $w_1 \dots w_N$. En general se puede considerar la aproximación de que para $t < -Q+1$, siendo Q un entero positivo suficientemente grande, se verifica que $E_c(\varepsilon_t) = 0$. Por tanto, la suma a minimizar en la práctica sería:

$$S^* = \sum_{t=-Q+1}^N E_c[(\varepsilon_t)^2] \quad (5-40)$$

Una vez establecida la expresión a minimizar, se va a examinar la forma en que se calculan las esperanzas condicionadas. En la exposición se utiliza un modelo ARMA(1, 1). Un modelo de este tipo se expresa así:

$$w_t - \phi_1 w_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (5-41)$$

En la expresión anterior, los valores de w_t se determinan en función de los valores pasados de w_t , y de ε_t y de valores pasados de ε_t . Es decir, los valores sucesivos de w_t se van determinando «hacia adelante» en función de valores pasados. En un proceso estacionario e invertible también cabe imaginar una determinación en sentido inverso, es decir, «hacia atrás», con lo que se tendría que

$$w_t - \phi w_{t+1} = \eta_t - \theta_1 \eta_{t+1} \quad (5-42)$$

donde η_t es una variable de ruido blanco con la misma varianza que ε_t .

Si se toman esperanzas condicionadas en (5-41) y en (5-42) se obtienen las siguientes expresiones:

$$E_c[w_t] - \phi_1 E_c[w_{t-1}] = E_c[\varepsilon_t] - \theta_1 E_c[\varepsilon_{t-1}] \quad (5-43)$$

$$E_c[w_t] - \phi_1 E_c[w_{t+1}] = E_c[\eta_t] - \theta_1 E_c[\eta_{t+1}] \quad (5-44)$$

Las esperanzas condicionadas anteriores se calculan tomando como condición a los datos de la muestra, es decir, a w_1, w_2, \dots, w_N .

Por tanto se verifica que

$$E_c(w_t) = w_t \quad t = 1, \dots, N \quad (5-45)$$

puesto que la esperanza condicionada de w en el momento t será justamente el valor que ha salido en la muestra para ese período.

Por otra parte, también se verifica que

$$E_c[\varepsilon_{N+j}] = E[\varepsilon_{N+j}] = 0 \quad j = 1, 2, \dots \quad (5-46)$$

ya que los valores futuros que toma el ruido ε_t no vienen condicionados por la muestra dada en un proceso «hacia adelante».

Análogamente si se contempla un proceso «hacia atrás» se verifica que

$$E_c[\eta_{t-j}] = E[\eta_{t-j}] = 0 \quad j = 1, 2, \dots \quad (5-47)$$

ya que en este caso los ruidos η_0, η_{-1}, \dots serán independientes de la muestra dada de w' .

Si se hacen supuestos acerca de los valores finales —de la misma forma que se hacen para los valores iniciales en el enfoque condicionado— dados unos valores de los parámetros y de la muestra $w_1 \dots w_N$, se puede calcular de forma recursiva «hacia atrás» la secuencia $E_c(\eta_N), E_c(\eta_{N-1}) \dots$

Así, para $t = N$

$$E_c[w_N] - \phi_1 E_c[w_{N+1}] = E_c[\eta_N] - \theta_1 E_c[\eta_{N+1}] \quad (5-48)$$

Si se hace el supuesto de que $E_c(w_{N+1}) = 0$ y de que $E_c(\eta_{N+1}) = 0$ se obtiene, se tiene que

$$E_c(\eta_N) = w_N$$

Para $t = N-1$, se obtiene

$$E_c(\eta_{N-1}) = w_{N-1} - \phi_1 w_N + \theta_1 E_c(\eta_N) = w_{N-1} - \phi_1 w_N + \theta_1 w_N \quad (5-49)$$

De forma análoga se calcularían $E_c(\eta_{N-2}), E_c(\eta_{N-3}) \dots$ hasta llegar a

$$E_c[\eta_1] = w_1 - \phi_1 w_2 + \theta_1 E_c[\eta_2] \quad (5-50)$$

Ahora bien, de acuerdo con (5-47), $E_c[\eta_0]$, $E_c[\eta_{-1}]$ son igual a 0. Continuando el cálculo «hacia atrás», mediante la aplicación de la fórmula (5-44) se obtendrán $E_c(w_0)$, $E_c(w_{-1})$, de la siguiente forma:

$$E_c(w_0) = \phi_1 w_1 + E_c[\eta_0] - \theta_1 E_c[\eta_1] = \phi_1 w_1 - \theta_1 E_c[\eta_1]$$

$$E_c(w_{-1}) = \phi_1 E_c(w_0)$$

$$E_c(w_{-2}) = \phi_1 E_c(w_{-1})$$

⋮

Los valores precedentes se irán atenuando, ya que no existe realimentación del ruido.

Una vez que se llega a valores cercanos a 0 se puede iniciar el cálculo hacia adelante que permite obtener los valores $E_c(\epsilon_t)$ en función de los parámetros de la muestra, y de los valores $E_c(w_t)$ obtenidos precisamente en el proceso recursivo «hacia atrás». En consecuencia, se puede evaluar S^* de (5-40) para valores dados de los parámetros.

En la estimación por mínimos cuadrados se evaluará (5-40) para unos valores iniciales β^0 . El proceso iterativo subsiguiente se efectúa de forma análoga que en el enfoque condicional con la diferencia de que ahora los valores iniciales no se consideran dados, sino que se calculan de forma recursiva «hacia atrás» a partir de la predicción hacia atrás— para aludir al enfoque no condicional.

Cuando se aplica el principio de máxima verosimilitud con el enfoque no condicional, se obtienen las funciones de verosimilitud exactas.

En el enfoque no condicional los estimadores de máxima verosimilitud coinciden con los estimadores minimocuadráticos; en el enfoque condicional, en cambio no se da esta igualdad. La función de verosimilitud de σ_ϵ^2 y β , dado el vector muestral w es del siguiente tipo:

$$L(\beta, \sigma_\epsilon^2 | w) = g_1(\beta, \sigma_\epsilon^2) e^{-\frac{1}{2\sigma_\epsilon^2} \sum E_c(\epsilon_i)^2} \quad (5-51)$$

donde g_1 es una función de los parámetros β y σ_ϵ^2 .

La maximización de la función (5-51) presenta problemas de varios tipos. Para un tratamiento de este tema puede verse el capítulo 7 de la obra de Box-Jenkins (1976), así como las aportaciones entre otros de Newbold (1974), Dent (1977) y Ansley (1979).

Ejercicios

[1] Sea la serie Y_t

t	1	2	3	4	5	6	7	8	9	10
Y_t	1,40	-0,65	-1,46	-2,53	-1,25	-1,88	-1,00	-1,12	0,55	-1,16

Teniendo en cuenta que ha sido generada por el proceso

$$Y_t = 0,8 Y_{t-1} + \epsilon_t$$

- Calcular la serie ϵ_t en un enfoque condicionado, tomando como valor inicial $\epsilon_0 = 0$.
- Calcular la serie ϵ_t en un enfoque condicionado, tomando como valor inicial $\epsilon_1 = 1,40$.
- Calcular la serie ϵ_t en un enfoque no condicionado.
- Comparar los resultados obtenidos en apartados anteriores.

[2] Sea la serie Y_t

t	0	1	2	3	4	5	6	7	8	9	10
Y_t	102,90	101,45	100,93	100,33	98,60	98,75	98,54	96,71	97,68	98,24	98,34

Teniendo en cuenta que ha sido generada por el proceso

$$\Delta Y_t = \epsilon_t - 0,6\epsilon_{t-1}$$

- Calcular la serie transformada $w_t = \Delta Y_t$.
- Calcular la serie ϵ_t en un enfoque condicionado.
- Calcular la serie ϵ_t en un enfoque no condicionado.
- Comparar los resultados obtenidos en los apartados anteriores.

[3] Sea la serie Y_t

t	1	2	3	4	5	6	7	8	9	10
Y_t	2,66	-1,26	-0,23	1,18	-1,30	-0,64	0,17	-0,98	0,12	2,24

Teniendo en cuenta que ha sido generada por el proceso:

$$Y_t = 0,8 Y_{t-1} + \epsilon_t - 0,6\epsilon_{t-1}$$

- Calcular la serie ϵ_t en un enfoque condicionado, tomando como valor inicial $\epsilon_0 = 0$.

- b) Calcular la serie ε_t en un enfoque no condicionado, tomando como valor inicial $\varepsilon_1 = 2,66$.
- c) Calcular la serie ε_t en un enfoque no condicionado.
- d) Componer los resultados obtenidos en apartados anteriores.

4 Sea el proceso

$$Y_t = \varepsilon_t - 0,6\varepsilon_{t-1}$$

y las observaciones

t	1	2	3	4	5
Y_t	3	2	0	2	1

Se pide:

- a) Calcular $S = \sum \varepsilon_t^2$ condicionada al valor de ε_0 .
- b) Calcular el valor de ε_0 que minimice S .

5 Sea la serie Y_t

t	1	2	3	4	5	6	7	8	9	10
Y_t	1,55	1,30	-1,06	-0,88	1,35	1,07	0,30	1,00	3,08	1,26

Suponga que se ha identificado —aunque con tan pocas observaciones sería difícil hacerlo— que estas observaciones de Y_t se han generado por un proceso AR(1).

Se pide:

- a) Estimar el modelo por el método de mínimos cuadrados condicionados indicando la suma de cuadrados de los residuos y la estimación del parámetro ϕ_1 obtenidas en cada iteración.
- b) Estimar el modelo por mínimos cuadrados utilizando «backforecasting», e indicando resultados parciales como en el apartado a).
- c) Calcular y representar gráficamente la suma de los cuadrados de los residuos condicionada obtenida para valores de $\phi_1 = -1,0, -0,9, -0,8, \dots, 0,8, 0,9, 1,0$.

6 Sea la serie Y_t

t	1	2	3	4	5	6	7	8	9	10
Y_t	0,60	0,72	-1,36	-0,34	-1,88	-1,20	-0,50	-2,05	-2,64	2,63

Suponga que se ha identificado que estas observaciones de Y_t se han generado por un proceso MA(1).

Se pide:

- a) Estimar el modelo por el método de mínimos cuadrados condicionados indicando la suma de cuadrados de los residuos y la estimación del parámetro θ_1 obtenidas en cada iteración.

- b) Estimar el modelo por mínimos cuadrados utilizando «backforecasting», e indicando resultados parciales como en el apartado a).
- c) Calcular y representar gráficamente la suma de los cuadrados de los residuos condicionada obtenida para valores de θ_1 variando desde -1,0 a 1,0 con incrementos de 0,1.

7 Sea la serie Y_t

t	1	2	3	4	5	6	7	8	9	10
Y_t	-0,82	-0,46	-0,49	0,31	-0,56	-1,84	-1,92	-1,79	-0,76	-1,74

Suponga que se ha identificado que estas observaciones de Y_t se han generado por un proceso ARMA(1, 1).

Se pide:

- a) Estimar el modelo por el método de mínimos cuadrados condicionados indicando la suma de cuadrados de los residuos y la estimación de los parámetros ϕ_1 y θ_1 obtenidas en cada iteración.
- b) Estimar el modelo por mínimos cuadrados utilizando «backforecasting», e indicando resultados parciales como en el apartado a).

8 Estime utilizando los 100 primeros datos de la serie C del apéndice de datos el modelo —o modelos alternativos— que haya identificado en el ejercicio 4.14.

9 Estime utilizando los 200 primeros datos de la serie D del apéndice de datos el modelo —o modelos alternativos— que haya identificado en el ejercicio 4.15.

10 Estime utilizando los 100 primeros datos de la serie E del apéndice de datos el modelo —o modelos alternativos— que haya identificado en el ejercicio 4.16.

CAPÍTULO 6

Validación

6.1. Introducción

El objetivo perseguido al elaborar un modelo ARIMA es encontrar un modelo que sea lo más adecuado posible para representar el comportamiento de la serie estudiada. Así, un modelo ideal sería el que cumpliera los siguientes requisitos:

- Los residuos del modelo estimado se aproximan al comportamiento de un ruido blanco.
- El modelo estimado es estacionario e invertible.
- Los coeficientes son estadísticamente significativos, y están poco correlacionados entre sí.
- Los coeficientes del modelo son suficientes para representar la serie.
- El grado de ajuste es elevado en comparación al de otros modelos alternativos.

La finalidad de la fase de validación, que se examina en este capítulo, consiste precisamente en analizar la adecuación entre el modelo y los datos, o, dicho de otra forma, en analizar en qué medida se cumplen los anteriores requisitos.

Conviene señalar que es esencial que se cumpla el requisito a), pues en caso contrario el modelo debe ser rechazado, ya que ello sería indicativo de que los residuos contienen información relevante para la predicción. En el epígrafe 6.2, se desarrollan procedimientos de análisis de los residuos que tienen por objeto determinar si dichos residuos se aproximan a un ruido blanco.

En el epígrafe 6.3 se procede a efectuar un análisis de los coeficientes estimados, la correlación existente entre ellos, así como el examen de las condiciones de estacionariedad e invertibilidad.

Con objeto de determinar si los coeficientes del modelo son suficientes para representar la serie, en el epígrafe 6.4 se examina el efecto que produce la introducción de parámetros adicionales en el modelo. El epígrafe 6.5 está dedicado al análisis de la bondad del ajuste del modelo en comparación con otros posibles modelos alternativos.

En el epígrafe 6.6 se plantea el análisis de estabilidad estructural del modelo, mediante la utilización del test de Chow.

Si el modelo estimado superase satisfactoriamente las etapas del proceso de va-

lidación que acabamos de describir, se estaría en condiciones de utilizarlo en la predicción de valores futuros de la variable. En caso contrario sería necesario proceder a la reformulación del modelo. Esta última cuestión será abordada en el epígrafe 6.7.

6.2. Análisis de los residuos

En el proceso de identificación se habrá formulado un modelo ARIMA del tipo:

$$\phi(L)(1-L)^d(Y_t^{(1)} - \mu) = \theta(L)\epsilon_t \quad (6-1)$$

Si en la expresión anterior se despeja ϵ_t se tiene que

$$\epsilon_t = \theta(L)^{-1} \phi(L)(1-L)^d(Y_t^{(1)} - \mu) \quad (6-2)$$

Si el modelo especificado es el correcto, y se conocieran los parámetros verdaderos y los valores iniciales, al dar valores en la expresión anterior a Y_1, Y_2, \dots, Y_T se obtiene una serie temporal de «ruido blanco» $\epsilon_1, \dots, \epsilon_T$. Según ha demostrado Anderson (1942) los coeficientes de autocorrelación muestrales procedentes de un proceso de ruido blanco se distribuyen, en muestras grandes, de la siguiente forma

$$r_t \sim N[0, 1/T] \quad \forall t \quad (6-3)$$

En consecuencia cabe esperar, bajo la hipótesis nula de que ϵ_t es un ruido blanco, que

$$|r_t \sqrt{T}| < 1,96$$

en un 95 % de los casos.

El problema que se plantea es que no conocemos ϵ_t .

Al sustituir en el segundo miembro de (6-2) los parámetros por sus correspondientes estimaciones, en el primer miembro se obtienen los residuos $\hat{\epsilon}_t$.

$$\hat{\epsilon}_t = [\hat{\theta}(L)]^{-1} [\hat{\phi}(L)][1-L]^d [Y_t^{(1)} - \hat{\mu}] \quad (6-4)$$

Obviamente, los análisis deben hacerse a partir de los residuos. Si el comportamiento de los residuos se asemeja al de una serie ruido blanco, existirá una adecuación entre el modelo identificado y los datos.

Contrastes individuales de los coeficientes

Para analizar la independencia de los residuos se utiliza la FACE y la FACPE. Lógicamente para que una serie de residuos se asemeje a una serie ruido blanco, los coeficientes de la FACE y de la FACPE no deben ser significativamente distintos de 0. El estudio puede hacerse bien analizando los coeficientes individuales, o bien mediante un contraste global.

Para el análisis de los coeficientes individuales de la FACE y de la FACPE, en la práctica se construyen bandas de confianza utilizando la distribución de una variable ruido blanco cuya varianza aproximada, según (6-3), viene dada por $1/T$. Sin embargo, este valor no constituye una buena aproximación ya que la distribución (6-3) no es aplicable a los coeficientes de autocorrelación de una serie de residuos, especialmente en los retardos de orden bajo. Así, cuando se ha estimado un AR(1), una aproximación más adecuada de la varianza de los coeficientes de autocorrelación es la siguiente:

$$\text{var}(r_k) \simeq \frac{1}{T} [1 - \phi_1^{2k-2} (1 - \phi_1^2)] \quad (6-5)$$

La diferencia entre esta aproximación y $1/T$ puede ser muy elevada en los coeficientes de autocorrelación de orden bajo. Así para $k = 1$

$$\text{var}(r_k) \simeq \frac{\phi_1^2}{T} \quad (6-6)$$

Considerando un valor hipotético de $\phi_1 = 0,7$, resulta que la fórmula anterior da un valor de $0,49/T$, que es aproximadamente la mitad de la varianza dada en (6-3).

Aunque la fórmula (6-5) es solamente adecuada para un proceso AR(1), en general puede considerarse que si se utiliza el valor $1/T$ se sobrevalora la varianza de las autocorrelaciones de los residuos, especialmente en los órdenes bajos de retardos. De esta forma, se llegaría a concluir que un coeficiente r_τ es estadísticamente no significativo cuando en realidad lo es. Por ejemplo, suponiendo que $r_1 = 0,15$ y que $T = 100$ y tomando como valor de la varianza $1/T$, se obtiene que

$$|r_1\sqrt{T}| = 1,5 < 1,96$$

con lo cual se establecería que r_1 no es estadísticamente significativo al 95 %, contrariamente a lo que ocurriría si la fórmula adecuada para aproximación de la varianza es (6-6) y $|\phi_1| < 0,77$.

Aunque en general no será factible calcular una aproximación adecuada para la varianza de r_τ , conviene tener presente este fenómeno de sobreestimación cuando se está analizando la significatividad de los coeficientes de autocorrelación, especialmente los relativos a los retardos de orden bajo. Pankratz (1983) considera que si $|r_\tau\sqrt{T}| > 1,25$ para $\tau = 1, 2$ y 3 se ha superado el umbral práctico de admisibilidad de la hipótesis nula; para los demás retardos fija este umbral en 1,60.

Contraste global

Para observaciones procedentes de una serie ruido blanco, como contraste global —o contraste de «portmanteau»— se utiliza el siguiente estadístico propuesto por Box y Pierce (1970)

$$Q = T \sum_{\tau=1}^M r_\tau^2 \quad (6-7)$$

que se distribuye asintóticamente como una χ^2_M con M grados de libertad bajo la hipótesis nula de que

$$H_0: \rho_1 = \rho_2 = \dots = \rho_M = 0 \quad (6-8)$$

La elección de M , es decir, del número de coeficientes de autocorrelación que se incluyen en el sumatorio, es arbitrario. Tomando M elevado se tiene la ventaja de que puede captar los r_τ significativos correspondientes a retardos elevados, pero con la contrapartida de que a medida que aumenta M disminuye la potencia del contraste.

La aplicación del estadístico Q a los residuos plantea los mismos problemas que el contraste de significatividad de un solo coeficiente. De todas formas Box y Pierce han demostrado que utilizando un M razonablemente elevado, el estadístico Q construido a partir de los residuos de un modelo ARMA es apropiado si se supone una distribución χ^2 con $M-p-q$ grados de libertad bajo la hipótesis nula de que los residuos son independientes. De esta forma al reducir los grados de libertad en razón al número de parámetros ajustados, la probabilidad de rechazar la hipótesis nula aumenta. El comportamiento de este estadístico en pequeñas muestras ha sido estudiado por Davies y otros (1977).

El estadístico Q de Box-Pierce ha sido refinado posteriormente, a fin de disminuir el sesgo en pequeñas muestras, por Ljung y Box (1978) que han propuesto el siguiente estadístico de «portmanteau»

$$Q^* = T(T+2) \sum_{\tau=1}^M (T-\tau)^{-1} r_\tau^2 \quad (6-9)$$

que se distribuye como una χ^2 con $M-p-q$ grados de libertad bajo la hipótesis nula de que los residuos son independientes.

Un problema: determinación del nivel de significación

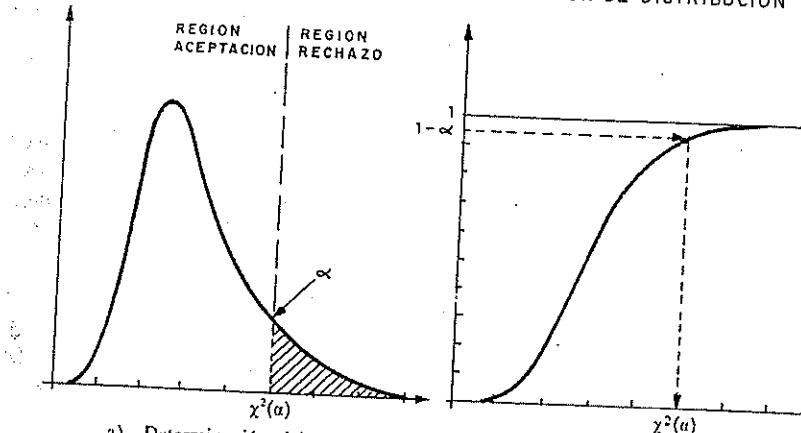
En los contrastes de hipótesis generalmente se fija de antemano un nivel de significación α —por ejemplo, $\alpha = 5\%$ —, en base al cual se busca el valor de la variable, o cuantil, en las tablas de la distribución teórica del estadístico utilizado, tomando los grados de libertad que correspondan. Cuando este valor es inferior —en los contrastes de una cola— al valor del estadístico calculado se rechaza la hipótesis nula. Así en nuestro caso, designado por $\chi^2_{M-p-q}(\alpha)$ al valor χ^2 de las tablas para un nivel de significación α y $M-p-q$ grados de libertad se procedería de la siguiente forma:

- Si $Q^* < \chi^2_{M-p-q}(\alpha)$ se acepta H_0
- Si $Q^* > \chi^2_{M-p-q}(\alpha)$ se rechaza H_0

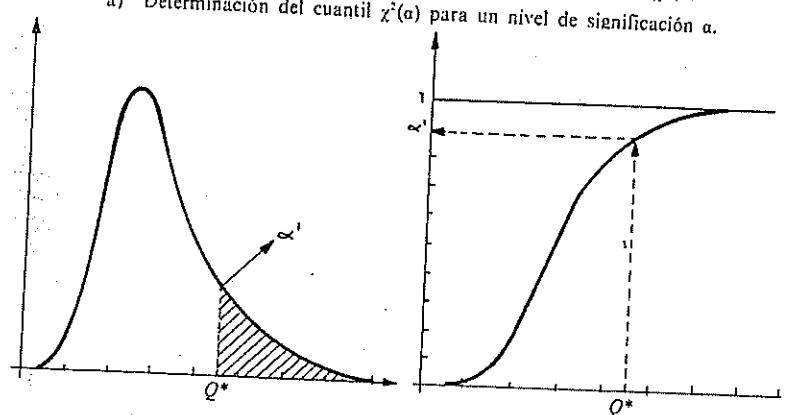
En la figura 6.1 a) se ilustra cómo a partir de un valor de α se determina el cuantil, o valor de la variable, que divide el espacio muestral en una zona de aceptación y una zona de rechazo de la hipótesis dada. Se ha representado, además de la función de densidad, que es la usual en este tipo de ilustraciones, la función de distribución que puede contribuir en este caso a clarificar la exposición.

VALIDACIÓN

FUNCION DE DENSIDAD

a) Determinación del cuantil $\chi^2(\alpha)$ para un nivel de significación α .

FUNCION DE DISTRIBUCION



b) Determinación del nivel de significación crítico (NSC).

Fig. 6.1. Formas alternativas de proceder en el contraste de hipótesis con una $\chi^2_{M-\alpha-\alpha}$.

En el proceso de contrastación que se ha expuesto, la parte más subjetiva sería la determinación a priori del nivel de significación. ¿En base a qué criterios se puede determinar α ? En general se trata de una decisión arbitraria, siendo $\alpha = 5\%$ y $\alpha = 1\%$ los niveles utilizados en la práctica. A veces se efectúa el contraste de hipótesis condicionando a distintos niveles de significación, aunque ésta pueda estar limitada por los

Con la utilización de los ordenadores, el contraste de hipótesis puede contemplarse desde otra perspectiva mucho más racional. Así, algunos programas de ordenador

ANÁLISIS DE LOS RESIDUOS

ofrecen, junto al estadístico muestral, el nivel de significación crítico (NSC) α' asociado al estadístico muestral. En nuestro caso α' se calcularía de forma que:

$$\text{Prob} [\chi^2 > Q^*] = \alpha' \quad (6-10)$$

En la figura 6.1. b) se ilustra el cálculo de α' a partir del estadístico muestral Q^* . Puede verse claramente observando el gráfico de la función de distribución que la operación de determinar α' es la inversa de buscar el valor de las tablas para un nivel de significación dado.

Determinado α' sabemos que se rechaza la hipótesis nula para todo nivel de significación α tal que $\alpha > \alpha'$; por el contrario, se acepta la hipótesis nula cuando $\alpha < \alpha'$.

El NSC es un indicador del nivel de admisibilidad de la hipótesis nula. Cuanto mayor sea el NSC, mayor confianza podemos depositar en la hipótesis nula.

La utilización del NSC implica dar la vuelta al problema del contraste de hipótesis: en lugar de fijar a priori un nivel de significación, se calcula un valor α' que permite determinar a posteriori para qué niveles de significación se puede rechazar la hipótesis nula.

Primera diferencia de los residuos

Para completar el análisis de independencia de los residuos algunos autores —por ejemplo, Vandaele (1983)— proponen tomar diferencias de primer orden en los residuos y calcular la FACE y la FACPE de la serie diferenciada.

Así, si la serie de residuos se aproxima a un ruido blanco:

$$\hat{\epsilon}_t \approx \epsilon_t$$

entonces

$$\Delta\hat{\epsilon}_t \approx \epsilon_t - \theta_1 \epsilon_{t-1} \quad \text{con } \theta_1 \approx 1$$

En este caso, la FACE se comportará¹ aproximadamente así:

$$r_1(\Delta\hat{\epsilon}_t) \approx -0,5$$

$$r_\tau(\Delta\hat{\epsilon}_t) \approx 0 \quad \text{para } \tau > 1$$

y la FACPE

$$\hat{o}_{11}(\Delta\hat{\epsilon}_t) \approx -0,5$$

$$\hat{o}_{22}(\Delta\hat{\epsilon}_t) \approx -0,333$$

$$\hat{o}_{33}(\Delta\hat{\epsilon}_t) \approx -0,250$$

.....

¹ En inglés, se utiliza la expresión P-value.

² Véase en los capítulos 3 y 4 respectivamente el comportamiento teórico de la FACT y FACPT de un MA(1).

En la medida que la FACE y la FACPE de cualquier serie de residuos diferenciados una vez, tengan un comportamiento semejante al que se acaba de ver, se tendrá una prueba adicional de que los residuos obtenidos en la estimación se aproximan a un ruido blanco.

Otros contrastes sobre los residuos

Los contrastes que se han examinado hasta ahora están diseñados para determinar si los residuos son o no independientes entre sí, pero no ofrecen información acerca de si la serie de residuos es o no heteroscedástica. Para esta finalidad es útil examinar el gráfico de los residuos donde puede apreciarse la evolución de la dispersión de los residuos a lo largo del tiempo. Si se advierte una tendencia creciente —o decreciente en su caso— en la evolución de dicha dispersión será indicativo de una posible heteroscedasticidad.

El gráfico de los residuos puede ser útil también para detectar la existencia de observaciones anómalas (*“outliers”* en terminología anglosajona). Pueden ser consideradas como tales las correspondientes a los residuos que superen en gran medida las bandas de confianza establecidas para los residuos. Así, si un residuo \hat{e}_t fuera tal que $|\hat{e}_t/\hat{\sigma}_{\hat{e}_t}| > 4$ sería indicativo de que puede existir un factor exógeno puntual que haya provocado tal anomalía. Box y Tiao (1975) han desarrollado procedimientos para tratar problemas de este tipo, que se engloban bajo la denominación genérica de análisis de intervención. Estos procedimientos van más allá del objetivo de la presente obra.

6.3. Análisis de los coeficientes estimados

En las salidas de ordenador de los programas ARIMA suele aparecer —junto al coeficiente estimado— la desviación típica, el estadístico t^* , y en algunos casos el NSC de cada coeficiente. Con estos elementos se pueden realizar distintos contrastes de hipótesis sobre los parámetros del modelo. El estadístico t^* está construido bajo la hipótesis nula de que el parámetro es igual a 0. Así para un coeficiente $-\phi_1$, por ejemplo, el estadístico t^* vendría dado por la siguiente expresión:

$$t^*_{N-p-q-\delta} = \frac{\hat{\phi}_1 - (\phi_1/H_0)}{\hat{\sigma}_{\hat{\phi}_1}} = \frac{\hat{\phi}_1}{\hat{\sigma}_{\hat{\phi}_1}} \quad (6-11)$$

donde $N-p-q-\delta$ son los grados de libertad y $\delta = 1$ si se incluye término independiente; en caso contrario, $\delta = 0$.

Conviene tener en cuenta no obstante que la distribución $|t|$ es una distribución exacta, y por tanto válida para pequeñas muestras, en el contexto del modelo lineal general. En los modelos ARIMA, en cambio, únicamente se pueden establecer propiedades asintóticas. Así, el segundo miembro de (6-11) tiene una distribución asintótica normal, aunque quizás en pequeñas muestras la $|t|$ siga siendo una buena aproximación. En cualquier caso los modelos ARIMA se aplican en general a muestras grandes, se pueden utilizar los niveles de significación de la normal. Así, a grosso modo se rechazaría la hipótesis nula de que $\phi_1 = 0$, para un nivel de significa-

ción del 5 % cuando $|t^*| > 2$. En cualquier caso, sigue siendo válido aquí lo que se dijo acerca del nivel de significación y del NSC al tratar del estadístico Q^* , con la única diferencia de que la aplicación del estadístico t^* requiere un contraste de dos colas, a no ser que la hipótesis alternativa fuera de tipo $H_1: \phi_1 > 0$ o $H_1: \phi_1 < 1$. Para establecer hipótesis alternativas de este último tipo se requiere una información adicional acerca de la naturaleza del parámetro; en los modelos ARIMA no se dispondrá en general de esta información adicional por lo que se adopta la hipótesis alternativa $H_1: \phi_1 \neq 0$.

Una cuestión importante en el análisis de los coeficientes es el examen de si se cumplen las condiciones de estacionariedad y de invertibilidad.

Para comprobar si el modelo estimado es estacionario, se procede a factorizar el polinomio de la parte autorregresiva, es decir,

$$(1 - \hat{\phi}_1 L - \dots - \hat{\phi}_p L^p) = (1 - \hat{\lambda}_1 L) \dots (1 - \hat{\lambda}_p L)$$

Si existe un coeficiente $\hat{\lambda}_i$ próximo a 1, es un indicio de que el modelo no es estacionario, siendo aconsejable en este caso tomar una diferencia adicional.

Por otra parte, para comprobar si el modelo estimado es invertible se efectúa la siguiente factorización:

$$(1 - \hat{\theta}_1 L - \dots - \hat{\theta}_q L^q) = (1 - \hat{\delta}_1 L) \dots (1 - \hat{\delta}_q L)$$

Cuando algún valor $\hat{\delta}_i$ esté próximo a 1, sería indicativo de que el modelo estimado es no invertible. En muchos casos, este problema se ha podido producir como consecuencia de una sobrediferenciación.

Para completar el análisis, es conveniente examinar la matriz de correlación entre los coeficientes estimados con objeto de detectar si existe un problema de multicolinealidad. Cuando la correlación entre dos coeficientes estimados esté próxima a 1 —por ejemplo, 0,9— sería un indicio suficiente para considerar que existe un problema grave de multicolinealidad. Si esto ocurre, los coeficientes estimados serían de carácter muy inestable con lo que podrían cambiar mucho de una realización a otra. Así, el modelo estimado para el periodo muestral puede diferir del que se obtendría para los períodos de predicción. Para evitar este problema puede ser conveniente eliminar algún parámetro del modelo con objeto de conseguir unos estimadores más estables aun a costa de que el grado de ajuste sea más pequeño. No obstante, si todos los coeficientes son significativos a pesar de que la multicolinealidad sea elevada, tampoco sería aconsejable eliminar coeficientes del modelo. En todo caso, la eliminación de parámetros no debe hacerse a expensas de que se pierda la independencia entre los residuos.

6.4. Introducción de parámetros adicionales en el modelo

El principio de la parametrización escueta implica que se utilice el mínimo número de parámetros posible para representar el hipotético proceso generador de la serie. Este principio se aplica naturalmente en la fase de identificación. Una vez estimados

los parámetros es aconsejable examinar si se puede mejorar la adecuación entre datos y modelo introduciendo parámetros adicionales. La introducción de parámetros adicionales estaría justificada si en el proceso de estimación los coeficientes resultaran estadísticamente significativos. En todo caso conviene no introducir parámetros de forma simultánea en la parte autorregresiva y en la parte de medias móviles con objeto de evitar posibles problemas de redundancia, según puede verse en el siguiente ejemplo.

Supóngase que para la serie Y_t se ha identificado inicialmente como un AR(1), y que este modelo es el correcto. Por tanto, la serie Y_t se puede representar mediante el

$$(1 - \phi_1 L) Y_t = \epsilon_t \quad (6-12)$$

Si se introduce un parámetro adicional tanto en la parte AR como en la parte MA se llegaría a un modelo ARMA(2, 1) que se representaría así:

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = (1 - \theta_1 L) \epsilon_t \quad (6-13)$$

Naturalmente, si el modelo (6-12) es el correcto, entonces el modelo (6-13) estaría claramente sobreparametrizado, con el inconveniente adicional de que se plantean problemas en el proceso de estimación, ya que una de las raíces del polinomio AR coincidiría teóricamente con $1/\theta_1$.

6.5. Bondad del ajuste

En la fase de identificación puede ocurrir que se especifiquen varios modelos alternativos y que, una vez estimados, cada uno de ellos satisfaga los contrastes examinados en los epígrafes anteriores. En ese caso, en la fase de validación se tendrá que seleccionar con arreglo a algún criterio de bondad del ajuste cuál es el más adecuado entre todos ellos.

Entre los criterios de bondad del ajuste más utilizados en econometría se encuentran los estadísticos R^2 y R^2 corregido (\bar{R}^2), cuyas fórmulas, aplicadas en el ajuste de un modelo ARMA(p, q) a la serie estacionaria $w_1 \dots w_N$ vienen dadas por las expresiones siguientes:

$$R^2 = 1 - \frac{\hat{\sigma}_\epsilon^2}{\hat{\sigma}_w^2} \quad (6-14)$$

$$\bar{R}^2 = 1 - \frac{\hat{\sigma}_\epsilon^2/(N-p-q)}{\hat{\sigma}_w^2/(N-1)} \quad (6-15)$$

donde

$$\hat{\sigma}_\epsilon^2 = \frac{\sum_{t=1}^N \hat{\epsilon}_t^2}{N} \quad \text{y} \quad \hat{\sigma}_w^2 = \frac{\sum (w_t - \bar{w})^2}{N}$$

La diferencia entre el R^2 y el \bar{R}^2 corregido se debe a que el segundo criterio penaliza la introducción de parámetros adicionales en el modelo. Por ello sería en principio más aconsejable utilizar el R^2 corregido ya que de alguna forma tiende a evitar la sobreparametrización.

Sin embargo, la aplicación de las fórmulas (6-14) y (6-15) en los modelos ARIMA presenta un inconveniente grave. En efecto, en el R^2 y en el \bar{R}^2 se compara la varianza del error ($\hat{\sigma}_\epsilon^2$) con la varianza del regresando ($\hat{\sigma}_w^2$), pero el valor de esta última varianza depende del número de diferencias que se hayan tomado. Por ello, estos estadísticos únicamente tendrían validez cuando se comparasen modelos en que se haya tomado idéntico número de diferencias. Para aclarar esta cuestión considérese el siguiente modelo teórico

$$(1 - \phi_1 L) Y_t = \epsilon_t \quad (6-16)$$

con ϕ_1 próximo a 1. Para este modelo se tendría que

$$\hat{\sigma}_\epsilon^2 = \frac{1}{1 - \phi_1^2} \sigma_\epsilon^2$$

con lo que R^2 cuadrado teórico vendría dado por:

$$R_{\text{teó}}^2 = 1 - \frac{\sigma_\epsilon^2}{\frac{\sigma_\epsilon^2}{1 - \phi_1^2}} = \phi_1^2 \approx 1 \quad (6-17)$$

de acuerdo con el supuesto adoptado de que ϕ_1 está próximo a 1.

Si se considera $w_t = (1 - L) Y_t$ y se sustituye en (6-16), se obtiene que

$$(1 - \phi_1 L) \frac{1}{1 - L} w_t = \epsilon_t \quad (6-18)$$

que de acuerdo con el supuesto adoptado se puede simplificar a

$$w_t \approx \epsilon_t \quad (6-19)$$

En este caso, el coeficiente de determinación teórico de w_n o ΔY_t , vendría dado por:

$$R_{\text{teó}}^2 = 1 - \frac{\sigma_\epsilon^2}{\sigma_w^2} = 0 \quad (6-20)$$

Es decir, si se utiliza el modelo sin diferenciar el ajuste sería perfecto, mientras que después de haber tomado una diferencia ocurre todo lo contrario. No son, pues, comparables el R^2 que se obtiene en uno y otro caso. En cambio, sí serían comparables la estimación de σ_ϵ^2 que se obtiene en ambos casos.

Dadas las limitaciones que presenta el R^2 que son extensibles al R^2 corregido

—excepto en lo que se refiere a la penalización por número de parámetros—, se han propuesto diversas medidas alternativas, entre las que debe destacarse por su aceptación el estadístico AIC (Akaike Information Criterion) formulado por Akaike (1974). El estadístico AIC del que existen diversas versiones está basado en la teoría de la información.

La formulación que se utilizará será la siguiente:

$$AIC = -2/N \ln L(y, \hat{\beta}) + \frac{2(p+q+d)}{N} \quad (6-21)$$

donde

$$(L(y, \hat{\beta}) = 2\pi\sigma_e^2)^{-N/2} e^{\frac{1}{2\sigma_e^2} \sum_{t=1}^N \epsilon_t^2} = (2\pi\sigma_e^2)^{-N/2} e^{-N/2} \quad (6-22)$$

es el valor que se obtiene al minimizar la función de verosimilitud, con respecto a los elementos del vector β y a σ_e^2 .

Como puede verse, en el segundo miembro de (6-21) aparecen dos términos: el primero, con signo negativo, que depende de $\hat{\sigma}_e^2$, y el segundo que depende del número de parámetros del modelo. Entre los parámetros se incluye también el número de diferencias que se toman. Véase a este respecto Harvey (1981).

El criterio de Akaike consiste en seleccionar aquel modelo para el que se obtenga un estadístico AIC más bajo. Este estadístico no presenta el inconveniente que presentaba el R^2 y el R^2 corregido, pues penaliza los modelos con mayor número de parámetros y además permite comparar modelos con transformaciones de Box-Cox diferentes. Vamos a examinar este último punto.

La relación entre la función de verosimilitud de (Y_1, \dots, Y_N) y la correspondiente a $(Y_1^{(1)}, \dots, Y_N^{(1)})$ vendrá dada por

$$L(y, \hat{\beta}) = L(y^{(1)}, \hat{\beta})/|J| \quad (6-23)$$

donde el jacobiano de la transformación $|J|$ viene dado por:

$$|J| = \left| \frac{\partial (Y_1^{(1)}, \dots, Y_N^{(1)})}{\partial (Y_1, \dots, Y_N)} \right| = \prod_{i=1}^N Y_i^{1-1} \quad (6-24)$$

teniendo en cuenta la definición de $Y_i^{(1)}$ dada en (3-118).

Si en (6-21) se introduce el valor de (6-19), se tendrá siempre definido el estadístico AIC en términos de las variables sin transformar, con lo que será posible la com-

³ El vector β , según se vio en el capítulo 5, está integrado por $(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$. Por otra parte:

$$\hat{\epsilon}_t = w_t - \hat{\phi}_1 w_{t-1} - \dots - \hat{\phi}_p w_{t-p} + \hat{\theta}_1 \hat{\epsilon}_{t-1} + \dots + \hat{\theta}_q \hat{\epsilon}_{t-q}$$

paración directa. De acuerdo con lo anterior, y teniendo en cuenta también (6-22), se obtiene la siguiente formulación

$$AIC = (2\pi\sigma_e^2 + 1) - 2(\lambda - 1)\sum \ln Y_t / N + 2(p + d + q) / N \quad (6-25)$$

Obsérvese que el segundo término es nulo en el caso de que no se efectúe transformación de Box-Cox o, en otras palabras, de que la transformación de Box-Cox seleccionada sea $\lambda = 1$.

6.6. Análisis de estabilidad

La construcción de un modelo ARIMA está motivada por su utilización para la predicción. Ahora bien, el modelo estimado para el período muestral, ¿sigue siendo válido para períodos futuros? Se plantea aquí un problema de estabilidad estructural. Naturalmente, a medida que se vayan efectuando nuevas observaciones, se irá comprobando si el modelo sigue siendo válido. Pero lo que importa es tener en el presente un indicador de si es admisible el supuesto de estabilidad.

Para esta finalidad se puede aplicar —como una aproximación— el test de Chow (1960). En este test se utiliza el siguiente estadístico:

$$F_{K, N-2K} = \frac{\left[\sum_{t=1}^N \hat{\epsilon}_t^2 - \left(\sum_{t=1}^{N_1} \hat{\epsilon}_{1t}^2 + \sum_{t=N_1+1}^N \hat{\epsilon}_{2t}^2 \right) \right] / K}{\left(\sum_{t=1}^{N_1} \hat{\epsilon}_{1t}^2 + \sum_{t=N_1+1}^N \hat{\epsilon}_{2t}^2 \right) / (N - 2K)} \quad (6-26)$$

$$N = N_1 + N_2 \quad K = p + q + \delta$$

siendo:

$\hat{\epsilon}_t$ = residuo del modelo utilizando en la estimación todo el período muestral (N).

$\hat{\epsilon}_{1t}$ = residuo del modelo utilizando en la estimación los N_1 primeros datos.

$\hat{\epsilon}_{2t}$ = residuo del modelo utilizando en la estimación los N_2 últimos datos.

Algunos autores aconsejan tomar como N_2 un tercio o un cuarto de la muestra, de forma que pueda contrastarse si el último tramo muestral (tomado más pequeño) ha estado generado por la misma estructura que el resto de las observaciones.

6.7. Reformulación del modelo

Si después de aplicar los contrastes y análisis de los epígrafes anteriores se llega a la conclusión de que el modelo seleccionado⁴ no es adecuado, se debe proceder a reformular el modelo. Para ello, junto a los estadísticos de la fase de identificación, es conveniente tener en cuenta los estadísticos y resultados obtenidos en esta fase de va-

⁴ Puede ocurrir que se hayan seleccionado varios modelos alternativos.

VALIDACIÓN

lidadación, puesto que pueden arrojar luz sobre la dirección en que debe reformularse el modelo. Ya se ha visto, por ejemplo, que cuando una raíz de la ecuación característica de la parte autorregresiva está próxima a la unidad sería aconsejable tomar una diferencia adicional. Por el contrario, cuando una raíz de la ecuación característica de la parte de medias móviles está próxima a 1, entonces puede ser indicativo de que se han tomado más diferencias de las necesarias.

La FACE y la FACPE de los residuos del modelo estimado son instrumentos valiosos a la hora de reformular el modelo. A efectos ilustrativos, supóngase que se ha estimado un modelo AR(1)

$$(1 - \hat{\phi}_1 L) Y_t = \hat{\epsilon}_t \quad (6-27)$$

Después de examinar la FACE y la FACPE de la serie $\hat{\epsilon}_t$, se llega a la conclusión de que sigue un MA(1), es decir,

$$\hat{\epsilon}_t = (1 - \theta_1 L) \epsilon_t \quad (6-28)$$

Si se sustituye la expresión anterior en (6-27) se tendría que

$$(1 - \hat{\phi}_1 L) Y_t = (1 - \hat{\theta}_1 L) \epsilon_t \quad (6-29)$$

con lo que podría concluirse que la serie Y_t sigue un modelo ARMA(1, 1). Si los residuos $\hat{\epsilon}_t$ siguieran un AR(1), es decir,

$$(1 - \phi_1^* L) \hat{\epsilon}_t = \epsilon_t \quad (6-30)$$

la sustitución $\hat{\epsilon}$ en el modelo estimado conduciría a

$$(1 - \hat{\phi}_1 L) Y_t = (1 - \phi_1^* L)^{-1} \epsilon_t$$

es decir, a

$$(1 - \hat{\phi}_1 L)(1 - \phi_1^* L) Y_t = \epsilon_t \quad (6-31)$$

lo que llevaría a la reformulación de un modelo AR(2).

No obstante conviene advertir que en muchas ocasiones el examen de FACE y de la FACPE de los residuos no proporciona una pista clara de cuál debe ser la acción a tomar.

EJERCICIOS

Ejercicios

- [1] A partir de una muestra de 100 datos se ha ajustado el siguiente modelo:

$$\Delta Y_t = (1 - 0,324L) \hat{\epsilon}_t$$

La FACE Y la FACPE de los residuos $\hat{\epsilon}_t$ vienen dadas por:

Retardo (t) →	1	2	3	4	5
r_t	-0,508	0,013	-0,032	0,015	0,043
$\hat{\theta}_{tt}$	-0,508	-0,329	-0,220	-0,143	-0,093

A la vista de la información anterior, ¿consideraría aceptable el modelo ajustado? En caso contrario, ¿cuál podría ser la reformulación del modelo?

- [2] Sea el siguiente modelo ajustado con una muestra de 64 observaciones

$$(1 - 0,8L) Y_t = (1 + 0,6L) \hat{\epsilon}_t$$

La FACE y la FACPE de los residuos vienen dadas por:

Retardo (t) →	1	2	3	4	5
r_t	0,380	0,160	0,054	0,021	-0,004
$\hat{\theta}_{tt}$	0,380	0,030	-0,033	0,025	0,026

A la vista de la información anterior, ¿consideraría aceptable el modelo ajustado? En caso contrario, ¿cuál podría ser la reformulación del modelo?

- [3] Una serie Y_t , conteniendo 300 términos, se considera que ha sido generada por un modelo ARMA(0, 1), sin término independiente. En la estimación de este modelo se han obtenido los siguientes resultados para distintos tamaños de muestra.

Muestra utilizada		Suma de cuadrados de los residuos
Período inicial	Período final	
1	200	2,354
201	300	3,824
1	300	6,513

Las observaciones 201-300, ¿han estado generadas por la misma estructura que el resto de las observaciones?

- 4) Sea Y_t una serie de tamaño N generada por un proceso ARMA(1, 0). Estimando el parámetro ϕ_1 mediante regresión, pruebe que la estimación obtenida $\hat{\phi}_1$ es aproximadamente igual al coeficiente de determinación.
- 5) Con los datos de la serie A del apéndice de datos, a la que se designará por Y_t , se han estimado los siguientes modelos utilizando la muestra correspondiente al período 1900-1980.
- $(1 - L) Y_t = (1 - \theta_1 L) \epsilon_t$
 - $(1 - \phi_1 L)(1 - L) Y_t = (1 - \theta_1 L) \epsilon_t$

Para cada una de las anteriores especificaciones se ha obtenido los resultados que se detallan a continuación:

Especificación a)

Operador	Coeficiente estimado	Desviación típica	Estadístico t
θ_1	0,3290	0,1074	3,06

$$\sigma^2_{\epsilon} = 1,896 \quad R^2 = 0,107 \quad \bar{R}^2 = 0,107 \quad \text{AIC} = 12,962$$

Residuos en que se verifica que $|\hat{\epsilon}/\hat{\sigma}_{\epsilon}| > 2$

Año	1938	1939	1940
$\hat{\epsilon}/\hat{\sigma}_{\epsilon}$	-2,48	-3,41	4,63

Coeficientes de autocorrelación y autocorrelación parcial de los residuos.

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	-0,11	0,00	0,15	-0,05	0,02	-0,25	-0,08	0,12	-0,13	-0,06
$\hat{\phi}_{tt}$	-0,11	-0,01	0,15	-0,01	0,02	-0,28	-0,14	0,11	-0,01	-0,07
Retardo	11	12	13	14	15	16	17	18	19	20
r_t	-0,04	0,19	-0,05	-0,02	-0,01	-0,10	0,03	-0,09	-0,02	0,09
$\hat{\phi}_{tt}$	-0,11	0,16	-0,05	0,03	-0,09	-0,18	-0,10	0,01	0,09	0,05

Estadístico Ljung-Box

$$Q^*(5) = 3,1; \quad Q^*(10) = 13; \quad Q^*(15) = 17; \quad Q^*(20) = 20$$

Entre paréntesis se indica el número de coeficientes de autocorrelación que se han tomado para el cálculo de Q^* .

Especificación b)

Operador	Coeficiente estimado	Desviación típica	Estadístico t
ϕ_1	0,1759	0,1074	0,56
θ_1	-0,2009	0,3065	-0,66

$$\hat{\sigma}^2_{\epsilon} = 1,918 \quad R^2 = 0,108 \quad \bar{R}^2 = 0,097 \quad \text{AIC} = 13,125$$

Estimación de correlación de los coeficientes estimados:

Operador	ϕ_1	θ_1
ϕ_1	1,000	0,932
θ_1	0,932	1,000

Residuos en que se verifica que $|\hat{\epsilon}/\hat{\sigma}_{\epsilon}| > 2$

Período	1938	1939	1940
$\hat{\epsilon}/\hat{\sigma}_{\epsilon}$	-2,49	-3,38	4,60

Coeficientes de autocorrelación y autocorrelación parcial de los residuos.

Retardo	1	2	3	4	5	6	7	8	9	10
r_t	-0,07	-0,05	0,15	-0,02	0,01	-0,27	-0,08	0,13	-0,12	-0,08
$\hat{\phi}_{tt}$	-0,07	-0,05	0,14	-0,01	0,02	-0,29	-0,12	0,11	-0,02	-0,07
Retardo	11	12	13	14	15	16	17	18	19	20
r_t	-0,02	0,20	-0,04	-0,02	0,01	-0,10	-0,04	-0,09	0,02	0,09
$\hat{\phi}_{tt}$	-0,10	0,16	-0,05	0,05	-0,10	-0,18	-0,10	0,01	0,09	0,04

Estadístico Ljung-Box

$$Q^*(5) = 2,5; \quad Q^*(10) = 13; \quad Q^*(15) = 17; \quad Q^*(20) = 20$$

Teniendo en cuenta los resultados anteriores, explique cuál de los dos modelos estimados le parece más adecuado para representar la serie Y_t . En el caso de rechazar ambos modelos proponga un modelo o modelos alternativos. ¿Qué explicación daría a los valores anómalos de los residuos que aparecen en ambas especificaciones?

VALIDACIÓN

- 6** Con los datos de la serie E del apéndice de datos, a la que se designará por Y_t , y utilizando la muestra correspondiente al período 1940-1980, se pide:

a) Estimar los siguientes modelos:

$$(1 - \phi_1 L)(1 - L) Y_t = \delta + \epsilon_t$$

$$(1 - \phi_1 L)(1 - L) \ln Y_t = \delta + \epsilon_t$$

$$(1 - \phi_1 L)(1 - L) Y_t = \delta + (1 - \theta_1 L)\epsilon_t$$

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3)(1 - L)^2 Y_t = \epsilon_t$$

$$(1 - L)^2 Y_t = (1 - \theta_1 L)\epsilon_t$$

- b) A la vista de los resultados que obtenga en el apartado anterior, seleccione el modelo o modelos más adecuados para representar la serie, o, en otro caso, proponga especificaciones alternativas a las indicadas.

- 7** Analice si el modelo estimado en el ejercicio 8 del capítulo 5 es adecuado para representar la serie C del apéndice de datos. En el caso que haya estimado varios modelos alternativos seleccione el más adecuado en caso de que lo hubiese.

- 8** Analice si el modelo estimado en el ejercicio 9 del capítulo 5 es adecuado para representar la serie D del apéndice de datos. En el caso que haya estimado varios modelos alternativos seleccione el más adecuado en caso de que lo hubiese.

- 9** Analice si el modelo estimado en el ejercicio 10 del capítulo 5 es adecuado para representar la serie E del apéndice de datos. En el caso que se haya estimado varios modelos alternativos seleccione el más adecuado en caso de que lo hubiese.

CAPÍTULO 7

Predicción

7.1. Introducción

Las tres primeras fases de la elaboración de un modelo ARIMA constituyen un proceso iterativo cuyo resultado final es la obtención de un modelo estimado que sea compatible con la estructura de los datos. Una vez que se ha conseguido este resultado, la fase siguiente consiste en utilizar este modelo estimado en la predicción de valores futuros de la variable objeto de estudio.

En la exposición de la fase de predicción se plantearán en primer lugar distintas formas de presentar un modelo ARIMA que serán útiles para deducciones posteriores. Como segundo punto se discutirá la obtención teórica de un predictor óptimo. En el epígrafe 7.4 se calcula de forma operativa el valor del predictor, aplicándose a distintos modelos. Finalmente, en el epígrafe 7.5 se formula un contraste de estabilidad estructural para la predicción.

7.2. Formas alternativas de presentar un modelo ARIMA

Un modelo ARIMA(p, d, q) de la variable Y_t viene dado por la siguiente expresión:

$$\phi_p(L)/\Delta^d Y_t = \theta_q(L)\epsilon_t \quad (7-1)$$

donde los subíndices p y q indican el orden del correspondiente operador polinomial de retardos.

Para simplificar, y sin pérdida de generalidad, se ha supuesto que Y_t ha sido obtenida como resultado de una transformación del tipo Box-Cox a una serie original con media nula. En la notación tampoco se añadirá ningún signo especial a los parámetros para señalar que se trata de estimadores.

Multiplicando los dos operadores que aparecen en el primer miembro de (7-1), se obtiene:

$$\begin{aligned}\phi_p(L)\Delta^d &= (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) \underbrace{(1 - L)(1 - L^2) \dots (1 - L^d)}_{d \text{ veces}} \\ &= 1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_{p+d} L^{p+d}\end{aligned}\quad (7-2)$$

A los coeficientes de $\varphi_{p+d}(L)$ —o simplemente $\varphi(L)$ — se les denomina coeficientes autorregresivos generalizados. De acuerdo con (7-2), el modelo (7-1) se puede expresar así

$$\varphi(L)Y_t = \theta(L)\varepsilon_t \quad (7-3)$$

El objetivo perseguido es predecir valores futuros de la variable Y_t . Se adoptará el supuesto de que se dispone de información hasta el momento T , y que para el período $T+l$, con $l > 0$, se desea predecir el valor Y_{T+l} . En el epígrafe siguiente se calculará el predictor para ese valor; ahora únicamente se va a representar de formas alternativas la observación Y_{T+l} , es decir, el valor que se trata de predecir.

Las formas alternativas son las siguientes:

a) Mediante un proceso ARMA.

Consiste simplemente en expresar (7-3) para el valor Y_{T+l} . Por tanto:

$$Y_{T+l} = \varphi_1 Y_{T+l-1} + \dots + \varphi_{p+d} Y_{T+l-p-d} + \varepsilon_{T+l} - \theta_1 \varepsilon_{T+l-1} - \dots - \theta_q \varepsilon_{T+l-q} \quad (7-4)$$

b) Mediante un proceso MA(∞).

Calculando $\psi(L)$ a partir de la relación

$$\varphi(L)\psi(L) = \theta(L) \quad (7-5)$$

donde $\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \dots$

se tiene que

$$Y_{T+l} = \varphi^{-1}(L)\theta(L)\varepsilon_{T+l} = \psi(L)\varepsilon_{T+l} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T+l-j} = \varepsilon_{T+l} + \psi_1 \varepsilon_{T+l-1} + \dots \quad (7-6)$$

A efectos de predicción conviene expresar (7-6) en forma truncada. Designando

$$C_T(l) = \sum_{j=0}^{\infty} \psi_{l+j} \varepsilon_{T-j} \quad (7-7)$$

la observación Y_{T+l} se expresará de la siguiente forma:

$$Y_{T+l} = \varepsilon_{T+l} + \psi_1 \varepsilon_{T+l-1} + \dots + \psi_{l-1} \varepsilon_{T+1} + C_T(l) \quad (7-8)$$

Así pues, la función complementaria de $C_T(l)$ engloba todos los ruidos correspondientes al período muestral o anteriores.

c) Mediante un proceso AR(∞).

Calculando $\pi(L)$ según la relación

$$\varphi(L) = \pi(L)\theta(L) \quad (7-9)$$

donde $\pi(L) = 1 - \pi_1 L - \pi_2 L^2 \dots$

se tiene que

$$\theta^{-1}(L)\varphi(L)Y_{T+l} = \pi(L)Y_{T+l} = \varepsilon_{T+l} \quad (7-10)$$

Desarrollando (7-10) se obtiene:

$$Y_{T+l} = \sum_{j=1}^{\infty} \pi_j Y_{T+l-j} + \varepsilon_{T+l} \quad (7-11)$$

Las formas de presentación que se utilizarán en los siguientes epígrafes son las desarrolladas en los apartados a) y b).

7.3. El predictor óptimo

Se designará por $\tilde{Y}_{T+l|T}$ el predictor óptimo para el período $T+l$ utilizando toda la información disponible para el período T .

En la obtención del predictor $\tilde{Y}_{T+l|T}$ se adoptarán los siguientes supuestos:

a) Los parámetros ϕ_i y θ_j son conocidos.

En el tratamiento de las series temporales se adopta este supuesto sin hacerlo explícito en la mayoría de los casos. En los modelos económicos, por el contrario, al calcular el error de predicción se tiene en cuenta el error debido a la estimación de los parámetros. Hay que destacar por otra parte que en los manuales de econometría no se suele hablar de los errores acumulativos en que se incurre cuando el modelo es dinámico, hecho que sí es tenido en cuenta en los modelos ARIMA.

Para evaluar correctamente la incidencia del supuesto adoptado, conviene tener presente que en series de tamaño superior a 60, que generalmente son las utilizadas en la elaboración de los modelos ARIMA, los errores de estimación de los parámetros son de cuantía reducida.

b) Se supone que todos los ruidos presentes y pasados $\varepsilon_T, \varepsilon_{T-1}, \dots, \varepsilon_1, \varepsilon_0 \dots$ son conocidos.

Como criterio de selección, el predictor $\tilde{Y}_{T+l|T}$ se determinará de forma que minimice la varianza del error de predicción. Por tanto se deberá verificar que:

$$E[Y_{T+l} - \tilde{Y}_{T+l|T}]^2 \leq E[Y_{T+l} - Y^*_{T+l|T}]^2 \quad (7-12)$$

donde $Y^*_{T+l|T}$ es un predictor arbitrario.

PREDICCIÓN

El predictor \tilde{Y}_{T+IT} que cumple la condición (7-12) se dice que es un estimador que goza de un error cuadrático medio mínimo (ECMM).

Teniendo en cuenta que ε_t es una variable normal por hipótesis, la búsqueda del predictor \tilde{Y}_{T+IT} queda restringida al campo de los estimadores lineales, ya que bajo la hipótesis de normalidad un estimador con ECMM es un estimador lineal.

El predictor se va a construir como una función lineal de todos los valores conocidos de ε_t , teniendo en cuenta el supuesto b). Por tanto,

$$\tilde{Y}_{T+IT} = \psi^* \varepsilon_T + \psi^*_{T+1} \varepsilon_{T-1} + \psi^*_{T+2} \varepsilon_{T-2} + \dots \quad (7-13)$$

donde ψ^* son coeficientes que se determinarán de forma que \tilde{Y}_{T+IT} tenga un ECMM.

El error de predicción, definido como diferencia entre el valor observado y el valor de predicción, a partir de (7-6) y (7-13) se puede expresar así:

$$\begin{aligned} Y_{T+I} - \tilde{Y}_{T+IT} &= \varepsilon_{T+I} + \psi_1 \varepsilon_{T+I-1} + \dots + \psi_{I-1} \varepsilon_{T+1} + \psi_I \varepsilon_T + \dots \\ &\quad - \psi^* \varepsilon_T - \psi^*_{T+1} \varepsilon_{T-1} - \psi^*_{T+2} \varepsilon_{T-2} \\ &= \varepsilon_{T+I} + \psi_1 \varepsilon_{T+I-1} + \dots + \psi_{I-1} \varepsilon_{T+1} \\ &\quad + (\psi_I - \psi^*) \varepsilon_T + (\psi_{I+1} - \psi^*_{I+1}) \varepsilon_{T-1} + \dots \end{aligned} \quad (7-14)$$

De acuerdo con (7-14), el error cuadrático medio (ECM) de un estimador \tilde{Y}_{T+IT} viene dado por:

$$\begin{aligned} \text{ECM}[\tilde{Y}_{T+IT}] &= E[Y_{T+I} - \tilde{Y}_{T+IT}]^2 \\ &= E[\varepsilon_{T+I} + \psi_1 \varepsilon_{T+I-1} + \dots + \psi_{I-1} \varepsilon_{T+1} + (\psi_I - \psi^*) \varepsilon_T + (\psi_{I+1} - \psi^*_{I+1}) \varepsilon_{T-1} + \dots]^2 \\ &= \sigma_\varepsilon^2 [1 + \psi_1^2 + \dots + \psi_{I-1}^2] + \sigma_\varepsilon^2 \sum_{j=0}^I (\psi_{I+j} - \psi^*_{I+j})^2 \end{aligned} \quad (7-15)$$

Los términos de la última igualdad de (7-15) aparecen agrupados en dos sumandos. Los términos del primer sumando no dependen de los pesos ψ_j^* y, por lo tanto, no vienen afectados por la elección que se haga de estos pesos. Concentrando la atención en el segundo sumando, se ve fácilmente que éste queda minimizado al hacer:

$$\psi_{I+j}^* = \psi_{I+j} \quad \text{para } j = 0, 1, 2, 3, \dots \quad (7-16)$$

ya que en ese caso este sumando vale 0.

Si se eligen los pesos ψ_{I+j}^* de acuerdo con (7-16) se habrá minimizado el ECM (\tilde{Y}_{T+IT}), y se habrá obtenido el predictor óptimo en el sentido de un estimador con ECMM. Este predictor quedará, pues, definido de la siguiente forma:

$$\tilde{Y}_{T+IT} = \sum_{j=0}^I \psi_{I+j} \varepsilon_{T-j} \quad (7-17)$$

CÁLCULO DEL PREDICTOR

y su ECM será:

$$\text{ECM}[\tilde{Y}_{T+IT}] = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{I-1}^2) \sigma_\varepsilon^2 \quad (7-18)$$

Obsérvese que como

$$E[Y_{T+I} - \tilde{Y}_{T+IT}] = 0 \quad (7-19)$$

entonces se verifica que el ECM coincide exactamente con la varianza del error de predicción.

Teniendo en cuenta que las ε_t se distribuyen normalmente se pueden establecer intervalos de confianza del siguiente tipo:

$$\text{Prob}[\tilde{Y}_{T+IT} - 1,96\sqrt{\text{ECM}[\tilde{Y}_{T+IT}]} < Y_{T+I} < \tilde{Y}_{T+IT} + 1,96\sqrt{\text{ECM}[\tilde{Y}_{T+IT}]}] = 0,95 \quad (7-20)$$

7.4. Cálculo del predictor

En la fórmula (7-17), el predictor óptimo está expresado en forma MA(∞). Comparando con (7-8), donde el valor observado Y_{T+I} también se expresa en forma MA(∞), se advierte que coincide exactamente con $C_T(I)$ (véase la expresión 7-7). En el predictor óptimo, los ruidos correspondientes a $T+1, \dots, T+I$ que son desconocidos se igualan a 0 que es su valor medio teórico. Así como la fórmula (7-18) es adecuada para calcular el ECM, la fórmula (7-17) no es operativa en cambio, salvo para el caso de que se trate de un modelo MA(q) finito, según puede verse a continuación en un ejemplo aplicado a un MA(2).

Ejemplo

Sea el modelo ARIMA (0, 0, 2):

$$\psi_1 = -\theta_1; \psi_2 = -\theta_2 \text{ y } \psi_j = 0 \quad \text{para } j > 2$$

Se calculan a continuación para varios períodos de predicción el valor real u observado, el predictor y el correspondiente error cuadrático medio.

Período	Valor Real	Predictor	ECM
$T+1$	$Y_{T+1} = \varepsilon_{T+1} - \theta_1 \varepsilon_T - \theta_2 \varepsilon_{T-1}$	$\tilde{Y}_{T+1\pi} = -\theta_1 \varepsilon_T - \theta_2 \varepsilon_{T-1}$	σ_ε^2
$T+2$	$Y_{T+2} = \varepsilon_{T+2} - \theta_1 \varepsilon_{T+1} - \theta_2 \varepsilon_T$	$\tilde{Y}_{T+2\pi} = -\theta_2 \varepsilon_T$	$(1 + \theta_1^2) \sigma_\varepsilon^2$
$T+3$	$Y_{T+3} = \varepsilon_{T+3} - \theta_1 \varepsilon_{T+2} - \theta_2 \varepsilon_{T+1}$	$\tilde{Y}_{T+3\pi} = 0$	$(1 + \theta_1^2 + \theta_2^2) \sigma_\varepsilon^2$

Puede verse que para $l > 2$, resultará que $\tilde{Y}_{T+l|T} = 0$ y que $\text{ECM}[\tilde{Y}_{T+l|T}] = (1 + \theta_1^2 + \theta_2^2)\sigma_e^2$. En un modelo MA(q), se verifica que $Y_{T+l|T} = 0$, y que el error cuadrático medio se estabiliza en el valor que toma para $l = q + 1$.

En general, para el cálculo de predicciones de valores futuros se utiliza el modelo en la forma original ARMA, donde la parte AR está constituida como en (7-4) por los coeficientes autorregresivos generalizados. En este caso, el predictor óptimo, que se determina recursivamente, viene dado por la siguiente expresión:

$$\begin{aligned}\tilde{Y}_{T+l|T} &= \varphi_1 \tilde{Y}_{T+l-1|T} + \dots + \varphi_{p+1} \tilde{Y}_{T+l-p-d|T} \\ &+ \tilde{\varepsilon}_{T+l} - \theta_1 \tilde{\varepsilon}_{T+l-1} - \dots - \theta_q \tilde{\varepsilon}_{T+l-q}\end{aligned}\quad (7-21)$$

donde

$$\begin{aligned}\tilde{Y}_{T+l-j|T} &= \begin{cases} Y_{T+l-j} & \text{si } l \leq j \\ \tilde{Y}_{T+l-j|T} & \text{si } l > j \end{cases} \\ \tilde{\varepsilon}_{T+l-i} &= \begin{cases} \varepsilon_{T+l-i} & \text{si } l \leq i \\ 0 & \text{si } l > i \end{cases}\end{aligned}\quad (7-22)$$

La interpretación de las anteriores fórmulas es la siguiente. Cuando el valor de $\tilde{Y}_{T+l-j|T}$ que aparece en el segundo miembro de (7-21) corresponde al período muestral, es decir, menor o igual a T , se utiliza el valor observado; en caso contrario se utiliza el valor previamente pronosticado. Naturalmente la aplicación de esta fórmula debe hacerse de forma recursiva obteniéndose sucesivamente $\tilde{Y}_{T+1|T}, \tilde{Y}_{T+2|T}, \dots, \tilde{Y}_{T+l-1|T}$ antes de obtener el valor $\tilde{Y}_{T+l|T}$. En lo que respecta a los ruidos, si corresponde al período muestral, se toma el valor observado, mientras que en los demás casos se toma el valor medio teórico, que es 0.

A continuación se examinan varios ejemplos.

Ejemplos.

a) Sea un modelo ARIMA (0, 1, 0), es decir, un modelo «random walk»

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

Al pasar a la forma MA(∞), se tiene que $\psi_i = 1$ para $i = 1, 2, 3, \dots$

Para los períodos $T + 1, T + 2$ y $T + 3$ se obtienen los siguientes resultados:

Período	Valor Real	Predictor	ECM
$T + 1$	$Y_{T+1} = Y_T + \varepsilon_{T+1}$	$\tilde{Y}_{T+1 T} = \tilde{Y}_T$	σ_e^2
$T + 2$	$Y_{T+2} = Y_{T+1} + \varepsilon_{T+2}$	$\tilde{Y}_{T+2 T} = \tilde{Y}_{T+1 T} = Y_T$	$2\sigma_e^2$
$T + 3$	$Y_{T+3} = Y_{T+2} + \varepsilon_{T+3}$	$\tilde{Y}_{T+3 T} = \dots = Y_T$	$3\sigma_e^2$

En un modelo «random walk» para todos los períodos futuros se predice el valor correspondiente al último período muestral. El ECM se incrementa en σ_e^2 para cada período adicional.

b) Sea un modelo ARIMA (1, 1, 0)

$$(1 - \phi_1 L)(1 - L)Y_t = \varepsilon_t$$

Pasando a un ARMA mediante la utilización de parámetros autorregresivos generalizados se tiene que

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t$$

donde $\varphi_1 = \phi_1 + 1$ y $\varphi_2 = -\phi_1$.

A partir de la relación:

$$\psi(L)\varphi(L) = 1$$

se tiene que $\psi_0 = 1$, y $\psi_1 = \varphi_1$. Tomando ψ_0 y ψ_1 como valores iniciales, los siguientes valores de ψ_i se determinan recursivamente mediante la fórmula:

$$\psi_k = \varphi_1 \psi_{k-1} + \varphi_2 \psi_{k-2}$$

Para los períodos $T + 1, T + 2$ y $T + 3$ se obtienen los siguientes resultados relativos a la predicción

Período	Valor Real	Predictor	ECM
$T + 1$	$Y_{T+1} = \varphi_1 Y_T + \varphi_2 Y_{T-1} + \varepsilon_{T+1}$	$\tilde{Y}_{T+1 T} = \varphi_1 Y_T + \varphi_2 Y_{T-1}$	σ_e^2
$T + 2$	$Y_{T+2} = \varphi_1 Y_{T+1} + \varphi_2 Y_T + \varepsilon_{T+2}$	$\tilde{Y}_{T+2 T} = \varphi_1 \tilde{Y}_{T+1 T} + \varphi_2 Y_T$	$(1 + \psi_1^2)\sigma_e^2$
$T + 3$	$Y_{T+3} = \varphi_1 Y_{T+2} + \varphi_2 Y_{T+1} + \varepsilon_{T+3}$	$\tilde{Y}_{T+3 T} = \varphi_1 \tilde{Y}_{T+2 T} + \varphi_2 \tilde{Y}_{T+1 T}$	$(1 + \psi_1^2 + \psi_2^2)\sigma_e^2$

c) Sea el modelo ARIMA (1, 0, 1)

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

El coeficiente $\psi_1 = \phi_1 - \theta_1$. Los siguientes se calculan de forma recursiva a partir de

$$\psi_k = \phi_1 \psi_{k-1}$$

Para este modelo, al predecir los períodos $T + 1, T + 2$ y $T + 3$ se obtienen los siguientes resultados:

Período	Valor Real	Predictor	ECM
$T + 1$	$Y_{T+1} = \phi_1 Y_T + \varepsilon_{T+1} - \theta_1 \varepsilon_T$	$\tilde{Y}_{T+1/T} = \phi_1 Y_T - \theta_1 \varepsilon_T$	σ_e^2
$T + 2$	$Y_{T+2} = \phi_1 Y_{T+1} + \varepsilon_{T+2} - \theta_1 \varepsilon_{T+1}$	$\tilde{Y}_{T+2/T} = \phi_1 \tilde{Y}_{T+1/T}$	$(1 + \psi_1^2) \sigma_e^2$
$T + 3$	$Y_{T+3} = \phi_1 Y_{T+2} + \varepsilon_{T+3} - \theta_1 \varepsilon_{T+2}$	$\tilde{Y}_{T+3/T} = \phi_1 \tilde{Y}_{T+2/T}$	$(1 + \psi_1^2 + \psi_2^2) \sigma_e^2$

7.5. Contraste de la estabilidad estructural en la predicción

Cuando se ha utilizado un modelo ARIMA, ya validado, para predecir los valores de la variable en períodos extramuestrales, se puede plantear el siguiente interrogante: ¿sigue siendo válido el modelo para los períodos de predicción?; en otras palabras, ¿se ha producido un cambio estructural en los períodos de predicción?

Con objeto de obtener una respuesta a esta cuestión se puede utilizar el siguiente estadístico

$$E_m^* = \frac{\sum_{l=0}^m e_{T+l/T+1}^2}{\hat{\sigma}_e^2} \quad (7-23)$$

donde

$e_{T+l/T+1}$ es el error de predicción de \tilde{Y}_{T+l+1} utilizando la información disponible en el momento $T + l$.

El estadístico E_m^* se distribuye asintóticamente como una χ^2 con m grados de libertad bajo la hipótesis nula de que no se ha dado una ruptura estructural al pasar a los períodos de predicción. Para comprender el significado de este contraste debe tenerse en cuenta que si el modelo sigue siendo correcto se verificará que

$$e_{T+l/T+1} \approx e_{T+l+1}$$

Por el contrario, en la medida que esto no ocurra, tenderá a aumentar el ECM de cada predicción, y por tanto el estadístico E_m^* , lo que llevaría a rechazar la hipótesis nula.

Ejercicios

[1] Sea el modelo ARIMA (0, 1, 1)

$$(1 - L) Y_t = (1 + 0,8L) \varepsilon_t$$

Se pide expresar la observación Y_{t+1} de tres formas alternativas:

- a) Mediante un proceso ARMA.
- b) Mediante un proceso MA(∞).
- c) Mediante un proceso AR(∞).

[2] Sea el modelo ARIMA (1, 1, 1)

$$(1 - 0,5L)(1 - L) Y_t = (1 + 0,6L) \varepsilon_t$$

Se pide expresar la observación Y_{t+1} de tres formas alternativas:

- a) Mediante un proceso ARMA.
- b) Mediante un proceso MA(∞).
- c) Mediante un proceso AR(∞).

[3] Sea el modelo

$$\phi(L)(1 - L)^d Y_t = \theta(L)$$

Con una muestra de T observaciones se ha realizado la estimación del modelo, y subsequentlyemente la previsión del período $T + l$ mediante el predictor \tilde{Y}_{t+lt} . Posteriormente, se dispone de la observación Y_{t+1} .

En base a esta nueva observación, ¿cuál será la predicción para el período $T + l$? Razona la respuesta.

[4] Sea

$e_{T+lt} = Y_{T+l} - \tilde{Y}_{T+lt}$; error de predicción para l períodos hacia adelante con origen en T .
 $e_{T+l+jT-j} = \tilde{Y}_{t+l-j} - \tilde{Y}_{t+l-jT-j}$; error de predicción para l períodos hacia adelante con origen en $T - j$.

¿Están correlacionados ambos errores de predicción?

[5] Sea:

$$e_{T+lt} = Y_{T+l} - \tilde{Y}_{T+lt}$$

$$e_{T+l+jT-j} = Y_{T+l+j} - \tilde{Y}_{T+l+jT-j}$$

Se pide calcular $E[e_{T+lt} e_{T+l+jT-j}]$.

- 6 Se ha estimado con una muestra de 110 observaciones de la variable Y_t el siguiente modelo:

$$Y_t = \varepsilon_t + 0,5\varepsilon_{t-1} + 0,4\varepsilon_{t-2} + 18; \quad \sigma^2 = 4$$

La información acerca de ε_t e Y_t se ha perdido excepto para los siguientes valores de Y_t :

$$Y_{106} = 20 \quad Y_{107} = 21 \quad Y_{108} = 19 \quad Y_{109} = 19 \quad Y_{110} = 17$$

Bajo este supuesto,

- Efectuar la previsión de Y para los períodos 111, 112 y 113 con origen en $t = 110$.
- Calcular el ECM para las anteriores previsiones.
- Suponiendo que se dispone de una observación adicional — $Y_{111} = 17$ —, actualizar la predicción de los períodos 112 y 113.

- 7 Sea el modelo estimado

$$(1 + 0,6L)(1 - L)Y_t = (1 - 0,5L)\varepsilon_t + 0,8$$

con $\sigma^2 = 1$

Se pide

- Dado $Y_{82} = 40$, $Y_{83} = 41$ y $\varepsilon_{80} = 0,2$, calcular las predicciones $\tilde{Y}_{81/80}$, $\tilde{Y}_{82/80}$ e $\tilde{Y}_{83/80}$.
- Calcular el ECM de las predicciones.
- Establecer una banda de confianza del 90 % para las predicciones.
- Suponiendo que se dispone de una observación adicional — $Y_{81} = 41$ —, actualizar la predicción de los períodos 82 y 83.

- 8 Utilizando el modelo seleccionado en el ejercicio 7 del capítulo 6, se pide

- Calcular predicciones para los períodos 101-105.
- Calcular un intervalo de confianza del 90 % para las predicciones del apartado anterior.
- Comparar los resultados obtenidos con las predicciones que habría obtenido aplicando el proceso utilizado en la generación de la serie y que viene dado por $(1 - L)Y_t = (1 + 0,6)\varepsilon_t$.

- 9 Utilizando el modelo seleccionado en el ejercicio 8 del capítulo 6, se pide

- Calcular predicciones para los períodos 201-210.
- Calcular un intervalo de confianza del 90 % para las predicciones del apartado anterior.
- Comparar los resultados obtenidos con las predicciones que habría obtenido aplicando el proceso utilizado en la generación de la serie y que viene dado por $(1 + 0,5L)(1 - L)Y_t = (1 - 0,8L)\varepsilon_t$.

- 10 Utilizando el modelo seleccionado en el ejercicio 9 del capítulo 6, se pide

- Calcular predicciones para los períodos 101-105.
- Calcular un intervalo de confianza del 90 % para las predicciones del apartado anterior.
- Comparar los resultados obtenidos con las predicciones que habría obtenido aplicando el proceso utilizado en la generación de la serie y que viene dado por $(1 - L)hY_t = (1 - 0,4L)\varepsilon_t$.

CAPÍTULO 8

Modelos estacionales

8.1. Introducción

En muchas ocasiones, los datos que se manejan en la elaboración de modelos ARIMA son datos estacionales, es decir, son datos que tienen unas oscilaciones estrechamente periódicas, donde el período es igual o inferior al año. En algunas magnitudes, estas oscilaciones obedecen a causas físicas y coinciden con las estaciones del año, como es el caso de la pluviosidad. En otras magnitudes las oscilaciones se deben a causas sociales, económicas, etc., es decir, a factores humanos que pueden tener a su vez una motivación de orden físico. Así, el índice de producción industrial en España desciende considerablemente todos los años en el mes de agosto, a causa de las vacaciones, que es un fenómeno de carácter social, si bien a su vez está condicionado por factores climatológicos.

En la elaboración de los modelos de series temporales, se debe tener en cuenta el factor estacional cuando los datos utilizados son de esta naturaleza. En este capítulo se examinarán sucesivamente, según el enfoque Box-Jenkins, las características de los modelos estacionales puros, multiplicativos y no estacionarios. Finalmente, se revisarán brevemente las fases de elaboración de modelos ARIMA teniendo en cuenta la introducción del factor estacional.

8.2. Modelos estacionales puros y estacionarios

Se designará por s al período estacional. Así en datos trimestrales $s = 4$, etc.

Un modelo estacional puro se caracteriza porque sólo existe relación entre las observaciones que distan entre sí s períodos o múltiplos de s . En un modelo estacional puro que además sea estacionario se puede calcular fácilmente la FACT y la FACPT, a partir de los resultados obtenidos en los capítulos 3 y 4.

A continuación se examina en particular el modelo AR(1) estacional, proporcionando algunos resultados para otros casos.

Modelo AR(1) estacional

Un modelo AR(1) estacional, o $AR(1)_s$, viene definido por

$$Y_t = \Phi_1 Y_{t-s} + \varepsilon_t \quad (8-1)$$

o, utilizando el operador polinomial de retardos, por¹

$$\Phi_1(L)Y_t = \varepsilon_t \quad (8-2)$$

En la figura 8.1 se ha representado el modelo (8-1) mediante un diagrama de flechas en el que se ha tomado $s=4$.

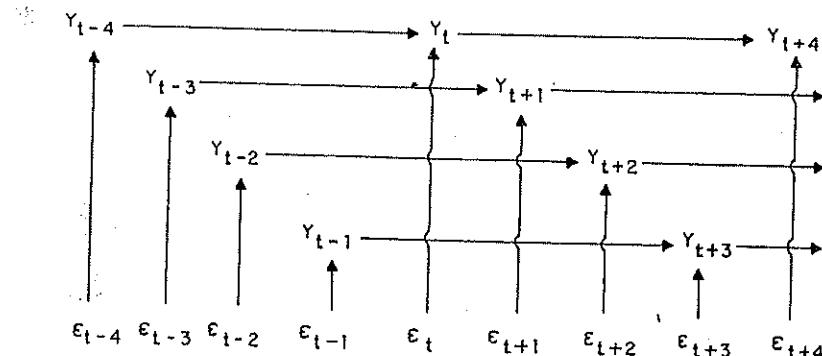


Fig. 8.1 Diagrama de flechas del modelo (8-1): $s=4$.

En el diagrama de flechas queda claramente de manifiesto que solamente existe relación entre las observaciones que están separadas por 4 períodos, como sería el caso de los datos trimestrales. El diagrama de flechas revela que en realidad coexisten 4 procesos autorregresivos independientes entre sí pero con el mismo parámetro Φ_1 .

Bajo el supuesto de estacionariedad, que exige que $|\Phi_1| < 1$, la varianza y las autocovarianzas se pueden calcular multiplicando ambos miembros de (8-1) por Y_{t-s} y tomando esperanzas.

Así, para $\tau=0$

$$\gamma_0 = E(Y_t^2) = \Phi_1^2 E(Y_{t-s})^2 + E(\varepsilon_t)^2 + 2E(\varepsilon_t Y_{t-s}) = \Phi_1^2 \gamma_0 + \sigma_\varepsilon^2$$

¹ Se designará con las letras mayúsculas Φ y Θ a los parámetros estacionales autorregresivos y de medias móviles respectivamente.

con lo que operando

$$\gamma_0 = \frac{1}{1 - \Phi_1^2} \sigma_\varepsilon^2$$

Para $\tau=1$, se obtiene

$$\gamma_1 = E(Y_t Y_{t-1}) = \Phi_1 E(Y_{t-s} Y_{t-1}) + E(\varepsilon_t Y_{t-1})$$

Las dos esperanzas del segundo miembro son igual a 0, según puede deducirse del examen del diagrama de flechas —por ejemplo, entre Y_{t-4} e Y_{t-1} no existe ninguna relación— con lo que

$$\gamma_1 = 0$$

Continuando dando valores sucesivos a τ , se irán obteniendo distintas autocovarianzas, el conjunto de las cuales puede expresarse de la siguiente forma:

$$\gamma_\tau = \begin{cases} \theta_1^2/(1-\Phi_1^2) & \text{para } \tau=0 \\ \Phi_1 \gamma_{\tau-s} & \text{para } \tau=s, 2s, 3s, \dots \\ 0 & \text{en los demás casos} \end{cases} \quad (8-3)$$

Dividiendo ambos miembros de (8-3) se obtiene la FACT

$$\rho_\tau = \begin{cases} 1 & \text{para } \tau=0 \\ \Phi_1 \rho_{\tau-s} & \text{para } \tau=s, 2s, 3s, \dots \\ 0 & \text{en los demás casos} \end{cases} \quad (8-4)$$

Como puede verse en el modelo $AR(1)_s$, son aplicables los resultados obtenidos en el capítulo 3 para el modelo $AR(1)$, con la única salvedad de que solamente hay que considerar los coeficientes correspondientes a los retardos $s, 2s, 3s, \dots$

Ejemplo

En un modelo $AR(1)$ con $\phi_1=0,8$ y en otro $AR(1)_s$ también con $\Phi_1=0,8$ y $s=4$, la FACT de cada uno vienen dadas en el cuadro 8.1 y en la figura 8.2.

Cuadro 8.1. FACT es un modelo AR(1) y en un modelo AR(1)_s ($\Phi_1 = \phi_1 = 0,8$).

Retardo	AR(1)	AR(1) _s
1	0,8000	0
2	0,6400	0
3	0,5120	0
4	0,4096	0,8000
5	0,3277	0
6	0,2621	0
7	0,2097	0
8	0,1678	0,6400
9	0,1342	0
10	0,1074	0
11	0,0859	0
12	0,0687	0,5120

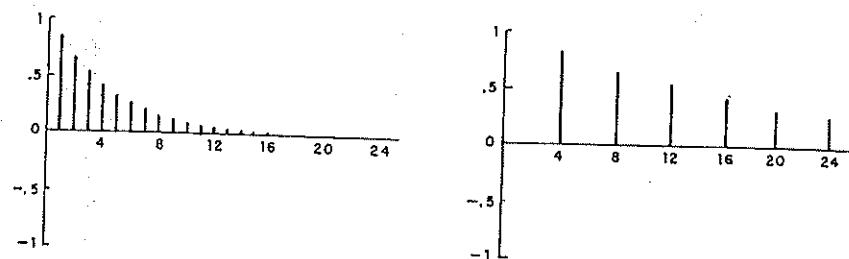


Fig. 8.2 Funciones de autocorrelación de modelos AR(1) y AR(1)_s ($s = 4$).

Otros modelos estacionales puros

Se recoge a continuación la FACT correspondiente a modelos AR(2)_s, MA(1)_s, MA(2)_s y ARMA(1, 1)_s. Su deducción no ofrece ninguna dificultad a partir de los resultados obtenidos en el capítulo 3.

Modelo AR(2)_s: $Y_t = \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \varepsilon_t$

$$\rho_\tau = \begin{cases} 1 & \tau = 0 \\ \Phi_1 \rho_{\tau-s} + \Phi_2 \rho_{\tau-2s} & \tau = s, 2s, \dots \\ 0 & \text{en los demás casos} \end{cases} \quad (8-5)$$

Modelo MA(1)_s: $Y_t = \varepsilon_t - \Theta_1 \varepsilon_{t-s}$

$$\rho_\tau = \begin{cases} 1 & \tau = 0 \\ \frac{-\Theta_1}{1 + \Theta_1^2} & \tau = s \\ 0 & \text{en los demás casos} \end{cases} \quad (8-6)$$

Modelo MA(2)_s: $Y_t = \varepsilon_t - \Theta_1 \varepsilon_{t-s} - \Theta_2 \varepsilon_{t-2s}$

$$\rho_\tau = \begin{cases} 1 & \tau = 0 \\ \frac{-\Theta_1 + \Theta_1 \Theta_2}{1 + \Theta_1^2 + \Theta_2^2} & \tau = s \\ \frac{-\Theta_2}{1 + \Theta_1^2 + \Theta_2^2} & \tau = 2s \\ 0 & \text{en los demás casos} \end{cases} \quad (8-7)$$

Modelo ARMA (1, 1)_s: $Y_t = \Phi_1 Y_{t-s} + \varepsilon_t - \Theta_1 \varepsilon_{t-s}$

$$\rho_\tau = \begin{cases} 1 & \tau = 0 \\ \frac{(1 - \Phi_1 \Theta_1)(\Phi_1 - \Theta_1)}{1 + \Theta_1^2 - 2\Phi_1 \Theta_1} & \tau = s \\ \Phi_1 \rho_{\tau-s} & \tau = 2s, 3s, \dots \\ 0 & \text{en los demás casos} \end{cases} \quad (8-8)$$

En la figura 8.3 se recoge la FACT de dos casos de cada uno de estos 4 últimos modelos.

La función de autocorrelación parcial teórica (FACPT) se puede calcular en los modelos que se han visto de igual forma a como se hizo en el capítulo 4 para modelos no estacionales. Las propiedades de la FACTP serían análogas en uno y otro caso.

En general, un modelo ARMA (P, Q) estacional puro tendrá la siguiente formulación.

$$Y_t - \Phi_1 Y_{t-s} - \dots - \Phi_p Y_{t-ps} = \varepsilon_t - \Theta_1 \varepsilon_{t-s} - \dots - \Theta_Q \varepsilon_{t-Qs}, \quad (8-9)$$

o, alternativamente, utilizando el operador polinomial de retardos

$$\Phi_s(L) Y_t = \Theta_s(L) \varepsilon_t$$

MODELOS ESTACIONALES

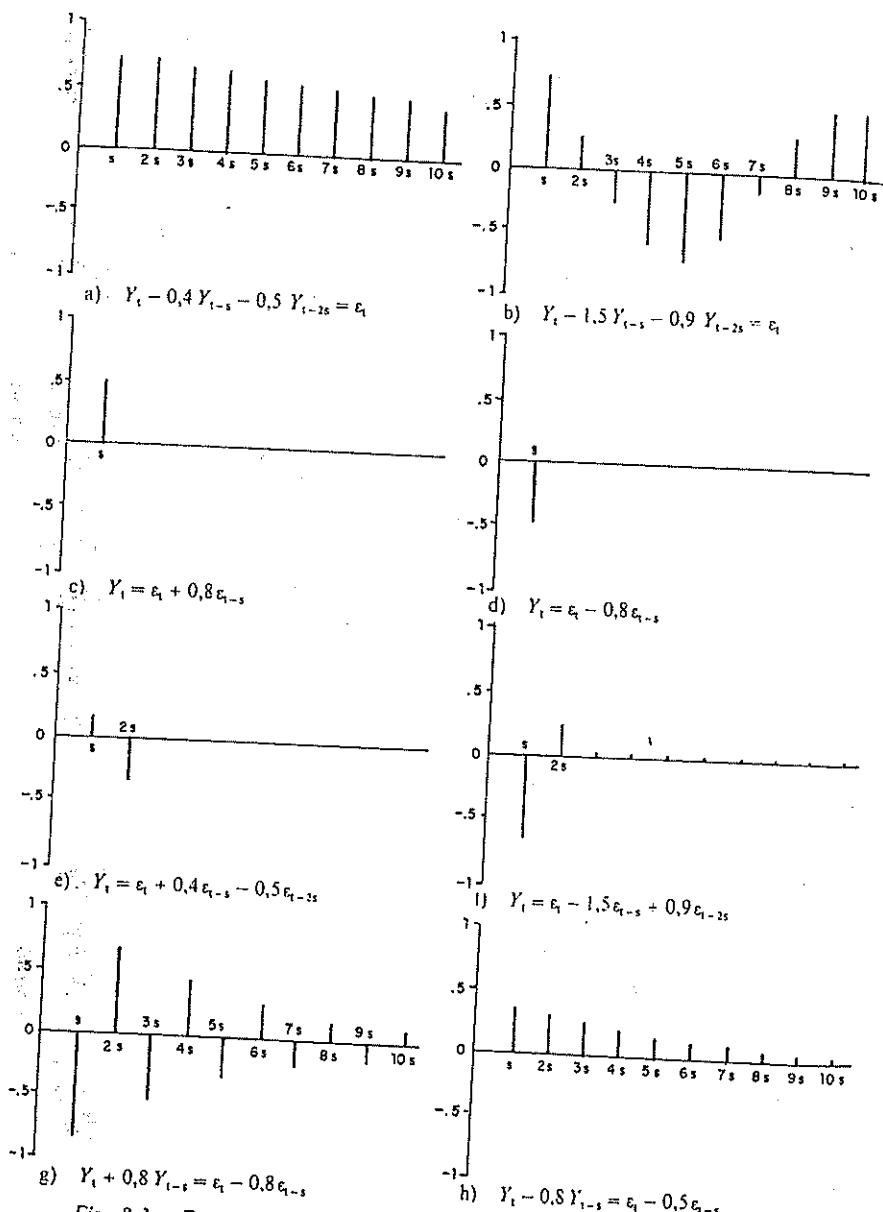


Fig. 8.3. Funciones de autocorrelación de modelos estacionales puros.

MODELOS ESTACIONALES MULTIPLICATIVOS ESTACIONARIOS

8.3. Modelos estacionales multiplicativos estacionarios

En el epígrafe anterior se han examinado modelos ARMA estacionales puros, mientras que en el capítulo 3 se introdujeron los modelos ARMA sin darles en aquel momento ningún calificativo especial, pero a los que a partir de ahora se les denominará ordinarios para diferenciarlos de los estacionales.

Combinando el modelo (8-10) con el modelo (3-77) se puede formular el siguiente modelo general mixto, o multiplicativo, ARMA(p, q) \times ARMA(P, Q):

$$\phi(L)\Phi_s(L)Y_t = \theta(L)\Theta_s(L)\epsilon_t \quad (8-11)$$

En (8-11), tanto en la parte autorregresiva como en la parte de medias móviles aparece el producto de dos operadores polinomiales: uno ordinario y otro estacional.

Para ver cómo operan los dos factores consideremos a título de ejemplo el modelo MA(1) \times MA(1)_s, cuya expresión será la siguiente tomando $s = 12$:

$$Y_t = (1 - \theta_1 L)(1 - \Theta_1 L^{12})\epsilon_t \quad (8-12)$$

Desarrollando el segundo miembro de (8-12) se tiene que

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \Theta_1 \epsilon_{t-12} + \theta_1 \Theta_1 \epsilon_{t-13} \quad (8-13)$$

Evidentemente, una formulación equivalente a (8-13) utilizando solamente parámetros ordinarios sería la siguiente

$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_{12} \epsilon_{t-12} - \theta_{13} \epsilon_{t-13} \quad (8-14)$$

Como puede observarse, el modelo (8-13) contiene solamente dos parámetros —uno ordinario y otro estacional—, reflejando el término correspondiente a ϵ_{t-13} la interacción entre ambos. En cambio, en (8-14) se ha introducido un nuevo parámetro para recoger la influencia de este último término. Naturalmente si existe la interacción entre la parte ordinaria y la parte estacional es preferible utilizar el modelo (8-13), ya que es más escueto.

El cálculo de la FACT correspondiente a los modelos multiplicativos puede llegar a ser laborioso. Peña (1979) ha deducido una fórmula general que permite aproximar los coeficientes de autocorrelación de un proceso ARMA(p, q) \times ARMA(P, Q). Esta aproximación viene dada por

$$p_t^T = p_\tau + \sum_{i=1}^s p_{\tau i}^s (\rho_\tau + s\chi_i + p_{s\chi i-\tau}) \quad (8-15)$$

donde

p_τ : coeficiente de autocorrelación de orden τ en un proceso ordinario ARMA(p, q).

p_τ^s : coeficiente de autocorrelación de orden τ en un proceso estacional puro ARMA(P, Q).

p_T^s : coeficiente de autocorrelación de orden τ en proceso multiplicativo ARMA(p, q) \times ARMA(P, Q).

La fórmula (8-15) es exacta en caso de que la parte ordinaria sea un proceso de medias móviles de orden inferior al período estacional. Si la parte ordinaria tiene términos autorregresivos, entonces es claro que p_s no es igual a 0; sin embargo, si el orden de la parte autorregresiva es bajo y si el período estacional es relativamente elevado —p. ej., en el caso de datos mensuales—, se verifica que $p_s \approx 0$ para $\tau \geq s$. Adoptando esta aproximación, la aplicación de la fórmula (8-15) permitirá obtener los siguientes resultados:

- a) En los retardos bajos —1, 2,..., $s/2$ — la única estructura que aparece es la correspondiente a la parte ordinaria

$$p_s^T \approx p_\tau \quad \text{para } \tau = 0, 1, 2, \dots, s/2 \quad (8-16)$$

ya que

$$p_{\tau+s} \approx 0 \quad y \quad p_{s+\tau} \approx 0 \quad \text{para } \tau = 1, 2, \dots$$

- b) En los retardos $\tau = s, 2s, 3s, \dots$, el único efecto importante es el correspondiente a la parte estacional. En efecto

$$p_s^T = p_s + p_s^s(p_{2s} + p_0) + p_{2s}^s(p_{3s} + p_s) \approx p_s^s p_0 = p_s^s \quad (8-17)$$

- c) En los retardos contiguos a múltiplos de s , se obtiene simétricamente la reproducción de la parte ordinaria tomando como referencia los valores de los retardos estacionales.

Así para $\tau = s+k$, siendo $k < s/2$, se obtiene el siguiente resultado:

$$p_{s+k}^T = p_{s+k} + p_s^s(p_{2s+k} + p_{-k}) + \dots \approx p_s^s p_{-k} = p_s^s p_k$$

Análogamente

$$p_{s-k}^T = p_{s-k} + p_s^s(p_{2s-k} + p_k) + \dots \approx p_s^s p_k$$

De acuerdo con lo anterior

$$p_{s+k}^T \approx p_{s-k}^T$$

En general se verifica que

$$p_{s+k}^T \approx p_{s-k}^T \approx p_s^s \times p_k \quad (8-18)$$

A los coeficientes de autocorrelación contiguos a múltiplos de s se los denomina a veces coeficientes satélites, caracterizándose porque recogen la interacción entre la parte ordinaria y la parte estacional, y por la propiedad de simetría. Conviene tener en cuenta que la aproximación dada en (8-18) y por tanto la propiedad de simetría puede perderse cuando el período estacional es corto —por ej., en datos trimestrales— y la parte ordinaria es autorregresiva con raíces cercanas al círculo unidad.

A continuación se examinan la FACT de algunos modelos estacionales multiplicativos.

Modelo ARMA(0, 1) × ARMA(0, 1),

Este proceso es el que se expuso a título de ejemplo en (8-12).

Como en la parte ordinaria no existen términos autorregresivos, las aproximaciones (8-16), (8-17) y (8-18) toman valores exactos. De acuerdo con (8-6) y (3-5c) la FACT toma los siguientes valores:

$$p_\tau^T = \begin{cases} 1 & \tau = 0 \\ \frac{-\theta_1}{1 + \theta_1^2} & \tau = 1 \\ p_1 \times p_s^s & \tau = s - 1 \\ \frac{-\theta_1}{1 + \theta_1^2} & \tau = s \\ p_1 \times p_s^s = p_{s-1}^T & \tau = s + 1 \\ 0 & \text{en los demás casos} \end{cases}$$

Vamos a calcular la FACT de una forma alternativa. Si en (8-13) se multiplican ambos miembros por $Y_{t-\tau}$ y se toman esperanzas se tiene que

$$E(Y_t Y_{t-\tau}) = E(\epsilon_t - \theta_1 \epsilon_{t-1} - \theta_1 \epsilon_{t-s} + \theta_1 \theta_1 \epsilon_{t-s-1})(\epsilon_{t-\tau} - \theta_1 \epsilon_{t-\tau-1} - \theta_1 \epsilon_{t-\tau-s} - \theta_1 \theta_1 \epsilon_{t-\tau-s-1}) \quad (8-20)$$

Para distintos valores de τ , se obtienen a partir de (8-20) autocovarianzas de distintos órdenes:

$$\begin{aligned} \gamma_0 &= (1 + \theta_1^2 + \theta_1^2 + \theta_1^2 \theta_1^2) \sigma_\epsilon^2 = (1 + \theta_1^2)(1 + \theta_1^2) \sigma_\epsilon^2 \\ \gamma_1 &= -\theta_1 (1 + \theta_1^2) \sigma_\epsilon^2 \\ \gamma_{s-1} &= \theta_1 \theta_1 \sigma_\epsilon^2 \\ \gamma_s &= -\theta_1 (1 + \theta_1^2) \sigma_\epsilon^2 \\ \gamma_{s+1} &= \theta_1 \theta_1 \sigma_\epsilon^2 \end{aligned} \quad (8-21)$$

Dividiendo las distintas ecuaciones de (8-21) por γ_0 se obtiene, como puede verse fácilmente, el sistema (8-19).

En la figura 8.4 a) y b) se ha representado la FACT correspondiente a dos conjuntos de valores paramétricos. Obsérvese que el signo de los coeficientes "satélites" se obtiene como producto de los signos correspondientes al coeficiente ordinario y al coeficiente estacional.

Modelo ARMA(0, 2) × ARMA(0, 1),

La expresión analítica de este modelo es la siguiente

$$Y_t = (1 - \theta_1 L - \theta_2 L^2)(1 - \Theta_1 L^s) \epsilon_t$$

MODELOS ESTACIONALES

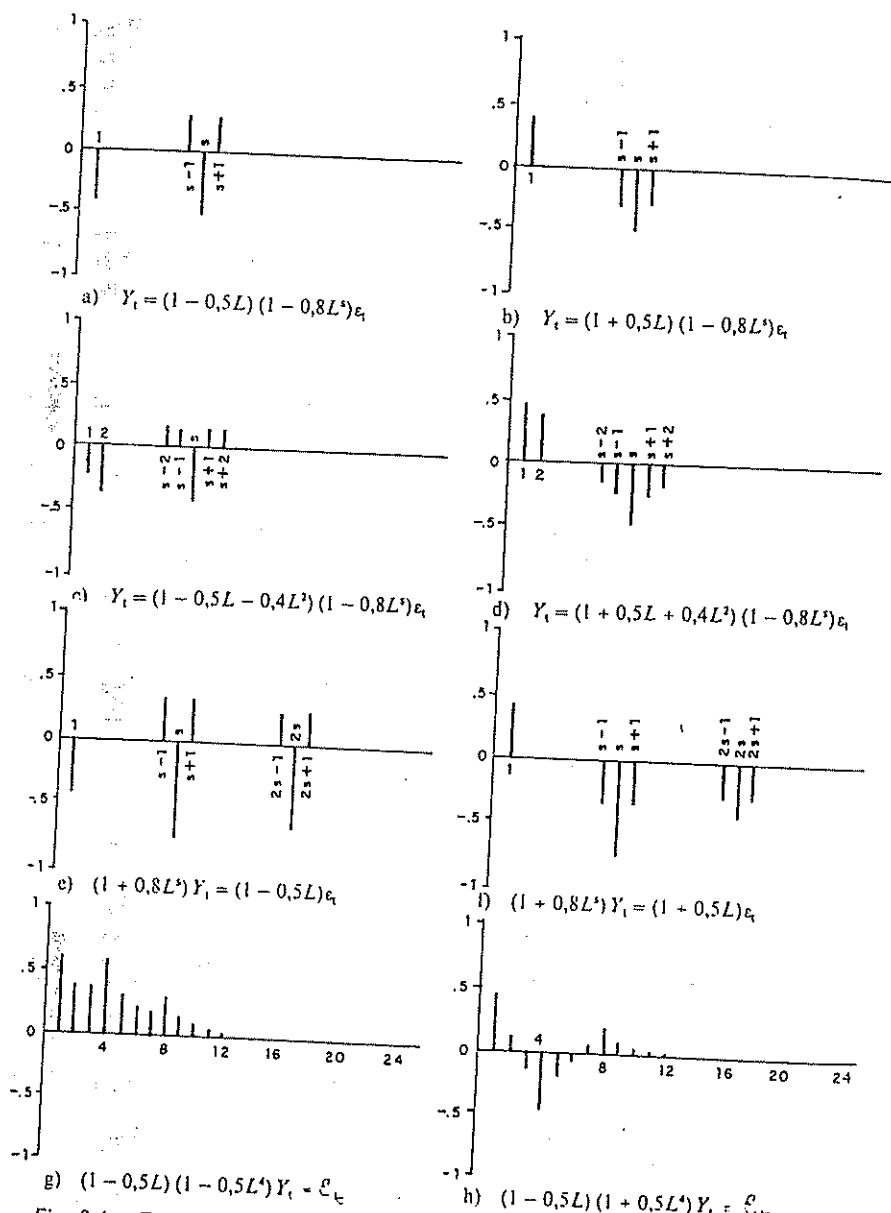


Fig. 8.4. Funciones de autocorrelación de modelos estacionales multiplicativos.

MODELOS ESTACIONALES MULTIPLICATIVOS ESTACIONARIOS

De acuerdo con (8-15), (8-7), (3-69) y (3-70), la FACT toma los siguientes valores:

$$\rho_{\tau}^T = \begin{cases} 1 & \tau = 0 \\ \frac{-\theta_1 + \theta_1 \theta_2}{1 + \theta_1^2 + \theta_2^2} & \tau = 1 \\ \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} & \tau = 2 \\ \rho_2 \rho_s^{\tau} & \tau = s-2 \\ \rho_1 \rho_s^{\tau} & \tau = s-1 \\ \frac{-\theta_1}{1 + \theta_1^2} & \tau = s \\ \rho_1 \rho_s^{\tau} = \rho_{s-1}^T & \tau = s+1 \\ \rho_2 \rho_s^{\tau} = \rho_{s-2}^T & \tau = s+2 \end{cases} \quad (8-23)$$

Puede verse en la figura 8.4 c) y d) la representación de dos casos correspondientes a este modelo. Los coeficientes satélites, que ahora son 4, están situados simétricamente también con respecto al coeficiente estacional.

Modelo ARMA (0, 1) \times ARMA (1, 0)_s

Analíticamente el modelo viene dado por la siguiente expresión:

$$(1 - \Phi_1 L^s) Y_t = (1 - \theta_1 L) \epsilon_t \quad (8-24)$$

De acuerdo con (8-15), (8-4) y (3-56), la FACT viene dada por

$$\rho_{\tau}^T = \begin{cases} 1 & \tau = 0 \\ \frac{-\theta_1}{1 + \theta_1^2} & \tau = 1 \\ \rho_1 \rho_s^{\tau} & \tau = s-1 \\ \rho_s^{\tau} = \Phi_1 & \tau = s \\ \rho_1 \rho_s^{\tau} = \rho_{s-1}^T & \tau = s+1 \\ \Phi_1 \rho_{\tau-s}^T & \tau = 2s-1, 3s-1, \dots \\ \tau = 2s, 3s & \tau = 2s+1, 2s+2, \dots \\ 0 & \text{en los demás casos} \end{cases} \quad (8-25)$$

Como la parte estacional es autorregresiva, la FACT de los múltiplos de s y los valores adyacentes se extiende hasta el infinito. En la figura 8.4 e) y f) pueden verse dos casos correspondientes a este modelo.

Modelo ARMA(1,0) × ARMA(1,0)_s

Un modelo de este tipo viene dado por la siguiente expresión

$$(1 - \phi_1 L)(1 - \Phi L^s) Y_t = \epsilon_t \quad (8-26)$$

En este caso no es fácil establecer los patrones de comportamiento específico de la FACT. Por ello, se utilizará directamente la fórmula general (8-15), así como las expresiones (3-17) y (8-4). En consecuencia

$$\begin{aligned} p_{\tau}^T &= p_{\tau} + p_s^s (p_{\tau+s} + p_{s-\tau}) + p_{2s}^s (p_{\tau+2s} + p_{2s-\tau}) + \dots \\ p_{\tau} &= \phi_1 p_{\tau-s} \\ p_{\tau}^s &= \Phi_1 p_{\tau-s} \end{aligned} \quad (8-27)$$

En la figura 8.4 g) y h) se ha representado la FACT de dos casos correspondientes a este modelo. Como puede verse, especialmente en el caso h), la FACT tiene una continuidad en la que es difícil percibir la interacción entre la parte ordinaria y la parte estacional. En cualquier caso se pierde la simetría perfecta de los coeficientes satélites.

En lo que se refiere al comportamiento de la FACT, los 4 modelos examinados puede resumirse diciendo que tiene un grado de dificultad creciente. Así cuando la parte ordinaria y la parte estacional son de medias móviles, los patrones de comportamiento son nítidos; se complican un poco las cosas cuando se introduce una parte estacional autorregresiva. Naturalmente, la situación es más difícil todavía cuando existe una parte autorregresiva ordinaria.

Aunque existen estudios acerca del comportamiento de la función de autocorrelación parcial teórica (FACPT) en un modelo estacional multiplicativo², los patrones son más difíciles de establecer. En general, puede decirse que la parte de medias móviles, ordinaria y/o estacional, introduce un decrecimiento exponencial con posibles oscilaciones sinusoidales. Por el contrario la FACPT de un proceso autorregresivo estacional presenta un corte brusco.

Como resumen de los modelos estacionales puede construirse el siguiente cuadro

Modelos	FACT	FACPT
AR(p) × AR(P) _s	Decaimiento exponencial y/u oscilaciones sinusoidales atenuadas	Se anula para $\tau > p + sP$
MA(q) × MA(Q) _s	Se anula para $\tau > q + sQ$	Decaimiento exponencial y/u oscilaciones sinusoidales atenuadas

²Véase Hamilton y Wattis (1978).

8.4. Modelos estacionales no estacionarios

De la misma forma que en los modelos no estacionales se planteaba el problema de no estacionariedad en la media, en los modelos estacionales se presenta un problema similar. Como introducción a este tema, vamos a considerar el siguiente modelo en que la parte estacional es determinista y la parte no estacional es un ruido blanco:

$$\begin{cases} Y_t = S_t + \epsilon_t \\ S_t = S_{t-s} \end{cases} \quad (8-28)$$

donde S_t es el componente estacional que permanece constante cuando va referido a una misma estación. Así, en datos mensuales $S_1, S_2, S_3, \dots, S_{12}$ tomarían valores que podrían ser distintos, pero se cumpliría que $S_{13} = S_1, S_{14} = S_2$, etc.

Conociendo S_j para $j = 1, 2, \dots, s$, se puede predecir exactamente, como es obvio, la evolución de la estacionalidad de la serie Y_t . Por esta razón el modelo (8-28) es un modelo estacional determinista.

Los factores estacionales S_j se pueden estimar aplicando mínimos cuadrados al siguiente modelo

$$Y_{it} = \sum_{h=1}^s S_h F_{ih} + \epsilon_{it} \quad \begin{matrix} j = 1, 2, \dots, s \\ i = 1, 2, 3, \dots \end{matrix} \quad (8-29)$$

donde

$$F_{ih} = \begin{cases} 1 & \text{para } h=j \\ 0 & \text{en los demás casos} \end{cases}$$

El modelo (8-29) es una forma de expresar (8-18) que incorpora en una sola ecuación a todos los factores estacionales.

También se puede modelizar una serie temporal estacional mediante modelos de regresión que utilizan términos trigonométricos como en el siguiente caso:

$$Y_t = \beta_0 + \sum_{j=1}^{s/2} (\beta_{1j} \sin \frac{2\pi jt}{12} + \beta_{2j} \cos \frac{2\pi jt}{12}) \epsilon_t \quad (8-30)$$

El problema que plantea tanto el modelo (8-28) —u (8-29)— como el (8-30) es que los esquemas de la estacionalidad que siguen son muy rígidos, ya que los factores estacionales permanecen constantes a lo largo del tiempo. Una forma de flexibilizar el modelo (8-28) consiste en introducir un componente aleatorio en la evolución de S_t , como se hace en la siguiente formulación:

$$\begin{cases} Y_t = S_t + \omega_t \\ S_t = S_{t-s} + \eta_t \end{cases} \quad (8-31)$$

donde ω_t y η_t son variables de ruido blanco independientes entre sí y con varianzas σ_{ω}^2 y σ_{η}^2 .

En (8-31) el componente estacional sigue un modelo similar al paseo aleatorio, con la diferencia de que la variable S en el segundo miembro aparece con s retardos en lugar de 1. En consecuencia, se trata de un modelo estacional no estacionario.

Si en la variable Y_t se toman diferencias estacionales de orden 1, el modelo (8-31) se transformaría en el siguiente:

$$w_t = \Delta_s Y_t = \eta_t + \omega_t - \omega_{t-s} \quad (8-32)$$

La serie $\Delta_s Y_t$ sigue un proceso MA(1)_s

$$\Delta_s Y_t = \varepsilon_t - \Theta_1 \varepsilon_{t-s} \quad (8-33)$$

donde $\varepsilon_t = \omega_t + \eta_t$ y Θ_1 depende de $m = \sigma_\eta^2 / \sigma_\omega^2$ según se verá a continuación.

En el modelo (8-33), el coeficiente de autocorrelación de orden s viene dado por

$$\rho_s = \frac{-\Theta_1}{1 + \Theta_1^2} \quad (8-34)$$

Por otra parte, operando en (8-32) se obtiene:

$$Y_0 = E[w_t^2] = \sigma_\eta^2 + 2\sigma_\omega^2 \quad (8-35)$$

$$\gamma_s = E[w_t w_{t-s}] = -\sigma_\omega^2$$

con lo que

$$\rho_s = \frac{\gamma_s}{Y_0} = \frac{-\sigma_\omega^2}{\sigma_\eta^2 + 2\sigma_\omega^2} = \frac{-1}{m+2} \quad (8-36)$$

Igualando el segundo miembro de (8-36) con el segundo miembro de (8-34) se obtiene una ecuación de segundo grado que permite calcular Θ_1 en función de m .

Volviendo a (8-32), ha quedado patente que al tomar diferencias estacionales de orden 1 en la serie Y_t , la serie resultante es estacionaria en el componente estacional de la misma forma que al tomar diferencias ordinarias se obtendría una serie estacionaria en el caso de modelos homogéneos no estacionarios en media.

En teoría el orden de las diferencias estacionales que se toman puede ser de orden más o menos elevado, aunque en la práctica en muy pocas ocasiones es superior a 1.

Se utiliza la FACE de la serie, o de otras transformaciones, como instrumento para determinar si es necesario o no tomar diferencias estacionales. Así, cuando la secuencia de los coeficientes de autocorrelación correspondientes a $s, 2s, 3s, \dots$ disminuyen lentamente, será un indicio de que el componente estacional es no estacionario.

Introduciendo en (8-11) esquemas no estacionarios, se puede formular el siguiente modelo multiplicativo

$$\phi(L)\Phi_s(L)\Delta^d \Delta_s^D Y_t^{(1)} = \theta(L)\Theta_t(L)\varepsilon_t \quad (8-37)$$

Por su generalidad, este modelo puede ser adecuado para el tratamiento de la mayor parte de las series que se manejan en la práctica.

Entre las diversas especificaciones concretas que se pueden derivar de la fórmula (8-37), se ha comprobado que frecuentemente es apropiado utilizar el esquema:

$$\Delta\Delta_{12} Y_t^{(1)} = (1 - \theta_1 L)(1 - \Theta_1 L^s)\varepsilon_t \quad (8-38)$$

donde la transformación de Box-Cox seleccionada es en general $\lambda = 0$, y por tanto, consiste en tomar logaritmos neperianos.

El modelo (8-38) corresponde a un esquema en que la serie, sometida a una transformación instantánea adecuada, se descompone de forma aditiva en tendencia, factor estacional y componente aleatorio pero teniendo un carácter estocástico, tanto la tendencia como el factor estacional. Un modelo de este tipo se formula de la siguiente forma:

$$\begin{aligned} Y^{(1)} &= M_t + S_t + \omega_t \\ M_t &= M_{t-1} + \xi_t \\ S_t &= S_{t-s} + \eta_t \end{aligned} \quad (8-39)$$

donde ω_t , ξ_t y η_t son variables de ruido blanco independientes entre sí.

Si se toman diferencias ordinarias de orden 1 sobre $Y^{(1)}$ para el período t y el período $t-s$ se obtiene:

$$\begin{aligned} Y_t^{(1)} - Y_{t-1}^{(1)} &= S_t - S_{t-1} + \xi_t + \omega_t - \omega_{t-1} \\ Y_{t-s}^{(1)} - Y_{t-s-1}^{(1)} &= S_{t-s} - S_{t-s-1} + \xi_{t-s} + \omega_{t-s} - \omega_{t-s-1} \end{aligned} \quad (8-40)$$

Restando en (8-40) la segunda ecuación de la primera se obtiene:

$$\Delta\Delta_s Y_t^{(1)} = (\omega_t + \xi_t + \eta_t) - (\omega_{t-1} + \eta_{t-1}) - (\omega_{t-s} + \xi_{t-s}) + \omega_{t-s-1} \quad (8-41)$$

Denominando $\varepsilon_t = \omega_t + \xi_t + \eta_t$, se puede pasar de la formulación (8-41) al modelo (8-38) donde los parámetros θ_1 y Θ_1 se determinan en función de la relación entre las varianzas σ_ω^2 , σ_ξ^2 y σ_η^2 .

8.5. Elaboración de un modelo ARIMA estacional

La elaboración de un modelo ARIMA estacional presenta unas características análogas a las que se vieron en el capítulo 4 para modelos ARIMA no estacionales. Por tanto se pueden distinguir las etapas de especificación, estimación, validación y

predicción. A continuación, se examina la incidencia que puede tener en las tres primeras etapas la inclusión del factor estacional.

Identificación

En la identificación de un modelo ordinario se analiza en primer lugar la estacionariedad de la serie en media y en varianza, y su conversión en estacionaria en caso de que no lo fuera, tomando diferencias ordinarias y la transformación de Box-Cox apropiada. Adicionalmente, y como ya se ha visto en el epígrafe anterior, en el modelo estacional hay que examinar si la serie es o no estacionaria en el componente estacional. En caso de que no lo sea, se toman diferencias de orden estacional. El problema es decidir si el modelo es estacionario en el componente estacional en base a unos pocos coeficientes de autocorrelación correspondientes a los retardos múltiples del período estacional. Imaginense, por ejemplo, una serie de 200 datos mensuales, que puede considerarse relativamente larga en el terreno de la información económica. Según la recomendación de Box y Jenkins admitida por muchos autores, el número de coeficientes de autocorrelación calculados no debe superar la cuarta parte del tamaño muestral; así pues, en nuestro ejemplo se dispondrá de los coeficientes estacionales para los retardos 12, 24, 36 y 48, y en base a esos pocos coeficientes se debe dictaminar si siguen un decrecimiento aritmético, un decrecimiento geométrico o son valores aislados. Solamente en el caso de un decrecimiento aritmético sería pertinente tomar diferencias estacionales.

Conviene señalar a la hora de tomar diferencias estacionales y no estacionales que una diferencia estacional para datos mensuales — $\Delta_{12} = 1 - L^{12}$ — se puede factorizar según sus doce raíces, de las cuales dos son reales y diez complejas³. Es decir:

$$(1 - L^{12}) = (1 - L)(1 + L)(1 - \sqrt{3}L + L^2)(1 - L + L^2) \times \\ \times (1 + L^2)(1 + L + L^2)(1 + \sqrt{3}L + L^2) \quad (8-42)$$

Como puede verse, entre estos factores se encuentra $(1 - L)$. Quiere esto decir que en ocasiones una diferencia estacional puede ser adecuada para eliminar tanto la tendencia como la no estacionariedad del componente estacional.

Por otra parte, como señalan Abraham y Box (1978), para eliminar la no estacionariedad del componente estacional, puede ser suficiente con aplicar a la serie uno de los factores complejos —por ejemplo, $1 - \sqrt{3}L + L^2$ — de la diferencia estacional.

Resumiendo lo que antecede: la aplicación del operador $(1 - L)(1 - L^{12})$ puede sustituirse en determinadas circunstancias por un operador del tipo $(1 - L)(1 - \sqrt{3}L + L^2)$ con lo que se perderían 3 observaciones iniciales en lugar de 12.

Estimación

Los métodos que se examinaron en el capítulo 5, son también válidos para la

³Véase el ejercicio 8.2.

estimación de modelos estacionales. Conviene no obstante hacer algunas consideraciones acerca de la inclusión de parámetros estacionales en el proceso de estimación.

Así, un proceso autorregresivo mixto —con componente estacional y no estacional— no se puede estimar por métodos lineales, a diferencia de lo que ocurriría con los procesos autorregresivos no estacionales.

Por otra parte, en la estimación condicionada de los procesos autorregresivos estacionales hay que tener presente que se perderán $D \times s$ observaciones iniciales. La utilización de métodos de estimación no condicionada requiere el cálculo hacia atrás de muchos más términos que en los procesos no estacionales. Por este motivo, en los modelos estacionales las diferencias entre una estimación condicionada y una estimación no condicionada pueden llegar a ser considerable, siendo preferible, como es lógico, utilizar este último procedimiento.

Validación

Para la validación de modelos estacionales se aplican las técnicas descritas en el capítulo 6. Específicamente, en modelos de este tipo se debe examinar si después de ajustado el modelo seleccionado, subsiste en los residuos algún componente de carácter estacional. Cuando la FACE de los residuos presente valores significativos en algún retardo múltiplo del período estacional será indicativo de que el modelo seleccionado no explica completamente el componente estacional. Una forma alternativa para detectar si aún subsisten periodicidades sin explicar es mediante el periodograma acumulativo normalizado, según proponen Box y Jenkins (1976).

Ejercicios

[1] Considere los modelos estacionales siguientes:

$$Y_t = (1 - \theta_1 L)(1 - \Theta_1 L^{12})\epsilon_t \quad (a)$$

$$Y_t = (1 - \theta_1 L - \theta_{12} L^{12} - \theta_{13} L^{13})\epsilon_t \quad (b)$$

Se pide:

¿Se puede distinguir entre la FACT del modelo (a) y la FACT del modelo (b)?

[2] Las 12 raíces del operador $\Delta_{12} = 1 - L^{12}$ son las siguientes:

- | | |
|----------|-----------------------|
| (1)-(2) | $(\sqrt{3} \pm i)/2$ |
| (3)-(4) | $(1 \pm i\sqrt{3})/2$ |
| (5)-(6) | $\pm i$ |
| (7)-(8) | $(-1 \pm \sqrt{3})/2$ |
| (9)-(10) | $(-\sqrt{3} \pm i)/2$ |
| (11) | -1 |
| (12) | +1 |

Se pide:

- Comprobar que $I - L^{12}$ se puede factorizar utilizando las raíces anteriores.
- Obtener para cada par de raíces complejas el periodo y la frecuencia en ciclos por año.
- Representar gráficamente las raíces de $I - L^{12} = 0$.

3 Sea el modelo

$$(I - L) PIBI_{4t+j} = (I - \theta_4 L^4) \varepsilon_{4t+j} \quad t = 0, 1, 2, \dots$$

donde

$PIBI_{4t+j}$ es el producto interior bruto industrial del año t y trimestre j .

Se ha estimado este modelo con una muestra correspondiente a 21 años completos ($t = 0, 1 \dots 20$), efectuándose una predicción de los siguientes valores

$$PIBI_{4x21+l/4x21} \quad l = 1, 2, 3, 4$$

A partir de estas predicciones se calcula una predicción global para el año 21

$$PIBI_{21}^* = \sum_{l=1}^4 PIBI_{4x21+l/4x21}$$

Se pide determinar el ECM de predicción global del año 21.

4 Se ha estimado el siguiente modelo a partir de una muestra de 120 observaciones

$$(I - 0.5L)(I + 0.7L^4)(I - L) Y_t = (I - 0.5L^4)\varepsilon_t$$

Se conoce también la siguiente información

t	110	111	112	113	114	115	116	118	119	120
Y_t	56	59	58	60	63	66	65	64	63	68

Se pide:

- Expresar el modelo mediante un proceso ARMA.
- Expresar el modelo mediante un proceso MA(∞).
- Efectuar la predicción de Y para los períodos 121, 122, 123, 124 y 125.
- Calcular el ECM para las previsiones del apartado c).
- Suponiendo que se conoce que $Y_{121} = 70$, actualizar la predicción para los períodos 122 y 123.

5 A partir de la serie F del apéndice 4 —índice de producción industrial (IPI)— y utilizando el período —enero 1965 a diciembre 1983— se pide

- Identificar el modelo —o modelos— adecuados para representar la serie.
- Estimar los modelos alternativos considerados en el apartado a).

- Efectuar la validación de los distintos modelos alternativos, seleccionando el más adecuado.
- Predecir el IPI para los períodos enero, febrero y marzo de 1984, comparando los resultados con los datos observados.

Nota: Sobre esta serie puede consultarse el estudio realizado por Sanz (1978).

6 A partir de la serie G del apéndice 4 —saldos medios mensuales de las disponibilidades líquidas (M3)—, y utilizando el período enero de 1975 a diciembre de 1983 se pide

- Identificar el modelo —o modelos— adecuados para representar la serie. En las figuras 8.5 y 8.6 se ha representado para distintas transformaciones la evolución de la serie, la FACE y la FACPE.
- Estimar los modelos alternativos considerados en el apartado a).
- Efectuar la validación de los distintos modelos alternativos, seleccionando el más adecuado.
- Predecir el M3 para los períodos enero, febrero y marzo de 1984, comparando los resultados obtenidos con los datos observados.

Nota: Sobre esta serie puede consultarse el estudio realizado por Espasa (1980).

7 A partir de la serie H del apéndice 4 —consumo de energía electiva en el sector industrial del País Valenciano (CEEPV)—, y utilizando el período enero 1976 a diciembre de 1983, se pide

- Identificar el modelo —o modelos— adecuados para representar la serie. En las figuras 8.7 y 8.8 se ha representado para distintas transformaciones la evolución de la serie, la FACE y la FACPE.
- Estimar los modelos alternativos considerados en el apartado a).
- Efectuar la validación de los distintos modelos alternativos, seleccionando el más adecuado.
- Predecir el CEEPV para los períodos enero, febrero y marzo de 1984, comparando con los datos observados.

Nota: Sobre esta serie puede consultarse el estudio realizado por Molto y Uriel (1984).

8 A partir de la serie simulada I del apéndice 4, y utilizando las observaciones 1-240, se pide

- Identificar el modelo —o modelos— adecuados para representar la serie.
- Estimar los modelos alternativos considerados en el apartado a).
- Efectuar la validación de los distintos modelos alternativos, seleccionando el más adecuado.
- Predecir los valores correspondientes a los períodos 241-252 comparando con los datos simulados y con las predicciones que se hubieran obtenido al utilizar el proceso generador de los datos que viene dado por $Y_t = (I - 0.5L)(I + 0.5L^{12})\varepsilon_t$.

DISPONIBILIDADES LIQUIDAS (M.M.)

NUMERO DE OBSERVACIONES UTILIZADAS 171

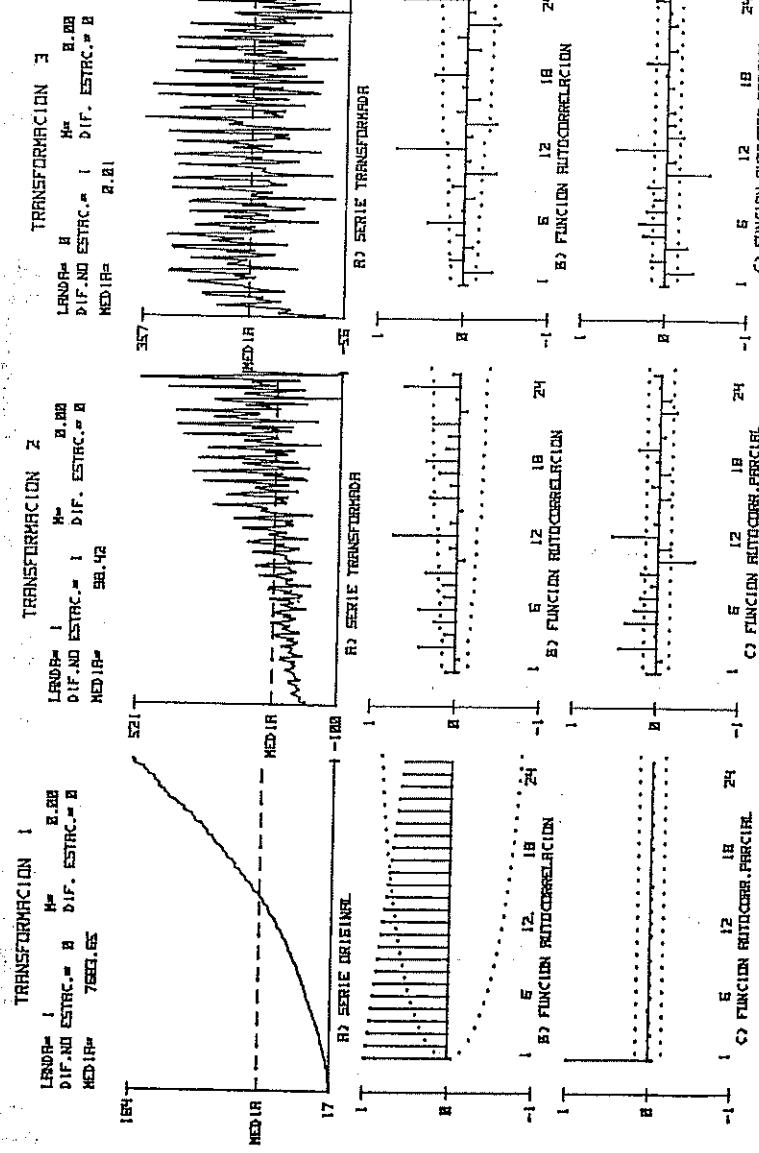


Fig. 8.5

DISPONIBILIDADES LIQUIDAS (M.M.)

NUMERO DE OBSERVACIONES UTILIZADAS 171

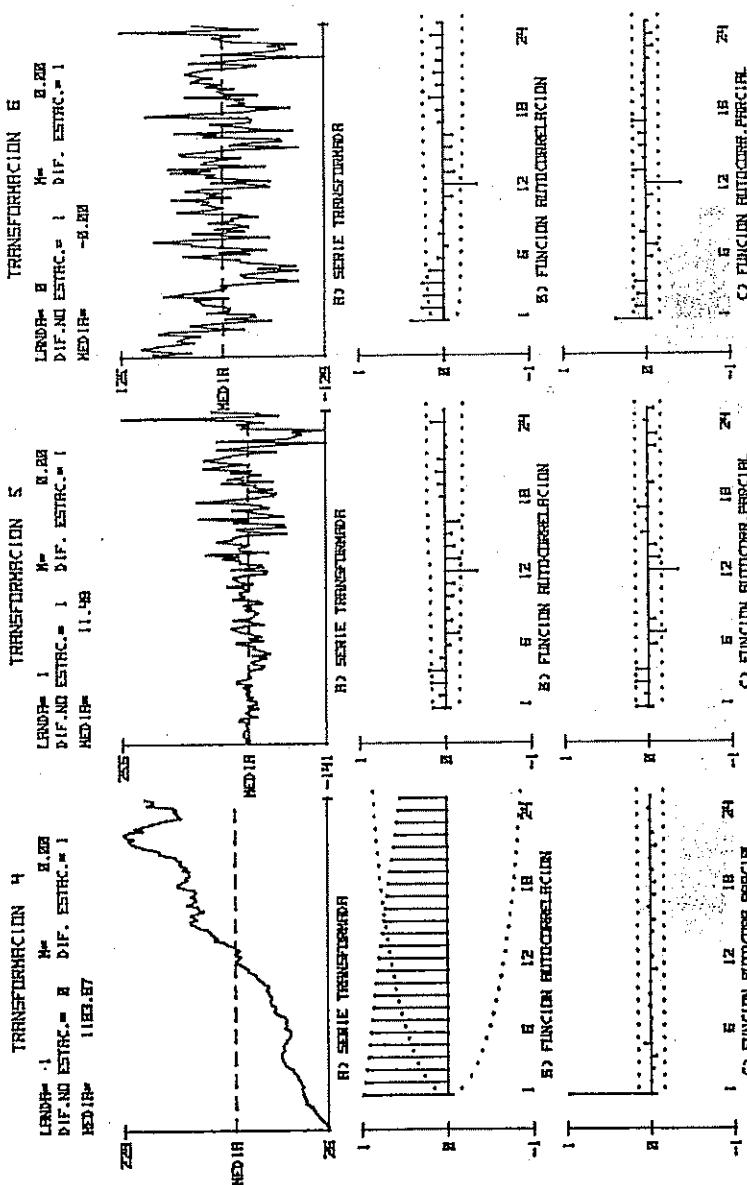
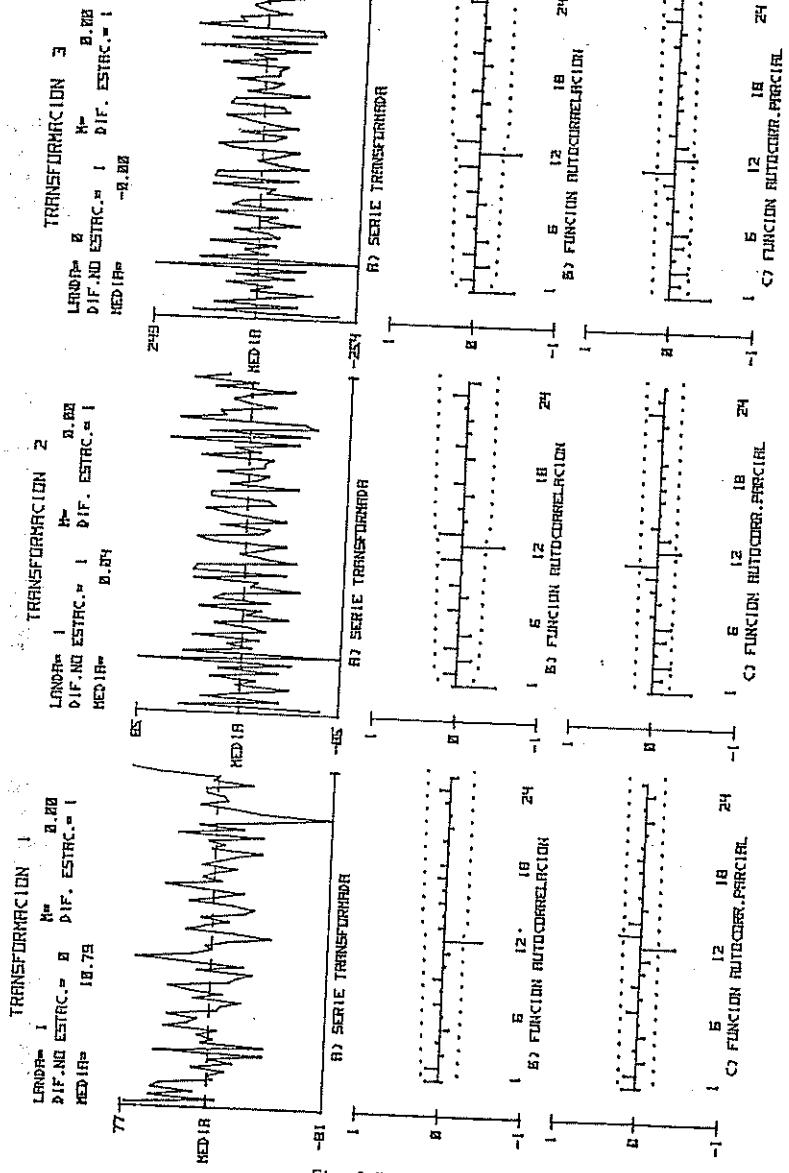


Fig. 8.6.

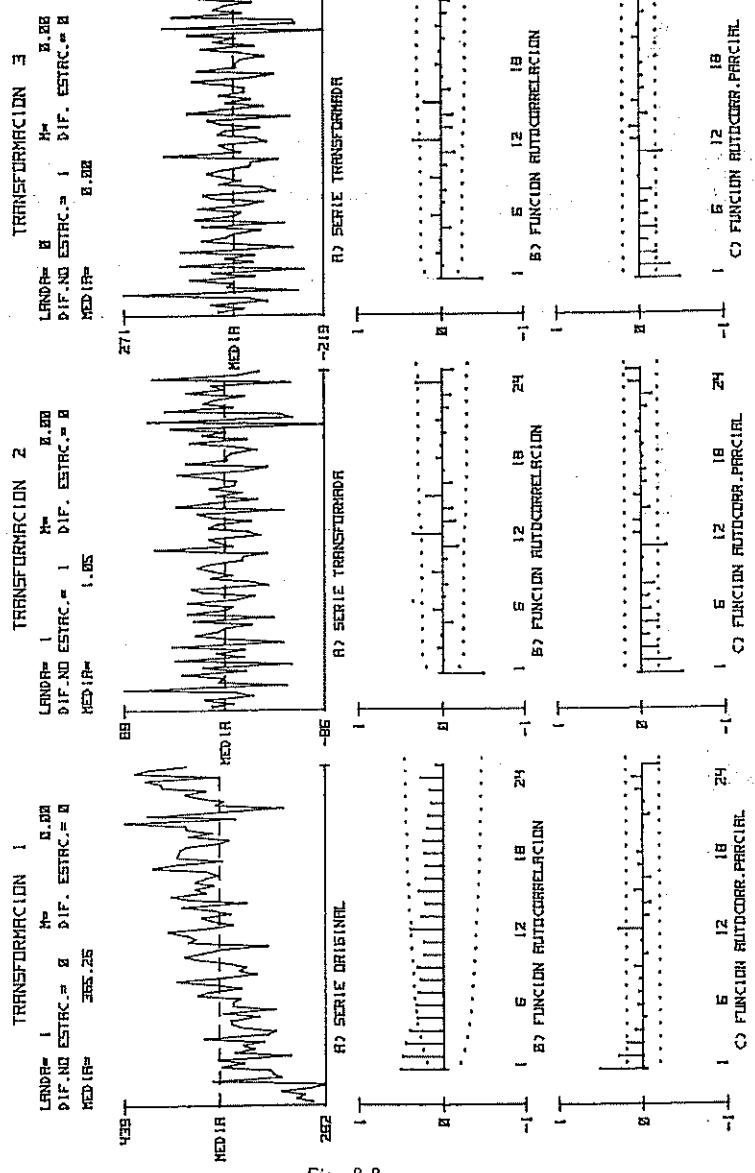
CONSUMO DE ENERGIA ELECTRICA EN EL SECTOR INDUSTRIAL

NUMERO DE OBSERVACIONES UTILIZADAS 96



CONSUMO DE ENERGIA ELECTRICA EN EL SECTOR INDUSTRIAL

NUMERO DE OBSERVACIONES UTILIZADAS 96



- 9] A partir de la serie simulada J del apéndice 4, y utilizando las observaciones 1-120, se pide
- Identificar el modelo —o modelos— adecuados para representar la serie.
 - Estimar los modelos alternativos considerados en el apartado a).
 - Efectuar la validación de los distintos modelos alternativos, seleccionando el más adecuado.
 - Predecir los valores correspondientes a los períodos 121-124 comparando con los datos simulados y con las predicciones que se hubieran obtenido al utilizar el proceso generador de los datos que viene dado por

$$(1 + 0,4L)(1 - L^4)Y_t = (1 - 0,7L^4)\varepsilon_t$$

APÉNDICE 1

Operador de diferencias y de retardos

El tratamiento de los modelos ARIMA queda facilitado en gran medida con la utilización de los operadores matemáticos de diferencias y de retardos. A continuación se examinan las características de estos operadores.

A1.1. Operador de diferencias Δ

Si se aplica el operador de diferencias Δ a una variable referida a un momento de tiempo, Y_t por ejemplo, tiene el siguiente efecto:

$$\Delta Y_t = Y_t - Y_{t-1} \quad (\text{A1-1})$$

Es decir, se obtiene la diferencia entre el valor referido al momento actual y el valor que toma la variable en el período inmediatamente anterior. Abreviadamente, se dice que se han calculado primeras diferencias.

Si se aplica el operador Δ a ΔY_t se obtendrá:

$$\Delta[\Delta Y_t] = \Delta[Y_t - Y_{t-1}] = \Delta Y_t - \Delta Y_{t-1} = [Y_t - Y_{t-1}] - [Y_{t-1} - Y_{t-2}] \quad (\text{A1-2})$$

A este resultado se le denomina segundas diferencias, y se expresa mediante la anteposición de Δ^2 a la variable Y_t .

En general, el exponente a que se eleva Δ indica el número de veces consecutivas que se ha aplicado el operador Δ . A este exponente se le denomina orden del retardo.

Ejemplo

t	Y_t	ΔY_t	$\Delta^2 Y_t$	$\Delta^3 Y_t$	$\Delta^4 Y_t$
1	423	—	—	—	—
2	432	9	—	—	—
3	477	45	36	—	—
4	450	-27	-72	-108	—
5	430	-20	7	79	187

Conviene tener en cuenta que en cada diferenciación, como puede verse en el ejemplo anterior, se pierde un dato.

En la aplicación del operador Δ se toman diferencias entre cada par de valores consecutivos, bien sean de la serie original o bien de los obtenidos en unas diferencias anteriores. A veces, sin embargo, interesa tomar diferencias no entre datos correspondientes a períodos consecutivos, sino con un intervalo mayor; por ejemplo, puede convenir tomar diferencias con un intervalo de 12 períodos cuando se están manejando datos mensuales. Este tipo de diferencias suele estar conectado con el problema de ARIMA se suele utilizar la terminología: «diferencias estacionales de período s », en datos mensuales, etc... La « s » aparece en este caso como subíndice del símbolo Δ , es decir, Δ_s , para distinguirlos de las diferencias entre períodos consecutivos.

A1.2. Operador de retardos L

El operador de retardos L aplicado a una variable Y_t significa que se *retarda* en 1 período el subíndice a que va referida la variable, es decir:

$$LY_t = Y_{t-1} \quad (\text{A1-3})$$

El significado de $L^2 Y_t$ es inmediato, ya que:

$$L^2 Y_t = L[LY_t] = [Y_{t-1}] = Y_{t-2}$$

En general:

$$L^k Y_t = Y_{t-k} \quad (\text{A1-4})$$

El operador L puede manipularse como si fuera una cantidad algebraica cualquiera. Así,

$$\begin{aligned} L(aY_t) &= aLY_t = aY_{t-1} \\ L^k L^s(Y_t) &= L^{k+s}(Y_t) = Y_{t-k-s} \end{aligned}$$

El operador de retardos identidad es $L^0 = I = 1$. La aplicación del operador de retardos unidad no altera el período de referencia.

$$L^0(Y_t) = Y_t$$

La relación entre el operador Δ y el operador L es inmediata, ya que $\Delta = 1 - L$. Por ejemplo: $\Delta Y_t = Y_t - Y_{t-1} = Y_t - LY_t = (1 - L)Y_t$.

Análogamente, se puede establecer la relación entre el operador estacional de diferencias Δ_s y el operador de retardos, ya que:

$$\Delta_s = 1 - L^s$$

Por ejemplo, $\Delta_{12} Y_t = Y_t - Y_{t-12} = (1 - L^{12})Y_t$.

El operador L se puede utilizar para expresar un modelo con retardos. Sea, por ejemplo, el siguiente modelo

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p} = \varepsilon_t \quad (\text{A1-5})$$

si se aplica el operador L se obtendrá

$$[1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p] Y_t = \varepsilon_t$$

A la expresión que aparece entre corchetes se le denomina operador polinomial de retardos, y se designa abreviadamente por $\phi(L)$. La letra que precede a (L) hace referencia a la utilizada para designar los coeficientes que figuran en el operador polinomial.

De forma compacta el modelo (A1-5) quedaría expresado así

$$\phi(L) Y_t = \varepsilon_t \quad (\text{A1-6})$$

Al operador polinomial igualado a 0, es decir $\phi(L) = 0$, se le denomina ecuación polinomial. En el apéndice A2 pueden verse aplicaciones.

Si se considera el modelo de medias móviles MA(∞),

$$Y_t = \sum_{j=0}^{\infty} [\phi L]^j \varepsilon_t \quad (\text{A1-7})$$

en caso de que $|\phi| < 1$, y tomando $|L| \leq 1$, el factor que precede a ε_t es la suma de infinitos términos de una progresión geométrica convergente. Por tanto:

$$Y_t = \frac{1}{1 - \phi L} \varepsilon_t \quad (\text{A1-8})$$

Eliminando el denominador, queda

$$(1 - \phi L) Y_t = \varepsilon_t \quad (\text{A1-9})$$

que es un modelo AR(1).

Así pues, se puede considerar que el modelo del tipo (A1-7) es una simple transformación del modelo AR(1), siempre que $|\phi| < 1$.

APÉNDICE 2

Ecuaciones ordinarias en diferencias finitas:
lineales y con coeficientes constantes

Una ecuación ordinaria en diferencias finitas es una ecuación que contiene una o más diferencias de una función desconocida cuyo argumento es el tiempo. Así,

$$f(Y_t, \Delta Y_t, \Delta^2 Y_t, \dots, \Delta^n Y_t) = 0 \quad (\text{A2-1})$$

donde Δ es el operador diferencias.

El calificativo de ordinaria hace referencia a que la función desconocida depende de un sólo argumento: el tiempo. Si se consideran varios argumentos pasarían a denominarse ecuaciones en diferencias finitas parciales.

Dentro de las ecuaciones ordinarias en diferencias finitas, la exposición se va a restringir a aquellas que son lineales y con coeficientes constantes¹. Por ejemplo:

$$a_0 \Delta^p Y_t + a_1 \Delta^{p-1} Y_t + \dots + a_{p-1} \Delta Y_t + a_p Y_t = h(t) \quad (\text{A2-2})$$

donde $h(t)$ es una función conocida que depende del tiempo.

La ecuación anterior es lineal, porque en cada sumando aparece un solo operador de diferencias. Por la definición dada, los coeficientes a_i son constantes.

Una ecuación en diferencias finitas también se puede expresar utilizando variables retardadas en el tiempo.

Así, la ecuación

$$a_0 Y_t + a_1 \Delta Y_t + a_2 \Delta^2 Y_t = h(t)$$

se puede expresar así:

$$a_0 Y_t + a_1 [Y_t - Y_{t-1}] + a_2 [(Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})] = h(t)$$

es decir:

$$(a_0 + a_1 + a_2) Y_t + (-a_1 - 2a_2) Y_{t-1} + a_2 Y_{t-2} = h(t)$$

¹ En lo sucesivo, la expresión de ecuaciones en diferencias finitas se referirá únicamente a ecuaciones ordinarias en diferencias finitas, lineales y con coeficientes constantes.

Haciendo:

$$\begin{aligned} b_0 &= a_0 + a_1 + a_2 \\ b_1 &= -a_1 - 2a_2 \\ b_2 &= a_2 \end{aligned}$$

La ecuación en diferencias quedaría finalmente así

$$b_0 Y_t + b_1 Y_{t-1} + b_2 Y_{t-2} = h(t)$$

En lo sucesivo se utilizará preferentemente esta formulación.

El orden de una ecuación en diferencias es el orden de la diferencia más elevada —o alternativamente, la variable de mayor retardo— que aparece en la ecuación. Así, una ecuación de orden p , utilizando la formulación de variables retardadas, vendría dada por

$$b_0 Y_t + b_1 Y_{t-1} + \dots + b_p Y_{t-p} = h(t) \quad (\text{A2-3})$$

Si dividimos ambos miembros por b_0 y denominamos $\phi_i = -b_i/b_0$, siendo $i = 1, 2, \dots, p$, obtendríamos una formulación normalizada, que estaría caracterizada porque el coeficiente de Y_t es igual a la unidad.

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = g(t) \quad (\text{A2-4})$$

donde $g(t) = h(t)/b_0$.

Utilizando el operador de retardos L , la anterior ecuación se puede expresar de la siguiente forma:

$$[1 - \phi_1 L - \dots - \phi_p L^p] Y_t = g(t)$$

o, de forma más compacta,

$$\phi(L) Y_t = g(t) \quad (\text{A2-5})$$

La ecuación (A2-5) es una ecuación no homogénea, por la presencia en el segundo miembro de la función $g(t)$, que no contiene términos actuales o retardados de Y_t . Si hacemos 0 al segundo miembro, obtenemos la siguiente ecuación homogénea:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = 0 \quad (\text{A2-6})$$

o, alternativamente:

$$\phi(L) Y_t = 0$$

En este momento, la cuestión que se plantea es resolver la ecuación (A2-5), es decir, obtener una expresión de Y_t en función del tiempo que cumpla dicha ecuación.

Como se verá no existe una solución única. Interesa, por tanto, deducir aquella expresión que recoja o englobe a todas las posibles soluciones, y a esta expresión se le denomina: *solución completa* (Y_t^c).

Como puede siempre comprobarse, la solución (más) completa es el resultado de sumar dos expresiones: la solución (más) general de la ecuación homogénea (Y_t^h) y una solución particular cualquiera de la ecuación no homogénea (Y_t^p).

$$Y_t^c = Y_t^h + Y_t^p \quad (A2-7)$$

A2.1. Solución general de la ecuación homogénea

La solución general de la ecuación homogénea debe ser una función del tiempo $f(t)$, que satisface la ecuación (A2-6). Con el calificativo de general se quiere indicar que debe ser válida cualquiera que sean las condiciones iniciales del proceso.

Ensayamos con una solución del tipo:

$$Y_t = \lambda^t$$

Para ver si es una solución adecuada, la sustituimos en (A2-6), es decir,

$$\lambda^t - \phi_1 \lambda^{t-1} - \dots - \phi_p \lambda^{t-p} = 0$$

Sacando factor común a λ^{t-p} obtendremos la siguiente expresión:

$$\lambda^{t-p} [\lambda^p - \phi_1 \lambda^{p-1} - \dots - \phi_p] = 0$$

Una solución trivial que satisface la anterior ecuación es $\lambda = 0$. Dejando aparte esta solución inicial, podemos obtener p raíces ($\lambda_1, \lambda_2, \dots, \lambda_p$) de la ecuación:

$$\lambda^p - \phi_1 \lambda^{p-1} - \dots - \phi_p = 0 \quad (A2-8)$$

que también satisfarán la ecuación en diferencias.

La ecuación (A2-8) recibe la denominación de ecuación característica.

Alternativamente, las raíces se pueden obtener a partir de la ecuación polinomial que, según se indicó en el apéndice A1, se define así:

$$\begin{aligned} \phi(L) &= 0 \\ 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p &= 0 \end{aligned} \quad (A2-9)$$

En la resolución de la anterior ecuación se obtendrán sus raíces L_1, L_2, \dots, L_p . Se puede comprobar que las raíces L_i son justamente los valores reciprocos de las raíces λ_i . En efecto, si L_i es una raíz de (A2-9) se verifica que

$$1 - \phi_1 L_i - \phi_2 L_i^2 - \dots - \phi_p L_i^p = 0 \quad (A2-10)$$

Por otra parte, según se ha enunciado, $\lambda_1 = \frac{1}{L_1}$ será una solución de (A2-8) con lo que

$$\left[\frac{1}{L_1} \right]^p - \phi_1 \left[\frac{1}{L_1} \right]^{p-1} - \dots - \phi_p = 0 \quad (A2-11)$$

El cumplimiento de (A2-11) se comprueba multiplicando ambos miembros por L_1^p , con lo que se obtiene precisamente la igualdad (A2-10).

Hemos visto que si λ_i es una raíz de la ecuación característica (A2-8), entonces $Y_i = A_i \lambda_i^t$ es una solución a la ecuación homogénea de diferencias. A esta solución se le denominará *solución básica*.

Veamos ahora dos importantes teoremas que se utilizan en la obtención de la solución general. El primer teorema dice que si λ_i^t es una solución y A_i es una constante arbitraria, entonces $Y_i = A_i \lambda_i^t$ es también una solución. En efecto, sustituyendo este valor en (A2-6), y sacando factor común a $A_i \lambda_i^{t-p}$, se obtiene que:

$$A_i \lambda_i^{t-p} [\lambda_i^p - \phi_1 \lambda_i^{p-1} - \dots - \phi_p] = 0$$

ya que la expresión entre corchetes es igual a 0 de acuerdo con la ecuación (A2-8).

En consecuencia, $A_1 \lambda_1^t, A_2 \lambda_2^t, \dots, A_m \lambda_m^t$ donde A_1, A_2, \dots, A_m son constantes arbitrarias, son también soluciones de la ecuación homogénea. Es decir, son soluciones aquellas que resultan de multiplicar una solución básica (λ_i^t) por una constante arbitraria.

El segundo teorema se enuncia diciendo que si λ_i^t y λ_j^t son soluciones a la ecuación homogénea, una combinación lineal / de ambas como

$$A_i \lambda_i^t + A_j \lambda_j^t$$

también es una solución de la ecuación homogénea.

Esta propiedad se verifica sustituyendo la anterior expresión en la ecuación homogénea. Después de sacar el factor común a $A_i \lambda_i^{t-p}$ por una parte y a $A_j \lambda_j^{t-p}$ por otra, se obtiene

$$A_i \lambda_i^{t-p} [\lambda_i^p - \dots - \phi_p] + A_j \lambda_j^{t-p} [\lambda_j^p - \dots - \phi_p] = 0$$

ya que las dos expresiones entre corchetes son igual a 0.

Este segundo teorema permite establecer que una combinación lineal de dos soluciones también es una solución. Precisamente la solución general de la ecuación homogénea se construye tomando las p soluciones básicas, multiplicadas cada una por una constante arbitraria. Es decir:

$$Y^h = Y_t = A_1 \lambda_1^t + A_2 \lambda_2^t + \dots + A_p \lambda_p^t \quad (A2-12)$$

que es la solución (más) general de la ecuación homogénea.

Es importante analizar el comportamiento de Y_t cuando t crece indefinidamente.

De acuerdo con la anterior expresión puede verse que si se verifica que $|\lambda_i| < 1 \forall i$, entonces cuando $t \rightarrow \infty$, sucede que $Y_t \rightarrow 0$ con independencia del valor que se asigne a las constantes arbitrarias. Se dice entonces que el sistema es estable. En el contexto de los procesos estocásticos se utiliza el término de estacionario. Una condición necesaria y suficiente de estabilidad o estacionariedad es por tanto que $|\lambda_i| < 1 \forall i$.

La solución (A2-12) es válida para cualquier valor que se asigne a las constantes arbitrarias. En aplicaciones concretas, se suelen asignar valores a las constantes arbitrarias obtenidos a partir de unas condiciones iniciales. Así, si se supone que conocemos los valores Y_0, Y_1, \dots, Y_{p-1} que corresponden a los períodos $t=0, t=1, \dots, t=p-1$ respectivamente, se puede plantear el siguiente sistema de $p-1$ ecuaciones:

$$\begin{aligned} Y_0 &= A_1 + A_2 + \dots + A_p \\ Y_1 &= A_1 \lambda_1 + A_2 \lambda_2 + \dots + A_p \lambda_p \\ &\vdots \\ Y_{p-1} &= A_1 \lambda_1^{p-1} + A_2 \lambda_2^{p-1} + \dots + A_p \lambda_p^{p-1} \end{aligned} \quad | \quad (A2-13)$$

En el sistema anterior, las incógnitas serían A_1, A_2, \dots, A_p , mientras que Y_0, Y_1, \dots, Y_{p-1} y $\lambda_1, \lambda_2, \dots, \lambda_p$ serían datos.

A continuación se va a examinar las características de las ecuaciones homogéneas en diferencias de 1.^o y 2.^o orden. Conociendo estos dos tipos de ecuaciones, no existe ningún problema conceptual para pasar a ecuaciones de orden superior.

Ecuaciones de primer orden: condiciones de estacionariedad

Una ecuación homogénea de primer orden presenta la siguiente forma:

$$Y_t - \phi_1 Y_{t-1} = 0 \quad (A2-14)$$

La solución de la anterior ecuación será del tipo λ^t . Por tanto

$$\lambda^t - \phi_1 \lambda^{t-1} = 0$$

o, dividiendo por λ^{t-1}

$$\lambda - \phi_1 = 0 \quad (A2-15)$$

Es decir, la raíz obtenida será

$$\lambda = \phi_1$$

La solución de la ecuación homogénea será:

$$Y_t = A \lambda^t = A \phi_1^t \quad (A2-16)$$

En la ecuación anterior para $t=0$, se tiene que $Y_0 = A$. Tomando como dato este valor de Y_0 , se obtiene el siguiente modelo:

$$Y_t = Y_0 \phi_1^t$$

Esta ecuación describe la trayectoria temporal de Y_t a partir del valor inicial dado. La condición necesaria y suficiente para que el modelo sea estacionario, o estable, es que $|\phi_1| < 1$. En este caso se verificaría que

$$\lim_{t \rightarrow \infty} Y_t = \lim_{t \rightarrow \infty} Y_0 \phi_1^t = 0$$

Alternativamente se puede utilizar la ecuación polinomial de retardos:

$$1 - \phi_1 L = 0 \quad (A2-17)$$

que daría una raíz:

$$L_1 = \frac{1}{\phi_1}$$

Si utilizamos (A2-17), la condición de estacionariedad o estabilidad se formularía diciendo que $|L_1| > 1$.

Si $0 < \phi_1 < 1$, entonces Y_t seguirá una trayectoria de decrecimiento exponencial a partir de la condición inicial. Cuando $-1 < \phi_1 < 0$, entonces Y_t decrecerá en valor absoluto de forma exponencial aunque con alternancia de los signos.

Ecuaciones de 2.^o orden: Condiciones de estacionariedad

Una ecuación homogénea de 2.^o orden toma la siguiente forma:

$$\begin{aligned} Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} &= 0 \\ (1 - \phi_1 L - \phi_2 L^2) Y_t &= 0 \end{aligned} \quad (A2-18)$$

La ecuación característica que corresponde a esta ecuación en diferencias es:

$$\lambda^2 - \phi_1 \lambda - \phi_2 = 0 \quad (A2-19)$$

Resolviendo esta ecuación de 2.^o grado se obtienen las raíces λ_1 y λ_2 .

$$\lambda_1, \lambda_2 = \frac{\phi_1}{2} \pm \frac{\sqrt{\phi_1^2 + 4\phi_2}}{2}$$

Denominemos $D = \sqrt{\phi_1^2 + 4\phi_2}$. Se utilizará también el término discriminante para designar D .

Según que el valor del discriminante sea positivo, nulo o negativo, se presentarán tres casos con características distintas.

a) $D > 0$

Las dos raíces λ_1 y λ_2 son distintas y reales. La solución de la ecuación homogénea tomará la siguiente forma

$$Y_t = A_1 \lambda_1^t + A_2 \lambda_2^t \quad (\text{A2-20})$$

b) $D = 0$

En este caso las dos raíces son reales e iguales $\lambda_1 = \lambda_2 = \phi_1/2$, es decir, se obtiene una raíz doble, que designaremos por λ_u . En esta circunstancia, es fácil comprobar que no sólo es solución λ_u^t , sino que también la expresión $t\lambda_u^t$, según se verá más adelante, cumple con la ecuación homogénea, siendo solución de la misma. Por lo tanto, la expresión (más) general que resuelve dicha ecuación es la combinación lineal de ambas

$$Y_t^H = A_1 \lambda_u^t + A_2 t \lambda_u^t$$

Comprobemos ahora que $t\lambda_u^t$ es una solución a la ecuación (A2-18). Después de eliminar el factor común λ_u^{t-2} se obtiene

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

Ahora bien, en este caso como $D = 0$ se verifica que $\phi_1^2 = -4\phi_2$. Sustituyendo ϕ_1^2 en la anterior expresión se comprueba fácilmente que se hace igual a 0.

c) $D < 0$

Se obtienen dos raíces complejas conjugadas.

Haciendo

$$\alpha = \frac{\phi_1}{2} ; \theta = \sqrt{-(\phi_1^2 + 4\phi_2)/2}$$

las dos raíces serían:

$$\lambda_1, \lambda_2 = \alpha \pm i\theta$$

donde $i = \sqrt{-1}$.

La solución general a la ecuación homogénea tomará la forma

$$Y_t = A_1 [\alpha + i\theta]^t + A_2 [\alpha - i\theta]^t \quad (\text{A2-21})$$

donde A_1 y A_2 son números complejos conjugados arbitrarios que se pueden determinar a partir de unas condiciones iniciales. Suponiendo que la variable Y toma los valores Y_0 e Y_1 para $t = 0$ y $t = 1$ respectivamente, se puede plantear el siguiente sistema:

$$\begin{aligned} Y_0 &= A_1 + A_2, \\ Y_1 &= A_1 [\alpha + i\theta] + A_2 [\alpha - i\theta] \end{aligned} \quad (\text{A2-22})$$

Como A_1 y A_2 son los números complejos conjugados, se pueden expresar de la siguiente forma:

$$A_1 = \frac{1}{2}B + \frac{1}{2}Ci \quad A_2 = \frac{1}{2}B - \frac{1}{2}Ci \quad (\text{A2-23})$$

donde el factor $\frac{1}{2}$ se ha introducido por conveniencia.

Sustituyendo estos valores en (A2-22) y resolviendo para B y C se obtiene

$$\begin{aligned} B &= Y_0 \\ C &= [\alpha Y_0 - Y_1]/\theta \end{aligned} \quad (\text{A2-24})$$

Puede observarse directamente de (A2-23) que

$$A_1 + A_2 = B \quad A_1 - A_2 = C i \quad (\text{A2-25})$$

Alternativamente, la solución (A2-21) se puede expresar en forma polar, efectuando las siguientes transformaciones:

$$a = r \cos \omega$$

$$\theta = r \operatorname{sen} \omega$$

Fácilmente se ve que:

$$a^2 + \theta^2 = r^2$$

El número r es precisamente el *módulo* o valor absoluto del número complejo.

Si las transformaciones de a y θ se sustituyen en (A2-21) la ecuación homogénea vendrá dada por la siguiente expresión:

$$\begin{aligned} Y_t &= A_1 [r \cos \omega + ir \operatorname{sen} \omega]^t + A_2 [r \cos \omega - ir \operatorname{sen} \omega]^t \\ &= A_1 r^t [\cos \omega + i \operatorname{sen} \omega]^t + A_2 r^t [\cos \omega - i \operatorname{sen} \omega]^t \end{aligned} \quad (\text{A2-26})$$

Aplicando el teorema de Moivre a la anterior expresión quedará

$$\begin{aligned} Y_t &= r^t [A_1 (\cos \omega t + i \operatorname{sen} \omega t) + A_2 (\cos \omega t - i \operatorname{sen} \omega t)] \\ &= r^t [(A_1 + A_2) \cos \omega t + (A_1 - A_2) i \operatorname{sen} \omega t] \end{aligned} \quad (\text{A2-27})$$

Teniendo en cuenta (A2-25) a la expresión anterior quedará

$$Y_t = r^t [B \cos \omega t - C \operatorname{sen} \omega t] \quad (\text{A2-28})$$

Alternativamente, se puede obtener otra expresión de la solución. Si se hace

$$B = A \cos \varepsilon \quad C = A \operatorname{sen} \omega$$

se puede expresar la solución general de la ecuación en diferencias de la siguiente forma:

$$Y_t = A r^t [\cos \varepsilon \cos \omega t - \operatorname{sen} \varepsilon \operatorname{sen} \omega t] \quad (\text{A2-29})$$

Teniendo en cuenta la propiedad trigonométrica de que

$$\cos \varepsilon \cos t - \operatorname{sen} \varepsilon \operatorname{sen} t = \operatorname{sen}(t - \varepsilon)$$

quedará finalmente

$$Y_t = Ar^t \operatorname{sen}(\omega t - \varepsilon) \quad (\text{A2-30})$$

La solución anterior comprende una única función trigonométrica que es fácilmente representable. Es por tanto una función cíclica del tiempo, en la que se pueden distinguir los siguientes elementos:

— *Amplitud: Ar^t* : Es el factor de escala que se aplica a la función seno para obtener el valor de Y_t . En este caso la amplitud es variable ya que depende de t . La condición necesaria y suficiente para que la ecuación sea estacionaria, o estable, es que $r < 1$. En ese caso, (A2-30) representará una oscilación sinusoidal que se irá atenuando de forma geométrica.

— *Frecuencia (o frecuencia angular): ω* : La frecuencia se mide en radianes e indica el número de ciclos que hay por unidad de tiempo. Recuérdese que un ciclo se completa cada vez que se alcanza 2π , o un múltiplo de 2π , radianes. El *periodo P_ε* es el recíproco de la frecuencia, y señala el número de períodos de tiempo que se necesitan para completar un ciclo. El período viene dado por tanto por la siguiente expresión:

$$P_\varepsilon = \frac{2\pi}{\omega} \quad (\text{A2-31})$$

Así, por ejemplo, una función en que $\omega = 2\pi/6$ significaría que el número de ciclos por unidad de tiempo es $1/6$. Alternativamente, $P_\varepsilon = 2\pi/(2\pi/6) = 6$, es decir, cada 6 períodos de tiempo se completa un ciclo completo.

— *Fase: ε* : Se mide también en radianes. Con la fase se determina para un momento dado del tiempo la situación del ciclo. Así, por ejemplo, si $\varepsilon = \pi/2$, en el momento $t = 0$, $Y_t = A(\operatorname{sen} \frac{\pi}{2})$.

En la figura A2.1, se ha representado la solución (A2-30) para $A = 100$, $r = 0.8$, $\omega = 2\pi/6$ y $\varepsilon = \pi/2$.

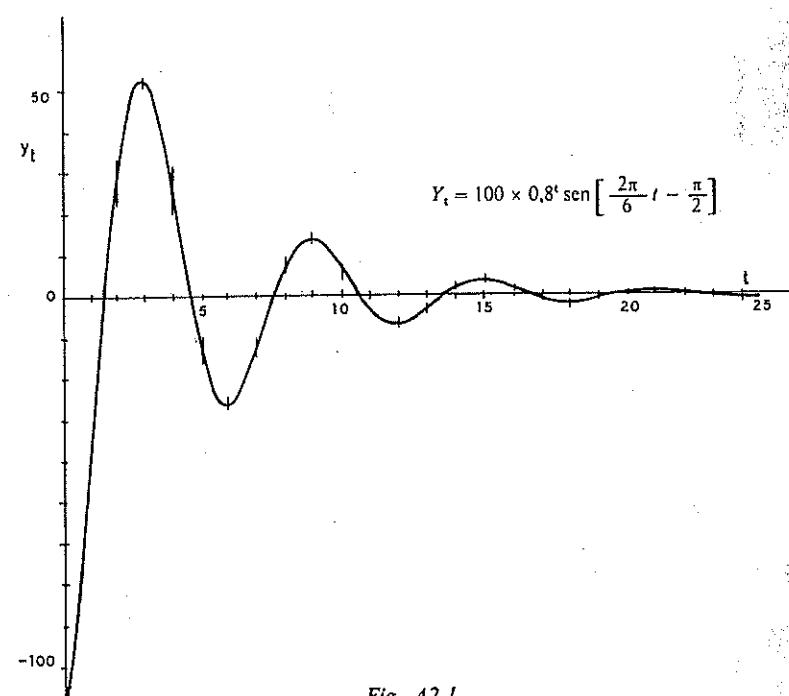


Fig. A2.1.

En una ecuación homogénea de 2.^o orden cuando las raíces son reales y distintas, la condición de estacionariedad implica que $|\lambda_1| < 1$ y $|\lambda_2| < 1$. Cuando es una raíz única debe verificarse que $|\lambda_u| < 1$. Si son raíces imaginarias se puede expresar diciendo que el módulo debe ser inferior a la unidad, es decir $r < 1$.

Como alternativa a la ecuación característica (A2-19), para obtener las raíces, se puede utilizar la ecuación polinomial:

$$1 - \phi_1 L - \phi_2 L^2 = 0$$

De acuerdo con en (A2-11), las raíces de la ecuación anterior, L_1 y L_2 , son recíprocas de las que se obtienen en (A2-19). Si las raíces son reales, se pueden representar utilizando solamente el eje de abcisas. Si son complejas, se necesita el eje de ordenadas para representar la parte imaginaria. Así, si L_1 y L_2 son raíces complejas conjugadas tomarán la forma:

$$L_1 = \alpha + i\theta$$

$$L_2 = \alpha - i\theta$$

Su representación gráfica para $\alpha = 0,8$ y $\theta = 1,2$ puede verse en la figura A2.2.

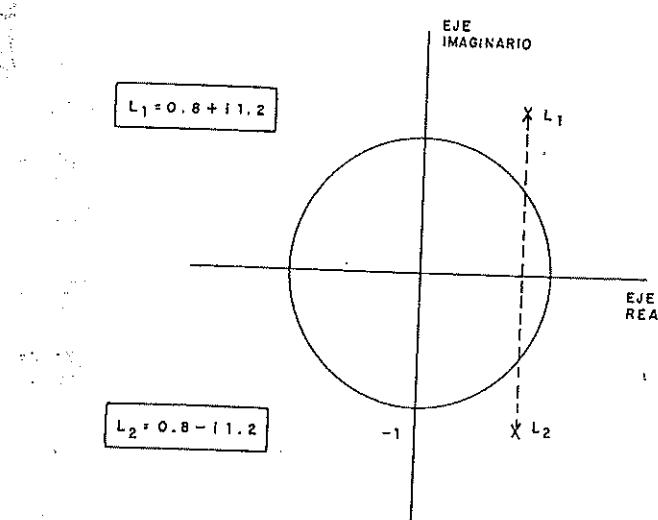


Fig. A2.2.

El módulo de las raíces de este ejemplo es $r = \sqrt{\alpha^2 + \theta^2} = 1,44$. Por tanto la ecuación es estacionaria, ya que este módulo es el recíproco del que hubiéramos obtenido si se hubiera utilizado la ecuación característica.

En el contexto de procesos estocásticos, y cuando se utiliza el operador polinomial de retardos, la condición de estacionariedad se suele formular diciendo que las raíces deben caer fuera del círculo unidad, según queda ilustrado en la figura A2.2. Lógicamente, en términos de las raíces λ_1 y λ_2 de la ecuación característica, deben caer dentro del círculo unidad para que la ecuación sea estacionaria.

Las condiciones de estacionariedad pueden definirse directamente a partir de los coeficientes de ϕ_1 y ϕ_2 , ya que después de realizadas ciertas manipulaciones algebraicas:

cas se puede demostrar que cuando las raíces L_1 —en una ecuación polinomial de 2.º orden— caen fuera del círculo unidad, se verifican las siguientes desigualdades:

$$\begin{aligned}\phi_1 + \phi_2 &< 1 \\ -\phi_1 + \phi_2 &< 1 \\ \phi_2 &> -1\end{aligned}\tag{A2-32}$$

Si se sustituye en las anteriores desigualdades el signo < por el signo =, se obtienen tres líneas rectas, que definen la zona de estacionariedad, que está constituida por todos los puntos en el interior del triángulo, representado en la figura A2.3. La zona sombreada corresponde a raíces estacionarias complejas.

En las ecuaciones homogéneas en diferencias de orden superior a 2, pueden aparecer raíces reales o complejas. Naturalmente, cada raíz compleja irá siempre acompañada de su conjugada. En el caso de que una raíz λ_i aparezca repetida n veces se toman como soluciones básicas a $\lambda_i, t\lambda_i, t^2\lambda_i, \dots, t^{n-1}\lambda_i$.

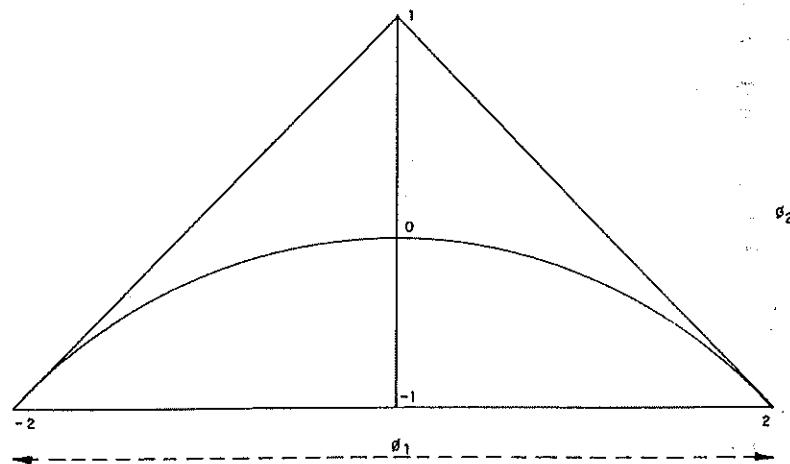


Fig. A2.3.

A2.2. Solución particular

La solución particular de la ecuación (A2-5) se busca analizando las características de la parte no homogénea $g(t)$ y asumiendo que Y_t se comporta de forma análoga a $g(t)$.

En el caso de que la ecuación homogénea sea estacionaria o estable, puede demostrarse que siempre existe un operador $\Psi(L)$ tal que si

$$\phi(L)Y_t = g(t)$$

entonces

$$Y_t = \Psi(L)g(t) \quad (\text{A2-33})$$

definiéndose a $\Psi(L)$ como operador inverso de $\phi(L)$:

$$\Psi(L) = \frac{1}{\phi(L)}$$

La formulación (A2-33) ofrece una solución analítica general, pero debido a que la definición de operador inverso no nos suministra un procedimiento útil para proceder a su cálculo, nos podemos encontrar en la práctica con dificultades operativas.

A continuación se van a considerar varios casos:

a) *Una constante: $g(t) = c$*

Se considera que Y_t se mantiene constante a lo largo del tiempo en un valor μ . Por tanto, sustituyendo Y_{t-1} por μ y haciendo $g(t) = c$ se obtiene la siguiente expresión para (A2-4)

$$\mu - \phi_1\mu - \phi_2\mu - \dots - \phi_p\mu = c$$

Despejando queda

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p} \quad (\text{A2-34})$$

Si se hubiera despejado directamente Y_t en (A2-5) se tendría que

$$Y_t = \frac{1}{\phi(L)}c = \frac{1}{\phi(1)}c = \frac{c}{1 - \phi_1 - \dots - \phi_p} \quad (\text{A2-35})$$

Mediante $\phi(1)$ se designa el valor del operador polinomial en el que se ha sustituido L por 1, ya que al aplicar L a una constante, queda ésta inalterada.

La solución completa será:

$$Y_t = Y_t^H + Y_t^P = A_1\lambda_1^t + \dots + A_p\lambda_p^t + \frac{c}{1 - \phi_1 - \dots - \phi_p} \quad (\text{A2-36})$$

Si la ecuación es estacionaria, cuando $t \rightarrow \infty$ se verificará que:

$$\lim_{t \rightarrow \infty} Y_t = \frac{c}{1 - \phi_1 - \dots - \phi_p}$$

es decir, Y_t tiende a igualarse a la constante obtenida en la solución particular.

b) *Una función lineal de tiempo: $g(t) = \gamma + \delta t$*

Se ensaya una solución del tipo

$$Y_t = a + \beta t \quad (\text{A2-37})$$

Para determinar los valores de a y β , se sustituye Y_t por $a + \beta t$ y $g(t)$ por $\gamma + \delta t$ en la ecuación (A2-4)

$$[a + \beta t] - \phi_1[a + \beta(t-1)] - \dots - \phi_p[a + \beta(t-p)] = \gamma + \delta t$$

Reagrupando los términos quedará:

$$[a(1 - \phi_1 - \dots - \phi_p) + \beta(\phi_1 + 2\phi_2 + \dots + p\phi_p) + \beta(1 - \phi_1 - \dots - \phi_p)]t = \gamma + \delta t$$

Igualando término a término podemos obtener los valores de β y a .

$$\beta = \frac{\delta}{1 - \phi_1 - \dots - \phi_p} \quad (\text{A2-38})$$

$$a = \frac{\gamma - \beta(\phi_1 + 2\phi_2 + \dots + p\phi_p)}{1 - \phi_1 - \dots - \phi_p}$$

La aplicación directa del inverso del operador polinomial $\gamma + \delta t$ es más complicada que en el caso a). Con objeto de simplificar los cálculos, se va a considerar esta aplicación en una ecuación homogénea de primer orden.

Sea

$$Y_t - \phi_1 Y_{t-1} = \gamma + \delta t \quad (\text{A2-39})$$

La solución particular será:

$$\begin{aligned} Y_t &= \frac{1}{1 - \phi_1 L} [\gamma + \delta t] = \\ &= \sum_{j=0}^{\infty} (\phi_1 L)^j [\gamma + \delta t] \\ &= \frac{\gamma}{1 - \phi_1} + \delta[t + \phi_1(t-1) + \phi_1^2(t-2) + \dots] \\ &= \frac{\gamma}{1 - \phi_1} - \delta[\phi_1 + 2\phi_1^2 + 3\phi_1^3 + \dots] + \delta[1 + \phi_1 + \phi_1^2 + \dots]t \end{aligned}$$

Evaluando las expresiones entre corchetes que aparecen en el segundo término se obtiene finalmente que

$$Y_t = \frac{\gamma}{1 - \phi_1} - \delta \frac{\phi_1}{(1 - \phi_1)^2} + \delta \frac{\phi_1}{1 - \phi_1} t \quad (\text{A2-40})$$

que es el mismo resultado que el obtenido al sustituir en (A2-37) los valores de α y β de (A2-38) particularizando para el caso de $p = 1$.

Para esta misma ecuación de primer orden, la solución completa de acuerdo con (A2-16) y (A2-37) será

$$Y_t = Y_0 \phi_1^t + \alpha + \beta t \quad (\text{A2-41})$$

El primer componente decrece geométricamente tomando Y_0 como valor inicial. El segundo componente por el contrario hace que en cada período Y_t experimente un incremento en valor absoluto igual a β . Merced a este segundo componente Y_t crecerá indefinidamente.

c) Una función exponencial del tiempo $g(t) = \beta d^t$

Se ensaya una función del tipo:

$$Y_t = \alpha d^t \quad (\text{A2-42})$$

Sustituyendo este valor en la ecuación de diferencias no homogénea (A2-4) se obtiene que:

$$[\alpha d^t - \phi_1 \alpha d^{t-1} - \dots - \phi_p \alpha d^{t-p}] = \beta d^t$$

Sacando factor común a αd^t

$$\alpha d^t [1 - \phi_1 d^{-1} - \dots - \phi_p d^{-p}] = \beta d^t$$

²Sea $S = \phi_1 + 2\phi_1^2 + 3\phi_1^3 + \dots$

Puede verse que

$$S(1 - \phi_1) = \phi_1 + \phi_1^2 + \phi_1^3 + \dots = S'$$

Ahora bien:

$$S' = \frac{\phi_1}{1 - \phi_1}$$

En consecuencia:

$$S = \frac{\phi_1}{(1 - \phi_1)^2}$$

de donde:

$$\alpha = \frac{\beta}{1 - \phi_1 d^{-1} - \dots - \phi_p d^{-p}} \quad (\text{A2-43})$$

Alternativamente, si se aplica directamente el operador de retardos a una ecuación de primer orden se obtendría que

$$\begin{aligned} Y_t &= \frac{1}{1 - \phi_1 L} \beta d^t \\ &= \beta \Sigma(\phi_1 L)^t d^t \\ &= \beta \Sigma [d^t + \phi_1 d^{t-1} + \phi_1^2 d^{t-2} + \dots] \\ &= \beta \frac{1}{1 - \phi_1 d^{-1}} d^t \end{aligned} \quad (\text{A2-44})$$

La solución completa para este último modelo, teniendo en cuenta (A2-16) y (A2-44) será:

$$Y_t = Y_0 \phi_1^t + \frac{\beta}{1 - \phi_1 d^{-1}} d^t \quad (\text{A2-45})$$

Si $|d| < 1$, entonces $Y_t \rightarrow 0$, al crecer t indefinidamente.

d) Una variable aleatoria tipo ruido blanco $g(t) = \epsilon_t$

La solución particular vendrá dada por:

$$Y_t = \frac{1}{\phi_1(L)} \epsilon_t \quad (\text{A2-46})$$

En el caso de que $\phi_1(L) = 1 - \phi_1 L$, la solución particular, suponiendo que el proceso arranca en $-N$ ³, será:

$$Y_t = \sum_{j=0}^{t-N-1} \phi_1^j \epsilon_{t-j} \quad (\text{A2-47})$$

La solución completa vendrá dada por:

$$Y_t = Y_{-N} \phi_1^{t+N} + \sum_{j=0}^{t-N-1} \phi_1^j \epsilon_{t-j} \quad (\text{A2-48})$$

³Para el momento $-N$, la condición inicial sería Y_{-N} . Para este período y para todos los anteriores ϵ_t tomará valor 0.

Puede verse que la solución completa tiene dos componentes: un componente sistemático que viene dado por la solución general y que depende de las condiciones iniciales — Y_{-N} en este caso— y un componente aleatorio, que es justamente la solución particular.

Si se hace tender N hacia infinito —es decir, si el proceso arranca desde $-\infty$ —, y el modelo es estacionario, la solución completa queda reducida a:

$$Y_t = \sum_{j=0}^{\infty} \phi_j^t \varepsilon_{t-j} \quad (\text{A2-49})$$

En el otro extremo si se toma como condición inicial el valor Y_{t-1} , la solución completa sería:

$$Y_t = Y_{t-1} \phi_1 + \varepsilon_t \quad (\text{A2-50})$$

En este caso, la solución particular vendría dada únicamente por ε_t .

APÉNDICE 3

Simulación con modelos ARIMA por el método de Montecarlo

Cuando se generan valores con un modelo por el método de Montecarlo se invierten los términos en relación a cuando se actúa sobre datos observados en el mundo real. En este último caso, el problema consiste en inferir, a partir de los datos observados, los parámetros del hipotético modelo generador. Naturalmente las inferencias realizadas son de carácter probabilístico, dado que se supone que el modelo es de naturaleza estocástica.

En un experimento por el método de Montecarlo se parte del conocimiento de una estructura especificada¹ de un modelo determinado. A partir de esta estructura de datos de variables conocidas, y de valores generados para los elementos estocásticos del modelo, se calculan los valores de la variable o variables endógenas del modelo. La denominación de Montecarlo proviene precisamente del nombre de la ciudad que es conocida especialmente por el casino en el que se practican juegos de azar.

Las previsiones de períodos futuros con modelos ARIMA se obtienen también en base al conocimiento (estimado) de una estructura, pero no pueden ser consideradas como ejercicios de simulación porque la intervención de las variables aleatorias siempre se contempla bajo el supuesto de que toman, en cada momento, su valor medio. En una simulación por el método de Montecarlo lo que se busca es que las variables aleatorias tomen valores al azar, de acuerdo con la distribución de probabilidad que les corresponda, según la «suerte» del ejercicio de generación de cada valor; y en esto radica, precisamente, la diferencia.

Las series simuladas son útiles para experimentar los métodos de identificación y de estimación de parámetros. Por otro lado, una reflexión sobre el significado de los ejercicios de simulación ayuda a entender mejor el carácter estocástico de las series, al comprender que el mismo modelo puede generar series diferentes, para un mismo período. Esto se traduce en que, al observar una serie histórica, hemos de advertir conscientemente que es una de las muchas series que podía haber producido el modelo que identifiquemos.

¹ Se entiende que una estructura está especificada cuando se conocen los valores de todos los parámetros del modelo, incluidos los correspondientes a las distribuciones probabilísticas de los elementos estocásticos.

Así pues, el método Montecarlo está asociado a la generación de números «al azar». Estos números se requieren en la simulación de modelos estocásticos: en el epígrafe A3.1 se analiza la generación de números aleatorios uniformes que constituyen, por decirlo de una forma gráfica, la materia prima que se utiliza en los experimentos de Montecarlo.

En los modelos de series temporales y en los modelos econométricos se suele asumir que los elementos estocásticos siguen una distribución normal. Un problema que plantea es el paso de unos números aleatorios uniformes a números aleatorios normales, o, en general, de cualquier otra distribución. En los epígrafes A3.2. y A3.3. se aborda esta cuestión.

Finalmente, en el epígrafe A3.4. se ilustra la generación de series a partir de modelos ARIMA.

A3.1 Generación de números aleatorios uniformes

Como introducción al tema, considérese la extracción de bolas de un sorteo de la lotería nacional. Antes de iniciarse el sorteo, todos los números —supóngase que son 100.000— tienen idéntica probabilidad de que se les adjudique el premio gordo. La probabilidad que tiene cada uno de obtener dicho premio es, por tanto, de 1/100.000. La figura A3.1a) refleja esto.

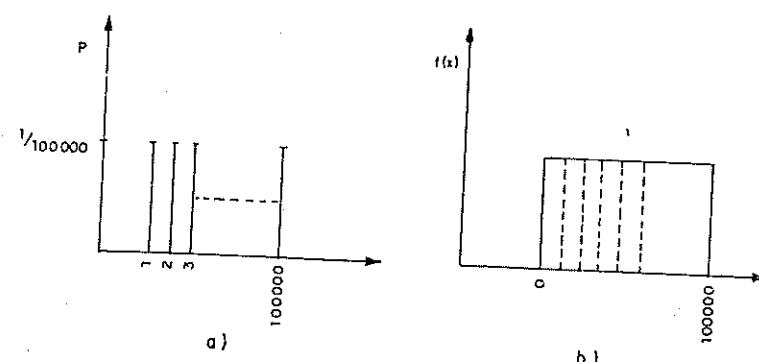


Fig. A3.1.

La extracción de una bola, o, en términos más técnicos, de un número aleatorio, puede considerarse que es equivalente a la selección de un intervalo de una distribución uniforme en la que el recorrido total se ha dividido en intervalos de igual tamaño. Obsérvese que la distribución uniforme es de carácter continuo, según puede verse en la figura A3.1 b), a diferencia del carácter discreto de la distribución representada en la figura A3.1 a).

Debe tenerse en cuenta que para realizar una extracción aleatoria de un número

uniforme debería dejarse que un puntero se situara al azar sobre el recorrido de la distribución. Sin embargo, al convertir esa posición en un número, ineludiblemente se pasa a una elección discreta; es decir, lo que se está seleccionando es en realidad un intervalo. Hecha esta precisión, a efectos prácticos señalaremos que se utilizan como equivalentes las expresiones «generación de números aleatorios» y «generación de números aleatorios uniformes».

La obtención de una secuencia de números aleatorios uniformes o, simplemente, números aleatorios puede efectuarse por muy distintos procedimientos. Estos procedimientos de acuerdo a la metodología empleada pueden agruparse en tres grupos:

- Métodos manuales.
- Consulta de tablas.
- Generación por ordenador.

Los métodos manuales consisten en realizar directamente experimentos u observar los resultados de experimentos aleatorios, como puede ser el lanzamiento de un dado, el juego de la ruleta, o la extracción de números de un bombo. Naturalmente, para que a partir de estas extracciones se pueda obtener una secuencia de números aleatorios uniformes, es preciso que el experimento se haya realizado en condiciones adecuadas. El problema de estos métodos es que son muy laboriosos.

Por el procedimiento de la consulta de tablas de números aleatorios se podría obtener una secuencia de estos números con más facilidad que por el procedimiento anterior. Así, son muy conocidas las tablas elaboradas por la RAND Corporation (1955) y, en España, las elaboradas por los estadísticos Royo y Ferrer (1955). Estos últimos autores las construyeron en base a los resultados obtenidos en la Lotería Nacional. Estas tablas pueden ser un instrumento útil para ciertos tipos de aplicaciones como pueden ser la selección de unidades en un muestreo aleatorio. En cambio, en la simulación con modelos estocásticos su utilidad es mucho más reducida por razones operativas, ya que en estos casos se requieren miles o millones de números aleatorios, lo que implica la utilización de un soporte informático.

El tercer grupo de procedimientos incluye los métodos de generación de números aleatorios mediante ordenador. En los métodos manuales se generan los números aleatorios por extracciones al azar. Por el contrario, cuando se emplea un ordenador digital se aplican fórmulas deterministas, lo que hace que a los números generados no se les pueda denominar en rigor estrictamente aleatorios. Algunos autores utilizan por ello la denominación seudoaleatorios. Ahora bien, existen fórmulas deterministas que permiten obtener secuencias de números que tienen unas propiedades que no se diferencian de los estrictamente aleatorios. Entre los diferentes métodos desarrollados debe destacarse el método congruencial multiplicativo, que es el que se examina a continuación.

El método congruencial multiplicativo se basa en el concepto de congruencia. Se dice que dos números enteros a y b son congruentes respecto a un módulo m cuando al ser divididos por m se obtiene el mismo resto.

La relación de congruencia se expresa por la notación:

$$a \equiv b \pmod{m} \quad (\text{A3-1})$$

que se lee « a es congruente a b en módulo m ».

De la relación de congruencia se deduce que $(b - a)$ es divisible por m . En efecto, por ser ambos números congruentes se cumple que

$$a = k_1 m + r$$

$$b = k_2 m + r$$

Restando la segunda igualdad de la primera se verifica que $b - a = (k_2 - k_1)m$.

Como puede observarse dado un número entero b , se pueden encontrar infinitos números que sean congruentes en módulo m . Si se hace $k_1 = 0$ se obtiene que a es igual precisamente al resto r . Por tanto, dado un número b se puede encontrar un número congruente a que verifique que:

$$0 \leq a < m \quad (\text{A3-2})$$

Cuando se impone la restricción (A3-2), la condición de congruencia (A3-1) se lee « a es igual a b reducido a módulo m ». En lo sucesivo interpretaremos la relación de congruencia en este sentido.

En el método congruencial multiplicativo los números aleatorios se generan mediante la siguiente fórmula recursiva

$$c_i \equiv \beta c_{i-1} \pmod{m} \quad i = 1, 2, 3, \dots \quad (\text{A3-3})$$

Para aplicar la fórmula (A3-3) se necesita determinar c_0 , que es el número inicial o raíz —«root» o «seed» en inglés—, el módulo m y el factor β .

Ejemplo.

Para $c_0 = 3$, $m = 16$ y $\beta = 5$ se obtiene:

$$c_0 = 3$$

$$c_1 \equiv 5 \times 3 \pmod{16} \quad c_1 = 15$$

$$c_2 \equiv 5 \times 15 \pmod{16} \quad c_2 = 11$$

$$c_3 \equiv 5 \times 11 \pmod{16} \quad c_3 = 7$$

$$c_4 \equiv 5 \times 7 \pmod{16} \quad c_4 = 3$$

Como puede verse, $c_4 = c_0$, con lo que $c_5 = c_1$, $c_6 = c_2$, etc. Así pues, con los parámetros seleccionados se ha generado un ciclo de sólo 4 números distintos. Para evitar el problema de un ciclo tan corto se toma un número m suficientemente elevado. Así en los ordenadores binarios que son los más conocidos se selecciona $m = 2^b$, donde b es el número de bits, que tiene una palabra en el ordenador utilizado. En muchos ordenadores b es igual a 32. Fijado m de esta forma, se puede obtener un ciclo integrado por $h = 2^{b-2}$ números aleatorios. En el ejemplo numérico se tiene que $m = 2^4$, con lo que $h = 2^2 = 4$ que es justamente el resultado que hemos obtenido. Por otra parte, β se elige próximo a $2^{b/2}$ según la fórmula $\beta = 8t \pm 3$, donde t es un número entero positivo. El número inicial c_0 debe ser primo en relación a 2^b ; este requisito lo cumple cualquier número entero impar.

Una vez calculada la secuencia de números c_0, c_1, c_2, \dots dividiendo cada m por uno se obtiene una secuencia de números aleatorios uniformes en el intervalo $(0,1)$, ya que éste es el intervalo usual. Aunque conociendo las fórmulas aplicadas se puede calcular de antemano cuál es el valor del número que ocupa el lugar i -ésimo, los números de la secuencia se comportan como si estuvieran distribuidos uniforme e independientemente cuando los parámetros están elegidos adecuadamente. Con objeto de poder determinar la «calidad» de los números aleatorios, se han diseñado para tal finalidad una batería de tests. Para más detalles sobre generación de números aleatorios puede consultarse a Naylor y otros (1966) y a Hernández y Asiaín (1977).

Casi todos los microordenadores llevan incorporados un procedimiento RND —del inglés random, que significa aleatorio— que permite generar directamente números aleatorios uniformes comprendidos en el intervalo $(0, 1)$.

A3.2. Generación de números aleatorios de distribuciones no uniformes

Como ya se ha indicado, los números aleatorios uniformes, y específicamente los tipificados $(0, 1)$, constituyen la materia prima que se utiliza en la generación de otros números aleatorios.

Veamos a continuación un procedimiento general que permite convertir un número aleatorio uniforme $(0, 1)$ en un número de otra distribución.

En la figura A3.2 se ha representado una función de distribución genérica de la variable aleatoria. De acuerdo con la definición de la función de distribución, para un valor de la variable x —p. ej., x_0 — se verifica que

$$F(x_0) = P(X \leq x_0) = r_0 \quad (\text{A3-4})$$

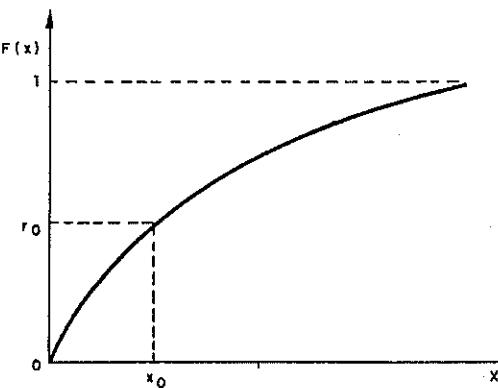


Fig. A3.2.

SIMULACIÓN CON MODELOS ARIMA POR EL MÉTODO DE MONTECARLO

Ahora bien, considerando la función inversa se puede establecer que

$$x_0 = F^{-1}(r_0) \quad (\text{A3-5})$$

De esta forma se determina el valor que toma la variable aleatoria, dado un valor de la función de distribución. Ahora bien, como la función de distribución varía entre 0 y 1, la selección de un número aleatorio uniforme $(0, 1)$ permitirá obtener —a través de (A3-5)— un número aleatorio de la distribución considerada.

El método que acabamos de describir se puede utilizar para obtener números aleatorios normales, o de cualquier otra distribución. Unicamente se puede plantear problemas de carácter numérico en caso de que la función inversa no sea fácil de determinar.

A3.3. Generación de números aleatorios de la distribución normal

Además del anterior método general, existen otros métodos particulares para obtener números aleatorios normales. A continuación se expone un método basado en el Teorema Central del Límite.

Como es sabido, el Teorema Central del Límite permite establecer que la distribución de una variable aleatoria que es la suma de n variables aleatorias x_i distribuidas independiente e idénticamente con media α y varianza σ^2 , cuando n es muy grande, se approxima asintóticamente a una distribución normal con media $n\alpha$ y varianza $n\sigma^2$. Es decir,

$$\sum_{i=1}^n x_i \xrightarrow{n \rightarrow \infty} N[n\alpha, \sigma\sqrt{n}] \quad (\text{A3-6})$$

o análogamente,

$$\frac{\sum_{i=1}^n x_i - n\alpha}{\sigma\sqrt{n}} \xrightarrow{n \rightarrow \infty} N(0, 1) \quad (\text{A3-7})$$

En una distribución uniforme $(0, 1)$ se tiene que:

$$\begin{aligned} \alpha &= 0,5 \\ \sigma^2 &= 1/12 \end{aligned} \quad (\text{A3-8})$$

Por tanto, si dispone de n números uniformes independientes $U(0, 1)$, aplicando (A3-7) se tiene que

$$\frac{\sum_{i=1}^n U_i(0, 1) - 0,5n}{\sqrt{n}/\sqrt{12}} \xrightarrow{n \rightarrow \infty} N(0, 1) \quad (\text{A3-9})$$

Ahora bien, ¿qué valor mínimo debe tener n al objeto de conseguir una buena

GENERACIÓN DE NÚMEROS ALEATORIOS DE DISTRIBUCIÓN NO UNIFORMES

aproximación a una distribución normal? Por los estudios realizados se ha visto que para $n \geq 10$ se consiguen unas aproximaciones razonables buenas. El valor $n = 12$ se utiliza bastante porque se simplifica la fórmula (A3-9). En efecto,

$$\sum_{i=1}^n U_i(0, 1) - 6 \approx N(0, 1) \quad (\text{A3-10})$$

Una vez obtenido un número $N(0, 1)$, el paso a un número $N(\alpha, \sigma)$ es inmediato, ya que

$$N(\alpha, \sigma) = \alpha + \sigma N(0, 1) \quad (\text{A3-11})$$

En el cuadro A3.1 figuran 600 números aleatorios uniformes $U(0, 1)$ y 50 números aleatorios $N(0, 1)$ calculados según la fórmula (A3-10). Los números aleatorios uniformes se han obtenido mediante la instrucción RND en un microordenador de 16 bits.

A3.4. Generación de series con modelos ARIMA

Dada una estructura de un modelo ARIMA, para generar una serie simulada únicamente se requiere disponer de un generador de números aleatorios normales y de unos criterios para fijar los valores iniciales. En lo que respecta al papel que juegan los valores iniciales debe tenerse en cuenta lo expuesto en el capítulo 5. En el ejemplo que se verá a continuación el valor inicial se determina igualando a 0 el valor medio teórico de la variable.

Ejemplo

Sea la estructura

$$\begin{aligned} Y_t &= 9 + 0,7 Y_{t-1} + \varepsilon_t \\ \varepsilon &\sim N(0, 4) \end{aligned}$$

Tomando $Y_0 = 30$, y los 20 primeros números aleatorios $N(0, 1)$ del cuadro A3.1, se obtiene una serie simulada de Y_t según puede verse en el cuadro A3.2.

Si en vez de utilizar los 20 primeros números aleatorios del cuadro A3.1 hubiéramos empleado, por ejemplo, los 20 siguientes, habríamos detenido una serie de valores distintos, para el mismo período, del mismo modelo.

A partir de series simuladas se puede ver cómo se comportan ciertos estadísticos en muestras finitas. Así, en las figuras 4.7 a 4.12 se representa la FACE y la FACPE correspondientes a series simuladas de distinto tamaño. A medida que aumenta el tamaño de la muestra es claro que estos estadísticos se aproximan a sus valores teóricos. Sin embargo, es ilustrativo examinar las distorsiones que aparecen cuando las muestras son pequeñas.

Cuadro A3.1. Números aleatorios uniformes y $N(0, 1)$

Números aleatorios uniformes													$N(0, 1)$
0,678	0,382	0,419	0,504	0,744	0,510	0,595	0,388	0,889	0,811	0,875	0,979	1,773	
0,405	0,315	0,039	0,697	0,415	0,862	0,991	0,073	0,782	0,893	0,200	0,273	-0,054	
0,783	0,913	0,082	0,934	0,694	0,852	0,929	0,428	0,733	0,339	0,343	0,074	1,605	
0,828	0,631	0,257	0,011	0,590	0,284	0,109	0,918	0,996	0,654	0,121	0,926	0,326	
0,378	0,939	0,076	0,563	0,816	0,314	0,397	0,109	0,912	0,236	0,411	0,231	-0,617	
0,435	0,095	0,938	0,157	0,890	0,730	0,568	0,454	0,790	0,051	0,824	0,714	0,645	
0,062	0,181	0,052	0,903	0,432	0,981	0,985	0,783	0,115	0,665	0,580	0,651	0,389	
0,179	0,839	0,420	0,067	0,118	0,866	0,036	0,449	0,452	0,486	0,906	0,734	-0,449	
0,435	0,191	0,280	0,572	0,737	0,989	0,437	0,838	0,107	0,076	0,204	0,830	-0,303	
0,241	0,731	0,122	0,888	0,503	0,964	0,404	0,879	0,633	0,551	0,118	0,280	0,315	
0,772	0,949	0,498	0,661	0,992	0,073	0,709	0,861	0,041	0,572	0,174	0,973	1,275	
0,500	0,746	0,469	0,536	0,639	0,266	0,727	0,195	0,805	0,770	0,272	0,407	0,331	
0,666	0,330	0,547	0,825	0,524	0,743	0,777	0,718	0,795	0,484	0,129	0,646	1,184	
0,886	0,007	0,711	0,570	0,701	0,558	0,556	0,732	0,524	0,900	0,619	0,740	1,504	
0,949	0,123	0,139	0,990	0,555	0,905	0,407	0,594	0,740	0,121	0,672	0,044	0,242	
0,849	0,668	0,452	0,822	0,422	0,804	0,184	0,685	0,744	0,343	0,306	0,344	0,621	
0,122	0,501	0,555	0,039	0,925	0,294	0,226	0,870	0,395	0,210	0,372	0,738	-0,753	
0,254	0,434	0,081	0,460	0,110	0,556	0,716	0,137	0,560	0,683	0,462	0,124	-1,423	
0,206	0,572	0,632	0,142	0,020	0,710	0,007	0,186	0,131	0,087	0,198	0,511	-2,599	
0,761	0,390	0,906	0,906	0,334	0,743	0,803	0,300	0,685	0,018	0,319	0,116	0,282	
0,493	0,311	0,601	0,224	0,066	0,102	0,479	0,483	0,599	0,137	0,447	0,627	-1,433	
0,102	0,635	0,746	0,651	0,534	0,962	0,257	0,720	0,878	0,707	0,789	0,918	1,899	
0,619	0,883	0,020	0,789	0,931	0,906	0,232	0,139	0,870	0,568	0,392	0,403	0,753	
0,190	0,786	0,144	0,171	0,681	0,430	0,465	0,445	0,886	0,804	0,762	0,144	-0,093	
0,027	0,530	0,512	0,634	0,911	0,008	0,775	0,748	0,083	0,595	0,302	0,806	-0,071	
0,670	0,529	0,520	0,066	0,458	0,258	0,806	0,895	0,399	0,411	0,918	0,998	0,957	
0,165	0,360	0,943	0,962	0,771	0,345	0,238	0,018	0,155	0,575	0,174	0,461	-0,834	
0,776	0,399	0,669	0,575	0,991	0,191	0,605	0,537	0,448	0,702	0,094	0,854	0,840	
0,600	0,392	0,622	0,793	0,230	0,581	0,480	0,275	0,255	0,252	0,611	0,948	-0,062	
0,332	0,953	0,627	0,049	0,574	0,082	0,868	0,166	0,928	0,869	0,261	0,753	0,403	
0,914	0,842	0,773	0,259	0,665	0,646	0,567	0,206	0,848	0,842	0,835	0,360	1,757	
0,057	0,270	0,815	0,364	0,711	0,608	0,927	0,253	0,881	0,641	0,742	0,838	1,114	
0,663	0,392	0,649	0,642	0,696	0,610	0,092	0,994	0,906	0,356	0,448	0,171	0,619	
0,990	0,287	0,970	0,090	0,555	0,630	0,141	0,466	0,644	0,322	0,586	0,498	0,177	
0,409	0,909	0,652	0,031	0,258	0,774	0,929	0,098	0,666	0,145	0,542	0,807	0,219	
0,493	0,872	0,190	0,722	0,559	0,081	0,030	0,036	0,870	0,372	0,008	0,540	-1,128	
0,305	0,573	0,581	0,618	0,436	0,039	0,015	0,848	0,331	0,095	0,945	0,479	-0,735	
0,877	0,805	0,412	0,693	0,051	0,220	0,691	0,138	0,406	0,563	0,962	0,132	-0,043	
0,806	0,757	0,590	0,977	0,676	0,113	0,968	0,075	0,969	0,058	0,687	0,583	1,258	
0,288	0,923	0,491	0,702	0,631	0,716	0,329	0,310	0,366	0,224	0,016	0,343	-0,662	
0,095	0,750	0,402	0,915	0,221	0,621	0,370	0,583	0,595	0,610	0,988	0,052	0,202	
0,117	0,741	0,878	0,713	0,487	0,976	0,839	0,338	0,274	0,784	0,390	0,814	1,351	
0,229	0,491	0,227	0,929	0,691	0,190	0,718	0,246	0,900	0,197	0,262	0,245	-0,674	
0,828	0,572	0,873	0,350	0,200	0,839	0,783	0,118	0,201	0,450	0,566	0,023	-0,197	
0,543	0,903	0,157	0,299	0,735	0,970	0,946	0,512	0,812	0,132	0,649	0,854	1,513	
0,184	0,607	0,588	0,904	0,659	0,165	0,651	0,772	0,341	0,081	0,440	0,138	-0,468	
0,416	0,033	0,413	0,691	0,812	0,132	0,684	0,119	0,309	0,497	0,744	0,123	-1,027	
0,432	0,685	0,694	0,438	0,091	0,357	0,323	0,040	0,868	0,253	0,422	0,646	-0,752	
0,964	0,412	0,380	0,435	0,738	0,640	0,937	0,986	0,268	0,058	0,111	0,249	0,177	
0,602	0,037	0,809	0,769	0,810	0,841	0,865	0,596	0,515	0,321	0,908	0,610	1,683	

Cuadro A3.2. Obtención de una serie simulada

t	$N(0, 1)$	ϵ_t	Y_t	t	$N(0, 1)$	ϵ_t	Y_t
1	1,773	3,547	33,547	11	1,275	2,549	32,979
2	-0,054	-0,109	32,374	12	0,331	0,662	32,748
3	1,605	3,210	34,872	13	1,184	2,367	34,290
4	0,326	0,651	34,061	14	1,504	3,007	36,010
5	-0,617	-1,233	31,610	15	0,242	0,483	34,690
6	0,645	1,290	32,417	16	0,621	1,241	34,526
7	-0,389	0,779	32,471	17	-0,753	-1,507	31,661
8	-0,449	-0,897	30,833	18	-1,423	-2,845	28,318
9	-0,303	-0,605	29,978	19	-2,599	-5,199	23,624
10	0,315	0,630	30,614	20	0,282	0,564	26,100

También las series simuladas se utilizan para comparar métodos alternativos de estimación de parámetros en los que está garantizado su buen comportamiento en el límite, pero para los que no se dispone de instrumentos analíticos que permitan estudiar su comportamiento en pequeñas muestras.

APÉNDICE 4

Datos

A. Tasa de natalidad por 1.000 habitantes en España. 1900-1980

Años										
1900-1909	33,93	34,87	35,49	36,23	34,12	34,97	33,68	33,25	33,61	33,00
1910-1919	32,60	31,43	31,69	30,49	29,82	30,75	28,98	28,94	29,94	27,78
1920-1929	29,36	30,31	30,34	30,33	29,60	28,93	29,47	27,97	29,01	28,17
1930-1939	28,19	27,43	28,06	27,69	26,20	25,74	24,74	22,59	20,02	16,45
1940-1949	24,37	19,55	20,15	22,86	22,51	23,06	21,43	21,38	23,14	21,52
1950-1959	20,02	19,98	20,71	20,43	19,88	20,44	20,60	21,73	21,79	21,65
1960-1969	21,60	21,13	21,08	21,32	21,98	21,13	20,70	20,81	20,03	19,79
1970-1979	19,50	19,51	19,36	19,21	19,47	18,85	18,83	18,05	17,32	16,22
1980	15,12									

Fuente: Anuario estadístico del INE. Varios años.

B. Renta nacional neta al coste de los factores deflactada.
1940-1980 (en millones de pesetas). Base 1970.

Años					
1940-1944	517.900	524.933	562.186	560.647	574.904
1945-1949	561.707	623.528	634.720	630.753	637.500
1950-1954	669.926	743.021	793.342	804.707	862.423
1955-1959	903.268	979.349	1.000.801	1.053.850	1.016.590
1960-1964	1.017.898	1.138.209	1.260.085	1.386.123	1.478.361
1965-1969	1.571.015	1.692.772	1.783.838	1.883.418	2.048.485
1970-1974	2.142.459	2.268.702	2.480.564	2.693.266	2.856.371
1975-1979	2.875.003	2.961.644	3.050.841	3.128.573	3.125.063
1980	3.160.703				

Fuentes: Para el periodo 1954-1981 se han tomado trabajo Uriel (1983), enlazándose con las series, elaboradas por Alcuide (1976) para el periodo 1940-1954.

C. Serie simulada

Observaciones										
I-10	-76	-260	-264	-271	-276	-216	-215	-230	280	398
11-20	-491	-396	-443	-545	-660	-821	-829	-973	1.034	999
21-30	-935	-1.093	-1.332	-1.498	-1.717	-1.826	-1.674	-1.575	1.565	1.545
31-40	-1.583	-1.552	-1.403	-1.373	-1.250	-933	-873	-989	-1.100	-1.186
41-50	-1.293	-1.277	-1.613	-1.773	-1.863	-1.877	-1.772	-1.741	1.840	1.834
51-60	-1.777	-1.862	-2.009	-2.210	-2.322	-2.340	-2.582	-2.830	-2.907	-2.775
61-70	-2.707	-2.782	-2.845	-2.906	-2.870	-2.962	-3.114	-3.132	-3.046	-3.083
71-80	-3.279	-3.470	-3.422	-3.343	-3.367	-3.450	-3.461	-3.569	-3.662	-3.770
81-90	-3.678	-3.631	-3.787	-3.842	-3.705	-3.722	-3.815	-3.936	-4.208	-4.369
91-100	-4.448	-4.528	-4.625	-4.758	-4.741	-4.627	-4.604	-4.612	-4.647	-4.642
101-110	-4.428	-4.442	-4.552	-4.509	-4.528	-4.486	-4.189	-4.060	-4.005	-4.017

D. Serie simulada

Observaciones										
1-10	-80	-50	223	-234	192	-78	-19	27	-61	-18
11-20	39	131	-279	203	-173	60	106	-273	322	-188
21-30	146	-314	190	-90	-45	142	95	-170	89	-35
31-40	-42	12	6	-13	24	-5	-260	124	-78	52
41-50	-74	168	-512	620	-490	409	-186	28	-98	184
51-60	-121	-54	18	-89	142	-58	-207	190	-26	165
61-70	-211	54	1	-27	113	-237	147	13	41	-154
71-80	-10	23	178	-163	37	-56	106	-199	180	-171
81-90	312	-319	64	95	56	-226	170	-174	-35	147
91-100	-102	52	-60	-5	147	-60	-47	42	-60	86
101-110	133	-357	253	-41	-84	163	110	-279	226	-219
111-120	130	-207	389	-425	240	-95	137	-159	127	-75
121-130	-28	-9	24	110	-24	28	-122	28	-13	106
131-140	-257	145	14	-35	-37	49	-136	106	-100	255
141-150	-55	-71	152	-198	-41	88	72	-89	57	-10
151-160	-55	198	-42	-153	25	158	-190	101	-215	325
161-170	-167	54	12	110	-273	435	-424	213	-42	-40
171-180	76	-7	-24	30	-41	-49	68	-117	214	-246
181-190	4	377	-309	239	-261	198	-283	198	-7	-4
191-200	15	-65	37	-227	62	27	103	-164	136	34
201-210	-7	-30	15	50	-38	156	-174	141	-439	493

E. Serie simulada

Observaciones										
1-10	704	813	828	277	955	928	615	330	1.600	917
11-20	531	910	474	87	148	128	226	1.070	606	1.380
21-30	345	376	378	532	314	1.470	444	132	546	318
31-40	186	988	8.840	3.850	666	394	302	2.040	387	129
41-50	1.200	977	1.280	945	2.370	2.250	6.680	2.820	9.340	1.950
51-60	1.470	6.630	5.050	2.700	816	2.300	1.500	173	966	1.280
61-70	486	163	810	107	135	469	180	789	137	829
71-80	157	364	250	83	139	289	138	332	529	234
81-90	197	699	447	217	323	606	462	328	436	569
91-100	574	222	251	295	152	211	123	145	730	945
101-110	984	470	376	284	110	287	517	357	118	938

F. Índice de producción industrial 1965-1984
Base año 1972 = 100

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
1965	46,534	49,26	51,89	50,54	52,11	50,26	51,48	43,24	51,40	52,85	54,63	55,21
1966	54,99	56,14	60,95	58,22	59,45	58,94	55,37	49,17	59,63	61,21	61,19	63,51
1967	60,04	57,85	61,57	62,81	61,92	62,92	58,83	50,32	59,38	60,20	62,23	62,25
1968	62,78	63,86	64,99	63,84	67,08	64,86	63,22	53,10	64,64	68,91	69,75	71,85
1969	72,91	71,11	78,54	76,53	78,43	76,56	73,48	61,88	76,55	80,49	77,57	78,55
1970	80,15	80,85	79,78	86,03	81,90	83,02	80,94	64,37	80,06	83,04	85,61	82,31
1971	82,44	83,84	88,88	86,29	87,54	89,03	87,25	69,68	86,39	88,47	91,37	91,03
1972	94,33	97,04	103,08	99,06	102,27	105,35	97,70	78,51	102,77	104,77	108,71	106,40
1973	110,28	108,40	115,37	109,39	115,85	117,09	108,13	85,90	111,32	116,73	119,38	114,23
1974	126,10	121,78	126,32	121,18	124,68	124,18	115,35	97,74	119,13	124,74	118,35	112,10
1975	111,78	114,10	119,80	116,44	117,42	116,66	111,19	77,52	120,13	129,47	121,57	118,82
1976	112,15	115,39	124,20	122,69	125,27	122,47	124,11	81,25	128,08	130,57	132,91	127,75
1977	122,97	128,75	140,77	128,47	135,62	128,91	119,52	82,47	133,90	134,00	133,17	134,74
1978	129,98	130,36	133,69	131,67	133,00	137,78	123,85	84,67	136,75	141,72	142,81	131,90
1979	134,78	125,91	137,56	126,56	142,59	136,94	131,02	86,25	133,68	142,02	144,82	128,34
1980	137,66	137,24	139,11	133,53	139,36	133,49	133,02	79,88	135,84	147,39	140,84	133,40
1981	130,37	132,74	140,15	133,32	136,88	136,76	139,54	76,94	133,49	142,30	139,40	132,90
1982	126,20	129,20	142,50	136,60	137,30	133,30	134,50	75,80	139,10	135,70	140,10	131,60
1983	133,50	133,00	144,60	134,90	142,60	138,70	131,60	79,70	140,10	138,60	141,70	138,30
1984	137,40	137,80	143,30									

Fuentes: Números índices de la producción industrial y Boletín de Estadística INE. Elaboración propia.

G. Disponibilidades líquidas: M³. Saldos medios mensuales
Miles de millones de pesetas

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
1970	1.723	1.721	1.734	1.745	1.755	1.778	1.814	1.838	1.852	1.875	1.896	1.940
1971	1.986	2.000	2.024	2.054	2.086	2.129	2.186	2.223	2.252	2.291	2.323	2.394
1973	2.468	2.479	2.513	2.559	2.601	2.640	2.710	2.743	2.768	2.820	2.857	2.943
1973	3.033	3.052	3.103	3.171	3.220	3.292	3.395	3.444	3.476	3.539	3.586	3.678
1974	3.754	3.755	3.806	3.880	3.921	3.971	4.077	4.097	4.110	4.175	4.221	4.335
1975	4.445	4.432	4.492	4.554	4.599	4.674	4.817	4.882	4.915	4.964	5.016	5.169
1976	5.286	5.273	5.327	5.390	5.451	5.532	5.695	5.746	5.771	5.860	5.931	6.111
1977	6.257	6.256	6.331	6.444	6.509	6.619	6.816	6.907	6.953	7.018	7.066	7.289
1978	7.418	7.405	7.506	7.647	7.772	7.914	8.202	8.268	8.342	8.438	8.501	8.749
1979	8.947	8.979	9.119	9.278	9.341	9.500	9.821	9.826	9.842	9.977	10.057	10.410
1980	10.552	10.552	10.694	10.895	10.930	11.075	11.411	11.524	11.590	11.720	11.765	12.073
1981	12.227	12.192	12.334	12.531	12.627	12.737	13.123	13.258	13.313	13.487	13.537	13.943
1982	14.164	14.139	14.347	14.638	14.801	14.956	15.367	15.459	15.582	15.777	15.785	16.050
1983	16.341	16.251	16.380	16.579	16.664	16.680	17.114	17.216	17.371	17.548	17.582	18.103
1984	18.347	18.248	18.456									

Fuente: Boletín Económico, Banco de España.

H. Consumo de energía eléctrica en el sector industrial del País Valenciano (CEEPV)
Millones de KWh.

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
1976	292	304	298	316	292	282	370	317	323	325	363	346
1977	366	309	353	338	346	330	377	327	321	348	355	356
1978	334	362	321	354	354	349	382	373	348	368	398	361
1979	336	349	344	355	380	392	378	363	327	389	384	390
1980	405	375	355	370	379	357	369	394	344	387	403	376
1981	384	370	384	376	367	375	417	399	363	398	397	396
1982	391	364	358	389	371	391	439	353	421	363	315	
1983	368	362	381	392	376	410	411	423	367	431	421	392
1984	383	384	421									

Fuente: Boletín de Coyuntura Industrial de las Cámaras de Comercio Regionales, Valencia.

I. Serie simulada

Observaciones												
1-10	-217	-27	-3	-104	135	-190	110	-8	30	-256		
11-20	-23	38	-206	85	123	-102	58	-90	12	76		
21-30	110	-205	151	142	-253	7	45	-45	-58	113		
31-40	-435	294	-119	80	119	34	-184	93	-29	-112		
41-50	-68	-52	-151	133	-317	78	57	126	-130	-3		
51-60	3	-88	67	-214	167	67	-63	-105	-92	-47		
61-70	116	-69	-36	-34	102	-218	38	-55	226	-174		
71-80	-164	75	191	-182	7	-118	-125	-2	-18	-60		
81-90	51	-115	63	79	2	-68	-24	-2	107	-197		
91-100	16	68	-110	76	273	-131	52	-80	-14	-126		
101-110	329	-299	13	46	39	-18	145	-79	-25	-85		
111-120	72	36	-36	34	36	-7	50	71	-151	-19		
121-130	33	-40	-57	68	-194	64	-46	171	116	7		
131-140	44	-85	-108	15	74	-23	-42	35	-75	253		
141-150	159	-143	5	110	-148	12	-117	151	27	5		
151-160	23	215	-119	211	-169	68	20	-47	-35	125		
161-170	6	19	7	-3	-62	49	59	-118	-112	302		
171-180	-25	126	-104	71	-178	-22	84	-38	84	-105		
181-190	-76	-52	-143	41	28	-79	-62	100	70	-9		
191-200	18	40	7	43	-104	61	-303	130	-17	-70		
201-210	-6	85	63	-32	-104	51	102	-184	-89	188		
211-220	-158	125	15	-307	165	213	-158	52	23	-87		
221-230	26	88	-158	-56	169	-424	163	30	57	97		
231-240	-84	3	-85	90	3	-23	144	-228	240	-8		
241-250	-13	-68	197	-69	-169	160	-108	78	-69	28		
251-252	50	-194										

J. Serie simulada

Observaciones										
1-10	52	2	25	12	18	17	25	78	24	2
11-20	81	24	121	3	21	40	47	1	70	50
21-30	48	2	13	2	93	3	15	32	93	98
31-40	18	2	43	1	34	1	124	64	54	2
41-50	15	85	18	3	13	2	10	2	377	3
51-60	16	1	78	4	26	39	185	2	3	3
61-70	64	2	3	2	34	1	6	2	7	3
71-80	6	21	57	2	15	98	117	38	22	3
81-90	12	59	26	1	140	1	16	1	41	2
91-100	78	12	203	5	20	2	107	3	5	5
101-110	24	1	38	2	60	1	1	1	29	4
111-120	5	5	22	73	3	65	29	3	8	2
121-128	57	100	44	74	91	2	11	61		

Soluciones a ejercicios

CAPÍTULO 3

[1]

- a) Es estacionario puesto que $\phi_1 = |0,8| < 1$.
 b) Por definición todo AR de orden finito es invertible.
 c) De acuerdo con (3-12)

$$\gamma_0 = \frac{\sigma^2}{1 - \phi_1^2} = \frac{2}{1 - 0,8^2} = 5,55$$

Las demás autocovarianzas se obtienen de forma recursiva a partir de (3-14):

$$\gamma_t = \phi_1 \gamma_{t-1}$$

tomando γ_0 como valor inicial.

Por tanto,

$$\begin{aligned}\gamma_1 &= 4,44 \\ \gamma_2 &= 3,56 \\ \gamma_3 &= 2,83 \\ \gamma_4 &= 2,29 \\ \gamma_5 &= 1,82\end{aligned}$$

- d) Los coeficientes de autocorrelación se obtienen directamente de forma recursiva de la fórmula

$$\rho_t = \phi_1 \rho_{t-1}$$

tomando $\rho_0 = 1$ como valor inicial.

Así pues:

$$\begin{aligned}\rho_1 &= 0,800 \\ \rho_2 &= 0,640 \\ \rho_3 &= 0,512 \\ \rho_4 &= 0,410 \\ \rho_5 &= 0,328\end{aligned}$$

- e) De acuerdo con (3-19) los coeficientes ψ_i coincidirán con los coeficientes p_i de idéntico retraso. Por tanto,

$$\begin{aligned}\psi_1 &= 0,800 \\ \psi_2 &= 0,640 \\ &\vdots\end{aligned}$$

- [2]** a) Es estacionario porque todos los modelos MA de un orden finito lo son.
 b) Es invertible porque $|0,9| < 1$.
 c) De acuerdo con (3-53), (3-54) y (3-55) se obtiene:

$$\begin{aligned}y_0 &= (1 + \theta_1^2)\sigma_e^2 = 7,24 \\ y_1 &= -\theta_1\sigma_e^2 = -3,6 \\ y_t &= 0 \quad t = 2, 3, 4, 5, \dots\end{aligned}$$

- d) De acuerdo con (3-56) y (3-57)

$$\begin{aligned}p_1 &= \frac{-\theta_1}{1 + \theta_1^2} = -0,497 \\ p_2 &= 0 \quad t = 2, 3, 4, 5, \dots\end{aligned}$$

- [3]** Para que un modelo AR(2) sea estacionario, las raíces de la ecuación característica deben ser menor que 1 en valor absoluto. Alternativamente, si para el cálculo se utiliza el polinomio característico, las raíces deben ser en valor absoluto mayor que 1, o lo que es lo mismo, deben estar situadas fuera del círculo unidad.

Utilizando la ecuación característica

$$\lambda^2 - 0,6\lambda - 0,3 = 0$$

las raíces de esta ecuación de segundo grado son las siguientes

$$\lambda_1, \lambda_2 = \frac{0,6}{2} \pm \frac{\sqrt{0,6^2 + 4 \times 0,3}}{2} = 0,3 \pm 0,62$$

Por tanto, $\lambda_1 = 0,92$ y $\lambda_2 = -0,32$. Son raíces reales y distintas.

Como $|\lambda_1| = |0,92| < 1$ y $|\lambda_2| = |-0,32| < 1$ el proceso propuesto es estacionario. Alternativamente, utilizando el polinomio característico

$$1 - 0,6L - 0,3L^2 = 0$$

se obtendrán las raíces

$$L_1, L_2 = \frac{0,6}{-2 \times 0,3} \pm \frac{\sqrt{0,6^2 + 4 \times 0,3}}{-2 \times 0,3} = -1 \pm 2,08$$

Como $L_1 = 1,08$ y $L_2 = -3,08$ están fuera del círculo unidad, el proceso es estacionario. Como puede comprobarse de forma inmediata se verifica que $\lambda_1 = 1/L_1$ y $\lambda_2 = 1/L_2$.

En un proceso AR(2) se puede comprobar si es estacionario o no, examinando si los coeficientes ϕ_1 y ϕ_2 cumplen las tres desigualdades siguientes:

$$\phi_1 + \phi_2 < 1$$

$$-\phi_1 + \phi_2 < 1$$

$$\phi_2 > -1$$

En nuestro caso, teniendo en cuenta que $\phi_1 = 0,6$ y $\phi_2 = 0,3$, se verifica que

$$0,6 + 0,3 < 1$$

$$-0,6 + 0,3 < 1$$

$$0,3 > -1$$

Si solamente se hace una comprobación de este tipo no se obtiene información acerca de la naturaleza de las raíces, es decir, no se conoce si son reales o imaginarias.

- b) Por definición, todo proceso AR finito es invertible.
 c) Para calcular y_0 se tiene en cuenta (3-24).

$$y_0 = \phi_1 y_1 + \phi_2 y_2 + \sigma_e^2$$

Por otra parte, dando a t los valores de 1 y 2 en (3-25) se obtiene

$$y_1 = \phi_1 y_0 + \phi_2 y_1$$

$$y_2 = \phi_1 y_1 + \phi_2 y_0$$

Resolviendo este sistema para y_1 y y_2 se obtiene

$$y_1 = \frac{\phi_1}{1 - \phi_2} y_0 \quad y_2 = \frac{\phi_1^2 + \phi_2 - \phi_1^2}{1 - \phi_2} y_0$$

Sustituyendo estos valores en (3-24) y operando se obtiene

$$y_0 = \frac{(1 - \phi_2)}{(1 + \phi_2)[(1 - \phi_2)^2 - \phi_1^2]} \sigma_e^2 = 12,42$$

Del sistema anterior

$$y_1 = \frac{\phi_1}{1 - \phi_2} y_0 = 10,64$$

Una vez determinados y_0 y y_1 , los restantes valores se pueden producir de forma recursiva de la fórmula (3-25)

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2}$$

Por tanto

$$\begin{aligned}\gamma_2 &= \phi_1 \gamma_1 + \phi_2 \gamma_0 = 10,11 \\ \gamma_3 &= \phi_1 \gamma_2 + \phi_2 \gamma_1 = 9,26 \\ \gamma_4 &= \phi_1 \gamma_3 + \phi_2 \gamma_2 = 8,59 \\ \gamma_5 &= \phi_1 \gamma_4 + \phi_2 \gamma_3 = 7,93\end{aligned}$$

- d) Como ya se conocen los correspondientes γ_τ , para calcular p_τ teniendo en cuenta que

$$p_\tau = \frac{\gamma_\tau}{\gamma_0}$$

se obtienen los siguientes resultados

$$p_1 = 0,86 \quad p_2 = 0,81 \quad p_3 = 0,75 \quad p_4 = 0,69 \quad p_5 = 0,64$$

Alternativamente, en la ecuación (3-26) haciendo $\tau = 1$ se obtiene

$$p_1 = \frac{\phi_1}{1 - \phi_2} = 0,86$$

Tomando $\rho_0 = 1$ y $\rho_1 = 0,86$ como valores iniciales, los restantes valores se pueden calcular de forma recursiva a partir de la fórmula (3-26).

- 4 a) Por definición todo proceso MA de orden finito es estacionario.
b) Para comprobar si cumple las condiciones de invertibilidad se calculan las raíces de la ecuación característica:

$$\lambda^2 - 0,4\lambda + 1,2 = 0$$

Operando se obtiene

$$\lambda_1, \lambda_2 = \frac{0,4}{2} \pm \frac{\sqrt{0,4^2 - 4 \times 1,2}}{2} = 0,2 \pm i1,08$$

Es decir, $\lambda_1 = 0,2 + i1,08$ y $\lambda_2 = 0,2 - i1,08$ son raíces complejas conjugadas. Su módulo es:

$$r = \sqrt{a^2 + b^2} = \sqrt{0,2^2 + 1,08^2} = 1,10$$

Como $r = |1,10| > 1$ el proceso no es invertible. Si se hubiera utilizado el polinomio

$$1 - 0,4L + 1,2L^2 = 0$$

se obtendrían las raíces $L_1, L_2 = 0,166 \pm i0,90$, cuyo módulo

$$r = \sqrt{0,166^2 + 0,90^2} = 0,91$$

que es menor que 1 en valor absoluto.

Alternativamente, en un MA(2) las condiciones de invertibilidad aplicadas directamente sobre los parámetros son las siguientes:

$$\begin{aligned}\theta_1 + \theta_2 &< 1 \\ -\theta_1 + \theta_2 &< 1 \\ \theta_2 &> -1\end{aligned}$$

En el modelo propuesto,

$$\begin{aligned}0,4 - 1,2 &= -0,8 < 1 \quad \text{Se cumple} \\ -0,4 - 1,2 &= -1,6 < 1 \quad \text{Se cumple} \\ -1,2 &< -1 \quad \text{No se cumple}\end{aligned}$$

luego por este tercer procedimiento queda también constatado que el proceso no es invertible.

- c) Aunque el proceso no sea invertible siguen siendo válidas las fórmulas (3-65) a (3-68) para el cálculo de la varianza y de las autocovarianzas de distinto orden. Por tanto,

$$\begin{aligned}\gamma_0 &= (1 + \theta_1^2 + \theta_2^2)\sigma_e^2 = 5,2 \\ \gamma_1 &= (-\theta_1 + \theta_1\theta_2)\sigma_e^2 = -1,76 \\ \gamma_2 &= -\theta_2\sigma_e^2 = 2,4 \\ \gamma_\tau &= 0 \quad \text{para } \tau = 3, 4, 5, \dots\end{aligned}$$

- d) A partir de la varianza y de las autocovarianzas, el cálculo es inmediato

$$\begin{aligned}p_1 &= -0,34 \\ p_2 &= 0,46 \\ p_\tau &= 0 \quad \text{para } \tau = 3, 4, 5, \dots\end{aligned}$$

- 5 a) Puesto que $|\phi_1| = 0,9 < 1$ el proceso es estacionario.
b) Análogamente, puesto que $|\phi_1| = 0,8 < 1$ el proceso es invertible.
c) Utilizando (3-99) se obtiene

$$\gamma_0 = \frac{1 - 2\theta_1\phi_1 + \theta_1^2}{1 - \phi_1^2} \sigma_e^2 = 5,26$$

Utilizando (3-97) y (3-98) se tiene

$$\begin{aligned}\gamma_1 &= \phi_1 \gamma_0 - \theta_1 \sigma_e^2 = 0,73 \\ \gamma_2 &= \phi_1 \gamma_1 = 0,66 \\ \gamma_3 &= \phi_1 \gamma_2 = 0,59 \\ \gamma_4 &= \phi_1 \gamma_3 = 0,53 \\ \gamma_5 &= \phi_1 \gamma_4 = 0,48\end{aligned}$$

d) La secuencia de coeficientes de autocorrelación se obtiene de forma inmediata

$$\rho_1 = 0,14 \quad \rho_2 = 0,13 \quad \rho_3 = 0,11 \quad \rho_4 = 0,10 \quad \rho_5 = 0,09$$

6] Expresando el modelo como un proceso de medias móviles se tiene que

$$Y_t = \frac{1}{1-0,9L} [\delta + \epsilon_t] = \frac{1}{1-0,9} \delta + \sum_{j=0}^{\infty} 0,9^j \epsilon_{t-j}$$

La media y la varianza de la serie son

$$E[Y_t] = \frac{\delta}{1-0,9}$$

$$E[Y_t - \frac{\delta}{1-0,9}]^2 = E\left[\sum_{j=0}^{\infty} 0,9^j \epsilon_{t-j}\right]^2 = \frac{\sigma_\epsilon^2}{1-0,9^2}$$

Por otra parte, la autocovarianza de orden τ vendrá dada por

$$\begin{aligned} \gamma &= E[Y_t - \frac{\delta}{1-0,9}] [Y_{t-\tau} - \frac{\delta}{1-0,9}] = E\left[\sum_{j=0}^{\infty} 0,9^j \epsilon_{t-j}\right] \left[\sum_{j=0}^{\infty} 0,9^{j-\tau} \epsilon_{t-\tau-j}\right] = \\ &= E[\epsilon_t + 0,9\epsilon_{t-1} + \dots + 0,9^\tau\epsilon_{t-\tau} + \dots] [\epsilon_{t-\tau} + 0,9\epsilon_{t-\tau-1} + 0,9^2\epsilon_{t-\tau-2} + \dots] = \\ &= 0,9^\tau \frac{1}{1-0,9^2} \sigma_\epsilon^2 \end{aligned}$$

Como se ha visto, tanto la varianza como las autocovarianzas no dependen de δ . En consecuencia los coeficientes de autocorrelación tampoco dependerán de los valores que se asignen a δ .

7] Se postula como primer ensayo un modelo AR(1). Como en un modelo AR(1), se verifica que $\phi_1 = \rho_1$ se tendría el siguiente modelo:

$$Y_t = 0,70 Y_{t-1} + \epsilon_t$$

En un modelo AR(1), los coeficientes de autocorrelación se generan de acuerdo con la ecuación en diferencias

$$\rho_\tau = \phi_1 \rho_{\tau-1}$$

siendo $\rho_0 = 1$

Aplicando de forma recursiva la ecuación anterior, se obtiene

$$\begin{aligned} \rho_1 &= 0,7 \\ \rho_2 &= 0,49 \\ \rho_3 &= 0,343 \\ \rho_4 &= 0,2401 \\ \rho_5 &= 0,16807 \end{aligned}$$

Las cifras calculadas —redondeadas a dos decimales— coinciden con las facilitadas. Por tanto, proceden del modelo AR(1) que se ha postulado al principio.

8] En un principio se descarta un MA de orden bajo, ya que en un MA(q) resulta que $\rho_\tau = 0$ para $\tau > q$.

Centrando la búsqueda en procesos AR y ARMA, se ensayarían sucesivamente los procesos AR(1), ARMA(1, 1) y AR(2), ya que éste sería el orden de menor a mayor complejidad en los cálculos.

En un proceso AR(1) se verifica que

$$\rho_\tau = \phi_1 \rho_{\tau-1} \quad \text{para } \tau > 0$$

Tomando $\phi_1 = \rho_1 = 0,8$ se tendrá

$$\rho_2 = 0,8 \times 0,8 = 0,64 \neq 0,82$$

Por tanto, no puede ser un proceso AR(1).

En un proceso ARMA(1, 1) se verifica que

$$\rho_\tau = \phi_1 \rho_{\tau-1} \quad \text{para } \tau > 1$$

Pero si la secuencia procediera de un ARMA(1, 1) se verificaría que

$$\phi_1 = \frac{\rho_2}{\rho_1} = \frac{0,82}{0,80} = 1,025$$

valor que no es admisible en un proceso estacionario.

Ensayando un AR(2), con las ecuaciones de Yule-Walker se tendrá

$$0,80 = \phi_1 + 0,80 \phi_2$$

$$0,82 = 0,80 \phi_1 + \phi_2$$

Resolviendo para ϕ_1 y ϕ_2 se obtiene que

$$\phi_1 = 0,4 ; \phi_2 = 0,5$$

Por tanto, utilizando la ecuación en diferencias

$$\rho_\tau = 0,4 \rho_{\tau-1} + 0,5 \rho_{\tau-2}$$

para el cálculo de ρ_3, ρ_4, ρ_5 se obtienen los siguientes resultados:

$$\rho_3 = 0,728$$

$$\rho_4 = 0,7012$$

$$\rho_5 = 0,6448$$

que coinciden salvo pequeñas diferencias con los coeficientes suministrados. Por tanto, se puede suponer que éstos proceden de un AR(2).

SOLUCIONES A EJERCICIOS

- 9 En un proceso estacionario, la primera ecuación de Yule-Walker de un AR(2) viene dada por

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

Como se conocen ρ_1 y ϕ_1 , despejando ϕ_2 en la ecuación anterior se tendrá

$$\phi_2 = \frac{\rho_1 - \phi_1}{\rho_1} = \frac{0,9 - 0,7}{0,9} = 0,22$$

- 10 Adoptando el supuesto de estacionariedad la primera ecuación de Yule-Walker en un modelo AR(2) viene dada por

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

Tomando $\rho_1 = 0,7$ y $\phi_1 = -0,4$, se obtiene que:

$$\phi_2 = \frac{\rho_1 - \phi_1}{\rho_1} = \frac{0,7 + 0,4}{0,7} = 1,57$$

Ahora bien, en este resultado $\phi_2 = 1,57 > 1$. Por tanto, este valor de ϕ_2 no corresponde a un AR(2) estacionario. Como a su vez las ecuaciones de Yule-Walker solamente son válidas para modelos estacionarios, la conclusión es que no se puede deducir el valor de ϕ_2 .

- 11 Se adoptará el supuesto de que el proceso es estacionario. Para poder deducir el valor de ϕ_2 será preciso utilizar las dos ecuaciones de Yule-Walker correspondientes a un AR(2).

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

En el anterior sistema se conocen ρ_1 y ϕ_1 ; por consiguiente las incógnitas serán ϕ_2 y ρ_2 . Si en la primera ecuación se despeja ρ_1 y se sustituye en la segunda; se tendrá

$$\rho_2 = \phi_1 \frac{\phi_1}{1 - \phi_2} + \phi_2$$

Operando

$$\phi_2^2 - (1 - \rho_2) \phi_2 - \rho_2 - \phi_1^2 = 0$$

Resolviendo la ecuación teniendo en cuenta que $\phi_1 = 0,8$ y $\rho_2 = 0,6$, se obtienen los valores

$$\phi_2^* = 1,62 \quad y \quad \phi_2^{**} = -0,02$$

a los que corresponden respectivamente los valores

$$\rho_1^* = -1,29 \quad y \quad \rho_1^{**} = 0,78$$

Como ϕ_2^* no cumple las condiciones de estacionariedad —y además el valor ρ_1^* no tiene sentido—, el valor de ϕ_2 consistente con el modelo, sería $\phi_2^{**} = -0,02$.

CAPITULO 3

- 12 En los procesos de medias móviles los coeficientes de autocorrelación tienen unos valores máximos —inferiores a 1— que no se pueden superar en ningún caso. Para un coeficiente de autocorrelación de un orden dado, el valor máximo alcanzable es no decreciente a medida que aumenta el orden del proceso. Así, para un coeficiente ρ_5 la cota es 0,50 en un MA(1) y un 0,98 en un MA(13) según puede verse en el cuadro 3.1. Siguiendo ese cuadro, los coeficientes ρ_1, ρ_2, ρ_3 y ρ_4 son inferiores a los valores correspondientes máximos alcanzables en un MA(13); por lo tanto, es verosímil que puedan proceder a un MA(15).

En lo que se refiere a ρ_5 , la cota máxima en un MA(15) vendrá dada por la fórmula:

$$\max \rho_5 = \cos[\pi/(H+2)] = \cos(\pi/5) = 0,81$$

donde $H = 3$, ya que $(15+1)/5 = 3,2$.

Como en el enunciado $\rho_5 = 0,75 < 0,81$, es posible que los coeficientes de autocorrelación suministrados procedan de un MA(15).

La cota en valor absoluto de los restantes coeficientes vendrá dada por:

$$\max \rho_6 = \cos(\pi/4) = 0,71$$

$$\max \rho_7 = \cos(\pi/4) = 0,71$$

$$\max \rho_8 = \cos(\pi/3) = 0,50$$

$$\max \rho_9 = \cos(\pi/3) = 0,50$$

⋮

$$\max \rho_{15} = \cos(\pi/3) = 0,50$$

- 13 Para examinar la estacionariedad, se calculan las raíces de la ecuación característica

$$\lambda^2 - 0,8\lambda + 0,17 = 0$$

obteniéndose

$$\lambda_1, \lambda_2 = \frac{0,8}{2} \pm \frac{\sqrt{0,64 - 4 \times 0,17}}{2} = 0,4 \pm 0,17$$

Por lo tanto, λ_1 y λ_2 son raíces imaginarias. El módulo es el siguiente:

$$r = \sqrt{a^2 + b^2} = \sqrt{0,4^2 + 0,17^2} = 0,4123$$

Como $|r| = 0,4123 < 1$ el proceso es estacionario.

b) Para el cálculo de la frecuencia angular se tiene en cuenta la siguiente relación:

$$a = r \cos \omega$$

Por tanto:

$$\omega = \arccos a/r = \arccos \frac{0,40}{0,4123} = 0,2448 \text{ radianes}$$

De forma análoga se puede calcular de la relación

$$\theta = r \operatorname{sen} \omega$$

es decir

$$\omega = \arcsen \theta/r = \arcsen \frac{0,10}{0,4123} = 0,2450 \text{ radianes}$$

La pequeña diferencia entre uno y otro cálculo se debe a los efectos del redondeo.

La solución de la parte homogénea del modelo se puede expresar de la siguiente forma de acuerdo con (A2-30)

$$Y_t = A(0,4123)^t \operatorname{sen}(0,2450t - \epsilon)$$

donde

A y ϵ dependen de las condiciones iniciales

c) El período de la oscilación sinusoidal es

$$P_\epsilon = \frac{2\pi}{\omega} = \frac{2\pi}{0,2450} = 25,65 \text{ meses}$$

como el índice es de carácter mensual, el período es compatible con la hipótesis.

d) La solución completa del modelo, prescindiendo del componente aleatorio, de acuerdo con (A2-28) puede ser del siguiente tipo

$$Y_t = r^t (B \cos \omega t - C \operatorname{sen} \omega t) + \mu$$

Ahora bien

$$\mu = \frac{40}{1 - 0,8 + 0,17} = 108$$

Por otra parte, teniendo en cuenta que $Y_0 = 100$ e $Y_1 = 102$, resulta

$$100 = B + 108$$

$$102 = 0,4123 (B \cos 0,2450 - C \operatorname{sen} 0,2450) + 108$$

de donde

$$B = -8$$

$$C = 28$$

Por tanto, una vez determinadas las constantes arbitrarias, y la solución particular, no estocástica, se tiene que

$$Y_t = (0,4123)^t (-8 \cos 0,2450t - 28 \operatorname{sen} 0,2450t) + 108$$

Como enero de 1984 se corresponde con la observación 157, sustituyendo t en la ecuación anterior por dicho valor se tiene:

$$Y_t \approx 108$$

ya que el primer sumando es prácticamente igual a 0 debido al efecto del factor $(0,4123)^{157}$.

- [14] a) Para aplicar el procedimiento de las fracciones parciales se deben calcular previamente las raíces de la ecuación

$$\lambda^2 - 1,2\lambda + 0,32 = 0$$

Resolviendo se obtiene

$$\lambda_1 = 0,8 \quad \text{y} \quad \lambda_2 = 0,4$$

El modelo se puede expresar así:

$$\epsilon_t = \frac{1}{(1 - 0,8L)(1 - 0,4L)} Y_t$$

Haciendo

$$\frac{1}{(1 - 0,8L)(1 - 0,4L)} = \frac{A}{(1 - 0,8L)} + \frac{B}{(1 - 0,4L)}$$

Eliminando denominadores resulta:

$$1 = (1 - 0,4L)A + (1 - 0,8L)B$$

o

$$1 = (A + B) - (0,4A + 0,8B)L = 0$$

Con el sistema

$$A + B = 1$$

$$0,4A + 0,8B = 0$$

se obtienen los valores

$$A = 2 \quad B = -1$$

Por tanto

$$\epsilon_t = \frac{2}{1 - 0,8L} - \frac{1}{1 - 0,4L} Y_t$$

$$= 2 \sum_{j=0}^{\infty} (0,8L)^j - \sum_{j=0}^{\infty} (0,4L)^j Y_t$$

$$= Y_t + 1,2Y_{t-1} + 1,12Y_{t-2} + 0,96Y_{t-3} + 0,70Y_{t-4} + 0,65Y_{t-5} + \dots$$

b) La igualdad $\pi(L)\theta(L) = \phi(L)$, por ser un proceso de medias móviles, queda reducido a

$$\pi(L)\theta(L) = 1$$

es decir,

$$(1 - \pi_1 L - \pi_2 L^2 - \pi_3 L^3 - \pi_4 L^4 \dots)(1 - \theta_1 L - \theta_2 L^2) = 1$$

Realizando operaciones:

$$1 - (\pi_1 + \theta_1)L - (\pi_2 - \pi_1\theta_1 + \theta_2)L^2 - (\pi_3 - \pi_2\theta_1 - \pi_1\theta_2)L^3 \dots = 1$$

Igualando miembro a miembro los coeficientes de cada potencia de L se tiene

$$\pi_1 = -\theta_1 = -1,2$$

$$\pi_2 = \pi_1\theta_1 - \theta_2 = -1,12$$

Para $j > 2$ puede comprobarse, continuando el desarrollo anterior, que los coeficientes π_j siguen la ecuación en diferencias

$$\pi_j = \theta_1\pi_{j-1} + \theta_2\pi_{j-2}$$

En consecuencia

$$\pi_3 = -1,2 \times 1,12 + 0,32 \times 1,2 = -0,96$$

$$\pi_4 = -1,2 \times 0,96 + 0,32 \times 1,12 = -0,79$$

$$\pi_5 = -1,2 \times 0,79 + 0,32 \times 0,96 = -0,64$$

[15] a) Aplicando el procedimiento de las fracciones parciales se obtiene

$$Y_1 = \left[\frac{1,75}{1 - 0,7L} - \frac{0,75}{1 - 0,3L} \right] e_t$$

$$= \left[1,75 \sum_{j=0}^{\infty} (0,7L)^j - 0,75 \sum_{j=0}^{\infty} (0,3L)^j \right] e_t =$$

$$Y_1 = e_t + e_{t-1} + 0,79e_{t-2} - 0,58e_{t-3} + 0,41e_{t-4} - 0,29e_{t-5} + \dots$$

b) A partir de la relación $\psi(L)\phi(L) = 1$, se obtienen las siguientes relaciones:

$$\psi_1 - \phi_1 = 0$$

$$\psi_2 - \psi_1\phi_1 - \phi_2 = 0$$

Para $j > 0$, se obtiene la ecuación en diferencias

$$\psi_j = \phi_1\psi_{j-1} + \phi_2\psi_{j-2}$$

Teniendo en cuenta que $\phi_1 = 1$ y $\phi_2 = -0,21$, se obtiene

$$\psi_1 = 1$$

$$\psi_2 = 1 \times 1 - 0,21 = 0,79$$

$$\psi_3 = 1 \times 0,79 - 0,21 \times 1 = 0,58$$

$$\psi_4 = 1 \times 0,58 - 0,21 \times 0,79 = 0,41$$

$$\psi_5 = 1 \times 0,41 - 0,21 \times 0,58 = 0,29$$

[16] a) Se establecen las relaciones

$$(1 + \psi_1 L + \psi_2 L^2 + \dots)(1 - \phi_1 L) = 1 - \theta_1 L$$

de donde

$$\psi_1 - \phi_1 = -\theta_1$$

y para $j > 1$

$$\psi_j - \phi_1\psi_{j-1} = 0$$

Por tanto

$$\psi_1 = 0,8 + 0,7 = 1,5$$

$$\psi_2 = 0,8 \times 1,5 = 1,2$$

$$\psi_3 = 0,8 \times 1,2 = 0,96$$

$$\psi_4 = 0,8 \times 0,96 = 0,77$$

$$\psi_5 = 0,8 \times 0,77 = 0,61$$

b) Se establecen las relaciones

$$(1 - \pi_1 L - \pi_2 L^2 - \dots)(1 - \theta_1 L) = 1 - \phi_1 L$$

de donde

$$\pi_1 + \theta_1 = \phi_1$$

y para $j > 1$

$$\pi_1 = 0,8 + 0,7 = 1,5$$

$$\pi_2 = -0,7 \times 1,5 = -1,05$$

$$\pi_3 = 0,7 \times 1,05 = 0,74$$

$$\pi_4 = -0,7 \times 0,74 = -0,52$$

$$\pi_5 = 0,7 \times 0,52 = 0,36$$

- [17] Las raíces de la ecuación característica son

$$\lambda_1 = \frac{\theta_1}{2} + \frac{\sqrt{\theta_1^2 + 4\theta_2}}{2}$$

$$\lambda_2 = \frac{\theta_1}{2} - \frac{\sqrt{\theta_1^2 + 4\theta_2}}{2}$$

A partir de las expresiones anteriores se puede ver que

$$\lambda_1 + \lambda_2 = \theta_1$$

$$\lambda_1 \lambda_2 = \frac{\theta_1}{4} - \frac{(\theta_1^2 + 4\theta_2)}{4} = -\theta_2$$

En consecuencia

$$\theta_1 = 0,90$$

$$\theta_2 = -0,18$$

- [18] a) De acuerdo con (3-101) y (3-102) se obtiene

$$\rho_1 = \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1 \theta_1} = \frac{(1 - 0,6^2)0}{1 + 0,6^2 - 2 \times 0,6^2} = 0$$

$$\rho_\tau = \phi_1 \rho_{\tau-1} = 0 \quad \tau > 1$$

es decir, todos los coeficientes de autocorrelación son igual a 0.

- b) El hecho de que $\rho_\tau = 0$ para $\forall \tau$, se debe a la existencia de un factor común en la parte autorregresiva y en la parte de medias móviles.

En efecto:

$$(1 - 0,6L)Y_t = (1 - 0,6L)\epsilon_t$$

Dividiendo ambos miembros por $(1 - 0,6L)$ el modelo queda reducido a

$$Y_t = \epsilon_t$$

que es un ruido blanco, que se caracteriza precisamente porque $\rho_\tau = 0$, para $\forall \tau$.

- [19] Para ver si se pueden introducir simplificaciones se debe proceder al cálculo de las raíces de las ecuaciones características de la parte autorregresiva y de la parte de medias móviles:

En la parte autorregresiva, esta ecuación característica viene dada por

$$\lambda_{AR}^2 - 0,9\lambda_{AR} + 0,2 = 0$$

obteniéndose

$$\lambda_{AR1} = 0,5 \text{ y } \lambda_{AR2} = 0,4$$

En la parte de medias móviles, la ecuación característica viene dada por

$$\lambda_{MA}^2 - 1,3\lambda_{MA} + 0,4 = 0$$

obteniéndose

$$\lambda_{MA1} = 0,8 \text{ y } \lambda_{MA2} = 0,5$$

En consecuencia el modelo en forma factorizada se podrá expresar así:

$$(1 - 0,5L)(1 - 0,4L)Y_t = (1 - 0,8L)(1 - 0,5L)\epsilon_t$$

Dividiendo ambos miembros por $(1 - 0,5L)$ resulta que

$$(1 - 0,4L)Y_t = (1 - 0,8L)\epsilon_t$$

es decir

$$Y_t = 0,4Y_{t-1} + \epsilon_t - 0,8\epsilon_{t-1}$$

con lo cual queda el modelo simplificado.

- [20] La ecuación que genera los coeficientes de autocorrelación es la ecuación en diferencias homogéneas de 2.^o orden:

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2}$$

Para hallar la solución se deben calcular las raíces de la ecuación característica

$$\lambda^2 - \phi_1 \lambda - \phi_2 = 0$$

Calculando las raíces en el modelo propuesto se obtiene una única raíz real:

$$\lambda_u = 0,9$$

En este caso particular, la solución de la ecuación homogénea es del tipo

$$\rho_t = A_1 \lambda_u^t + A_2 t \lambda_u$$

donde A_1 y A_2 se determinan a partir de los valores iniciales ρ_0 y ρ_1 . El valor de ρ_1 viene dado por

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} = \frac{1,8}{1 + 0,81} = 0,99$$

Por tanto, para $\tau = 0$ y $\tau = 1$ se tendrá:

$$\begin{aligned} I &= A_1 \\ 0,99 &= A_1 \cdot 0,9 + A_1 \times I \times 0,9 \end{aligned}$$

En consecuencia

$$A_1 = 1 \quad A_2 = 0,1$$

y la solución vendrá dada por

$$p_\tau = 0,9^\tau + 0,1 \times \tau \times 0,9^\tau$$

Así para $\tau = 10$, se obtiene

$$p_{10} = 0,9^{10} + 0,1 \times 10 \times 0,9^{10} = 0,70$$

Para $\tau = 20$, se obtiene

$$p_{20} = 0,36$$

- [21]** a) Las raíces del polinomio característico

$$1 - 0,1L - 0,9L^2 = 0$$

son las siguientes

$$L_1 = 1 \quad L_2 = -1,11$$

Como una de las raíces $L_1 = 1$, el proceso es no estacionario.

b) De acuerdo con las raíces obtenidas el modelo se puede factorizar de la siguiente forma:

$$(1 + 0,9L)(1 - L)Y_t = \varepsilon_t$$

Haciendo:

$$w_t = \Delta Y_t = (1 - L)Y_t$$

El proceso correspondiente a la variable transformada w_t es el siguiente:

$$(1 + 0,9L)w_t = \varepsilon_t$$

es decir:

$$w_t + 0,9w_{t-1} = \varepsilon_t$$

Como $|0,9| < 1$, el proceso correspondiente a w_t es estacionario.

- [22]** Utilizando el operador de retardo el modelo puede expresarse así

$$(1 + L)Y_t = \varepsilon_t$$

Por otra parte

$$w_t = \Delta Y_t = (1 - L)Y_t$$

de donde

$$Y_t = (1 - L)^{-1}w_t$$

Sustituyendo en el modelo, se tiene que

$$(1 + L)(1 - L)^{-1}w_t = \varepsilon_t$$

es decir

$$(1 + L)w_t = (1 - L)\varepsilon_t$$

$$w_t = -w_{t-1} + \varepsilon_t - \varepsilon_{t-1}$$

El proceso transformado w_t no cumple las condiciones de estacionariedad.

- [23]** a) La media del modelo es:

$$\mu_t = E(Y_t) = E[\beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t] = \beta_0 + \beta_1 t + \beta_2 t^2$$

Como la media es distinta para cada t , el modelo no estacionario en media.

- b) El modelo en el período t es

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t$$

y en el período $t-1$

$$Y_{t-1} = \beta_0 + \beta_1(t-1) + \beta_2(t-1)^2 + \varepsilon_{t-1}$$

Restando miembro a miembro se obtiene:

$$\Delta Y_t = \beta_1 - \beta_2 + 2\beta_2 t + \varepsilon_t - \varepsilon_{t-1}$$

Expresando la ecuación anterior para el período $t-1$, se obtiene

$$\Delta Y_{t-1} = \beta_1 - \beta_2 + 2\beta_2(t-1) + \varepsilon_{t-1} - \varepsilon_{t-2}$$

Restando miembro a miembro se obtiene

$$w_t = \Delta^2 Y_t = 2\beta_2 + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}$$

El proceso w_t es estacionario, ya que es un proceso MA(2) y todos los procesos de medias móviles de orden finito son estacionarios.

- c) La ecuación característica de la parte MA del proceso resultante viene dada por

$$\lambda^2 - 2\lambda + 1 = 0$$

Las raíces de esta ecuación son

$$\lambda_1 = \lambda_2 = 1$$

Por tanto, el proceso es no invertible.

- [24] a) Un modelo ARIMA (0, 1, 1) se puede expresar así:

$$\begin{aligned} \epsilon_t &= \frac{1}{(1 - \theta_1 L)} (Y_t - Y_{t-1}) = \\ &= (Y_t + \theta_1 Y_{t-1} + \theta_1^2 Y_{t-2} + \dots) - (Y_{t-1} + \theta_1 Y_{t-2} + \theta_1^2 Y_{t-3} + \dots) = \\ &= Y_t - (1 - \theta_1) Y_{t-1} - (1 - \theta_1) \theta_1 Y_{t-2} - (1 - \theta_1) \theta_1^2 Y_{t-3} - \dots \end{aligned}$$

Teniendo en cuenta que $\theta_1 = 0,09$, el modelo se expresaría así:

$$Y_t - 0,91 Y_{t-1} - 0,082 Y_{t-2} - 0,007 Y_{t-3} - \dots = \epsilon_t$$

- b) La expresión teórica de un AR(∞) sería la siguiente

$$Y_t - \pi_1 Y_{t-1} - \pi_2 Y_{t-2} - \pi_3 Y_{t-3} \dots = \epsilon_t$$

Igualando cada π_i con la expresión obtenida en el apartado a), se tiene que

$$\pi_1 = (1 - \theta_1)$$

$$\pi_2 = (1 - \theta_1) \theta_1$$

$$\pi_3 = (1 - \theta_1) \theta_1^2$$

En consecuencia

$$\sum_{i=1}^{\infty} \pi_i = (1 - \theta_1) \sum_{i=0}^{\infty} \theta_1^i = (1 - \theta_1) \frac{1}{1 - \theta_1} = 1$$

El sumatorio $\sum \pi_i = 1$, cualquiera que sea el valor de θ_1 .

- [25] Antes de contestar a los distintos apartados se va a analizar la distribución probabilística de e^{ϵ_t} .

Si $\epsilon_t \sim N(0, \sigma_t)$ y se denomina

$$v_t = e^{\epsilon_t}$$

entonces como $\ln v_t = \epsilon_t$, se dice que v_t tiene distribución log-normal. (Una variable tiene distribución log-normal cuando el logaritmo neperiano de esa variable tiene una distribución normal.)

La media y la varianza de v_t vienen dadas por las siguientes expresiones

$$\begin{aligned} E(v_t) &= e^{\frac{1}{2}\sigma_t^2} \\ E(v_t - e^{\frac{1}{2}\sigma_t^2}) &= e^{\sigma_t^2}(e^{\sigma_t^2} - 1) \end{aligned}$$

- a) Para analizar si es estacionario en media y en varianza, se van a calcular la media y la varianza de Y_t :

$$\begin{aligned} E[Y_t] &= [\beta_0 + \beta_1 t] E(v_t) = (\beta_0 + \beta_1 t) e^{\frac{1}{2}\sigma_t^2} \\ E[Y_t - (\beta_0 + \beta_1 t) e^{\frac{1}{2}\sigma_t^2}]^2 &= E[(\beta_0 + \beta_1 t)v_t - (\beta_0 + \beta_1 t)e^{\frac{1}{2}\sigma_t^2}]^2 \\ &= (\beta_0 + \beta_1 t)^2 E(v_t - e^{\frac{1}{2}\sigma_t^2})^2 = (\beta_0 + \beta_1 t)^2 e^{\sigma_t^2}(e^{\sigma_t^2} - 1) \end{aligned}$$

El modelo no es, por tanto, estacionario en media, ni estacionario en varianza, ya que ambos parámetros varían al variar t .

- b) Si se hace $w_t = \Delta Y_t$, el modelo transformado será:

$$w_t = Y_t - Y_{t-1} = (\beta_0 + \beta_1 t) e^{\epsilon_t} - [\beta_0 + \beta_1 (t-1)] e^{\epsilon_{t-1}}$$

La media y la varianza de w_t son las siguientes:

$$\begin{aligned} E(w_t) &= (\beta_0 + \beta_1 t) e^{\frac{1}{2}\sigma_t^2} - [\beta_0 + \beta_1 (t-1)] e^{\frac{1}{2}\sigma_{t-1}^2} = \beta_1 e^{\frac{1}{2}\sigma_t^2} \\ E[w_t - \beta_1 e^{\frac{1}{2}\sigma_t^2}]^2 &= \\ &= E[(\beta_0 + \beta_1 t) e^{\epsilon_t} - (\beta_0 + \beta_1 (t-1)) e^{\epsilon_{t-1}}]^2 = [(\beta_0 + \beta_1 t) e^{\frac{1}{2}\sigma_t^2} - (\beta_0 + \beta_1 (t-1)) e^{\frac{1}{2}\sigma_{t-1}^2}]^2 \end{aligned}$$

Al calcular esta última expresión se ve claramente que la varianza de w_t varía al variar t .

En consecuencia w_t es estacionario en media, pero no es estacionario en varianza. Además, la distribución probabilística de w_t no es normal sino lognormal.

- c) Definido el proceso

$$Z_t = \Delta \ln Y_t = \ln(\beta_0 + \beta_1 t) + \epsilon_t - \ln[\beta_0 + \beta_1 (t-1)] - \epsilon_{t-1}$$

su media y varianza serán las siguientes

$$E(Z_t) = \ln(\beta_0 + \beta_1 t) - \ln[\beta_0 + \beta_1 (t-1)] = \ln \frac{\beta_0 + \beta_1 t}{\beta_0 + \beta_1 (t-1)}$$

Al crecer t , $E(Z_t)$ tiende hacia 0.

$$\text{var}(Z_t) = E[\ln(\beta_0 + \beta_1 t) + \epsilon_t - \ln[\beta_0 + \beta_1 (t-1)] - \epsilon_{t-1} - \ln(\beta_0 + \beta_1 t) +$$

$$\ln[\beta_0 + \beta_1 (t-1)]]^2 = E(\epsilon_t - \epsilon_{t-1})^2 = 2\sigma_t^2$$

El proceso Z_t es estacionario en varianza y tiende a ser estacionario en media al crecer t .

CAPÍTULO 4

- [1] a) En primer lugar se calcula la media del proceso con los 10 datos disponibles

$$\bar{Y} = 2$$

Para la obtención de r_1 , r_2 y r_3 se aplican las fórmulas siguientes:

$$r_1 = \frac{\sum_{t=2}^{10} (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=1}^{10} (Y_t - \bar{Y})^2} = \frac{-296}{386} = -0,767$$

$$r_2 = \frac{\sum_{t=3}^{10} (Y_t - \bar{Y})(Y_{t-2} - \bar{Y})}{\sum_{t=1}^{10} (Y_t - \bar{Y})^2} = \frac{-266}{386} = 0,689$$

$$r_3 = \frac{\sum_{t=4}^{10} (Y_t - \bar{Y})(Y_{t-3} - \bar{Y})}{\sum_{t=1}^{10} (Y_t - \bar{Y})^2} = \frac{-272}{386} = -0,705$$

Obsérvese que el denominador es el mismo para r_1 , r_2 y r_3 y que la media calculada con las 10 observaciones disponibles es la que se ha utilizado en todos los casos.

b) Para analizar la significatividad de los coeficientes de autocorrelación se adoptarán los siguientes supuestos:

$$\text{Para } \begin{cases} r_1 \\ r_2 \\ r_3 \end{cases} \text{ se considerará que } p_\tau = 0 \text{ para } \begin{cases} \tau \geq 1 \\ \tau \geq 2 \\ \tau \geq 3 \end{cases}$$

En consecuencia, utilizando las fórmulas aproximadas de Barlett se obtendrán las siguientes estimaciones de las varianzas de los coeficientes de autocorrelación:

$$\widehat{\text{var}}(r_1) = \frac{1}{T} = \frac{1}{10} = 0,10$$

$$\widehat{\text{var}}(r_2) = \frac{1}{T} [1 + 2r_1^2] = \frac{1}{10} [1 + 2 \times (-0,767)^2] = 0,218$$

$$\widehat{\text{var}}(r_3) = \frac{1}{T} [1 + 2(r_1^2 + r_2^2)] = \frac{1}{10} [1 + 2[(-0,767)^2 + 0,689^2]] = 0,313$$

Bajo el supuesto de que $p_\tau = 0$, se tiene que aproximadamente

$$\frac{r_\tau}{\sqrt{\text{var}(r_\tau)}} \sim N[0, 1]$$

Por tanto, para que sean significativamente distintos de 0 para un nivel de confianza del 95 %, se tiene que verificar que:

$$\left| \frac{r_\tau}{\sqrt{\text{var}(r_\tau)}} \right| > 1,96$$

Para cada uno de los coeficientes, los resultados que se obtienen son:

Para $\tau = 1$ $| -2,427 | > 1,96 \quad r_1 \text{ es estadísticamente significativo.}$

Para $\tau = 2$ $| 1,476 | < 1,96 \quad r_2 \text{ no es estadísticamente significativo.}$

Para $\tau = 3$ $| 1,260 | < 1,96 \quad r_3 \text{ no es estadísticamente significativo.}$

Aunque los tres coeficientes tienen unos valores relativamente elevados, solamente el primero es estadísticamente significativo debido a que el tamaño de la muestra es muy pequeño.

- [2] A partir de la ecuación de Yule-Walker para un AR(1) se tiene que:

$$p_1 = \phi_{11} = 0,9$$

Las ecuaciones de Yule-Walker, calculadas para un AR(2), permiten obtener los coeficientes ϕ_{21} y ϕ_{22} :

$$\begin{bmatrix} \phi_{21} \\ \phi_{22} \end{bmatrix} = \begin{bmatrix} 1 & p_1 \\ p_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 1 & 0,9 \\ 0,9 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,9 \\ 0,8 \end{bmatrix} = \begin{bmatrix} 0,947 \\ 0,053 \end{bmatrix}$$

Como el proceso estocástico generador es un AR(2), se verifica que $\phi_{rr} = 0$ para $\tau > 2$. En consecuencia, $\phi_{33} = 0$.

- [3] a) $\bar{Y} = 5$

Aplicando las mismas fórmulas que en el ejercicio 1 se obtiene

$$r_1 = \frac{61}{80} = 0,763$$

$$r_2 = \frac{29}{80} = 0,363$$

$$r_3 = \frac{-4}{80} = -0,05$$

- b) Aplicando la fórmula de Barlett y adoptando idénticos supuestos que en el ejercicio 1 se obtiene:

$$\widehat{\text{var}}(r_1) = 0,100$$

$$\widehat{\text{var}}(r_2) = 0,216$$

$$\widehat{\text{var}}(r_3) = 0,243$$

Las bandas de confianza del 95 % bajo el supuesto de que $\rho_t = 0$, para $t = 1, 2$ y 3 son las siguientes:

$$\text{Para } t = 1 \quad [-1,96 \times \sqrt{0,10}; 1,96 \times \sqrt{0,10}] \quad [-0,619; 0,619]$$

$$\text{Para } t = 2 \quad [-1,96 \times \sqrt{0,216}; 1,96 \times \sqrt{0,216}] \quad [-0,911; 0,911]$$

$$\text{Para } t = 3 \quad [-1,96 \times \sqrt{0,243}; 1,96 \times \sqrt{0,243}] \quad [-0,966; 0,966]$$

Solamente r_1 cae fuera de la banda de confianza. Por tanto, se puede afirmar que r_1 es estadísticamente significativo.

- c) Aplicando sucesivamente las ecuaciones de Yule-Walker para un AR(1), un AR(2), y un AR(3) se tiene que:

$$\phi_{11} = 0,763$$

$$\begin{bmatrix} \hat{\phi}_{11} \\ \hat{\phi}_{21} \end{bmatrix} = \begin{bmatrix} 1 & 0,763 \\ 0,763 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,763 \\ 0,363 \end{bmatrix} = \begin{bmatrix} 1,163 \\ -0,525 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{22} \\ \hat{\phi}_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0,763 & 0,363 \\ 0,763 & 1 & 0,763 \\ 0,363 & 0,763 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,763 \\ 0,363 \\ -0,050 \end{bmatrix} = \begin{bmatrix} 1,038 \\ -0,248 \\ -0,238 \end{bmatrix}$$

- d) Bajo el supuesto de que $\phi_{rr} = 0$, aproximadamente

$$\phi_{rr} \sim N[0,1/T]$$

En consecuencia, las bandas de confianza del 95 % para ϕ_{11} , ϕ_{21} , y ϕ_{31} serán

$$[-1,96 \times \sqrt{1/T}; 1,96 \times \sqrt{1/T}] = [-0,619; 0,619]$$

El coeficiente $\hat{\phi}_{11}$ cae fuera de este intervalo, siendo por tanto estadísticamente significativo.

4

- a) $\bar{Y} = 0$

$$r_1 = \frac{20}{44} = 0,455$$

$$r_2 = \frac{-11}{44} = -0,250$$

$$r_3 = \frac{-26}{44} = -0,591$$

- b) Aplicando sucesivamente las ecuaciones de Yule-Walker para un AR(1), un AR(2) y un AR(3), se obtienen los siguientes resultados:

$$\phi_{11} = 0,455$$

$$\begin{bmatrix} \hat{\phi}_{11} \\ \hat{\phi}_{21} \end{bmatrix} = \begin{bmatrix} 1 & 0,455 \\ 0,455 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,455 \\ -0,250 \end{bmatrix} = \begin{bmatrix} 0,717 \\ -0,576 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{22} \\ \hat{\phi}_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0,455 & -0,250 \\ 0,455 & 1 & 0,455 \\ -0,250 & 0,455 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,455 \\ -0,250 \\ -0,591 \end{bmatrix} = \begin{bmatrix} 0,555 \\ -0,374 \\ -0,282 \end{bmatrix}$$

- c) Por regresión se ajustarían sucesivamente los modelos.

$$Y_t = \phi_{11} Y_{t-1} + u_{1t}$$

$$Y_t = \phi_{21} Y_{t-1} + \phi_{11} Y_{t-2} + u_{2t}$$

$$Y_t = \phi_{31} Y_{t-1} + \phi_{21} Y_{t-2} + \phi_{11} Y_{t-3} + u_{3t}$$

Los resultados obtenidos son los siguientes:

$$\hat{\phi}_{11} = \left[\sum_{t=2}^{10} Y_{t-1}^2 \right]^{-1} \left[\sum_{t=2}^{10} Y_t Y_{t-1} \right] = \frac{1}{40} \cdot 20 = 0,5$$

$$\begin{bmatrix} \hat{\phi}_{21} \\ \hat{\phi}_{31} \end{bmatrix} = \begin{bmatrix} \sum_{t=3}^{10} Y_{t-1}^2 & \sum_{t=3}^{10} Y_{t-1} Y_{t-2} \\ \sum_{t=3}^{10} Y_{t-1} Y_{t-2} & \sum_{t=3}^{10} Y_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=3}^{10} Y_t Y_{t-1} \\ \sum_{t=3}^{10} Y_t Y_{t-2} \end{bmatrix}$$

$$= \begin{bmatrix} 36 & 22 \\ 22 & 39 \end{bmatrix}^{-1} \begin{bmatrix} 14 \\ -11 \end{bmatrix} = \begin{bmatrix} 0,857 \\ -0,765 \end{bmatrix}$$

Análogamente

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{22} \\ \hat{\phi}_{32} \end{bmatrix} = \begin{bmatrix} 27 & 16 & -7 \\ 16 & 35 & 20 \\ -7 & 20 & 35 \end{bmatrix}^{-1} \begin{bmatrix} 11 \\ -13 \\ -26 \end{bmatrix} = \begin{bmatrix} 0,489 \\ -0,337 \\ -0,453 \end{bmatrix}$$

- d) En primer lugar se examinará la diferencia para $\hat{\phi}_{11}$. Calculado por Yule-Walker:

$$\hat{\phi}_{11}^* = \frac{\sum_{t=2}^{10} Y_t Y_{t-1}}{\sum_{t=1}^{10} Y_t^2}$$

Calculando por regresión

$$\hat{\phi}_{11}^* = \frac{\sum_{t=2}^{10} Y_t Y_{t-1}}{\sum_{t=2}^{10} Y_{t-1}^2}$$

La diferencia radica en que por Yule-Walker, en el denominador existe un sumando más, concretamente $Y_{10}^2 = 4$, ya que

$$\sum_{t=1}^{10} Y_t^2 = \sum_{t=1}^9 Y_t^2 + Y_{10}^2 = \sum_{t=2}^{10} Y_{t-1}^2 + Y_{10}^2$$

Examinaremos ahora las diferencias para $\hat{\phi}_{11}$, que son fácilmente generalizables para un orden mayor.

Las ecuaciones de Yule-Walker se pueden expresar así:

$$\begin{bmatrix} \hat{\phi}_{11} \\ \hat{\phi}_{22} \end{bmatrix} = \begin{bmatrix} 1 & r_1 \\ r_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} \frac{\sum_{t=1}^{10} Y_t^2}{\sum_{t=1}^{10} Y_t^2} & \frac{\sum_{t=2}^{10} Y_t Y_{t-1}}{\sum_{t=1}^{10} Y_t^2} \\ \frac{\sum_{t=2}^{10} Y_t Y_{t-1}}{\sum_{t=1}^{10} Y_t^2} & \frac{\sum_{t=3}^{10} Y_t Y_{t-2}}{\sum_{t=1}^{10} Y_t^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\sum_{t=2}^{10} Y_t Y_{t-1}}{\sum_{t=1}^{10} Y_t^2} \\ \frac{\sum_{t=3}^{10} Y_t Y_{t-2}}{\sum_{t=1}^{10} Y_t^2} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{t=1}^{10} Y_t^2 & \sum_{t=2}^{10} Y_t Y_{t-1} \\ \sum_{t=2}^{10} Y_t Y_{t-1} & \sum_{t=3}^{10} Y_t Y_{t-2} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=2}^{10} Y_t Y_{t-1} \\ \sum_{t=3}^{10} Y_t Y_{t-2} \end{bmatrix}$$

Comparando esta última expresión con la fórmula aplicada en la regresión se observa que coincide en ambas expresiones el término $\sum_{t=3}^{10} Y_t Y_{t-2}$. En todos los demás casos difieren en uno o dos sumandos. Así, por ejemplo:

$$\sum_{t=1}^{10} Y_t^2 = Y_1^2 + \sum_{t=3}^{10} Y_{t-1}^2 + Y_{10}^2 = \sum_{t=3}^{10} Y_{t-2}^2 + Y_9^2 + Y_{10}^2$$

Las diferencias entre los cálculos por ambos procedimientos —Yule-Walker y regresión— se van atenuando a medida que crece el tamaño de la muestra, siendo equivalentes asintóticamente.

- 5) a) Para cada r se considerará el supuesto de que $\rho_k = 0$ para $k \geq \tau$. Por tanto, aplicando la fórmula aproximada de Barlett se obtienen los siguientes resultados:

$$\begin{aligned} \text{var}(r_1) &= 0,010 \\ \text{var}(r_2) &= 0,026 \\ \text{var}(r_3) &= 0,040 \\ \text{var}(r_4) &= 0,050 \\ \text{var}(r_5) &= 0,059 \end{aligned}$$

Para que un coeficiente sea estadísticamente significativo al 95 % se deberá verificar que

$$\left| \frac{r_\tau}{\sqrt{\text{var}(r_\tau)}} \right| > 1,96$$

Con los datos propuestos se obtienen los siguientes resultados:

$$\begin{aligned} \text{Para } \tau = 1 & | -9,02 | > 1,96. \text{ Es estadísticamente significativo.} \\ \text{Para } \tau = 2 & | 5,06 | > 1,96 \quad " \quad " \quad " \\ \text{Para } \tau = 3 & | -3,57 | > 1,96 \quad " \quad " \quad " \\ \text{Para } \tau = 4 & | 3,00 | > 1,96 \quad " \quad " \quad " \\ \text{Para } \tau = 5 & | 2,40 | > 1,96 \quad " \quad " \quad " \end{aligned}$$

$$\hat{\phi}_{11} = 0,902$$

$$\begin{bmatrix} \hat{\phi}_{11} \\ \hat{\phi}_{22} \end{bmatrix} = \begin{bmatrix} 1 & -0,902 \\ -0,902 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -0,902 \\ 0,816 \end{bmatrix} = \begin{bmatrix} -0,890 \\ 0,013 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\phi}_{11} \\ \hat{\phi}_{22} \\ \hat{\phi}_{33} \end{bmatrix} = \begin{bmatrix} 1 & -0,902 & 0,816 \\ -0,902 & 1 & -0,902 \\ 0,816 & -0,902 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,902 \\ 0,816 \\ -0,714 \end{bmatrix} = \begin{bmatrix} -0,892 \\ 0,128 \\ 0,130 \end{bmatrix}$$

- c) Como distribución aproximada se tendrá en cuenta que

$$\hat{\phi}_{11} \sim N[0,1/T]$$

Los resultados obtenidos son los siguientes

- Para $t = 1$ $| -9,02 | > 1,96$ Es estadísticamente significativo.
 Para $t = 2$ $| 0,13 | < 1,96$ No es estadísticamente significativo.
 Para $t = 3$ $| 1,30 | < 1,96$ No es estadísticamente significativo.

- d) Los coeficientes de autocorrelación simple son todos significativos y decrecen de forma aproximadamente exponencial con alternancia de signos; por el contrario, después del primer coeficiente de autocorrelación parcial se advierte un corte bursco. En consecuencia, puede postularse un modelo AR(1), en el que

$$\hat{\phi}_1 = -0,902$$

Con este valor de $\hat{\phi}_{11}$, se obtendrían los siguientes coeficientes de correlación teóricos:

$$\rho_1 = -0,902 \quad \rho_2 = 0,814 \quad \rho_3 = -0,734 \quad \rho_4 = 0,662 \quad \rho_5 = -0,597$$

Como puede observarse las discrepancias entre estos coeficientes y los coeficientes suministrados son mínimas, confirmado así la identificación adoptada.

- [6] a) Adoptando los supuestos utilizados en anteriores ejercicios se obtiene la siguiente tabla donde el número que aparece entre paréntesis debajo de cada coeficiente es la desviación típica estimada. Con un asterisco se han señalado los coeficientes que son estadísticamente significativos para un nivel de confianza del 95 %.

Retardo	r_r	$\hat{\phi}_{rr}$
1	-0,490*	(0,08)
2	0,040	(0,10)
3	-0,026	(0,10)
4	0,009	(0,10)
5	0,013	(0,10)

- b) Es muy claro el corte que se observa en los coeficientes de autocorrelación simple después del primer retardo, mientras que los coeficientes de autocorrelación parcial disminuyen de forma relativamente rápida. En consecuencia puede postularse un modelo MA(1).

Teóricamente:

$$\rho_t = \frac{-\theta_1}{1 + \theta_1^2}$$

o, lo que es lo mismo:

$$\rho_1 \theta_1^2 + \theta_1 + \rho_1 = 0$$

Sustituyendo ρ_1 por -0,490 y resolviendo se obtienen las siguientes estimaciones

para θ_{11} :

$$\hat{\theta}_{111} = 0,817$$

$$\hat{\theta}_{112} = 1,223$$

Estas dos raíces corresponden a un modelo invertible y a otro no invertible, verificándose que

$$\hat{\theta}_{111} = \frac{1}{\hat{\theta}_{112}}$$

- [7] Procediendo de forma análoga al ejercicio 6, se obtiene la siguiente tabla:

Retardo	r_r	$\hat{\phi}_{rr}$
1	-0,909*	(0,07)
2	0,727*	(0,12)
3	-0,545	(0,14)
4	0,454	(0,15)
5	-0,364	(0,15)

A la vista de los resultados anteriores el modelo más escueto que cabe postular sería un ARMA (1,1).

Teóricamente, los coeficientes de autocorrelación en un ARMA (1,1) sigue el esquema

$$\rho_r = \phi_1 \rho_{r-1} \quad \text{tomando } \rho_1 \text{ como valor inicial.}$$

Por tanto una estimación aproximada de ϕ_1 se puede obtener de la siguiente relación:

$$\hat{\phi}_1 = \frac{r_2}{r_1} = \frac{0,727}{-0,909} = -0,800$$

Análogamente, se pueden obtener otras estimaciones de ϕ_1 de las siguientes relaciones:

$$\hat{\phi}_1 = \frac{r_3}{r_2} = \frac{-0,545}{0,727} = -0,750$$

$$\hat{\phi}_1 = \frac{r_4}{r_3} = \frac{0,454}{-0,545} = -0,833$$

$$\hat{\phi}_1 = \frac{r_5}{r_4} = \frac{-0,364}{10,454} = -0,802$$

Como puede observarse, todas las estimaciones anteriores están muy próximas entre sí, lo que confirma la hipótesis adoptada.

- [8] Procediendo de forma análoga que en el ejercicio 6, se obtiene la siguiente tabla:

Retardo	r_{rr}	$\hat{\phi}_{rr}$
1	0,777* (0,07)	0,777* (0,07)
2	0,277 (0,11)	-0,859* (0,07)
3	-0,268 (0,11)	-0,082 (0,07)
4	-0,468* (0,11)	0,041 (0,07)
5	0,701* (0,11)	-0,057 (0,07)

En los coeficientes de autocorrelación parcial se advierte un corte brusco después del segundo retardo; en cambio, los coeficientes de autocorrelación simple son todos significativos. Por tanto, parece razonable postular que el proceso generador es un AR(2). El hecho de que en valor absoluto r_1 y r_2 sean más elevados que r_3 y r_4 quedaría justificado si las raíces de la ecuación característica son complejas.

A partir de las ecuaciones de Yule-Walker para un AR(2) se puede obtener una estimación aproximada de ϕ_1 y ϕ_2 .

$$\begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0,777 \\ 0,777 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,777 \\ 0,277 \end{bmatrix} = \begin{bmatrix} 1,418 \\ -0,825 \end{bmatrix}$$

Las raíces de la ecuación característica

$$\lambda^2 - 1,418\lambda + 0,825 = 0$$

son las siguientes

$$\lambda_1 = 0,709 + 0,568i \quad \lambda_2 = 0,709 - 0,568i$$

Así pues, las raíces son complejas.

En un modelo AR(2) con $\phi_1 = 1,418$ y $\phi_2 = -0,825$ en que $\rho_1 = 0,777$, se obtendrán los siguientes valores para ρ_2 , ρ_3 , ρ_4 y ρ_5 :

$$\rho_2 = 1,418 \times 0,777 - 0,825 = 0,277$$

$$\rho_3 = 1,418 \times 0,277 - 0,825 \times 0,777 = -0,248$$

$$\rho_4 = 1,418 \times (-0,248) - 0,825 \times 0,277 = -0,580$$

$$\rho_5 = 1,418 \times (-0,580) - 0,825 \times (-0,248) = -0,618$$

Estos valores no están muy alejados de los valores muestrales suministrados, confirmado por tanto la identificación propuesta.

- [9] Procediendo como en el ejercicio 6, se obtiene la siguiente tabla:

Retardo	r_r	$\hat{\phi}_{rr}$
1	0,736* (0,10)	0,736* (0,10)
2	0,232 (0,14)	-0,541* (0,10)
3	0,035 (0,15)	0,359* (0,10)
4	-0,115 (0,15)	-0,191 (0,10)
5	0,013 (0,15)	0,047 (0,10)

La tabla anterior sugiere que el modelo sea un MA(2), ya que se produce un corte brusco en la FACE después del retardo 2, mientras que la FACPE tiene un decrecimiento progresivo en valor absoluto. No obstante conviene advertir que r_2 no es estadísticamente significativo para un nivel de confianza del 95 %, por lo que podría pensarse en un MA(1). Ahora bien, en un MA(1) se verifica que $\rho_1 \leq 0,5$, por lo que se debe descartar este supuesto al ser $r_1 = 0,736$. En un MA(2) se debe verificar que $\rho_1 \leq 0,71$; aunque $r_1 > 0,736$ puede atribuirse este hecho a errores de muestreo, ya que la diferencia con la cota máxima alcanzable es muy pequeña.

- [10]
- Los coeficientes de autocorrelación simples correspondientes a la serie original Y_t descienden lentamente de valor a medida que aumenta el orden del retardo, por lo que se puede suponer que la serie original procede de un proceso no estacionario. Este comportamiento queda eliminado al tomar primeras diferencias. Por tanto, parece correcto transformar la serie tomando primeras diferencias.
 - En la serie transformada Y_t , únicamente r_1 es estadísticamente significativo. Como $\hat{\phi}_{rr}$ decrece de forma relativamente rápida, puede postularse que el proceso gene-

rador es un MA(1) aplicado a la serie transformada ΔY_t . En consecuencia el posible modelo generador sería:

$$\Delta Y_t = \epsilon_t - \theta_1 \epsilon_{t-1}$$

- [1] Se adoptará el supuesto de que Y_t y w_t son estacionarios. Esto exige que $|\phi_1| < 1$. En el proceso original, se tiene que:

$$Y_t = \sum_{j=0}^{\infty} \phi_1^j \epsilon_{t-j}$$

La varianza teórica sería:

$$E[Y_t]^2 = E\left[\sum_{j=0}^{\infty} \phi_1^j \epsilon_{t-j}\right] = \frac{1}{1 - \phi_1^2} \sigma_\epsilon^2$$

En la serie transformada

$$w_t = Y_t - Y_{t-1} = \phi_1 Y_{t-1} + \epsilon_t - \phi_1 Y_{t-2} - \epsilon_{t-1} = \phi_1 w_{t-1} + \epsilon_t - \epsilon_{t-1}$$

operando queda:

$$\begin{aligned} w_t &= \epsilon_t - (1 - \phi_1)\epsilon_{t-1} - (1 - \phi_1)\phi_1 \epsilon_{t-2} \dots \\ &= \epsilon_t - (1 - \phi_1) \sum_{j=0}^{\infty} \phi_1^j \epsilon_{t-j-1} \end{aligned}$$

La varianza teórica de w_t será

$$E[w_t]^2 = \sigma_\epsilon^2 + (1 - \phi_1)^2 \frac{1}{1 - \phi_1^2} \sigma_\epsilon^2 = \frac{2(1 - \phi_1)}{1 - \phi_1^2} \sigma_\epsilon^2$$

$$\begin{array}{lll} \frac{E[w_t]^2}{E[Y_t]^2} & > 1 & \text{si } \phi_1 < 0,5 \\ & & \\ \frac{E[w_t]^2}{E[Y_t]^2} & = 1 & \text{si } \phi_1 = 0,5 \\ & & \\ & < 1 & \text{si } \phi_1 > 0,5 \end{array}$$

En consecuencia, si $\phi_1 > 0,5$ se podría esperar que, de promedio, la serie transformada tenga menor varianza que la serie original. Si $\phi_1 < 0,5$ se esperaría el resultado opuesto.

CAPÍTULO 5

- [1] a) Se obtiene a partir de la siguiente ecuación haciendo $Y_0 = 0$

$$\epsilon_t = Y_t - 0,8 Y_{t-1} \quad t = 1, 2, \dots, 10$$

Véase columna (1) del cuadro adjunto.

- b) Se obtiene a partir de la siguiente ecuación haciendo $Y_1 = 1,40$

$$\epsilon_t = Y_t - 0,8 Y_{t-1} \quad t = 2, 3, \dots, 10$$

Véase columna (2) del cuadro adjunto.

- c) Como se trata de un AR puro no hace falta calcular la secuencia hacia atrás de $E_c(\eta_t)$. Por ello se comienza calculando directamente

$$E_c(Y_t) = 0,8 E_c(Y_{t-1}) + E_c(\eta_t) \quad t = 0, -1, -2, \dots$$

Por hipótesis

$$E_c(\eta_t) = 0 \quad t = 0, -1, -2, -3, \dots$$

con lo que

$$E_c(Y_t) = 0,8 E_c(Y_{t+1}) \quad \text{tomando } E_c(Y_1) = Y_1$$

En la columna (3) del cuadro adjunto se han recogido estos cálculos. En realidad, solamente se requiere para su utilización posterior el cálculo $E_c(Y_0)$.

t	Y _t	Condicionado a valores iniciales		No condicionado	
		ε _t (1)	ε _t (2)	E _c (Y _t) (3)	E _c (ε _t) (4)
-2				0,72	—
-1				0,90	0,32
-0		Y ₀ = 0	Y ₁ = 1,40	1,12	0,40
1	1,40	1,40	—	Y ₁	0,50
2	-0,65	-1,77	-1,77	•	-1,77
3	-1,46	-0,94	-0,94	•	-0,94
4	-2,53	-1,36	-1,36	•	-1,36
5	-1,25	0,77	0,77	•	-0,77
6	-1,88	-0,88	-0,88	•	-0,88
7	-1,00	0,50	0,50	•	0,50
8	-1,12	-0,32	-0,32	•	-0,32
9	-0,55	0,35	0,35	•	0,35
10	-1,16	-0,72	-0,72	Y ₁₀	-0,72

Finalmente, se calcula $E_c(\epsilon_t)$ según la fórmula

$$E_c(\epsilon_t) = E_c(Y_t) - 0,8 E_c(Y_{t-1}) \quad t = 1, 2, \dots, 10$$

donde $E_c(Y_t) = Y_t$ para $t = 1, 2, \dots, 10$.

En la columna (4) del cuadro se recogen estos cálculos.

d) Como se trata de un AR(1), los tres procedimientos utilizados conducen a los mismos resultados excepto en lo que se refiere a ϵ_t .

[2]

- a) Se calcula en la columna (1) del cuadro adjunto.
- b) Se calcula en la columna (2) del cuadro adjunto.
- c) En primer lugar se calcula la secuencia hacia atrás de los errores

$$E_c[\eta_t] = w_t + 0,6E_c[\eta_{t+1}] \quad t = 10, 9, 8, \dots, 2, 1$$

Estos cálculos están recogidos en la columna (3).

A continuación se calcula en la columna (4) el valor

$$E_c[w_t] = E_c(\eta_0) - 0,6E_c[\eta_{-1}] = -0,6E_c[\eta_{-1}]$$

No se requieren para cálculos posteriores más valores hacia atrás de $E_c[w_t]$.

En la columna (5) se ha recogido el cálculo de $E_c[\epsilon_t]$.

d) Puede observarse que a diferencia de lo que sucedería con un AR, el enfoque condicionado y no condicionado dan lugar a series distintas del error. Sin embargo, a medida que crece t , las diferencias se van atenuando.

t	Y_t	$w_t = \Delta Y_t$ (1)	Enfoque condicionado ϵ_t (2)	Enfoque no condicionado		
				$E_c(\eta_t)$ (3)	$E_c(w_t)$ (4)	$E_c(\epsilon_t)$ (5)
-1				$E(\eta_{-1}) = 0$		
0	102,90		$\epsilon_0 = 0$	$E(\eta_0) = 0$	1,438	$E(\epsilon_0) = 0$
1	101,45	-1,45	-1,450	-2,396		
2	100,93	-0,52	-1,390	-1,577	w_1	-0,587
3	100,33	-0,60	-1,434	-1,762		-0,872
4	98,60	-1,73	-2,390	-1,937		-1,123
5	98,75	0,15	-1,404	-0,345		-2,404
6	98,54	-0,21	-1,053	-0,825		-1,292
7	96,71	-1,83	-2,462	-1,025		-0,985
8	97,68	0,97	-0,507	1,342		-2,421
9	98,24	0,56	0,256	0,620	w_{10}	0,483
10	98,34	0,10	0,254	0,100		0,270
						0,262

[3]

- a) Se ha calculado en la columna (1) del cuadro adjunto.
- b) Se ha calculado en la columna (2) del cuadro adjunto.
- c) En la columna (3) se ha calculado la $E_c(\eta_t)$ para $t = 9, \dots, 1$; en la columna (4) se

ha calculado la $E_c(Y_t)$ para todo $t = 0, -1, \dots, -6$; finalmente, en la columna (5) se ha calculado la $E_c[\epsilon_t]$ desde $t = -5$ hasta $t = 10$.

d) Al ser un proceso mixto existe realimentación tanto al calcular la $E_c(\eta_t)$ como al calcular la $E_c(Y_t)$. Como puede verse, la recursión hacia atrás de $E_c(Y_t)$ podría haberse continuado hasta llegar a un valor aproximadamente igual a 0. Conviene tener en cuenta que los valores que se obtienen de $E_c[\epsilon_t]$ varían al cambiar el último valor en el cálculo hacia atrás.

t	Y_t	Condicionado a valores iniciales		No condicionado		
		ϵ_t (1)	ϵ_t (2)	$E_c(\epsilon_t)$ (3)	$E_c(Y_t)$ (4)	$E_c(\epsilon_t)$ (5)
-6					$E(\eta_{-6}) = 0$	0,171
-5					$E(\eta_{-5}) = 0$	0,214
-4					$E(\eta_{-4}) = 0$	0,267
-3					$E(\eta_{-3}) = 0$	0,334
-2					$E(\eta_{-2}) = 0$	0,418
-1					$E(\eta_{-1}) = 0$	0,522
-0		$(Y_0 = 0, \epsilon_0 = 0)$	$(Y_1 = 2,66, \epsilon_1 = 0)$	$E(\eta_0) = 0$	$E(\eta_0) = 0$	0,652
1	2,66	2,660			2,459	2,406
2	1,26	0,728	-0,868		1,346	0,575
3	-0,23	-0,801	-1,759		-0,164	-0,893
4	1,18	0,883	0,309		1,683	0,828
5	-1,30	-1,714	-2,059		-0,895	-1,747
6	-0,64	-0,628	-0,835		-0,178	-0,648
7	0,17	0,305	0,181		0,997	0,293
8	-0,98	-0,933	-1,007		0,071	-0,940
9	0,12	0,344	0,300		1,912	0,340
10	-2,24	-2,129	-2,156	$\eta_{10} = 0$	$\eta_{10} = 0$	-2,132

[4]

- a) $c_1 = Y_1 + 0,6\epsilon_0 = c_1 + 0,6\epsilon_0 = 3 + 0,6\epsilon_0$
 - $c_2 = Y_2 + 0,6c_1 = Y_2 + 0,6Y_1 + 0,6^2\epsilon_0 = Y_2 + 0,6c_1 + 0,6^2\epsilon_0 = c_1 + 0,6^2\epsilon_0 = 3,8 + 0,6^2\epsilon_0$
 - $c_3 = Y_3 + 0,6c_2 = Y_3 + 0,6c_2 + 0,6^3\epsilon_0 = c_1 + 0,6^3\epsilon_0 = 2,28 + 0,6^3\epsilon_0$
 - $c_4 = Y_4 + 0,6c_3 = Y_4 + 0,6c_3 + 0,6^4\epsilon_0 = c_4 + 0,6^4\epsilon_0 = 3,368 + 0,6^4\epsilon_0$
 - $c_5 = Y_5 + 0,6c_4 = Y_5 + 0,6c_4 + 0,6^5\epsilon_0 = c_4 + 0,6^5\epsilon_0 = 3,0208 + 0,6^5\epsilon_0$
- La expresión S condicionada a ϵ_0 viene dada por:

$$S = \sum_{t=1}^5 [\epsilon_t/\epsilon_0]^2 = \sum_{t=1}^5 [c_t + 0,6^t\epsilon_0]^2$$

donde

$$c_t = \sum_{j=0}^{t-1} 0,6^j Y_{t-j}$$

- b) Para minimizar S se deriva la expresión obtenida con respecto a ε_0 y se iguala a 0.

$$\frac{\partial S}{\partial \varepsilon_0} = 2 \sum_{t=1}^5 [c_t + 0,6\varepsilon_0] 0,6^t = 0$$

con lo que

$$\varepsilon_0 = \frac{-\sum c_t 0,6^t}{\sum 0,6^{2t}} = 7,748$$

- 5 a) Utilizando el programa de ordenador BMDP2T¹ se obtienen los siguientes resultados, tomando como valor inicial $\hat{\phi}_1^0 = 0,6$:

Iteración	SCR	$\hat{\phi}_1$
1	14,19	0,483
2	14,19	0,482

Obsérvese que prácticamente coinciden las SCR y la estimación de $\hat{\phi}_1$ en las dos iteraciones. Desde el punto de vista teórico como se trata de un modelo autorregresivo estimado por un método condicionado, en la primera iteración se debe obtener directamente el resultado final, ya que es un modelo lineal.

- b) Utilizando el programa de ordenador BMDP2T se obtienen los siguientes resultados:

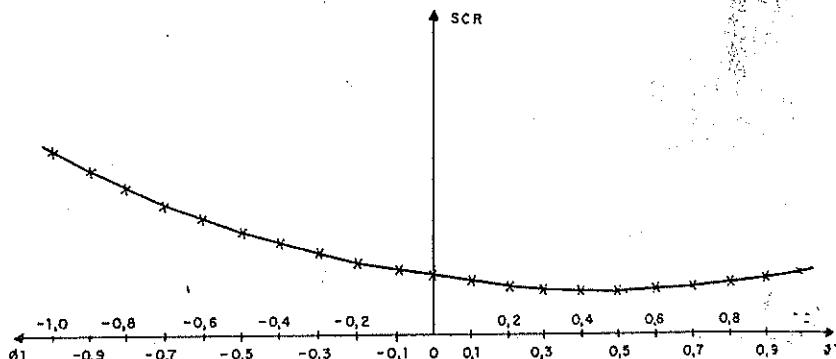
Iteración	SCR	$\hat{\phi}_1$
1	16,15	0,529
2	16,15	0,543
3	16,15	0,548
4	16,15	0,549

SCR (excluidos los obtenidos hacia atrás) = 14,29.

c)

$\hat{\phi}_1$	-1,0	-0,9	-0,8	-0,7	-0,6	-0,5	-0,4	-0,3	-0,2	-0,1	0,0
SCR	57,01	51,48	46,27	41,46	37,04	33,01	29,37	26,12	23,26	20,80	18,72
$\hat{\phi}_1$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	
SCR	17,03	15,74	14,83	14,32	14,19	14,46	15,12	16,17	17,61	19,44	

¹BMDP: Biomedical Computer Programs, Series P, edición 1981. Department of Biomathematics, University of California, Los Angeles.



- 6 a) Utilizando el programa de ordenador BMDP2T se obtienen los siguientes resultados, tomando como valor inicial $\hat{\phi}_1^0 = -0,8$:

Iteración	SCR	$\hat{\phi}_1$
1	26,11	-0,454
2	25,75	-0,320
3	25,67	-0,262
4	25,66	-0,232
5	25,65	-0,218
6	25,65	-0,212

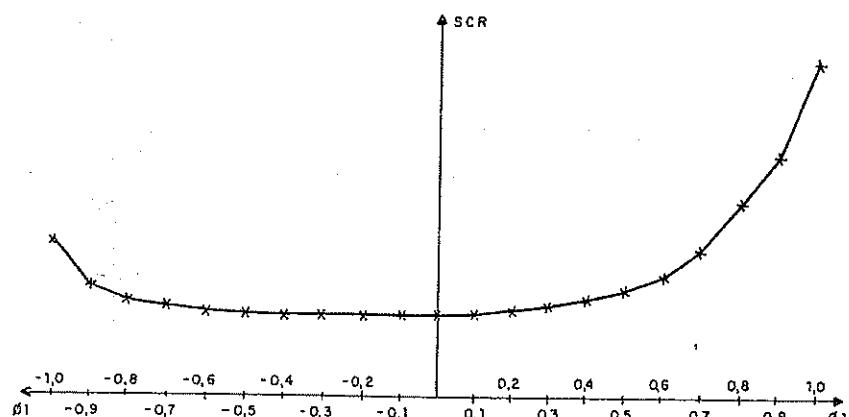
- b) Utilizando, como valor inicial, el calculado en la última iteración de la estimación condicionada se obtiene el siguiente resultado:

Iteración	SCR	$\hat{\phi}_1$
1	25,64	-0,210

SCR (excluido los obtenidos hacia atrás) = 25,64.

c)

θ_1	-1,0	-0,9	-0,8	-0,7	-0,6	-0,5	-0,4	-0,3	-0,2	-0,1	0,0
SCR	60,91	42,22	33,33	29,16	27,23	26,35	25,92	25,72	25,65	25,76	26,16
θ_1	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	
SCR	26,99	28,49	31,00	35,02	41,35	51,24	66,66	90,67	127,51	182,19	



- 7] a) Utilizando el programa BMDP2T se obtienen los siguientes resultados tomando como valores iniciales $\hat{\phi}_0^0 = 0,5$ y $\hat{\theta}_0^0 = 0,4$

Iteración	SCR	$\hat{\phi}_1$	$\hat{\theta}_1$
1	5,063	0,884	-0,046
2	5,041	0,889	0,023
3	5,039	0,898	0,047
4	5,038	0,901	0,058
5	5,038	0,903	0,063
6	5,038	0,903	0,066

- b) Utilizando como valores iniciales los calculados en la última iteración de la estimación condicionada, se obtienen los siguientes resultados:

Iteración	SCR	$\hat{\phi}_1$	$\hat{\theta}_1$
1	5,661	0,927	0,127
2	5,636	0,951	0,174
3	5,435	0,984	0,217
4	5,364	0,981	0,257
5	5,361	0,991	0,273
6	5,360	0,992	0,284
7	5,360	0,992	0,292
8	3,359	0,992	0,298
9	3,359	0,992	0,303
10	3,359	0,992	0,306

SCR (excluidos de los obtenidos hacia atrás) = 5,017.

CAPÍTULO 6

- 1] Utilizando como una primera aproximación que

$$\text{var}(r_t) = 1/T$$

$$\text{var}(\hat{\epsilon}_t) = 1/T$$

resultan estadísticamente significativos para un nivel del 5%: r_1 , $\hat{\phi}_{11}$, $\hat{\phi}_{22}$, $\hat{\phi}_{33}$.

Teniendo en cuenta también que en $\hat{\phi}_{44}$ y $\hat{\phi}_{55}$ continúa el decrecimiento de carácter exponencial iniciado en los coeficientes anteriores, se puede deducir que el ruido $\hat{\epsilon}_t$ sigue un proceso aproximado MA(1).

Como en un MA(1) con $\theta_1 = 1$, se verifica que $\rho_1 = 0,5$ se tiene que aproximadamente

$$\hat{\epsilon}_t = (1 - L)\epsilon_t$$

siendo ϵ_t un ruido blanco.

Sustituyendo en el modelo ajustado se tiene que $(1 - L)Y_t = (1 - 0,324L)(1 - L)\epsilon_t$.

Simplificando queda

$$Y_t = (1 - 0,324L)\epsilon_t$$

En consecuencia el modelo reformulado sería un ARIMA (0, 0, 1) en lugar del modelo ajustado ARIMA (0, 1, 1).

- [2]** Tomando como una primera aproximación de la varianza de r_i , y de $\hat{\phi}_{ii}$ a $1/T$, resulta que solamente se pueden considerar significativos para un nivel del 5 % a r_i y a $\hat{\phi}_{ii}$. No obstante, se observa que la FACE sigue un decrecimiento exponencial por lo que se podría formular el siguiente modelo aproximado para el ruido

$$(1 - 0,380L)\hat{\epsilon}_t = \epsilon_t$$

con lo que sustituyendo en el modelo ajustado, quedaría

$$(1 - 0,8L)Y_t = (1 + 0,6L)(1 - 0,380L)^{-1}\epsilon_t$$

es decir,

$$(1 - 0,8L)(1 - 0,380L)Y_t = (1 + 0,6L)\epsilon_t$$

En consecuencia el modelo reformulado sería del tipo ARIMA (2, 0, 1).

- [3]** Aplicando el test de Chow propuesto en (6-26) se tiene que

$$F_{1,298}^* = \frac{[6,513 - (2,354 + 3,824)]}{(2,354 + 3,824)/298} = 16,16$$

Como el número de grados de libertad del denominador es muy elevado se puede considerar que es una $F_{1,\infty}$ o, lo que es lo mismo, una χ^2_1 . Para un nivel de significación del 1 %, el valor de las tablas es $F_{1,\infty}(0,01) = 6,63$.

Como

$$F_{1,298}^* > F_{1,\infty}(0,01)$$

se rechaza para un nivel del 1 % la H_0 de que las dos submuestas han sido generadas por el mismo proceso generador.

- [4]** Si el parámetro ϕ_i del proceso

$$Y_t = \phi_i Y_{t-1} + \epsilon_t$$

se estima por MC, haciendo el supuesto de que $\bar{Y} = 0$ se tiene que

$$R^2 = \frac{\sum \hat{Y}_t^2}{\sum Y_t^2} = \frac{\sum [\hat{\phi}_i Y_{t-1}]^2}{\sum Y_t^2} = \hat{\phi}_i^2 \frac{\sum Y_{t-1}^2}{\sum Y_t^2} \approx \hat{\phi}_i^2$$

CAPÍTULO 7

[1]

a)

$$Y_{T+i} = Y_{T+i} + \epsilon_{T+i} + 0,8\epsilon_{T+i-1}$$

- b) Para expresar el modelo propuesto como un MA(∞) es preciso calcular $\psi(L)$ a partir de la relación

$$\phi(L)\psi(L) = \theta(L)$$

es decir,

$$(1 - L)(1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots) = (1 + 0,8L)$$

$$1 + (\psi_1 - 1)L + (\psi_2 - \psi_1)L^2 + (\psi_3 - \psi_2)L^3 + \dots = 1 + 0,8L$$

Igualando cada uno de los coeficientes de las potencias de L en ambos miembros se obtiene:

$$\psi_1 = 1,8; \quad \psi_2 = 1,8; \quad \psi_3 = 1,8\dots$$

Por tanto,

$$Y_{T+i} = \epsilon_{T+i} + 1,8\epsilon_{T+i-1} + 1,8\epsilon_{T+i-2} + \dots$$

- c) Para expresar el modelo propuesto como un AR(∞), se calculan los coeficientes de $\pi(L)$ utilizando la relación

$$\phi(L) = \pi(L)\theta(L)$$

obteniéndose

$$\pi_1 = 1,8 \quad \pi_2 = -1,44 \quad \pi_3 = 1,152\dots$$

En consecuencia

$$Y_{T+i} = \epsilon_{T+i} + 1,8Y_{T+i-1} - 1,44Y_{T+i-2} + 1,152Y_{T+i-3}$$

- [2]** Los parámetros autorregresivos generalizados se obtienen efectuando el producto:

$$\begin{aligned} \phi(L) &= \phi(L)(1 - L) \\ &= (1 - 0,5L)(1 - L) \\ &= 1 - 1,5L + 0,5L^2 \end{aligned}$$

- a) $Y_{T+i} = 1,5Y_{T+i-1} - 0,5Y_{T+i-2} + \epsilon_{T+i} + 0,6\epsilon_{T+i-1}$
 b) $Y_{T+i} = \epsilon_{T+i} + 2,1\epsilon_{T+i-1} + 2,65\epsilon_{T+i-2} + 2,925\epsilon_{T+i-3} + \dots$
 c) $Y_{T+i} = \epsilon_{T+i} + 2,1Y_{T+i-1} - 1,76Y_{T+i-2} + 1,056Y_{T+i-3} + \dots$

3 Expresando el predictor $\tilde{Y}_{T+1/T}$ en función de los coeficientes ψ_i se tiene que

$$\tilde{Y}_{T+1/T} = \psi_1 e_T + \psi_{l+1} e_{T+1} + \dots$$

Al conocerse la observación Y_{T+l} , la minimización del ECM según (7-15) y (7-16) lleva al predictor

$$\tilde{Y}_{T+1/T+1} = \psi_{l-1} e_{T+1} + \psi_l e_T + \psi_{l+1} e_{T+2} + \dots$$

Comparando con el predictor $\tilde{Y}_{T+1/T}$ se tiene que

$$\tilde{Y}_{T+1/T-1} = \tilde{Y}_{T+1/T} + \psi_{l-1} e_{T+1}$$

Es decir, a la predicción anterior se le añade el error correspondiente a la observación $T+1$ ponderado por el peso ψ_{l-1} . A esta operación se le denomina actualización —updating— de la predicción. Teniendo en cuenta que

$$e_{T+1} = Y_{T+1} - \tilde{Y}_{T+1/T}$$

el predictor para el periodo $T+l$ quedaría expresado de la siguiente forma:

$$\tilde{Y}_{T+1/T+1} = \tilde{Y}_{T+1/T} + \psi_{l-1} (Y_{T+1} - \tilde{Y}_{T+1/T})$$

que sería más operativa.

4 Para $l > j$

$$\begin{aligned} E(e_{T+1/T} e_{T+j-T-j}) &= E(Y_{T+1} - \tilde{Y}_{T+1/T}) (Y_{T+j-T-j} - \tilde{Y}_{T+j-T-j}) \\ &= E(e_{T+1} + \psi_1 e_{T+1-1} + \dots + \psi_j e_{T+j-1} + \dots + \psi_{j-1} e_{T+1} \times \\ &\quad \times (e_{T+1-j} + \psi_1 e_{T+1-j-1} + \dots + \psi_{l-1} e_{T+1-j})) \\ &= \sigma_e^2 \sum_{i=j}^{l-1} \psi_i \psi_{l-i} \end{aligned}$$

siendo $\psi_0 = 1$.

Para $l \leq j$

$$E(e_{T+1/T} e_{T+j-T-j}) = 0$$

ya que en ese caso no existe ningún solapamiento de los ruidos que integran cada uno de los errores de predicción.

$$\begin{aligned} \text{5} \quad E(e_{T+1/T} e_{T+j-T}) &= E(Y_{T+1} - \tilde{Y}_{T+1/T}) (Y_{T+j-T} - \tilde{Y}_{T+j-T}) = \\ &= E(e_{T+1} + \psi_1 e_{T+1-1} + \dots + \psi_{l-1} e_{T+1}) = \\ &\quad \times (e_{T+j-T} + \psi_1 e_{T+j-T-1} + \dots + \psi_{l-1} e_{T+j-T} + \dots + \psi_{l-j-1} e_{T+1}) \\ &= \sigma_e^2 \sum_{i=0}^{l-1} \psi_i \psi_{l-i} \end{aligned}$$

a) Para $t = 11$, el predictor será

$$\tilde{Y}_{11/11/10} = 18 + 0,5 e_{110} + 0,4 e_{109}$$

No se dispone de información directa acerca de e_{109} y e_{110} , pero se pueden estimar a partir de la información suministrada de Y . El procedimiento, haciendo el supuesto de que $e_{104} = 0$ y $e_{105} = 0$, podría ser el siguiente:

$$e_{105} = Y_{105} - 18 - 0,5 e_{104} - 0,4 e_{105} = 2,00$$

$$e_{107} = Y_{107} - 18 - 0,5 e_{106} - 0,4 e_{107} = 2,00$$

$$e_{108} = Y_{108} - 18 - 0,5 e_{107} - 0,4 e_{108} = -0,80$$

$$e_{109} = Y_{109} - 18 - 0,5 e_{108} - 0,4 e_{109} = 0,60$$

$$e_{110} = Y_{110} - 18 - 0,5 e_{109} - 0,4 e_{110} = -0,98$$

En consecuencia

$$\tilde{Y}_{11/11/10} = 17,75$$

$$\tilde{Y}_{11/21/10} = 18 + 0,4 e_{110} = 17,608$$

$$\tilde{Y}_{11/31/10} = 18$$

b) $\text{ECM}(\tilde{Y}_{11/11/10}) = \sigma_e^2 = 4$

$$\text{ECM}(\tilde{Y}_{11/21/10}) = (1 + \psi_1^2) \sigma_e^2 = 5$$

$$\text{ECM}(\tilde{Y}_{11/31/10}) = (1 + \psi_1^2 + \psi_2^2) \sigma_e^2 = 5,64$$

c) De acuerdo con los resultados obtenidos en el ejercicio 3 se tiene que

$$\begin{aligned} \tilde{Y}_{T+2/T+1} &= \tilde{Y}_{11/21/10} = \tilde{Y}_{11/31/10} + \psi_1 (Y_{111} - \tilde{Y}_{11/11/10}) = \\ &= 17,608 + 0,5(17 - 17,75) = 17,233 \\ \tilde{Y}_{T+3/T+1} &= \tilde{Y}_{11/31/10} = 18 + 0,4(17 - 17,75) = 17,7 \end{aligned}$$

a) $\tilde{Y}_{81/80} = 0,4 Y_{80} + 0,6 Y_{79} + 0,8 - 0,5 e_{80} = 41,100$

$$\tilde{Y}_{82/80} = 0,4 \tilde{Y}_{81/80} + 0,6 Y_{80} + 0,8 = 41,840$$

$$\tilde{Y}_{83/80} = 0,4 \tilde{Y}_{82/80} + 0,6 \tilde{Y}_{81/80} + 0,8 = 42,196$$

b) Los pesos ψ_i que se obtienen son los siguientes:

$$\psi_1 = -0,1 \quad \psi_2 = 0,56$$

Por tanto

$$\text{ECM}(\tilde{Y}_{81/80}) = \sigma_e^2 = 1$$

$$\text{ECM}(\tilde{Y}_{82/80}) = (1 + \psi_1^2) \sigma_e^2 = 1,01$$

$$\text{ECM}(\tilde{Y}_{83/80}) = (1 + \psi_1^2 + \psi_2^2) \sigma_e^2 = 1,324$$

- c) Para Y_{T+1} $41,100 \pm 1,65 \sqrt{1} \rightarrow (39,450; 42,750)$
 Para Y_{T+2} $41,840 \pm 1,65 \sqrt{1,01} \rightarrow (40,182; 43,498)$
 Para Y_{T+3} $42,196 \pm 1,65 \sqrt{1,324} \rightarrow (40,297; 44,095)$
- d) $\tilde{Y}_{s2/s1} = \tilde{Y}_{s2/s0} + \psi_1(Y_{s1} - \tilde{Y}_{s1/s0}) = 41,85$
 $\tilde{Y}_{s3/s1} = \tilde{Y}_{s3/s0} + \psi_2(Y_{s1} - \tilde{Y}_{s1/s0}) = 42,14$

CAPITULO 8

1 La FACT correspondiente al modelo a) viene dada en (8-19) con sólo sustituir s por 12.

Para calcular la FACT del modelo b) se multiplican ambos miembros del modelo por Y_{t+1} , se toman esperanzas, y se dan distintos valores a τ . Después dividiendo por y_0 , los resultados que se obtienen son los siguientes:

$$\rho_\tau = \begin{cases} \frac{-\theta_1 + \theta_{12}\theta_{13}}{1 + \theta_1^2 + \theta_{12}^2 + \theta_{13}^2} & \tau = 1 \\ \frac{\theta_1\theta_{12}}{1 + \theta_1^2 + \theta_{12}^2 + \theta_{13}^2} & \tau = 11 \\ \frac{-\theta_{12} + \theta_1\theta_{13}}{1 + \theta_1^2 + \theta_{12}^2 + \theta_{13}^2} & \tau = 12 \\ \frac{-\theta_{13}}{1 + \theta_1^2 + \theta_{12}^2 + \theta_{13}^2} & \tau = 13 \end{cases}$$

Es claro que el modelo a) es un caso particular del modelo b) en que:

$$\theta_{13} = -\theta_1\theta_{12}$$

Fuera de este caso particular, en el modelo b) se verificará que $\rho_{11} \neq \rho_{13}$ con lo que se pierde la propiedad de simetría que tiene el modelo a).

2 A partir de un par de raíces complejas $a \pm i\theta$, se pueden obtener los coeficientes de la ecuación de 2.º grado

$$1 - \phi_1 L - \phi_2 L^2 = 0$$

haciendo:

$$\phi_1 = 2a \quad \phi_2 = \frac{-4\theta^2 - \phi_1^2}{4}$$

De esta forma, y teniendo en cuenta además las raíces reales (11) y (12), se obtienen los siguientes factores:

$$\begin{aligned} (1)-(2) & 1 - \sqrt{3}L + L^2 \\ (3)-(4) & 1 - L + L^2 \\ (5)-(6) & 1 + L^2 \\ (7)-(8) & 1 + L + L^2 \\ (9)-(10) & 1 + \sqrt{3}L + L^2 \\ (11) & 1 + L \\ (12) & 1 - L \end{aligned}$$

Efectuando el producto, ordenando adecuadamente los factores, se tiene que:

$$\begin{aligned} & (1 - \sqrt{3}L + L^2)(1 + \sqrt{3}L + L^2) \times (1 - L + L^2)(1 + L + L^2) \times (1 + L^2)(1 + L)(1 - L) \\ & = (1 - L^2 + L^4)(1 + L^2 + L^4)(1 - L^4) = \\ & = (1 - L^2 + L^4)(1 + L^2 - L^6 - L^8) \\ & = (1 - L^{12}) \end{aligned}$$

b) Cada par de raíces complejas se puede expresar de forma polar:

$$r^2 = \sqrt{a^2 + \theta^2} \quad (\text{módulo})$$

$$\omega = \text{arc cos } a/r \quad (\text{frecuencia angular})$$

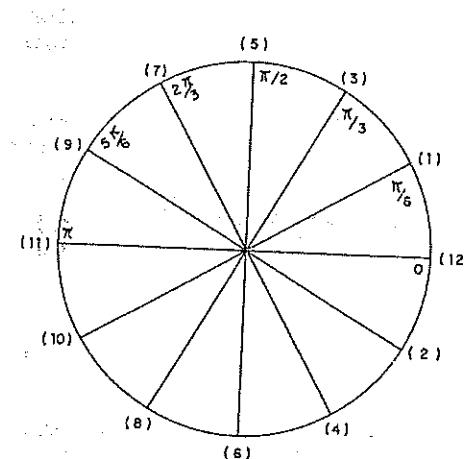
A partir de aquí se puede definir el período (Pe) y las frecuencias en ciclos por año (FCA)

$$Pe = 2\pi/\omega$$

$$FCA = 12/Pe$$

	r^2	ω	Período (Pe)	Frecuencia ciclos por año (FCA)
(1)-(2)	$3/4 + 1/4 = 1$	$\text{arcos } \sqrt{3}/2 = \pi/6$	12	1
(3)-(4)	$1/4 + 3/4 = 1$	$\text{arcos } 1/2 = 2\pi/6$	6	2
(5)-(6)	$0 + 1 = 1$	$\text{arcos } 0 = \pi/2$	4	3
(7)-(8)	$1/4 + 3/4 = 1$	$\text{arcos } (-1/2) = 2\pi/3$	3	4
(9)-(10)	$3/4 + 1/4 = 1$	$\text{arcos } (-\sqrt{3}/2) = 5\pi/6$	12/5	5
(11)	1	$\text{arcos } (1) = 0$	2	6
(12)	1	$\text{arcos } (1) = 0$	—	constante

c)



3] El error de predicción global para el año 21 vendrá dado por

$$e_{21/20}^g = \text{PIBI}_{21}^* - \text{PIBI}_{21/20}^* =$$

$$= \sum_{l=1}^4 [\text{PIBI}_{4 \times 21 + l/4 \times 21} - \text{PIBI}_{4 \times 21 + l}] =$$

$$= \sum_{l=1}^4 \varepsilon_{4 \times 21 + l}$$

Por tanto, se deduce que

$$\begin{aligned} \text{ECM} &= E[e_{21/20}^g] = E\left[\sum_{l=1}^4 \varepsilon_{4 \times 21 + l}\right]^2 = \\ &= 4\sigma_e^2 \end{aligned}$$

Referencias bibliográficas

ABRAHAM, B., y BOX, G. E. P. (1978): *Deterministic and forecast adaptive time dependent models*. Applied Statist., 27, págs. 120-130.

ABRAHAM, B., y LEDOLTER, J. (1983): *Statistical Methods for Forecasting*. John Wiley.

AKAIKE, H. (1974): *A new look at the statistical models identification*. IEEE Transactions on Automatic Control, AC-19, págs. 716-723.

ALCAIDE, J. (1976): *Una revisión urgente de la serie de renta nacional española en el siglo XX*, en la obra «Datos básicos para la historia financiera de España, 1850-1975», vol. I, págs. 1127-1150. Instituto de Estudios Fiscales. Madrid.

AMARA, R. C. (1972): *Methodology: A Note on Cross-Impact Analysis*. Futures, 4, págs. 267-271.

AMENT, R. H. (1970): *Comparison of Delphi Forecasting Studies in 1964 and 1969*. Futures, 3, págs. 35-44.

ANDERSON, R. L. (1942): *Distribution of the serial correlation coefficient*. The Annals of Mathematical Statistics, 13, págs. 1-13.

ANSLEY, C. F. (1979): *An algorithm for the exact likelihood of a mixed autoregressive moving average process*. Biometrika, 66, págs. 59-65.

BARTLETT, M. S. (1946): *On the theoretical specification of sampling properties of autocorrelated time series*. Journal of Royal Statistical Society, B, 8, págs. 27-41.

BEGUIN, J. M.; GOURIEROUX, C., y MONFORT, A. (1980): *Identification of a mixed autoregressive-moving average process: the corner method*. En «Time Series», ed O.D. Anderson, North-Holland, págs. 423-436.

BOX, G. E. P. y COX, D. R. (1964): *An Analysis of Transformations*. Journal of the Royal Statistical Society, B, 26, págs. 211-252.

- BOX, G. E. P., y JENKINS, G. M. (1976): *Time series Analysis. Forecasting and control.* Holden Day, segunda edición. (La primera edición es de 1970.)
- BOX, G. E. P., y PIERCE, D. A. (1970): *Distribution of residual autocorrelations in autoregressive integrated moving average time series models.* Journal of the American Statistical Association, 65, págs. 1509-1526.
- BOX, G. E. P., y TIAO, G. C. (1975): *Intervention Analysis with Applications to Economic and Environmental Problems.* Journal of the American Statistical Association, 70, págs. 70-79.
- BROWN, R. G. (1956): *Exponential Smoothing for Predicting Demand.* Comunicación presentada en el Tenth National Meeting of the Operations Research Society of America. San Francisco.
- BROWN, R. G. (1959): *Statistical Forecasting for Inventory Control.* McGraw-Hill.
- BROWN, R. G. (1963): *Smoothing Forecasting and Prediction of Discrete Time Series.* Prentice-Hall.
- BROWN, R. G., y MEYRES, R. F. (1961): *The Fundamental Theorem of Exponential Smoothing.* Operations Research, 9, págs. 673-685.
- CLEVELAND, D. S. (1972): *The inverse autocorrelations of a time series and their applications.* Technometrics, 14, págs. 277-293.
- CHOW, G. C. (1960): *Test of equality between sets of coefficients in two linear regressions.* Econometrica, 28, págs. 591-605.
- DAGUM, E. D. (1975): *Seasonal factor forecasts from ARIMA models.* Comunicación presentada en la sesión 40 del International Statistical Institute. Varsovia. Polonia.
- DAGUM, E. B. (1978): *Modelling, Forecasting and Seasonally adjusting Economic Time Series with the X-II ARIMA Method.* The Statistician, 27, páginas 203-216.
- DAVIES, N; PATE, M. B., y FROST, M. G. (1974): *Maximum autocorrelations for moving average processes.* Biometrika, 61, págs. 191-200.
- DAVIES, N; TRIGGS, C. M., y NEWBOLD, P. (1977): *Significance Levels of the Box-Pierce Portmanteau Statistic in Finite Samples.* Biometrika, 64, páginas 517-522.
- DENT, W. (1977): *Computation of the exact likelihood function of an ARIMA process.* Journal of Statist. Comput. Simulation, 5, págs. 193-206.

- DURBIN, J. (1960): *The fitting of time series models.* Review of International Statistical Institute, 28, págs. 233-244.
- ENGLE, R. F. (1978): *Estimating structural models of seasonality.* En Zellner, A (ed). «Seasonal Analysis of Economic Time Series», págs. 281-308, Bureau of Census, Washington.
- ESPASA, A. (1979): *Modelos ARIMA univariantes, con análisis de intervención para las series de agregados monetarios (Saldos medios mensuales) M3 y M2.* Banco de España.
- FERNÁNDEZ, R. E. (1984): *Análisis de identificación de modelos ARIMA mediante simulación con métodos de Montecarlo.* Tesis doctoral. Universidad de Santiago de Compostela.
- GEA, J.; y URIEL, E. (1984): *Identificación automática mediante la función de autocorrelación extendida.* Comunicación presentada al IX Simposium de Teoría Económica y Econometría. Universidad Autónoma de Barcelona.
- GOLDBERGER, A. S. (1964): *Econometric theory.* John Wiley.
- GORDON, T. J., y HAYWARD, H. (1968): *Initial Experiment with the Cross-Impact Matrix Method of Forecasting.* Futures, 1.
- GRANGER, C. W. J., y NEWBOLD, P. (1977): *Forecasting Economic Time Series.* Academic Press.
- GRAY, W.; KELLEY, G. D., y McINTIRE, D. D. (1978): *A new approach to ARMA modeling.* Communications in Statistics, Simulation and Computation, B7-1, páginas 1-77.
- HAMILTON, D., y WATTS, D. (1978): *Interpreting partial autocorrelation functions of seasonal time series models.* Biometrika, 65, págs. 135-140.
- HARVEY, A. C. (1981): *Time Series Models.* Philip Allan. Deddington. Oxford.
- HARVEY, A. C. (1984): *A Unified View of Statistical Forecasting Procedures.* Journal of Forecasting, 3, págs. 245-275.
- HELMER, O. (1966b): *The Delphi Method Systematizing Judgements about the Future. Innovations.* Santa Mónica, California. The RAND Corp.
- HELMER, O. (1966b): *The Delphi Method Systematizing judgements about the Future.* University of California. Los Angeles.
- HELMER, O., y RESCHER, N. (1959): *On the Epistemology of the Inexact Sciences.* Management Science, 6.

REFERENCIAS BIBLIOGRÁFICAS

- HERNÁNDEZ, D., y ASIAIN, J. (1977): *Técnicas de simulación en los procesos de planificación*. Cuadernos del Instituto de Planificación. Presidencia del Gobierno. Madrid.
- HIPEL, K. W., McLEOD, A. I., y LENNOX, W. C. (1977): *Advances in Box-Jenkins modelling I. Model construction*. Water Resources Research, 14, págs. 567-575.
- HOLT, C. C. (1957): *Forecasting Trends on Seasonals by Exponentially Weighted Moving Averages*. Carnegie Institute of Technology, Pennsylvania.
- JENKINS, G. M. (1979): *Practical Experiences with Modelling and Forecasting Time Series*. Jersey. Channel Islands. A GPJ Publication.
- JENKINS, G. M., y ALAVI, A. S. (1981): *Some aspects of Modelling and Forecasting Multivariate Time Series*. Journal of Time Series Analysis, 2, 1-47.
- JENKINS, G. M., y WATTS, D. G. (1968): *Spectral Analysis and Its Applications*. Holden Day.
- LINDLEY, D. V. (1971): *Making decisions*. John Wiley. Existe traducción castellana de J. M. Bernardo con el título *Principios de la teoría de la decisión*, 1977. Editorial Vicens Vives.
- LINSTONE, H. A., y TUROFF, M. (1975): *The Delphi Method: Techniques and Applications*. Addison-Wesley. Reading. Massashussets.
- LJUNG, G. M., y BOX, G. E. P. (1978): *On a measure of lack of fit in time series models*. Biometrika, 66, págs. 297-303.
- MAKRIDAKIS, S., y WHEELWRIGHT, S. C. (1978): *Forecasting Methods and Applications*. John Wiley.
- MARQUARDT, D. W. (1963): *An algorithm for least squares estimation of non linear models*. J. Soc. Ind. Appl. Math., 11, págs. 431-444.
- MOLTO, M. L., y URIEL, E. (1984): *Relación entre nivel de actividad en el sector industrial y expectativas empresariales en la Comunidad Valenciana*. Comunicación Presentada al I Congreso de Economía Valenciana.
- MUTH, R. F. (1960): *Optimal Properties of Exponentially Weighted Forecasts*. Journal of American Statistical Association, 55, págs. 299-306.
- NAYLOR, H. T., BALINTFY, J. L., BURDICK, D. S., y CHU, K. (1966): *Computer Simulation Techniques*. John Wiley.
- NELSON, C. R., y GRANGER, C. W. J. (1979): *Experience with using the Box-Cox transformation when forecasting economic time series*. Journal of Econometrics, 10, págs. 57-69.

REFERENCIAS BIBLIOGRÁFICAS

- NERLOVE, M., GREther, D. M., y CARVALHO, J. L. (1978): *Analysis of Economic Time Series*. Academic Press.
- NEWBOLD, P. (1974): *The exact likelihood function for a mixed autoregressive moving average process*. Biometrika, 61, págs. 423-426.
- PANKRATZ, A. (1983): *Forecasting with Univariate Box-Jenkins Model. Concepts and cases*. John Wiley.
- PEÑA, D. (1979): *Interacción en la identificación de los modelos ARIMA univariantes*. Cuadernos Económicos del ICE, 11-12, págs. 9-35.
- PROTHERO, D. L., y WALLIS, K. F. (1976): *Modelling Macroeconomic Time Series*. Journal of the Royal Statistical Society, A, 139, págs. 468-486.
- PULIDO, A. (1983): *Modelos econométricos*. Ediciones Priárnide. Madrid.
- QUENQUILLE, M. H. (1949): *Approximate test of correlation time series*. Journal of the Royal Statistical Society Series B, 11, págs. 68-84.
- RAND CORPORATION (1955): *A million Random Digits with 100000 Normal Deviates*. Glencoe. Illinois. The Free Press.
- ROCHBERG, R., GORDON, T. J., y HELMER, O. (1970): *The Use of Cross-Impact Matrices for Forecasting and Planning*. The Institute for the future, IFF Report R-10. Middletown, Connecticut.
- ROYO, E., y FERRER, S. (1955): *Tablas auxiliares de estadística*. Consejo Superior de Investigaciones Científicas. Madrid.
- SACKMAN, H. (1975): *Delphi Critique*. Lexington Books, Lexington, Massachusetts.
- SANZ, R. (1978): *Modelización del índice de producción industrial y su relación con el consumo de energía eléctrica*. Documento de trabajo n.º 8403, Servicio de Estudios. Banco de España.
- SHISKIN, J., YOUNG, A. H., y MUSGRAVE, J. C. (1967): *The variant of Census II Method Seasonal adjustment Program*. Bureau of the Census. Technical Paper, 15. Washington.
- STRALKOWSKI, C. M., WU, S. M., y DEVOR, R. E. (1970): *Charts for the interpretation and estimation of the second order autoregressive model*. Technometrics, 12, págs. 669-685.

REFERENCIAS BIBLIOGRÁFICAS

- STRALKOWSKI, C. M., WU, S. M., y DEVOR, R. E. (1974): *Charts for the interpretation and estimation of the second order moving average and mixed first order auto-regressive-moving average models.* Technometrics, 16, págs. 275-285.
- TIAO, G. C., y BOX, G. E. P. (1981): *Modeling multiple time series with applications.* Journal of the American Statistical Association, 76, págs. 802-816.
- TIAO, G. C., y TSAY, R. S. (1981): *Consistency properties of least squares estimates of autoregressive parameters in ARMA models.* Technical Report 658. University of Wisconsin. Madison.
- TINTNER, G. (1940): *The Variate Difference Method.* Principia Press. Bloomington. Indiana.
- TINTNER, G., y RAO, J. N. K. (1963): *On the variate difference method.* Australian Journal of Statistics, 5, págs. 106 y ss.
- TREADWAY, A. B., GARCÍA PARDO, J., y CARBAJO, A (1978): *Efectos sobre la economía española de una devaluación de la peseta.* Fundación Ramón Areces. Madrid.
- TRIGG, D. W. (1964): *Monitoring a Forecasting System.* Operational Research Quarterly, 15, págs. 271-274.
- TRIGG, D. W., y LEACH, D. H. (1967): *Exponential Smoothing with an Adaptive Response Rate.* Operational Research Quaterly, 18, págs. 53-59.
- TSAY, R. S., y TIAO, G. C. (1981): *Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary ARMA models.* Technical Report, 638. University of Wisconsin. Madison.
- URIEL, E. (1983): *Serie enalzadas de la contabilidad nacional de España: 1954-1980.* (Documento de Trabajo). Universidad Literaria de Valencia.
- URIEL, E., y CASADO, A. (1981): *Series temporales,* en *Manual de Investigación Comercial*, coordinada por E. Ortega. Editorial Pirámide.
- VALLS, M., y PRAT, A. (1983): *Identificación automática de series temporales.* Comunicación presentada en el VII Simposium de Teoría Económica y Econometría. Sitges.
- VANDAELE, W. (1983): *Applied Time Series and Box-Jenkins Models.* Academic Press.
- WALKER, G. (1931): *On periodicity in series of related terms.* Proceedings Royal Society London, A, 131, págs. 518-532.
- WINTERS, P. R. (1960): *Forecasting Sales by Exponentially Weighted Moving Averages.* Management Science, 6, págs. 324-342.

REFERENCIAS BIBLIOGRÁFICAS

- WOLD, H. (1938): *A Study in the Analysis of Stationary Time Series.* Segunda edición 1954. Almqvist and Wiksell. Upsala.
- YANGLOM, A. M. (1955): *The correlation theory of processes whose nth difference constitute a stationary process.* Matem. Sb, 37, págs. 141-196.
- YULE, G. U. (1921): *On the time-correlation problems with special reference to the variate difference correlation method.* Journal of Royal Statistical Society, 84, págs. 497-526.
- YULE, G. U. (1926): *Why do we sometimes get nonsense correlations between time series?: a study in sampling and the nature of time series.* Journal or Royal Statistical Society, 80, págs. 1-64.
- YULE, G. U. (1927): *On a method for investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers.* Phil. Trans. Roy. Soc. London A, 226,

Índice alfabético de materias

- AIC (Akaike Information Criterion), 136-137
Alisado exponencial, 17
Amplitud, 188
Análisis Causal, 18-19
Análisis residuos, 127-132
Análisis univariante, 15-19
Autocorrelación, coeficientes de:
 (Véase también Función autocorrelación teórica)
 — definición, 25
 — paseo aleatorio, 60
Autocovarianzas
 — AR(1), 35
 — AR(1), 154-155
 — A(2), 38
 — AR(p), 43
 — ARMA (1, 1), 55-56
 — ARMA (p, q), 54
 — MA(1), 45-46
 — MA(2), 50
 — MA(q), 52
 — paseo aleatorio, 60
 — proceso estacionario, 28
 — proceso no estacionario, 25

Backforecasting, 122
Bondad del Ajuste, 134-137
Box-Cox, transformación de, 62, 93, 96-98
Box-Pierce, estadístico de, 128-129
Brain-storming, 12

Coeficiente de determinación (R^2), 134-136
Coeficiente de determinación corregido (\tilde{R}^2), 134-136
Congruencia, 199-200

Contrastes:
 — coeficientes autocorrelación (individuales), 127-128
 — coeficientes autocorrelación (global), 128-129
 — determinación nivel significación, 129-131
 — estabilidad período muestral, 137
 — estabilidad predicción, 150
 — primera diferencia residuos, 131-132
Conversión de un ARMA(p, q)
 — en un AR(ϕ), 55
 — en un MA(ω), 53-54
Conversión de un ARIMA(p, d, q)
 — en un AR(ϕ), 145
 — en un ARMA, 144
 — en un MA(ω), 144-145
Correlograma, 28
Cross-impact, 13

Datos de corte transversal, 20-23
Decisión, teoría de la, 1-2
Delphi, 12-13
Derivadas, cálculo de:
 — de forma analítica, 116
 — de forma numérica, 116
Descomposición, métodos de, 16-17
Desestacionalización, 16-17
Distribuciones probabilísticas
 — Chi-cuadrado (χ^2), 128, 131, 150
 — log-normal, 96-97
 — normal, 28, 110, 132, 133
 — t de Student, 132-133

Ecuación característica, 182, 183