

Sarriko-On

Análisis de series temporales: Modelos ARIMA

ISBN: 978-84-692-3814-1

María Pilar González Casimiro

04-09



Facultad de Ciencias Económicas y Empresariales
Ekonomia eta Enpresal-Zientzien Fakultatea

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

Análisis de Series Temporales: Modelos ARIMA

Pilar González Casimiro

Departamento de Economía Aplicada III (Econometría y Estadística)

Facultad de Ciencias Económicas y Empresariales

Universidad del País Vasco (UPV-EHU)

Contenido

1. Introducción	1
2. Procesos estocásticos	11
2.1. Proceso estocástico y serie temporal	11
2.2. Características de un proceso estocástico	12
2.3. Procesos estocásticos estacionarios	13
2.3.1. Función de autocovarianzas y de autocorrelación	14
2.3.2. Estimación de los momentos	17
2.4. Proceso Ruido Blanco	18
3. Modelos lineales estacionarios	19
3.1. Modelo Lineal General	19
3.2. Procesos autorregresivos: AR(p).	22
3.2.1. Proceso AR(1).	22
3.2.2. Proceso AR(p).	27
3.3. Procesos de Medias Móviles: MA(q).	31
3.3.1. Proceso MA(1).	31
3.3.2. Proceso MA(q).	34
3.4. Procesos Autorregresivos de Medias Móviles: ARMA(p,q).	38
4. Modelos lineales no estacionarios	43
4.1. No estacionariedad en varianza	43
4.2. No estacionariedad en media.	44
4.3. Modelos ARIMA(p,d,q).	45

4.3.1. Modelo de paseo aleatorio	47
4.3.2. Modelo de paseo aleatorio con deriva	48
5. Modelización ARIMA	49
5.1. Estrategia de modelización ARIMA	49
5.2. Identificación	51
5.2.1. Análisis de estacionariedad	52
5.2.2. Identificación del modelo estacionario	61
5.3. Estimación	69
5.4. Validación del modelo	73
5.4.1. Análisis de coeficientes estimados	73
5.4.2. Análisis de residuos.	77
6. Predicción óptima con modelos ARIMA(p, d, q)	85
6.1. Predicción con modelos estacionarios	86
6.1.1. Predicción con modelos MA(q).	88
6.1.2. Predicción con modelos AR(p)	90
6.1.3. Predicción con modelos ARMA(p,q).	92
6.1.4. Predicciones con modelos estacionarios estimados	94
6.2. Predicción con modelos no estacionarios.	97
7. Modelos ARIMA estacionales	101
7.1. Modelos estacionales puros	102
7.2. Modelos ARMA estacionales no estacionarios	107
7.3. Modelos ARIMA estacionales multiplicativos	109
7.4. Modelización ARIMA estacional.	118
8. Ejercicios	135
8.1. Cuestiones	135
8.2. Problemas	137

Capítulo 1

Introducción

El objetivo de este documento es presentar un conjunto de técnicas para el análisis de series temporales en economía. Por lo tanto, vamos a comenzar definiendo qué se entiende por serie temporal, cuáles son sus características específicas y por qué es necesario desarrollar métodos específicos para su análisis.

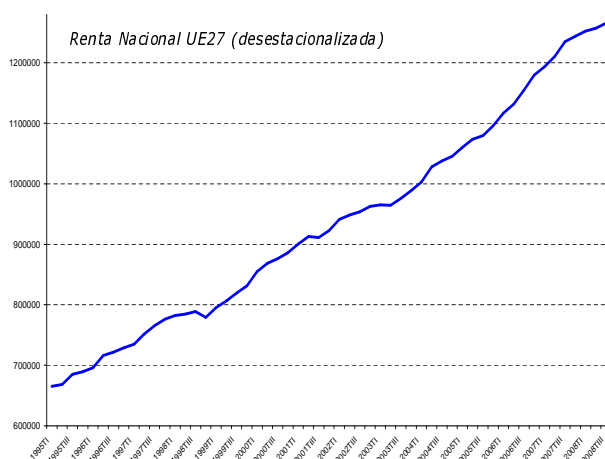
Una *serie temporal* es una secuencia ordenada de observaciones cada una de las cuales está asociada a un momento de tiempo. Ejemplos de series temporales las podemos encontrar en cualquier campo de la ciencia. En Economía cuando buscamos datos para estudiar el comportamiento de una variable económica y su relación con otras a lo largo del tiempo, estos datos se presentan frecuentemente en forma de series temporales. Así, podemos pensar en series como los precios diarios de las acciones, las exportaciones mensuales, el consumo mensual, los beneficios trimestrales, etc. En Meteorología, tenemos series temporales de temperatura, cantidad de lluvia caída en una región, velocidad del viento, etc. En Marketing son de gran interés las series de ventas mensuales o semanales. En Demografía se estudian las series de Población Total, tasas de natalidad, etc. En Medicina, los electrocardiogramas o electroencefalogramas. En Astronomía, la actividad solar, o en Sociología, datos como el número de crímenes, etc.

La mayoría de los métodos estadísticos elementales suponen que las observaciones individuales que forman un conjunto de datos son realizaciones de variables aleatorias mutuamente independientes. En general, este supuesto de independencia mutua se justifica por la atención prestada a diversos aspectos del experimento, incluyendo la extracción aleatoria de la muestra de una población más grande, la asignación aleatoria del tratamiento a cada unidad experimental, etc. Además en este tipo de datos (tomamos una muestra aleatoria simple de una población más grande) el orden de las observaciones no tiene mayor importancia. Sin embargo, en el caso de las series temporales, hemos de tener en cuenta, sin embargo, que:

- el orden es fundamental: tenemos un conjunto de datos ordenado
- el supuesto de independencia no se sostiene ya que, en general, las observaciones son dependientes entre sí y la naturaleza de su dependencia es de interés en sí misma

Consideremos algunos ejemplos de series temporales cuya evolución es bastante habitual cuando se trabaja con variables económicas.

Gráfico 1.1: Renta Nacional UE 27 miembros



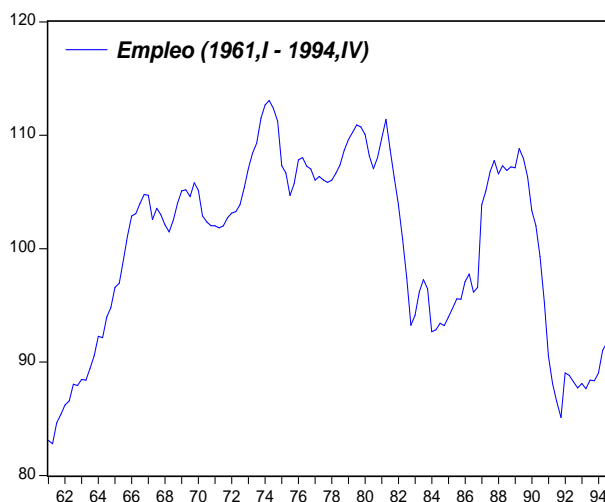
La serie representada en el gráfico 1.1 es la Renta Nacional de la Unión Europea (27 países miembros). Es una serie trimestral que comienza el primer trimestre de 1995 y finaliza el tercer trimestre de 2008. En el gráfico no se presentan los datos brutos de la serie, sino lo que se denomina serie desestacionalizada, es decir, la serie sin comportamiento estacional.

Observando este gráfico se puede concluir que la renta nacional de la UE ha seguido en este periodo una evolución continuamente creciente, es decir, la serie presenta una tendencia creciente. Ahora bien, también hay que señalar que el ritmo de crecimiento de la serie no siempre es el mismo. Por ejemplo, al final de la muestra se puede observar que la serie estaba creciendo a una tasa muy fuerte a partir de finales de 2005 y que este ritmo se suaviza a partir de finales del año 2007. Incluso más, a pesar de que la inspección visual de este gráfico lleva a decir que el comportamiento general, en promedio, de la serie es creciente, se pueden aislar momentos específicos en que la renta nacional decrece, por ejemplo, principios de 1999, el tercer trimestre de 2001 o el tercer trimestre de 2003.

El gráfico 1.2 representa el empleo en el estado español. Es también una serie trimestral que comienza el primer trimestre de 1961 y termina el cuarto trimestre de 1994. En este caso el comportamiento de la serie es muy variable, no se puede decir que tenga un comportamiento a largo plazo ni creciente ni decreciente sino que está continuamente cambiando de nivel. Así, al principio de la muestra, hasta el año 1967, la serie crece. De 1967 hasta principios de los 80, el empleo crece en promedio suavemente, aunque con fuertes oscilaciones. Con la crisis de 1980, se observa una fuerte caída del empleo hasta el año 1985 cuando comienza a recuperarse rápidamente para estabilizarse en los años 1988-90. De nuevo, el gráfico refleja claramente el efecto en el empleo de la crisis de comienzos de los años 90 con un descenso espectacular en solo dos años. Al final de la muestra, la serie comienza a crecer tímidamente.

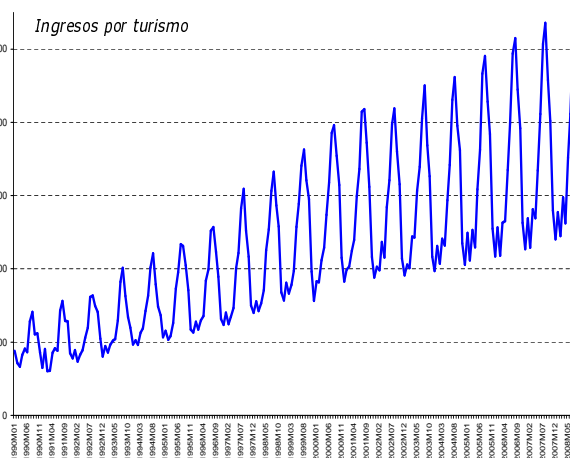
En el gráfico 1.3 podemos encontrar los Ingresos por turismo en el estado español obtenidos de la Balanza de Pagos. Es una serie mensual que comienza el mes de enero de 1990 y termina

Gráfico 1.2: Empleo en el Estado Español



el mes de agosto de 2008. Si tuvieramos que describir brevemente la evolución de esta serie, resaltaríamos, en primer lugar, que a largo plazo crece ya que en media el nivel de ingresos es superior al final de la muestra que al principio de la misma, es decir, la serie tiene una tendencia creciente. Pero esta variable presenta también un tipo de comportamiento cíclico que no observabamos en las dos anteriores. Analizando con más cuidado el gráfico observamos que, dentro de cada año, los picos más altos de los ingresos son siempre en verano (julio y agosto) y los más bajos en invierno. Por lo tanto, si queremos estudiar esta serie y predecirla hemos de tener en cuenta todas sus características, incluyendo los ciclos que se observan con un periodo anual y que se donominan, estacionalidad.

Gráfico 1.3: Ingresos por turismo en el estado español



Debido a estas características específicas de los datos de series temporales se han de desarrollar modelos específicos que recojan y aprovechen la dependencia entre las observaciones ordenadas de una serie temporal.

El conjunto de técnicas de estudio de series de observaciones dependientes ordenadas en el tiempo se denomina Análisis de Series Temporales. El instrumento de análisis que se suele utilizar es un *modelo* que permita reproducir el comportamiento de la variable de interés.

Los Modelos de Series Temporales pueden ser:

- Univariantes: sólo se analiza una serie temporal en función de su propio pasado
- Multivariantes: se analizan varias series temporales a la vez. Un ejemplo muy popular en la literatura son las series de número de pieles de visón y rata almizclera capturadas en Canadá. Se sabe que existe una relación víctima-depredador entre ambos animales lo que se supone que afecta a la dinámica de ambas series. La forma de reflejar estas interacciones dinámicas entre ambas series es construir un modelo multivariante. Cuando se construye un modelo multivariante, para casos como éste, suponemos que hay cierta dependencia o relación entre los pasados de las diversas series.

Estas interacciones dinámicas también aparecen cuando construimos modelos multivariantes para variables económicas, tales como la renta, consumo e inversión que, cómo es bien sabido, influyen las unas en las otras.

En este libro se van a considerar únicamente los modelos de series temporales univariantes, por lo que se va a centrar en el análisis de las series temporales univariantes.

Una serie temporal univariante consiste en un conjunto de observaciones de una variable Y . Si hay T observaciones, se denota por

$$Y_t, \quad t \in \mathfrak{S} \qquad Y_t, \quad t = 1, \dots, T$$

El subíndice t indica el tiempo en que se observa el dato Y_t . Los datos u observaciones se suelen recoger a intervalos iguales de tiempo, es decir, equidistantes los unos de los otros; es el caso de series mensuales, trimestrales, etc.

Cuando las observaciones se recogen solo en momentos determinados de tiempo, generalmente a intervalos iguales, nos referimos a una *serie temporal discreta*. Puede darse el caso de que los datos se generen de forma continua y se observen de forma continua, como, por ejemplo, la temperatura, que se observa de forma continua en el tiempo por medio de aparatos físicos. En este caso denotamos la serie temporal por

$$Y_t, \quad t \in R$$

y contamos con un número infinito de observaciones. En este caso nos referiremos a una *serie temporal continua*. Sin embargo, la mayoría de las series disponibles, en particular, en las ciencias sociales, se observan en tiempo discreto a intervalos iguales, aunque se puedan suponer generadas por algún proceso en tiempo continuo. Por lo tanto, este libro se centrará en el estudio de variables discretas tomadas a intervalos regulares dejando de lado las variables continuas y las variables discretas tomadas a intervalos irregulares.

Cada uno de los datos, Y_t , puede representar o una acumulación sobre un intervalo de tiempo de alguna cantidad subyacente, por ejemplo, lluvia diaria, o bien un valor tomado en un momento dado. Las *variables flujo* son aquellas que se miden respecto a un intervalo de tiempo. Por ejemplo, el consumo mensual de petróleo. Las *variables stock* son aquellas que se miden en un momento determinado de tiempo, por ejemplo, la temperatura, las cotizaciones de bolsa, etc. Los métodos de análisis de series temporales son, en general, los mismos para ambos tipos de variables, pero puede haber ocasiones en que la distinción sea de interés debido a las distintas características que tienen las observaciones.

Los objetivos que se pueden plantear al analizar una serie temporal son muy diversos, pero entre los más importantes se encuentran:

- a) Describir las características de la serie, en términos de sus componentes de interés. Por ejemplo, podemos desear examinar la tendencia para ver cuales han sido los principales movimientos de la serie. Por otro lado, también el comportamiento estacional es de interés, y para algunos propósitos, nos puede interesar extraerlo de la serie: desestacionalizar.

Esta descripción puede consistir en algunos estadísticos resumen (media, varianza, etc.) pero es más probable que incluya una o más representaciones gráficas de los datos. La complejidad de una serie temporal (opuesta a una m.a.s.) es tal que a menudo requiere una función, más que un simple número, para señalar las características fundamentales de las series; por ejemplo, una función μ_t en vez de un número μ para recoger el valor medio de la serie.

- b) Predecir futuros valores de las variables. Un modelo de series temporales univariante se formula únicamente en términos de los valores pasados de Y_t , y/o de su posición con respecto al tiempo (ningún modelo univariante puede ser tomado seriamente como un mecanismo que describe la manera en que la serie es generada: no es un proceso generador de datos). Las predicciones obtenidas a partir de un modelo univariante no son por lo tanto más que extrapolaciones de los datos observados hasta el momento T . En este sentido se dice que son naïve; sin embargo, son en muchas ocasiones muy efectivas y nos proporcionan un punto de referencia con el que comparar el funcionamiento de otros modelos más sofisticados.

Cuando las observaciones sucesivas son dependientes, los valores futuros pueden ser predichos a partir de las observaciones pasadas. Si una serie temporal se puede predecir exactamente, entonces se diría que es una serie determinista. Pero la mayoría de las series son estocásticas en que el futuro sólo se puede determinar parcialmente por sus valores pasados, por lo que las predicciones exactas son imposibles y deben ser reemplazadas por la idea de que los valores futuros tienen una distribución de probabilidad que está condicionada al conocimiento de los valores pasados.

Ambos objetivos pueden conseguirse a muy distintos niveles. Es evidente que, dada la muestra, calcular la media y la desviación típica de las observaciones supone describir características de la serie. De la misma manera podemos predecir que los valores futuros de la variable van a ser iguales al último valor observado. Sin embargo, en ninguno de los dos casos se usa la información muestral de una manera sistemática. El uso sistemático de la información muestral pasa normalmente por la formulación de modelos que pueden describir la evolución de la serie.

Los modelos utilizados para describir el comportamiento de las variables económicas de interés, siempre responden a la misma estructura:

$$Y_t = PS_t + a_t$$

donde: PS_t = Parte sistemática o comportamiento regular de la variable, y a_t es la parte aleatoria, también denominada *innovación*.

En los modelos de series temporales univariantes la PS_t se determina únicamente en función de la información disponible en el pasado de la serie:

$$PS_t = f(Y_t, Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots)$$

El análisis de series temporales se basa en dos nociones fundamentales:

Componentes no observados. Se basa en la idea de que una serie temporal puede ser considerada como la superposición de varios componentes elementales no observables: tendencia, estacionalidad y ciclo. La estimación de tendencias y el ajuste estacional atraen mucho la atención debido a su importancia práctica para el análisis de series económicas.

Modelos ARIMA . Son modelos paramétricos que tratan de obtener la representación de la serie en términos de la interrelación temporal de sus elementos. Este tipo de modelos que caracterizan las series como sumas o diferencias, ponderadas o no, de variables aleatorias o de las series resultantes, fue propuesto por Yule y Slutsky en la década de los 20. Fueron la base de los procesos de medias móviles y autorregresivos que han tenido un desarrollo espectacular tras la publicación en 1970 del libro de Box-Jenkins sobre modelos ARIMA.

El instrumento fundamental a la hora de analizar las propiedades de una serie temporal en términos de la interrelación temporal de sus observaciones es el denominado *coeficiente de autocorrelación* que mide la correlación, es decir, el grado de asociación lineal que existe entre observaciones separadas k periodos. Estos coeficientes de autocorrelación proporcionan mucha información sobre como están relacionadas entre sí las distintas observaciones de una serie temporal, lo que ayudará a construir el modelo apropiado para los datos.

Recordemos que el coeficiente de correlación entre dos variables x_t e y_t , ρ_{xy} , mide el grado de asociación lineal entre ambas variables y se define como:

$$\rho_{xy} = \frac{\text{cov}(x, y)}{\sqrt{V(x)V(y)}}$$

Este coeficiente no tiene unidades por definición y toma valores $-1 \leq \rho_{xy} \leq 1$. Si $\rho_{xy} = 0$, no existe relación lineal entre x e y . Si $\rho_{xy} = 1$, existe relación lineal perfecta y positiva entre x e y . Si $\rho_{xy} = -1$, existe relación lineal perfecta y negativa entre x e y .

Una vez que tenemos una muestra de las variables x e y , se puede estimar este coeficiente

poblacional mediante su correspondiente coeficiente muestral:

$$r_{xy} = \frac{\sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^T (x_t - \bar{x})^2 \sum_{t=1}^T (y_t - \bar{y})^2}}$$

El coeficiente de correlación se puede utilizar como instrumento para analizar la estructura de correlación existente entre las observaciones de una serie temporal Y_t .

Si tenemos T observaciones Y_1, \dots, Y_T , podemos crear $(T-1)$ pares de observaciones $(Y_1, Y_2), \dots, (Y_{T-1}, Y_T)$. Si consideramos Y_1, \dots, Y_{T-1} como una variable y Y_2, \dots, Y_T como otra, podemos definir la correlación entre ambas variables Y_t, Y_{t+1} :

$$r_{Y_t Y_{t+1}} = r_1 = \frac{\sum_{t=1}^{T-1} (Y_t - \bar{Y}_{(1)})(Y_{t+1} - \bar{Y}_{(2)})}{\sqrt{\sum_{t=1}^{T-1} (Y_t - \bar{Y}_{(1)})^2 \sum_{t=1}^{T-1} (Y_{t+1} - \bar{Y}_{(2)})^2}}$$

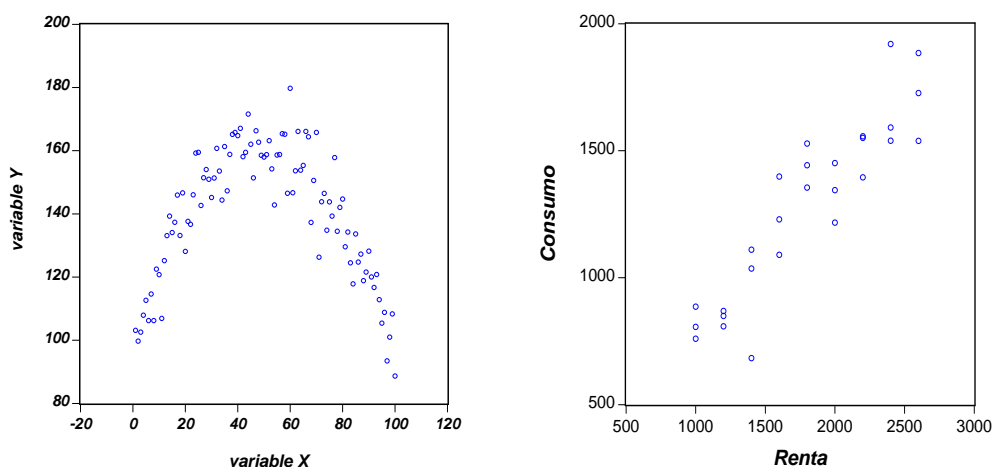
donde $\bar{Y}_{(1)}$ es la media muestral de las $T-1$ primeras observaciones e $\bar{Y}_{(2)}$ es la media muestral de las $T-1$ últimas observaciones. Esta fórmula se suele aproximar por:

$$r_1 = \frac{\sum_{t=1}^{T-1} (Y_t - \bar{Y})(Y_{t+1} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

El coeficiente r_1 mide la correlación, grado de asociación lineal, entre observaciones sucesivas; se le denomina coeficiente de correlación simple o coeficiente de correlación serial, o coeficiente de autocorrelación de primer orden. En general, se puede estimar la correlación entre observaciones separadas por k intervalos, es decir, el coeficiente de autocorrelación de orden k , como sigue:

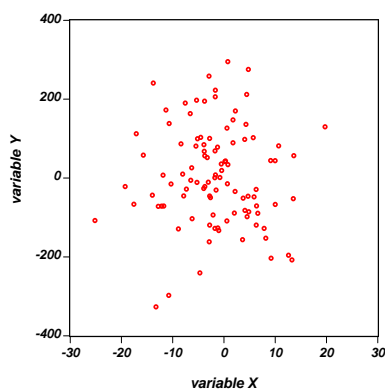
$$r_k = \frac{\sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

Ejercicio 1.1. ¿Qué valores esperas para el coeficiente de correlación muestral en estos casos?



El correlograma es el gráfico de los coeficientes de autocorrelación de orden k contra el retardo k . Este gráfico es muy útil para interpretar el conjunto de los coeficientes de autocorrelación de una serie temporal. La interpretación del correlograma no es fácil pero vamos a dar unas pistas al respecto.

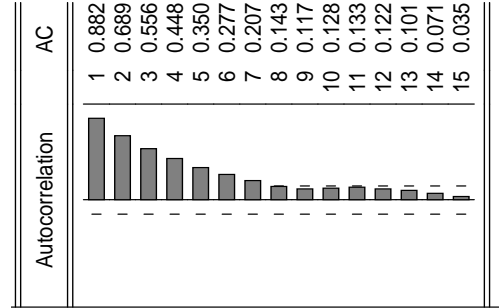
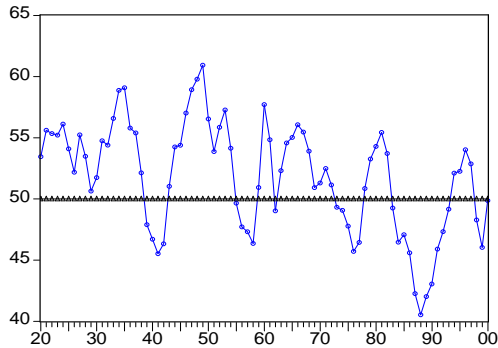
(A) Si una serie es puramente aleatoria, entonces para valores de T grandes, $r_k \simeq 0$, para cualquier $k \neq 0$.



(B.1) Correlación a corto plazo.

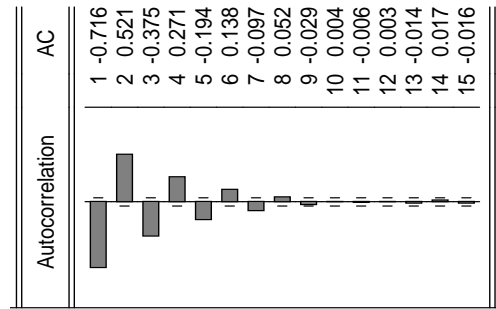
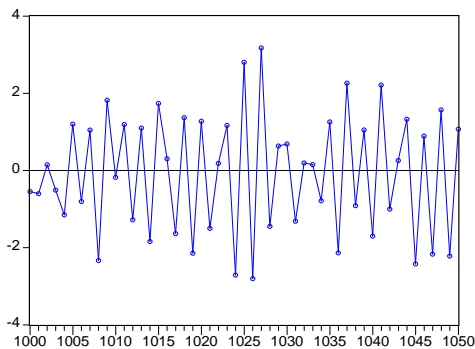
Sea una serie sin tendencia, que oscila en torno a una media constante, pero cuyas observaciones sucesivas están correlacionadas positivamente, es decir, una serie en la que a una observación por encima de la media, le suele seguir otra o más observaciones por encima de la media (lo mismo, para observaciones por debajo de la media).

El correlograma es de la siguiente forma: un valor relativamente alto de r_1 , seguido de algunos coeficientes r_k distintos de cero, pero cada vez más pequeños y valores de r_k aproximadamente cero para k grande.



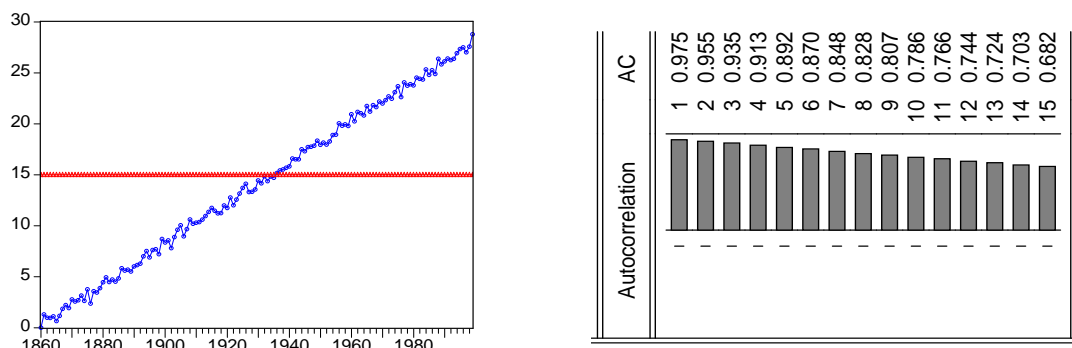
(B.2) Correlación a corto plazo.

Las series sin tendencia que oscilan en torno a una media constante pero que alternan valores con observaciones sucesivas a diferentes lados de la media general, presentan un correlograma que también suele alternar los signos de sus coeficientes.

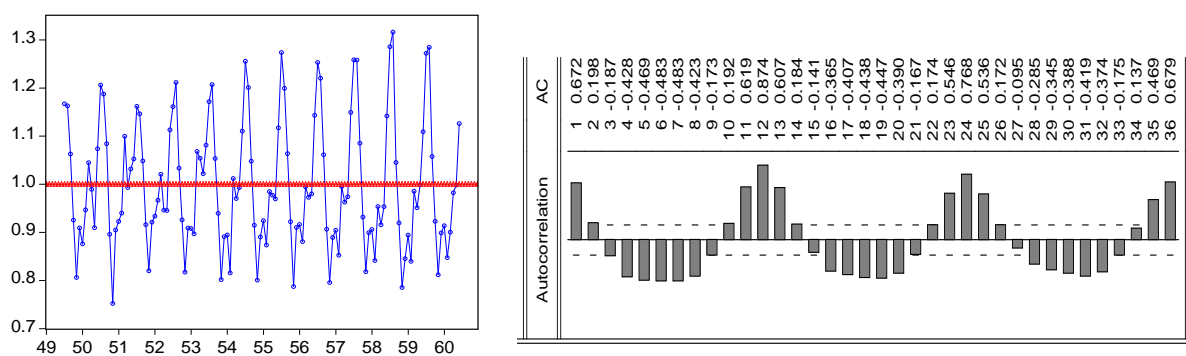


Por ejemplo, si el valor de r_1 es negativo, el valor de r_2 puede ser, sin embargo, positivo porque las observaciones separadas dos periodos tienden a estar al mismo lado de la media.

(C) Si la serie contiene una tendencia, es decir, cambia continuamente de nivel, los valores de r_k no van a decrecer hacia cero rápidamente. Esto es debido a que una observación por encima de la media (o por debajo) de la media general es seguida de muchas observaciones por el mismo lado de la media debido a la tendencia. A lo largo del curso, se estudiará como para que el correlograma sea informativo habremos de eliminar esta tendencia.



(D) Si una serie presenta algún tipo de ciclo, el correlograma también presentará una oscilación a la misma frecuencia. Por ejemplo, para series mensuales, r_6 será grande y negativo, y r_{12} será grande y positivo. En general, el correlograma va a reproducir el comportamiento cíclico presente en la serie. En las series con un comportamiento cíclico permanente, el correlograma da de nuevo poca información porque lo domina el comportamiento cíclico presente en los datos. Si eliminamos la variación cíclica de los datos el correlograma podría ser más útil.



Los modelos de series temporales ARIMA van a ser el objetivo central del libro. Estos modelos se basan en la teoría de los procesos estocásticos que se desarrolla en el capítulo 2. Los modelos ARMA estacionarios se explican con detalle en el capítulo 3 y los modelos ARIMA no estacionarios en el capítulo 4. La metodología de modelización ARIMA con las fases de identificación, estimación y contraste de diagnósticos se explica detalladamente en el capítulo 5 y el capítulo 6 se dedica a la predicción. El capítulo 7 trata sobre la especificación de los modelos de series temporales adecuados a los datos estacionales. Por último, el capítulo 8 reúne una colección de ejercicios para el trabajo propio del lector.

Capítulo 2

Procesos estocásticos

Una serie temporal univariante consiste en un conjunto de observaciones de una variable de interés. Un modelo de series temporales debe reproducir las características de la serie. Si contamos con T observaciones, $Y_t, t = 1, 2, \dots, T$, el modelo univariante de series temporales se formulará en términos de los valores pasados de Y_t y/o su posición en relación con el tiempo. Ahora bien, no existe un único modelo para conseguir este modelo. En este capítulo se va a estudiar la teoría de procesos estocásticos para describir la estructura probabilística de una secuencia de observaciones en el tiempo y para predecir.

2.1. Proceso estocástico y serie temporal

Un *proceso estocástico* es una familia de variables aleatorias que, en general, están relacionadas entre sí y siguen una ley de distribución conjunta. Se denota por:

$$Y_t(\omega), t = \dots, t-2, t-1, t, t+1, t+2, \dots,$$
$$\text{ó } \dots, Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}, \dots, \quad \text{ó } \{Y_t\}_{-\infty}^{+\infty} \quad \text{ó simplemente } Y_t$$

Dado que el objetivo del libro es el análisis de series temporales, consideraremos únicamente secuencias de variables aleatorias ordenadas en el tiempo, pero no tiene por qué ser así. Se podría pensar, por ejemplo, en una ordenación de tipo espacial.

Si se fija el tiempo, $t = t_0$, se tiene una variable aleatoria determinada de la secuencia, $Y_{t_0}(\omega)$, con su distribución de probabilidad univariante.

Si se fija la variable aleatoria $\omega = \omega_0$, se tiene una *realización* del proceso, es decir, una muestra de tamaño 1 del proceso formada por una muestra de tamaño 1 de cada una de las variables aleatorias que forman el proceso:

$$\dots, Y_{t-2}(\omega_0), Y_{t-1}(\omega_0), Y_t(\omega_0), Y_{t+1}(\omega_0), Y_{t+2}(\omega_0), \dots$$
$$\dots, Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}, \dots$$

En el marco estadístico de los procesos estocásticos, una serie temporal, Y_1, Y_2, \dots, Y_T , se puede interpretar como una realización muestral de un proceso estocástico que se observa únicamente para un número finito de periodos, $t = 1, 2, \dots, T$.

En este contexto, por lo tanto, una serie temporal es una sucesión de observaciones en la que cada una de ellas corresponde a una variable aleatoria distinta, y la ordenación de la sucesión de observaciones es esencial para el análisis de la misma, no se puede alterar, porque se cambiarían las características de la serie que se quiere estudiar. Sin embargo, en la teoría del muestreo aleatorio, una muestra es un conjunto de observaciones de una única variable aleatoria y, por lo tanto, su orden es irrelevante. Así, si se considera el proceso estocástico *Consumo Privado*, denotado por $\{C_t\}_{-\infty}^{+\infty}$, la serie temporal dada por los valores del consumo anuales publicados por la agencia estadística correspondiente:

$$C_{1996} \quad C_{1997} \quad C_{1998} \quad C_{1999} \quad C_{2000} \quad C_{2001} \quad C_{2002} \quad C_{2003} \quad C_{2004} \quad C_{2005} \quad C_{2006} \quad C_{2007}$$

es una realización del proceso estocástico, en principio infinito, *Consumo Privado* que se observa únicamente de 1996 a 2007. Sin embargo, si se considera el consumo semanal de una determinada clase de familias de una ciudad en un momento dado, la muestra tomada para N familias :

$$C_1 \quad C_2 \quad C_3 \quad C_4 \quad \dots \quad C_N$$

es un conjunto de N observaciones de una variable aleatoria consumo familiar, $C \sim (\mu, \sigma^2)$.

2.2. Características de un proceso estocástico

Un proceso estocástico se puede caracterizar bien por su función de distribución o por sus momentos.

Función de distribución. Para conocer la función de distribución de un proceso estocástico es necesario conocer las funciones de distribución univariantes de cada una de las variables aleatorias del proceso, $F[Y_{t_i}], \forall t_i$, y las funciones bivariantes correspondientes a todo par de variables aleatorias del proceso, $F[Y_{t_i}, Y_{t_j}], \forall (t_i, t_j)$, y todas las funciones trivariantes, ...

En resumen, la función de distribución de un proceso estocástico incluye todas las funciones de distribución para cualquier subconjunto finito de variables aleatorias del proceso:

$$F[Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}], \quad \forall (t_1, t_2, \dots, t_n), \quad \text{siendo } n \text{ finito}$$

Momentos del proceso estocástico. Como suele ser muy complejo determinar las características de un proceso estocástico a través de su función de distribución se suele recurrir a caracterizarlo a través de los dos primeros momentos.

El primer momento de un proceso estocástico viene dado por el conjunto de las *medias* de todas las variables aleatorias del proceso:

$$E(Y_t) = \mu_t < \infty, \quad t = 0, \pm 1, \pm 2, \dots,$$

El segundo momento centrado del proceso viene dado por el conjunto de las *varianzas* de todas las variables aleatorias del proceso y por las *covarianzas* entre todo par de variables aleatorias:

$$\begin{aligned} V(Y_t) &= E[Y_t - \mu_t]^2 = \sigma_t^2, < \infty \quad t = 0, \pm 1, \pm 2, \dots, \\ \text{cov}(Y_t, Y_s) &= E[Y_t - \mu_t][Y_s - \mu_s] = \gamma_{t,s}, \quad \forall t, s \quad (t \neq s) \end{aligned}$$

Si la distribución del proceso es normal y se conocen sus dos primeros momentos (medias, varianzas y covarianzas), el proceso está perfectamente caracterizado y se conoce su función de distribución.

2.3. Procesos estocásticos estacionarios

En el análisis de series temporales el objetivo es utilizar la teoría de procesos estocásticos para determinar qué proceso estocástico ha sido capaz de generar la serie temporal bajo estudio con el fin de caracterizar el comportamiento de la serie y predecir en el futuro. Si se quieren conseguir métodos de predicción consistentes, no se puede utilizar cualquier tipo de proceso estocástico, sino que es necesario que la estructura probabilística del mismo sea estable en el tiempo.

La filosofía que subyace en la teoría de la predicción es siempre la misma: se aprende de las regularidades del comportamiento pasado de la serie y se proyectan hacia el futuro. Por lo tanto, es preciso que los procesos estocásticos generadores de las series temporales tengan algún tipo de estabilidad. Si, por el contrario, en cada momento de tiempo presentan un comportamiento diferente e inestable, no se pueden utilizar para predecir. A estas condiciones que se les impone a los procesos estocásticos para que sean estables para predecir, se les conoce como *estacionariedad*.

El concepto de estacionariedad se puede caracterizar bien en términos de la función de distribución o de los momentos del proceso. En el primer caso, se hablará de estacionariedad en sentido estricto y, en el segundo, de estacionariedad de segundo orden o en covarianza.

Estacionariedad estricta. Un proceso estocástico, Y_t , es estacionario en sentido estricto si y solo si:

$$F[Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}] = F[Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k}] \quad \forall (t_1, t_2, \dots, t_n) \text{ y } k$$

es decir, si la función de distribución de cualquier conjunto finito de n variables aleatorias del proceso no se altera si se desplaza k periodos en el tiempo.

Estacionariedad en covarianza. Un proceso estocástico, Y_t , es estacionario en covarianza si y solo si:

- a) Es estacionario en media, es decir, todas las variables aleatorias del proceso tienen la misma media y es finita:

$$E(Y_t) = \mu < \infty, \quad \forall t$$

- b) Todas las variables aleatorias tienen la misma varianza y es finita, es decir, la dispersión en torno a la media constante a lo largo del tiempo es la misma para todas las variables del proceso

$$V(Y_t) = E[Y_t - \mu]^2 = \sigma_Y^2 < \infty, \quad \forall t$$

- c) Las autocovarianzas solo dependen del número de periodos de separación entre las variables y no del tiempo, es decir, la covarianza lineal entre dos variables aleatorias del proceso que

disten k periodos de tiempo es la misma que existe entre cualesquiera otras dos variables que estén separadas también k periodos, independientemente del momento concreto de tiempo al que estén referidas

$$\text{cov}(Y_t, Y_s) = E[Y_t - \mu][Y_s - \mu] = \gamma_{|t-s|} = \gamma_k < \infty, \quad \forall k$$

Por lo tanto, un proceso estocástico es estacionario en covarianza si y solo si:

$$(1) \quad E(Y_t) = \mu < \infty$$

$$(2) \quad \text{Cov}(Y_t, Y_s) = \begin{cases} V(Y_t) = \sigma_Y^2 < \infty & t = s \\ \gamma_{|t-s|} = \gamma_k < \infty & t \neq s \end{cases}$$

Un proceso estocástico estacionario en covarianza está caracterizado si se conoce:

$$\mu, \quad V(Y_t) \text{ y } \gamma_k, \quad k = \pm 1, \pm 2, \dots,$$

Si un proceso estocástico es estacionario en covarianza y su distribución es Normal, es estacionario en sentido estricto.

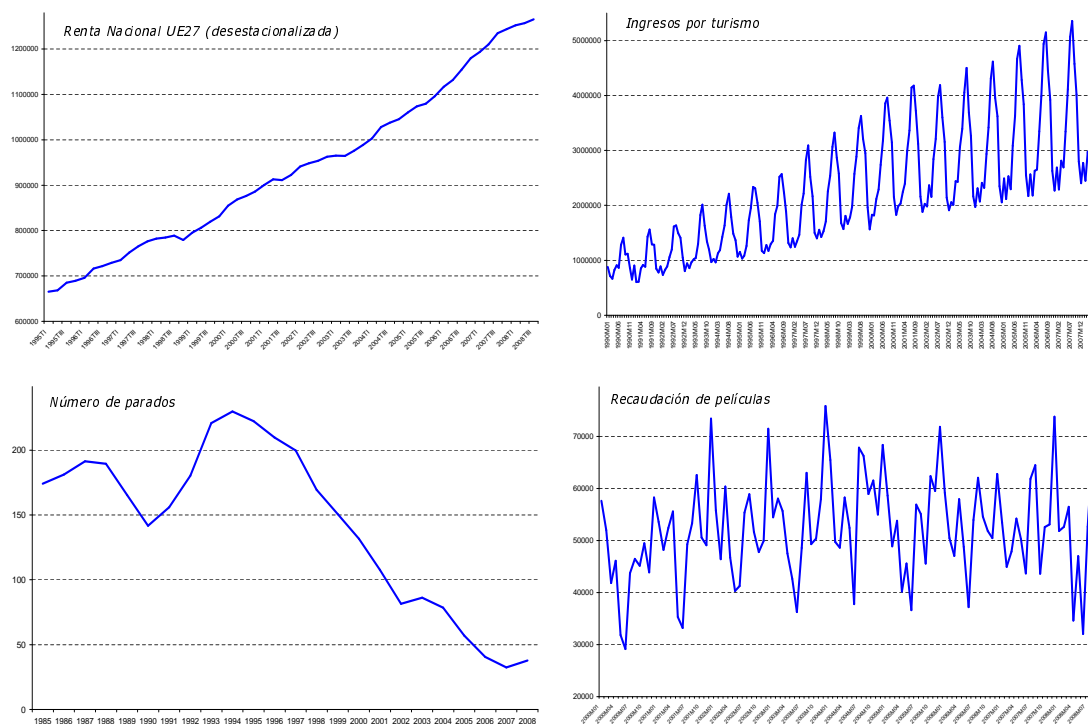
La pregunta relevante ahora es si las series económicas presentan, en general, un comportamiento estacionario. Una simple inspección visual de muchas series temporales económicas, permite comprobar que presentan cambios de nivel, tendencias crecientes, estacionalidades. En las figuras del gráfico 2.1 se observa, por ejemplo, como la serie de *Renta Nacional* crece sistemáticamente o como la serie de *Número de parados* cambia constantemente de nivel, creciendo y decreciendo. En ambos casos no se puede sostener el supuesto de que la media de todas las variables de los procesos que han generado estas series sea la misma, por lo que no son estacionarias en media. La serie de *Ingresos por turismo* tampoco refleja un comportamiento estacionario en media no sólo por su crecimiento a largo plazo sino por la presencia de la estacionalidad que hace que el comportamiento promedio de cada mes sea diferente. La serie de *Recaudación de películas* parece oscilar en torno a una media constante y, a pesar de ser mensual, no tiene comportamiento estacional. Sin embargo, tampoco presenta un comportamiento estacionario dado que se puede observar que la variabilidad de la serie no es la misma para todo el periodo muestral, luego no sería estacionaria en varianza.

Se puede concluir, por lo tanto, que la mayoría de las series económicas no van a presentar comportamientos estacionarios. De todas formas, se procederá a proponer modelos para los procesos estocásticos estacionarios, en primer lugar, para luego pasar a analizar cómo se puede generalizar este tipo de modelos para incluir comportamientos no estacionarios del tipo observado en las series económicas.

2.3.1. Función de autocovarianzas y de autocorrelación

En principio, si se considera el proceso estocástico teórico, Y_t , que comienza en algún momento del pasado lejano y acaba en un futuro indeterminado se pueden calcular un número indefinido de autocovarianzas, por lo que conviene definir una función que las agrupe a todas.

Gráfico 2.1: Series temporales económicas



Función de autocovarianzas (FACV)

La función de autocovarianzas de un proceso estocástico estacionario es una función de k (número de periodos de separación entre las variables) que recoge el conjunto de las autocovarianzas del proceso y se denota por:

$$\gamma_k, \quad k = 0, 1, 2, 3, \dots,$$

Características de la función de autocovarianzas:

- Incluye la varianza del proceso para $k = 0$: $\gamma_0 = E[Y_t - \mu][Y_t - \mu] = V(Y_t)$
- Es una función simétrica: $\gamma_k = E[Y_t - \mu][Y_{t+k} - \mu] = E[Y_t - \mu][Y_{t-k} - \mu] = \gamma_{-k}$

La función de autocovarianzas de un proceso estocástico recoge toda la información sobre la estructura dinámica lineal del mismo. Pero depende de las unidades de medida de la variable, por lo que, en general, se suele utilizar la función de autocorrelación.

Función de autocorrelación (FAC)

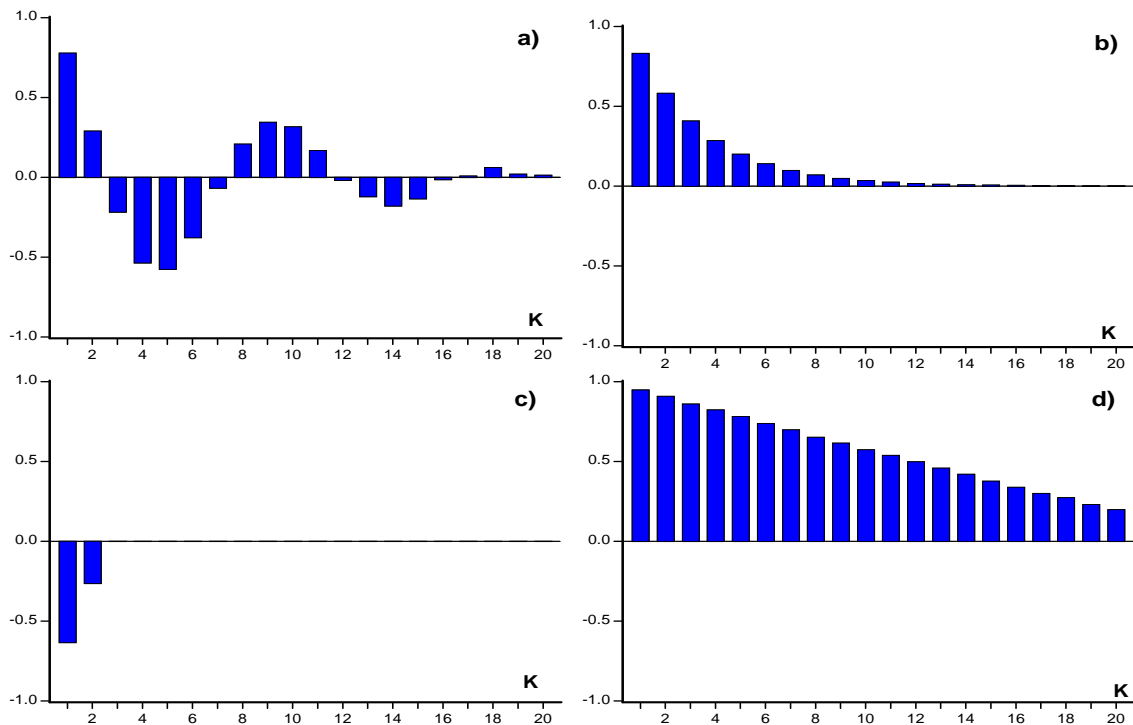
El coeficiente de autocorrelación de orden k de un proceso estocástico estacionario mide el grado de asociación lineal existente entre dos variables aleatorias del proceso separadas k periodos:

$$\rho_k = \frac{\text{cov}(Y_t, Y_{t+k})}{\sqrt{V(Y_t) V(Y_{t+k})}} = \frac{\gamma_k}{\sqrt{\gamma_0 \gamma_0}} = \frac{\gamma_k}{\gamma_0}$$

Por ser un coeficiente de correlación, no depende de unidades y $|\rho_k| \leq 1, \forall k$.

La función de autocorrelación de un proceso estocástico estacionario es una función de k que recoge el conjunto de los coeficientes de autocorrelación del proceso y se denota por $\rho_k, k = 0, 1, 2, 3, \dots$. La función de autocorrelación se suele representar gráficamente por medio de un gráfico de barras denominado *correlograma*.

Gráfico 2.2: Funciones de autocorrelación



Las características de la función de autocorrelación de un proceso estocástico estacionario:

- El coeficiente de autocorrelación de orden 0 es, por definición, 1. Por eso, a menudo, no se le incluye explícitamente en la función de autocorrelación.

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$$

- Es una función simétrica: $\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_{-k}}{\gamma_0} = \rho_{-k}$. Por ello, en el *correlograma* se representa la función de autocorrelación solamente para los valores positivos del retardo k .
- La función de autocorrelación de un proceso estocástico estacionario tiende a cero rápidamente cuando k tiende a ∞ .

La función de autocorrelación va a ser el principal instrumento utilizado para recoger la estructura dinámica lineal del modelo. Los gráficos de la figura 2.2 muestran 4 correlogramas correspondientes a diferentes series temporales. Los correlogramas a), b) y c) decrecen rápidamente hacia cero conforme aumenta k : exponencialmente en los casos a) y b) y truncándose en el caso c). Son, por lo tanto, correlogramas correspondientes a series estacionarias. Por el contrario, los coeficientes de autocorrelación del correlograma d) decrecen lentamente, de forma lineal, por lo que no corresponden a una serie estacionaria.

Para caracterizar un proceso estocástico estacionario por sus momentos existen dos opciones:

- Media y función de autocovarianzas: $\mu \quad \gamma_k, \quad k = 0, 1, 2, 3, \dots$
- Media, varianza y función de autocorrelación: $\mu \quad \gamma_0 \quad \rho_k, \quad k = 1, 2, 3, \dots$

2.3.2. Estimación de los momentos

Los momentos poblacionales de un proceso estocástico estacionario se estiman a través de los correspondientes momentos muestrales.

$$\text{a) Media: } \hat{\mu} = \bar{Y} = \frac{1}{T} \sum_{t=1}^T Y_t$$

$$\text{b) Varianza: } \hat{\gamma}_0 = \hat{V}(Y) = \frac{1}{T} \sum_{t=1}^T (Y_t - \bar{Y})^2$$

$$\text{c) Función de Autocovarianzas: } \hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}), \quad k = 1, 2, \dots, K$$

$$\text{d) Función de Autocorrelación: } \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, \quad k = 1, 2, \dots, K$$

Aunque para el proceso estocástico teórico se cuenta con un número indefinido de autocovarianzas y coeficientes de autocorrelación, cuando se dispone de una serie temporal finita de tamaño T , como máximo se pueden estimar $T - 1$ coeficientes de autocorrelación, pero, en la práctica, se van a estimar muchos menos. Se recomienda un máximo de $T/3$. Esto es debido a que cuanto mayor sea k menos información hay para estimar ρ_k y la calidad de la estimación es menor.

La estimación de los momentos poblacionales de un proceso estocástico proporciona otra razón por la que es necesario imponer la restricción de estacionariedad de un proceso. Si el proceso que ha generado la serie (Y_1, Y_2, \dots, Y_T) no fuera estacionario, sería preciso estimar T medias

diferentes, μ_t , T varianzas diferentes, σ_t^2 , y un gran número de autocorrelaciones, lo que no es posible contando con T observaciones solamente. La solución habitual a este tipo de problema es buscar más datos pero, en el caso de las series temporales, esta solución no es válida. Si conseguimos una serie temporal más larga, por ejemplo, M observaciones más, tendríamos que estimar M medias más, M varianzas más, etc. con lo cual estaríamos en la misma situación.

2.4. Proceso Ruido Blanco

El proceso estocástico más sencillo es el denominado *Ruido Blanco* que es una secuencia de variables aleatorias de media cero, varianza constante y covarianzas nulas. Se denotará habitualmente por a_t , $t = 0, \pm 1, \pm 2, \dots$:

$$E(a_t) = 0, \forall t \quad V(a_t) = \sigma^2, \forall t \quad Cov(a_t a_s) = 0, \forall t \neq s$$

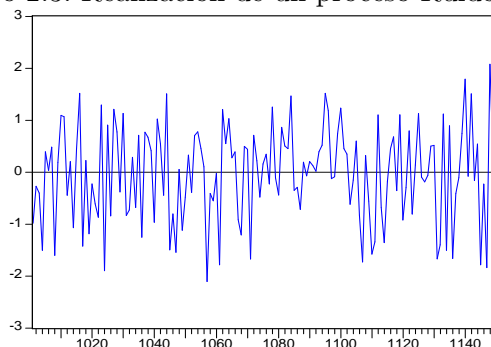
Así, un proceso ruido blanco, $a_t \sim RB(0, \sigma^2)$, es estacionario si la varianza σ^2 es finita con función de autocovarianzas (FACV):

$$\gamma_k = \sigma^2, k = 0 \quad y \quad \gamma_k = 0, k > 0$$

y función de autocorrelación (FAC):

$$\gamma_k = 1, k = 0 \quad y \quad \gamma_k = 0, k > 0$$

Gráfico 2.3: Realización de un proceso Ruido Blanco



El gráfico 2.3 muestra una realización de tamaño 150 de un proceso ruido blanco normal con varianza unidad, $a_t \sim RBN(0, 1)$. Se observa que la serie simulada oscila en torno a un nivel constante nulo con una dispersión constante y de forma aleatoria, sin ningún patrón de comportamiento, como corresponde a la ausencia de correlación en el tiempo entre sus observaciones. Dado que la característica fundamental de las series temporales es la dependencia temporal entre sus observaciones, por lo que no es de esperar que muchas series económicas presenten un comportamiento que responda a una ausencia de correlación temporal en el sentido de que lo que ocurre hoy no tiene relación lineal con lo sucedido en el pasado. De todas formas, el proceso ruido blanco es muy útil en el análisis de series temporales porque es la base para la construcción de los modelos $ARIMA(p, d, q)$.

Capítulo 3

Modelos lineales estacionarios

3.1. Modelo Lineal General

La metodología de la modelización univariante es sencilla. Dado que el objetivo es explicar el valor que toma en el momento t una variable económica que presenta dependencia temporal, una forma de trabajar es recoger información sobre el pasado de la variable, observar su evolución en el tiempo y explotar el patrón de regularidad que muestran los datos. La estructura de dependencia temporal de un proceso estocástico está recogida en la función de autocovarianzas (FACV) y/o en la función de autocorrelación (FAC). En este contexto, se trata de utilizar la información de estas funciones para extraer un patrón sistemático y, a partir de éste, un modelo que reproduzca el comportamiento de la serie y se pueda utilizar para predecir. Este procedimiento se hará operativo mediante los modelos ARMA que son una aproximación a la estructura teórica general.

En un modelo de series temporales univariante se descompone la serie Y_t en dos partes, una que recoge el patrón de regularidad, o parte sistemática, y otra parte puramente aleatoria, denominada también *innovación*:

$$Y_t = PS_t + a_t \quad t = 1, 2, \dots$$

La *parte sistemática* es la parte predecible con el conjunto de información que se utiliza para construir el modelo, es decir, la serie temporal Y_1, Y_2, \dots, Y_T . La *innovación* respecto al conjunto de información con el que se construye el modelo, es una parte aleatoria en la que sus valores no tienen ninguna relación o dependencia entre sí. La innovación en el momento t no está relacionada, por lo tanto, ni con las innovaciones anteriores ni con las posteriores, ni con la parte sistemática del modelo. Es impredecible, es decir, su predicción es siempre cero. A la hora de construir un modelo estadístico para una variable económica, el problema es formular la parte sistemática de tal manera que el elemento residual sea una innovación, en el mundo Normal, un ruido blanco.

Dada una serie temporal de media cero, como el valor de Y en el momento t depende de su pasado, un modelo teórico capaz de describir su comportamiento sería:

$$Y_t = f(Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots) + a_t \quad t = 1, 2, \dots$$

donde se exige que el comportamiento de Y_t sea función de sus valores pasados, posiblemente infinitos.

Dentro de los procesos estocásticos estacionarios se considerará únicamente la clase de procesos lineales que se caracterizan porque se pueden representar como una combinación lineal de variables aleatorias. De hecho, en el caso de los procesos estacionarios con distribución normal y media cero, la teoría de procesos estocásticos señala que, bajo condiciones muy generales, Y_t se puede expresar como combinación lineal de los valores pasados infinitos de Y más una innovación ruido blanco:

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \pi_3 Y_{t-3} + \dots + a_t \quad \forall t \quad t = 1, 2, \dots \quad (3.1)$$

Las condiciones generales que ha de cumplir el proceso son:

- a) Que el proceso sea *no anticipante*, es decir, que el presente no venga determinado por el futuro, luego el valor de Y en el momento t no puede depender de valores futuros de Y o de las innovaciones a .
- b) Que el proceso sea *invertible*, es decir, que el presente dependa de forma convergente de su propio pasado lo que implica que la influencia de Y_{t-k} en Y_t ha de ir disminuyendo conforme nos alejamos en el pasado. Esta condición se cumple si los parámetros del modelo general (3.1) cumplen la siguiente restricción:

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty$$

El modelo (3.1) se puede escribir de forma más compacta en términos del operador de retardos:

$$\begin{aligned} Y_t &= (\pi_1 L + \pi_2 L^2 + \dots) Y_t + a_t \\ \longrightarrow (1 - \pi_1 L - \pi_2 L^2 - \dots) Y_t &= a_t \quad \longrightarrow \quad \Pi_{\infty}(L) Y_t = a_t \end{aligned}$$

Otra forma alternativa de escribir el modelo (3.1) es:

$$\begin{aligned} Y_t &= \frac{1}{\Pi_{\infty}(L)} a_t = \Psi_{\infty}(L) a_t = (1 + \psi_1 L + \psi_2 L^2 + \dots) a_t \\ Y_t &= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \dots \quad t = 1, 2, \dots \quad (3.2) \end{aligned}$$

Es decir, el valor Y_t se puede representar como la combinación lineal del ruido blanco a_t y su pasado infinito. El modelo (3.2) cumple la condición de estacionariedad si los parámetros del modelo satisfacen la siguiente restricción:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

Esta restricción implica que el valor presente depende de forma convergente de las innovaciones pasadas, es decir, la influencia de la innovación a_{t-k} va desapareciendo conforme se aleja en el pasado.

Tanto la representación (3.1) como la (3.2) son igualmente válidas para los procesos estocásticos estacionarios siempre que se cumplan las dos condiciones antes mencionadas, o sea que el modelo sea no anticipante e invertible.

Como en la práctica se va a trabajar con series finitas, los modelos no van a poder expresar dependencias infinitas sin restricciones sino que tendrán que especificar una dependencia en el tiempo acotada y con restricciones. Será necesario simplificar las formulaciones del modelo general (3.1) y/o (3.2) de forma que tengan un número finito de parámetros. Acudiendo a la teoría de polinomios, bajo condiciones muy generales, se puede aproximar un polinomio de orden infinito mediante un cociente de polinomios finitos:

$$\Pi_{\infty}(L) \simeq \frac{\phi_p(L)}{\theta_q(L)}$$

donde $\phi_p(L)$ y $\theta_q(L)$ son polinomios en el operador de retardos finitos de orden p y q , respectivamente:

$$\begin{aligned}\phi_p(L) &= 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \\ \theta_q(L) &= 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q\end{aligned}$$

Sustituyendo en el modelo (3.1), se obtiene:

$$\Pi_{\infty}(L) Y_t \simeq \frac{\phi_p(L)}{\theta_q(L)} Y_t = a_t \quad \rightarrow \quad \phi_p(L) Y_t = \theta_q(L) a_t$$

Por lo tanto, el modelo lineal general admite tres representaciones, todas igualmente válidas bajo los supuestos señalados:

- Representación *puramente autorregresiva* (3.1), $AR(\infty)$: el valor presente de la variable se representa en función de su propio pasado más una innovación contemporánea.
- Representación *puramente de medias móviles* (3.2), $MA(\infty)$: el valor presente de la variable se representa en función de todas las innovaciones presente y pasadas.
- Representación *finita*:

$$\begin{aligned}\phi_p(L) Y_t &= \theta_q(L) a_t \\ (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) Y_t &= (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) a_t \\ Y_t &= \underbrace{\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p}}_{\text{parte autorregresiva}} + a_t \underbrace{-\theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}}_{\text{parte medias móviles}}\end{aligned}$$

En este modelo finito, el valor de Y_t depende del pasado de Y hasta el momento $t - p$ (parte autorregresiva), de la innovación contemporánea y su pasado hasta el momento

$t - q$ (parte medias móviles). Este modelo se denomina *Autorregresivo de Medias Móviles* de orden (p, q) , y se denota por $ARMA(p, q)$.

Esta formulación finita es una representación más restringida que las representaciones generales (3.1)-(3.2).

Dos casos particulares del modelo $ARMA(p, q)$ de gran interés son:

- $AR(p)$. Modelo que sólo presenta parte autorregresiva, es decir, el polinomio de medias móviles es de orden 0:

$$Y_t \sim AR(p) \quad \rightarrow \quad Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t$$

- $MA(q)$. Modelo que sólo presenta parte medias móviles, es decir, el polinomio autorregresivo es de orden 0:

$$Y_t \sim MA(q) \quad \rightarrow \quad Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Cuando el modelo es conocido se puede utilizar cualquiera de las tres representaciones dependiendo de los intereses. Si el modelo no es conocido y hay que especificarlo y estimarlo a partir de una serie temporal concreta, hay que utilizar necesariamente la formulación finita.

Cuando se construye un modelo de series temporales univariante el objetivo no es conseguir el “verdadero” modelo. Es preciso ser conscientes de que estamos tratando de modelar una realidad compleja y el objetivo es lograr un modelo parsimonioso y suficientemente preciso que represente adecuadamente las características de la serie recogidas fundamentalmente en la función de autocorrelación. Los modelos $ARMA(p, q)$, $AR(p)$ y $MA(q)$ son aproximaciones al modelo lineal general. Comenzaremos por caracterizar sus funciones de autocorrelación para conocer sus propiedades y posteriormente utilizarlos para modelar series y predecir.

Ejercicio 3.1. Deriva el modelo finito a partir del modelo (3.2).

3.2. Procesos autorregresivos: AR(p).

El modelo autorregresivo finito de orden p , $AR(p)$ es una aproximación natural al modelo lineal general (3.1). Se obtiene un modelo finito simplemente truncando el modelo general:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad t = 1, 2, \dots$$

Para estudiar sus características, comenzaremos por el modelo autorregresivo de orden 1, $AR(1)$.

3.2.1. Proceso AR(1).

En el proceso $AR(1)$ la variable Y_t viene determinada únicamente por su valor pasado, Y_{t-1} , y la perturbación contemporánea, a_t :

$$Y_t = \phi Y_{t-1} + a_t \quad t = 1, 2, \dots \quad a_t \sim RB(0, \sigma^2) \quad (3.3)$$

La estructura de correlación temporal representada por el modelo $AR(1)$ es como sigue:

$$\begin{array}{ccccccccc} \dots & Y_{t-2} & \rightarrow & Y_{t-1} & \rightarrow & Y_t & \rightarrow & Y_{t+1} & \rightarrow & Y_{t+2} & \dots \\ & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & \\ \dots & a_{t-2} & & a_{t-1} & & a_t & & a_{t+1} & & a_{t+2} & \dots \end{array}$$

es decir, la perturbación a_t entra en el sistema en el momento t e influyendo en Y_t y en su futuro, pero no en su pasado¹.

Para demostrar si el modelo $AR(1)$ es estacionario para cualquier valor del parámetro ϕ es preciso comprobar las condiciones de estacionariedad.

a) Estacionario en media:

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

Para que el proceso sea estacionario, la media ha de ser constante y finita, lo que implica:

$$E(Y_t) = \phi E(Y_t) \quad \longrightarrow \quad (1 - \phi) E(Y_t) = 0 \quad \longrightarrow \quad E(Y_t) = \frac{0}{1 - \phi} = 0$$

Por lo tanto, para que el proceso sea estacionario es necesario que el parámetro $\phi \neq 1$.

b) Estacionario en covarianza:

Para que el proceso sea estacionario, la varianza ha de ser constante y finita:

$$\begin{aligned} \gamma_0 &= E(Y_t - E(Y_t))^2 = E(\phi Y_{t-1} + a_t - 0)^2 = \\ &= \phi^2 E(Y_{t-1})^2 + E(a_t)^2 + 2\phi E(Y_{t-1} a_t) = \phi^2 V(Y_{t-1}) + \sigma^2 + 0 \end{aligned}$$

Dada la estructura de autocorrelación del proceso

$$E(Y_{t-1} a_t) = E[(Y_{t-1} - 0)(a_t - 0)] = cov(Y_{t-1} a_t) = 0$$

Bajo el supuesto de que el proceso es estacionario,

$$E(Y_{t-1})^2 = V(Y_{t-1}) = V(Y_t) = \gamma_0$$

por lo que:

$$\gamma_0 = \phi \gamma_0 + \sigma^2 \quad \longrightarrow \quad (1 - \phi^2) \gamma_0 = \sigma^2 \quad \longrightarrow \quad \gamma_0 = \frac{\sigma^2}{1 - \phi^2}$$

Para que el proceso sea estacionario, varianza constante y finita, es necesario que $|\phi| < 1$.

¹ Recuérdese además que suponemos que el proceso es no anticipante, es decir, el futuro no influye en el pasado.

En lo que se refiere a la función de autocovarianzas, la autocovarianza de orden k es:

$$\begin{aligned}\gamma_k &= E(Y_t - E(Y_t))(Y_{t-k} - E(Y_{t-k})) = E(Y_t Y_{t-k}) = E[(\phi Y_{t-1} + a_t) Y_{t-k}] \\ &= \phi E(Y_t Y_{t-k}) + E(a_t Y_{t-k}) = \phi \gamma_{k-1}\end{aligned}$$

Por lo que:

$$\begin{aligned}\gamma_1 &= \phi \gamma_0 \\ \gamma_2 &= \phi \gamma_1 \\ \gamma_3 &= \phi \gamma_2 \\ \vdots &= \vdots\end{aligned}$$

Por lo tanto, si el parámetro autorregresivo es $|\phi| < 1$, la varianza es finita y el resto de las autocovarianzas también lo son, y además dependen únicamente de los periodos de separación entre las variables y no del tiempo, por lo que el proceso es estacionario.

Se puede concluir, por lo tanto, que el proceso $AR(1)$ es estacionario si y sólo si $|\phi| < 1$.

La función de autocovarianzas de un proceso $AR(1)$ estacionario es:

$$\gamma_k = \begin{cases} \frac{\sigma^2}{1 - \phi^2} & k = 0 \\ \phi \gamma_{k-1} & k > 0 \end{cases}$$

Los coeficientes de autocorrelación de un proceso estacionario $AR(1)$ son:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi \gamma_{k-1}}{\gamma_0} = \phi \rho_{k-1}$$

La función de autocorrelación de un proceso $AR(1)$ estacionario es:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \phi \rho_{k-1} & k > 0 \end{cases}$$

Se puede demostrar fácilmente que la FAC de un $AR(1)$ es una función exponencial:

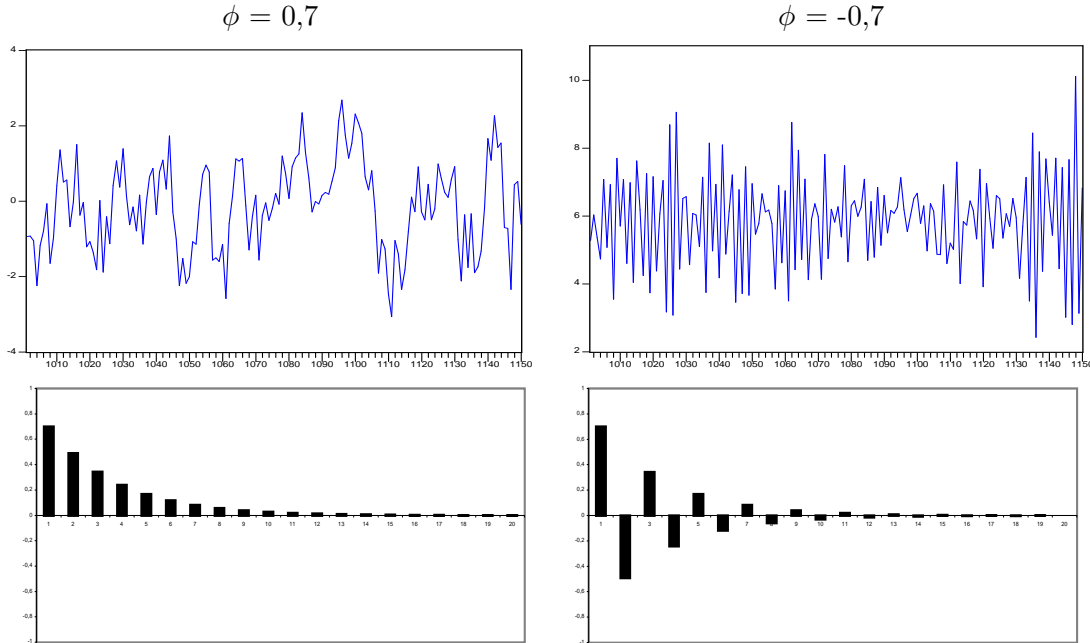
$$\begin{aligned}\rho_1 &= \phi \rho_0 = \phi \\ \rho_2 &= \phi \rho_1 = \phi^2 \\ \rho_3 &= \phi \rho_2 = \phi^3 \\ \vdots &= \vdots\end{aligned}$$

Por lo que:

$$\rho_k = \phi^k \quad k = 1, 2, 3, \dots$$

Dado que para procesos estacionarios el parámetro en valor absoluto es menor que la unidad, la función de autocorrelación decrece exponencialmente con k , es decir, se va aproximando a cero conforme $k \rightarrow \infty$ pero no llega nunca a valer cero.

Gráfico 3.1: Procesos Autorregresivos de orden 1: AR(1)



Las figuras del gráfico 3.1 muestran dos realizaciones de procesos $AR(1)$ y sus correspondientes correlogramas. Ambas han sido simuladas en base a la misma realización del ruido blanco a_t , por lo que sólo se diferencian en el valor del parámetro autorregresivo. El correlograma del proceso $AR(1)$ de parámetro $\phi = 0,7$ tiene todos los coeficientes de autocorrelación positivos con decrecimiento exponencial. Como era de esperar esto supone una serie que oscila en torno a cero (su media) con una dispersión estable, pero con rachas de observaciones por encima de la media seguidas de rachas de observaciones por debajo de la media. Este comportamiento por rachas de la serie será más pronunciado o persistente cuanto más se acerque el parámetro ϕ a la unidad. Para el proceso $AR(1)$ de parámetro negativo, $\phi = -0,7$, el correlograma presenta un decrecimiento exponencial con coeficientes que alternan de signo. La serie correspondiente a esta estructura de autocorrelación oscila en torno a cero (su media) con una dispersión estable, pero con un comportamiento muy ruidoso, cruzando constantemente la media.

Ejercicio 3.2. Compara la evolución de las realizaciones $AR(1)$ del gráfico 3.1 con la del ruido blanco de la gráfico 2.3. ¿Qué diferencias observas entre ellas? ¿A qué son debidas?

Una forma alternativa de escribir el modelo $AR(1)$ es la siguiente:

$$Y_t = \phi Y_{t-1} + a_t \quad \longrightarrow \quad (1 - \phi L) Y_t = a_t \quad \longrightarrow \quad Y_t = \frac{1}{1 - \phi L} a_t$$

Como un cociente de polinomios es un polinomio infinito y, en particular,

$$\frac{1}{1 - \phi L} = 1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \phi^4 L^4 + \dots$$

tenemos que:

$$\begin{aligned} Y_t &= (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \phi^4 L^4 + \dots) a_t \\ Y_t &= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + \phi^4 a_{t-4} + \dots \end{aligned} \quad (3.4)$$

Por lo tanto, el modelo $AR(1)$, no solo es una versión restringida del modelo lineal general (3.1), sino que también es una versión restringida del modelo general de medias móviles (3.2), en la que los infinitos coeficientes ψ_i dependen de un sólo parámetro ϕ .

El modelo (3.3) se puede generalizar para representar procesos estocásticos estacionarios con media distinta de cero:

$$Y_t = \delta + \phi Y_{t-1} + a_t \quad t = 1, 2, \dots \quad a_t \sim RB(0, \sigma^2) \quad (3.5)$$

Las características del modelo (3.5) son las mismas que las del modelo (3.3) salvo por la media que no es nula. Se puede demostrar que:

- El modelo (3.5) es estacionario sí y solo sí $|\phi| < 1$.
- Media:

$$E(Y_t) = E(\delta + \phi Y_{t-1} + a_t) = \delta + \phi E(Y_{t-1})$$

Si el proceso es estacionario, la media es constante y:

$$(1 - \phi) E(Y_t) = \delta \quad \longrightarrow \quad E(Y_t) = \frac{\delta}{1 - \phi} < \infty$$

- La función de autocovarianzas y la función de autocorrelación son las mismas que las del modelo (3.3):

$$\gamma_k = \begin{cases} \frac{\sigma^2}{1 - \phi^2} & k = 0 \\ \phi \gamma_{k-1} & k > 0 \end{cases} \quad \rho_k = \begin{cases} 1 & k = 0 \\ \phi \rho_{k-1} & k > 0 \end{cases}$$

Esto es lógico, ya que si restamos la media la media a ambos lados del modelo (3.5):

$$\begin{aligned} Y_t - E(Y_t) &= \delta + \phi Y_{t-1} - E(Y_t) + a_t \\ Y_t - \frac{\delta}{1 - \phi} &= \delta + \phi Y_{t-1} - \frac{\delta}{1 - \phi} + a_t = \delta \frac{1 - \phi}{1 - \phi} + \phi Y_{t-1} - \frac{\delta}{1 - \phi} + a_t = \\ &= \frac{\delta}{1 - \phi} - \phi \frac{\delta}{1 - \phi} + \phi Y_{t-1} - \frac{\delta}{1 - \phi} + a_t = \phi \left(Y_{t-1} - \frac{\delta}{1 - \phi} \right) + a_t \end{aligned}$$

Por lo que²:

$$Y_t^* = \phi Y_{t-1}^* + a_t, \quad \text{donde } Y_t^* = Y_t - E(Y_t) \quad \text{con } E(Y_t^*) = 0$$

$$\begin{aligned} V(Y_t) &= E(Y_t - E(Y_t))^2 = E(Y_t^*)^2 = E(Y_t^* - 0)^2 = V(Y_t^*) \\ \gamma_k(Y) &= cov(Y_t, Y_{t+k}) = E(Y_t - E(Y_t))(Y_{t+k} - E(Y_{t+k})) = E(Y_t^*)(Y_{t+k}^*) = E(Y_t^* - 0)(Y_{t+k}^* - 0) = \gamma_k(Y^*) \end{aligned}$$

3.2.2. Proceso AR(p).

El proceso autorregresivo de orden p , expresa Y_t en función de su pasado hasta el retardo $t-p$ y una innovación contemporánea:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad a_t \sim RB(0, \sigma^2) \quad t = 1, 2, \dots \quad (3.6)$$

En términos del operador de retardos,

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) Y_t = a_t \quad \longrightarrow \quad \phi_p(L) Y_t = a_t$$

donde $\phi_p(L)$ recibe el nombre de polinomio autorregresivo y $(\phi_1, \phi_2, \dots, \phi_p)$ es el vector de parámetros autorregresivos.

En primer lugar es preciso comprobar si el proceso $AR(p)$ cumple las condiciones de estacionariedad para cualquier valor de los parámetros. Esta comprobación que es relativamente sencilla para el modelo $AR(1)$, se complica para modelos autorregresivos de orden mayor cuyos parámetros han de satisfacer restricciones complejas para ser estacionarios. El siguiente teorema proporciona condiciones necesarias y suficientes para que el modelo $AR(p)$ sea estacionario.

Teorema: Un proceso autorregresivo finito $AR(p)$ es estacionario sí y solo sí el modulo de las raíces del polinomio autorregresivo $\phi_p(L)$ está fuera del círculo unidad.

Ejemplo 3.1. Condiciones de estacionariedad.

Modelo $AR(1)$: $Y_t = \phi Y_{t-1} + a_t \quad \Rightarrow \quad (1 - \phi L) Y_t = a_t$

Polinomio autorregresivo: $\phi_1(L) = 1 - \phi L$

Raíces: $1 - \phi L = 0 \quad \Rightarrow \quad L = \frac{1}{\phi}$

Condición de estacionariedad: $|L| = \left| \frac{1}{\phi} \right| > 1 \quad \Rightarrow \quad |\phi| < 1$

Modelo $AR(2)$: $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t \quad \Rightarrow \quad (1 - \phi_1 L - \phi_2 L^2) Y_t = a_t$

Polinomio autorregresivo: $\phi_2(L) = 1 - \phi_1 L - \phi_2 L^2$

Raíces: $1 - \phi_1 L - \phi_2 L^2 = 0 \quad \Rightarrow \quad L_1, L_2 = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4 \times \phi_2}}{-2 \times \phi_2}$

Condición de estacionariedad:

$$|L_1| = \left| \frac{\phi_1 + \sqrt{\phi_1^2 + 4 \times \phi_2}}{-2 \times \phi_2} \right| > 1 \quad y \quad |L_2| = \left| \frac{\phi_1 - \sqrt{\phi_1^2 + 4 \times \phi_2}}{-2 \times \phi_2} \right| > 1$$

¿Cómo se calcula el modulo de estas raíces?

- Si el radicando es positivo, $(\phi_1^2 + 4 \times \phi_2) > 0$, las raíces son reales y el modulo es el valor absoluto.
- Si el radicando es negativo, $(\phi_1^2 + 4 \times \phi_2) < 0$, las raíces son complejas de la forma

$$L_1, L_2 = a \pm b i$$

y su modulo es igual a:

$$|L_1| = |L_2| = \sqrt{a^2 + b^2}$$

Las otras dos condiciones que tiene que cumplir el modelo lineal general y, por lo tanto, también el $AR(p)$ como aproximación al mismo, es que el modelo fuera *no anticipante e invertible*.

Todo modelo autorregresivo finito, $AR(p)$, cumple estas dos condiciones para cualquier valor de los parámetros autorregresivos ya que, por definición, está escrito en forma autorregresiva. Es decir, es *no anticipante* porque su formulación hace depender al valor de Y_t de su pasado y no de su futuro y es *invertible* porque su formulación finita hace que se cumpla obligatoriamente la condición

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty$$

El modelo $AR(p)$ (3.6) es una versión restringida del modelo (3.1) ya que supone que todos los parámetros $\pi_i = 0$, $\forall i > p$. Por otro lado, el modelo $AR(p)$ se puede escribir como un medias móviles de orden infinito:

$$\phi_p(L) Y_t = a_t \rightarrow Y_t = \frac{1}{\phi_p(L)} a_t = \Psi_{\infty}(L) a_t = (1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots) a_t$$

$$Y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \dots$$

donde los parámetros ψ_i están restringidos a depender del vector de parámetros autorregresivos $(\phi_1, \phi_2, \dots, \phi_p)$. Por lo tanto, el modelo (3.6) es también una versión restringida del modelo lineal general (3.2).

Las características del proceso estacionario $AR(p)$ estacionario son:

a) Media:

$$\begin{aligned} E(Y_t) &= E(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t) = \\ &= \phi_1 E(Y_{t-1}) + \phi_2 E(Y_{t-2}) + \dots + \phi_p E(Y_{t-p}) + E(a_t) \end{aligned}$$

Como el proceso es estacionario, la media es constante:

$$(1 - \phi_1 - \phi_2 - \dots - \phi_p) E(Y_t) = 0 \quad \rightarrow \quad E(Y_t) = 0$$

Este modelo se puede generalizar para representar series con media distinta de cero. El siguiente modelo $AR(p)$:

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t$$

tiene media:

$$(1 - \phi_1 - \phi_2 - \dots - \phi_p) E(Y_t) = \delta \quad \rightarrow \quad E(Y_t) = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$$

- b) La función de autocorrelación, ρ_k , $k = 0, 1, 2, \dots$ de un proceso $AR(p)$ tiene la misma estructura que la de un proceso $AR(1)$, es decir, decrece exponencialmente hacia cero sin truncarse. Ahora bien, para valores de $p > 1$ al ser el modelo más complejo, la estructura de la FAC es también más rica y puede presentar más variedad de formas.

Ejemplo 3.2. Características del modelo $AR(2)$.

Consideremos el modelo autorregresivo de orden 2 estacionario:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t \quad a_t \sim RB(0, \sigma^2) \quad t = 1, 2, \dots \quad (3.7)$$

Las características de este proceso son las siguientes:

- a) Media. Como el proceso es estacionario, la media es constante:

$$(1 - \phi_1 - \phi_2) E(Y_t) = 0 \quad \rightarrow \quad E(Y_t) = 0$$

- b) Función de autocovarianzas:

$$\begin{aligned} \gamma_0 &= E(Y_t - E(Y_t))^2 = E(Y_t)^2 = E(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t)^2 = \\ &= \phi_1^2 E(Y_{t-1})^2 + \phi_2^2 E(Y_{t-2})^2 + E(a_t)^2 + 2\phi_1 \phi_2 E(Y_{t-1} Y_{t-2}) + \\ &+ 2\phi_1 E(Y_{t-1} a_t) + 2\phi_2 E(Y_{t-2} a_t) = \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + \sigma^2 + 2\phi_1 \phi_2 \gamma_1 \end{aligned}$$

$$\rightarrow (1 - \phi_1^2 - \phi_2^2) \gamma_0 = \sigma^2 + 2\phi_1 \phi_2 \gamma_1$$

$$\rightarrow \gamma_0 = \frac{\sigma^2 + 2\phi_1 \phi_2 \gamma_1}{1 - \phi_1^2 - \phi_2^2}$$

$$\gamma_1 = E(Y_t - E(Y_t))(Y_{t-1} - E(Y_{t-1})) = E(Y_t Y_{t-1}) =$$

$$= E[(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t) Y_{t-1}] =$$

$$= \phi_1 E(Y_{t-1})^2 + \phi_2 E(Y_{t-2} Y_{t-1}) + E(a_t Y_{t-1}) = \phi_1 \gamma_0 + \phi_2 \gamma_1$$

$$\rightarrow \gamma_1 = \frac{\phi_1 \gamma_0}{1 - \phi_2}$$

Estas dos ecuaciones proporcionan las dos primeras autocovarianzas, γ_0, γ_1 en función de los parámetros autorregresivos (ϕ_1, ϕ_2) y de σ^2 , la varianza del ruido blanco.

Las autocovarianzas, $\gamma_k, \forall k > 1$ son:

$$\begin{aligned}\gamma_k &= E(Y_t - E(Y_t))(Y_{t-k} - E(Y_{t-k})) = E(Y_t Y_{t-k}) = \\ &= E[(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t) Y_{t-k}] = \\ &= \phi_1 E(Y_{t-1} Y_{t-k}) + \phi_2 E(Y_{t-2} Y_{t-k}) + E(a_t Y_{t-k}) = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}\end{aligned}$$

La función de autocovarianzas de un modelo $AR(2)$ es, por lo tanto:

$$\gamma_k = \begin{cases} \gamma_0 & k = 0 \\ \gamma_1 & k = 1 \\ \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} & k > 1 \end{cases}$$

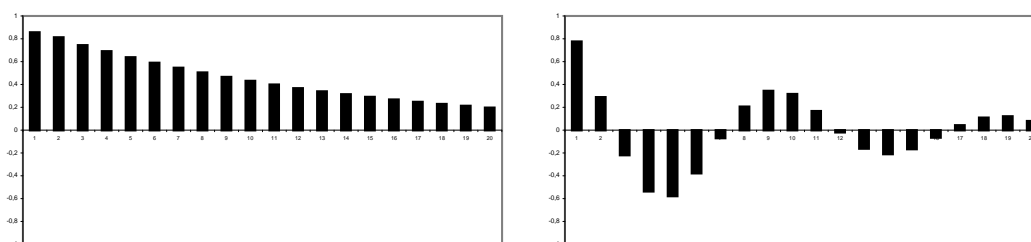
c) Los coeficientes de autocorrelación son:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}}{\gamma_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \quad k = 1, 2, 3, \dots$$

La función de autocorrelación de un modelo $AR(2)$ es:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} & k > 0 \end{cases}$$

Gráfico 3.2: Procesos Autorregresivos de orden 2



El gráfico 3.2 muestra diferentes estructuras de funciones de autocorrelación para modelos $AR(2)$ estacionarios. Cuando las raíces del polinomio autorregresivo $\phi_2(L)$ son reales, la función de autocorrelación es una función que decae exponencialmente con todos los coeficientes positivos o con alternancia de signos, como para el modelo $AR(1)$. Si las raíces del polinomio autorregresivo son complejas entonces la función de autocorrelación decae exponencialmente hacia cero pero con forma de onda seno-coseno.

3.3. Procesos de Medias Móviles: MA(q).

El modelo de medias móviles de orden finito q , $MA(q)$, es una aproximación natural al modelo lineal general (3.2). Se obtiene un modelo finito por el simple procedimiento de truncar el modelo de medias móviles de orden infinito:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad a_t \sim RB(0, \sigma^2)$$

Para estudiar sus características, comenzaremos por el más sencillo, el modelo medias móviles de primer orden, $MA(1)$.

3.3.1. Proceso MA(1).

El modelo $MA(1)$ determina el valor de Y en el momento t en función de la innovación contemporánea y su primer retardo:

$$Y_t = a_t - \theta a_{t-1} \quad a_t \sim RB(0, \sigma^2) \quad t = 1, 2, \dots \quad (3.8)$$

Los procesos de medias móviles se suelen denominar procesos de memoria corta, mientras que a los autorregresivos se les denomina procesos de memoria larga. Esto es debido a que, como se puede observar en la estructura de correlación temporal del modelo $MA(1)$ representada en el esquema siguiente, la perturbación a_t aparece en el sistema en el momento t e influye en Y_t e Y_{t+1} únicamente, por lo que su memoria es de un solo periodo. Sin embargo, en un proceso autorregresivo la perturbación a_t aparece en el sistema en el momento t , influye en Y_t y a través de Y_t en las observaciones futuras, permaneciendo su influencia en el sistema un periodo más largo.

$$\begin{array}{cccccccc} \dots & Y_{t-2} & & Y_{t-1} & & Y_t & & Y_{t+1} & & Y_{t+2} & \dots \\ \nearrow & \uparrow & \nearrow & \uparrow & \nearrow & \uparrow & \nearrow & \uparrow & \nearrow & \uparrow & \nearrow \\ \dots & a_{t-2} & & a_{t-1} & & a_t & & a_{t+1} & & a_{t+2} & \dots \end{array}$$

Vamos a comprobar si el modelo $MA(1)$ cumple las condiciones de estacionariedad para cualquier valor del parámetro θ .

a) Estacionario en media:

$$E(Y_t) = E(a_t - \theta a_{t-1}) = 0$$

Por tanto, el modelo $MA(1)$ es estacionario en media para todo valor del parámetro θ .

b) Estacionario en covarianza:

$$\begin{aligned} \gamma_0 &= E(Y_t - E(Y_t))^2 = E(Y_t)^2 = E(a_t - \theta a_{t-1})^2 = \\ &= E(a_t)^2 + \theta^2 E(a_{t-1})^2 - 2\theta E(a_t a_{t-1}) = \sigma^2 + \theta^2 \sigma^2 - 0 \\ \gamma_0 &= (1 + \theta^2) \sigma^2 < \infty \end{aligned}$$

Es decir, la varianza del modelo $MA(1)$ es constante y finita para cualquier valor de θ .

Las autocovarianzas, γ_k , $k > 0$:

$$\begin{aligned}\gamma_1 &= E(Y_t - E(Y_t))(Y_{t-1} - E(Y_{t-1})) = E(a_t - \theta a_{t-1})(a_{t-1} - \theta a_{t-2}) = \\ &= E(a_t a_{t-1}) - \theta E(a_{t-1})^2 - \theta E(a_t a_{t-2}) + \theta^2 E(a_{t-1} a_{t-2}) = -\theta \sigma^2 < \infty\end{aligned}$$

$$\begin{aligned}\gamma_2 &= E(Y_t - E(Y_t))(Y_{t-2} - E(Y_{t-2})) = E(a_t - \theta a_{t-1})(a_{t-2} - \theta a_{t-3}) = \\ &= E(a_t a_{t-2}) - \theta E(a_{t-1} a_{t-2}) - \theta E(a_t a_{t-3}) + \theta^2 E(a_{t-1} a_{t-3}) = 0\end{aligned}$$

La función de autocovarianzas de un $MA(1)$ es:

$$\gamma_k = \begin{cases} \gamma_0 &= (1 + \theta^2) \sigma^2 & k = 0 \\ \gamma_1 &= -\theta \sigma^2 & k = 1 \\ \gamma_k &= 0 & k > 1 \end{cases}$$

La función de autocovarianzas es finita y depende sólo de k y no del tiempo, para cualquier valor del parámetro θ .

Por lo tanto, se puede concluir que no es necesario poner restricciones sobre el valor del parámetro θ para que el modelo $MA(1)$ sea estacionario.

La función de autocorrelación de un proceso $MA(1)$ es:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \frac{-\theta}{1 + \theta^2} & k = 1 \\ 0 & k > 1 \end{cases}$$

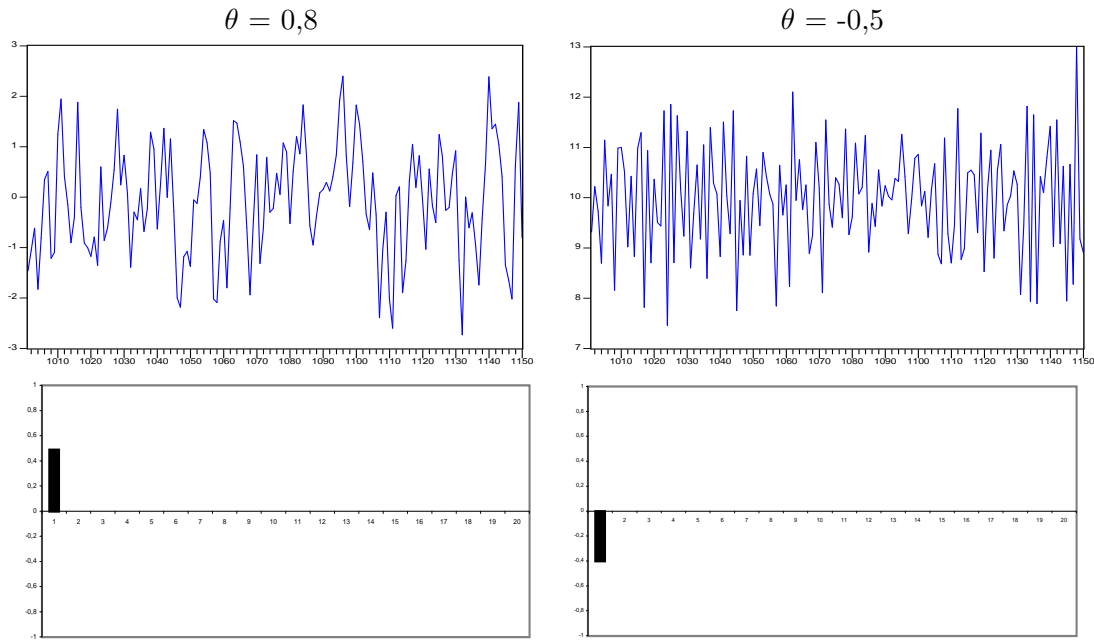
La función de autocorrelación de un proceso $MA(1)$ se caracteriza por ser una función truncada que sufre un corte brusco tomando el valor cero a partir del retardo 1.

El gráfico 3.3 recoge dos realizaciones de los siguientes procesos $MA(1)$ con sus funciones de autocorrelación correspondientes:

$$\begin{aligned}(a) \quad Y_t &= a_t + 0,8 a_{t-1} \\ (b) \quad Y_t &= a_t - 0,5 a_{t-1}\end{aligned}$$

Se puede observar que ambas funciones de autocorrelación se truncan en el primer retardo como corresponde a un proceso medias móviles de orden 1. El coeficiente de autocorrelación, ρ_1 , es positivo o negativo dependiente del signo del parámetro de medias móviles. En cuanto a las series generadas con estos modelos, se nota que ambas oscilan en torno a la media cero con una

Gráfico 3.3: Procesos Medias Móviles de orden 1: MA(1)



varianza constante y que la serie generada por el modelo (a) con parámetro positivo es más suave que la generada por el modelo (b) con parámetro negativo.

Los modelos lineales tienen que cumplir, además, las condiciones de ser modelos *no anticipantes* e *invertibles*. Estas condiciones se comprueban fácilmente en modelos escritos de forma autorregresiva. Además, la representación autorregresiva siempre es más intuitiva en el sentido de que si el objetivo es predecir parece lógico que el modelo relacione lo que sucede en el presente con lo sucedido en el pasado de manera que las predicciones de futuras observaciones se construyan en base al pasado y presente observado. En principio, un modelo de medias móviles no parece estar escrito de esta forma ya que Y_t no depende directamente de las observaciones pasadas sino de la innovación y de sus valores pasados. Ahora bien, los modelos de medias móviles también se pueden escribir de forma autorregresiva. Escribiendo el modelo (3.8) en términos del operador de retardos, se tiene que:

$$Y_t = (1 - \theta L) a_t \quad \rightarrow \quad \frac{1}{1 - \theta L} Y_t = a_t \quad \rightarrow \quad \Pi_\infty(L) Y_t = a_t$$

$$(1 - \theta L + \theta^2 L^2 - \theta^3 L^3 + \dots) Y_t = a_t$$

$$Y_t = a_t + \theta Y_{t-1} - \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \dots$$

Esta representación $AR(\infty)$ es restringida ya que, como se puede observar, todos los parámetros autorregresivos dependen del único parámetro de medias móviles del modelo, θ , como sigue, $\pi_i = (-\theta)^i$.

El modelo $MA(1)$ es *no anticipante* porque el futuro no influye en el pasado. Para que el modelo $MA(1)$ sea *invertible* es preciso que su representación autorregresiva sea convergente, es decir,

que se cumpla que la influencia de Y_{t-k} va siendo menor conforme nos alejamos en el pasado. Esta condición se cumple si:

$$\sum_{i=1}^{\infty} (-\theta^i)^2 < \infty$$

Para que se cumpla la condición anterior es necesario y suficiente que $|\theta| < 1$, por lo que el modelo $MA(1)$ no es siempre invertible sino que hay que poner restricciones sobre el valor del parámetro θ .

Bajo las condiciones de invertibilidad, en este caso, $|\theta| < 1$, el modelo $MA(1)$, y cualquier modelo de medias móviles, en general, se puede escribir en forma autorregresiva, por lo que si deseamos una representación autorregresiva para el modelo, no hace falta empezar por un $AR(p)$. Siempre vamos a utilizar modelos invertibles porque de esta forma siempre se puede expresar el valor contemporáneo de Y_t en función de su pasado observado.

El modelo (3.8) se puede generalizar para representar procesos con media distinta de cero:

$$Y_t = \delta + a_t - \theta a_{t-1} \quad a_t \sim RB(0, \sigma^2) \quad (3.9)$$

Este modelo tiene la misma función de autocovarianzas y, por lo tanto, de autocorrelación, que el modelo (3.8). La única diferencia es que su media es distinta de cero:

$$E(Y_t) = E[\delta + a_t - \theta a_{t-1}] = \delta$$

3.3.2. Proceso $MA(q)$.

El modelo medias móviles finito de orden q , es la aproximación natural al modelo lineal general (3.2), ya que supone su truncamiento bajo el supuesto de que:

$$\psi_i = 0 \quad \forall i > q$$

Este modelo expresa el valor de Y_t en función de la innovación contemporánea y de su pasado hasta el retardo q :

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad a_t \sim RB(0, \sigma^2) \quad (3.10)$$

En términos del operador de retardos queda como sigue:

$$Y_t = (1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3 - \dots - \theta_q L^q) a_t \quad \rightarrow \quad Y_t = \theta_q(L) a_t$$

donde $\theta_q(L)$ se denomina polinomio de medias móviles y $(\theta_1, \theta_2, \theta_3, \dots, \theta_q)$ es el vector de parámetros de medias móviles.

El modelo $MA(q)$ es una generalización del modelo $MA(1)$, por lo tanto sus características van a ser muy similares. Ahora bien, al introducir más retardos de la perturbación en el modelo, la memoria aumenta y la estructura dinámica representada por el modelo es más rica. Si una perturbación entra en el momento t , influye en Y_t , en Y_{t+1} , y en valores posteriores hasta Y_{t+q} .

Por lo tanto, la perturbación a_t en un modelo $MA(q)$ permanece q periodos en el sistema. Esta memoria más larga del $MA(q)$ se verá reflejada en la estructura de la función de autocovarianzas y / o autocorrelación.

Vamos a comprobar si el modelo $MA(q)$ es estacionario para cualquier valor de los parámetros de medias móviles. Para ello sería preciso comprobar si se cumplen las condiciones de estacionariedad. Pero teniendo en cuenta que el modelo medias móviles finito de orden q no es más que el resultado de truncar el modelo lineal general (3.2) a partir del retardo q , será estacionario bajo las mismas condiciones que el modelo general, es decir, si se cumple la condición de que la sucesión de los parámetros del modelo es convergente:

$$\sum_{i=1}^{\infty} \psi_i^2 = \sum_{i=1}^q \theta_i^2 < \infty$$

Como el número de parámetros de un $MA(q)$ es finito, esta condición siempre se cumple y, por lo tanto, todos los modelos de medias móviles finitos son siempre estacionarios para cualquier valor de sus parámetros.

El modelo de $MA(q)$ (3.10) tiene media constante y cero, varianza constante y finita y su función de autocovarianzas está truncada a partir del retardo q , es decir,

$$\rho_k = \begin{cases} \rho_k \neq 0 & k = 1, 2, \dots, q \\ \rho_k = 0 & k > q \end{cases}$$

El modelo (3.10) se puede generalizar fácilmente para representar modelos con media no nula:

$$Y_t = \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad a_t \sim RB(0, \sigma^2) \quad (3.11)$$

Ejemplo 3.3. Características del modelo $MA(2)$.

Consideremos el modelo de medias móviles de orden 2:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad a_t \sim RB(0, \sigma^2)$$

Este proceso es estacionario para cualquier valor de θ_1, θ_2 y sus características son:

a) Media: $E(Y_t) = E(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}) = 0$

b) Función de autocovarianzas, γ_k , $k = 0, 1, 2, 3, \dots$:

$$\begin{aligned} \gamma_0 &= E[Y_t - E(Y_t)]^2 = E(Y_t)^2 = E(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})^2 = \\ &= E(a_t)^2 + \theta_1^2 E(a_{t-1})^2 + \theta_2^2 E(a_{t-2})^2 - 2\theta_1 E(a_t a_{t-1}) - 2\theta_2 E(a_t a_{t-2}) + \\ &+ 2\theta_1 \theta_2 E(a_{t-1} a_{t-2}) = \sigma^2 + \theta_1^2 \sigma^2 + \theta_2^2 \sigma^2 = (1 + \theta_1^2 + \theta_2^2) \sigma^2 \end{aligned}$$

$$\begin{aligned}
\gamma_1 &= E[(Y_t - E(Y_t))(Y_{t-1} - E(Y_t))] = E(Y_t Y_{t-1}) = \\
&= E[(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})(a_{t-1} - \theta_1 a_{t-2} - \theta_2 a_{t-3})] = \\
&= E(a_t a_{t-1}) - \theta_1 E(a_t a_{t-2}) - \theta_2 E(a_t a_{t-3}) - \theta_1 E(a_{t-1})^2 + \theta_1^2 E(a_{t-1} a_{t-2}) + \\
&+ \theta_1 \theta_2 E(a_{t-1} a_{t-3}) - \theta_2 E(a_{t-1} a_{t-2}) + \theta_2 \theta_1 E(a_{t-2})^2 + \theta_2^2 E(a_{t-2} a_{t-3}) = \\
&= -\theta_1 \sigma^2 + \theta_1 \theta_2 \sigma^2 = (-\theta_1 + \theta_1 \theta_2) \sigma^2 \\
\gamma_2 &= E[(Y_t - E(Y_t))(Y_{t-2} - E(Y_t))] = E(Y_t Y_{t-2}) = \\
&= E[(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})(a_{t-2} - \theta_1 a_{t-3} - \theta_2 a_{t-4})] = \\
&= E(a_t a_{t-2}) - \theta_1 E(a_t a_{t-3}) - \theta_2 E(a_t a_{t-4}) - \theta_1 E(a_{t-1} a_{t-2}) + \\
&+ \theta_1^2 E(a_{t-1} a_{t-3}) + \theta_1 \theta_2 E(a_{t-1} a_{t-4}) - \theta_2 E(a_{t-2})^2 + \\
&+ \theta_2 \theta_1 E(a_{t-2} a_{t-3}) + \theta_2^2 E(a_{t-2} a_{t-4}) = -\theta_2 \sigma^2 \\
\gamma_3 &= E[(Y_t - E(Y_t))(Y_{t-3} - E(Y_t))] = E(Y_t Y_{t-3}) = \\
&= E[(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})(a_{t-3} - \theta_1 a_{t-4} - \theta_2 a_{t-5})] = \\
&= E(a_t a_{t-3}) - \theta_1 E(a_t a_{t-4}) - \theta_2 E(a_t a_{t-5}) - \theta_1 E(a_{t-1} a_{t-3}) + \\
&+ \theta_1^2 E(a_{t-1} a_{t-4}) + \theta_1 \theta_2 E(a_{t-1} a_{t-5}) - \theta_2 E(a_{t-2} a_{t-3}) + \\
&+ \theta_2 \theta_1 E(a_{t-2} a_{t-4}) + \theta_2^2 E(a_{t-2} a_{t-5}) = 0
\end{aligned}$$

La función de autocovarianzas de un $MA(2)$ es:

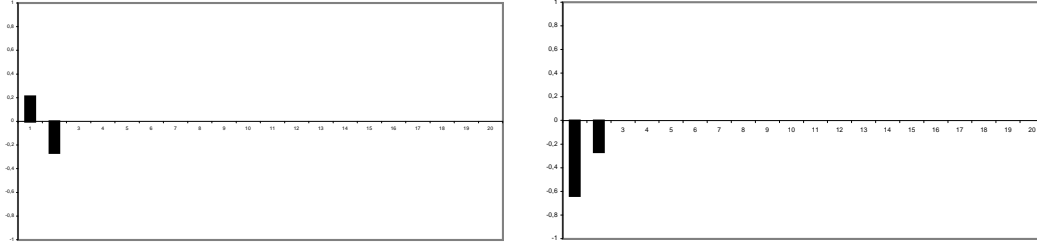
$$\gamma_k = \begin{cases} \gamma_0 &= (1 + \theta_1^2 + \theta_2^2) \sigma^2 & k = 0 \\ \gamma_1 &= (-\theta_1 + \theta_1 \theta_2) \sigma^2 & k = 1 \\ \gamma_2 &= -\theta_2 \sigma^2 & k = 2 \\ \gamma_k &= 0 & k > 2 \end{cases}$$

La función de autocorrelación de un $MA(2)$ es:

$$\rho_k = \begin{cases} \rho_1 &= \frac{-\theta_1 + \theta_1 \theta_2}{1 + \theta_1^2 + \theta_2^2} & k = 1 \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} & k = 2 \\ \rho_k &= 0 & k > 2 \end{cases}$$

La FAC de un proceso $MA(2)$ es una función truncada en el retardo 2. Los coeficientes de autocorrelación, ρ_1 , ρ_2 , pueden ser positivos, negativos, o de distinto signo dependiendo de

Gráfico 3.4: Procesos Medias Móviles de orden 2: MA(2)



los valores de los parámetros autorregresivos. El gráfico 3.4 muestra dos ejemplos de FAC de un $MA(2)$.

El modelo $MA(q)$ (3.11) tiene también una representación $AR(\infty)$ restringida:

$$\begin{aligned}
 Y_t &= (1 - \theta L - \theta_2 L^2 - \dots - \theta_q L^q) a_t \\
 &\rightarrow \frac{1}{1 - \theta L - \theta_2 L^2 - \dots - \theta_q L^q} Y_t = a_t \\
 &\rightarrow \Pi_\infty(L) Y_t = a_t \quad \rightarrow (1 + \pi_1 L + \pi_2 L^2 + \pi_3 L^3 + \dots) Y_t = a_t \\
 &\rightarrow Y_t = a_t - \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \pi_3 Y_{t-3} + \dots
 \end{aligned}$$

Esta representación $AR(\infty)$ es restringida ya que todos los parámetros autorregresivos, π_i , dependen del vector de parámetros de medias móviles del modelo $(\theta_1, \theta_2, \dots, \theta_q)$.

El modelo $MA(q)$ es *no anticipante* porque el futuro no influye en el pasado y será *invertible* si su representación autorregresiva es tal que la influencia de Y_{t-k} es menor conforme nos alejamos en el pasado. Esta condición se cumple si:

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty$$

Por lo tanto, el modelo $MA(q)$ no es invertible para cualquier valor del vector de parámetros de medias móviles, sino que estos tendrán que cumplir algunas restricciones. Derivar estas restricciones fue muy sencillo para el $MA(1)$, pero se complica mucho al aumentar el orden del modelo de medias móviles. El siguiente teorema proporciona condiciones necesarias y suficientes para el un modelo de medias móviles sea invertible.

Teorema: Un proceso de medias móviles finito $MA(q)$ es invertible sí y solo sí el modulo de las raíces del polinomio de medias móviles $\theta_q(L)$ está fuera del círculo unidad.

Ejemplo 3.4. Condiciones de invertibilidad.

Modelo MA(1): $Y_t = a_t - \theta a_{t-1} \Rightarrow Y_t = (1 - \theta L) a_t$

Polinomio medias móviles: $\theta_1(L) = 1 - \theta L$

Raíces: $1 - \theta L = 0 \Rightarrow L = \frac{1}{\theta}$

Condición de invertibilidad: $|L| = \left| \frac{1}{\theta} \right| > 1 \Rightarrow |\theta| < 1$

Modelo MA(2): $Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \Rightarrow Y_t = (1 - \theta_1 L - \theta_2 L^2) a_t$

Polinomio de medias móviles: $\theta_2(L) = 1 - \theta_1 L - \theta_2 L^2$

Raíces: $1 - \theta_1 L - \theta_2 L^2 = 0 \Rightarrow L_1, L_2 = \frac{\theta_1 \pm \sqrt{\theta_1^2 + 4 \times \theta_2}}{-2 \times \theta_2}$

Condición de invertibilidad:

$$|L_1| = \left| \frac{\theta_1 + \sqrt{\theta_1^2 + 4 \times \theta_2}}{-2 \times \theta_2} \right| > 1 \quad y \quad |L_2| = \left| \frac{\theta_1 - \sqrt{\theta_1^2 + 4 \times \theta_2}}{-2 \times \theta_2} \right| > 1$$

Ejercicio 3.3. Calcula la media y la varianza del modelo (3.11).

3.4. Procesos Autorregresivos de Medias Móviles: ARMA(p,q).

Los procesos autorregresivos de medias móviles determinan Y_t en función de su pasado hasta el retardo p , de la innovación contemporánea y el pasado de la innovación hasta el retardo q :

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad a_t \sim RB(0, \sigma^2) \quad (3.12)$$

Este modelo se puede escribir en términos del operador de retardos como sigue:

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

$$\phi_p(L) Y_t = \theta_q(L) a_t$$

donde $\phi_p(L)$ es el polinomio autorregresivo y $\theta_q(L)$ es el polinomio medias móviles.

El modelo (3.12) es una aproximación finita al modelo lineal general tanto en su forma $AR(\infty)$ (3.1) como $MA(\infty)$ (3.2). De hecho, si es estacionario su representación $MA(\infty)$ es

$$Y_t = \frac{\theta_q(L)}{\phi_p(L)} a_t \rightarrow Y_t = \psi_\infty(L) a_t \rightarrow Y_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \dots$$

y si es invertible una representación $AR(\infty)$

$$\frac{\phi_p(L)}{\theta_p(L)} Y_t = a_t \quad \rightarrow \quad \pi_\infty(L) Y_t = a_t \quad \rightarrow \quad Y_t = a_t + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \pi_3 Y_{t-3} + \dots$$

Tanto los pesos de la representación MA infinita, como de la forma AR infinita, están restringidos a depender del vector finito de parámetros ARMA: $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$.

¿Es estacionario el modelo $ARMA(p, q)$ para cualquier valor de sus parámetros?

Teorema: Un proceso autorregresivo de medias móviles finito $ARMA(p, q)$ es estacionario sí y solo sí el modulo de las raíces del polinomio autorregresivo $\phi_p(L)$ está fuera del círculo unidad.

Las condiciones de estacionariedad del modelo $ARMA(p, q)$ vienen impuestas por la parte autorregresiva, dado que la parte de medias móviles finita siempre es estacionaria.

Para comprobar si el modelo $ARMA(p, q)$ es no anticipante e invertible, se estudia su representación autorregresiva general.

Teorema: Un proceso autorregresivo de medias móviles finito $ARMA(p, q)$ es invertible sí y solo sí el modulo de las raíces del polinomio medias móviles $\theta_p(L)$ está fuera del círculo unidad.

Las condiciones de invertibilidad del modelo $ARMA(p, q)$ vienen impuestas por la parte de medias móviles, dado que la parte autorregresiva finita siempre es invertible porque está directamente escrita en forma autorregresiva

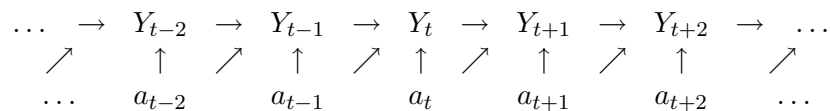
El modelo $ARMA(p, q)$ va a compartir las características de los modelos $AR(p)$ y $MA(q)$ ya que contiene ambas estructuras a la vez. El modelo $ARMA(p, q)$ tiene media cero, varianza constante y finita y una función de autocovarianzas infinita. La función de autocorrelación es infinita decreciendo rápidamente hacia cero pero sin truncarse.

Ejemplo 3.5. Modelo $ARMA(1, 1)$.

Consideremos el modelo $ARMA(1, 1)$ en el Y_t se determina en función de su pasado hasta el primer retardo, la innovación contemporánea y el pasado de la innovación hasta el retardo 1:

$$Y_t = \phi Y_{t-1} + a_t - \theta a_{t-1} \quad a_t \sim RB(0, \sigma^2) \quad t = 1, 2, \dots \quad (3.13)$$

La estructura de dependencia temporal representada por este modelo es:



La memoria de este proceso es larga debido a la presencia de estructura autorregresiva. La perturbación a_t influye directamente en Y_t y en Y_{t-1} , pero de forma indirecta, a través de la cadena de dependencia temporal de las Y , influye en todo el futuro del proceso.

Para comprobar la estacionariedad, se calculan las raíces del polinomio autorregresivo:

$$1 - \phi L = 0 \quad \rightarrow \quad L = \frac{1}{\phi} \quad \rightarrow \quad |L| = \left| \frac{1}{\phi} \right| \quad \rightarrow \quad |\phi| < 1$$

Para comprobar la invertibilidad, se calculan las raíces del polinomio medias móviles:

$$1 - \theta L = 0 \quad \rightarrow \quad L = \frac{1}{\theta} \quad \rightarrow \quad |L| = \left| \frac{1}{\theta} \right| \quad \rightarrow \quad |\theta| < 1$$

Las características del proceso $ARMA(1,1)$ estacionario son:

a) Media: $E(Y_t) = E(\phi Y_{t-1} + a_t - \theta a_{t-1}) = \phi E(Y_t)$

$$\rightarrow (1 - \phi) E(Y_t) = 0 \quad \rightarrow \quad E(Y_t) = 0$$

b) Función de autocovarianzas:

$$\begin{aligned} \gamma_0 &= E(Y_t - E(Y_t))^2 = E(Y_t)^2 = E(\phi Y_{t-1} + a_t - \theta a_{t-1})^2 = \\ &= \phi^2 E(Y_{t-1})^2 + E(a_t)^2 + \theta^2 E(a_{t-1})^2 + 2\phi E(Y_{t-1} a_t) - 2\phi\theta E(Y_{t-1} a_{t-1}) - \\ &\quad - 2\theta E(a_t a_{t-1}) = \phi^2 \gamma_0 + \sigma^2 + \theta^2 \sigma^2 - 2\phi\theta \sigma^2 = \\ &= \phi^2 \gamma_0 + (1 + \theta^2 - 2\phi\theta) \sigma^2 \quad \rightarrow \quad \gamma_0 = \frac{(1 + \theta^2 - 2\phi\theta) \sigma^2}{1 - \phi^2} \end{aligned}$$

$$\begin{aligned} \gamma_1 &= E(Y_t - E(Y_t))(Y_{t-1} - E(Y_{t-1})) = \\ &= E(Y_t Y_{t-1}) = E[(\phi Y_{t-1} + a_t - \theta a_{t-1}) Y_{t-1}] = \\ &= \phi E(Y_{t-1})^2 + E(Y_{t-1} a_t) - \theta E(Y_{t-1} a_{t-1}) = \phi \gamma_0 - \theta \sigma^2 \end{aligned}$$

$$\begin{aligned} \gamma_2 &= E(Y_t - E(Y_t))(Y_{t-2} - E(Y_{t-2})) = E(Y_t Y_{t-2}) = \\ &= E[(\phi Y_{t-1} + a_t - \theta a_{t-1}) Y_{t-2}] = \\ &= \phi E(Y_{t-1} Y_{t-2}) + E(Y_{t-2} a_t) - \theta E(Y_{t-2} a_{t-1}) = \phi \gamma_1 \end{aligned}$$

Para derivar los resultados siguientes, hay que tener en cuenta que:

$$E(Y_{t-1} a_{t-1}) = E[(\phi Y_{t-2} + a_{t-1} - \theta a_{t-2}) a_{t-1}] = E(a_{t-1})^2 = \sigma^2$$

La función de autocovarianzas de un $ARMA(1,1)$ es:

$$\gamma_k = \begin{cases} \gamma_0 = \frac{(1 + \theta^2 - 2\phi\theta)\sigma^2}{1 - \phi^2} & k = 0 \\ \gamma_1 = \phi\gamma_0 - \theta\sigma^2 & k = 1 \\ \gamma_k = \phi\gamma_{k-1} & k > 1 \end{cases}$$

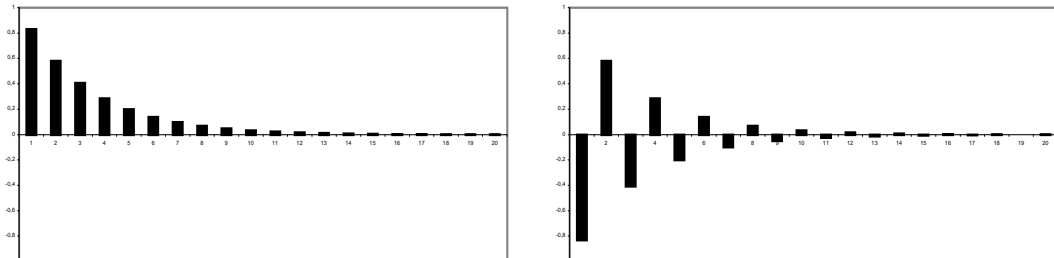
Como se puede observar, la varianza cuenta con una parte que proviene de la parte de medias móviles, otra que proviene de la parte autorregresiva y una tercera que es la interacción entre ambas partes del modelo. La autocovarianza de orden 1, γ_1 , es la suma de la autocovarianza de orden 1 de la parte $AR(1)$ y de la autocovarianza de orden 1 de la parte $MA(1)$. A partir del retardo 1, la parte medias móviles del modelo no aparece de forma explícita en la FACV, dependiendo ésta sólo de la estructura autorregresiva. Esta FACV es una función infinita, que depende de los parámetros AR , ϕ , y MA , θ , hasta $k = 1$ y luego decrece exponencialmente, siguiendo la estructura marcada por la parte autorregresiva de primer orden.

La función de autocorrelación de un $ARMA(1,1)$ es:

$$\rho_k = \begin{cases} \rho_1 = \phi - \frac{\theta\sigma^2}{\gamma_0} & k = 0 \\ \rho_k = \phi\rho_{k-1} & k > 1 \end{cases}$$

La FAC presenta la misma estructura que la función de autocovarianzas, es decir, es una función infinita cuyo primer coeficiente, ρ_1 , depende de los parámetros autorregresivos y de medias móviles, pero a partir del retardo 2, decrece exponencialmente, siguiendo la estructura dada por la parte autorregresiva de orden 1. En el gráfico 3.5, se observan dos ejemplos de funciones de autocorrelación de modelos $ARMA(1,1)$ para diferentes valores de los parámetros. Aunque las FAC que aquí se muestran son como las de un $AR(1)$, las estructuras posibles para un $ARMA(1,1)$ son más variadas, debido a la contribución de la parte medias móviles del modelo en la determinación del primer coeficiente.

Gráfico 3.5: Procesos Autorregresivo de Medias Móviles: $ARMA(1,1)$



Los resultados obtenidos para el modelo $ARMA$ más sencillo, el $ARMA(1,1)$, se pueden generalizar para el modelo $ARMA(p,q)$ y concluir que su función de autocorrelación será infinita,

con los q primeros coeficientes, ρ_1, \dots, ρ_q , dependiendo de los parámetros autorregresivos y de medias móviles. A partir del retardo q , que es el orden de la parte media móvil del modelo, sus parámetros no aparecen de forma explícita en la FAC y está decrece rápidamente hacia cero siguiendo la estructura marcada por la parte autorregresiva de orden p , es decir, como funciones exponenciales si todas las raíces del polinomio autorregresivo, $\phi_p(L)$, son reales, o en forma onda seno-coseno, si este polinomio tiene raíces complejas.

El modelo $ARMA(p, q)$ (3.12) se puede generalizar para que su media no sea nula. Basta con añadir una constante al proceso estacionario:

$$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad a_t \sim RB(0, \sigma^2)$$

Este modelo tiene las mismas características dinámicas que el modelo (3.12), la única diferencia es que su media no es cero:

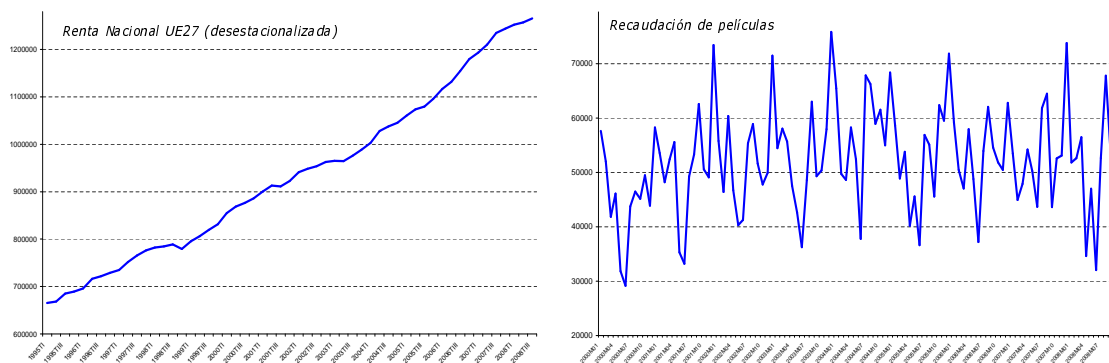
$$\begin{aligned} E(Y_t) &= E(\delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}) \\ \rightarrow E(Y_t) &= \delta + \phi_1 E(Y_t) + \dots + \phi_p E(Y_t) \\ \rightarrow (1 - \phi_1 - \dots - \phi_p) E(Y_t) &= \delta \\ \rightarrow E(Y_t) &= \frac{\delta}{1 - \phi_1 - \dots - \phi_p} \end{aligned}$$

Capítulo 4

Modelos lineales no estacionarios

Los modelos presentados en el capítulo 3 se basan en el supuesto de estacionariedad en covarianza, es decir, en procesos donde la media y la varianza son constantes y finitas y las autocovarianzas no dependen del tiempo sino sólo del número de periodos de separación entre las variables. Pero la mayoría de las series económicas no se comportan de forma estacionaria, bien porque suelen ir cambiando de nivel en el tiempo (serie Renta Nacional) o porque la varianza no es constante (serie Recaudación de películas).

Gráfico 4.1: Series temporales económicas



4.1. No estacionariedad en varianza

Cuando una serie no es estacionaria en varianza, es decir, no se puede sostener el supuesto de que ha sido generada por un proceso con varianza constante en el tiempo, la solución es transformar la serie mediante algún método que estabilice la varianza. El comportamiento habitual en las series económicas suele ser que la varianza cambie conforme el nivel de la serie cambia. En estos casos, suponemos que la varianza del proceso la podemos expresar como alguna función del nivel:

$$V(Y_t) = k f(\mu_t)$$

siendo k una constante positiva y f una función conocida. El objetivo es conseguir alguna función que transforme la serie de forma que $h(Y_t)$ tenga varianza constante. Utilizando la

extensión de Taylor de primer orden alrededor de μ_t :

$$h(Y_t) \simeq h(\mu_t) + (Y_t - \mu_t) h'(\mu_t)$$

donde $h'(\mu_t)$ es la primer derivada de $h(Y_t)$ evaluada en μ_t . La varianza de la transformación $h(Y_t)$ se puede aproximar como sigue:

$$\begin{aligned} V[h(Y_t)] &\simeq V[h(\mu_t) + (Y_t - \mu_t) h'(\mu_t)] = \\ &= [h'(\mu_t)]^2 V(Y_t) = [h'(\mu_t)]^2 k f(\mu_t) \end{aligned}$$

Por lo tanto, para que la transformación $h(Y_t)$ tenga varianza constante, la función h se ha de elegir de tal forma que:

$$h'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}}$$

Si la desviación típica de la serie Y_t es proporcional a su nivel, se tiene que:

$$f(\mu_t) = \mu_t^2 \quad \longrightarrow \quad h(\mu_t) \text{ ha de cumplir que } h'(\mu_t) = \frac{1}{\mu_t} \quad \Longrightarrow \quad h(\mu_t) = \ln(\mu_t)$$

Por lo tanto, la transformación logarítmica proporcionaría una varianza constante. Ahora bien, si es la varianza de la serie Y_t la que es proporcional a su nivel:

$$f(\mu_t) = \mu_t \quad \longrightarrow \quad h(\mu_t) \text{ ha de cumplir que } h'(\mu_t) = \frac{1}{\sqrt{\mu_t}} \quad \Longrightarrow \quad h(\mu_t) = \sqrt{\mu_t}$$

y habría que tomar la raíz cuadrada de la serie para obtener una varianza constante.

En general, para estabilizar la varianza se utilizan las transformaciones Box-Cox:

$$Y_t^{(\lambda)} = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln(Y_t) & \lambda = 0 \end{cases}$$

donde λ es el parámetro de transformación. Es interesante señalar que, usualmente, las transformaciones Box-Cox no sólo estabilizan la varianza sino que también mejoran la aproximación a la distribución normal del proceso Y_t .

4.2. No estacionariedad en media.

Una de las características dominantes en las series económicas es la tendencia. La tendencia es el movimiento a largo plazo de la serie una vez eliminados los ciclos y el término irregular. En economía esta tendencia se suele producir debido a la evolución de las preferencias, la tecnología, de la demografía, etc. Este comportamiento tendencial puede ser creciente o decreciente, exponencial o aproximadamente lineal. Las series que presentan un comportamiento sistemático de este tipo (veáse la figura izquierda del gráfico 4.1) no son estacionarias, no evolucionan en torno a un nivel constante.

La no estacionariedad en media se puede modelar de diferentes maneras. Por un lado, es posible modelar tendencias, cambios sistemáticos de nivel, mediante modelos globales en los que se especifica la tendencia como una función del tiempo:

$$Y_t = T_t + u_t$$

donde $T_t = f(t)$, es una función determinista del tiempo (lineal, cuadrática, exponencial, ...) y u_t es un proceso estocástico estacionario con media cero. Por ejemplo, el modelo con tendencia lineal sería:

$$Y_t = a + b_t + u_t$$

Estos modelos para la tendencia suponen que la serie evoluciona de una forma perfectamente predecible, y se denominan *modelos de tendencia determinista*.

Por otro lado, si un proceso no es estacionario en media también se puede modelar dentro de la clase de modelos $ARMA(p, q)$. De hecho, un modelo $ARMA(p, q)$ no es estacionario si las raíces de su polinomio AR no satisfacen la condición de estacionariedad, es decir, si alguna de sus raíces no está fuera del círculo unidad:

- a) El módulo de alguna raíz está dentro del círculo unidad: $\exists i \mid |L_i| < 1$. El comportamiento de las series generadas por estos procesos es explosivo, crecen o decrecen a gran velocidad hacia infinito. Esta no es la evolución temporal que se suele observar en series económicas por lo que este tipo de no estacionariedad no es de interés en nuestro campo.
- b) El módulo de alguna raíz es exactamente igual a la unidad: $\exists i \mid |L_i| = 1$. Este tipo de modelos genera comportamientos no estacionarios que no son explosivos, sino que las realizaciones se comportan de forma similar a lo largo del tiempo, salvo porque cambian de nivel. Este comportamiento sí es el que observamos en series económicas, por lo que modelaremos series no estacionarias mediante modelos $ARMA(p, q)$ no estacionarios porque tienen al menos una raíz exactamente igual a la unidad, denominada *raíz unitaria*.

4.3. Modelos ARIMA(p,d,q).

Supongamos el siguiente modelo $ARMA(p, q)$:

$$\Phi_p(L) Y_t = \Theta_q(L) a_t \quad (4.1)$$

donde el polinomio AR se puede factorizar en función de sus p raíces L_1, L_2, \dots, L_p ,

$$\Phi_p(L) = (1 - L_1^{-1} L) (1 - L_2^{-1} L) \dots (1 - L_p^{-1} L)$$

Supongamos que $(p - 1)$ raíces son estacionarias (con módulo fuera del círculo unidad) y una de ellas es unitaria, $L_i = 1$. Entonces, el polinomio AR , se puede reescribir como sigue:

$$\Phi_p(L) = (1 - L_1^{-1} L) (1 - L_2^{-1} L) \dots (1 - L_p^{-1} L) = \varphi_{p-1}(L) (1 - (1)^{-1} L)$$

$$\Phi_p(L) = \varphi_{p-1}(L) (1 - L)$$

donde el polinomio $\varphi_{p-1}(L)$ resulta del producto de los $(p-1)$ polinomios de orden 1 asociados a las raíces L_i con módulo fuera del círculo unidad.

Sustituyendo en el modelo $ARMA(p, q)$ (4.1) se tiene que:

$$\varphi_{p-1}(L) (1 - L) Y_t = \Theta_q(L) a_t \quad \rightarrow \quad \varphi_{p-1}(L) \Delta Y_t = \Theta_q(L) a_t \quad (4.2)$$

donde el polinomio $\varphi_{p-1}(L)$ es estacionario porque todas sus raíces tienen modulo fuera del círculo unidad y el polinomio $\Delta = (1 - L)$ es el que recoge la raíz unitaria.

El modelo (4.2) representa el comportamiento de un proceso Y_t que no es estacionario porque tiene una raíz unitaria. A un proceso Y_t con estas características se le denomina *proceso integrado de orden 1*.

En general, el polinomio AR del modelo (4.1) puede contener más de una raíz unitaria, por ejemplo, d , entonces se puede descomponer como:

$$\Phi_p(L) = \varphi_{p-d}(L) (1 - L)^d$$

y sustituyendo, de nuevo, en el modelo $ARMA(p, q)$ (4.1), se tiene:

$$\varphi_{p-d}(L) \Delta^d Y_t = \Theta_q(L) a_t$$

donde el polinomio $\varphi_{p-d}(L)$ es estacionario porque sus $(p-d)$ raíces tienen modulo fuera del círculo unidad, y el polinomio $\Delta^d = (1 - L)^d$, de orden d , contiene las d raíces unitarias no estacionarias. A un proceso Y_t con estas características se le denomina *proceso integrado de orden d* y se denota por $Y_t \sim I(d)$.

Definición: Un proceso Y_t es integrado de orden d , $Y_t \sim I(d)$, si Y_t no es estacionario, pero su diferencia de orden d , $\Delta^d Y_t$, sigue un proceso $ARMA(p-d, q)$ estacionario e invertible.

El orden de integración del proceso es el número de diferencias que hay que tomar al proceso para conseguir la estacionariedad en media, o lo que es lo mismo, el número de raíces unitarias del proceso. En la práctica, los procesos que surgen más habitualmente en el análisis de las series temporales económicas son los $I(0)$ e $I(1)$, encontrándose los $I(2)$ con mucha menos frecuencia.

En general, si una serie Y_t es integrada de orden d , se puede representar por el siguiente modelo:

$$\Phi_p(L) \Delta^d Y_t = \delta + \Theta_q(L) a_t \quad (4.3)$$

donde el polinomio autorregresivo estacionario $\Phi_p(L)$ y el invertible de medias móviles $\Theta_q(L)$ no tienen raíces comunes.

El modelo (4.3) se denomina modelo Autorregresivo Integrado de Medias Móviles de orden (p, d, q) o $ARIMA(p, d, q)$, donde p es el orden del polinomio autorregresivo *estacionario*, d es el *orden de integración* de la serie, es decir, el número de diferencias que hay que tomar a la serie para que sea estacionaria, y q es el orden del polinomio de medias móviles *invertible*.

Para comprender mejor el comportamiento representado por los modelos $ARIMA(p, d, q)$ vamos a estudiar detalladamente dos de los modelos más sencillos.

4.3.1. Modelo de paseo aleatorio

El modelo de paseo aleatorio es simplemente un modelo $AR(1)$ con parámetro $\phi = 1$:

$$\begin{aligned} Y_t &= Y_{t-1} + a_t & a_t &\sim RBN(0, \sigma^2) \\ \Delta Y_t &= a_t \end{aligned} \quad (4.4)$$

En el modelo de paseo aleatorio el valor de Y en el momento t es igual a su valor en el momento $t - 1$ más una perturbación aleatoria. El modelo de paseo aleatorio no es estacionario porque la raíz del polinomio AR no tiene módulo mayor que la unidad:

$$1 - L = 0 \quad \longrightarrow \quad L = 1$$

Pero, sin embargo, como la primera diferencia de la serie, ΔY_t es un ruido blanco, se tiene que $Y_t \sim I(1)$. En la figura 4.2 se muestra el gráfico de una realización simulada de 150 observaciones de un proceso de paseo aleatorio con $\sigma^2 = 1$. Se puede observar que la evolución de la serie no tiene una característica presente en los procesos estacionarios y que se denomina *reversión a la media*: la serie se mueve aleatoriamente hacia arriba y hacia abajo, sin mostrar ninguna inclinación a dirigirse a ningún punto en particular. Si una perturbación aumenta el valor de Y_t no hay ninguna tendencia a volver a disminuir otra vez, sino que en principio permanece en un valor alto.

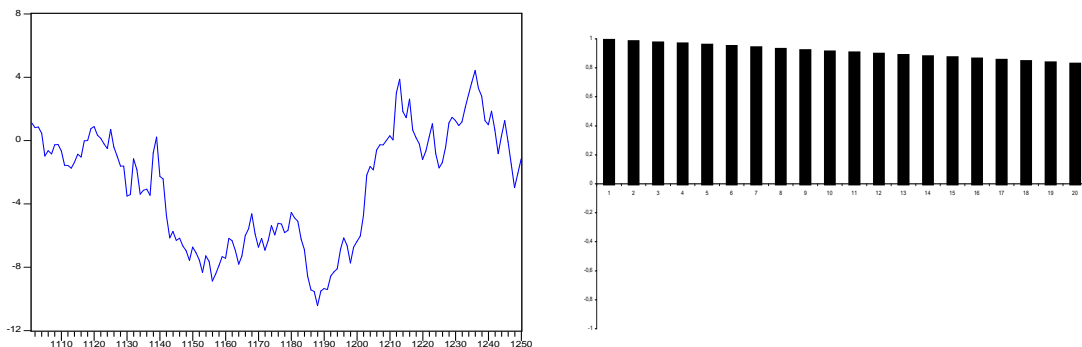
Tomando esperanzas condicionadas a la información anterior al modelo (4.4), se obtiene:

$$E[Y_t | Y_{t-1}, Y_{t-2}, \dots] = Y_{t-1}$$

lo que implica que el nivel condicionado de la serie en el momento t es aleatorio. Es decir, para este modelo el nivel promedio de la serie va cambiando de forma estocástica a lo largo del tiempo, por lo que se dice que este modelo tiene tendencia estocástica.

En lo que se refiere a la FAC del modelo (4.4), se puede comprobar que se caracteriza porque sus coeficientes decrecen muy lentamente (véase la figura derecha del gráfico 4.2).

Gráfico 4.2: Modelo de paseo aleatorio



4.3.2. Modelo de paseo aleatorio con deriva

El modelo de paseo aleatorio con deriva surge de añadir una constante al modelo de paseo aleatorio. Por lo tanto, al igual que para el modelo de paseo aleatorio, se tiene que $Y_t \sim I(1)$.

$$Y_t = Y_{t-1} + \delta + a_t \quad \rightarrow \quad \Delta Y_t = \delta + a_t \quad (4.5)$$

La única diferencia entre los modelos (4.4) y (4.5) es la presencia de la constante δ . Ahora bien, así como en los modelos estacionarios la inclusión de una constante solo afecta al nivel promedio constante de la serie, en los modelos no estacionarios su presencia en el modelo tiene efectos más relevantes. Suponiendo que el modelo (4.5) empieza en algún momento t_0 y, sustituyendo repetidamente, se obtiene:

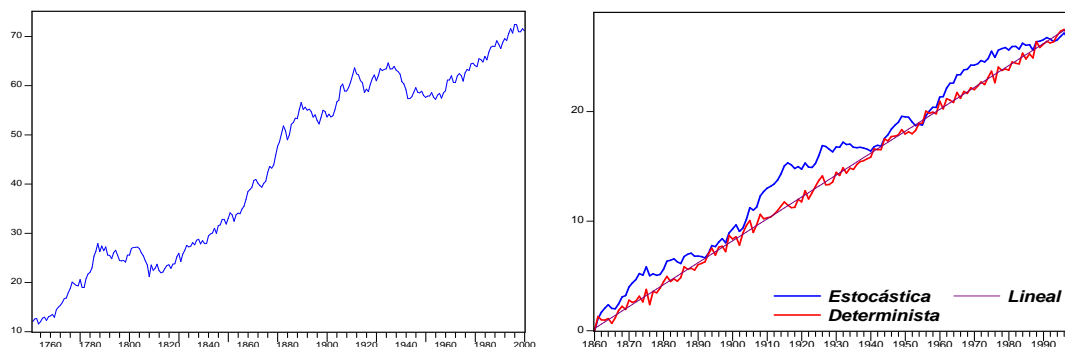
$$Y_t = Y_0 + (t - t_0)\delta + \sum_{i=t_0+1}^t a_i$$

Se puede observar que la introducción de la constante en el modelo implica la inclusión de una tendencia determinista con pendiente δ , junto con la tendencia estocástica. Tomando esperanzas condicionadas a la información pasada en el modelo (4.5), se obtiene:

$$E[Y_t | Y_{t-1}, Y_{t-2}, \dots] = Y_{t-1} + \delta$$

es decir, el nivel promedio de la serie se ve afectado por un elemento estocástico a lo largo del tiempo a través de Y_{t-1} así como por un componente determinista a través del parámetro δ .

Gráfico 4.3: Tendencias deterministas vs. Tendencias estocásticas



La figura izquierda del gráfico 4.3 muestra una realización de tamaño 150 simulada por el modelo de paseo aleatorio con deriva con parámetros $\delta = 0,2$ y $\sigma = 1$. El modelo (4.5) muestra un crecimiento promedio en cada periodo de δ . Este parámetro de deriva tiene el mismo papel que el parámetro de la pendiente en el modelo determinista lineal. Pero, de la misma manera que el modelo de paseo aleatorio no tiene un nivel promedio al que volver, el modelo de paseo aleatorio con deriva no tiene una tendencia lineal a la que retornar cuando un shock estacionario lo aleja de la misma. El gráfico 4.3 muestra en su figura de la derecha la diferente evolución que se puede observar entre tendencias estocásticas y deterministas. La serie con tendencia determinista oscila aleatoriamente en torno a una tendencia lineal de pendiente constante y, si se separa de ella, tiende a retornar rápidamente, mientras que si la serie con tendencia estocástica se separa de la línea de tendencia de pendiente δ no tiende a volver.

Capítulo 5

Modelización ARIMA

En este capítulo se van a desarrollar de forma detallada cada una de las etapas de las que consta la metodología de modelización ARIMA. Para apoyar esta explicación se va a modelar una serie clásica en la literatura *Pieles de lince* (1821-1930) y se va a proponer al lector como ejercicio la modelización de una serie artificial que se ha denominado *Producción de tabaco* (1872-1984). Los datos de ambas series se encuentran en el apéndice A.

Los gráficos y resultados de este análisis se han obtenido mediante el software econométrico Gretl. Este es un software que se puede obtener de forma gratuita de la siguiente dirección:

<http://gretl.sourceforge.net>

5.1. Estrategia de modelización ARIMA

En los capítulos 3 y 4 se han analizado las propiedades de los procesos $ARIMA(p, d, q)$, tanto estacionarios ($d = 0$) como no estacionarios ($d > 0$):

$$\Phi(L)(1 - L)^d Y_t = \delta + \Theta(L)a_t \quad (5.1)$$

Si se conocen los parámetros del modelo teórico (5.1) para el proceso Y_t , a partir de una realización concreta del ruido blanco, a_1, a_2, \dots, a_T , y de valores iniciales para Y y a se genera la serie temporal Y_1, Y_2, \dots, Y_T , que es una realización de tamaño T del proceso estocástico. La estructura es, por lo tanto,

$$\text{Proceso estocástico} \begin{cases} \Phi(L)(1 - L)^d Y_t = \delta + \Theta(L) a_t \\ a_t \sim RBN(0, \sigma^2) \end{cases}$$

↓

Generación

↓

Realización: Y_1, Y_2, \dots, Y_T

A partir de una misma estructura $ARIMA(p, d, q)$ se pueden obtener infinitas realizaciones. Para cada realización, la serie del ruido blanco, a_1, a_2, \dots, a_T , será diferente y la serie temporal generada, Y_1, Y_2, \dots, Y_T , también. Ahora bien, todas las series temporales generadas que proceden de una misma estructura conservarán entre sí una similitud de comportamiento dinámico.

En la aplicación de la metodología Box-Jenkins el punto de partida es el contrario: se conocen los valores de la serie temporal Y_1, Y_2, \dots, Y_T y se trata de determinar la estructura $ARIMA(p, d, q)$ que la ha podido generar.

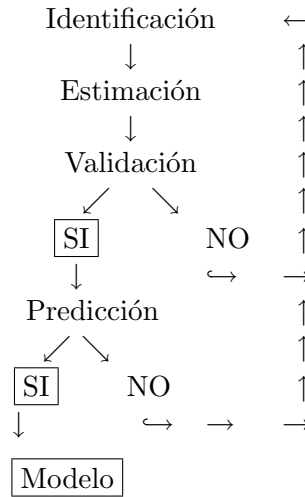
$$\begin{array}{c}
 \text{Realización: } Y_1, Y_2, \dots, Y_T \\
 \downarrow \\
 \text{Inferencia} \\
 \downarrow \\
 \text{Proceso estocástico: } \left\{ \begin{array}{l} \Phi(L)(1-L)^d Y_t = \delta + \Theta(L) a_t \\ a_t \sim RBN(0, \sigma^2) \end{array} \right.
 \end{array}$$

La construcción de modelos $ARIMA$ se lleva a cabo de forma iterativa mediante un proceso en el que se pueden distinguir cuatro etapas:

- a) **Identificación.** Utilizando los datos y/o cualquier tipo de información disponible sobre cómo ha sido generada la serie, se intentará sugerir una subclase de modelos $ARIMA(p, d, q)$ que merezca la pena ser investigada. El objetivo es determinar los órdenes p, d, q que parecen apropiados para reproducir las características de la serie bajo estudio y si se incluye o no la constante δ . En esta etapa es posible identificar más de un modelo candidato a haber podido generar la serie.
- b) **Estimación.** Usando de forma eficiente los datos se realiza inferencia sobre los parámetros condicionada a que el modelo investigado sea apropiado.
Dado un determinado proceso propuesto, se trata de cuantificar los parámetros del mismo, $\theta_1, \dots, \theta_q, \phi_1, \dots, \phi_p, \sigma^2$ y, en su caso, δ .
- c) **Validación.** Se realizan contrastes de diagnóstico para comprobar si el modelo se ajusta a los datos, o, si no es así, revelar las posibles discrepancias del modelo propuesto para poder mejorarlo.
- d) **Predicción.** Obtener pronósticos en términos probabilísticos de los valores futuros de la variable. En esta etapa se tratará también de evaluar la capacidad predictiva del modelo.

Esta metodología se basa, fundamentalmente, en dos principios:

- Selección de un modelo en *forma iterativa*. En cada etapa se plantea la posibilidad de rehacer las etapas previas.



- Principio de parametrización escueta, también denominado parsimonia. Se trata de proponer un modelo capaz de representar la serie con el mínimo de parámetros posibles y únicamente acudir a una ampliación del mismo en caso de que sea estrictamente necesario para describir el comportamiento de la serie.

5.2. Identificación

El objetivo de esta etapa de la modelización es seleccionar el modelo $ARIMA(p, d, q)$ apropiado para la serie, es decir, que reproduce las características de la serie. La identificación del modelo se lleva a cabo en dos fases:

- Análisis de estacionariedad, en el que se determinan las transformaciones que son necesarias aplicar para obtener una serie estacionaria. Incluye, a su vez, dos apartados:
 - Estacionariedad en varianza: transformaciones estabilizadoras de varianza.
 - Estacionariedad en media: número de diferencias d que hay que tomar para lograr que la serie sea estacionaria en media.
- Elección de los órdenes p y q . Una vez obtenida la serie estacionaria, el objetivo es determinar el proceso estacionario $ARMA(p, q)$ que la haya generado.

Los instrumentos que se van a utilizar en estas dos fases de la identificación del modelo son fundamentalmente los siguientes:

- Gráfico y correlogramas muestrales de la serie original.

- Gráfico y correlogramas muestrales de determinadas transformaciones de la serie: logaritmos, diferencias,...
- Contrastes de raíces unitarias.

5.2.1. Análisis de estacionariedad

Estacionariedad en varianza.

Una serie será *estacionaria en varianza* cuando pueda mantenerse el supuesto de que existe una única varianza para toda la serie temporal, es decir, cuando la variabilidad de la serie en torno a su media se mantenga constante a lo largo del tiempo. Si la serie no es estacionaria en varianza, se utilizan las transformaciones estabilizadoras de varianza, es decir, las transformaciones Box-Cox,

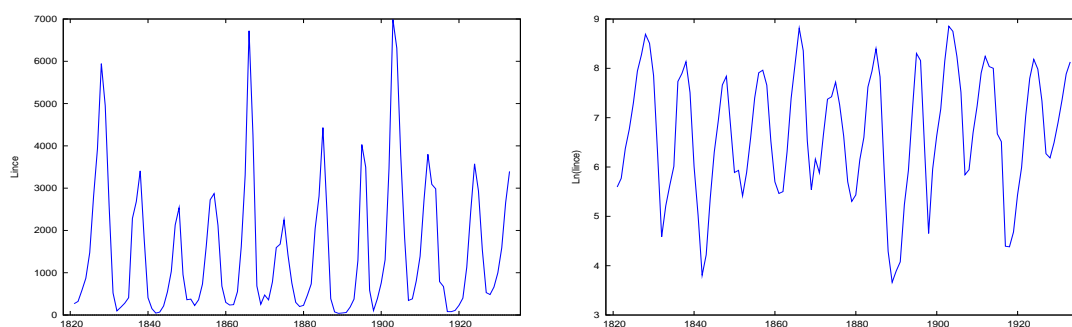
$$Y_t^{(\lambda)} = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln Y_t & \lambda = 0 \end{cases}$$

Las transformaciones Box-Cox incluye una familia infinita de funciones: raíz cuadrada, inversa, etc. Como las series económicas suelen ser positivas y sin valores cero, la transformación más utilizada en la práctica económica es la logarítmica.

Los instrumentos que se utilizan para analizar la estacionariedad en varianza de una serie son el gráfico de la serie original y de las transformaciones correspondientes.

Ejemplo 6.1. Pieles de Lince

Gráfico 5.1: Pieles de lince (1821-1930)

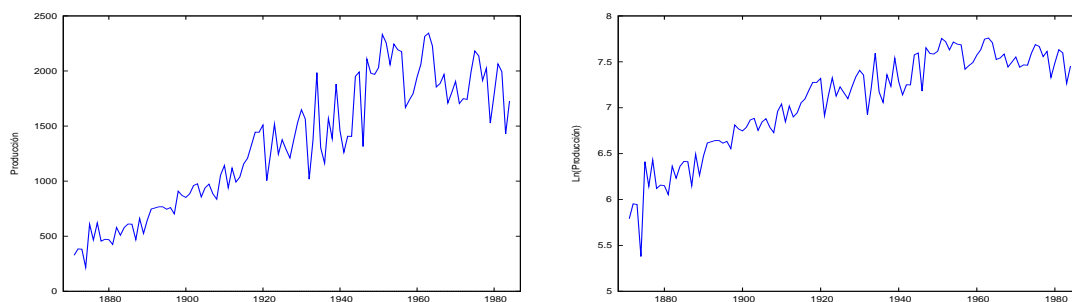


En el gráfico 5.1 se presenta la serie *Pieles de lince* capturadas de 1821 a 1930. En la figura izquierda del gráfico se puede observar que la serie en niveles presenta un comportamiento cíclico en el que la amplitud de los ciclos es variable en el tiempo. Por lo tanto, no parece adecuado suponer que la serie es estacionaria en varianza. Tomando logaritmos a la serie (la transformación Box-Cox más utilizada para las series económicas) se observa que la amplitud de estos ciclos se ha homogeneizado y que la variabilidad de la serie es mucho más estable en el tiempo (la figura derecha del gráfico). Se puede concluir que la serie estacionaria en varianza es $Z_t = \ln(\text{lince})$.

Ejercicio 6.1. *Producción de tabaco.*

En el gráfico 5.2 se muestra la serie de *Producción de tabaco* desde 1872 a 1984 tanto en niveles (figura izquierda) como en logaritmos (figura derecha).

Gráfico 5.2: Producción de tabaco (1872-1984)



Pregunta: *¿Cuál es la serie estacionaria en varianza?. Razona tu respuesta.*

Estacionariedad en media.

En este apartado, se ha de identificar si la serie es estacionaria en media, es decir, si oscila en torno a un nivel constante o no. Para tomar esta decisión nos basaremos en las características que diferencian las series estacionarias de las no estacionarias.

Características de una *serie estacionaria*:

- Una serie es estacionaria en media cuando se puede mantener el supuesto de que existe una única media para toda la serie temporal, es decir, cuando fluctúa alrededor de una media constante.
- La función de autocorrelación teórica de un proceso estacionario en media decae rápidamente, exponencialmente.

Características de una *serie no estacionaria*:

- La serie no es estacionaria en media cuando presenta tendencia o varios tramos con medias diferentes.
- En general, un proceso con alguna raíz unitaria presenta una función de autocorrelación muestral con un decaimiento muy lento, no siendo necesario que se mantenga próximo a la unidad.

Como se ha visto en el capítulo 4, si la serie no es estacionaria en media se puede lograr la estacionariedad transformándola tomando diferencias. Así, si la serie no es estacionaria en media, se tomarán d sucesivas diferencias de orden 1 sobre la serie hasta obtener una serie estacionaria:

$$Z_t = (1 - L)^d Y_t$$

Ahora bien, el problema es cómo identificar el número de diferencias exacto que es preciso tomar a la serie para que sea estacionaria ya que pueden aparecer los siguientes problemas:

- Raíz autorregresiva próxima a la unidad. Consideremos el siguiente modelo $AR(1)$

$$Y_t = 0,95 Y_{t-1} + a_t \quad \rightarrow \quad 1 - 0,95L = 0 \quad \rightarrow \quad L = \frac{1}{0,95} = 1,05 > 1$$

que es un proceso estacionario. Sin embargo, en la práctica puede resultar muy difícil distinguir una realización de este proceso de una serie generada por un paseo aleatorio, $Y_t = Y_{t-1} + a_t$, de raíz unitaria porque está muy próximo a él.

- Sobrediferenciación: se dice que la serie está sobrediferenciada si se toman más diferencias de las necesarias, por ejemplo, si se elige un orden de integración d cuando la serie $\Delta^{d-1} Y_t$ ya es estacionaria.

En este punto conviene recordar que si se diferencia un proceso estacionario sigue siendo estacionario. Por lo tanto, en principio, el hecho de que $\Delta^d Y_t$ sea estacionario, no significa necesariamente que $\Delta^{d-1} Y_t$ no lo sea, por lo que hay tener cuidado ya que el objetivo en esta fase de la modelización *ARIMA* es determinar el menor número de diferencias d capaz de convertir a una serie en estacionaria.

Sabemos que para las series económicas los valores de d más habituales son $d = 0, 1, 2$ y para decidir cuál es el más apropiado para la serie bajo estudio utilizaremos los siguientes instrumentos:

- a) Gráfico de la serie original y las transformaciones correspondientes, para observar si se cumple o no la condición de estacionariedad de oscilar en torno a un nivel constante.
- b) Correlograma estimado de la serie original y de las transformaciones correspondientes, para comprobar si decrece rápidamente hacia cero o no.
- c) Contrastes de raíces unitarias.

Los contrastes de raíces unitarias proporcionan unos contrastes estadísticos que permiten, a partir del conjunto de información, hacer inferencia sobre la existencia o no de una raíz unitaria en una serie, es decir, sobre la no estacionariedad de la serie. Si se rechaza la hipótesis nula de existencia de raíz unitaria en $\Delta^{d-1} Y_t$, no se diferenciará más la serie. En caso contrario, si no se rechaza la hipótesis nula se tomará una diferencia más de orden 1.

A continuación, se desarrollan brevemente, los contrastes de raíces unitarias más utilizados en la práctica del análisis de series temporales económicas y que, a la vez, fueron los pioneros en este campo.

Contraste de Dickey-Fuller. Supongamos el siguiente proceso $AR(1)$:

$$Y_t = \phi Y_{t-1} + a_t, \quad a_t \sim RBN(0, \sigma^2) \quad (5.2)$$

que es estacionario si $|\phi| < 1$. Si, por el contrario, $\phi = 1$ el proceso Y_t no es estacionario porque tiene una raíz unitaria, de hecho sería un modelo de paseo aleatorio:

$$Y_t = Y_{t-1} + a_t \quad \rightarrow \quad 1 - L = 0 \quad \rightarrow \quad L = 1 \quad \rightarrow \quad Y_t \sim I(1)$$

Si se desea contrastar alguna hipótesis nula sobre el valor del parámetro ϕ en el modelo (5.2), por ejemplo, $H_0 : \phi = \phi_0$, el procedimiento habitual es estimar la regresión de Y_t sobre Y_{t-1} y después utilizar el estadístico t estándar para contrastar la hipótesis nula. Como es bien sabido, si $|\phi| < 1$, el modelo $AR(1)$ es estacionario y el estimador de ϕ por Mínimos Cuadrados Ordinarios, $\hat{\phi}$, es igual al estimador Máximo Verosímil y sigue asintóticamente una distribución normal. Además, el estadístico t tiene la siguiente distribución asintótica bajo la hipótesis nula:

$$t_\phi = \frac{\hat{\phi} - \phi_0}{\sqrt{V_{\hat{\phi}}}} \sim N(0, 1)$$

Para muestras pequeñas este estadístico se distribuye aproximadamente como un t de Student de $(T - 1)$ grados de libertad.

Por lo tanto, se podría pensar en contrastar la hipótesis de existencia de raíz unitaria en la serie Y_t (no estacionariedad) contrastando la hipótesis nula $H_0 : \phi = 1$ en el modelo (5.2). El problema surge porque no se sabe *a priori* si el modelo es estacionario o no y cuando $\phi = 1$ el modelo $AR(1)$ no es estacionario y los resultados anteriores no se mantienen. Se puede demostrar que el estimador $\hat{\phi}$ está sesgado hacia abajo y que el estadístico t bajo la hipótesis nula de raíz unitaria no sigue una distribución t de Student ni en el límite cuando el tamaño muestral tiende a infinito. Además si se utilizan los valores críticos de esta distribución se tiende a rechazar la H_0 más veces de las deseadas.

El modelo (5.2) se puede escribir como sigue restando Y_{t-1} en ambos lados,

$$\begin{aligned} Y_t - Y_{t-1} &= \phi Y_{t-1} - Y_{t-1} + a_t && \text{o bien} \\ \Delta Y_t &= \beta Y_{t-1} + a_t && \text{con } \beta = \phi - 1 \end{aligned} \quad (5.3)$$

Ahora, contrastar la hipótesis nula de raíz unitaria (paseo aleatorio) equivale a plantear el siguiente contraste de hipótesis en el modelo (5.3):

$$\begin{aligned} H_0 : \beta &= 0 \\ H_a : \beta &< 0 \end{aligned}$$

La hipótesis alternativa de estacionariedad se plantea únicamente en términos de β negativo, porque un valor positivo de β supondría un modelo no estacionario de comportamiento explosivo. El modelo (5.3) es un modelo de regresión y el estadístico habitual para realizar este tipo de contraste es:

$$t = \frac{\hat{\beta} - 0}{\sqrt{\hat{V}(\hat{\beta})}} \quad (5.4)$$

donde:

$$\hat{\beta} = \frac{\sum_{t=2}^T y_{t-1} \Delta y_t}{\sum_{t=2}^T y_{t-1}^2} \quad \text{y} \quad \hat{V}(\hat{\beta}) = \frac{\hat{\sigma}^2}{\sum_{t=2}^T y_{t-1}^2} \quad (5.5)$$

Al estadístico t se le denomina **estadístico de Dickey-Fuller**. Este estadístico t no sigue ninguna distribución conocida bajo la hipótesis nula de no estacionariedad por lo que Dickey-Fuller calcularon los percentiles de este estadístico bajo la H_0 proporcionando las tablas con los niveles críticos correctos para el estadístico en función del tamaño muestral (T) y el nivel de significación (α) (las tablas de estos valores críticos se encuentran en el apéndice B.).

Se rechaza H_0 , es decir, la existencia de raíz unitaria a un nivel de significación α si el valor muestral del estadístico t es menor que el valor crítico de las tablas de Dickey-Fuller (DF_α):

$$t < DF_\alpha$$

Hasta el momento se ha explicado como contrastar la hipótesis nula de raíz unitaria (paseo aleatorio) frente a la alternativa de un proceso estacionario $AR(1)$ de media cero. Para las series económicas, sería de interés considerar hipótesis alternativas más generales que puedan incluir estacionariedad alrededor de una constante o de una tendencia lineal.

Si el modelo $AR(1)$ incluye una **constante**:

$$Y_t = c + \phi Y_{t-1} + a_t, \quad a_t \sim RBN(0, \sigma^2) \quad (5.6)$$

el contraste de existencia de raíz unitaria se realiza de forma similar. Restando Y_{t-1} en ambos lados del proceso se obtiene:

$$\begin{aligned} \Delta Y_t &= c + \beta Y_{t-1} + a_t \\ \beta &= \phi - 1 \end{aligned} \quad (5.7)$$

El contraste de la hipótesis nula de raíz unitaria se plantea en este modelo (5.7) como sigue:

$$\begin{aligned} H_0 &: \beta = 0 \\ H_a &: \beta < 0 \end{aligned}$$

y el estadístico de contraste es:

$$t_c = \frac{\hat{\beta}}{\sqrt{\hat{V}(\hat{\beta})}}$$

donde $\hat{\beta}$ es el estimador MCO de β y $\hat{V}(\hat{\beta})$ el de su varianza. Como anteriormente, Dickey-Fuller construyeron las tablas que proporcionan los niveles críticos correctos para el estadístico t_c en función del tamaño muestral, T , y el nivel de significación, α .

Si el modelo $AR(1)$ incluye una **tendencia lineal**:

$$Y_t = c + \gamma t + \phi Y_{t-1} + a_t, \quad a_t \sim RBN(0, \sigma^2)$$

el modelo que se estima por MCO es:

$$\Delta Y_t = c + \gamma t + \beta Y_{t-1} + a_t \quad (5.8)$$

y se utiliza el estadístico t que denotaremos por t_γ para contrastar la hipótesis nula de raíz unitaria $H_0 : \beta = 0$ frente $H_a : \beta < 0$.

Contraste de Dickey-Fuller aumentado. Los contrastes anteriores de Dickey-Fuller se han derivado partiendo del supuesto restrictivo de que la serie sigue un proceso $AR(1)$. Pero la serie puede seguir procesos más generales $ARMA(p, q)$. Como es bien conocido, todo proceso $ARMA(p, q)$ se puede aproximar hasta el grado de bondad necesario mediante un $AR(p)$:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t \quad a_t \sim RBN(0, \sigma^2)$$

Este modelo se puede reparametrizar como sigue:

$$\Delta Y_t = \beta Y_{t-1} + \alpha_1 \Delta Y_{t-1} + \dots + \alpha_{p-1} \Delta Y_{t-p+1} + a_t \quad (5.9)$$

donde

$$\beta = \sum_{i=1}^p \phi_i - 1 \quad \text{y} \quad \alpha_i = \sum_{j=1}^i \phi_{p-i+j}$$

Dado que un proceso $AR(p)$ tiene una raíz unitaria cuando $\sum_{i=1}^p \phi_i = 1$, contrastar la hipótesis nula de existencia de raíz unitaria es equivalente a contrastar la $H_0 : \beta = 0$ en la regresión (5.9).

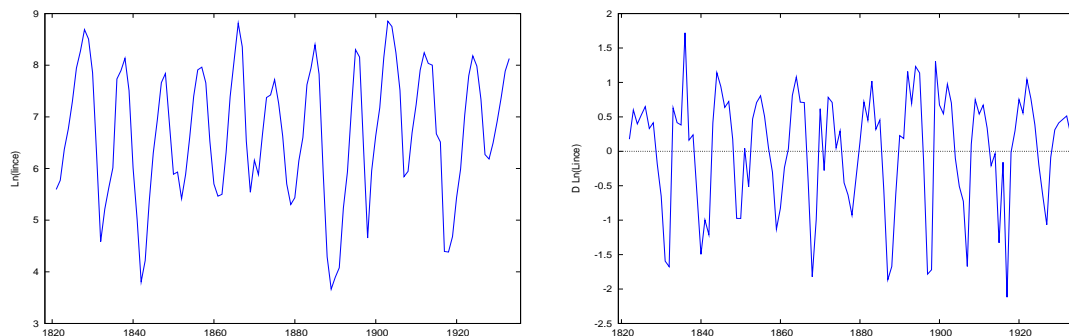
Este contraste de raíz unitaria se denomina **Dickey-Fuller Aumentado (ADF)** y se basa en la estimación MCO del parámetro β en el modelo (5.9) y en el correspondiente estadístico t . Este estadístico tiene la misma distribución que para el caso del modelo $AR(1)$ y, por lo tanto, podemos utilizar los mismos valores críticos tabulados por Dickey-Fuller.

El modelo (5.9) se puede también generalizar introduciendo una constante y/o una tendencia lineal:

$$\begin{aligned} \Delta Y_t &= \delta + \beta Y_{t-1} + \alpha_1 \Delta Y_{t-1} + \dots + \alpha_{p-1} \Delta Y_{t-p+1} + a_t \\ \Delta Y_t &= \delta + \gamma t + \beta Y_{t-1} + \alpha_1 \Delta Y_{t-1} + \dots + \alpha_{p-1} \Delta Y_{t-p+1} + a_t \end{aligned}$$

procediéndose a la realización del contraste de la hipótesis nula de raíz unitaria por un procedimiento similar al desarrollado para el modelo $AR(1)$.

Gráfico 5.3: Gráficos de Ln(Lince) y su primera diferencia.



Cuadro 5.1: Contrates de raíces unitarias

```

Contrastes aumentados de Dickey-Fuller, orden 4, para Ln(lince):
tamaño muestral 108 hipótesis nula de raíz unitaria: a = 1
contraste con constante
modelo: (1 - L)y = b0 + (a-1)*y(-1) + ... + e
Coef. de autocorrelación de primer orden de e: -0,019
valor estimado de (a - 1): -0,451665
Estadístico de contraste: tau_c(1) = -5,13165
valor p asintótico 1,089e-005

```

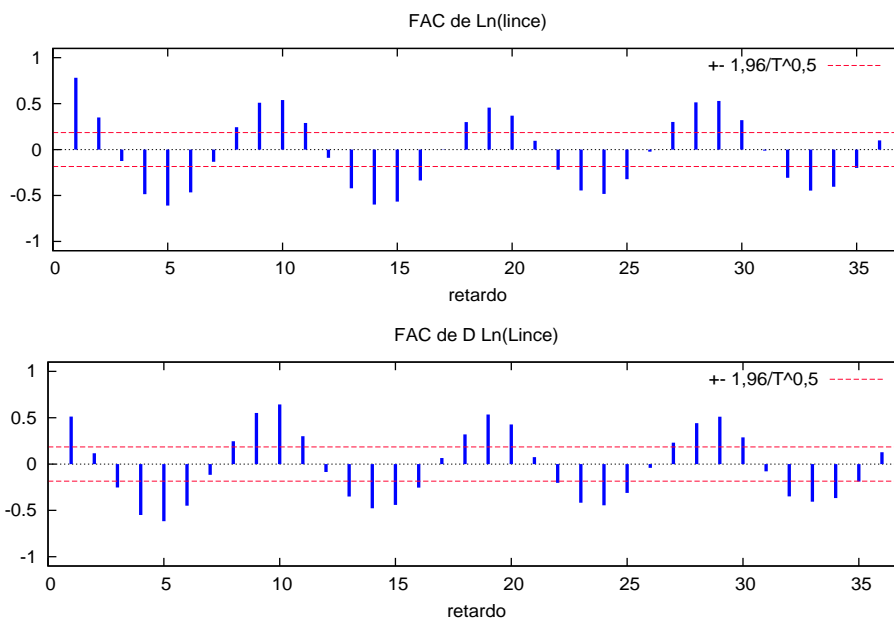
Ejemplo 6.2. *Pieles de lince.*

Analicemos la estacionariedad en media de la serie que es estacionaria en varianza, $Z_t = \ln(Y_t) = \ln(Lince)$.

En el gráfico 5.3 se representan la serie en logaritmos ($\ln(Lince)$) y su primera diferencia ($D \ln(Lince)$). Se puede observar que la serie Z_t oscila cíclicamente en torno a un nivel promedio más o menos constante y al diferenciarla no cambia su comportamiento salvo que pasa a oscilar en torno a un nivel próximo a cero. Se puede concluir, por lo tanto, que Z_t es ya estacionaria en media y no es preciso realizar más transformaciones. Analizando los correlogramas de Z_t y su primera diferencia recogidos en el gráfico 5.4 se obtiene la misma conclusión ya que el correlograma de Z_t decrece rápidamente hacia cero en forma de onda seno-coseno amortiguada. El tomar una diferencia a la serie no cambia el comportamiento del correlograma, como era de esperar, ya que si diferenciamos un proceso estacionario sigue siendo estacionario.

Para contrastar la existencia de raíz unitaria en la serie $Z_t = \ln(Lince)$ se lleva a cabo el contraste ADF aumentado. Los resultados del cuadro 5.1 muestran que el valor del estadístico ADF, $t_c = -5,13$, es menor que el valor crítico $DF_{0,05} = -2,89$ (véase apéndice B). Por lo tanto, se rechaza la hipótesis nula de existencia de raíz unitaria a un nivel de significación del 5% y la transformación estacionaria en media y varianza de la serie *Pieles de Lince* es $Z_t = \ln(Y_t) = \ln(Lince)$.

Gráfico 5.4: Correlogramas de Ln(Lince) y su primera diferencia.

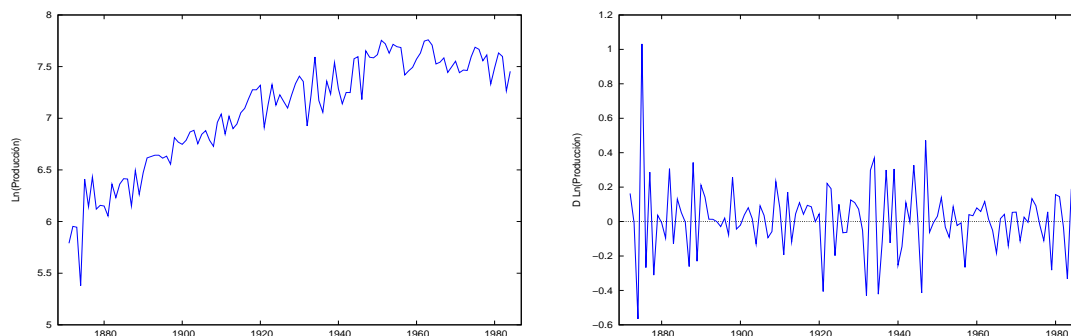


Ejercicio 6.2. *Producción de Tabaco.*

Considera la siguiente información sobre la serie *Producción de Tabaco*.

- a) Gráfico 5.5: muestra en la figura de la izquierda la serie Ln(Producción) y en la figura de la derecha la serie $D \ln(Producción) = \Delta \ln(Producción)$.
- b) Gráfico 5.6: muestra en la figura superior el correlograma de la serie Ln(Producción) y en la figura inferior el correlograma de la serie $D \ln(Producción) = \Delta \ln(Producción)$.
- c) Cuadros 5.2 y 5.3 muestran los resultados de los contrastes de raíces unitarias ADF para las series Ln(Producción) y $D \ln(Producción)$ respectivamente.

Gráfico 5.5: Gráficos de Ln(Producción) y su primera diferencia.



Pregunta: *Analiza la estacionariedad en media de la serie en logaritmos.*

Cuadro 5.2: Contrates de raíces unitarias

Contrastes aumentados de Dickey-Fuller, orden 4, para Ln(Producción):
tamaño muestral 109 hipótesis nula de raíz unitaria: $a = 1$

contraste con constante

modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,029

valor estimado de $(a - 1)$: -0,0776783

Estadístico de contraste: $\tau_{a,c}(1) = -2,61077$

valor p asintótico 0,0907

contraste con constante y tendencia

modelo: $(1 - L)y = b_0 + b_1*t + (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,020

valor estimado de $(a - 1)$: -0,127884

Estadístico de contraste: $\tau_{a,ct}(1) = -1,46933$

valor p asintótico 0,8402

Cuadro 5.3: Contrates de raíces unitarias

Contrastes aumentados de Dickey-Fuller, orden 4, para D Ln(Producción):
tamaño muestral 108 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante

modelo: $(1 - L)y = (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,019

valor estimado de $(a - 1)$: -2,37034

Estadístico de contraste: $\tau_{a,nc}(1) = -6,5554$

valor p asintótico 2,213e-010

contraste con constante

modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$

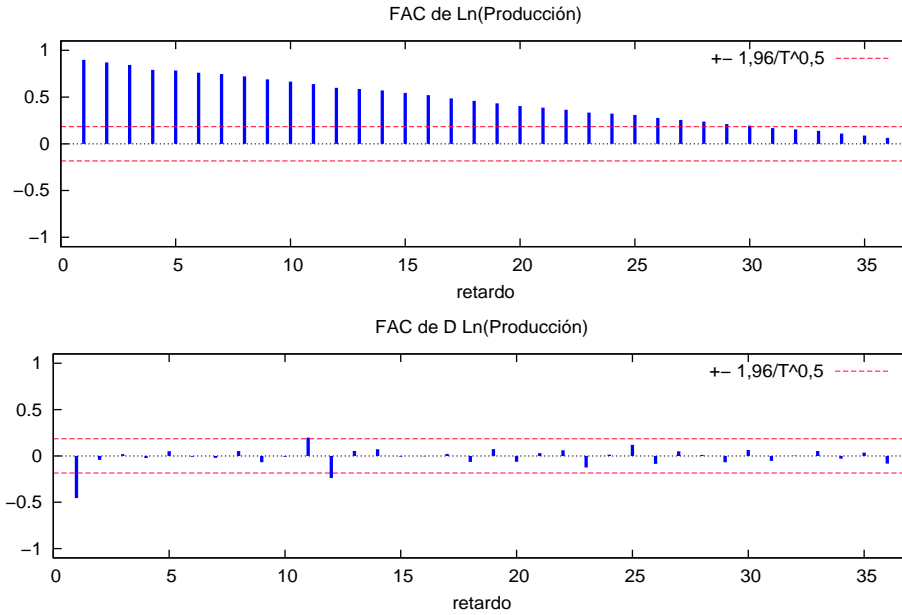
Coef. de autocorrelación de primer orden de e: -0,020

valor estimado de $(a - 1)$: -2,69183

Estadístico de contraste: $\tau_{a,c}(1) = -7,06954$

valor p asintótico 2,145e-010

Gráfico 5.6: Correlogramas de Ln(Producción) y su primera diferencia.



5.2.2. Identificación del modelo estacionario

Una vez determinado el orden de diferenciación d , se tiene la transformación estacionaria de la serie $Z_t = (1 - L)^d Y_t$ que puede representarse mediante un proceso $ARMA(p, q)$ estacionario. En esta fase se trata de identificar los órdenes p y q del proceso que puede reproducir las características de la serie estacionaria y de analizar la conveniencia de la incorporación del parámetro asociado a la media, δ .

A. Identificación de los órdenes p , q .

Las características dinámicas del proceso estacionario están recogidas en la función de autocorrelación, FAC, por lo que ésta será el instrumento básico para identificar los órdenes p y q del modelo ARMA adecuado para representar las características de la serie estacionaria Z_t .

Los coeficientes de autocorrelación muestral de Z_t son:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{T^*} (Z_t - \bar{Z})(Z_{t-k} - \bar{Z})}{\sum_{t=1}^{T^*} (Z_t - \bar{Z})^2} \quad k = 1, 2, \dots, T^*/3 \quad (5.10)$$

donde $T^* = T - d$ es la longitud de la serie estacionaria Z_t .

Para identificar los órdenes p y q , se compararán las funciones de autocorrelación muestrales con las FAC teóricas de los modelos ARMA cuyas características conocemos:

	FAC
$MA(q)$	Se anula para $j > q$
$AR(p)$	Decrecimiento rápido No se anula
$ARMA(p, q)$	Decrecimiento rápido No se anula

Si el correlograma muestral de la serie Z_t presenta un corte a partir de un retardo finito j , la identificación del proceso adecuado para la misma es sencilla, ya que se correspondería con la FAC teórica de un $MA(j)$. Pero si el correlograma muestral no presenta ningún corte sino que parece decrecer rápidamente siguiendo una estructura exponencial o de onda seno-coseno, la identificación no es tan clara, ya que basándose únicamente en la FAC podría corresponder a un modelo teórico AR o $ARMA$ de cualquier orden.

Para ayudarnos en la identificación de modelos $ARMA(p, q)$ en estos casos acudimos a otro instrumento complementario, la función de autocorrelación parcial.

El coeficiente de autocorrelación parcial de orden k , denotado por p_k , mide el grado de asociación lineal existente entre las variables Y_t e Y_{t-k} una vez ajustado el efecto lineal de todas las variables intermedias, es decir:

$$p_k = \rho_{Y_t Y_{t-k} \cdot Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}}$$

Por lo tanto, el coeficiente de autocorrelación parcial p_k es el coeficiente de la siguiente regresión lineal:

$$Y_t = \alpha + p_1 Y_{t-1} + p_2 Y_{t-2} + \dots + p_k Y_{t-k} + e_t$$

Las propiedades de la función de autocorrelación parcial (FACP), p_k , $k = 0, 1, 2, 3, \dots$, son equivalentes a las de la FAC:

- $p_0 = 1$ y $p_1 = \rho_1$
- Los coeficientes p_k no dependen de unidades y son menores que la unidad en valor absoluto.
- La FACP es una función simétrica.
- La FACP de un proceso estocástico estacionario decrece rápidamente hacia cero cuando $k \rightarrow \infty$.

La función de autocorrelación parcial se puede estimar a partir de los datos de la serie como una función de los coeficientes de autocorrelación simples estimados, $\hat{\rho}_k$.

La estructura de la función de autocorrelación parcial para un modelo estacionario $ARMA(p, q)$ es como sigue:

▷ Modelo $AR(p)$:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \dots + \phi_p Y_{t-p} + a_t$$

La estructura de la FACP es la siguiente:

$$p_k = \begin{cases} p_k \neq 0 & k = 1, 2, \dots, p \\ p_k = 0 & k = p + 1, p + 2, \dots \end{cases}$$

Si un proceso sigue un modelo $AR(p)$, Y_t depende directamente de su pasado hasta el retardo p de forma que las variables aleatorias separadas un periodo, dos, y hasta p periodos mantienen una relación lineal directa aunque se elimine el efecto de las variables intermedias, por lo que $p_k \neq 0$. Sin embargo, para variables separadas más de p periodos el coeficiente de autocorrelación parcial es cero porque no existe relación lineal directa entre ellas: existe relación lineal, $\rho_k \neq 0$, pero es indirecta a través de las variables intermedias por lo que, una vez controlado el efecto de estas, la correlación entre Y_t e Y_{t-k} es nula y $p_k = 0$.

▷ Modelo $MA(q)$:

$$Y_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} + \dots + \theta_q a_{t-q}$$

La FACP de un modelo de medias móviles finito es infinita y decrece rápidamente hacia cero, de forma exponencial cuando las raíces del polinomio de medias móviles son reales y en forma de onda seno-coseno amortiguada si las raíces del citado polinomio son complejas. Esta estructura es lógica dado que la representación AR de un modelo $MA(q)$ es infinita.

▷ Modelo $ARMA(p, q)$:

$$Y_t = a_t + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}$$

La FACP de un modelo $ARMA(p, q)$ es infinita y los p primeros coeficientes de la FACP dependen de los parámetros autorregresivos y de los de medias móviles y, a partir del retardo $p + 1$, depende únicamente de la estructura de la parte de medias móviles, de forma que, decrece rápidamente hacia cero, de forma exponencial cuando las raíces del polinomio de medias móviles son reales y en forma de onda seno-coseno amortiguada si las raíces del citado polinomio son complejas.

Comparando la estructura de las funciones de autocorrelación simple y parcial estimadas con las características básicas de las funciones de autocorrelación teóricas de la tabla siguiente, se puede identificar el / los procesos que podrían haber generado la serie bajo estudio.

	FAC	FACP
$MA(q)$	Se anula para $j > q$	Decrecimiento rápido No se anula
$AR(p)$	Decrecimiento rápido No se anula	Se anula para $j > p$
$ARMA(p, q)$	Decrecimiento rápido No se anula	Decrecimiento rápido No se anula

Ahora bien, para la identificación del modelo apropiado sólo se cuenta con las estimaciones de la FAC y de la FACP a partir de los datos de la serie Y_t . Los estimadores de los coeficientes de autocorrelación $\hat{\rho}_k$ y \hat{p}_k son variables aleatorias que tienen una distribución de probabilidad y tomarán diferentes valores para diferentes realizaciones de Y_t . Por lo tanto, para identificar el orden del proceso $ARMA(p, q)$ adecuado para la serie, es preciso determinar la estructura de las FAC y FACP estimadas realizando contrastes sobre la significación individual de los coeficientes de autocorrelación simple y parcial estimados.

El estimador de los coeficientes de autocorrelación $\hat{\rho}_k$ dado por (5.10) es una variable aleatoria que se distribuye asintóticamente como sigue bajo el supuesto de que Z_t es gaussiano y $k > 0$,

$$\hat{\rho}_k \stackrel{a}{\sim} N(\rho_k, V(\hat{\rho}_k)) \quad \text{con} \quad V(\hat{\rho}_k) \simeq \frac{1}{T^*}$$

Se desea realizar el siguiente contraste de hipótesis para $k = 1, 2, \dots, T^*/3$

$$H_0 : \rho_k = 0$$

$$H_a : \rho_k \neq 0$$

Bajo la hipótesis nula el estadístico de contraste se distribuye asintóticamente como

$$\frac{\hat{\rho}_k}{\sqrt{V(\hat{\rho}_k)}} \stackrel{a}{\sim} N(0, 1)$$

Se rechaza la H_0 a un nivel de significación del 5% si

$$\left| \frac{\hat{\rho}_k}{\sqrt{V(\hat{\rho}_k)}} \right| \geq N_{0,025} = 1,96$$

o análogamente,

$$|\hat{\rho}_k| \geq 1,96 \sqrt{\hat{\rho}_k} \simeq \frac{2}{\sqrt{T^*}}$$

Por lo tanto, $\pm \frac{2}{\sqrt{T^*}}$ son las bandas de confianza que delimitan la zona de significatividad del coeficiente ρ_k .

En lo que se refiere a la función de autocorrelación parcial, los coeficientes de autocorrelación parcial muestrales \hat{p}_k , con $k > p$, se distribuyen asintóticamente como sigue para procesos $AR(p)$:

$$\hat{p}_k \stackrel{a}{\sim} N(0, V(\hat{p}_k)) \quad \text{con} \quad V(\hat{p}_k) \simeq \frac{1}{T^*}$$

En la práctica, para obtener las bandas de confianza para \hat{p}_k se aplica esta distribución independientemente del tipo de proceso que haya generado la serie.

Se desea realizar el siguiente contraste de hipótesis para $k = 1, 2, \dots, T^*/3$

$$H_0 : p_k = 0$$

$$H_a : p_k \neq 0$$

Bajo la hipótesis nula el estadístico de contraste se distribuye asintóticamente como

$$\frac{\hat{p}_k}{\sqrt{V(\hat{p}_k)}} \stackrel{a}{\sim} N(0, 1)$$

Se rechaza la H_0 de no significatividad del coeficiente de autocorrelación a un nivel de significación del 5% si

$$\left| \frac{\hat{p}_k}{\sqrt{V(\hat{p}_k)}} \right| \geq N_{0,025} = 1,96 \quad \text{o análogamente,} \quad |\hat{p}_k| \geq 1,96 \sqrt{V(\hat{p}_k)} \simeq \frac{2}{\sqrt{T^*}}$$

Por lo tanto, $\pm \frac{2}{\sqrt{T^*}}$ son las bandas que delimitan la zona de significatividad del coeficiente p_k .

Cuando se procede a la identificación de un modelo $ARMA(p, q)$ hay que tener en cuenta que identificar el modelo a través de las funciones de autocorrelación simple y parcial estimadas no es una tarea sencilla. En esta primera etapa no se trata tanto de identificar *el* modelo correcto, sino de acotar un subconjunto de modelos ARIMA que han podido generar la serie. Posteriormente, en la fase de estimación y validación, dependiendo de los resultados, se vuelve a plantear la identificación del proceso.

A la hora de interpretar las funciones de autocorrelación estimadas no se debe dar demasiada importancia a cada detalle que pueda aparecer en los datos. En general, se trata de buscar los modelos más sencillos que reproduzcan las características de la serie. Es preciso tener en cuenta que, por un lado, hay cierta probabilidad de obtener algún coeficiente significativo aunque los datos fueran generados por un ruido blanco y, por otro lado, que al determinar qué retardos pueden ser significativamente distintos de cero, también hay que considerar la interpretación que se les pueda dar. Además, los coeficientes de autocorrelación muestral están correlacionados entre sí, por lo que en el correlograma muestral pueden aparecer ciertas ondulaciones que no se corresponden con el correlograma teórico. La tarea de identificación será más fácil cuanto mayor sea el tamaño de muestra.

Ejemplo 6.3. *Pieles de Lince.*

El gráfico 5.7 muestra las funciones de autocorrelación simple y parcial estimadas para la serie estacionaria $Z_t = \ln(Y_t) = \text{Ln}(\text{lince})$. La FAC estimada muestra una estructura infinita decreciendo en forma de onda seno-coseno amortiguada que sugiere un modelo $AR(p)$ o $ARMA(p, q)$ con parte autorregresiva de orden mayor o igual que 2. El análisis de la FACP muestra, según el resultado de los contrastes de significación individual, que los tres primeros coeficientes son significativamente distintos de cero y el resto no. Según esta interpretación, la FACP estaría truncada en el retardo 3 y el modelo adecuado para haber generado la serie sería un $AR(3)$. Ahora bien, estudiando el comportamiento general de la FACP, también se puede interpretar que tiene una estructura infinita que 2 decrece rápidamente a partir del retardo con forma cíclica. En este caso el modelo más sencillo que recogería esta estructura sería un $ARMA(2, 1)$.

Por lo tanto, se podrían identificar dos modelos adecuados, en principio, para la serie estacionaria:

$$a) \quad Z_t \sim AR(3) \qquad b) \quad Z_t \sim ARMA(2, 1)$$

Gráfico 5.7: FAC y FACP estimadas de la serie Ln(lince).

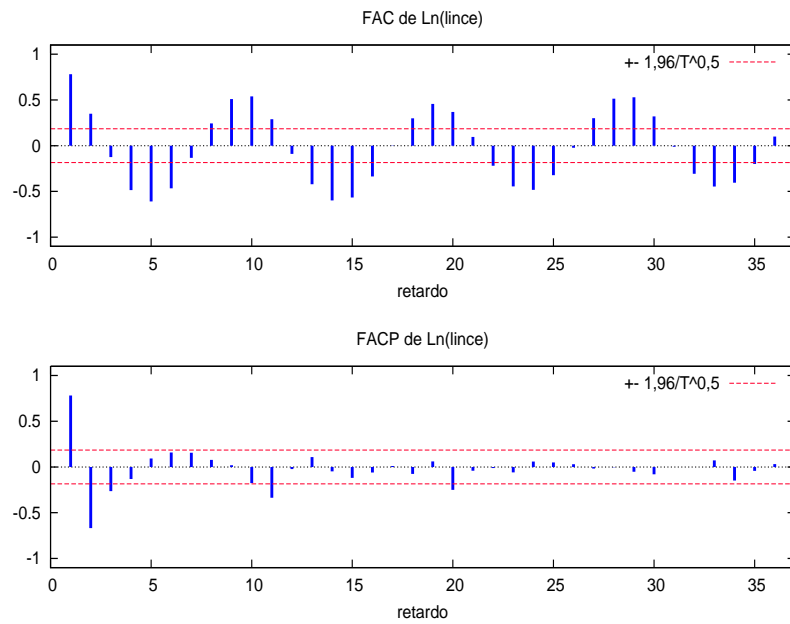
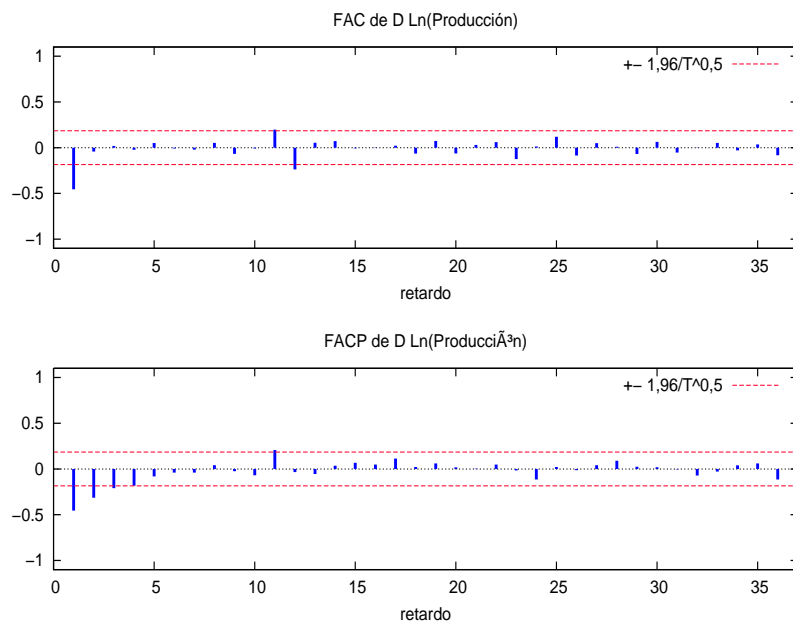


Gráfico 5.8: FAC y FACP estimadas de la serie (1-L) Ln(Produccion).



Ejercicio 6.3. *Producción de tabaco.*

La figura 5.8 muestra las funciones de autocorrelación simple y parcial estimadas para la transformación estacionaria de la serie *Producción de Tabaco*, $Z_t = \Delta \ln(Y_t) = D \ln(\text{Produccion})$.

Pregunta: *¿Qué modelo o modelos ARMA(p, q) propones para la serie estacionaria?*

B. Inclusión del término independiente

Recordemos que la media de un proceso $ARMA(p, q)$ estacionario está directamente relacionada con la constante δ . Si esta constante es cero, entonces la media del proceso es cero. Para decidir si se incluye un término independiente no nulo en el modelo, se contrastará si la media de la serie estacionaria es o no cero,

$$\begin{aligned} H_0: & E(Z_t) = 0 \\ H_a: & E(Z_t) \neq 0, \end{aligned}$$

El estadístico de contraste es:

$$t = \frac{\bar{Z}}{\hat{\sigma}_{\bar{Z}}} \stackrel{H_0}{\sim} t(T^* - 1)$$

donde $\hat{\sigma}_{\bar{Z}}^2$ es el estimador de la varianza de la media muestral \bar{Z} que viene dado por:

$$\hat{\sigma}_{\bar{Z}}^2 = \frac{C_0}{T^*} (1 + 2\hat{\rho}_1 + 2\hat{\rho}_2 + \dots + 2\hat{\rho}_n) \quad (5.11)$$

donde $C_0 = \sum (Z_t - \bar{Z})^2 / (T - 1)$ es la varianza muestral de la serie estacionaria y $(\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_n)$ representan las n primeras autocorrelaciones muestrales significativas de Z_t .

En ocasiones, para calcular esta varianza, se aplica la siguiente aproximación:

$$\hat{\sigma}_{\bar{Z}}^2 = \frac{C_0}{T^*}$$

Se rechazará la $H_0 : E(z_t) = 0$ a un nivel de significación α y, por lo tanto, se incluirá el parámetro δ en el modelo si:

$$t > t_{\alpha/2}(T^* - 1)$$

Ejemplo 6.4. *Pieles de Lince.*

La figura de la izquierda del gráfico 5.3 muestra la serie estacionaria $\ln(\text{lince})$ que parece oscilar en torno a una media diferente de cero. La siguiente tabla muestra los estadísticos descriptivos principales de la serie estacionaria $Z_t = \ln(Y_t)$ que se van a utilizar para contrastar si la media es o no nula.

Estadísticos principales, usando las observaciones 1821 - 1933
para la variable Ln(lince) (113 observaciones válidas)

Media	Mediana	Mínimo	Máximo
6,68671	6,66441	3,66356	8,85238
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
1,27356	0,190462	-0,374282	-0,668913

La distribución del estadístico de contraste bajo la hipótesis nula $H_0: E(Z_t) = 0$ es:

$$t = \frac{\bar{Z}}{\hat{\sigma}_{\bar{Z}}} \sim t(180 - 1)$$

Según los estadísticos principales, $\bar{Z} = 6,66$ y:

$$\hat{\sigma}_{\bar{Z}} = \frac{(1,276)^2}{180} (1 + 2 * 0,78 + 2 * 0,35 + \dots - 2 * 0,192) = 0,0095$$

por lo tanto,

$$t = \frac{6,66}{0,0095} = 701,05 > t_{0,025}(180 - 1) \approx 2$$

Se rechaza la $H_0: E(Z_t) = 0$ para un α del 5% y se incluye un término constante en el modelo.

Resumiendo, se proponen los dos modelos siguientes para la serie estacionaria:

- (1) $AR(3): Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + a_t$
- (2) $ARMA(2,1): Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t + \theta a_{t-1}$

Ejercicio 6.4. Serie *Producción de tabaco*.

En la figura derecha del gráfico 5.5 se presenta la transformación estacionaria de la serie *Producción de Tabaco*, $Z_t = \Delta \ln(Y_t)$ que parece oscilar en torno a cero y la siguiente tabla muestra los principales estadísticos descriptivos de la misma.

Estadísticos principales, usando las observaciones 1871 - 1984
para la variable D Ln(Produccion) (113 observaciones válidas)

Media	Mediana	Mínimo	Máximo
0,0147324	0,0199342	-0,565523	1,03192
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,203842	13,8363	0,711632	4,94746

Preguntas:

¿Es preciso incluir un término independiente en la especificación del modelo?

Escribe el/los modelos ARMA(p, q) que podrían haber generado la serie.

5.3. Estimación

Una vez identificados el/los procesos estocásticos que han podido generar la serie temporal Y_t , la siguiente etapa consiste en estimar los parámetros desconocidos de dichos modelos:

$$\beta = (\delta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)' \quad \text{y} \quad \sigma_a^2$$

Estos parámetros se pueden estimar de forma consistente por Mínimos Cuadrados o Máxima Verosimilitud. Ambos métodos de estimación se basan en el cálculo de las innovaciones, a_t , a partir de los valores de la serie estacionaria. El método de Mínimos Cuadrados minimiza la suma de cuadrados:

$$\text{Min} \sum_t a_t^2 \quad (5.12)$$

La función de verosimilitud se puede derivar a partir de la función de densidad conjunta de las innovaciones, a_1, a_2, \dots, a_T , que, bajo el supuesto de normalidad, es como sigue:

$$f(a_1, a_2, \dots, a_T) \propto \sigma^{-T^*} \exp \left\{ - \sum_{t=1}^{T^*} \frac{a_t^2}{2\sigma^2} \right\} \quad (5.13)$$

Para resolver el problema de estimación, las ecuaciones (5.12) y (5.13) se han de expresar en función del conjunto de información y de los parámetros desconocidos del modelo. Para un modelo $ARMA(p, q)$, la innovación se puede escribir como:

$$a_t = Z_t - \delta - \sum_{i=1}^p \phi_i Z_{t-i} - \sum_{i=1}^q \theta_i a_{t-i} \quad (5.14)$$

Por lo tanto, para calcular las innovaciones a partir de un conjunto de información y de un vector de parámetros desconocidos, se necesitan un conjunto de valores iniciales Z_0, Z_1, \dots, Z_{p-1} y a_0, a_1, \dots, a_{q-1} .

En la práctica, el procedimiento seguido es aproximar las innovaciones imponiendo una serie de condiciones sobre los valores iniciales, con lo que se obtienen los denominados estimadores de *Mínimos Cuadrados Condicionados* y *Máximo verosímiles Condicionados*. La condición que se impone habitualmente sobre los valores iniciales es que las p primeras observaciones de Z_t son los valores iniciales y las innovaciones previas son cero. En este caso, se calculan las innovaciones según la ecuación (5.14) de $t = p + 1$ en adelante. Los estimadores Máximo verosímiles condicionados a las p primeras observaciones son iguales a los estimadores Mínimos Cuadrados condicionados.

Existen métodos para obtener estimadores Máximos verosímiles no condicionados, basados en la maximización de la función de verosimilitud exacta que se obtiene como una combinación de la función de verosimilitud condicionada y la función de densidad no condicionada para los valores iniciales.

Ejemplo 6.5. *Pieles de lince.*

En las siguientes tablas se encuentran los resultados de la estimación de los dos modelos propuestos para la transformación estacionaria $Z_t = \ln(Y_t) = \ln(\text{lince})$.

Modelo 1: estimaciones ARMA utilizando las 113 observaciones 1821–1933

		Variable dependiente: Ln(lince)			
	Coefficiente	Desv. típica	estadístico t	valor p	
const	6,68527	0,114106	58,5883	0,0000	
ϕ_1	1,14741	0,0902207	12,7179	0,0000	
ϕ_2	-0,340902	0,137053	-2,4874	0,0129	
ϕ_3	-0,258269	0,0903199	-2,8595	0,0042	
Media de la var. dependiente			6,68671		
D.T. de la variable dependiente			1,27356		
media de las innovaciones			-0,00562629		
Varianza de las innovaciones			0,296613		
Log-verosimilitud			-92,887		

		Real	Imaginaria	Módulo	Frecuencia
AR	Raíz 1	0,9064	-0,6438	1,1117	-0,0983
	Raíz 2	0,9064	0,6438	1,1117	0,0983
	Raíz 3	-3,1327	0,0000	3,1327	0,5000

Modelo 2: estimaciones ARMA utilizando las 113 observaciones 1821–1933

		Variable dependiente: Ln(lince)			
	Coefficiente	Desv. típica	estadístico t	valor p	
const	6,68440	0,106330	62,8649	0,0000	
ϕ_1	1,49195	0,0648607	23,0023	0,0000	
ϕ_2	-0,817975	0,0578566	-14,1380	0,0000	
θ_1	-0,328985	0,0982730	-3,3477	0,0008	
Media de la var. dependiente			6,68671		
D.T. de la variable dependiente			1,27356		
media de las innovaciones			-0,00587872		
Varianza de las innovaciones			0,296583		
Log-verosimilitud			-92,876		

		Real	Imaginaria	Módulo	Frecuencia
AR	Raíz 1	0,9120	-0,6252	1,1057	-0,0956
	Raíz 2	0,9120	0,6252	1,1057	0,0956
MA	Raíz 1	3,0397	0,0000	3,0397	0,0000

Los modelos estimados son:

$$(1) \quad AR(3): \quad Z_t = \hat{\delta} + \underset{(0,10)}{1,14} Z_{t-1} - \underset{(0,14)}{0,34} Z_{t-2} - \underset{(0,09)}{0,26} Z_{t-3} + \hat{a}_t$$

$$(2) \quad ARMA(2,1): \quad Z_t = \hat{\delta} + \underset{(0,07)}{1,50} Z_{t-1} - \underset{(0,08)}{0,82} Z_{t-2} + \hat{a}_t - \underset{(0,13)}{0,34} \hat{a}_{t-1}$$

La estimación de la constante dada por **Gretl** es la media estimada de la serie estacionaria:

- si el modelo estimado es un $MA(q)$: $\widehat{const} = \hat{C} = \hat{\mu} = \hat{\delta}$
- si el modelo estimado es un $AR(p)$ o $ARMA(p, q)$:

$$\widehat{const} = \hat{C} = \hat{\mu} = \frac{\hat{\delta}}{1 - \hat{\phi}_1 - \hat{\phi}_2 - \dots - \hat{\phi}_p}$$

Para la serie de *Pieles de lince*:

$$(1) \quad \hat{C} = \hat{\mu} = 6,68 = \frac{\hat{\delta}}{1 - \hat{\phi}_1 - \hat{\phi}_2 - \hat{\phi}_3} = \frac{\hat{\delta}}{1 - 1,14 + 0,34 + 0,26} \rightarrow \hat{\delta} = 6,68 \times 0,46 = 3,07$$

$$(2) \quad \hat{C} = \hat{\mu} = 6,68 = \frac{\hat{\delta}}{1 - \hat{\phi}_1 - \hat{\phi}_2} = \frac{\hat{\delta}}{1 - 1,50 + 0,82} \rightarrow \hat{\delta} = 6,68 \times 0,32 = 2,13$$

Por último, la estimación de la varianza del ruido blanco, σ^2 , viene dada por:

$$\hat{\sigma}^2 = \frac{\sum \hat{a}^2}{T^* - p - q}$$

Por lo tanto,

$$(1) \quad \hat{\sigma}^2 = \frac{\sum \hat{a}^2}{T^* - p - q} = \frac{32,88223}{110 - 3 - 0} = 0,30731$$

$$(2) \quad \hat{\sigma}^2 = \frac{\sum \hat{a}^2}{T^* - p - q} = \frac{32,88891}{110 - 2 - 1} = 0,30737$$

Por lo que los modelos estimados para la serie de *Pieles de lince* son:

$$(1) \quad Z_t = 3,07 + \underset{(0,10)}{1,14} Z_{t-1} - \underset{(0,14)}{0,34} Z_{t-2} - \underset{(0,09)}{0,26} Z_{t-3} + \hat{a}_t \quad \sigma^2 = 0,30731$$

$$(2) \quad Z_t = 2,13 + \underset{(0,07)}{1,50} Z_{t-1} - \underset{(0,08)}{0,82} Z_{t-2} + \hat{a}_t - \underset{(0,13)}{0,34} \hat{a}_{t-1} \quad \sigma^2 = 0,30737$$

Ejercicio 6.5. *Producción de tabaco.*

En las siguientes tablas se encuentran los resultados de la estimación de varios modelos para la transformación estacionaria, $Z_t = \Delta \ln(Y_t) = D\ln(\text{Produccion})$, de la serie *Producción de Tabaco*.

Pregunta: *Escribe los correspondientes modelos estimados.*

Modelo 1: estimaciones ARIMA utilizando las 113 observaciones 1872–1984

Variable dependiente: $(1 - L) \ln(\text{Produccion})$

	Coefficiente	Desv. típica	estadístico t	valor p
const	0,0142529	0,00443753	3,2119	0,0013
θ_1	-0,719692	0,0613815	-11,7249	0,0000

Media de la var. dependiente	0,0147324
D.T. de la variable dependiente	0,203842
media de las innovaciones	0,00227119
Varianza de las innovaciones	0,0270321
Log-verosimilitud	43,3014

	Real	Imaginaria	Módulo	Frecuencia
MA Raíz 1	1,3895	0,0000	1,3895	0,0000

Modelo 2: estimaciones ARIMA utilizando las 113 observaciones 1872–1984

Variable dependiente: $(1 - L) \ln(\text{Produccion})$

	Coefficiente	Desv. típica	estadístico t	valor p
const	0,0142527	0,00444005	3,2100	0,0013
θ_1	-0,720531	0,0991205	-7,2692	0,0000
θ_2	0,00102293	0,0948359	0,0108	0,9914

Media de la var. dependiente	0,0147324
D.T. de la variable dependiente	0,203842
media de las innovaciones	0,00226524
Varianza de las innovaciones	0,0270320
Log-verosimilitud	43,3014

	Real	Imaginaria	Módulo	Frecuencia
MA Raíz 1	1,3906	0,0000	1,3906	0,0000
Raíz 2	702,9871	0,0000	702,9871	0,0000

5.4. Validación del modelo

En la etapa de validación, diagnosis o chequeo se procede a evaluar la adecuación de los modelos estimados a los datos. Se tiene en cuenta:

- a) Si las estimaciones de los coeficientes del modelo son significativas y cumplen las condiciones de estacionariedad e invertibilidad que deben satisfacer los parámetros del modelo.
- b) Si los residuos del modelo tienen un comportamiento similar a las innovaciones, es decir, si son ruido blanco.

5.4.1. Análisis de coeficientes estimados

En primer lugar, se han de realizar los contrastes habituales de significación individual de los coeficientes *AR* y *MA*:

$$\beta = (\delta, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \dots, \theta_q)$$

para comprobar si el modelo propuesto está sobreidentificado, es decir, si se ha incluido estructura que no es relevante.

En el caso más general de un *ARMA*(p, q) con constante se plantean los siguientes contrastes:

$$H_0 : \delta = 0 \quad \text{frente a} \quad H_a : \delta \neq 0 \quad (5.15)$$

$$H_0 : \phi_i = 0 \quad \text{frente a} \quad H_a : \phi_i \neq 0 \quad (5.16)$$

$$H_0 : \theta_i = 0 \quad \text{frente a} \quad H_a : \theta_i \neq 0 \quad (5.17)$$

En general, la distribución asintótica de los estimadores es la siguiente:

$$\hat{\beta}_i \sim N(\beta_i, V(\hat{\beta}_i)) \quad \forall i$$

con la varianza dada por la inversa de la matriz de información. De forma que, para contrastar la H_0 de no significatividad individual de los parámetros se utilizará el estadístico t habitual que sigue asintóticamente una distribución normal:

$$t = \frac{\hat{\beta}_i - 0}{\sqrt{V(\hat{\beta}_i)}} \sim N(0, 1)$$

Se rechazará la hipótesis nula a un nivel de significación $\alpha = 5\%$, cuando:

$$\left| \frac{\hat{\beta}_i}{\sqrt{V(\hat{\beta}_i)}} \right| > N_{\alpha/2}(0, 1) \simeq 2$$

Por otro lado, es importante comprobar si las condiciones de estacionariedad e invertibilidad se satisfacen para el modelo propuesto. Para ello, se calculan las raíces del polinomio autorregresivo,

$\hat{\phi}(L) = 0$, y las raíces del polinomio de medias móviles, $\hat{\theta}(L) = 0$. Si alguna de estas raíces está próxima a la unidad podría indicar falta de estacionariedad y/o invertibilidad.

Por último, es conveniente también analizar la matriz de covarianzas entre los parámetros estimados con el fin de detectar la posible presencia de una correlación excesivamente alta entre las estimaciones de los mismos lo que puede ser una indicación de la presencia de *factores comunes* en el modelo.

Ejemplo 6.6. Piel de lince.

Para realizar los contrastes de diagnóstico sobre los coeficientes, se analizan los resultados de la estimación de los modelos 1 y 2 y las matrices de covarianzas siguientes.

Matriz de covarianzas de los coeficientes (Modelo AR(3))

const	phi_1	phi_2	phi_3	
0,0130202	2,48347e-05	-4,80732e-05	5,56654e-05	const
	0,00813977	-0,0108657	0,00573862	phi_1
		0,0187835	-0,0109509	phi_2
			0,00815768	phi_3

Matriz de covarianzas de los coeficientes (Modelo ARMA(2,1))

const	phi_1	phi_2	theta_1	
0,0113060	-6,16782e-05	5,83053e-05	0,000117743	const
	0,00420691	-0,00318008	-0,00379432	phi_1
		0,00334739	0,00256617	phi_2
			0,00965758	theta_1

Modelo AR(3). Análisis de los coeficientes.

- Contrastes de significatividad individual:

$$\begin{aligned} H_0: C = 0 & & H_0: \phi_i = 0 \\ H_a: C \neq 0 & , & H_a: \phi_i \neq 0 , \quad i = 1, 2, 3 \end{aligned}$$

$$C: \left| \frac{\hat{C}}{\hat{\sigma}_{\hat{C}}} \right| = \left| \frac{6,68}{0,12} \right| = |55,87| > N_{0,025} \approx 2$$

Se rechaza $H_0 : C = 0$ para un nivel de significación del 5 %.

$$\phi_1: \left| \frac{\hat{\phi}_1}{\hat{\sigma}_{\hat{\phi}_1}} \right| = \left| \frac{1,14}{0,10} \right| = |12,03| > N_{0,025} \approx 2$$

Se rechaza $H_0 : \phi_1 = 0$ para un nivel de significación del 5 %.

$$\phi_2: \left| \frac{\hat{\phi}_2}{\hat{\sigma}_{\hat{\phi}_2}} \right| = \left| \frac{-0,34}{0,14} \right| = |-2,35| > N_{0,025} \approx 2$$

Se rechaza $H_0 : \phi_2 = 0$ para un nivel de significación del 5 %.

$$\phi_3: \left| \frac{\hat{\phi}_3}{\hat{\sigma}_{\hat{\phi}_3}} \right| = \left| \frac{-0,26}{0,09} \right| = |-2,78| > N_{0,025} \approx 2$$

Se rechaza $H_0 : \phi_3 = 0$ para un nivel de significación del 5 %.

- Estacionariedad del modelo $AR(3)$.

Los módulos de las raíces del polinomio autorregresivo cumplen la condición de estacionariedad:

$$|L_1| = \left| \frac{1}{-0,32} \right| > 1 \quad |L_2| = |L_3| = \left| \frac{1}{\sqrt{0,73^2 + (-0,52)^2}} \right| = \left| \frac{1}{0,89} \right| > 1$$

- Las correlaciones entre los coeficientes estimados que muestran la matriz de covarianzas no son muy altas:

$$\text{Corr}(\hat{\phi}_1, \hat{\phi}_2) = \frac{\text{Cov}(\hat{\phi}_1, \hat{\phi}_2)}{\sqrt{V(\hat{\phi}_1)V(\hat{\phi}_2)}} = \frac{-0,01}{\sqrt{0,009 * 0,02}} = -0,745$$

$$\text{Corr}(\hat{\phi}_1, \hat{\phi}_3) = \frac{\text{Cov}(\hat{\phi}_1, \hat{\phi}_3)}{\sqrt{V(\hat{\phi}_1)V(\hat{\phi}_3)}} = \frac{0,006}{\sqrt{0,009 * 0,009}} = 0,666$$

$$\text{Corr}(\hat{\phi}_2, \hat{\phi}_3) = \frac{\text{Cov}(\hat{\phi}_2, \hat{\phi}_3)}{\sqrt{V(\hat{\phi}_2)V(\hat{\phi}_3)}} = \frac{-0,01}{\sqrt{0,02 * 0,009}} = -0,745$$

Modelo ARMA(2,1). Análisis de los coeficientes.

- Contrastes de significatividad individual:

$$\begin{array}{lll} H_0: C = 0 & H_0: \phi_i = 0 & H_0: \theta = 0 \\ H_a: C \neq 0 & H_a: \phi_i \neq 0, \quad i = 1, 2 & H_a: \theta \neq 0 \end{array}$$

$$C: \left| \frac{\hat{C}}{\hat{\sigma}_{\hat{C}}} \right| = \left| \frac{6,67}{0,11} \right| = |60,87| > N_{0,025} \approx 2$$

Se rechaza $H_0 : C = 0$ para un nivel de significación del 5 %.

$$\phi_1: \left| \frac{\hat{\phi}_1}{\hat{\sigma}_{\hat{\phi}_1}} \right| = \left| \frac{1,50}{0,07} \right| = |20,62| > N_{0,025} \approx 2$$

Se rechaza $H_0 : \phi_1 = 0$ para un nivel de significación del 5 %.

$$\phi_2: \left| \frac{\hat{\phi}_2}{\hat{\sigma}_{\hat{\phi}_2}} \right| = \left| \frac{-0,82}{0,06} \right| = |-12,87| > N_{0,025} \approx 2$$

Se rechaza $H_0 : \phi_2 = 0$ para un nivel de significación del 5 %.

$$\theta: \left| \frac{\hat{\theta}}{\hat{\sigma}_{\hat{\theta}}} \right| = \left| \frac{-0,34}{0,13} \right| = |-2,70| > N_{0,025} \approx 2$$

Se rechaza $H_0 : \theta = 0$ para un nivel de significación del 5 %.

- Estacionariedad e invertibilidad del modelo $ARMA(2, 1)$.

Los módulos de las raíces del polinomio autorregresivo cumplen la condición de estacionariedad:

$$|L_1| = |L_2| = \left| \frac{1}{\sqrt{0,75^2 + (-0,51)^2}} \right| = \left| \frac{1}{0,91} \right| > 1$$

El módulo de la raíz del polinomio medias móviles cumple la condición de invertibilidad:

$$|L| = \left| \frac{1}{0,34} \right| > 1$$

- Las correlaciones entre los coeficientes estimados que se recogen en la matriz de covarianzas no son muy altas.

$$\text{Corr}(\hat{\phi}_1, \hat{\phi}_2) = \frac{\text{Cov}(\hat{\phi}_1, \hat{\phi}_2)}{\sqrt{V(\hat{\phi}_1)V(\hat{\phi}_2)}} = \frac{-0,004}{\sqrt{0,005 * 0,004}} = -0,894$$

$$\text{Corr}(\hat{\phi}_1, \hat{\theta}) = \frac{\text{Cov}(\hat{\phi}_1, \hat{\theta})}{\sqrt{V(\hat{\phi}_1)V(\hat{\theta})}} = \frac{-0,006}{\sqrt{0,005 * 0,016}} = -0,671$$

$$\text{Corr}(\hat{\phi}_2, \hat{\theta}) = \frac{\text{Cov}(\hat{\phi}_2, \hat{\theta})}{\sqrt{V(\hat{\phi}_2)V(\hat{\theta})}} = \frac{0,004}{\sqrt{0,004 * 0,016}} = 0,5$$

5.4.2. Análisis de residuos.

Si el modelo $ARMA(p, q)$ elegido para la serie estacionaria Z_t (por simplicidad, suponemos que $EZ_t = 0$)

$$\Phi_p(L) Z_t = \Theta_q(L) a_t \quad (5.18)$$

es correcto, entonces $a_t = \frac{\Phi_p(L)}{\Theta_q(L)} Z_t$ es un proceso ruido blanco. Los residuos del modelo estimado

$$\hat{a}_t = \frac{\hat{\Phi}_p(L)}{\hat{\Theta}_q(L)} Z_t \quad (5.19)$$

son estimaciones de a_t . El análisis de residuos consiste en una serie de contrastes de diagnóstico con el objetivo de determinar si los residuos replican el comportamiento de un ruido blanco, es decir, si su media es cero, su varianza constante y las autocorrelaciones nulas.

- a) Para comprobar si la *media es cero*, se realiza un análisis gráfico, representando los residuos a lo largo del tiempo y observando si los valores oscilan alrededor de cero.

Además se puede llevar a cabo el siguiente contraste de hipótesis:

$$\begin{aligned} H_0: & E(a_t) = 0 \\ H_a: & E(a_t) \neq 0, \end{aligned}$$

El estadístico de contraste se distribuye bajo la hipótesis nula como sigue:

$$t = \sqrt{T} \frac{\bar{\hat{a}}}{\sqrt{C_0(\hat{a})}} \stackrel{a}{\sim} N(0, 1) \quad (5.20)$$

donde $\bar{\hat{a}}$ y $C_0(\hat{a})$ son, respectivamente, la media y la varianza muestrales de los residuos. No se rechaza la hipótesis nula al nivel de significación del 5% si $|t| \leq 1,96$.

- b) *Varianza constante*. Si en el gráfico de los residuos la dispersión de los mismos es constante, concluiremos que la varianza de a_t permanece constante.

- c) *Ausencia de correlación serial*.

Si los residuos se comportaran como un ruido blanco, los coeficientes de la FAC y FACP muestrales deben ser prácticamente nulos para todos los retardos. Para comprobarlo, se pueden llevar a cabo:

- Contrastos de significatividad *individual* sobre los coeficientes de autocorrelación:

$$\begin{cases} H_0 : \rho_k(a) = 0 \\ H_a : \rho_k(a) \neq 0 \end{cases}$$

Bajo la hipótesis nula, se tiene que

$$\hat{\rho}_k(a) \stackrel{a}{\sim} N\left(0, \frac{1}{T^*}\right)$$

Se dirá que a_t es un ruido blanco si los coeficientes de autocorrelación estimados están dentro del intervalo de no significación, es decir, si

$$|\hat{\rho}_k(a)| \leq \frac{2}{\sqrt{T^*}}$$

para todo k o al menos para el 95 % de los coeficientes estimados.

- Otra alternativa es realizar un contraste de significatividad sobre un *conjunto* de coeficientes de autocorrelación, siendo la hipótesis nula,

$$H_0 : \rho_1(a) = \rho_2(a) = \dots = \rho_M(a) = 0$$

y la hipótesis alternativa que algún coeficiente ρ_k sea no nulo, para $k = 1, \dots, M$.

El estadístico más utilizado para contrastar esta hipótesis es el propuesto por Ljung-Box(1978):

$$Q^*(M) = T^*(T^* + 2) \sum_{k=1}^M \frac{\hat{\rho}_k^2(a)}{T^* - k} \stackrel{H_0, a}{\sim} \chi^2(M - p - q) \quad (5.21)$$

Se rechaza la hipótesis nula de ausencia de correlación serial, para un nivel de significación del 5 %, para valores grandes del estadístico, es decir, si

$$Q^*(M) > \chi_{0,05}^2(M - p - q)$$

En este caso, existiría correlación serial en los residuos del modelo por lo que se concluye que el modelo no ha sido capaz de reproducir el patrón de comportamiento sistemático de la serie y habría que reformularlo.

- d) *Contraste de normalidad.* El estadístico más utilizado es el de Jarque-Bera, que se define:

$$JB = \frac{T^* - p - q}{6} \left(Asimetria^2 + \frac{1}{4}(Kurtosis - 3)^2 \right)$$

y que bajo la hipótesis nula de normalidad se distribuye asintóticamente como $\chi^2(2)$. Se rechaza la hipótesis nula de normalidad al nivel de significación del 5 % si $JB > 5,99$.

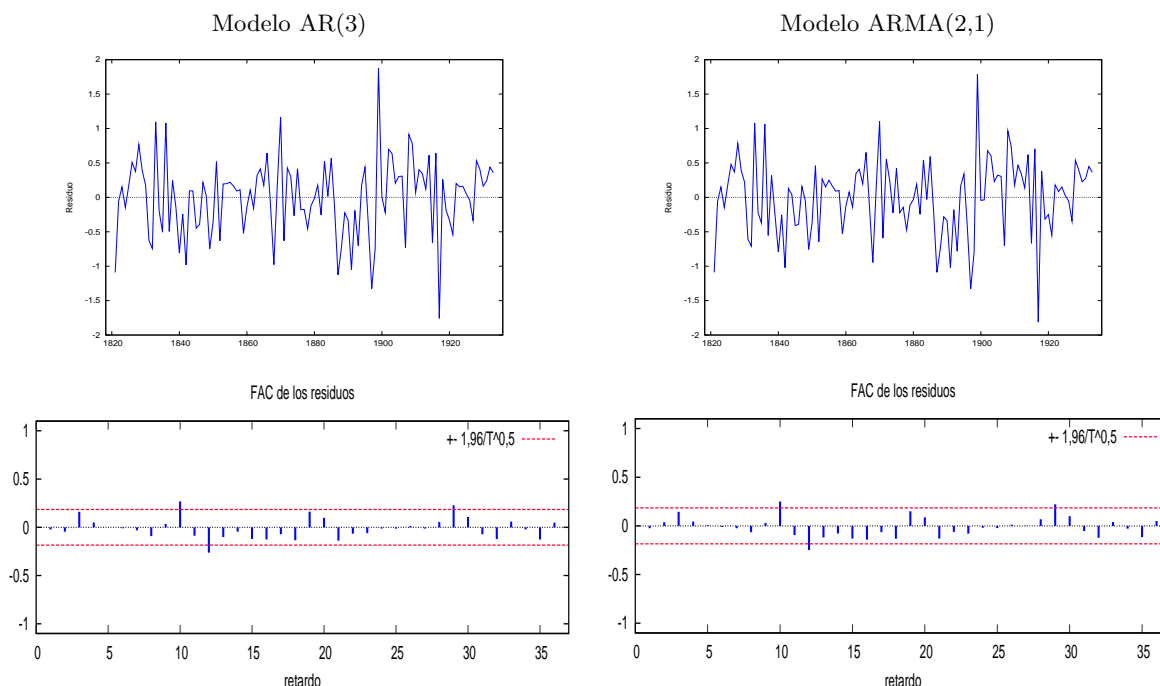
Ejemplo 6.7. Piel de Lince.

El gráfico 5.9 muestra el gráfico y el correlograma de los residuos para cada uno de los modelos propuestos y estimados para esta serie.

Se puede observar que el gráfico de los residuos de los dos modelos propuestos y estimados es muy similar y oscila en torno a cero con una varianza bastante homogénea.

En lo que se refiere al correlograma de los residuos, también en los correlogramas de los residuos de los dos modelos se observa, que la práctica totalidad de los coeficientes de autocorrelación

Gráfico 5.9: Ln(pieles de linco). Contrastes de diagnóstico.



simple estimados se encuentran dentro de las bandas de no significación, sobre todo los de los retardos más bajos.

Las siguientes tablas muestran los estadísticos descriptivos principales de los residuos de los dos modelos estimados para esta series.

Estadísticos principales, usando las observaciones 1821 - 1933 para la variable uhat1 (113 observaciones válidas) AR(3)			
Media	Mediana	Mínimo	Máximo
-0,00562629	0,0931793	-1,7580	1,87878
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,554851	98,6175	-0,128958	1,00983

Estadísticos principales, usando las observaciones 1821 - 1933 para la variable uhat2 (113 observaciones válidas) ARMA(2,1)			
Media	Mediana	Mínimo	Máximo
-0,00587872	0,0845680	-1,8127	1,78708
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,554805	94,3750	-0,203604	0,871798

El estadístico de Box-Ljung proporciona los siguientes resultados:

Modelo $AR(3)$	Modelo $ARMA(2, 1)$
$Q(15) = 22,42 < \chi_{0,05}^2(12) = 28,30$	$Q(15) = 21,51 < \chi_{0,05}^2(12) = 28,30$
$Q(36) = 54,36 < 57,67 \chi_{0,05}^2(33) = 28,30$	$Q(36) = 51,70 < \chi_{0,05}^2(33) = 57,67$

Por lo que no se rechaza la hipótesis nula de que los coeficientes de autocorrelación son cero.

En lo que se refiere al contraste de Jarque-Bera de Normalidad, se tiene que:

Modelo $AR(3)$	Modelo $ARMA(2, 1)$
$JB = 4,71 < \chi_{0,05}^2(2) = 5,99$	$JB = 4,05 < \chi_{0,05}^2(2) = 5,99$
No se rechaza H_0	No se rechaza H_0

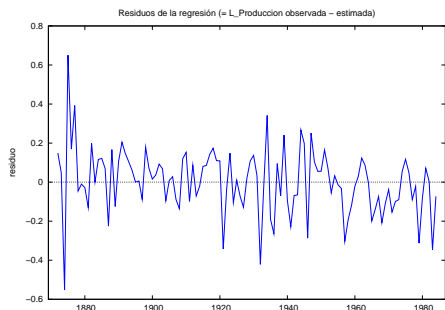
Ambos modelos propuestos satisfacen los contrastes de diagnóstico por lo que ambos recogen la estructura de la serie *Pieles de Lince*

Ejercicio 6.6. Producción de tabaco.

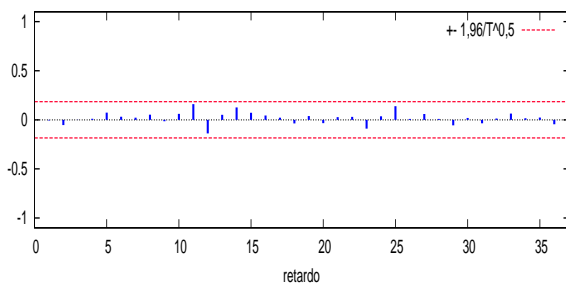
Estos son los resultados de estimar dos modelos para la serie $\text{Ln}(\text{Produccion})$.

Estadísticos principales, usando las observaciones 1871 - 1984 para la variable uhat1 (113 observaciones válidas) MA(1)			
Media	Mediana	Mínimo	Máximo
0,00227119	0,00881448	-0,550063	0,650369
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,166397	73,2644	-0,0241193	2,14855
Estadísticos principales, usando las observaciones 1871 - 1984 para la variable uhat2 (113 observaciones válidas) MA(2)			
Media	Mediana	Mínimo	Máximo
0,00226524	0,00864620	-0,550058	0,649804
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,166395	73,4557	-0,0253168	2,14227

Modelo ARIMA(0,1,1)



FAC de los residuos

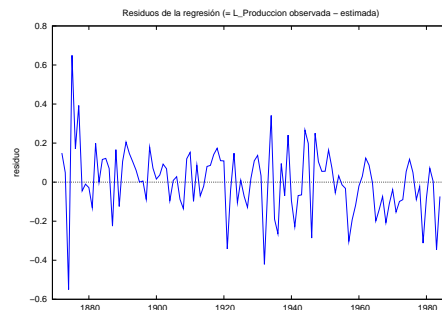


$$\hat{a} = -0,0025 \quad \hat{\sigma}_{\hat{a}_t} = 0,145$$

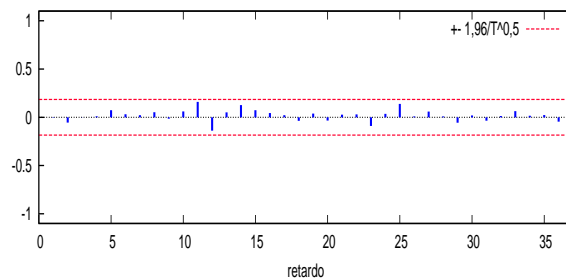
Normalidad: JB = 2,49

Autocorrelación: Q(15) = 10,86 Q(36) = 18,59

Modelo ARIMA(0,1,2)



FAC de los residuos



$$\hat{a} = -0,0028 \quad \hat{\sigma}_{\hat{a}_t} = 0,145$$

Normalidad: JB = 2,59

Autocorrelación: Q(15) = 10,86 Q(36) = 18,58

Matriz de covarianzas de los coeficientes (Modelo MA(1)).

	const	theta_1	
	1,96917e-05	6,82024e-07	const
		0,00376769	theta_1

Matriz de covarianzas de los coeficientes (Modelo MA(2)).

	const	theta_1	theta_2	
	1,97140e-05	1,90397e-06	-1,50197e-06	const
		0,00982488	-0,00737726	theta_1
			0,00899385	theta_2

Pregunta: ¿Qué modelo de los dos propuestos seleccionarías para la serie Producción de tabaco?

Apéndice A. Datos.

Cuadro 5.4: Pieles de lince (1821-1930)

Año	Pieles	Año	Pieles	Año	Pieles	Año	Pieles	Año	Pieles	Año	Pieles
1821	269	1840	409	1859	684	1878	299	1897	587	1916	674
1822	321	1841	151	1860	299	1879	201	1898	105	1917	81
1823	585	1842	45	1861	236	1880	229	1899	387	1918	80
1824	871	1843	68	1862	245	1881	469	1900	758	1919	108
1825	1475	1844	213	1863	552	1882	736	1901	1307	1920	229
1826	2821	1845	546	1864	1623	1883	2042	1902	3465	1921	1132
1827	3928	1846	1033	1865	3311	1884	2811	1903	6991	1922	2432
1828	5943	1847	2129	1866	6721	1885	4431	1904	6313	1923	3574
1829	4950	1848	2536	1867	4254	1886	2511	1905	3794	1924	2935
1830	2577	1849	957	1868	687	1887	389	1906	1836	1925	1537
1831	523	1850	361	1869	255	1888	73	1907	345	1926	529
1832	98	1851	377	1870	473	1889	39	1908	382	1927	485
1833	184	1852	225	1871	358	1890	49	1909	808	1928	662
1834	279	1853	360	1872	784	1891	59	1910	1388	1929	1000
1835	409	1854	731	1873	1594	1892	188	1911	2713	1930	1590
1836	2285	1855	1638	1874	1676	1893	377	1912	3800		
1837	2685	1856	2725	1875	2251	1894	1292	1913	3091		
1838	3409	1857	2871	1876	1426	1895	4031	1914	2985		
1839	1824	1858	2119	1877	756	1896	3495	1915	790		

Cuadro 5.5: Produccion de tabaco (1872-1984)

Año	Pieles	Año	Pieles	Año	Pieles	Año	Pieles	Año	Pieles	Año	Pieles
1872	327	1891	648	1910	1054	1929	1373	1948	2107	1967	1887
1873	385	1892	747	1911	1142	1930	1533	1949	1980	1968	1968
1874	382	1893	757	1912	941	1931	1648	1950	1969	1969	1710
1875	217	1894	767	1913	1117	1932	1565	1951	2030	1970	1804
1876	609	1895	767	1914	992	1933	1018	1952	2332	1971	1906
1877	466	1896	745	1915	1037	1934	1372	1953	2256	1972	1705
1878	621	1897	760	1916	1157	1935	1985	1954	2059	1973	1749
1879	455	1898	703	1917	1207	1936	1302	1955	2244	1974	1742
1880	472	1899	909	1918	1326	1937	1163	1956	2193	1975	1990
1881	469	1900	870	1919	1445	1938	1569	1957	2176	1976	2182
1882	426	1901	852	1920	1444	1939	1386	1958	1668	1977	2137
1883	579	1902	886	1921	1509	1940	1881	1959	1736	1978	1914
1884	509	1903	960	1922	1005	1941	1460	1960	1796	1979	2025
1885	580	1904	976	1923	1254	1942	1262	1961	1944	1980	1527
1886	611	1905	857	1924	1518	1943	1408	1962	2061	1981	1786
1887	609	1906	939	1925	1245	1944	1406	1963	2315	1982	2064
1888	469	1907	973	1926	1376	1945	1951	1964	2344	1983	1994
1889	661	1908	886	1927	1289	1946	1991	1965	2228	1984	1429
1890	525	1909	836	1928	1211	1947	1315	1966	1855		

Apéndice B. Contraste de Dickey-Fuller**Tablas de valores críticos del estadístico de Dickey-Fuller****Modelo sin constante**

T	Nivel de significación α							
	0,01	0,025	0,05	0,10	0,9	0,95	0,975	0,99
25	-2,66	- 2,26	-1,95	-1,60	0,92	1,33	1,70	2,16
50	-2,62	- 2,25	-1,95	-1,61	0,91	1,31	1,66	2,08
100	-2,60	- 2,24	-1,95	-1,61	0,90	1,29	1,64	2,03
250	-2,58	- 2,24	-1,95	-1,62	0,89	1,29	1,63	2,01
500	-2,58	- 2,23	-1,95	-1,62	0,89	1,28	1,62	2,00
∞	-2,58	- 2,23	-1,95	-1,62	0,89	1,28	1,62	2,00

Por ejemplo, para un tamaño de muestra $T = 25$ y $\alpha = 0,01$, se tiene que
 $Prob(t < -2,66) = 0,01$.

Modelo con constante

T	Nivel de significación α							
	0,01	0,025	0,05	0,10	0,9	0,95	0,975	0,99
25	-3,75	- 3,33	-3,00	-2,63	-0,37	0,00	0,34	0,72
50	-3,58	- 3,22	-2,93	-2,60	-0,40	-0,03	0,29	0,66
100	-3,51	- 3,17	-2,89	-2,58	-0,42	-0,05	0,26	0,63
250	-3,46	- 3,14	-2,88	-2,57	-0,42	-0,06	0,24	0,62
500	-3,44	- 3,13	-2,87	-2,57	-0,43	-0,07	0,24	0,61
∞	-3,43	- 3,12	-2,86	-2,57	-0,44	-0,07	0,23	0,60

Por ejemplo, para un tamaño de muestra $T = 25$ y $\alpha = 0,01$, se tiene que
 $Prob(t_c < -3,75) = 0,01$.

Capítulo 6

Predicción óptima con modelos ARIMA(p, d, q)

El objetivo es obtener predicciones óptimas de Y_t en algún momento futuro basadas en un conjunto de información dado que en el caso del análisis de series temporales univariante esta formado por el pasado disponible de la serie temporal

$$I_T = \{Y_T, Y_{T-1}, Y_{T-2}, Y_{T-2}, \dots\} \quad (6.1)$$

Supongamos que se observa una serie temporal denotada por Y_t , para $t = 0$ hasta $t = T$, donde T es la última observación de que disponemos. La predicción de series temporales supone decir algo sobre el valor que tomará la serie en momentos futuros $T + \ell$, donde ℓ representa el número de periodos en el futuro que estamos considerando. A la predicción de $Y_{T+\ell}$ con información hasta el momento T la vamos a denotar por $Y_T(\ell)$. Si $\ell = 1$, entonces se predice el valor de Y_{T+1} y se calcula lo que se denomina *predicción un periodo hacia adelante*. Si $\ell = 3$, entonces se predice el valor de Y_{T+3} y se calcula la *predicción tres periodos hacia adelante*, etc.

Como estamos considerando que la serie temporal es una realización de un proceso estocástico estacionario, el valor que se quiere predecir, $Y_{T+\ell}$, es también una variable aleatoria. Para predecir una variable aleatoria, se debería predecir su función de distribución y así poder hacer afirmaciones tales como: $\text{prob}(163 < Y_{T+\ell} \leq 190) = 0,42$. Ahora bien, en general, va a ser muy difícil determinar completamente la forma de la función de densidad sin hacer supuestos muy fuertes y poco realistas sobre la forma de esta función. Un objetivo menos ambicioso sería diseñar unos intervalos de confianza alrededor del valor $Y_{T+\ell}$, que nos permitan decir que $\text{prob}(B < Y_{T+\ell} \leq A) = 0,95$. Estos valores A y B permiten poner unos límites al valor que se quiere predecir con un grado de confianza de estar en lo cierto suficientemente alto.

Es interesante distinguir entre *predicción por intervalo* que conlleva la construcción de estos intervalos de predicción y la *predicción por punto*, que implica simplemente asignar un valor a $Y_{T+\ell}$ que de alguna manera represente a toda la distribución de valores. Por ejemplo, una predicción por punto sería afirmar que “se predice un número de ocupados en el sector industrial el primer trimestre de 2010 de 120.000”. Un ejemplo de la predicción por intervalo, sería decir: “ se predice un número de ocupados en el sector industrial el primer trimestre de 2010 entre 100.000 y 140.000 con una probabilidad de 0,95”. Si el método de predicción es bueno y si pudiéramos hacer un secuencia de predicciones de este tipo, es de esperar que el número real de ocupaados estaría fuera de los intervalos construidos solo en un 5% de los casos.

Por predictor óptimo (o predicción óptima) se denomina a áquel que es la mejor en el sentido de que minimiza una determinada función de pérdida. Lo más usual es minimizar el Error Cuadrático Medio de Predicción, por lo que diremos que $Y_T(\ell)$ es un predictor óptimo si minimiza el ECMP, es decir, si cumple que:

$$E[Y_{T+\ell} - Y_T(\ell)]^2 \leq E[Y_{T+\ell} - Y_T^*(\ell)]^2 \quad \forall Y_T^*(\ell)$$

Se puede demostrar que, bajo condiciones de regularidad muy débiles, el predictor por punto óptimo viene dado por la esperanza condicionada al conjunto de información:

$$Y_T(\ell) = E[Y_{T+\ell}|I_T] = E[Y_{T+\ell}|Y_T, Y_{T-1}, Y_{T-2}, Y_{T-3}, \dots] = E_T[Y_{T+\ell}]$$

es decir, por el valor esperado de la distribución de $Y_T(\ell)$ condicionada la información disponible.

Nada garantiza que esta esperanza condicionada sea una función lineal del pasado de la serie. Pero si el proceso sigue una distribución normal, se puede demostrar que la esperanza condicionada se puede expresar como una función lineal del conjunto de información, I_T . Por lo tanto, bajo el supuesto de normalidad, el predictor óptimo en el sentido de minimizar el ECMP es lineal. Si no se cumple este supuesto, la proyección lineal de $Y_{T+\ell}$ en su pasado proporcionaría el predictor óptimo dentro de la clase de predictores lineales.

La predicción óptima por intervalo se construirá a partir de la distribución del error de predicción que, bajo el supuesto de que $a_t \sim RBN(0, \sigma^2)$, es la siguiente:

$$e_T(\ell) = Y_{T+\ell} - Y_T(\ell) \sim N(0, V(e_T(\ell)))$$

Tipificando se obtiene:

$$\frac{Y_{T+\ell} - Y_T(\ell) - 0}{\sqrt{V(e_T(\ell))}} \sim N(0, 1)$$

De forma que el intervalo de predicción de probabilidad $(1 - \alpha)\%$ es:

$$\left[Y_T(\ell) - N_{\alpha/2} \sqrt{V(e_T(\ell))}, \quad Y_T(\ell) + N_{\alpha/2} \sqrt{V(e_T(\ell))} \right]$$

6.1. Predicción con modelos estacionarios

Consideremos el modelo lineal general (3.2), $MA(\infty)$ y el conjunto de información dado por (6.1). La estrategia de predicción se va a basar en escribir el valor que se desea predecir, $Y_{T+\ell}$, tal y como se genera en función del modelo para luego obtener la predicción óptima calculando la esperanza condicionada al conjunto de información. Para la representación medias móviles general, $Y_{T+\ell}$ viene dado por:

$$Y_{T+\ell} = a_{T+\ell} + \psi_1 a_{T+\ell-1} + \psi_2 a_{T+\ell-2} + \dots + \psi_{\ell-1} a_{T+1} + \psi_{\ell} a_T + \psi_{\ell+1} a_{T-1} + \psi_{\ell+2} a_{T-2} + \dots$$

Tomando la esperanza condicionada al conjunto de información, se obtiene:

$$Y_T(\ell) = E_T[Y_{T+\ell}] = \psi_{\ell} a_T + \psi_{\ell+1} a_{T-1} + \psi_{\ell+2} a_{T-2} + \psi_{\ell+3} a_{T-3} + \dots$$

dado que:

$$E_T(a_{T+j}) = \begin{cases} a_{T+j} & j \leq 0 \\ E(a_{T+j}) = 0 & j > 0 \end{cases}$$

La perturbación a_t es la innovación en el momento t . Si, dado el conjunto de información I_T , se conoce el verdadero valor de Y_t , como la parte sistemática se puede predecir mediante el modelo, la perturbación $a_t = Y_t - \widehat{P}S_t$ está determinada, es fija. Si, dado I_T , no se conoce el verdadero valor de Y_t , entonces la innovación a_t no está determinada por el conjunto de información, con lo que su media condicionada será la misma que su media no condicionada, es decir, cero.

Los errores de predicción son:

$$\begin{aligned} e_T(1) &= Y_{T+1} - \widehat{Y}_T(1) = a_{T+1} + \psi_1 a_T + \psi_2 a_{T-1} + \psi_3 a_{T-2} + \dots - \\ &\quad - (\psi_1 a_T + \psi_2 a_{T-1} + \psi_3 a_{T-2} + \dots) = a_{T+1} \\ e_T(2) &= Y_{T+2} - \widehat{Y}_T(2) = a_{T+2} + \psi_1 a_{T+1} + \psi_2 a_T + \psi_3 a_{T-1} + \dots - \\ &\quad - (\psi_2 a_T + \psi_3 a_{T-1} + \psi_4 a_{T-2} + \dots) = a_{T+2} + \psi_1 a_{T+1} \\ \dots &= \dots \\ e_T(\ell) &= Y_{T+\ell} - \widehat{Y}_T(\ell) = a_{T+\ell} + \psi_1 a_{T+\ell-1} + \psi_2 a_{T+\ell-2} + \dots + \psi_{\ell-1} a_{T+1} + \\ &\quad + \psi_\ell a_T + \psi_{\ell+1} a_{T-1} + \psi_{\ell+2} a_{T-2} + \dots - \\ &\quad - (\psi_\ell a_T + \psi_{\ell+1} a_{T-1} + \psi_{\ell+2} a_{T-2} + \psi_{\ell+3} a_{T-3} + \dots) = \\ &= a_{T+\ell} + \psi_1 a_{T+\ell-1} + \psi_2 a_{T+\ell-2} + \dots + \psi_{\ell-1} a_{T+1} \end{aligned}$$

Los errores de predicción son una combinación lineal de las perturbaciones futuras $a_{T+\ell}$, $\ell = 1, 2, \dots$ con valor medio cero:

$$E_T(e_T(\ell)) = E_T[a_{T+\ell} + \psi_1 a_{T+\ell-1} + \psi_2 a_{T+\ell-2} + \dots + \psi_{\ell-1} a_{T+1}] = 0$$

La varianza del error de predicción o Error Cuadrático Medio de Predicción viene dado por:

$$\begin{aligned} V_T(e_T(1)) &= E_T[e_T(1)]^2 = E_T[a_{T+1}]^2 = \sigma^2 \\ V_T(e_T(2)) &= E_T[e_T(2)]^2 = E_T[a_{T+2} + \psi_1 a_{T+1}]^2 = (1 + \psi_1^2) \sigma^2 \\ \dots &= \dots \\ V(e_T(\ell)) &= E_T[e_T(\ell)]^2 = E_T[a_{T+\ell} + \psi_1 a_{T+\ell-1} + \psi_2 a_{T+\ell-2} + \dots + \psi_{\ell-1} a_{T+1}]^2 = \\ &= (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{\ell-1}^2) \sigma^2 = \sigma^2 \sum_{i=0}^{\ell-1} \psi_i^2 \end{aligned} \quad (6.2)$$

Como se puede observar la varianza del error de predicción va creciendo conforme nos alejamos en el futuro. Ahora bien, si el proceso es estacionario se cumple que:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

por lo que esta varianza no crece indefinidamente, sino que tiene una cota máxima finita.

Se puede observar que la perturbación o innovación a_t y su varianza σ^2 tienen una nueva interpretación:

- $a_t = Y_t - Y_{t-1}(1)$, es el error de predicción un periodo hacia adelante.
- $V(a_t) = \sigma^2$, es la varianza del error de predicción un periodo hacia adelante.

Si el proceso ruido blanco sigue una distribución normal, se tiene que:

$$e_T(\ell) = Y_{T+\ell} - Y_T(\ell) \sim N[0, V(e_T(\ell))]$$

Por lo que el intervalo de predicción de probabilidad $(1 - \alpha)\%$ es:

$$\begin{aligned} Y_{T+1} : & \left[Y_T(1) - N_{\alpha/2} \sqrt{\sigma^2} \quad ; \quad Y_T(1) + N_{\alpha/2} \sqrt{\sigma^2} \right] \\ Y_{T+2} : & \left[Y_T(2) - N_{\alpha/2} \sqrt{\sigma^2 (1 + \psi_1^2)} \quad ; \quad Y_T(2) + N_{\alpha/2} \sqrt{\sigma^2 (1 + \psi_1^2)} \right] \\ & \vdots \quad \vdots \\ Y_{T+\ell} : & \left[Y_T(\ell) - N_{\alpha/2} \sqrt{\sigma^2 \sum_{i=0}^{\ell-1} \psi_i^2} \quad ; \quad Y_T(\ell) + N_{\alpha/2} \sqrt{\sigma^2 \sum_{i=0}^{\ell-1} \psi_i^2} \right] \end{aligned}$$

6.1.1. Predicción con modelos MA(q).

Comencemos por un modelo de medias móviles sencillo, por ejemplo, el MA(2) de media cero:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad a_t \sim RBN(0, \sigma^2) \quad t = 1, 2, \dots$$

La función de predicción es:

$$\begin{aligned} Y_{T+1} &= a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1} \\ Y_T(1) &= E_T[Y_{T+1}] = E_T[a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1}] = -\theta_1 a_T - \theta_2 a_{T-1} \\ Y_{T+2} &= a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T \\ Y_T(2) &= E_T[Y_{T+2}] = E_T[a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T] = -\theta_2 a_{T-1} \end{aligned}$$

$$\begin{aligned}
Y_{T+3} &= a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1} \\
Y_T(3) &= E_T[Y_{T+3}] = E_T[a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1}] = 0 \\
Y_T(\ell) &= E_T[Y_{T+\ell}] = 0 \quad (= E(Y_t)) \quad \forall \ell > 2
\end{aligned}$$

Por lo tanto, la función de predicción de un $MA(2)$, depende del conjunto de información, I_T , para $\ell = 1, 2$. A partir de $\ell > 2$, la predicción óptima viene dada por la media del proceso.

Estos resultados se pueden generalizar fácilmente para el modelo $MA(q)$:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad a_t \sim RBN(0, \sigma^2) \quad t = 1, 2, \dots$$

La función de predicción es:

$$Y_T(\ell) = \begin{cases} Y_T(1) = -\theta_1 a_T - \theta_2 a_{T-1} - \dots - \theta_q a_{T+1-q} \\ Y_T(2) = -\theta_2 a_T - \theta_3 a_{T-2} - \dots - \theta_q a_{T+2-q} \\ \dots \quad \dots \\ Y_T(q) = -\theta_q a_T \\ Y_T(\ell) = 0 \quad \forall \ell = q+1, q+2, \dots \end{cases}$$

Como el modelo $MA(2)$ está escrito directamente en forma medias móviles, se obtiene la varianza del error de predicción aplicando la expresión (6.2) con $\psi_i = \theta_i$, $i = 1, 2, \dots, q$ y $\psi_i = 0$, $\forall i > q$:

$$V(e_T(\ell)) = \begin{cases} V(e_T(1)) = \sigma^2 \\ V(e_T(2)) = (1 + \theta_1^2) \sigma^2 \\ \dots \quad \dots \\ V(e_T(q)) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_{q-1}^2) \sigma^2 \\ V(e_T(\ell)) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma^2 \quad (= V(Y_t)) \quad \ell = q+1, q+2, \dots \end{cases}$$

Aunque la varianza del error de predicción es una función creciente de ℓ , el horizonte de predicción, tiene una cota máxima que viene dada por la varianza no condicionada del proceso y que se alcanza para $\ell = q$.

Se puede concluir que para un modelo $MA(q)$ las predicciones para los q primeros horizontes de predicción, $\ell = 1, 2, \dots, q$, dependen del conjunto de información a través de los errores de predicción un periodo hacia adelante $a_T, a_{T-1}, \dots, a_{T+1-q}$, con lo que se mejora la predicción respecto de la media no condicionada del proceso porque se predice con una varianza del error de predicción menor que la varianza no condicionada del proceso. A partir de $\ell = q$, el conjunto de información no aporta nada a la predicción porque las predicciones óptimas son la media no condicionada del proceso y la varianza del error de predicción es la varianza no condicionada del proceso. Esto significa que, condicionando al conjunto de información, se obtienen los mismos resultados que sin condicionar, luego a partir de $\ell = q$, I_T ya no es informativo.

La predicción por intervalo viene dada por:

$$\begin{aligned}
\ell = 1 & \quad \left[-\theta_1 a_T - \theta_2 a_{T-1} - \dots - \theta_q a_{T+1-q} \pm N_{\alpha/2} \sqrt{\sigma^2} \right] \\
\ell = 2 & \quad \left[-\theta_2 a_T - \theta_3 a_{T-2} - \dots - \theta_q a_{T+2-q} \pm N_{\alpha/2} \sqrt{\sigma^2 (1 + \theta_1^2)} \right] \\
& \quad \dots \quad \dots \\
\ell = q & \quad \left[-\theta_q a_T \pm N_{\alpha/2} \sqrt{\sigma^2 (1 + \theta_1^2 + \dots + \theta_{q-1}^2)} \right] \\
\ell > q & \quad \left[0 \pm N_{\alpha/2} \sqrt{\sigma^2 (1 + \theta_1^2 + \dots + \theta_{q-1}^2 + \theta_q^2)} \right]
\end{aligned}$$

La amplitud de los intervalos de predicción va creciendo con ℓ , con el límite impuesto por

$$\pm N_{\alpha/2} \sqrt{\sigma^2 (1 + \theta_1^2 + \dots + \theta_{q-1}^2 + \theta_q^2)} = \pm N_{\alpha/2} \sqrt{V(Y_t)}$$

6.1.2. Predicción con modelos AR(p)

Consideremos el modelo autorregresivo más sencillo, el $AR(1)$.

$$Y_t = \phi Y_{t-1} + a_t \quad a_t \sim RBN(0, \sigma^2) \quad t = 1, 2, \dots$$

La función de predicción es:

$$\begin{aligned}
Y_{T+1} &= \phi Y_T + a_{T+1} \\
Y_T(1) &= E_T[Y_{T+1}] = E_T[\phi Y_T + a_{T+1}] = \phi Y_T \\
Y_{T+2} &= \phi Y_{T+1} + a_{T+2} \\
Y_T(2) &= E_T[Y_{T+2}] = E_T[\phi Y_{T+1} + a_{T+2}] = \phi E_T[Y_{T+1}] = \phi Y_T(1) \\
Y_{T+3} &= \phi Y_{T+2} + a_{T+3} \\
Y_T(3) &= E_T[Y_{T+3}] = E_T[\phi Y_{T+2} + a_{T+3}] = \phi E_T[Y_{T+2}] = \phi Y_T(2)
\end{aligned}$$

De forma que la función de predicción es:

$$Y_T(\ell) = \phi Y_T(\ell), \quad \ell = 1, 2, 3, \dots \quad (6.3)$$

dado que:

$$E_T(Y_{T+j}) = \begin{cases} Y_{T+j} & j \leq 0 \\ E(Y_T(j)) & j > 0 \end{cases}$$

La función de predicción (6.3) recoge una regla de cadena para obtener las predicciones de un proceso autorregresivo unas en función de las otras hasta un futuro indefinido. La trayectoria de la función de predicción depende de la estructura de la parte autorregresiva:

$$\begin{aligned} Y_T(1) &= \phi Y_T \\ Y_T(2) &= \phi Y_T(1) = \phi \phi Y_T = \phi^2 Y_T \\ Y_T(3) &= \phi Y_T(2) = \phi \phi^2 Y_T = \phi^3 Y_T \\ Y_T(4) &= \phi Y_T(3) = \phi \phi^3 Y_T = \phi^4 Y_T \\ Y_T(\ell) &= \phi^\ell Y_T \quad \ell = 1, 2, 3, \dots \end{aligned}$$

Como el proceso autorregresivo es estacionario, $|\phi| < 1$, y por lo tanto cuando nos alejamos en el futuro la función de predicción tiende hacia la media no condicionada del proceso:

$$\lim_{\ell \rightarrow \infty} Y_T(\ell) = 0 (= E(Y_t))$$

Para construir los intervalos de predicción, se ha de obtener la varianza del error de predicción. Para ello es preciso partir del modelo escrito en forma medias móviles. En el caso del $AR(1)$:

$$\begin{aligned} (1 - \phi L) Y_t = a_t &\rightarrow Y_t = \frac{1}{1 - \phi L} a_t \\ &\rightarrow Y_t = (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \dots) a_t \\ &\rightarrow Y_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + \dots \end{aligned}$$

Por lo que la varianza del error de predicción se obtiene aplicando la fórmula general (6.2) con $\psi_1 = \phi^i, \forall i$:

$$V(e_T(\ell)) = \begin{cases} V(e_T(1)) = \sigma^2 \\ V(e_T(2)) = (1 + \phi^2) \sigma^2 \\ V(e_T(3)) = (1 + \phi^2 + (\phi^2)^2) \sigma^2 \\ V(e_T(4)) = (1 + \phi^2 + (\phi^2)^2 + (\phi^3)^2) \sigma^2 \\ V(e_T(\ell)) = (1 + \phi^2 + (\phi^2)^2 + \dots + (\phi^{\ell-1})^2) \sigma^2 \end{cases}$$

La varianza del error de predicción es monotonamente creciente conforme nos alejamos en el futuro. Como el proceso es estacionario, esta varianza no crece indefinidamente sino que tiene una cota superior dada por la varianza no condicionada del proceso:

$$\lim_{\ell \rightarrow \infty} V(e_T(\ell)) = \lim_{\ell \rightarrow \infty} \sigma^2 \left[1 + \phi^2 + (\phi^2)^2 + \dots + (\phi^{\ell-1})^2 \right] = \frac{\sigma^2}{1 - \phi^2} = V(Y_t)$$

La predicción por intervalo es:

$$\begin{aligned} \ell = 1 & \quad \left[\phi Y_T \pm N_{\alpha/2} \sqrt{\sigma^2} \right] \\ \ell = 2 & \quad \left[\phi Y_T(1) \pm N_{\alpha/2} \sqrt{\sigma^2 (1 + \phi^2)} \right] \\ \ell = 3 & \quad \left[\phi Y_T(2) \pm N_{\alpha/2} \sqrt{\sigma^2 (1 + \phi^2 + (\phi^2)^2)} \right] \\ & \quad \dots \quad \dots \\ \ell & \quad \left[\phi Y_T(\ell - 1) \pm N_{\alpha/2} \sqrt{\sigma^2 (1 + \phi^2 + (\phi^2)^2 + \dots + (\phi^{\ell-1})^2)} \right] \end{aligned}$$

La amplitud de los intervalos de predicción va creciendo con ℓ , con el límite impuesto por

$$\pm N_{\alpha/2} \sqrt{\frac{\sigma^2}{1 - \phi^2}} = \pm N_{\alpha/2} \sqrt{V(Y_t)}$$

Los resultados obtenidos para el modelo $AR(1)$ se pueden extender para el modelo $AR(p)$. En general, las funciones de predicción de procesos autorregresivos puros, se obtendrán a partir de reglas de cadena:

$$Y_T(\ell) = \phi_1 Y_T(\ell - 1) + \phi_2 Y_T(\ell - 2) + \phi_3 Y_T(\ell - 3) + \dots + \phi_p Y_T(\ell - p), \quad \ell = 1, 2, 3, \dots$$

La función de predicción de un proceso $AR(1)$ utiliza la última observación Y_T para obtener la predicción un periodo hacia adelante y luego, a partir de ésta, se obtienen el resto de las predicciones. En el caso de un autorregresivo de orden p autorregresivo de orden p , se utilizarán las p últimas observaciones para obtener las predicciones para $\ell = 1, 2, \dots, p$, y el resto se obtienen a partir de las p primeras.

6.1.3. Predicción con modelos ARMA(p,q).

Consideremos un modelo $ARMA(p, q)$ sencillo, el $ARMA(1, 2)$:

$$Y_t = \delta + \phi Y_{t-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad a_t \sim RBN(0, \sigma^2) \quad t = 1, 2, \dots \quad (6.4)$$

La media de este proceso no es cero si $\delta \neq 0$:

$$E(Y_t) = \frac{\delta}{1 - \phi}$$

Las predicciones por punto son:

$$\begin{aligned}
Y_{T+1} &= \delta + \phi Y_T + a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1} \\
Y_T(1) &= E_T[Y_{T+1}] = E_T[\delta + \phi Y_T + a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1}] = \\
&= \delta + \phi Y_T - \theta_1 a_T - \theta_2 a_{T-1} \\
Y_{T+2} &= \delta + \phi Y_{T+1} + a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T \\
Y_T(2) &= E_T[Y_{T+2}] = E_T[\delta + \phi Y_{T+1} + a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T] = \\
&= \delta + \phi Y_T(1) - \theta_2 a_T \\
Y_{T+3} &= \delta + \phi Y_{T+2} + a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1} \\
Y_T(3) &= E_T[Y_{T+3}] = E_T[\delta + \phi Y_{T+2} + a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1}] = \\
&= \delta + \phi Y_T(2) \\
\rightarrow Y_T(\ell) &= E_T[Y_{T+\ell}] = \delta + \phi Y_T(\ell - 1) \quad \forall \ell > 2
\end{aligned}$$

La estructura de la función de predicción es la siguiente. Las dos primeras predicciones dependen de la última observación Y_T (parte autorregresiva) y de los últimos errores de predicción un periodo hacia adelante a_T y a_{T-1} (parte medias móviles). Para $\ell > 2$, la parte medias móviles no aparece de forma explícita en la función de predicción, y cada predicción se va obteniendo de las anteriores siguiendo una *regla en cadena* marcada por la parte autorregresiva. Esta función se va acercando a la media del proceso conforme nos alejamos en el futuro:

$$\begin{aligned}
Y_T(3) &= \delta + \phi Y_T(2) \\
Y_T(4) &= \delta + \phi Y_T(3) = \delta + \phi(\delta + \phi Y_T(2)) = \delta(1 + \phi) + \phi^2 Y_T(2) \\
Y_T(5) &= \delta + \phi Y_T(4) = \delta + \phi(\delta(1 + \phi) + \phi^2 Y_T(2)) = \delta(1 + \phi + \phi^2) + \phi^3 Y_T(2) \\
&\dots \quad \dots \\
Y_T(\ell) &= \delta + \phi Y_T(\ell - 1) = \delta(1 + \phi + \phi^2 + \dots + \phi^{\ell-3}) + \phi^{\ell-2} Y_T(2)
\end{aligned}$$

de forma que como el modelo $ARMA(2, 1)$ es estacionario, $|\phi| < 1$ y:

$$\lim_{\ell \rightarrow \infty} Y_T(\ell) = \delta \sum_{i=0}^{\ell-3} \phi^i = \frac{\delta}{1 - \phi} (= E(Y_t))$$

Para obtener la varianza del error de predicción y, por lo tanto, las predicciones por intervalo, se deriva la representación medias móviles infinita:

$$\begin{aligned}
(1 - \phi L) Y_t &= (1 - \theta_1 L - \theta_2 L^2) a_t \\
\rightarrow Y_t &= \frac{1 - \theta_1 L - \theta_2 L^2}{1 - \phi L} a_t = (1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots) a_t
\end{aligned}$$

de donde:

$$\frac{1 - \theta_1 L - \theta_2 L^2}{1 - \phi L} = 1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots$$

$$\rightarrow 1 - \theta_1 L - \theta_2 L^2 = (1 - \phi L)(1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots)$$

e igualando coeficientes:

$$\begin{aligned} L \quad -\theta_1 L &= (\psi_1 - \phi) L \quad \Rightarrow \quad -\theta_1 = \psi_1 - \phi \quad \Rightarrow \quad \psi_1 = \phi - \theta_1 \\ L^2 \quad -\theta_2 L^2 &= (\psi_2 - \phi\psi_1) L^2 \quad \Rightarrow \quad -\theta_2 = \psi_2 - \phi\psi_1 \quad \Rightarrow \quad \psi_2 = \phi\psi_1 - \theta_2 = \phi(\phi - \theta_1) - \theta_2 \\ L^3 \quad 0 L^3 &= (\psi_3 - \psi_2\phi) L^3 \quad \Rightarrow \quad 0 = \psi_3 - \psi_2\phi \quad \Rightarrow \quad \psi_3 = \phi\psi_2 \\ \dots \quad \dots & \end{aligned}$$

Los pesos de la forma medias móviles infinita son:

$$\psi_i = \begin{cases} k = 0 & \psi_0 = 1 \\ k = 1 & \psi_1 = \phi - \theta_1 \\ k = 2 & \psi_2 = \phi\psi_1 - \theta_2 = \phi(\phi - \theta_1) - \theta_2 \\ k > 2 & \psi_k = \phi\psi_{k-1} \end{cases}$$

con estos pesos se pueden construir los intervalos de predicción. Como el proceso $ARMA(1, 2)$ es estacionario, la amplitud de los intervalos irá creciendo conforme nos alejamos en el futuro pero con una cota máxima dada por $\left[\pm N_{\alpha/2} \times \sqrt{V(Y_t)} \right]$.

6.1.4. Predicciones con modelos estacionarios estimados

Habitualmente no se conoce el proceso que ha generado la serie temporal Y_t por lo que hay que estimarlo con los datos disponibles, obteniendo:

$$\hat{\phi}_p(L) Y_t = \hat{\theta}_q(L) \hat{a}_t \quad t = 1, 2, \dots, T$$

donde \hat{a}_t son los residuos del modelo, pero también una estimación del error de predicción un periodo hacia adelante.

Por ejemplo, en el caso del modelo (6.4), la función de predicción estimada sería:

$$Y_T(\ell) = \begin{cases} \ell = 1 & Y_T(1) = \hat{\delta} + \hat{\phi} Y_T - \hat{\theta}_1 \hat{a}_T - \hat{\theta}_2 \hat{a}_{T-1} \\ \ell = 2 & Y_T(2) = \hat{\delta} + \hat{\phi} Y_T(1) - \hat{\theta}_2 \hat{a}_T \\ \ell > 2 & Y_T(\ell) = \hat{\delta} + \hat{\phi} Y_T(\ell - 1) \end{cases}$$

Ejemplo 4.1. Modelo $MA(2)$

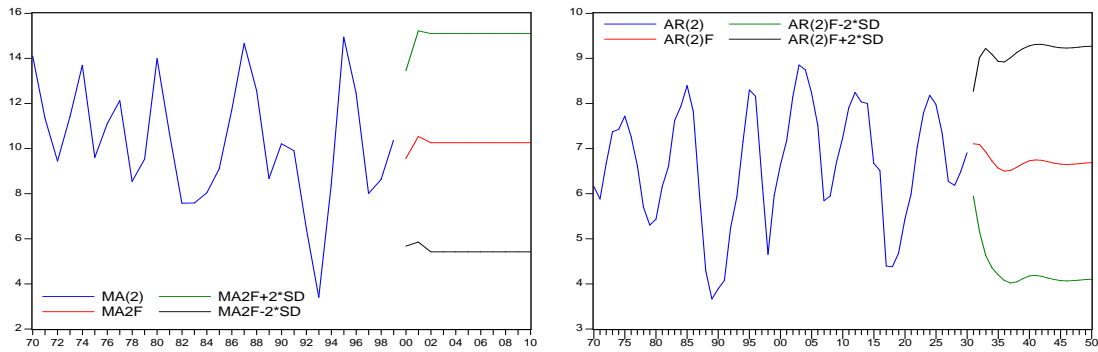
$$Y_t = 10,3 + \hat{a}_t + 0,67 \hat{a}_{t-1} - 0,31 \hat{a}_{t-2} \quad t = 1, 2, \dots, T$$

Función de predicción:

$$Y_T(\ell) = \begin{cases} \ell = 1 & Y_T(1) = 10,3 + 0,67 \hat{a}_T - 0,31 \hat{a}_{T-1} \\ \ell = 2 & Y_T(2) = 10,3 - 0,31 \hat{a}_T \\ \ell > 2 & Y_T(\ell) = 10,3 (= E(\hat{Y}_t)) \end{cases}$$

Todos los procesos medias móviles finitos son estacionarios y, como se puede observar en la figura de la izquierda del gráfico 6.1, a partir de $\ell = 2$ la función de predicción permanece constante en la media del proceso con una amplitud constante del intervalo de predicción.

Gráfico 6.1: Proceso $MA(2)$ versus Proceso $AR(2)$



Ejemplo 4.2. Modelo $AR(2)$

$$Y_t = 2,4 + 1,33 Y_{t-1} - 0,69 Y_{t-2} + \hat{a}_t \quad t = 1, 2, \dots, T$$

Función de predicción:

$$Y_T(\ell) = \begin{cases} \ell = 1 & Y_T(1) = 2,4 + 1,33 Y_T - 0,69 Y_{T-1} \\ \ell = 2 & Y_T(2) = 2,4 + 1,33 Y_T(1) - 0,69 Y_T \\ \ell = 1, 2, 3, \dots & Y_T(\ell) = 2,4 + 1,33 Y_T(\ell - 1) - 0,69 Y_T(\ell - 2) \end{cases}$$

El proceso $AR(2)$ es estacionario:

$$1 - 1,33 L + 0,69 L^2 = 0$$

$$\rightarrow L_{1,2} = \frac{1,33 \pm \sqrt{1,33^2 - 4 \times 0,69}}{2 \times 0,69} = \frac{1,33 \pm \sqrt{-0,99}}{1,38} = \frac{1,33}{1,38} \pm \frac{\sqrt{0,99}}{1,38} i = a \pm bi$$

$$|L_1| = |L_1| = \sqrt{a^2 + b^2} = \sqrt{\left(\frac{1,33}{1,38}\right)^2 + \left(\frac{\sqrt{0,99}}{1,38}\right)^2} = \sqrt{1,4486} = 1,2 > 1$$

Por lo que la función de predicción tiende a su media cuando nos alejamos en el futuro:

$$\lim_{\ell \rightarrow \infty} Y_T(\ell) = \hat{E}(Y_t) = \frac{\hat{\delta}}{1 - \hat{\phi}_1 - \hat{\phi}_2} = \frac{2,4}{1 - 1,33 + 0,69} = 6,67$$

Este resultado se puede observar claramente en la figura de la derecha del gráfico 6.1. La serie presenta un comportamiento cíclico que se refleja en la función de predicción que presenta un ciclo amortiguado antes de dirigirse sistemáticamente hacia la media del proceso.

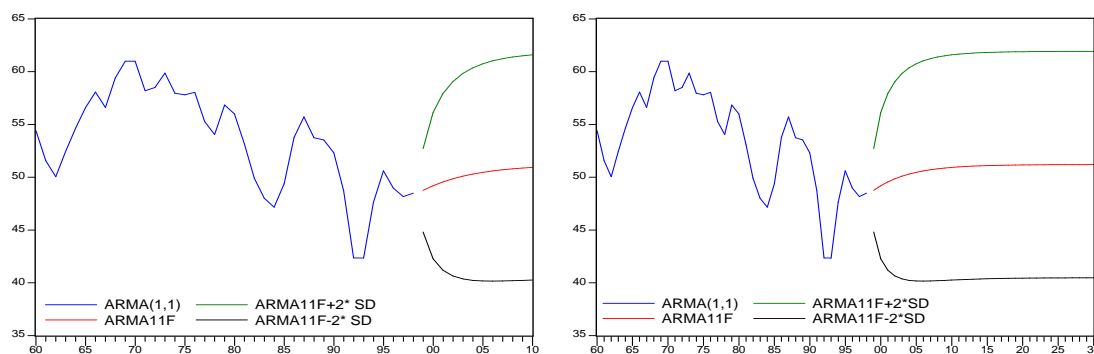
Ejemplo 4.3. Modelo ARMA(1,1)

$$Y_t = 9,3 + 0,82Y_{t-1} + \hat{a}_t + 0,62\hat{a}_{t-1} \quad t = 1, 2, \dots, T$$

La función de predicción es:

$$Y_T(\ell) = \begin{cases} \ell = 1 & Y_T(1) = 9,3 + 0,82 Y_T + 0,62 \hat{a}_T \\ \ell > 2 & Y_T(\ell) = 9,3 + 0,82 Y_T(\ell - 1) \end{cases}$$

Gráfico 6.2: Función de predicción: ARMA(1,1) estacionario



Como el proceso es estacionario dado que cumple la condición de que la raíz del polinomio autorregresivo es, en valor absoluto, mayor que la unidad,

$$1 - 0,82L = 0 \quad \rightarrow \quad L = \frac{1}{0,82} = 1,22 \quad \rightarrow \quad |L| = |1,22| > 1$$

la función de predicción tiende a su media cuando nos alejamos en el futuro:

$$\lim_{\ell \rightarrow \infty} Y_T(\ell) = \hat{E}(Y_t) = \frac{\hat{\delta}}{1 - \hat{\phi}} = \frac{9,3}{1 - 0,82} = 51,67$$

En las figuras del gráfico 6.2 se puede observar el comportamiento de la función de predicción por punto y por intervalo. Las últimas observaciones se encuentran por debajo de la media lo que lleva a que la función de predicción parta de un valor inferior a la media estimada del proceso. De forma que la trayectoria de la función de predicción es creciente al principio para dirigirse hacia la media del proceso donde se va estabilizando. Asimismo, se observa el comportamiento de los intervalos de predicción, de amplitud creciente al principio hasta estabilizarse con la varianza del proceso.

6.2. Predicción con modelos no estacionarios.

La predicción con modelos no estacionarios $ARIMA(p, d, q)$ se lleva a cabo de la misma manera que con los modelos estacionarios $ARMA(p, q)$. El predictor por punto óptimo de $Y_{T+\ell}$ viene dado por la esperanza condicionada al conjunto de información $Y_T(\ell) = E_T[Y_{T+\ell}]$. Para obtener esta esperanza condicionada basta con escribir el modelo en forma de ecuación en diferencias y obtener las esperanzas condicionadas, sabiendo que:

$$E_T[Y_{T+j}] = \begin{cases} Y_{T+j} & j \leq 0 \\ Y_T(j) & j > 0 \end{cases} \quad E_T[a_{T+j}] = \begin{cases} a_{T+j} & j \leq 0 \\ 0 & j > 0 \end{cases}$$

Para construir los intervalos de predicción,

$$\left[Y_T(\ell) \pm N_{\alpha/2} \sqrt{V(e_T(\ell))} \right] \quad \text{donde} \quad V(e_T(\ell)) = \sigma^2 \sum_{j=0}^{\ell-1} \psi_j^2$$

el modelo ha de estar escrito en forma $MA(\infty)$ ya que ψ_j son los pesos del modelo $ARIMA$ escrito en forma medias móviles.

Para analizar las características de la función de predicción para modelos no estacionarios $ARIMA(p, d, q)$, consideremos varios ejemplos sencillos.

Ejemplo 5.1. Modelo $ARIMA(0, 1, 1)$.

Consideremos que la serie Y_t ha sido generada por el siguiente modelo:

$$\begin{aligned} (1 - L)Y_t &= (1 + \theta L)a_t \\ Y_t &= Y_{t-1} + a_t + \theta a_{t-1} \end{aligned}$$

La función de predicción es:

$$\begin{aligned} Y_T(1) &= E_T[Y_{T+1}] = Y_T + \theta a_T \\ Y_T(2) &= E_T[Y_{T+2}] = Y_T(1) = Y_T + \theta a_T \\ Y_T(3) &= E_T[Y_{T+3}] = Y_T(2) = Y_T(1) = Y_T + \theta a_T \\ &\vdots \\ Y_T(\ell) &= E_T[Y_{T+\ell}] = Y_T(\ell - 1) = Y_T + \theta a_T \quad \ell = 1, 2, 3, \dots \end{aligned}$$

Por lo tanto, la función de predicción es una línea horizontal: pasa por la predicción un periodo hacia adelante, $Y_T(1)$, que depende del conjunto de información a través de Y_T y a_T y del parámetro del modelo, θ , y permanece allí conforme ℓ crece.

En el caso del modelo de paseo aleatorio, como $\theta = 0$, la función de predicción es:

$$Y_T(\ell) = E_T[Y_{T+\ell}] = Y_T \quad \ell = 1, 2, 3, \dots$$

Por lo tanto, las predicciones óptimas vienen dadas por la última observación independientemente del horizonte de predicción (veáse la figura izquierda del gráfico 6.3).

Se puede demostrar que si la serie Y_t ha sido generada por un proceso integrado de orden 1, de forma que se puede representar mediante un modelo $ARIMA(p, 1, q)$ sin término independiente ($\delta = 0$), la función de predicción cuando $\ell \rightarrow \infty$, tiende a una constante:

$$Y_T(\ell) \xrightarrow{\ell \rightarrow \infty} K^T$$

Hay que tener en cuenta que K^T no es la media del proceso porque como no es estacionario no tiene una media hacia dónde ir, sino que es una constante que depende del conjunto de información y de los parámetros AR y MA del modelo.

Obtengamos la función de predicción para un proceso integrado de orden 1 con constante, por ejemplo, el modelo de paseo aleatorio con deriva (4.5).

$$\begin{aligned} Y_T(1) &= E_T[Y_{T+1}] = Y_T + \delta \\ Y_T(2) &= E_T[Y_{T+2}] = Y_T(1) + \delta = Y_T + 2\delta \\ Y_T(3) &= E_T[Y_{T+3}] = Y_T(2) + \delta = Y_T + 3\delta \\ &\vdots \\ Y_T(\ell) &= E_T[Y_{T+\ell}] = Y_T(\ell-1) + \delta = Y_T + \ell\delta \quad \ell = 1, 2, 3, \dots \end{aligned}$$

La función de predicción es una línea recta de pendiente δ (veáse la figura derecha del gráfico 6.3).

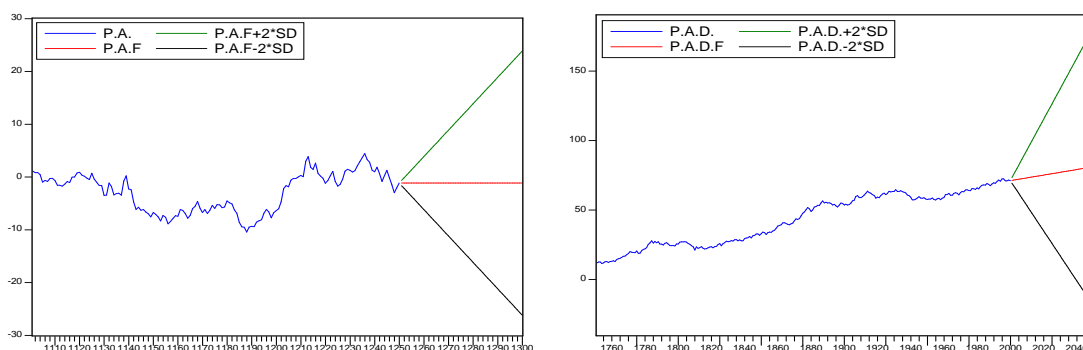
Se puede demostrar que si la serie Y_t ha sido generada por un proceso integrado de orden 1, de forma que se puede representar mediante un modelo $ARIMA(p, 1, q)$ con término independiente ($\delta \neq 0$), la función de predicción tiende a una línea recta cuando $\ell \rightarrow \infty$,

$$Y_T(\ell) \xrightarrow{\ell \rightarrow \infty} K^T + \delta \ell$$

La pendiente de la función de predicción viene dada por la constante δ y el intercepto, K^T , depende del conjunto de información y de los parámetros del modelo.

En lo que se refiere a la predicción por intervalo para modelos no estacionarios, las figuras del gráfico 6.3 muestran como la amplitud de los intervalos de predicción para los modelos $ARIMA(p, d, q)$ crece indefinidamente conforme el horizonte de predicción ℓ se hace mayor. Hay que tener en cuenta que cuando el proceso no es estacionario el límite $\lim_{\ell \rightarrow \infty} V[e_T(\ell)]$ no existe.

Gráfico 6.3: Modelos ARIMA: Funciones de predicción.



Para calcular la $V[e_T(\ell)]$ de un modelo no estacionario $ARIMA(p, d, q)$ es necesario escribir el modelo en forma $MA(\infty)$. Pongamos como ejemplo el modelo de paseo aleatorio (4.4) que, escrito en forma $MA(\infty)$ queda como sigue:

$$Y_t = a_t + a_{t-1} + a_{t-2} + a_{t-3} + \dots \quad \text{de forma que } \psi_j = 1 \quad \forall j$$

La varianza del error de predicción es:

$$\begin{aligned} V[e_T(1)] &= \sigma^2 \sum_{j=0}^0 \psi_j^2 = \sigma^2 \\ V[e_T(2)] &= \sigma^2 \sum_{j=0}^1 \psi_j^2 = \sigma^2 (1 + 1) = 2\sigma^2 \\ V[e_T(3)] &= \sigma^2 \sum_{j=0}^2 \psi_j^2 = \sigma^2 (1 + 1 + 1) = 3\sigma^2 \\ &\vdots \\ V[e_T(\ell)] &= \sigma^2 \sum_{j=0}^{\ell-1} \psi_j^2 = \sigma^2 (1 + 1 + \dots + 1) = \ell\sigma^2 \quad \ell = 1, 2, 3, \dots \end{aligned}$$

que no tiene límite finito.

Por último, considerando el caso de procesos integrados de orden 2, $ARIMA(p, 2, q)$, su función de predicción final toma la forma:

$$Y_T(\ell) \xrightarrow{\ell \rightarrow \infty} K_1^T + K_2^T \ell$$

Es decir, conforme el horizonte de predicción aumenta y nos alejamos en el futuro la función de predicción tiende a una línea recta donde tanto el intercepto, K_1^T , como la pendiente, K_2^T , dependen del conjunto de información y de los parámetros del modelo. Por lo tanto, aunque la

estructura de esta función de predicción es la misma que la de los modelos $ARIMA(p, 1, q)$ con término independiente (por ejemplo, la función de predicción del paseo aleatorio con deriva), esta función de predicción es más flexible.

Capítulo 7

Modelos ARIMA estacionales

Muchas series económicas si se observan varias veces a lo largo del año, trimestral o mensualmente, presentan comportamiento estacional. Este tipo de comportamiento puede ser debido a factores meteorológicos, tales como la temperatura, pluviosidad, etc. que afectan a muchas actividades económicas como el turismo; a costumbres sociales como las Navidades que están muy relacionadas con cierto tipo de ventas como juguetes; a fenómenos sociales como el caso del índice de producción industrial que en este país desciende considerablemente todos los años el mes de agosto debido a las vacaciones, etc.

A la hora de elaborar el modelo ARIMA adecuado para una serie temporal se ha de tener en cuenta el comportamiento estacional, si lo hubiere, porque implica que la observación de un mes y observación del mismo mes del año anterior tienen una pauta de comportamiento similar por lo que estarán temporalmente correlacionadas. Por lo tanto, el modelo de series temporales ARIMA apropiado para este tipo de series deberá recoger las dos clases de dependencia intertemporal que presentan, a saber

- la relación lineal existente entre observaciones sucesivas (comportamiento tendencial o regular) y
- la relación lineal existente entre observaciones del mismo mes en años sucesivos (comportamiento estacional).

El objetivo de este tema es extender los modelos ARIMA estudiados en los capítulos 3 y 4 para las series con comportamiento estacional.

Supongamos que la serie Y_t presenta un componente estacional y se especifica un modelo $ARIMA(p, d, q)$ general:

$$\phi_p(L) Y_t = \theta_q(L) a_t$$

Para recoger las dos estructuras de correlación anteriormente mencionadas, la regular y la estacional, bastaría con añadir al modelo los retardos de Y_t y a_t necesarios. Así, si la serie estacional es mensual y se quiere representar, además de la dependencia entre observaciones consecutivas, la autocorrelación entre observaciones del mismo mes separadas un año, dos años, etc. será necesario incluir en el modelo retardos hasta de orden 12, 24, etc. con lo que esto supone de aumento en el número de los parámetros del modelo. Así, por ejemplo, un modelo sencillo para una serie estacional, un $ARMA(12, 12)$, contiene 24 parámetros desconocidos. Como se ha discutido en el capítulo 5, uno de los criterios que se ha de seguir al construir un modelo $ARIMA$ es el de

parsimonia, es decir, buscar el modelo más sencillo y con el menor número de parámetros posible que reproduzca las características de la serie. Por lo tanto, en este capítulo se derivarán modelos capaces de recoger el comportamiento estacional pero con un número pequeño de parámetros, es decir, más parsimoniosos.

Antes de especificar un modelo apropiado, dentro del marco *ARIMA*, para una serie con tendencia y estacionalidad comenzaremos por estudiar las características de la dependencia lineal estacional a través de unos modelos muy sencillos, *los modelos estacionales puros*.

7.1. Modelos estacionales puros

Se entiende por modelo estacional puro aquel que recoge únicamente relaciones lineales entre observaciones del mismo mes para años sucesivos, es decir, entre observaciones separadas s periodos o múltiplos de s , donde $s = 4$ si la serie es trimestral y $s = 12$ si la serie es mensual. Por lo tanto, en estos modelos, para simplificar, se parte del supuesto de que no existe estructura regular, es decir, correlación entre observaciones consecutivas. Como este supuesto es poco realista, este tipo de modelos no va a ser muy útil en la práctica pero su estudio permitirá identificar en qué retardos de la función de autocovarianzas y/o autocorrelación se refleja la estructura de tipo estacional de una serie. Se denotarán, en general, $ARMA(P, Q)_s$, donde P es el orden del polinomio autorregresivo estacionario y Q es el orden del polinomio medias móviles invertible.

Vamos a analizar las características de los modelos estacionales puros específicos para un proceso Y_t estacionario y con estacionalidad trimestral, $s = 4$.

Modelo $MA(1)_4$

$$Y_t = (1 - \Theta L^4) a_t \qquad Y_t = a_t - \Theta a_{t-4}$$

En este proceso, el valor de Y en el momento t depende de la perturbación contemporánea y de la perturbación del mismo trimestre del año anterior. Como este proceso es un modelo de medias móviles finito es estacionario para cualquier valor de Θ y su media es cero. El proceso será invertible si y solo si las cuatro raíces del polinomio de medias móviles tienen modulo fuera del círculo unidad.

Las autocovarianzas vienen dada por:

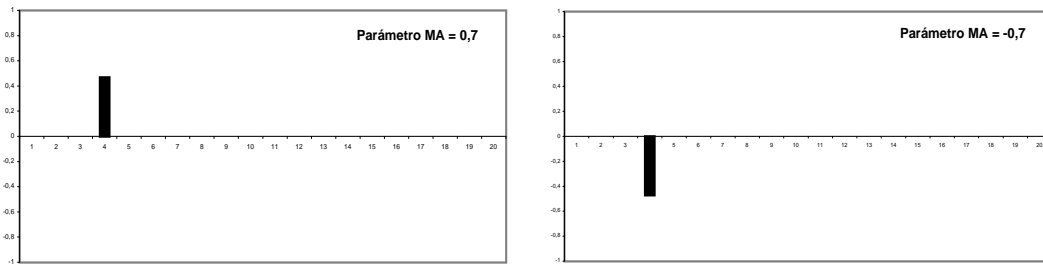
$$\begin{aligned} \gamma_0 &= V(Y_t) = E[Y_t]^2 = E[a_t - \Theta a_{t-4}]^2 = (1 + \Theta^2) \sigma^2 \\ \gamma_1 &= Cov(Y_t Y_{t-1}) = E[Y_t Y_{t-1}] = E[(a_t - \Theta a_{t-4})(a_{t-1} - \Theta a_{t-5})] = 0 \\ \gamma_2 &= Cov(Y_t Y_{t-2}) = E[Y_t Y_{t-2}] = E[(a_t - \Theta a_{t-4})(a_{t-2} - \Theta a_{t-6})] = 0 \\ \gamma_3 &= Cov(Y_t Y_{t-3}) = E[Y_t Y_{t-3}] = E[(a_t - \Theta a_{t-4})(a_{t-3} - \Theta a_{t-7})] = 0 \\ \gamma_4 &= Cov(Y_t Y_{t-4}) = E[Y_t Y_{t-4}] = E[(a_t - \Theta a_{t-4})(a_{t-4} - \Theta a_{t-8})] = -\Theta \sigma^2 \\ \gamma_k &= Cov(Y_t Y_{t-k}) = E[Y_t Y_{t-k}] = E[(a_t - \Theta a_{t-4})(a_{t-k} - \Theta a_{t-4-k})] = 0, \quad \forall k > 4 \end{aligned}$$

Por lo tanto, la función de autocovarianzas y la función de autocorrelación son las siguientes:

Retardo	FACV	FAC
$k = 0$	$\gamma_0 = (1 + \Theta^2) \sigma^2$	$\rho_0 = 1$
$k = 4$	$\gamma_4 = -\Theta \sigma^2$	$\rho_4 = \frac{-\Theta}{(1 + \Theta^2)}$
$k \neq 4$	$\gamma_k = 0$	$\rho_k = 0$

Se puede observar que, como para todo modelo de medias móviles finito, la función de autocorrelación se trunca y se hace cero a partir de un determinado retardo. Como el orden de la media móvil es uno, solo un coeficiente de autocorrelación es distinto de cero y el resto son nulos, pero como la estructura de correlación temporal es estacional, este coeficiente de autocorrelación no nulo es el asociado a $k = s = 4$ (veáse el gráfico 7.1). Denominemos *retardos estacionales* a los asociados con s y múltiplos de s : $k = s, 2s, 3s, \dots$, para distinguirlos así de la estructura regular, existente entre las observaciones sucesivas y que se refleja fundamentalmente para los retardos más cortos: $k = 1, 2, 3, \dots$. Por lo tanto, la estructura de la función de autocorrelación de un $MA(1)_4$ es la siguiente: función truncada en el primer retardo estacional.

Gráfico 7.1: Función de autocorrelación. Modelo $MA(1)_4$



Generalizando el resultado obtenido para el modelo $MA(1)_4$, se puede demostrar que la estructura de la función de autocorrelación para un modelo de medias móviles, $MA(Q)_s$:

$$Y_t = \Theta_Q(L^s) a_t = (1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_Q L^{Qs}) a_t$$

$$Y_t = a_t + \Theta_1 a_{t-s} + \Theta_2 a_{t-2s} + \dots + \Theta_Q a_{t-Qs}$$

es como sigue:

$$\rho_k = \begin{cases} k = 0 & \rho_0 = 1 \\ k = s, 2s, \dots, Qs & \rho_k \neq 0 \\ k \neq s, 2s, \dots, Qs & \rho_k = 0 \end{cases}$$

Es decir, un proceso de medias móviles de orden Q , tendrá Q coeficientes de autocorrelación diferentes de cero, correspondientes a los Q primeros retardos estacionales, $k = s, 2s, \dots, Qs$, y el resto son cero. La función de autocorrelación es, por lo tanto, finita, truncándose a partir del Q -ésimo retardo estacional, es decir, para $k = Qs$.

Modelo $AR(1)_4$:

$$(1 - \Phi L^4) Y_t = a_t \quad Y_t = \Phi Y_{t-4} + a_t$$

En este proceso, el valor de Y en el momento t depende de la perturbación contemporánea y del valor pasado de la serie Y en el mismo trimestre del año anterior. Como es un modelo autorregresivo finito es invertible para cualquier valor de Φ y su media es cero. El proceso será estacionario sí y solo sí las cuatro raíces del polinomio autorregresivo tienen modulo fuera del círculo unidad.

Las autocovarianzas son:

$$\begin{aligned} \gamma_0 &= V(Y_t) = E[Y_t]^2 = E[(\Phi Y_{t-4} + a_t)^2] = \Phi^2 \gamma_0 + \sigma^2 = \frac{\sigma^2}{1 - \Phi^2} \\ \gamma_1 &= Cov(Y_t Y_{t-1}) = E[Y_t Y_{t-1}] = E[(\Phi Y_{t-4} + a_t) Y_{t-1}] = 0 \\ \gamma_2 &= Cov(Y_t Y_{t-2}) = E[Y_t Y_{t-2}] = E[(\Phi Y_{t-4} + a_t) Y_{t-2}] = 0 \\ \gamma_3 &= Cov(Y_t Y_{t-3}) = E[Y_t Y_{t-3}] = E[(\Phi Y_{t-4} + a_t) Y_{t-3}] = 0 \\ \gamma_4 &= Cov(Y_t Y_{t-4}) = E[Y_t Y_{t-4}] = E[(\Phi Y_{t-4} + a_t) Y_{t-4}] = \Phi \gamma_0 \\ \gamma_5 &= Cov(Y_t Y_{t-5}) = E[Y_t Y_{t-5}] = E[(\Phi Y_{t-4} + a_t) Y_{t-5}] = 0 \\ \gamma_6 &= Cov(Y_t Y_{t-6}) = E[Y_t Y_{t-6}] = E[(\Phi Y_{t-4} + a_t) Y_{t-6}] = 0 \\ \gamma_7 &= Cov(Y_t Y_{t-7}) = E[Y_t Y_{t-7}] = E[(\Phi Y_{t-4} + a_t) Y_{t-7}] = 0 \\ \gamma_8 &= Cov(Y_t Y_{t-8}) = E[Y_t Y_{t-8}] = E[(\Phi Y_{t-4} + a_t) Y_{t-8}] = \Phi \gamma_4 = \Phi^2 \gamma_0 \\ \dots &= \dots \end{aligned}$$

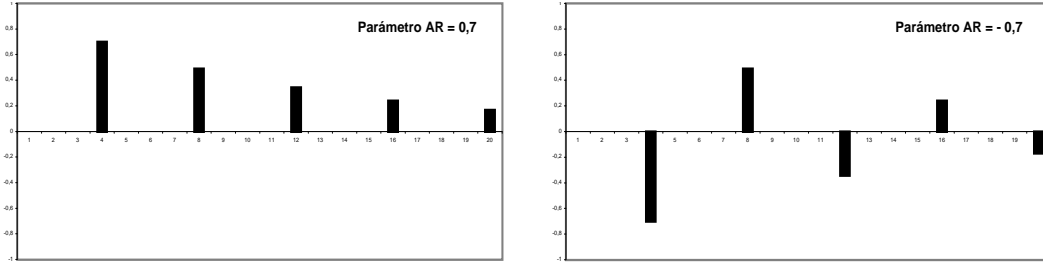
De lo que se deduce que la función de autocovarianzas γ , y, como consecuencia, la función de autocorrelación son:

Retardo	FACV	FAC
$k = 0$	$\gamma_0 = \frac{\sigma^2}{1 - \Phi^2}$	$\rho_0 = 1$
$k = 4j, j = 1, 2, \dots$	$\gamma_k = \Phi^{k/4} \gamma_0$	$\rho_k = \Phi^{k/4}$
$k \neq 4j, j = 1, 2, \dots$	$\gamma_k = 0$	$\rho_k = 0$

La función de autocorrelación es, por lo tanto, infinita y decrece como una función exponencial del parámetro Φ , siendo nulos los coeficientes de autocorrelación asociados a valores de k no estacionales, $k \neq j4$. Como se puede observar en el gráfico 7.2 si el parámetro es positivo, todos los coeficientes de autocorrelación no nulos son positivos, mientras que si Φ es negativo los coeficientes van alternando de signo.

El modelo $AR(P)_s$, $\Phi_P(L^s) Y_t = a_t$, se puede escribir en forma extendida como:

$$Y_t = \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \dots + \Phi_P Y_{t-Ps} + a_t$$

Gráfico 7.2: Función de autocorrelación. Modelo $AR(1)_4$ 

A partir de los resultados para el modelo $AR(1)_s$ se puede generalizar que su función de autocorrelación es infinita y decrece exponencialmente o en forma onda seno-coseno amortiguada, pero con valores no nulos únicamente para los retardos estacionales, $k = s, 2s, 3s, \dots$

Modelo $ARMA(1, 1)_4$.

$$(1 - \Phi L^4) Y_t = (1 - \Theta L^4) a_t \quad Y_t = \Phi Y_{t-4} + a_t - \Theta a_{t-4}$$

En los procesos generados por este modelo ARMA estacional puro el valor de Y en el momento t depende del valor pasado de Y del mismo trimestre del año anterior, de la perturbación contemporánea y su retardo correspondiente al mismo trimestre de hace un año. Este modelo será estacionario cuando las cuatro raíces del polinomio autorregresivo tengan módulo fuera del círculo unidad y será invertible cuando las cuatro raíces del polinomio de medias móviles tengan módulo fuera del círculo unidad.

Es fácil demostrar que la media de este proceso estacionario es cero y sus autocovarianzas

$$\begin{aligned} \gamma_0 &= V(Y_t) = E[Y_t]^2 = E[\Phi Y_{t-4} + a_t - \Theta a_{t-4}]^2 = \\ &= \Phi^2 \gamma_0 + \sigma^2 + \Theta^2 \sigma^2 - 2\Phi\Theta\sigma^2 \\ \gamma_0 &= \frac{\sigma^2(1 + \Theta^2 - 2\Phi\Theta)}{1 - \Phi^2} \\ \gamma_1 &= Cov(Y_t Y_{t-1}) = E[Y_t Y_{t-1}] = E[(\Phi Y_{t-4} + a_t - \Theta a_{t-4}) Y_{t-1}] = 0 \\ \gamma_2 &= Cov(Y_t Y_{t-2}) = E[Y_t Y_{t-2}] = E[(\Phi Y_{t-4} + a_t - \Theta a_{t-4}) Y_{t-2}] = 0 \\ \gamma_3 &= Cov(Y_t Y_{t-3}) = E[Y_t Y_{t-3}] = E[(\Phi Y_{t-4} + a_t - \Theta a_{t-4}) Y_{t-3}] = 0 \\ \gamma_4 &= Cov(Y_t Y_{t-4}) = E[Y_t Y_{t-4}] = E[(\Phi Y_{t-4} + a_t - \Theta a_{t-4}) Y_{t-4}] = \Phi \gamma_0 - \Theta \sigma^2 \\ \gamma_5 &= Cov(Y_t Y_{t-5}) = E[Y_t Y_{t-5}] = E[(\Phi Y_{t-4} + a_t - \Theta a_{t-4}) Y_{t-5}] = 0 \\ \gamma_6 &= Cov(Y_t Y_{t-6}) = E[Y_t Y_{t-6}] = \gamma_7 = Cov(Y_t Y_{t-7}) = E[Y_t Y_{t-7}] = 0 \\ \gamma_8 &= Cov(Y_t Y_{t-8}) = E[Y_t Y_{t-8}] = E[(\Phi Y_{t-4} + a_t - \Theta a_{t-4}) Y_{t-8}] = \Phi \gamma_4 \\ \dots &= \dots \end{aligned}$$

De donde se deducen la función de autocovarianzas y de autocorrelación siguientes:

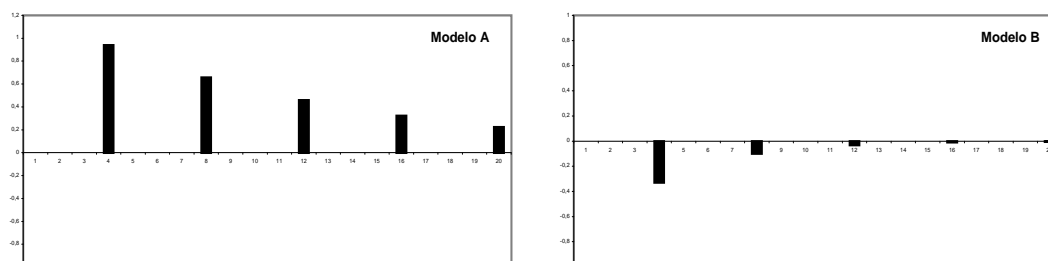
Retardo	FACV	FAC
$k = 0$	$\gamma_0 = \frac{(1 + \Theta^2 - 2\Phi\Theta) \sigma^2}{1 - \Phi^2}$	$\rho_0 = 1$
$k = 4$	$\gamma_4 = \Phi \gamma_0 - \Theta \sigma^2$	$\rho_4 = \frac{(\phi - \Theta)(1 - \Phi\Theta)}{1 + \Theta^2 - 2\Phi\Theta}$
$k = 4j, j = 2, 3, \dots$	$\gamma_k = \Phi \gamma_{k-4}$	$\rho_k = \Phi \rho_{k-4}$
$k \neq 4j, j = 1, 2, \dots$	$\gamma_k = 0$	$\rho_k = 0$

En el gráfico 7.3 se representan las funciones de autocorrelación para los siguientes modelos $ARMA(1,1)_4$:

$$(A) Y_t = 0,7Y_{t-4} + a_t + 0,4a_{t-4}$$

$$(B) Y_t = 0,3Y_{t-4} + a_t - 0,8a_{t-4}$$

Gráfico 7.3: Función de autocorrelación. Modelos $ARMA(1,1)_4$



En ambas figuras de este gráfico se observa que la función es infinita y decrece exponencialmente tomando valores no nulos únicamente en los retardos estacionales. La estructura que puede tomar esta FAC infinita es más compleja que la del $AR(1)_4$ porque el primer coeficiente estacional de la FAC depende tanto de la parte autorregresiva como medias móviles del modelo. Luego, a partir del segundo retardo estacional la trayectoria de la FAC depende sólo de la parte autorregresiva. Por eso, podemos tener FAC con todos los coeficientes positivos o todos negativos, o alternando de signo.

A partir de estos resultados se puede generalizar que la función de autocorrelación de un modelo $ARMA(P, Q)_s$:

$$(1 - \Phi_P L^s) Y_t = (1 - \Theta_Q L^s) a_t$$

$$(1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_P L^{Ps}) Y_t = (1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_Q L^{Qs}) a_t$$

$$Y_t = \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \dots + \Phi_P Y_{t-Ps} + a_t \Theta_1 a_{t-s} + \Theta_2 a_{t-2s} + \dots + \Theta_Q a_{t-Qs}$$

La FAC de este modelo es infinita y decrece exponencialmente pero con valores no nulos únicamente para los retardos estacionales, $k = s, 2s, \dots$. Los Q primeros coeficientes de autocorrelación $\rho_k, k = s, 2s, 3s, \dots, Qs$ dependen de los parámetros autorregresivos y de medias móviles; a partir de $k = s(Q+1)$ las autocorrelaciones se derivan de la estructura autorregresiva exclusivamente, por lo que decrecerán exponencialmente o en forma de onda seno-coseno amortiguada dependiendo de si las s raíces del polinomio autorregresivo $(1 - \Phi_P L^s)$ son reales o complejas.

En lo que se refiere a la estructura de la función de autocorrelación parcial de los modelos estacionales puros que estamos estudiando es análoga a la estudiada en el capítulo 3 para los modelos $ARMA(p, q)$ no estacionales, pero en estos modelos aparece sólo en los retardos estacionales. Así, se puede demostrar que la FACP de un modelo $AR(P)_s$ es finita, toma valores no nulos únicamente en los P primeros retardos estacionales, es decir, y se trunca a partir de $k = Ps$. La FACP de un modelo $MA(Q)_s$ es infinita y decrece exponencialmente tomando valores no nulos únicamente para los retardos no estacionales. Por último, la FACP de un modelo $ARMA(P, Q)_s$ será también infinita con valores no nulos sólo en los retardos estacionales y tal que los coeficientes de autocorrelación parcial, p_k , asociados a los P primeros retardos estacionales dependen de los parámetros autorregresivos y de medias móviles y a partir de $k = (P+1)s$, su estructura dependerá de la parte de medias móviles.

Por lo tanto, podemos concluir que la estructura estacional, es decir, la dependencia temporal de las observaciones separadas por un año, aparece exclusivamente en los coeficientes de autocorrelación simple y parcial asociados a los retardos estacionales, $k = s, 2s, 3s, \dots$, siendo nulos para el resto de los valores $k \neq s, 2s, 3s, \dots$.

7.2. Modelos ARMA estacionales no estacionarios

Hasta el momento hemos supuesto que la serie Y_t es estacionaria, quizás tras algún tipo de transformación. En la práctica, dado que por estacionalidad entendemos una pauta regular de comportamiento cíclico de periodo 1 año de la serie que implica que, en promedio, cada mes tiene un comportamiento diferente, las series estacionales suelen presentar problemas de falta de estacionariedad en media o lo que es lo mismo cambios sistemáticos en el nivel de la serie.

Si la estacionalidad fuese siempre exactamente periódica, $S_t = S_{t-s}$ se podría eliminar de la serie como un componente determinista previamente estimado (especificado, por ejemplo, en función de variables ficticias estacionales). Ahora bien, este esquema estacional es muy rígido porque exige que los factores estacionales permanezcan constantes a lo largo del tiempo. Sin embargo, la mayoría de las series no se comportan de una forma tan regular y, en general, el componente estacional será estocástico y estará correlacionado con la tendencia. Por lo tanto, un esquema de trabajo más flexible es suponer que la estacionalidad es sólo aproximadamente constante y que evoluciona estocásticamente:

$$Y_t = S_t + \eta_t \quad \text{con} \quad S_t = S_{t-s} + \nu_t$$

donde ν_t es un proceso estocástico estacionario.

Para solucionar la no estacionariedad en media que genera el comportamiento estacional, se toman diferencias entre observaciones separadas s periodos, que llamaremos *diferencias estacio-*

nales:

$$\Delta_s Y_t = (1 - L^s) Y_t = S_t + \eta_t - S_{t-s} - \eta_{t-s} = \nu_t + \eta_t - \eta_{t-s}$$

de forma que $\Delta_s Y_t \sim MA(1)_s$ estacionario e invertible.

En conclusión, el operador Δ_s convierte un proceso estacional no estacionario en estacionario, de forma análoga a como de un proceso no estacionario en media obteníamos una serie estacionaria aplicando el operador de diferencias de orden 1 o *regular*, $(1 - L)$.

La formulación general de un modelo para una serie estacional pura no estacionaria es:

$$\Phi_P(L^s) \Delta_s^D Y_t = \Theta_Q(L^s) a_t \quad (7.1)$$

En teoría, D , el número de las diferencias estacionales que se han de aplicar para convertir a la serie en estacionaria, puede tomar cualquier valor dependiendo de las características de la serie, aunque en la práctica nunca es superior a 1.

El modelo (7.1) es un modelo estacional puro que se denomina $ARIMA(P, D, Q)_s$, donde P es el orden del polinomio autorregresivo estacional estacionario, Q es el orden del polinomio medias móviles estacional estacionario y D es el número de diferencias estacionales $(1 - L^s)$ que es necesario aplicar a la serie Y_t para que sea estacionaria, es decir, el orden de integración estacional de la serie. De hecho, si la serie Y_t sigue el modelo (7.1) se dice que es integrada estacional de orden D , $Y_t \sim I(D)_s$.

El modelo (7.1) no es estacionario, no tiene una media constante y, al igual que para los modelos lineales no estacionarios en la parte regular, su función de autocorrelación no va a decrecer rápidamente hacia cero, sino lentamente.

En el gráfico 7.4 se representan dos series de 100 observaciones, así como sus correspondientes correlogramas, generadas a partir de los siguientes modelos:

$$(A) \quad Y_t = a_t - 0,4 a_{t-4}$$

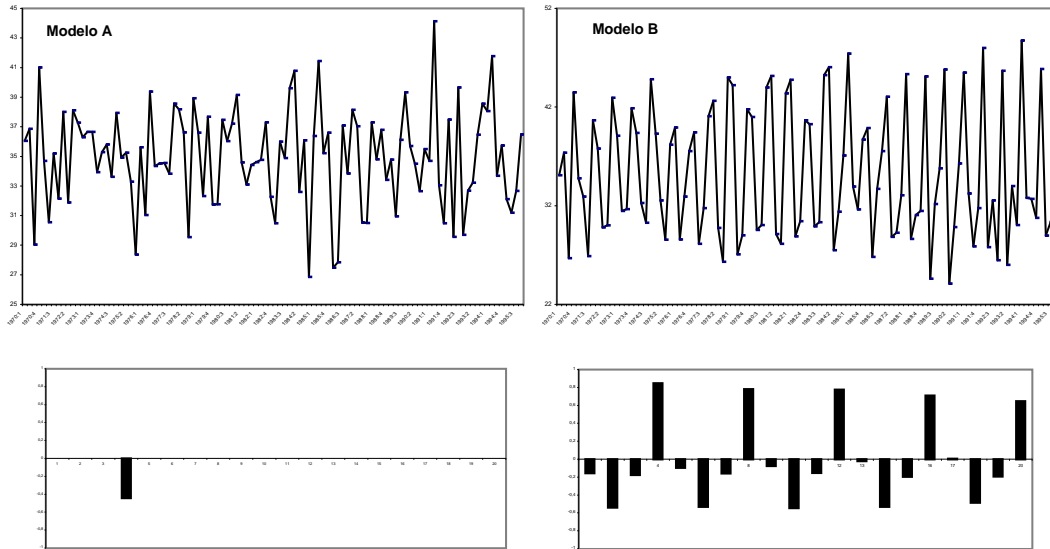
$$(B) \quad (1 - L^4) Y_t = a_t - 0,4 a_{t-4}$$

Aunque ambos modelos representan series con estructura de dependencia temporal entre observaciones separadas un año, el modelo (A) es un $MA(1)_4$ estacionario y el modelo (B) es un $ARIMA(0, 1, 1)_4$, lo que implica que la serie Y_t no es estacionaria y hay que tomar una diferencia estacional de orden 4 para transformarla en estacionaria.

El hecho de que la estructura estacional sea estacionaria o no se refleja en las características de cada proceso mostradas en las figuras del gráfico 7.4. En cuanto a las realizaciones de ambos procesos, ambas oscilan en torno a una media constante con un comportamiento cíclico de periodo anual. Pero mientras, la evolución cíclica estacionaria del modelo (A) es más irregular y aleatoria, la mostrada por el modelo (B) es más persistente y sistemática, reflejando claramente el hecho de que cada mes tiene una media diferente. Analizando los correlogramas se puede concluir que el correspondiente al modelo (A) decrece rápidamente hacia cero como corresponde a un modelo estacionario (de hecho se trunca en el retardo 4, primer retardo estacional), mientras que si nos fijamos en los coeficientes asociados a los retardos estacionales del correlograma del modelo (B), se observa que decrecen muy lentamente. Además, la no estacionariedad del ciclo estacional hace que la estructura cíclica sistemática de la serie se reproduzca en el correlograma,

apareciendo correlaciones negativas en los retardos $js/2, j = 1, 2, \dots$ que decrecen lentamente y recogen la relación lineal inversa entre los valles y los picos del ciclo, etc.

Gráfico 7.4: Modelos estacionales estacionarios vs. no estacionarios



7.3. Modelos ARIMA estacionales multiplicativos

En general, las series económicas además del componente estacional, presentan otros componentes, como tendencia, etc. de los que, por otro lado, el componente estacional no es independiente. Se trata de proponer un modelo *ARIMA* que recoja no solo las relaciones entre periodos (años) sino también las relaciones dentro de los periodos (intraanuales) y la interacción entre ambas estructuras.

Consideremos una serie estacional Y_t con periodo s conocido, por ejemplo, una serie mensual observada durante N años, de forma que en total contamos con $T = 12N$ observaciones. Si la serie es estacional podemos dividirla en 12 subseries, una por mes, de N observaciones que denotaremos:

$$y_\tau^{(1)}, y_\tau^{(2)}, y_\tau^{(3)}, \dots, y_\tau^{(12)}, \quad \tau = 1, 2, \dots, N$$

La relación entre estas subseries y la serie de partida es:

$$y_\tau^{(j)} = Y_{j+12(\tau-1)} = Y_t \quad \tau = 1, 2, \dots, N \quad j = 1, 2, \dots, 12 \quad (7.2)$$

Cada una de estas doce subseries no presenta comportamiento estacional por lo que podemos representarlas mediante los modelos *ARIMA*(p, d, q) derivados en el capítulo 4. Supongamos que el modelo *ARIMA* adecuado para las doce subseries $y_\tau^{(j)}$ es el mismo:

$$(1 - \Phi_1 L - \dots - \Phi_p L^p) (1 - L)^D y_\tau^{(j)} = (1 - \Theta_1 L - \dots - \Theta_q L^q) u_\tau^{(j)} \quad \tau = 1, 2, \dots, N \quad (7.3)$$

Si hay estacionalidad en la serie de partida Y_t se cumple que $D \geq 1$. Notése que si $D = 0$ y, por lo tanto, las series $y_\tau^{(j)}$ fueran estacionarias, todos los modelos (7.3), al ser el mismo, tendrían la misma media, lo que es incompatible con el supuesto de estacionalidad que implica que la media de cada mes, o sea de cada subserie $y_\tau^{(j)}, j = 1, 2, \dots, 12$, es diferente.

Los modelos para las doce subseries, al ser todos iguales, se pueden escribir conjuntamente en función de la serie de partida Y_t , teniendo en cuenta que, dada la relación (7.2):

$$L y_\tau^{(j)} = y_{\tau-1}^{(j)} = Y_{j+12(\tau-2)} = Y_{j-12+12(\tau-1)} = L^{12} Y_{j+12(\tau-1)}$$

Lo que implica que aplicar el operador L a $y_\tau^{(j)}$ es equivalente a aplicar el operador L^{12} a la serie original $Y_{j+12(\tau-1)}$.

Además habrá que definir una serie de ruido común para las doce subseries, α_t , asignando a cada mes t el ruido del modelo univariante correspondiente a dicho mes. En consecuencia, α_t se obtendría a partir de las doce subseries $u_\tau^{(j)}$ como sigue:

$$u_\tau^{(j)} = \alpha_{j+12(\tau-1)} \quad \tau = 1, 2, \dots, N$$

Aplicando ambos resultados al modelo (7.3), se obtiene el modelo conjunto para toda la serie mensual, Y_t :

$$(1 - \Phi_1 L^{12} - \dots - \Phi_p L^{12p}) (1 - L^{12})^D y_\tau^{(j)} = (1 - \Theta_1 L^{12} - \dots - \Theta_q L^{12q}) \alpha_t, \quad t = 1, 2, \dots, 12N$$

Para cada uno de los modelos (7.3) las series $u_\tau^{(j)}$ son, por construcción, ruido blanco, pero la serie conjunta, $\alpha_t, t = 1, 2, \dots, T$, no tiene por qué serlo, en general, ya que la mayoría de las veces existirá dependencia entre las observaciones contiguas que, como no ha sido todavía tenida en cuenta, quedará recogida en esta serie de ruido. Es lo que se denomina *estructura regular* o estructura asociada a los intervalos naturales de medida de la serie (meses en nuestro caso). Suponiendo que α_t sigue el modelo *ARIMA* no estacional siguiente:

$$\phi_p(L) (1 - L)^d \alpha_t = \theta_q(L) a_t \quad a_t \sim RBN(0, \sigma^2)$$

Sustituyendo este modelo en el general para Y_t se obtiene el modelo completo para la serie observada:

$$\Phi_P(L^s) \phi_p(L) \Delta^d \Delta_s^D Y_t = \theta_q(L) \Theta_Q(L^s) a_t \quad (7.4)$$

donde $\phi_p(L)$ y $\theta_q(L)$ son los polinomios autorregresivos y medias móviles de la parte regular y d es el orden de integración de la parte regular; mientras que $\Phi_P(L)$ y $\Theta_Q(L)$ son los polinomios autorregresivos y medias móviles de la parte estacional y D es el orden de integración de la parte estacional.

Los *modelos ARIMA estacionales multiplicativos*, $ARIMA(p, d, q) \times (P, D, Q)_s$. Estos modelos son flexibles en el sentido de que especifican estacionalidades estocásticas, tendencias estocásticas y además recogen la posible interacción entre ambos componentes.

Esta clase de modelos, como hemos visto, se basa en la hipótesis central de que la relación de dependencia estacional es la misma para todos los periodos. Este supuesto no se tiene por qué cumplir siempre pero, de todas maneras, estos modelos son capaces de representar muchos fenómenos

estacionales que encontramos en la práctica de una forma muy simple. Supongamos, por ejemplo, que Y_t es una serie estacional mensual que sigue el siguiente modelo $ARIMA(0, 1, 1) \times (0, 1, 1)_{12}$, denominado *Modelo de Líneas Aéreas* y que es muy utilizado en la práctica,

$$\Delta\Delta_{12}Y_t = (1 - \theta_1 L)(1 - \Theta_{12}L^{12})a_t = (1 - \theta_1 L - \Theta_{12}L^{12} + \theta_1 \Theta_{12}L^{13})a_t$$

Notése que este modelo para la transformación estacionaria, es un $MA(13)$ con restricciones, de forma que solo contiene dos parámetros desconocidos, en vez de los trece de un $MA(13)$ general. Lo que muestra que los modelos $ARIMA(p, d, q) \times (P, D, Q)_s$ son modelos muy parsimoniosos.

A continuación, se derivan las características de algunos modelos sencillos bajo el supuesto de que Y_t es una serie trimestral con comportamiento estacional y estacionaria, posiblemente tras aplicar los operadores de diferencias oportunos.

Modelo $ARMA(0, 1) \times (0, 1)_4$.

Sea Y_t un proceso estocástico estacionario de media cero que sigue el modelo:

$$\begin{aligned} Y_t &= (1 - \theta_1 L)(1 - \Theta_4 L^4)a_t \\ Y_t &= a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5} \end{aligned}$$

que se puede considerar la transformación estacionaria del *Modelo de Líneas Aéreas* para una serie trimestral.

La media es cero y sus autocovarianzas son:

$$\begin{aligned} \gamma_0 &= V(Y_t) = E[Y_t - 0]^2 = E[a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5}]^2 \\ &= (1 + \theta_1^2 + \Theta_4^2 + \theta_1^2 \Theta_4^2) \sigma^2 = (1 + \theta_1^2) (1 + \Theta_4^2) \sigma^2 \end{aligned}$$

$$\begin{aligned} \gamma_1 &= Cov(Y_t Y_{t-1}) = E[(Y_t - 0)(Y_{t-1} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5})(a_{t-1} - \theta_1 a_{t-2} - \Theta_4 a_{t-5} + \theta_1 \Theta_4 a_{t-6})] \\ &= (-\theta_1 - \theta_1 \Theta_4^2) \sigma^2 = -\theta_1 (1 + \Theta_4^2) \sigma^2 \end{aligned}$$

$$\begin{aligned} \gamma_2 &= Cov(Y_t Y_{t-2}) = E[(Y_t - 0)(Y_{t-2} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5})(a_{t-2} - \theta_1 a_{t-3} - \Theta_4 a_{t-6} + \theta_1 \Theta_4 a_{t-7})] = 0 \end{aligned}$$

$$\begin{aligned} \gamma_3 &= Cov(Y_t Y_{t-3}) = E[(Y_t - 0)(Y_{t-3} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5})(a_{t-3} - \theta_1 a_{t-4} - \Theta_4 a_{t-7} + \theta_1 \Theta_4 a_{t-8})] \\ &= \theta_1 \Theta_4 \sigma^2 \end{aligned}$$

$$\begin{aligned} \gamma_4 &= Cov(Y_t Y_{t-4}) = E[(Y_t - 0)(Y_{t-4} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5})(a_{t-4} - \theta_1 a_{t-5} - \Theta_4 a_{t-8} + \theta_1 \Theta_4 a_{t-9})] \\ &= (-\Theta_4 - \theta_1^2 \Theta_4) \sigma^2 = -\Theta_4 (1 + \theta_1^2) \sigma^2 \end{aligned}$$

$$\begin{aligned}
\gamma_5 &= Cov(Y_t Y_{t-5}) = E[(Y_t - 0)(Y_{t-5} - 0)] = \\
&= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5})(a_{t-5} - \theta_1 a_{t-6} - \Theta_4 a_{t-9} + \theta_1 \Theta_4 a_{t-10})] \\
&= \theta_1 \Theta_4 \sigma^2
\end{aligned}$$

$$\begin{aligned}
\gamma_k &= Cov(Y_t Y_{t-k}) = E[(Y_t - 0)(Y_{t-k} - 0)] = \\
&= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5}) \\
&\quad (a_{t-k} - \theta_1 a_{t-k-1} - \Theta_4 a_{t-k-4} + \theta_1 \Theta_4 a_{t-k-4-1})] = 0 \quad \forall k > 5
\end{aligned}$$

Las funciones de autocovarianzas y la de autocorrelación son:

Retardo	FACV	FAC
$k = 0$	$\gamma_0 = (1 + \theta_1^2)(1 + \Theta_4^2) \sigma^2$	$\rho_0 = 1$
$k = 1$	$\gamma_1 = -\theta_1 (1 + \Theta_4^2) \sigma^2$	$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2}$
$k = 2$	$\gamma_2 = 0$	$\rho_2 = 0$
$k = 3$	$\gamma_3 = \theta_1 \Theta_4 \sigma^2$	$\rho_3 = \frac{\theta_1 \Theta_4}{(1 + \theta_1^2)(1 + \Theta_4^2)}$
$k = 4$	$\gamma_4 = -\Theta_4 (1 + \theta_1^2) \sigma^2$	$\rho_4 = \frac{-\Theta_4}{1 + \Theta_4^2}$
$k = 5$	$\gamma_5 = \theta_1 \Theta_4 \sigma^2$	$\rho_5 = \frac{\theta_1 \Theta_4}{(1 + \theta_1^2)(1 + \Theta_4^2)}$
$k > 5$	$\gamma_k = 0$	$\rho_k = 0$

Modelo ARMA(1,0) × (0,1)₄.

Sea Y_t un proceso estocástico estacionario que sigue el modelo:

$$(1 - \phi L) Y_t = (1 - \Theta_4 L^4) a_t \quad Y_t = \phi Y_{t-1} + a_t - \Theta_4 a_{t-4}$$

La media es cero y su varianza es:

$$\begin{aligned}
\gamma_0 &= V(Y_t) = E[Y_t - 0]^2 = E[\phi Y_{t-1} + a_t - \Theta_4 a_{t-4}]^2 \\
&= \phi^2 \gamma_0 + \sigma^2 (1 + \Theta_4^2) - 2\phi \Theta_4 E(Y_{t-1} a_{t-4})
\end{aligned}$$

Es necesario calcular la covarianza entre a_{t-4} e Y_{t-1} :

$$\begin{aligned}
E(Y_{t-1} a_{t-4}) &= E[\phi Y_{t-2} + a_{t-1} - \Theta_4 a_{t-5}] a_{t-4} = \phi E(Y_{t-2} a_{t-4}) \\
E(Y_{t-2} a_{t-4}) &= E[\phi Y_{t-3} + a_{t-2} - \Theta_4 a_{t-6}] a_{t-4} = \phi E(Y_{t-3} a_{t-4})
\end{aligned}$$

$$\begin{aligned} E(Y_{t-3} a_{t-4}) &= E[\phi Y_{t-4} + a_{t-3} - \Theta_4 a_{t-7}] a_{t-4} = \phi E(Y_{t-4} a_{t-4}) \\ E(Y_{t-4} a_{t-4}) &= E[\phi Y_{t-5} + a_{t-4} - \Theta_4 a_{t-8}] a_{t-4} = \sigma^2 \end{aligned}$$

Por lo que,

$$E(Y_{t-1} a_{t-4}) = \phi^3 \sigma^2$$

y la varianza del proceso es:

$$\gamma_0 = \frac{\sigma^2 (1 + \Theta_4^2 - 2\phi^4 \Theta_4)}{1 - \phi^2}$$

El resto de las autocovarianzas vienen dadas por:

$$\begin{aligned} \gamma_1 &= Cov(Y_t Y_{t-1}) = E[(Y_t Y_{t-1})] = E[(\phi Y_{t-1} + a_t - \Theta_4 a_{t-4}) Y_{t-1}] = \phi \gamma_0 - \Theta_4 \phi^3 \sigma^2 \\ \gamma_2 &= Cov(Y_t Y_{t-2}) = E[(Y_t Y_{t-2})] = E[(\phi Y_{t-1} + a_t - \Theta_4 a_{t-4}) Y_{t-2}] = \phi \gamma_1 - \Theta_4 \phi^2 \sigma^2 \\ \gamma_3 &= Cov(Y_t Y_{t-3}) = E[(Y_t Y_{t-3})] = E[(\phi Y_{t-1} + a_t - \Theta_4 a_{t-4}) Y_{t-3}] = \phi \gamma_2 - \Theta_4 \phi \sigma^2 \\ \gamma_4 &= Cov(Y_t Y_{t-4}) = E[(Y_t Y_{t-4})] = E[(\phi Y_{t-1} + a_t - \Theta_4 a_{t-4}) Y_{t-4}] = \phi \gamma_3 - \Theta_4 \sigma^2 \\ \gamma_5 &= Cov(Y_t Y_{t-5}) = E[(Y_t Y_{t-5})] = E[(\phi Y_{t-1} + a_t - \Theta_4 a_{t-4}) Y_{t-5}] = \phi \gamma_4 \\ \gamma_k &= Cov(Y_t Y_{t-k}) = E[(Y_t Y_{t-k})] = E[(\phi Y_{t-1} + a_t - \Theta_4 a_{t-4}) Y_{t-k}] = \phi \gamma_{k-1} \quad \forall k \geq 5 \end{aligned}$$

y las funciones de autocovarianzas y de autocorrelación son:

Retardo	FACV	FAC
$k = 0$	$\gamma_0 = \frac{\sigma^2 (1 + \Theta_4^2 - 2\phi^4 \Theta_4)}{1 - \phi^2}$	$\rho_0 = 1$
$k = 1$	$\gamma_1 = \phi \gamma_0 - \Theta_4 \phi^3 \sigma^2$	$\rho_1 = \phi - \frac{\Theta_4 \phi^3 \sigma^2}{\gamma_0}$
$k = 2$	$\gamma_2 = \phi \gamma_1 - \Theta_4 \phi^2 \sigma^2$	$\rho_2 = \phi \rho_1 - \frac{\Theta_4 \phi^2 \sigma^2}{\gamma_0}$
$k = 3$	$\gamma_3 = \phi \gamma_2 - \Theta_4 \phi \sigma^2$	$\rho_3 = \phi \rho_2 - \frac{\Theta_4 \phi \sigma^2}{\gamma_0}$
$k = 4$	$\gamma_4 = \phi \gamma_3 - \Theta_4 \sigma^2$	$\rho_4 = \phi \rho_3 - \frac{\Theta_4 \sigma^2}{\gamma_0}$
$k \geq 5$	$\gamma_k = \phi \gamma_{k-1}$	$\rho_k = \phi \rho_{k-1}$

Modelo $ARMA(0,1) \times (0,2)_4$.

Sea Y_t un proceso estocástico estacionario que sigue el modelo:

$$\begin{aligned} Y_t &= (1 - \theta_1 L)(1 - \Theta_4 L^4 - \Theta_8 L^8)a_t \\ Y_t &= a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9} \end{aligned}$$

La media es cero y sus autocovarianzas son:

$$\begin{aligned} \gamma_0 &= V(Y_t) = E[Y_t - 0]^2 = E[a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}]^2 \\ &= (1 + \theta_1^2 + \Theta_4^2 + \Theta_8^2 + \theta_1^2 \Theta_4^2 + \theta_1^2 \Theta_8^2) \sigma^2 = (1 + \theta_1^2) (1 + \Theta_4^2 + \Theta_8^2) \sigma^2 \\ \gamma_1 &= Cov(Y_t Y_{t-1}) = E[(Y_t - 0)(Y_{t-1} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-1} - \theta_1 a_{t-2} - \Theta_4 a_{t-5} - \Theta_8 a_{t-9} + \theta_1 \Theta_4 a_{t-6} + \theta_1 \Theta_8 a_{t-10})] = -\theta_1 (1 + \Theta_4^2 + \Theta_8^2) \sigma^2 \\ \gamma_2 &= Cov(Y_t Y_{t-2}) = E[(Y_t - 0)(Y_{t-2} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-2} - \theta_1 a_{t-3} - \Theta_4 a_{t-6} - \Theta_8 a_{t-10} + \theta_1 \Theta_4 a_{t-7} + \theta_1 \Theta_8 a_{t-11})] = 0 \\ \gamma_3 &= Cov(Y_t Y_{t-3}) = E[(Y_t - 0)(Y_{t-3} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-3} - \theta_1 a_{t-4} - \Theta_4 a_{t-7} - \Theta_8 a_{t-11} + \theta_1 \Theta_4 a_{t-8} + \theta_1 \Theta_8 a_{t-12})] = \theta_1 \Theta_4 (1 - \Theta_8) \sigma^2 \\ \gamma_4 &= Cov(Y_t Y_{t-4}) = E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-4} - \theta_1 a_{t-5} - \Theta_4 a_{t-8} - \Theta_8 a_{t-12} + \theta_1 \Theta_4 a_{t-9} + \theta_1 \Theta_8 a_{t-13})] = \\ &= (-\Theta_4 + \Theta_4 \Theta_8) (1 + \theta_1^2) \sigma^2 \\ \gamma_5 &= Cov(Y_t Y_{t-5}) = E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-5} - \theta_1 a_{t-6} - \Theta_4 a_{t-9} - \Theta_8 a_{t-13} + \theta_1 \Theta_4 a_{t-10} + \theta_1 \Theta_8 a_{t-14})] = \theta_1 \Theta_4 (1 - \Theta_8) \sigma^2 \\ \gamma_6 &= Cov(Y_t Y_{t-6}) = E[(Y_t - 0)(Y_{t-6} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-6} - \theta_1 a_{t-7} - \Theta_4 a_{t-10} - \Theta_8 a_{t-14} + \theta_1 \Theta_4 a_{t-11} + \theta_1 \Theta_8 a_{t-15})] = 0 \\ \gamma_7 &= Cov(Y_t Y_{t-7}) = E[(Y_t - 0)(Y_{t-7} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-7} - \theta_1 a_{t-8} - \Theta_4 a_{t-11} - \Theta_8 a_{t-15} + \theta_1 \Theta_4 a_{t-12} + \theta_1 \Theta_8 a_{t-16})] = \theta_1 \Theta_8 \sigma^2 \\ \gamma_8 &= Cov(Y_t Y_{t-8}) = E[(Y_t - 0)(Y_{t-8} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-8} - \theta_1 a_{t-9} - \Theta_4 a_{t-12} - \Theta_8 a_{t-16} + \theta_1 \Theta_4 a_{t-13} + \theta_1 \Theta_8 a_{t-17})] = -\Theta_8 (1 + \theta_1^2) \sigma^2 \end{aligned}$$

$$\begin{aligned}\gamma_9 &= \text{Cov}(Y_t Y_{t-9}) = E[(Y_t - 0)(Y_{t-9} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-9} - \theta_1 a_{t-10} - \Theta_4 a_{t-13} - \Theta_8 a_{t-17} + \theta_1 \Theta_4 a_{t-14} + \theta_1 \Theta_8 a_{t-18})] = \theta_1 \Theta_8 \sigma^2\end{aligned}$$

$$\begin{aligned}\gamma_{10} &= \text{Cov}(Y_t Y_{t-10}) = E[(Y_t - 0)(Y_{t-10} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} - \Theta_8 a_{t-8} + \theta_1 \Theta_4 a_{t-5} + \theta_1 \Theta_8 a_{t-9}) \\ &\quad (a_{t-10} - \theta_1 a_{t-11} - \Theta_4 a_{t-14} - \Theta_8 a_{t-18} + \theta_1 \Theta_4 a_{t-15} + \theta_1 \Theta_8 a_{t-19})] = \theta_1 \Theta_8 \sigma^2\end{aligned}$$

$$\begin{aligned}\gamma_k &= \text{Cov}(Y_t Y_{t-k}) = E[(Y_t - 0)(Y_{t-k} - 0)] = \\ &= E[(a_t - \theta_1 a_{t-1} - \Theta_4 a_{t-4} + \theta_1 \Theta_4 a_{t-5}) \\ &\quad (a_{t-k} - \theta_1 a_{t-k-1} - \Theta_4 a_{t-k-4} + \theta_1 \Theta_4 a_{t-k-4-1})] = 0 \quad \forall k > 10\end{aligned}$$

Las funciones de autocovarianzas y de autocorrelación son las siguientes:

Retardo	FACV	FAC
$k = 0$	$\gamma_0 = (1 + \theta_1^2)(1 + \Theta_4^2 + \Theta_8^2) \sigma^2$	$\rho_0 = 1$
$k = 1$	$\gamma_1 = (-\theta_1)(1 + \Theta_4^2 + \Theta_8^2) \sigma^2$	$\rho_1 = \frac{\theta_1}{1 + \theta_1^2}$
$k = 2$	$\gamma_2 = 0$	$\rho_2 = 0$
$k = 3$	$\gamma_3 = \theta_1 \Theta_4 (1 - \Theta_8) \sigma^2$	$\rho_3 = \frac{\theta_1 \Theta_4 (1 - \Theta_8)}{(1 + \theta_1^2)(1 + \Theta_4^2 + \Theta_8^2)}$
$k = 4$	$\gamma_4 = (1 + \theta_1^2)(-\Theta_4 + \Theta_4 \Theta_8) \sigma^2$	$\rho_4 = \frac{-\Theta_4(1 - \Theta_8)}{1 + \Theta_4^2 + \Theta_8^2}$
$k = 5$	$\gamma_5 = \theta_1 \Theta_4 (1 - \Theta_8) \sigma^2$	$\rho_5 = \frac{\theta_1 \Theta_4 (1 - \Theta_8)}{(1 + \theta_1^2)(1 + \Theta_4^2 + \Theta_8^2)}$
$k = 6$	$\gamma_6 = 0$	$\rho_6 = 0$
$k = 7$	$\gamma_7 = \theta_1 \Theta_4 \sigma^2$	$\rho_7 = \frac{\theta_1 \Theta_4}{(1 + \theta_1^2)(1 + \Theta_4^2 + \Theta_8^2)}$
$k = 8$	$\gamma_8 = -\Theta_8 (1 + \theta_1^2) \sigma^2$	$\rho_8 = \frac{-\Theta_8}{1 + \Theta_4^2 + \Theta_8^2}$
$k = 9$	$\gamma_9 = \theta_1 \Theta_4 \sigma^2$	$\rho_9 = \frac{\theta_1 \Theta_4}{(1 + \theta_1^2)(1 + \Theta_4^2 + \Theta_8^2)}$
$k > 9$	$\gamma_k = 0$	$\rho_k = 0$

En las cuatro figuras del gráfico 7.5 se muestran las funciones de autocorrelación de los siguientes modelos para la serie Y_t estacionaria posiblemente tras alguna transformación:

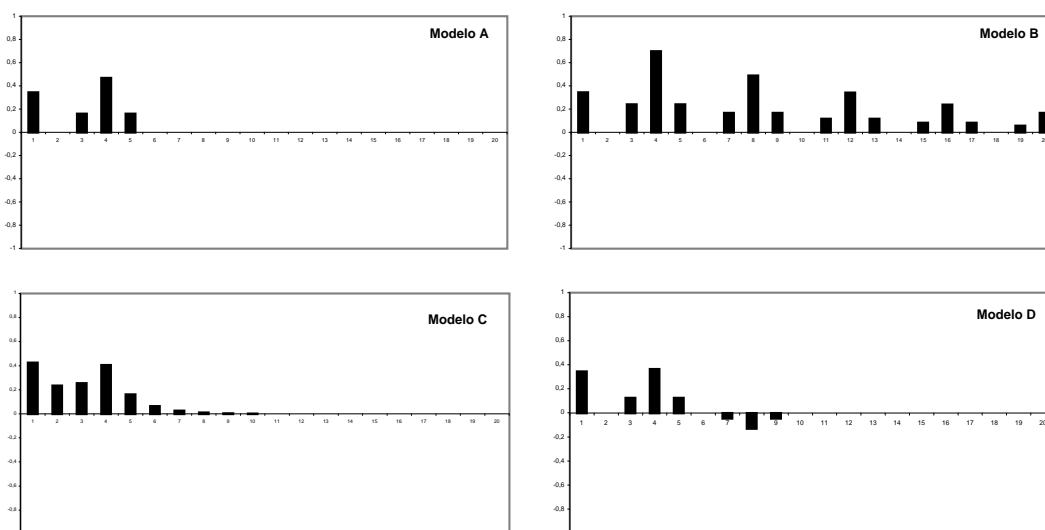
$$(A) \quad Y_t = (1 + 0,4L)(1 + 0,7L^4)a_t$$

$$(B) \quad (1 - 0,7L^4)Y_t = (1 + 0,4L)a_t$$

$$(C) \quad (1 - 0,4L)Y_t = (1 + 0,8L^4)a_t$$

$$(D) \quad Y_t = (1 + 0,7L - 0,2L^2)(1 + 0,4L^4)a_t$$

Gráfico 7.5: Función de autocorrelación. Modelos ARIMA multiplicativos.



Con los resultados derivados para los dos modelos presentados y la ayuda de la figura 7.5, se puede aproximar la estructura general de la FAC del proceso *ARIMA* estacional multiplicativo (7.4):

- El cálculo de la FAC puede llegar a ser muy laborioso, sobre todo si incluye parte autorregresiva y los órdenes de los polinomios autorregresivos regular y/o estacional son altos.
- La FAC del proceso *ARIMA* multiplicativo es una mezcla de las FAC correspondientes a su estructura regular y a su estructura estacional por separado, con términos de interacción porque ambas estructuras no son independientes.
- En los retardos más bajos, $k = 1, 2, \dots$, que recogen la estructura regular (observaciones contiguas) se reproduce la estructura marcada por los polinomios $\phi_p(L)$ y $\theta_q(L)$ de la parte regular. Así, si es un *MA*(1), como en los modelos (A), (B) y (D), se tiene que $\rho_1 \neq 0$ y luego el resto son cero. Mientras que si es un *AR*(1) como en el modelo (C) la parte regular es infinita y decrece exponencialmente hacia cero.
- En los retardos estacionales, $k = 2, 2s, 3s, \dots$, se representa la estructura estacional generada por los polinomios $\Phi_P(L)$ y $\Theta_Q(L)$. Para estructuras medias móviles estacionales

como las de los modelos (A) y (C) de forma $MA(1)_4$, se observa que $\rho_4 \neq 0$ y luego los siguientes coeficientes estacionales son cero, es decir, $\rho_8 = \rho_{12} = \dots = 0$ son cero. Para el modelo (D), $MA(2)_4$, sin embargo, los dos primeros coeficientes estacionales, ρ_4, ρ_8 , son distintos de cero y el resto son cero. En los modelos con estacionalidad autorregresiva la FAC es infinita para los retardos estacionales y decrece exponencialmente (véase modelo (B)).

- e) Alrededor de los retardos estacionales se observa la interacción entre la estructura *regular* y la *estacional* que se manifiesta en la repetición a ambos lados de cada retardo estacional de la función de autocorrelación simple de la parte regular.

En los modelos (A), (B) y (D) la parte regular es un $MA(1)$, por lo que $\rho_{s-1} = \rho_{s+1} \neq 0$. Si la parte regular fuera, en cambio, un $MA(2)$, aparecerían a ambos lados de cada retardo estacional no nulo 2 coeficientes distintos de cero. Por último, si la parte regular es un $AR(1)$ como en el modelo (C), a ambos lados de los retardos estacionales se observa el decrecimiento exponencial impuesto por esta estructura autorregresiva. Notése que para este modelo los coeficientes ρ_1, ρ_2, ρ_3 y ρ_4 recogen por un lado la estructura regular de la serie y la de interacción entre la parte regular y la parte estacional.

Para completar la descripción de las características de los procesos estacionales multiplicativos, se va a tratar la estructura de la función de autocorrelación parcial (FACP) de un modelo $ARMA(p, q) \times (P, Q)_s$:

- a) La FACP del proceso multiplicativo es una mezcla de las FACP correspondientes a su estructura regular y a su estructura estacional por separado, con un componente de interacción porque ambas no son independientes.
- b) En los retardos más bajos, $k = 1, 2, \dots$, que recogen la estructura regular se reproduce la estructura marcada por los polinomios $\phi_p(L)$ y $\theta_q(L)$ de la parte regular. Si es un $MA(q)$ o un $ARMA(p, q)$ presentaría el decaimiento rápido y continuado característico de estos modelos, mientras que si fuera un $AR(p)$ se truncaría en el retardo p .
- c) En los retardos estacionales, $k = s, 2s, 3s, \dots$, se representa la estructura estacional generada por los polinomios $\Phi_P(L)$ y $\Theta_Q(L)$. Si es un $MA(Q)_s$ o un $ARMA(P, Q)_s$ presentaría el decaimiento rápido y continuado característico de estos modelos, mientras que si fuera un $AR(P)$ se truncaría en el retardo $k = Ps$.
- d) Alrededor de los retardos estacionales se observa la interacción entre la parte *regular* y la *estacional* que se manifiesta como sigue:
- A la derecha de cada coeficiente estacional, $k = js + 1, js + 2, \dots$, se reproduce la FACP de la parte regular. Si el coeficiente estacional es positivo la FACP regular aparece invertida, mientras que si es negativo la FACP regular aparece con su propio signo.
 - A la izquierda de los coeficientes estacionales, $k = js - 1, js - 2, \dots$, se reproduce la FAC de la parte regular.

7.4. Modelización ARIMA estacional.

La metodología para construir el modelo $ARIMA(p, d, q)(P, D, Q)_s$ apropiado para la serie estacional Y_1, Y_2, \dots, Y_T consta de las fases de Identificación, Estimación, Validación y Predicción. Se ilustrará el proceso de construcción del modelo con una serie clásica en la literatura, *Pasajeros de Líneas Aéreas*, serie mensual observada desde Octubre de 1949 a Septiembre de 1961. Al final de esta sección se propone como ejercicio al lector la modelización de la serie trimestral índice de producción industrial austriaco. Los datos de ambas series se encuentran en el apéndice C.

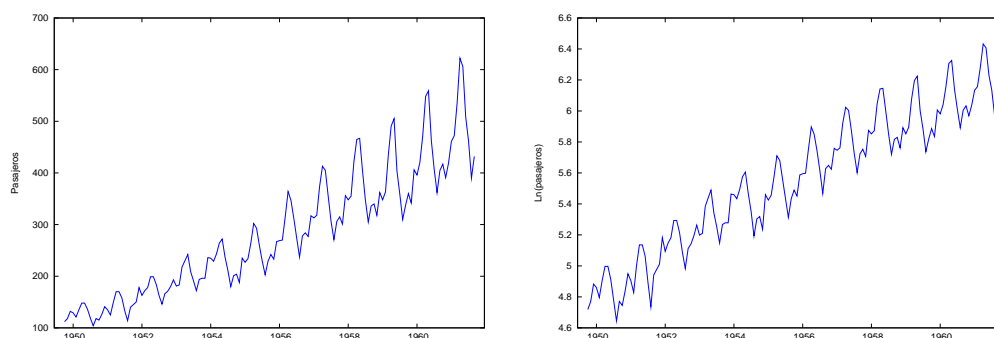
IDENTIFICACIÓN.

Se trata de proponer el/los modelos $ARIMA(p, d, q)(P, D, Q)_s$ que puedan representar la evolución de la serie Y_t . En primer lugar, se analiza la estacionariedad de la serie tanto en varianza como en media y, después, para la serie original o aquella transformación estacionaria de la misma, se seleccionan los ordenes (p, q) de la estructura regular estacionaria y $(P, Q)_s$ de la estructura estacional estacionaria. Por último, se estudia la necesidad de incluir o no un término independiente en el modelo.

A. Análisis de estacionariedad

El instrumento utilizado para estudiar la estacionariedad en varianza de la serie es su gráfico que permite observar si la variabilidad de la serie es homogénea a lo largo del tiempo o no. En las series con comportamiento estacional suele suceder que la variabilidad o amplitud del ciclo estacional crece con la tendencia. Este es el caso de la serie *Pasajeros de Líneas Aéreas* (Pasajeros) como muestra la figura de la izquierda del gráfico 7.6. Como la serie original no es estacionaria en varianza, tomamos la transformación logarítmica que presenta una variabilidad homogénea en toda la muestra (véase la figura de la derecha del gráfico 7.6). Por lo tanto, consideramos que la serie $\ln(\text{Pasajeros})$ es estacionaria en varianza.

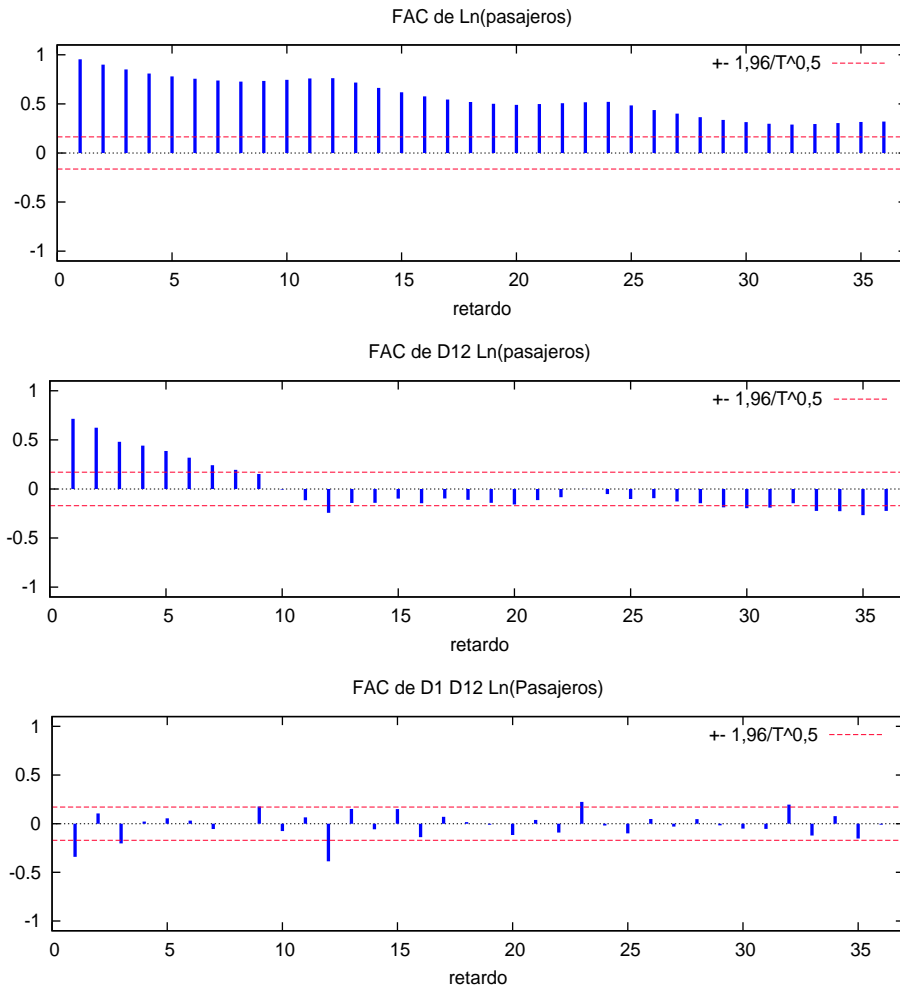
Gráfico 7.6: Pasajeros de Líneas Aéreas (datos originales y logaritmos).



El análisis de la estacionariedad en media de una serie se lleva a cabo, en general, utilizando tres instrumentos: el gráfico, la función de autocorrelación y los contrastes de raíces unitarias.

Inspeccionando en primer lugar el gráfico de la serie $\text{Ln}(\text{Pasajeros})_t$, se observa que no es estacionaria en media porque no oscila en torno a un nivel constante. De hecho, no es estacionaria porque, por un lado, su tendencia (comportamiento a largo plazo) es creciente y, por otro lado, presenta un ciclo estacional sistemático con los meses de verano en los valores máximos del ciclo y los de invierno en los mínimos, con lo que la media de cada mes es distinta. Por otro lado, la figura superior del gráfico 7.7 muestra como el correlograma de la serie $\text{Ln}(\text{Pasajeros})$ decrece lentamente hacia cero tanto en la estructura regular como en la estacional: noté el lento decrecimiento para los tres retardos estacionales: 12, 24 y 36.

Gráfico 7.7: Correlogramas de la $\text{Ln}(\text{Pasajeros})$ y algunas transformaciones.



Podemos concluir que la serie en logaritmos no es estacionaria en media ni en la estructura regular ni en la estacionalidad. Para buscar una transformación de la serie que sí sea estacionaria, se suele proceder a diferenciar la serie. Tomaremos, en primer lugar, diferencias estacionales para solucionar el problema de la no estacionariedad estacional.

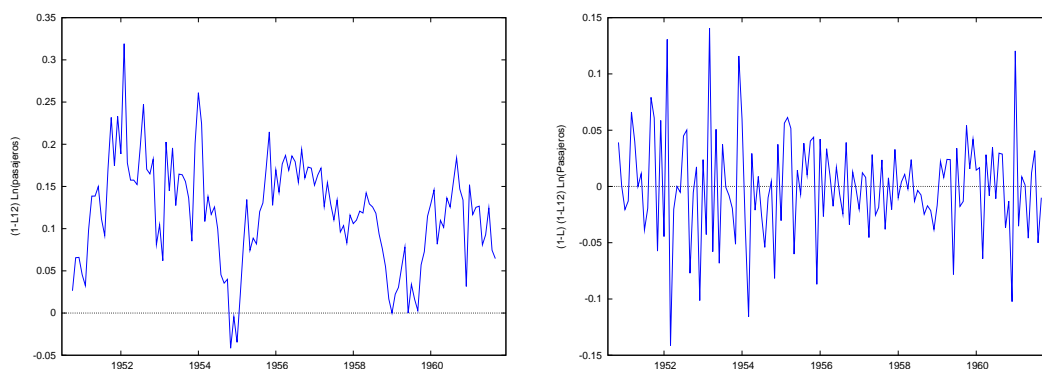
La figura izquierda del gráfico 7.8 muestra la serie $(1 - L^{12}) \text{Ln}(\text{Pasajeros})$. Ya no se

aprecia un comportamiento estacional sistemático y no se observan diferencias de media por cada mes. Por otro lado, hay que señalar también que ha desaparecido el comportamiento tendencial creciente de la serie. Esto es debido a que el operador de diferencias estacional se puede escribir como:

$$(1 - L^{12}) = (1 - L) S_{12}(L) = (1 - L)(1 + L + L^2 + \dots + L^{11})$$

lo que implica que al tomar diferencias estacionales, estamos eliminando, por un lado, la estacionalidad no estacionaria mediante el operador $S_{12}(L)$ que suma las observaciones de cada año, pero al estar aplicando a la vez el operador de diferencias regular, $(1 - L)$, se está actuando también sobre la no estacionariedad en tendencia.

Gráfico 7.8: Gráficos de algunas diferencias de la serie $\text{Ln}(\text{Pasajeros})$.



La serie $(1 - L^{12}) \text{Ln}(\text{Pasajeros})$ ya no crece sistemáticamente sino que parece oscilar en torno a un nivel pero con grandes rachas de observaciones por encima y por debajo del promedio. No parece el gráfico de una serie estacionaria en media sino el de una serie que va cambiando de nivel. El análisis de su correlograma (veáse la figura intermedia del gráfico 7.7) muestra un decrecimiento hacia cero que no es exponencial en la parte regular, aunque ya no hay problemas en los retardos estacionales. Este resultado parece indicar que la serie aún no es estacionaria en la parte regular.

Por último, el contraste ADF para la hipótesis nula de existencia de raíz unitaria en la serie $(1 - L^{12}) \text{Ln}(\text{Pasajeros})$, proporciona el siguiente resultado (veáse cuadro 7.1):

$$t_c = -2,71 > DF_{0,05} = -2,89$$

Por lo que no se rechaza la hipótesis nula de existencia de raíz unitaria.

Concluimos que la serie $(1 - L^{12}) \text{Ln}(\text{Pasajeros})$ no es estacionaria en la parte regular, por lo que se toma otra diferencia regular y se analiza la posible estacionariedad de la serie $(1 - L)(1 - L^{12}) \text{Ln}(\text{Pasajeros})$. Se observa que el gráfico de esta transformación de la serie (veáse la figura derecha del gráfico 7.8) oscila en torno a una media cero y que su correlograma (figura inferior del gráfico 7.7) decrece rápidamente hacia cero tanto en la estructura regular como en la estacional.

Cuadro 7.1: Contrastes de raíces unitarias. $(1 - L^{12})Ln(Pasajeros)$.

Contrastes aumentados de Dickey-Fuller, orden 12, para $(1-L^{12}) Ln(Pasajeros)$:
 tamaño muestral 119 hipótesis nula de raíz unitaria: $a = 1$
 contraste con constante
 modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$
 Coef. de autocorrelación de primer orden de e : -0,015
 valor estimado de $(a - 1)$: -0,249907
 Estadístico de contraste: $\tau_{a,c}(1) = -2,70958$
 valor p asintótico 0,07233

Cuadro 7.2: Contrastes de raíces unitarias. $(1 - L)(1 - L^{12})Ln(Pasajeros)$

Contrastes aumentados de Dickey-Fuller, orden 12, $(1-L)(1-L^{12}) Ln(Pasajeros)$:
 tamaño muestral 118 hipótesis nula de raíz unitaria: $a = 1$
 contraste sin constante
 modelo: $(1 - L)y = (a-1)*y(-1) + \dots + e$
 Coef. de autocorrelación de primer orden de e : -0,006
 valor estimado de $(a - 1)$: -2,22523
 Estadístico de contraste: $\tau_{a,nc}(1) = -4,46639$
 valor p asintótico $8,807e-006$

contraste con constante
 modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$
 Coef. de autocorrelación de primer orden de e : -0,006
 valor estimado de $(a - 1)$: -2,22377
 Estadístico de contraste: $\tau_{a,c}(1) = -4,44332$
 valor p asintótico 0,0001

En lo que se refiere al contraste de raíces unitarias sobre la serie $(1-L)(1-L^{12}) Ln(Pasajeros)$, los resultados del cuadro 7.2 muestran que se:

$$t_c = -4,44 < DF_{0,05} = -2,89$$

Por lo que se rechaza la hipótesis nula de existencia de raíz unitaria.

Podemos concluir que la serie $(1-L)(1-L^{12}) Ln(Pasajeros)$ es estacionaria, por lo que la serie $Ln(Pasajeros)$ es integrada de orden 1 en la parte regular, $d = 1$ e integrada de orden 1 en la parte estacional, $D = 1$.

B. Selección de los órdenes (p, q) y (P, Q)

La elección del / los modelos apropiados para la serie estacionaria se realiza estudiando su función de autocorrelación simple y parcial (gráfico 7.9). El gráfico de la función de autocorrelación simple muestra que el primer coeficiente regular significativamente distinto de cero y que en la parte estacional también solo el primer coeficiente, ρ_{12} , es significativamente distinto de cero. En la función de autocorrelación parcial se observa que el primer coeficiente es significativamente distinto de cero, aunque se aprecia una estructura general exponencial decreciente, y que en la parte estacional el coeficiente p_{12} es significativamente distinto de cero y decrecen.

En principio, como tanto en la parte regular como en la estacional observamos que el primer coeficiente de la FAC es distinto de cero y el resto no son significativos proponemos una estructura medias móviles de orden 1 para ambas, osea, $ARMA(0, 1)(0, 1)_{12}$. Obsérvese como a ambos lados de los coeficientes estacionales de la FAC se reproduce la estructura de la parte regular.

Por último, para acabar de identificar el/los modelos que pueden haber generado la serie *Pasajeros de líneas aéreas*, es preciso decidir si es necesario incluir o no un término independiente. El gráfico de la serie (figura derecha del gráfico 7.9) parece sugerir que la media de esta transformación estacionaria es cero y que, por lo tanto, no hay que añadir ninguna constante. Para confirmar esta conclusión, contrastamos la hipótesis nula de que la media de la serie $Z_t = \Delta\Delta_{12} Ln(Pasajeros)_t$ es cero, con el estadístico:

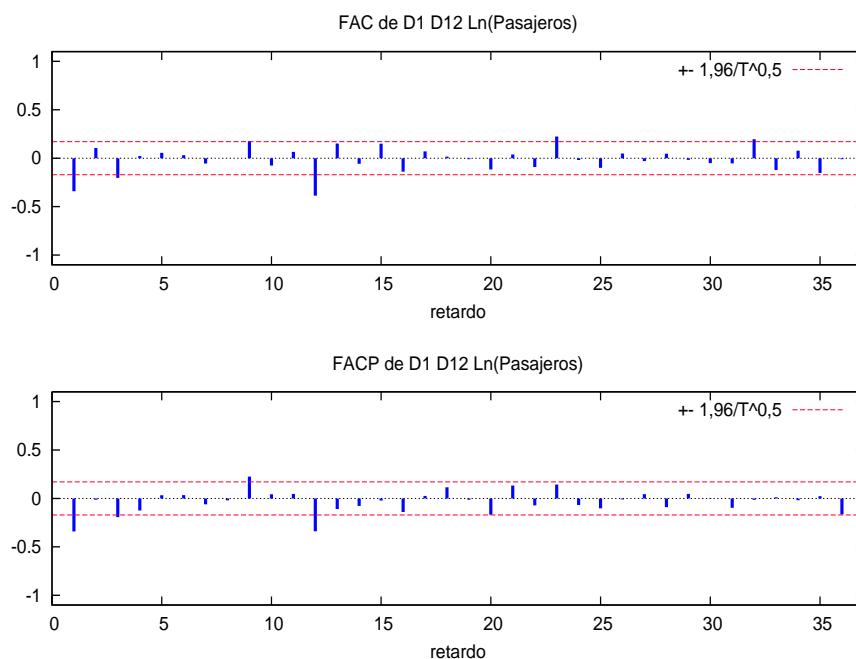
$$t = \frac{\bar{Z}}{\sigma_{\bar{Z}}}$$

Para calcular este estadístico se utilizan los siguientes datos sobre la serie estacionaria $(1-L)(1-L^{12})Ln(Pasajeros)$

Estadísticos principales, usando las observaciones 1949:10 - 1963:09 para la variable $(1-L)(1-L^{12})Ln(Pasajeros)$ (131 observaciones válidas)

Media	Mediana	Mínimo	Máximo
0,000290880	0,000000	-0,141343	0,140720
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,0458483	157,619	0,0381906	1,14778

Gráfico 7.9: FAC y FACP estimadas de la serie estacionaria.



Así, se obtiene:

$$\sigma_{\bar{Z}}^2 = \hat{V}(\bar{Z}) (1 + 2\hat{\rho}_1 + 2\hat{\rho}_{12}) = \frac{0,0458^2}{(131 - 1)} (1 + 2 \cdot 0,34 + 2 \cdot 0,387) = 0,0000395$$

$$t = \frac{0,000291}{0,006287} = 0,04628 < 2 \approx t_{0,025}(130)$$

Por lo tanto, no se rechaza la hipótesis nula y no se incluye un término independiente en el modelo.

En la etapa de Identificación se ha propuesto el siguiente modelo candidato a haber generado la serie *Pasajeros de líneas aéreas*:

$$ARIMA(0, 1, 1)(0, 1, 1)_{12} : \quad \Delta\Delta_{12} \text{Ln}(\text{Pasajeros})_t = (1 - \theta L)(1 - \Theta L^{12}) a_t$$

ESTIMACIÓN.

El modelo estimado según el output obtenido con Gretl que se presenta a continuación es:

$$\Delta\Delta_{12} \text{Ln}(\text{Pasajeros})_t = (1 - 0,4018 L)(1 - 0,5569 L^{12}) a_t$$

Modelo 1: estimaciones ARIMA utilizando las 131 observaciones 1950:11–1961:09

Variable dependiente: $(1 - L)(1 - L^{12})Ln(Pasajeros)$

Desviaciones típicas basadas en el Hessiano

	Coefficiente	Desv. típica	estadístico t	valor p
θ_1	-0,401823	0,0896453	-4,4824	0,0000
Θ_1	-0,556937	0,0731050	-7,6183	0,0000

Media de la var. dependiente	0,000290880
D.T. de la variable dependiente	0,0458483
media de las innovaciones	0,00100838
Varianza de las innovaciones	0,00134810
Log-verosimilitud	244,696

		Real	Imaginaria	Módulo	Frecuencia	
MA	Raíz	1	2,4887	0,0000	2,4887	0,0000
MA (estacional)	Raíz	1	1,7955	0,0000	1,7955	0,0000

VALIDACIÓN.

En esta etapa se trata de comprobar de que el modelo propuesto se ajusta bien a los datos y reproduce la estructura de comportamiento de la serie.

Comenzando por el análisis de los coeficientes, comprobaremos, en primer lugar, si son o no estadísticamente significativos.

Contrastes de significatividad individual:

$$H_0: \theta = 0 \quad H_0: \Theta = 0$$

$$H_a: \theta \neq 0 \quad , \quad H_a: \Theta \neq 0$$

$$\theta: \left| \frac{\hat{\theta}}{\hat{\sigma}_{\hat{\theta}}} \right| = \left| \frac{-0,265}{0,0699} \right| = |-3,79| > N_{0,025} \approx 2$$

Se rechaza $H_0: \theta = 0$ para un nivel de significación del 5%.

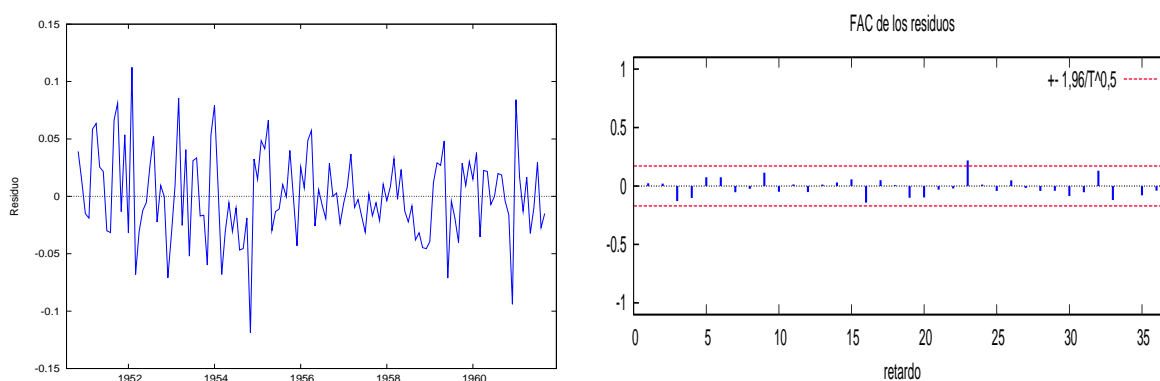
$$\Theta: \left| \frac{\hat{\Theta}}{\hat{\sigma}_{\hat{\Theta}}} \right| = \left| \frac{-0,527}{0,072} \right| = |-7,346| > N_{0,025} \approx 2$$

Se rechaza $H_0: \Theta = 0$ para un nivel de significación del 5%.

El modelo propuesto es estacionario porque es un medias móviles finito y es invertible porque el módulo de las trece raíces medias móviles está fuera del círculo unidad.

Comenzamos el análisis de los residuos con el análisis de la figura izquierda del gráfico 7.10 en la que se observa que los residuos oscilan en torno a cero con una variabilidad bastante homogénea. Por otro lado, el correlograma de los residuos muestra que ningún coeficiente de autocorrelación está fuera de la región crítica, por lo que ninguno es significativamente distinto de cero.

Gráfico 7.10: Serie: $\Delta\Delta_{12} Ln(Pasajeros)$. Estimación y diagnósticos.



En cuanto al contraste de significación conjunto de la siguiente hipótesis nula:

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_{36} = 0$$

El estadístico de Box-Ljung, $Q = 43,75 < \chi_{0,05}^2$, por lo que no se rechaza la hipótesis nula de que los 36 primeros coeficientes de autocorrelación son conjuntamente cero.

Estadísticos principales, usando las observaciones 1949:10 - 1963:09 para la variable uhat1 (131 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
0,00100838	-0,00280376	-0,118741	0,112436
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,0376848	37,3715	0,0770597	0,547118

En lo que se refiere al contraste de Jarque-Bera de Normalidad, se tiene que:

$$JB = 32,46 > \chi_{0,05}^2(2) = 5,99$$

Por lo tanto, podemos concluir que los residuos se comportan como un ruido blanco si bien no parecen seguir una distribución normal.

Ejercicio 7.1. Índice de Producción Industrial austriaco.

Dados los siguientes gráficos y tablas, construye del modelo $ARIMA(p, d, q)(P, D, Q)_s$ apropiado para la serie trimestral del Índice de producción industrial austriaco (IPI) para la que se cuenta con datos desde el primer trimestre de 1975 hasta el cuarto de 1988 (los datos originales se encuentran en el apéndice C).

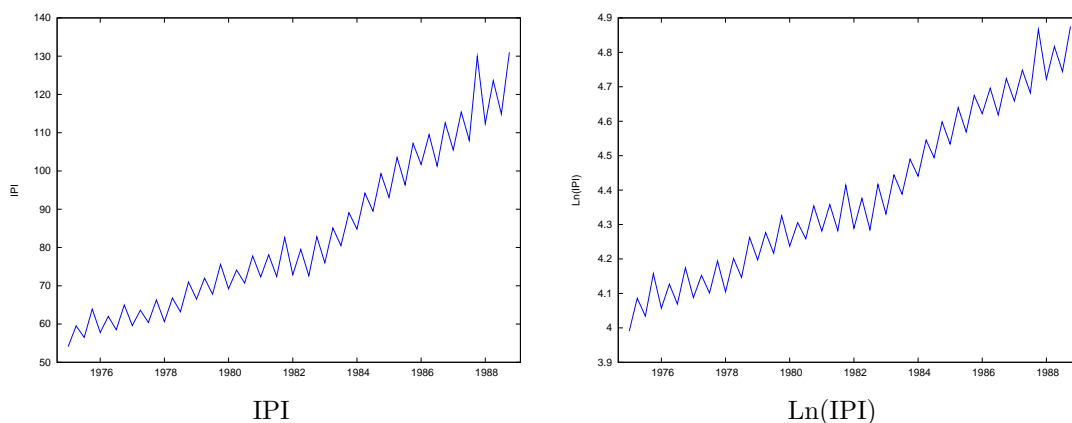
IDENTIFICACIÓN.

A. Análisis de estacionariedad.

Pregunta:

¿Es la serie estacionaria en varianza? Busca la transformación estacionaria.

Gráfico 7.11: Serie: Índice de Producción Industrial (1975:I - 1988:IV)

**Pregunta:**

¿Es la serie estacionaria en media? Busca la transformación estacionaria.

Gráfico 7.12: Gráficos de diferencias de la serie Ln(IPI).

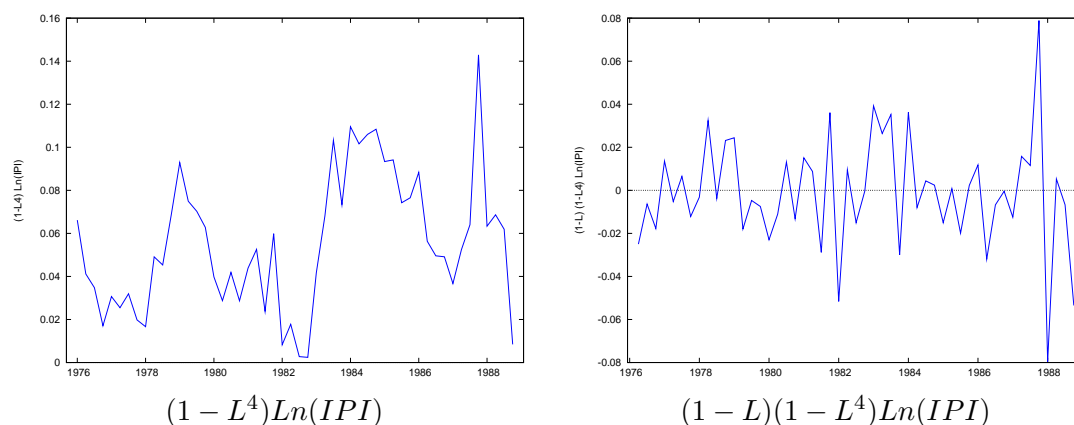
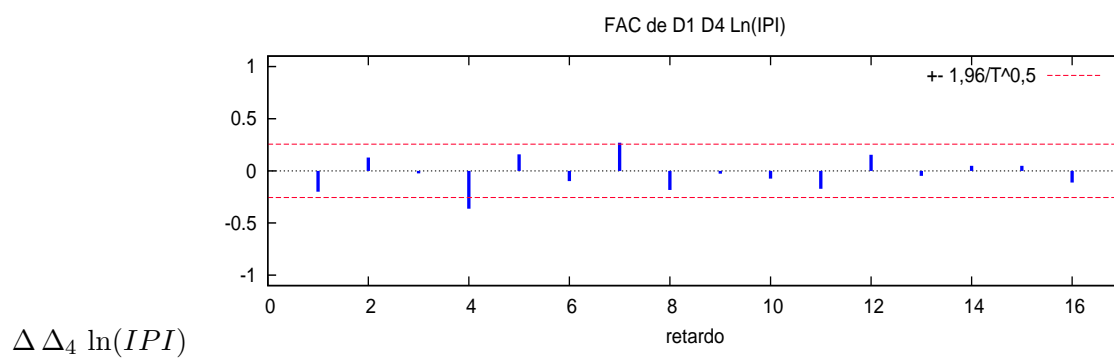
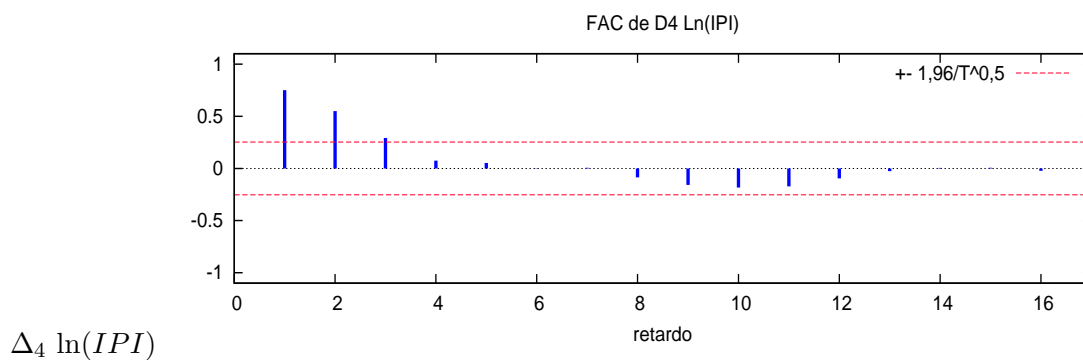
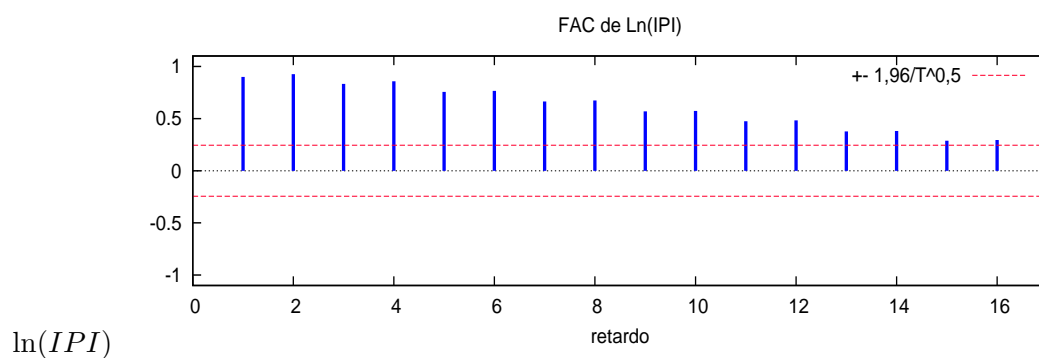


Gráfico 7.13: Correlogramas de diferencias de la serie Ln(IPI).



Cuadro 7.3: Contrastes de raíces unitarias.

Contrastes aumentados de Dickey-Fuller, orden 4, para $(1 - L^4) \ln(\text{IPI})$:
 tamaño muestral 47 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante
 modelo: $(1 - L)y = (a-1)y(-1) + \dots + e$
 Coef. de autocorrelación de primer orden de e : -0,022
 valor estimado de $(a - 1)$: -0,0473233
 Estadístico de contraste: $\tau_{nc}(1) = -0,799128$
 valor p asintótico 0,3701

contraste con constante
 modelo: $(1 - L)y = b_0 + (a-1)y(-1) + \dots + e$
 Coef. de autocorrelación de primer orden de e : 0,005
 valor estimado de $(a - 1)$: -0,353237
 Estadístico de contraste: $\tau_c(1) = -2,39611$
 valor p asintótico 0,1428

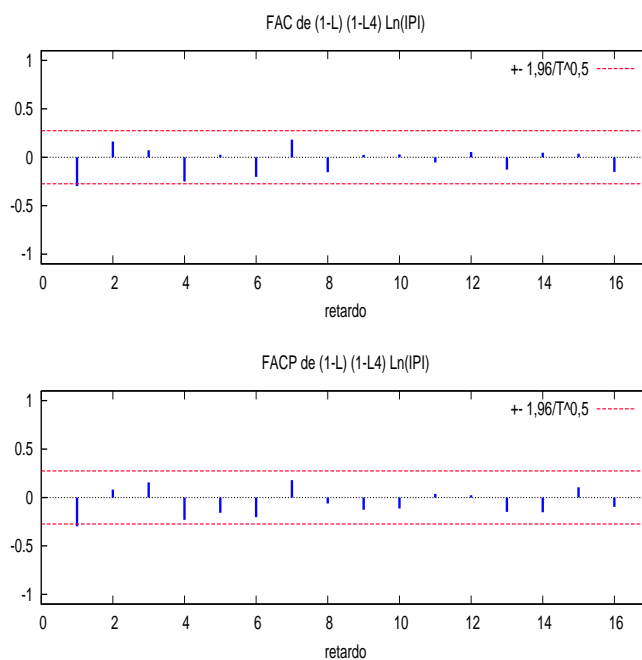
Contrastes aumentados de Dickey-Fuller, orden 4, para $(1 - L)(1 - L^4) \ln(\text{IPI})$:
 tamaño muestral 46 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante
 modelo: $(1 - L)y = (a-1)y(-1) + \dots + e$
 Coef. de autocorrelación de primer orden de e : -0,022
 valor estimado de $(a - 1)$: -1,64814
 Estadístico de contraste: $\tau_{nc}(1) = -3,70536$
 valor p asintótico 0,0001

contraste con constante
 modelo: $(1 - L)y = b_0 + (a-1)y(-1) + \dots + e$
 Coef. de autocorrelación de primer orden de e : -0,022
 valor estimado de $(a - 1)$: -1,65542
 Estadístico de contraste: $\tau_c(1) = -3,65382$
 valor p asintótico 0,004838

B. Selección de los ordenes (p, q) y (P, Q) , y del término independiente.

Gráfico 7.14: FAC y FACP estimadas de la serie $\Delta\Delta_4 \ln IPI_t$.



Estadísticos principales, usando las observaciones 1975:1 - 1988:4 para la variable D1D4.LnIPI (51 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
-0,00113181	-0,00316331	-0,0795702	0,0788784
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,0260694	23,0334	-0,0446416	1,73581

Pregunta:

¿Qué modelo/modelos ARIMA estacionales propones para la serie IPI?

ESTIMACIÓN.

Un investigador ha propuesto los siguientes modelos para la serie IPI:

$$\text{Modelo A : } (1 - \phi L) \Delta \Delta_4 \ln IPI_t = (1 - \Theta L^4) a_t$$

$$\text{Modelo B : } \Delta \Delta_4 \ln IPI_t = (1 - \theta L)(1 - \Theta L^4) a_t$$

Pregunta: *¿Estás de acuerdo? ¿Por qué?*Con los resultados que se presentan a continuación, *escribe los modelos estimados*.**Modelo A:** estimaciones ARIMA utilizando las 51 observaciones 1976:2–1988:4Variable dependiente: $(1 - L)(1 - L^4) \ln(\text{IPI})$

Variable	Coefficiente	Desv. típica	Estadístico t	valor p
ϕ_1	-0,353517	0,131443	-2,6895	0,0072
Θ_1	-0,868740	0,177300	-4,8998	0,0000
	media de las innovaciones		0,00171446	
	Varianza de las innovaciones		0,000401395	
	Log-verosimilitud		124,237	

			Real	Imaginaria	Módulo	Frecuencia
AR	Raíz	1	-2,8287	0,0000	2,8287	0,5000
MA (estacional)	Raíz	1	1,1511	0,0000	1,1511	0,0000

Modelo B: estimaciones ARIMA utilizando las 51 observaciones 1976:2–1988:4Variable dependiente: $(1 - L)(1 - L^4) \ln(\text{IPI})$

Variable	Coefficiente	Desv. típica	Estadístico t	valor p
θ_1	-0,296070	0,115774	-2,5573	0,0105
Θ_1	-0,855573	0,161817	-5,2873	0,0000
	media de las innovaciones		0,00184053	
	Varianza de las innovaciones		0,000413225	
	Log-verosimilitud		123,660	

			Real	Imaginaria	Módulo	Frecuencia
MA	Raíz	1	3,3776	0,0000	3,3776	0,0000
MA (estacional)	Raíz	1	1,1688	0,0000	1,1688	0,0000

VALIDACIÓN

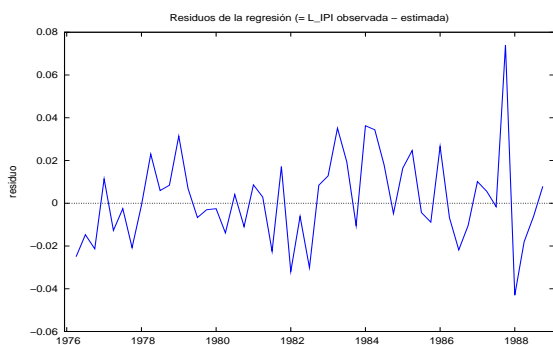
Con los resultados mostrados en los gráficos y tablas del cuadro 7.4 contesta a las siguientes pregunta:

Pregunta:

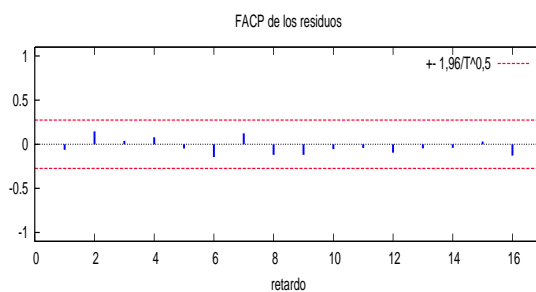
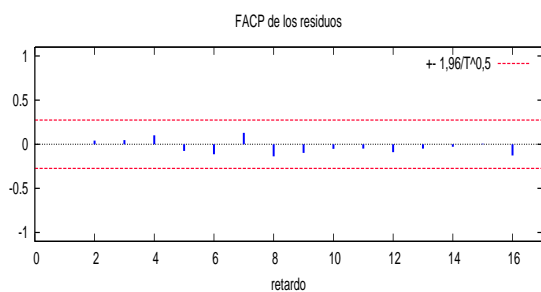
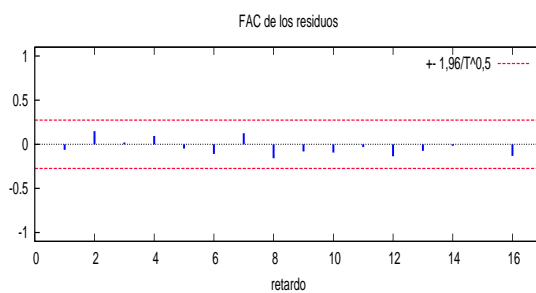
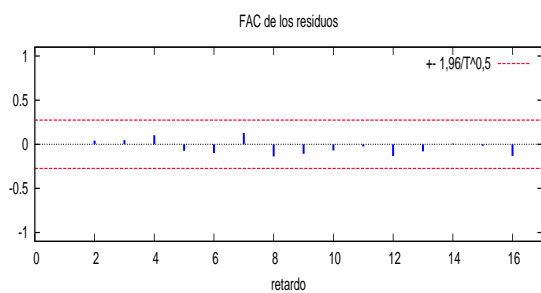
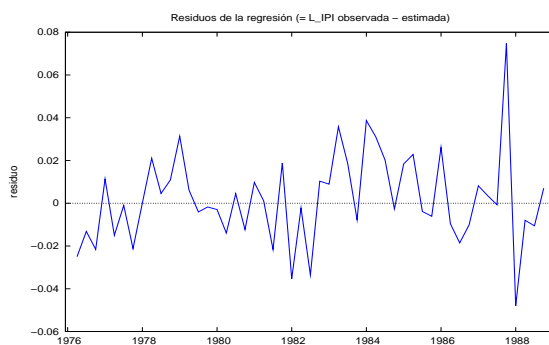
¿Se ajustan los modelos a los datos? ¿Qué modelo seleccionarías para predecir?

Cuadro 7.4: Serie: $\Delta\Delta_4 \ln IPI_t$. Diagnóstico.

Modelo $ARIMA(1, 1, 0)(0, 1, 1)_4$



Modelo $ARIMA(0, 1, 1)(0, 1, 1)_4$



$Q(24) = 16,465$

$Q(24) = 18,00$

Modelo $ARIMA(1, 1, 0)(0, 1, 1)_4$			
Estadísticos principales, usando las observaciones 1975:1 - 1988:4 para la variable uhatA (51 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
0,00171446	-0,00158269	-0,0430086	0,0740437
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,0208678	12,1716	0,706177	1,47041
Modelo $ARIMA(0, 1, 1)(0, 1, 1)_4$			
Estadísticos principales, usando las observaciones 1975:1 - 1988:4 para la variable uhatB (51 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
0,00184053	-0,000717917	-0,0479773	0,0747304
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,0211028	11,4656	0,588185	1,67900

Apéndice C. Datos.

Cuadro 7.5: Pasajeros de líneas aéreas (Octubre 1949 - Septiembre 1961)

Año	Ene.	Febr.	Marzo	Abril	Mayo	Jun.	Jul.	Agos.	Sept.	Oct.	Nov.	Dic.
1949										112	118	132
1950	129	121	135	148	148	136	119	104	118	115	126	141
1951	135	125	149	170	170	158	133	114	140	145	150	178
1952	163	172	178	199	199	184	162	146	166	171	180	193
1953	181	183	218	230	242	209	191	172	194	196	196	236
1954	235	229	243	264	272	237	211	180	201	204	188	235
1955	227	234	264	302	293	259	229	203	229	242	233	267
1956	269	270	315	364	347	312	274	237	278	284	277	317
1957	313	318	374	413	405	355	306	271	306	315	301	356
1958	348	355	422	465	467	404	347	305	336	340	318	362
1959	348	363	435	491	505	404	359	310	337	360	342	406
1960	396	420	472	548	559	463	407	362	405	417	391	419
1961	461	472	535	622	606	508	461	390	432			

Cuadro 7.6: Índice de producción industrial (1º trimestre 1975 - 4º trimestre 1988)

Fecha	IPI	Fecha	IPI	Fecha	IPI	Fecha	IPI	Fecha	IPI
1975:1	54,1	1978:1	60,6	1981:1	72,3	1984:1	84,8	1987:1	105,5
1975:2	59,5	1978:2	66,8	1981:2	78,1	1984:2	94,2	1987:2	115,4
1975:3	56,5	1978:3	63,2	1981:3	72,4	1984:3	89,5	1987:3	108,0
1975:4	63,9	1978:4	71,0	1981:4	82,6	1984:4	99,3	1987:4	129,9
1976:1	57,8	1979:1	66,5	1982:1	72,9	1985:1	93,1	1988:1	112,4
1976:2	62,0	1979:2	72,0	1982:2	79,5	1985:2	103,5	1988:2	123,6
1976:3	58,5	1979:3	67,8	1982:3	72,6	1985:3	96,4	1988:3	114,9
1976:4	65,0	1979:4	75,6	1982:4	82,8	1985:4	107,2	1988:4	131,0
1977:1	59,6	1980:1	69,2	1983:1	76,0	1986:1	101,7		
1977:2	63,6	1980:2	74,1	1983:2	85,1	1986:2	109,5		
1977:3	60,4	1980:3	70,7	1983:3	80,5	1986:3	101,3		
1977:4	66,3	1980:4	77,8	1983:4	89,1	1986:4	112,6		

Capítulo 8

Ejercicios

8.1. Cuestiones

C1. Sea Y_t un proceso estocástico tal que $Y_t = a + bX$, para todo t , donde a, b son constantes y X una variable aleatoria con media μ y varianza σ^2 . Demuestra que el proceso estocástico Y_t es estacionario en covarianza. ¿Cuál es la función de autocovarianzas del proceso Y_t ?

C2. Supongamos que la función de autocovarianzas del proceso Y_t no depende de t , pero sin embargo, la media del proceso viene dada por $E(Y_t) = 4t$. ¿Es el proceso Y_t estacionario?. Considera ahora el proceso $Z_t = 10 - 4t + Y_t$, ¿es estacionario?

C3. Considera el proceso estocástico $Z_t = X + Y_t$, donde Y_t es un proceso estocástico estacionario en covarianza y X es una variable aleatoria independiente de Y_t . Suponiendo conocidas la media y varianza de X y la media y la función de autocorrelación simple de Y_t , obtén la media y la función de autocorrelación simple del proceso Z_t .

C4. Sea Y_t un proceso estocástico estacionario con función de autocovarianzas γ_k . Demuestra que:

- $Z_t = \Delta Y_t = Y_t - Y_{t-1}$ es estacionario.
- $Z_t = \Delta^2 Y_t$ es estacionario.
- $Z_t = \Delta_4 Y_t$ es estacionario.
- en general, $Z_t = \sum_{j=0}^N a_j Y_{t-j}$ es estacionario.

¿Qué conclusión general puedes obtener de los resultados obtenidos en los cuatro apartados anteriores?

C5. Demuestra que el proceso $MA(\infty)$ siguiente:

$$Y_t = a_t + c(a_{t-1} + a_{t-2} + \dots)$$

donde c es una constante, no es estacionario. Demuestra también que las primeras diferencias del proceso:

$$W_t = Y_t - Y_{t-1}$$

son un proceso MA(1) estacionario. Busca la función de autocorrelación simple de W_t .

C6. Considera el proceso MA estacionario $Y_t = (1 - \theta L)a_t$ y sea $Z_t = Y_t - Y_{t-1}$. ¿Es el proceso Z_t estacionario? ¿Es la varianza del proceso Z_t mayor o menor que la del proceso Y_t ? ¿Cuál es la función de autocovarianzas del proceso diferenciado?

C7. Considera el proceso AR estacionario $(1 - \phi L)Y_t = a_t$ y sea $Z_t = Y_t - Y_{t-1}$. ¿Es el proceso Z_t estacionario? ¿Es la varianza del proceso Z_t mayor o menor que la del proceso Y_t ? ¿Cuál es la función de autocovarianzas del proceso diferenciado?

C8. Si un analista presenta el siguiente modelo AR(4) para una muestra dada:

$$(1 + 0,48L + 0,22L^2 + 0,14L^3 + 0,05L^4)Y_t = a_t$$

¿Podrías sugerirle un modelo más parsimonioso? ¿Por qué?

C9. Contesta a las siguientes preguntas:

- ¿Cómo se obtienen las funciones de autocorrelación simple y parcial teóricas?
- ¿Cómo se obtienen las funciones de autocorrelación simple y parcial estimadas o muestrales?
- ¿Las facs y facp estimadas se corresponden exactamente con las teóricas? ¿Por qué?

C10. Sea el proceso $Y_t = \mu + a_t + \theta a_{t-1}$, donde μ es una constante, demuestra que la f.a.c.s. de Y_t no depende de μ .

C11. Si I_T es un conjunto de información que incluye los valores pasados y presente de la serie que queremos predecir, Y_t , y $\hat{Y}_T(1)$ es la predicción óptima basada en el citado conjunto de información I_T , de forma que el error de predicción es $e_T(1)$. ¿Por qué esperas que $e_T(1)$ sea ruido blanco?

C12. Comenta las siguientes frases:

- Un modelo $ARIMA(p, d, q)$ bien construido tiene autocorrelaciones residuales que son todas cero.
- Un modelo ARIMA estimado con residuos que están significativamente autocorrelacionados no es apropiado.

8.2. Problemas

Problema 1. Contamos con la siguiente serie temporal estacionaria:

1,5 0,8 0,9 1,2 0,5 1,3 0,8 1,6 0,8 1,2 0,5 0,9 1,1 1,1 0,6 1,2

- En base al gráfico de las observaciones, sugiere un valor aproximado para el coeficiente de autocorrelación muestral de orden 1, $\hat{\rho}_1$.
- Dibuja Y_t contra Y_{t+1} e Y_t contra Y_{t+2} y, en base a estos gráficos, sugiere valores aproximados para $\hat{\rho}_1$ y $\hat{\rho}_2$.
- Comprueba si tus sugerencias eran apropiadas calculando $\hat{\rho}_k$, $k = 0, 1, 2, 3, 4, 5$

Problema 2. Considera los siguientes procesos estocásticos:

$$1. \quad Y_t = 1,05 Y_{t-1} - 0,5 Y_{t-2} + a_t$$

$$2. \quad Y_t = -1,03 Y_{t-1} + a_t$$

$$3. \quad Y_t = 0,7 Y_{t-1} + a_t - 0,45 a_{t-1}$$

$$4. \quad Y_t = a_t + 0,7 a_{t-1} - 0,5 a_{t-2}$$

$$5. \quad Y_t = a_t - 1,05 a_{t-2}$$

$$6. \quad Y_t = 0,09 Y_{t-2} + a_t - 0,3 a_{t-1}$$

$$7. \quad Y_t = -0,6 Y_{t-1} + a_t + 0,4 a_{t-2}$$

¿Qué modelo $ARMA(p, q)$ son? ¿Son estacionarios e invertibles? ¿Por qué?

Problema 3. Sean los modelos $AR(1)$ siguientes:

$$(a) \quad Y_t = 0,8 Y_{t-1} + a_t \quad \sigma_a^2 = 2$$

$$(b) \quad Y_t = -0,5 Y_{t-1} + a_t$$

- ¿Son estacionarios? ¿Son invertibles?
- Calcula las autocovarianzas, $\gamma_0, \gamma_1, \dots, \gamma_5$.
- Calcula los coeficientes de autocorrelación simple, $\rho_1, \rho_2, \dots, \rho_5$. Representalos gráficamente.
- Calcula los coeficientes de autocorrelación parcial, p_1, p_2, p_3, p_4, p_5 . Representalos gráficamente.
- Calcula los pesos $\psi_1, \psi_2, \dots, \psi_5$ del modelo en forma $MA(\infty)$ en que, si es posible, puede transformarse los dos modelos $AR(1)$ dados.

Problema 4. Considera el proceso $AR(1)$ siguiente:

$$Y_t = 20 + 0,6 Y_{t-1} + a_t \quad \sigma_a^2 = 4$$

- ¿Es estacionario? ¿Es invertible?
- Calcula la media y la función de autocovarianzas del proceso. Dibuja la función de autocorrelación teórica.
- Si $Y_{50} = 21$, ¿es de esperar que Y_{51} sea mayor o menor que la media?

Problema 5. Sea a_t un proceso ruido blanco. Encuentra la función de autocorrelación simple de los siguientes procesos lineales:

$$(1) \quad Y_t = a_t + 0,4 a_{t-1} \qquad (2) \quad Y_t = a_t + 2,5 a_{t-1}$$

¿Crees que estos dos modelos lineales se pueden distinguir en base a observaciones de $\{Y_t\}$?

Problema 6. Sean los modelos $MA(1)$ siguientes:

$$\begin{aligned} 1. \quad Y_t &= a_t - 0,9 a_{t-1} & \sigma_a^2 &= 4 \\ 2. \quad Y_t &= a_t + 0,5 a_{t-1} \end{aligned}$$

- ¿Son estacionarios? ¿Son invertibles?
- Calcula las autocovarianzas, $\gamma_0, \gamma_1, \dots, \gamma_5$.
- Calcula los coeficientes de autocorrelación simple, $\rho_1, \rho_2, \dots, \rho_5$. Representalos gráficamente.
- Calcula los pesos $\pi_1, \pi_2, \dots, \pi_5$ del modelo en forma $AR(\infty)$ en que, si es posible, puede transformarse los dos modelos $MA(1)$ dados.

Problema 7. Considera el proceso $MA(1)$ siguiente:

$$Y_t = 5 + a_t + 0,2 a_{t-1} \quad \sigma_e^2 = 3$$

- ¿Es estacionario? ¿Es invertible?
- Calcula la media y la función de autocovarianzas del proceso. Dibuja la función de autocorrelación teórica.
- Si Y_{100} es mayor que la media del proceso, ¿es de esperar que Y_{101} sea también mayor?

Problema 8. Sean los modelos $AR(2)$ siguientes:

$$\begin{aligned} 1. \quad Y_t &= 0,6 Y_{t-1} + 0,3 Y_{t-2} + a_t & \sigma^2 &= 3 \\ 2. \quad Y_t &= 1,3 Y_{t-1} - 0,4 Y_{t-2} + a_t \\ 3. \quad Y_t &= 1,2 Y_{t-1} - 0,8 Y_{t-2} + a_t \\ 4. \quad Y_t &= 0,8 Y_{t-1} - 0,5 Y_{t-2} + a_t \end{aligned}$$

- ¿Son estacionarios? ¿Son invertibles?
- ¿Cual es la estructura de la función de autocorrelación simple y de la función de autocorrelación parcial?
- Calcula las autocovarianzas, $\gamma_0, \gamma_1, \dots, \gamma_5$. Calcula los coeficientes de autocorrelación simple, $\rho_1, \rho_2, \dots, \rho_5$. Representalos gráficamente.
- Calcula los pesos $\psi_1, \psi_2, \dots, \psi_5$ del modelo en forma $MA(\infty)$ en que, si es posible, puede transformarse los cuatro modelos $AR(2)$ dados.

Problema 9. Considera el proceso $AR(2)$ siguiente:

$$Y_t = 10 + 0,8 Y_{t-1} - 0,2 Y_{t-2} + a_t \quad \sigma_a^2 = 1$$

- ¿Es estacionario? ¿Es invertible?
- Calcula su media y función de autocovarianzas. Dibuja la función de autocorrelación teórica.
- Si $Y_{45} = 12$, ¿es de esperar que Y_{46} sea mayor o menor que la media?

Problema 10. Sean los modelos $MA(2)$ siguientes:

$$\begin{aligned} 1. \quad Y_t &= a_t - 0,4 a_{t-1} + 1,2 a_{t-2} & \sigma_a^2 &= 2 \\ 2. \quad Y_t &= a_t + 0,7 a_{t-1} - 0,2 a_{t-2} \\ 3. \quad Y_t &= a_t - 1,3 a_{t-1} + 0,4 a_{t-2} \\ 4. \quad Y_t &= a_t - 0,1 a_{t-1} + 0,21 a_{t-2} \end{aligned}$$

- ¿Son estacionarios? ¿Son invertibles?
- ¿Cual es la estructura de la función de autocorrelación simple y de la función de autocorrelación parcial?
- Calcula las autocovarianzas, $\gamma_0, \gamma_1, \dots, \gamma_5$. Calcula los coeficientes de autocorrelación simple, $\rho_1, \rho_2, \dots, \rho_5$. Representalos gráficamente.
- Calcula los pesos $\pi_1, \pi_2, \dots, \pi_5$ del modelo en forma $AR(\infty)$ en que, si es posible, puede transformarse los cuatro modelos $MA(2)$ dados.

Problema 11. Sean los modelos $ARMA(1, 1)$ siguientes:

$$(a) \quad Y_t = 0,9Y_{t-1} + a_t + 0,8a_{t-1} \quad \sigma^2 = 5$$

$$(b) \quad Y_t = 0,6Y_{t-1} + a_t - 0,6a_{t-1}$$

- ¿Son estacionarios? ¿Son invertibles?
- Calcula las autocovarianzas, $\gamma_0, \gamma_1, \dots, \gamma_5$.
- Calcula los coeficientes de autocorrelación simple, $\rho_1, \rho_2, \dots, \rho_5$. Representalos gráficamente.
- ¿Pueden simplificarse estos modelos?
- Calcula los pesos $\pi_1, \pi_2, \dots, \pi_5$ del modelo en forma $AR(\infty)$ en que, si es posible, puede transformarse los dos modelos $ARMA(1, 1)$ dados.
- Calcula los pesos $\psi_1, \psi_2, \dots, \psi_5$ del modelo en forma $MA(\infty)$ en que, si es posible, puede transformarse los dos modelos $ARMA(1, 1)$ dados.

Problema 12. Supongamos que la serie Y_t ha sido generada como sigue:

$$Y_t = -0,2Y_{t-1} + 0,48Y_{t-2} + a_t + 0,6a_{t-1} - 0,16a_{t-2}$$

- ¿Qué proceso lineal sigue la serie Y_t ? ¿Es estacionario? ¿Es invertible?
- ¿Está el modelo sobreparametrizado? En caso afirmativo, escribe un modelo equivalente con menos parámetros.

Problema 13. Sea el modelo $ARMA(2,2)$ siguiente:

$$Y_t = 0,9Y_{t-1} - 0,2Y_{t-2} + a_t - 1,3a_{t-1} + 0,4a_{t-2}$$

¿Puede simplificarse este modelo?

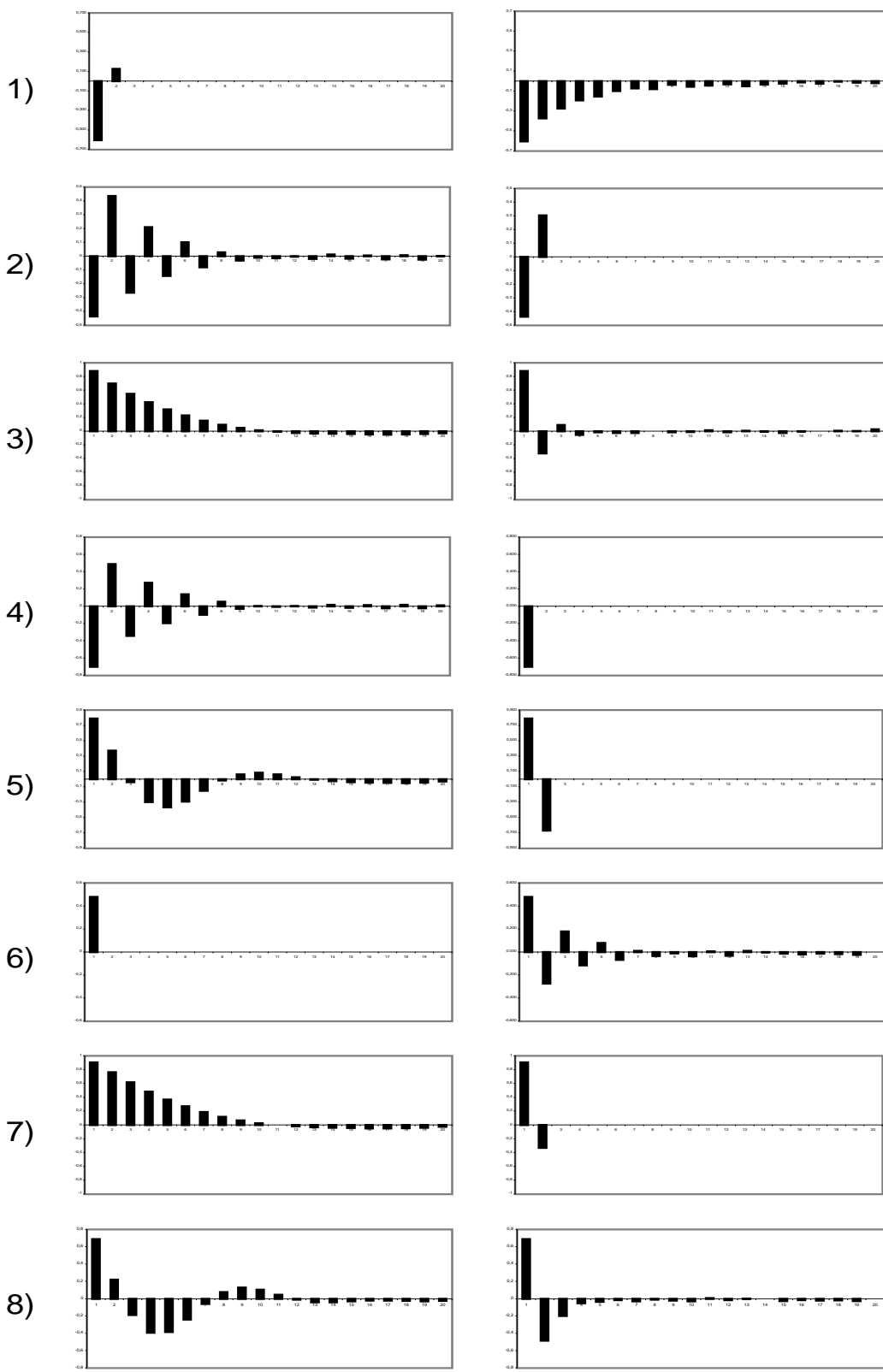
Problema 14. Encuentra un proceso invertible tal que:

$$\rho_0 = 1 \quad \rho_1 = 0,49 \quad \rho_k = 0 \quad k > 1$$

Problema 15. Indica a qué tipo de modelo $ARMA(p, q)$ se puede asociar cada uno de los siguientes pares de funciones de autocorrelación simples y parciales teóricas, explicando por qué.

F.AC.S.

F.AC.P.



Problema 16. Considera la siguiente secuencia de coeficientes de autocorrelación:

$$\rho_1 = 0,75 \quad \rho_2 = 0,56 \quad \rho_3 = 0,42 \quad \rho_4 = 0,31 \quad \rho_5 = 0,23$$

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

Problema 17. Las funciones de autocorrelación simple y parcial estimadas para una serie temporal de 110 observaciones son las siguientes:

k	$\hat{\rho}_k$	\hat{p}_k
1	-0,38	-0.38
2	-0.16	-0.37
3	0.12	-0.20
4	-0.08	-0.18
5	-0.06	-0.03

- Calcula la desviación típica asintótica para cada coeficiente estimado, $\hat{\rho}_k$ y \hat{p}_k bajo el supuesto de que se trata de una realización que procede de un ruido blanco.
- ¿Qué modelo $ARMA(p, q)$ ha podido generar la serie de datos?

Problema 18. La función de autocorrelación estimada para una serie Y_t de 144 observaciones es:

k	0	1	2	3	4	5	6
r_k	1	0.84	0.57	0.43	0.34	0.25	0.18

- Sugiere un modelo $ARMA(p, q)$ posible para Y_t .
- Estima él/los parámetros del modelo propuesto (*Nota: Recuerda las expresiones para la función de autocorrelación simple en función de los parámetros del modelo tanto para los modelos AR como MA.*)
- Suponiendo que el error de predicción un periodo hacia adelante cometido al predecir Y_6 con información hasta el momento $t = 5$ es de $-0,36$ y si $Y_7 = 1$, obtén la predicción para Y_8 .

Problema 19. A partir de una muestra de 64 observaciones se ha ajustado el siguiente modelo:

$$(1 - 0,8L)Y_t = (1 + 0,6L)\hat{a}_t$$

Las funciones de autocorrelación simple y parcial de los residuos, \hat{a}_t , vienen dadas por:

k	1	2	3	4	5
$\hat{\rho}_k$	0.390	0.160	0.054	0.021	-0.004
\hat{p}_k	0.390	0.030	-0.033	0.025	0.026

A la vista de la información anterior, ¿consideras aceptable el modelo ajustado? En caso contrario, ¿cual podría ser la reformulación del modelo?

Problema 20. Se ha estimado con una muestra de 110 observaciones de la variable Y_t el siguiente modelo:

$$Y_t = a_t + 0,5 a_{t-1} + 0,4 a_{t-2} + 18 \quad \hat{\sigma}^2 = 4$$

Se cuenta con la siguiente información acerca de a_t e Y_t :

$$Y_{106} = 20 \quad Y_{107} = 21 \quad Y_{108} = 19 \quad Y_{109} = 19 \quad Y_{110} = 17$$

- Obtén la predicción de Y_t por punto para los periodos 111, 112 y 113.
- Calcula los errores cuadráticos medios de predicción.
- Obtén la predicción de Y_t por intervalo para los periodos 111, 112 y 113.
- Suponiendo que se dispone de la información adicional: $Y_{111} = 17$, actualiza la predicción de los periodos 112 y 113.

Problema 21. Consideremos el proceso AR(1) siguiente:

$$(1 - \phi L)(Y_t - \mu) = a_t$$

donde: $\phi = 0,6$ $\sigma_a^2 = 0,1$ $\mu = 9$.

Supongamos que se cuenta con las siguientes observaciones:

$$Y_{100} = 8,9 \quad Y_{99} = 9 \quad Y_{98} = 9 \quad Y_{97} = 9,6$$

Predice $Y_{101}, Y_{102}, Y_{103}, Y_{104}$ tanto por punto como por intervalo.

Si se dispone de la información adicional: $Y_{101} = 9,7$, actualiza la predicción de los periodos 102, 103 y 104.

Problema 22. Para predecir la serie Y_t se han utilizado dos modelos diferentes:

$$\begin{aligned} Y_t &= 20 + 0,5 Y_{t-1} + v_t \\ Y_t &= 50 + a_t + 0,6 a_{t-1} \end{aligned}$$

donde a_t y v_t son procesos ruido blanco con media cero. En el momento $t = 20$, las predicciones de ambos modelos para Y_{21} han sido de 110 y 90 respectivamente. ¿Qué predicciones para Y_{22} proporcionarían ambos modelos con información hasta $t = 21$, sabiendo que $Y_{21} = 100$?

Problema 23. El economista del servicio de planificación del Ayuntamiento quiere predecir los ingresos mensuales por impuestos en base a los datos recogidos los últimos 100 meses, Y_t . Analizando el correlograma estimado de la serie Y_t , propone los siguientes modelos:

$$Y_t = 90 + 0,7Y_{t-1} + a_t \quad (8.1)$$

$$Y_t = 300 + a_t + 0,62a_{t-1} + 0,5a_{t-2} \quad (8.2)$$

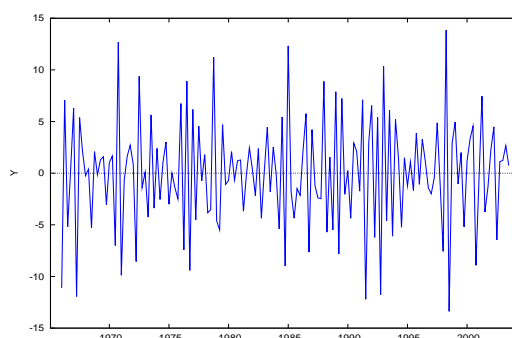
Sabiendo que $Y_{100} = 280$ y $a_{98} = -0,5, a_{99} = 2, a_{100} = -10$

- Obtén las predicciones por punto de Y_{101}, \dots, Y_{110} con ambos modelos.
- Representa ambas predicciones y compara sus características. ¿Cuál es la estructura de la función de predicción cuando $\ell \rightarrow \infty$ para ambos modelos?
- Sabiendo que los verdaderos valores para Y son:

t	101	102	103	104	105	106	107	108	109	110
y_t	295	310	302	275	320	310	290	283	315	303

Representa los errores de predicción cometidos con ambos métodos de predicción. ¿Qué modelo proporciona las “mejores” predicciones?. Utiliza para compararlos aquellos criterios de precisión de la predicción que te parezcan más convenientes?

Problema 24. La serie temporal representada en el gráfico corresponde a una variable observada trimestralmente desde el primer trimestre de 1966 al tercer trimestre de 2003. Se desea

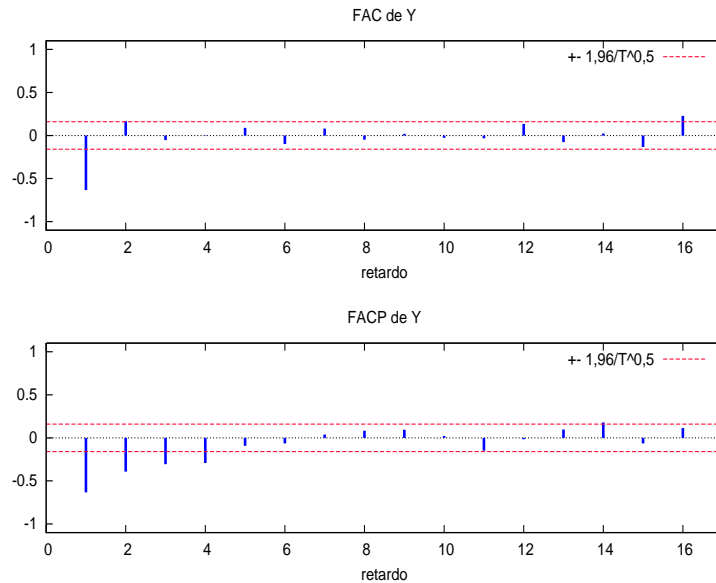


predecir la evolución de la variable a lo largo de 2004.

Basándose en la información proporcionada por los correlogramas simple y parcial estimados, un experto ha propuesto el siguiente modelo para predecir Y_t :

$$Y_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \quad a_t \sim NID(0, \sigma^2)$$

¿Crees que la elección es correcta? ¿Qué modelo hubieras propuesto tú? ¿Por qué?



La estimación del modelo propuesto ha producido los siguientes resultados:

$$\hat{\mu} = 0,0275 \quad \hat{\theta}_1 = -1,37 \quad \hat{\theta}_2 = 0,51$$

Calcula las predicciones por punto para los cuatro primeros meses de 2004, con los siguientes datos adicionales:

$$Y_{145} = 23,62 \quad Y_{146} = 23,18 \quad Y_{147} = 24,41$$

Problema 25. Una empresa tiene datos de sus ventas anuales y propone el siguiente modelo:

$$V_t = 100 + y_t$$

donde V_t son las ventas del trimestre t -ésimo, μ son las ventas medias e y_t es un proceso estocástico.

Las correlaciones estimadas para los valores de y con datos de los últimos años son:

k	0	1	2	3	4	5	6	7
$\hat{\rho}_k$	1	0,8	0,66	0,50	0,40	0,33	0,25	0,2

siendo la media de y durante el periodo muestral cero.

El estadístico de la compañía sugiere que el modelo apropiado para la y es:

$$y_t = 0,8y_{t-1} + a_t$$

donde a_t es un ruido blanco con media cero. ¿Estás de acuerdo con su decisión? ¿Por qué?

Dado que las ventas de año 2008 han sido 213, predice las ventas para el año 2009.

Problema 26. Sea $Y_t = 7 + t + X_t$, donde X_t es un proceso estacionario de media cero y función de autocovarianzas γ_k :

- ¿Cuál es la función media de Y_t ? ¿Cuál es la función de autocovarianzas de Y_t ? ¿Es Y_t estacionario?
- ¿Es ΔY_t estacionario?

Problema 27. Sea una serie anual, Y_t que está generada por el siguiente proceso estocástico:

$$Y_t = a + bt + ct^2 + a_t$$

donde a_t es ruido blanco y a, b, c son constantes.

- ¿Es el proceso Y_t estacionario en covarianza? ¿Es el proceso ΔY_t estacionario en covarianza?
- ¿Qué operador de diferencias se ha de aplicar para convertir al proceso Y_t en estacionario?. Discute la invertibilidad del proceso resultante.
- ¿Existe algún otro método para transformar al proceso Y_t en estacionario?

Problema 28. Sea Y_t una serie mensual estacional con factores estacionales constantes, $S_t = S_{t-12}$ y a_t una serie de perturbaciones estacionarias.

- Si suponemos que la serie Y_t consta de una tendencia global lineal y una estacionalidad aditiva, el modelo apropiado es:

$$Y_t = a + bt + S_t + a_t$$

¿Es la serie Y_t estacionaria? ¿Es la serie ΔY_t estacionaria? Demuestra que la aplicación del operador Δ_{12} a la serie Y_t la convierte en estacionaria.

- Si suponemos que la serie Y_t consta de una tendencia global lineal y una estacionalidad multiplicativa, el modelo apropiado es:

$$Y_t = (a + bt)S_t + a_t$$

¿Convierte el operador Δ_{12} en estacionaria a la serie Y_t ? Si no es así encuentra el operador que lo hace.

Problema 29. Para cada uno de los siguientes procesos estocásticos:

$$(1) \quad Y_t = 0,7Y_{t-1} + a_t \quad a_t \sim RBN(0, 9)$$

$$(2) \quad Y_t = 5 + 0,7Y_{t-1} + a_t \quad a_t \sim RBN(0, 9)$$

$$(3) \quad Y_t = Y_{t-1} + a_t - 0,7a_{t-1} \quad a_t \sim RBN(0, 9)$$

$$(4) \quad Y_t = 5 + Y_{t-1} + a_t - 0,7a_{t-1} \quad a_t \sim RBN(0, 9)$$

- ¿Es Y_t estacionario? Si no lo es, ¿existe alguna transformación que lo convierta en estacionario? ¿Qué modelo $ARIMA(p, d, q)$ sigue?
- Comenta las características de cada uno de los procesos por separado.
- Compara las características de los distintos procesos en las siguientes líneas:
 - ▷ Funciones de autocorrelación
 - ▷ Evolución de las series generadas por los distintos modelos.
- Simula una serie temporal de tamaño 50, 100, 200. Comenta los resultados.

Problema 30. Considera los siguientes procesos:

$$1. \quad (1 - 0,99L)(1 - L)Y_t = (1 - 0,5L)a_t$$

$$5. \quad (1 - 0,8L)Y_t = (1 - 0,12L)a_t$$

$$2. \quad (1 - L)^2 Y_t = a_t - 0,81a_{t-1} + 0,38a_{t-2}$$

$$6. \quad (1 - L)Y_t = (1 - 1,5L)a_t$$

$$3. \quad (1 - 0,6L)Y_t = (1 - 1,2L + 0,2L^2)a_t$$

$$7. \quad (1 - L)Y_t = a_t + 0,5a_{t-1}$$

$$4. \quad (1 - 1,1L + 0,8L^2)Y_t = (1 - 1,5L + 0,72L^2)a_t$$

$$8. \quad Y_t = -Y_{t-2} + a_t + 0,2a_{t-1}$$

- ¿Es el proceso Y_t estacionario? ¿Es invertible?
- Si no son estacionarios busca una transformación que los convierta en estacionarios. ¿Qué modelo $ARIMA(p, d, q)$ son?

Problema 31. Sean los modelos siguientes:

$$Y_t = 0,1Y_{t-1} + 0,9Y_{t-2} + a_t$$

$$Y_t = 0,3Y_{t-1} + 0,7Y_{t-2} + a_t - a_{t-1}$$

$$Y_t = -Y_{t-1} + a_t$$

- Analiza las condiciones de estacionariedad e invertibilidad del proceso Y_t .
- En el caso de que no sea estacionario, ¿puede hacerse alguna transformación para convertirlo en estacionario?
- ¿Qué modelo $ARIMA(p, d, q)$ son?

Problema 32.

A. Escribe estos siguientes modelos en forma $ARIMA(p, d, q)$ y en forma de ecuación en diferencias:

1. $(1 - \phi L)(1 - L)Y_t = a_t$
2. $(1 - L)^2 Y_t = (1 - \theta L) a_t$
3. $(1 - \phi L) Y_t = (1 - \theta L) a_t$
4. $(1 - L)Y_t = (1 - \theta_1 L - \theta_2 L^2) a_t$
5. $Y_t = (1 - \theta L^2) a_t$

B. Escribe los siguientes modelos en términos del operador de retardos L :

1. $ARIMA(1, 1, 1)$
2. $ARIMA(0, 2, 1)$
3. $ARIMA(2, 0, 2)$
4. $Y_t = \mu (1 - \phi_1 - \phi_2) + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t$
5. $Y_t = \mu(1 - \phi_1) + \phi_1 Y_{t-1} - \theta_1 a_t + a_t$
6. $Y_t = Y_{t-1} - \theta_1 a_{t-1} - \theta_2 a_{t-2} + a_t$
7. $Y_t = 2Y_{t-1} - Y_{t-2} - \theta_1 a_{t-1} + a_t$
8. $Y_t = Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2}) + \phi_2 (Y_{t-2} - Y_{t-3}) + a_t$

Problema 33. Sea el modelo $ARIMA(2, 1, 0)$ siguiente:

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

- ¿Es el proceso Y_t estacionario?
- ¿Es el proceso $W_t = \Delta Y_t$ estacionario? ¿Qué proceso lineal sigue la serie W_t ?
- Calcula la media y la función de autocorrelación simple teórica del proceso W_t .

Problema 34. A partir de una muestra de 100 datos se ha ajustado el siguiente modelo:

$$\Delta Y_t = (1 - 0,324L) \hat{a}_t$$

Las funciones de autocorrelación simple y parcial de los residuos vienen dadas por:

k	1	2	3	4	5
$\hat{\rho}_k$	-0,508	0,013	-0,032	0,015	0,043
\hat{p}_k	-0,508	-0,329	-0,220	-0,143	-0,093

A la vista de la información anterior, ¿consideras aceptable el modelo ajustado? En caso contrario, ¿cual podría ser la reformulación del modelo?

Problema 35. Considera los siguientes modelos:

1. $(1 - \phi_1 L - \phi_2 L^2) Y_t = (1 - \theta_1 L) a_t$

$$T = 100, \quad Y_{99} = 53, \quad Y_{100} = 56, \quad \hat{a}_{100} = 1,4, \quad \hat{\phi}_1 = 1,4, \quad \hat{\phi}_2 = -0,7, \quad \hat{\sigma}_a^2 = 3,3$$

$$\hat{\theta}_1 = 0,3, \quad \hat{\mu} = 50$$

2. $Y_t = (1 - \theta_1 L - \theta_2 L^2) a_t$

$$T = 100, \quad \hat{a}_{99} = 1,3, \quad \hat{a}_{100} = -2,6, \quad \hat{\theta}_1 = 0,7, \quad \hat{\theta}_2 = -0,5, \quad |\hat{m}u| = 100, \quad \hat{\sigma}_a^2 = 2,5$$

$$\hat{\theta}_1 = 0,3, \quad \hat{\mu} = 50$$

3. $(1 - \phi_1 L)(1 - L) Y_t = a_t$

$$T = 100, \quad Y_{99} = 217, \quad Y_{100} = 232, \quad \hat{\phi}_1 = 0,7, \quad \hat{\sigma}_a^2 = 8$$

4. $(1 - L) Y_t = (1 - \theta_1 L) a_t$

$$T = 100, \quad Y_{100} = 28, \quad \hat{\theta}_1 = 0,5, \quad \hat{a}_{100} = -0,7, \quad \hat{\sigma}_a^2 = 6,7$$

5. $(1 - \phi_1 L) Y_t = (1 - \theta_1 L - \theta_2 L^2) a_t$

$$T = 100, \quad Y_{100} = 103, \quad \hat{a}_{100} = 1,2, \quad \hat{a}_{99} = 0,3, \quad \hat{\phi}_1 = 0,6$$

$$\hat{\theta}_1 = 0,8, \quad \hat{\theta}_2 = -0,3, \quad \hat{\sigma}_a^2 = 2,5$$

6. $(1 - \phi_1 L) (\Delta Y_t - \mu) = a_t$

$$T = 100, \quad Y_{100} = 1, \quad Y_{99} = 1,5, \quad \hat{a}_{100} = 0,005, \quad \hat{a}_{99} = 0,3, \quad \hat{\phi}_1 = 0,8, \quad \hat{\mu} = 0,2$$

7. $(1 - \phi_1 L)(1 - L) Y_t = (1 - \theta_1 L) a_t$

$$T = 100, \quad Y_{100} = 23, \quad Y_{99} = 18, \quad \hat{a}_{100} = 0,2, \quad \hat{a}_{99} = -0,3, \quad \hat{\phi}_1 = 0,8, \quad \hat{\theta}_1 = 0,6$$

Para cada uno de ellos, obtén las predicciones por punto y por intervalo para $\ell = 1, 2, 3, 4, 5$.

Problema 36. Sea el modelo estimado:

$$(1 + 0,6L)(1 - L) Y_t = (1 - 0,5L) a_t + 0,8 \quad \hat{\sigma}_a^2 = 1$$

- Dado $Y_{79} = 40$, $Y_{80} = 41$, $Y_{81} = 41$ y $a_{80} = 0,2$, predice Y_t para los periodos 81, 82 y 83.
- Calcula los errores cuadráticos medios de predicción y establece una banda de confianza del 95 % para la predicción.
- Suponiendo que $Y_{81} = 41$, actualiza la predicción de los periodos 82 y 83.

Problema 37. Contamos con 221 datos para analizar la evolución temporal de la tasa de cambio de Hungría, $Tipo_t$. Con ayuda de los gráficos y la información siguientes, contesta a las preguntas:

- ¿Es estacionaria la serie $Tipo_t$? ¿Por qué?
- ¿Es preciso tomar alguna diferencia a la serie $Tipo_t$ para que sea estacionaria?
- ¿Qué modelo ARMA(p,q) propones para la serie que has elegido como estacionaria?

Estadísticos principales, usando las observaciones 1 - 221 para la variable Tipo (221 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
19,9115	21,0973	10,5407	24,8438
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
3,24292	0,162867	-0,756369	-0,312477
Estadísticos principales, usando las observaciones 1 - 221 para la variable DTipo (220 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
-0,0174748	-0,00373450	-2,4866	2,31328
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,810736	46,3946	0,0622331	0,357541

Contrastes aumentados de Dickey-Fuller, orden 2, para Tipo: tamaño muestral 218 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante

modelo: $(1 - L)y = (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,002

valor estimado de $(a - 1)$: -0,00164723

Estadístico de contraste: $\tau_{nc}(1) = -0,602228$

valor p asintótico 0,4568

contraste con constante

modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,001

valor estimado de $(a - 1)$: -0,0248391

Estadístico de contraste: $\tau_c(1) = -1,43424$

valor p asintótico 0,567

Contrastes aumentados de Dickey-Fuller, orden 2, para (1-L) Tipo:
 tamaño muestral 217 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante

modelo: $(1 - L)y = (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,001

valor estimado de $(a - 1)$: -1,1125

Estadístico de contraste: $\tau_{nc}(1) = -8,98227$

valor p asintótico 1,124e-016

contraste con constante

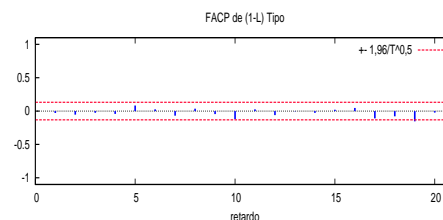
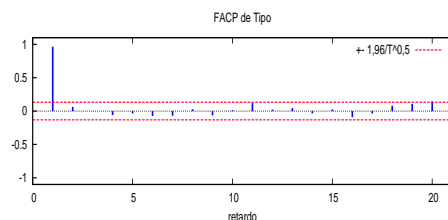
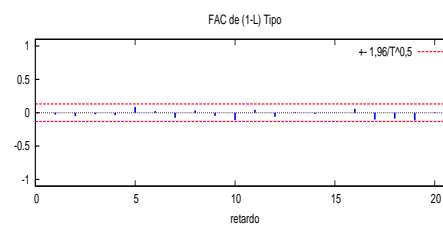
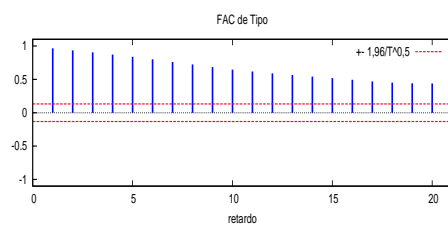
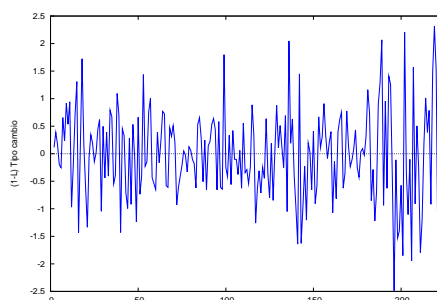
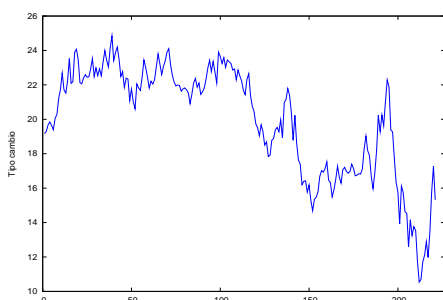
modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,001

valor estimado de $(a - 1)$: -1,11455

Estadístico de contraste: $\tau_c(1) = -8,974$

valor p asintótico 5,749 e-016



Problema 38. Los siguientes datos corresponden a la serie Y_t a la que se ha ajustado el siguiente modelo:

$$\Delta Y_t = 0,5 + a_t - a_{t-1} + 0,5a_{t-2} \quad \hat{\sigma}_a^2 = 0,04$$

t	96	97	98	99	100
Y_t	172	169	164	168	172

- Obtén las predicciones para $Y_{101}, Y_{102}, Y_{103}, Y_{104}, Y_{105}, Y_{106}$.
- ¿Cuál es el comportamiento de la función de predicción conforme $\ell \rightarrow \infty$?
- Calcula los Errores Cuadráticos Medios de Predicción y obtén los intervalos de confianza del 95 % para las predicciones.
- Suponiendo que $Y_{101} = 174$, actualiza las predicciones para $Y_{102}, Y_{103}, \dots, Y_{106}$. Representa gráficamente la serie, las predicciones obtenidas con información hasta $T = 100$ y las actualizadas con información hasta $T = 101$.

Problema 39. Con una muestra de 110 observaciones se ha estimado el siguiente modelo:

$$(1 - 0,5L) \Delta Y_t = (1 + 0,7L) a_t$$

También se dispone de la siguiente información:

t	106	107	108	109	110
y_t	66	65	64	63	68

- Expresa el modelo en la forma de ecuación en diferencias finitas.
- Calcula las predicciones de Y_t para los periodos 111, 112, 113, 114 y 115, sabiendo que $a_{110} = -0,5$.
- Representa la función de predicción. ¿Cuál es su estructura cuando $\ell \rightarrow \infty$?
- Expresa el modelo en la forma $MA(\infty)$.
- Calcula los errores cuadráticos medios de predicción sabiendo que $\sigma_a^2 = 2$.
- Calcula las predicciones por intervalo.
- Suponiendo que $Y_{121} = 70$ actualiza las predicciones.

Problema 40. El director del Servicio de Estudios de una División de la Unión Europea le ha pedido a uno de sus economistas que prediga la evolución de la población de la Unión Europea hasta el año 2000. Con los datos de los últimos 100 años (1896-1995), el economista estima el siguiente modelo:

$$\Delta P_t = (1 - 0,7L + 1,2L^2)a_t$$

- ¿Qué modelo $ARIMA(p, d, q)$ ha propuesto el economista?
- Basándote en este modelo obtén las predicciones de la población hasta el año 2000 por punto y por intervalo, sabiendo que los últimos datos de la población y de la predicción obtenida con un año de antelación son:

Mes	Valor real de Población	Predicción hecha el mes anterior
	P_t	$P_{t-1}(1)$
1992	105	104
1993	109	110.5
1994	111	110
1995	113	113.5

- Dadas las predicciones obtenidas y las características del modelo propuesto, ¿cómo se comporta la población a largo plazo?

Problema 41. Supongamos que el modelo:

$$\Delta Y_t = (1 - 0,5L)e_t \quad (8.3)$$

se utiliza para predecir una serie que de hecho ha sido generada por el modelo:

$$\Delta Y_t = (1 - 0,9L + 0,2L^2)a_t \quad (8.4)$$

- Obtén la función de autocorrelación simple del error de predicción un periodo hacia adelante e_t , obtenidos del modelo $ARIMA(0, 1, 1)$.
- Comenta como esta función de autocorrelación puede utilizarse para identificar un modelo para la serie e_t , llevándonos a la identificación de un $ARIMA(0, 1, 2)$ para la serie Y_t .

Problema 42. El servicio de estudios de un Banco quiere predecir el índice de Bolsa de New York mediante un modelo $ARIMA(p, d, q)$. Para ello cuenta con 78 observaciones diarias del índice de bolsa ($Y_t = DowJones_t$). Con la información presentada en los gráficos y tablas siguientes:

Estadísticos principales, usando las observaciones 1 - 78 para la variable DowJones (78 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
115,683	113,755	108,540	124,140
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
5,50606	0,0475960	0,246937	-1,6263
Estadísticos principales, usando las observaciones 1 - 78 para la variable (1-L) DowJones (77 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
0,133636	0,130000	-0,770000	1,53000
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,426950	3,19487	0,521255	0,939981
Estadísticos principales, usando las observaciones 1 - 78 para la variable $(1 - L)^2$ DowJones (76 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
-0,00684211	-0,0100000	-1,2100	1,23000
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,447974	65,4731	-0,112599	0,141083

Contrastes aumentados de Dickey-Fuller, orden 2, para (1-L) DowJones: tamaño muestral 74 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante

modelo: $(1 - L)y = (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,018

valor estimado de $(a - 1)$: -0,386798

Estadístico de contraste: $\tau_{nc}(1) = -2,95062$

valor p asintótico 0,00309

contraste con constante

modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: -0,008

valor estimado de $(a - 1)$: -0,46437

Estadístico de contraste: $\tau_c(1) = -3,17663$

valor p asintótico 0,02141

Contrastes aumentados de Dickey-Fuller, orden 2, para $(1 - L)^2$ DowJones: tamaño muestral 73 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante

modelo: $(1 - L)y = (a-1)*y(-1) + \dots + e$

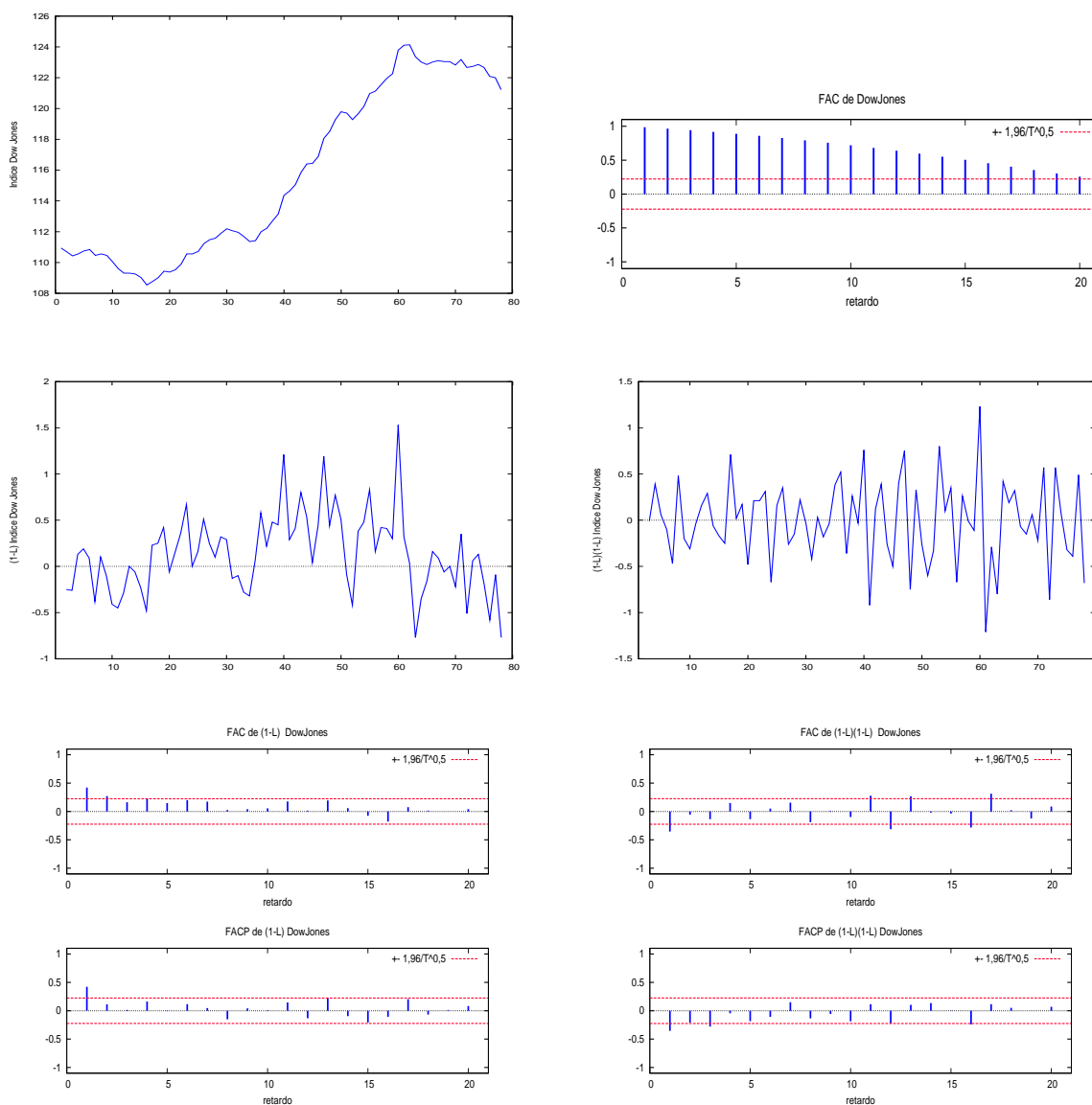
Coef. de autocorrelación de primer orden de e: -0,021

valor estimado de $(a - 1)$: -2,14076

Estadístico de contraste: $\tau_{nc}(1) = -8,09521$

valor p asintótico 2,654e-014

Gráfico 8.1: Gráficos y correlogramas



- Analiza la estacionariedad de la serie Y_t , es decir, ¿es la serie Y_t estacionaria en media o es preciso hacer alguna transformación para convertirla en estacionaria? ¿Por qué?
- ¿Qué modelo (o modelos) $ARMA(p, q)$ propones para la serie estacionaria? ¿Por qué?

El economista consultado por el servicio de estudios, basándose en la misma información de la que tú dispones, ha identificado los siguientes modelos:

Modelo 1	$\Delta Y_t \sim AR(1)$
Modelo 2	$\Delta Y_t \sim ARMA(1,1)$
Modelo 3	$\Delta^2 Y_t \sim MA(1)$
Modelo 4	$\Delta^2 Y_t \sim ARMA(1,1)$

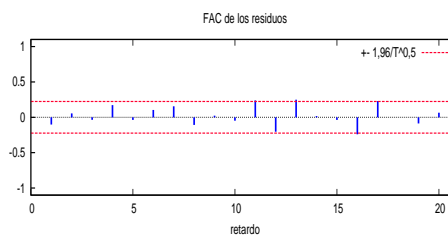
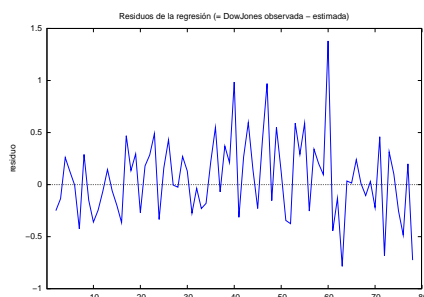
- Discute si te parecen razonables los cuatro modelos que él propone.
- Analiza los resultados que obtiene en la estimación de cada uno de los modelos. ¿Se ajustan los modelos bien a los datos?
- ¿Cuál de los cuatro modelos estimados elegirías para predecir el futuro de la serie Y_t ? ¿Por qué?

Modelo 1: estimaciones ARIMA utilizando las 77 observaciones 2–78
Variable dependiente: $(1 - L)\text{DowJones}$
Desviaciones típicas basadas en el Hessiano

	Coefficiente	Desv. típica	estadístico t	valor p
ϕ_1	0,499168	0,100102	4,9866	0,0000

Media de la var. dependiente	0,133636
D.T. de la variable dependiente	0,426950
media de las innovaciones	0,0619386
Varianza de las innovaciones	0,149332
Log-verosimilitud	-36,190

	Real	Imaginaria	Módulo	Frecuencia
AR				
Raíz	1	2,0033	0,0000	2,0033
				0,0000



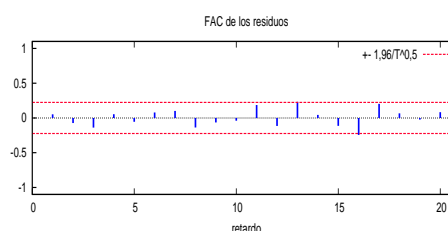
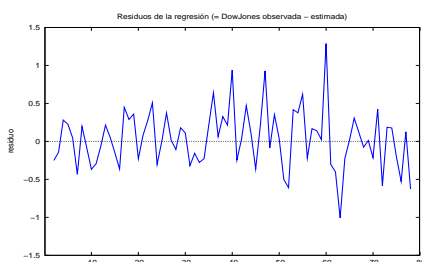
Modelo 2: estimaciones ARIMA utilizando las 77 observaciones 2–78

Variable dependiente: $(1 - L)\text{DowJones}$
 Desviaciones típicas basadas en el Hessiano

	Coefficiente	Desv. típica	estadístico t	valor p
ϕ_1	0,850965	0,138465	6,1457	0,0000
θ_1	-0,526265	0,254785	-2,0655	0,0389

Media de la var. dependiente	0,133636
D.T. de la variable dependiente	0,426950
media de las innovaciones	0,0341779
Varianza de las innovaciones	0,143363
Log-verosimilitud	-34,689

			Real	Imaginaria	Módulo	Frecuencia
AR						
	Raíz	1	1,1751	0,0000	1,1751	0,0000
MA						
	Raíz	1	1,9002	0,0000	1,9002	0,0000



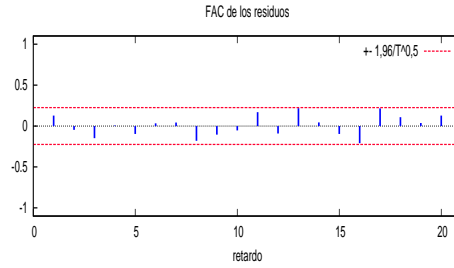
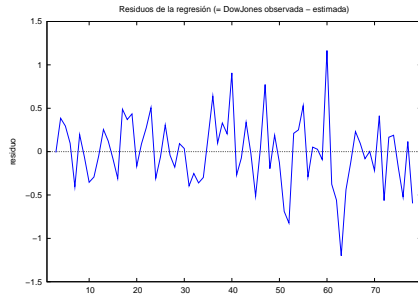
Modelo 3: estimaciones ARIMA utilizando las 76 observaciones 3–78

Variable dependiente: $(1 - L)^2\text{DowJones}$
 Desviaciones típicas basadas en el Hessiano

	Coefficiente	Desv. típica	estadístico t	valor p
θ_1	-0,715732	0,113333	-6,3153	0,0000

Media de la var. dependiente	-0,00684211
D.T. de la variable dependiente	0,447974
media de las innovaciones	-0,00664944
Varianza de las innovaciones	0,150368
Log-verosimilitud	-36,201

			Real	Imaginaria	Módulo	Frecuencia
MA						
	Raíz	1	1,3972	0,0000	1,3972	0,0000

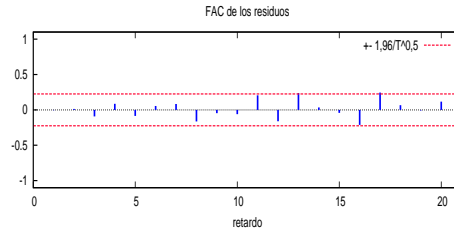
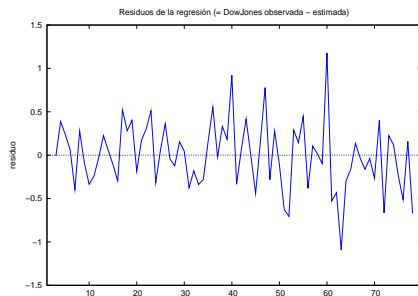


Modelo 4: estimaciones ARIMA utilizando las 76 observaciones 3–78
 Variable dependiente: $(1 - L)^2 \text{DowJones}$
 Desviaciones típicas basadas en el Hessiano

	Coefficiente	Desv. típica	estadístico t	valor p
ϕ_1	0,247838	0,144741	1,7123	0,0868
θ_1	-0,839131	0,0819245	-10,2427	0,0000

Media de la var. dependiente	-0,00684211
D.T. de la variable dependiente	0,447974
media de las innovaciones	-0,00467602
Varianza de las innovaciones	0,144924
Log-verosimilitud	-34,848

			Real	Imaginaria	Módulo	Frecuencia
AR	Raíz	1	4,0349	0,0000	4,0349	0,0000
	Raíz	1	1,1917	0,0000	1,1917	0,0000



Problema 43. Disponemos de 250 observaciones, desde octubre de 1978 hasta julio de 1999, para analizar y predecir la serie de tipo de interés a largo plazo ($Y_t = Interes_t$).

- Analiza la estacionariedad en media de la serie Y_t en base a los gráficos y la información presentada. ¿Es preciso hacer alguna transformación a la serie Y_t para que sea estacionaria? ¿Por qué?
- Propón razonadamente uno (o varios) modelos $ARIMA(p, d, q)$ para la serie estacionaria.

Estadísticos principales, usando las observaciones 1978:10 - 1999:07 para la variable Interes (250 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
11,4799	11,2492	5,35361	17,2649
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
3,41329	0,297328	-0,0707997	-1,0438

Estadísticos principales, usando las observaciones 1978:10 - 1999:07 para la variable (1-L) Interes (249 observaciones válidas)			
Media	Mediana	Mínimo	Máximo
-0,0477041	-0,0499643	-0,130352	0,0475899
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
0,0360135	0,754934	0,287116	-0,273057

Contrastes aumentados de Dickey-Fuller, orden 2, para (1-L) Interes: tamaño muestral 246 hipótesis nula de raíz unitaria: $a = 1$

contraste sin constante

modelo: $(1 - L)y = (a-1)*y(-1) + \dots + e$

Coef. de autocorrelación de primer orden de e: 0,057

valor estimado de $(a - 1)$: -0,124977

Estadístico de contraste: $\tau_{nc}(1) = -3,81022$

valor p asintótico 0,0001392

contraste con constante

modelo: $(1 - L)y = b_0 + (a-1)*y(-1) + \dots + e$

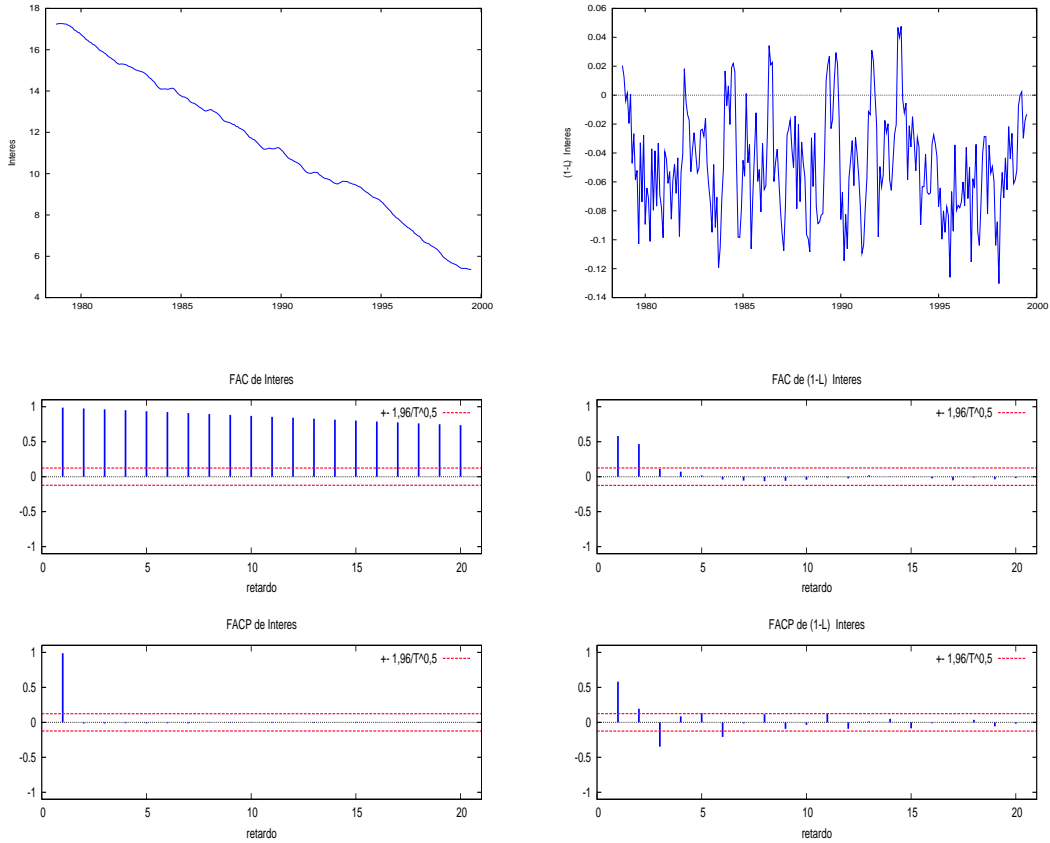
Coef. de autocorrelación de primer orden de e: 0,034

valor estimado de $(a - 1)$: -0,453041

Estadístico de contraste: $\tau_c(1) = -7,90484$

valor p asintótico 9,481e-013

Gráfico 8.2: Gráficos y correlogramas



Problema 44. Siguiendo con los datos del Problema anterior, el economista de un servicio de estudios ha propuesto los siguientes modelos para el tipo de interés:

$$\text{Modelo 1} \quad Y_t \sim ARIMA(1, 1, 0)$$

$$\text{Modelo 2} \quad Y_t \sim ARIMA(1, 1, 1)$$

$$\text{Modelo 3} \quad Y_t \sim ARIMA(0, 1, 2)$$

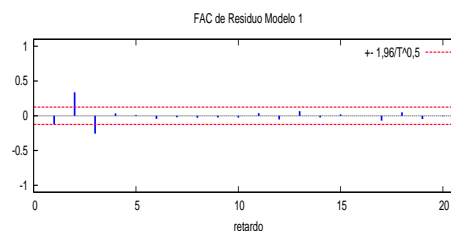
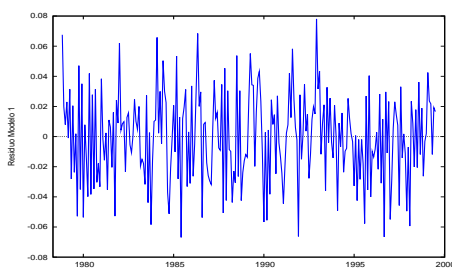
- ¿Te parecen apropiados los tres modelos propuestos? ¿Por qué?
- Basándote en los resultados de la estimación de los tres modelos (tablas y correlogramas simple y parcial de los residuos) que se presentan a continuación, ¿cuál de los tres modelos elegirías para predecir el futuro de la serie Y_t ?

Modelo 1: estimaciones ARIMA utilizando las 249 observaciones 1978:11–1999:07
Variable dependiente: $(1 - L)$ Interes

Variable	Coefficiente	Desv. típica	Estadístico t	valor p
const	-0,0471185	0,00447042	-10,5401	0,0000
ϕ_1	0,589559	0,0515826	11,4294	0,0000

Media de la var. dependiente	-0,0477041
D.T. de la variable dependiente	0,0360135
media de las innovaciones	-0,000160249
Varianza de las innovaciones	0,000847352
Log-verosimilitud	527,108

	Real	Imaginaria	Módulo	Frecuencia
AR				
Raíz	1	1,6962	0,0000	1,6962
				0,0000

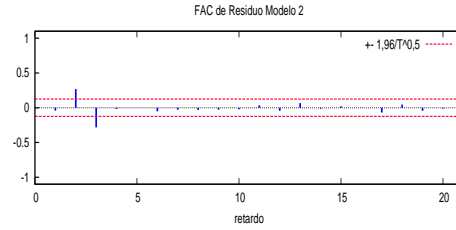
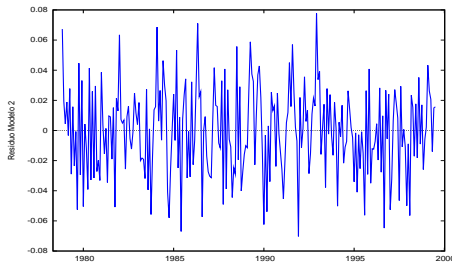


Modelo 2: estimaciones ARIMA utilizando las 249 observaciones 1978:11–1999:07
Variable dependiente: $(1 - L)$ Interes

Variable	Coefficiente	Desv. típica	Estadístico t	valor p
const	-0,0468605	0,00506103	-9,2591	0,0000
ϕ_1	0,699453	0,0628214	11,1340	0,0000
θ_1	-0,161248	0,0736799	-2,1885	0,0286

Media de la var. dependiente	-0,0477041
D.T. de la variable dependiente	0,0360135
media de las innovaciones	-0,000220568
Varianza de las innovaciones	0,000831130
Log-verosimilitud	529,499

	Real	Imaginaria	Módulo	Frecuencia
AR				
Raíz	1	1,4297	0,0000	1,4297
				0,0000
MA				
Raíz	1	6,2016	0,0000	6,2016
				0,0000

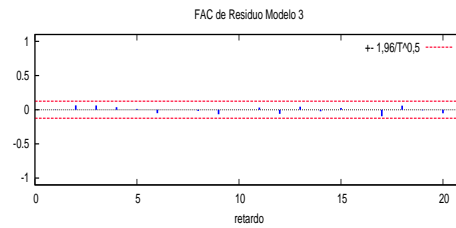
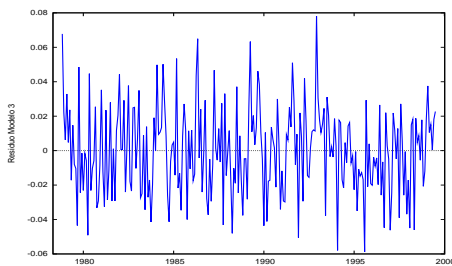


Modelo 3: estimaciones ARIMA utilizando las 249 observaciones 1978:11–1999:07
Variable dependiente: $(1 - L)$ Interes

Variable	Coefficiente	Desv. típica	Estadístico t	valor p
const	-0,0473130	0,00375620	-12,5960	0,0000
θ_1	0,615351	0,0437706	14,0586	0,0000
θ_2	0,746517	0,0402836	18,5315	0,0000

Media de la var. dependiente	-0,0477041
D.T. de la variable dependiente	0,0360135
media de las innovaciones	-1,85507e-05
Varianza de las innovaciones	0,000634269
Log-verosimilitud	562,501

		Real	Imaginaria	Módulo	Frecuencia
MA					
Raíz	1	-0,4121	-1,0815	1,1574	-0,3079
Raíz	2	-0,4121	1,0815	1,1574	0,3079



Sabiendo que:

$$Y_{98,9} = 5,37 \quad Y_{98,10} = 5,35 \quad a_{98,9} = 0,018 \quad a_{98,10} = 0,023$$

predice el tipo de interés a largo plazo desde Noviembre de 1998 hasta Diciembre de 1999. Calcula los errores cuadráticos medios de predicción ¿Hacia donde tiende la predicción del tipo de interés a largo plazo?

Problema 45. Supongamos que la serie Y_t sigue el siguiente modelo:

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

- ¿Qué modelo sigue la serie Y_t ? ¿Es la serie Y_t estacionaria?
- ¿Es estacionaria la serie $Z_t = \Delta Y_t$? ¿Qué modelo sigue?
- Obtén la función de autocorrelación simple teórica para la serie Z_t , ρ_k , $k = 1, 2, \dots$. Explica cual sería la estructura de la función de autocorrelación parcial teórica.
- En base a una muestra de 100 observaciones para Z_t , hemos obtenido que los dos primeros coeficientes de autocorrelación simple estimados son $r_1 = 0,93$ y $r_2 = 0,81$. Estima los parámetros del modelo para Z_t .
- Supongamos que las últimas observaciones de la serie Y_t son:

t	95	96	97	98	99	100
Y_t	145	141	142	148	150	143

y que el error de predecir Y_{100} con información hasta $t = 99$ es -0.5. Basándote en el modelo para Y_t con parámetros estimados:

- ▷ Predice los valores de Y_t para $t = 101, 102, 103, 104, 105, 106$.
- ▷ ¿Cual es el comportamiento de la función de predicción a largo plazo, es decir, cuando $\ell \rightarrow \infty$?
- ▷ Calcula los errores cuadráticos medios de predicción, sabiendo que la varianza del error de predicción un periodo hacia adelante es 2.
- ▷ Si sabemos que $Y_{101} = 155$, actualiza el resto de las predicciones.

Problema 46. Compara los modelos estacionales siguientes:

$$1. Y_t = (1 - \theta_1 L)(1 - \Theta_1 L^{12}) a_t \quad 2. Y_t = (1 - \theta_1 L - \theta_{12} L^{12} - \theta_{13} L^{13}) a_t$$

¿Se puede distinguir entre las funciones de autocorrelación del modelo (1) y del modelo (2)?

Problema 47. Con una muestra de 120 observaciones se ha estimado el siguiente modelo:

$$(1 - 0,5L)(1 + 0,7L^4)(1 - L) Y_t = (1 - 0,5L^4) a_t$$

También se dispone de la siguiente información:

t	110	111	112	113	114	115	116	117	118	119
Y_t	56	59	58	60	63	66	65	64	63	68

Se pide:

- Expresar el modelo en la forma de ecuación en diferencias finitas.
- Expresar el modelo en la forma $MA(\infty)$.
- Efectuar la predicción de Y_t para los periodos 121, 122, 123, 124 y 125.
- Calcular los errores cuadráticos medios de predicción y las predicciones por intervalo.
- Suponiendo que $Y_{121} = 70$, actualiza las predicciones para los periodos 122, 123, 124 y 125.

Problema 48. Deriva la función de autocorrelación teórica para los siguientes procesos:

$$1. Y_t = (1 - 0,8L^{12})(1 + 0,6L) a_t \quad 2. Y_t = (1 - 0,9L^4)(1 - 0,7L) a_t$$

Problema 49. ¿Cuáles de los siguientes modelos estimados son estacionarios e invertibles?

1. $(1 - 0,8L)(1 - L^4)Y_t = (1 - 0,8L^4) \hat{a}_t$
2. $Y_t = (1 - 0,4L^{12} - 0,3L^{24})(1 - 0,5L^2) \hat{a}_t$
3. $(1 - 1,2L + 0,5L^2)(1 - 0,5L^{12})(1 - L)Y_t = \hat{a}_t$

Problema 50. Escribe los siguientes modelos en términos del operador de retardos L y en forma de ecuación en diferencias:

1. $ARIMA(0, 1, 1)(0, 1, 1)_s$
2. $ARIMA(1, 1, 0)(0, 1, 1)_4$
3. $ARIMA(2, 0, 0)(0, 0, 1)_{12}$
4. $ARIMA(0, 2, 2)(2, 1, 2)_{12}$
5. $ARIMA(0, 1, 0)(0, 1, 1)_7$
6. $ARIMA(1, 1, 1)(1, 1, 1)_4$

Lecturas recomendadas

- a) Aznar, A. y F.J. Trívez (1993). *Métodos de Predicción en Economía*. Volumen II. Ed. Ariel. Barcelona.
- b) Box, G.E.P. y G.M. Jenkins (1970). *Time series analysis: forecasting and control*. San Francisco, Holden Day.
- c) Franses, P.H. (1998). *Time series models for business and economic forecasting*. Cambridge University Press. Cambridge.
- d) Peña, D. (2005). *Análisis de series temporales*. Alianza editorial. Madrid.
- e) Uriel, E. (2000). *Introducción al análisis de series temporales*. Editorial AC. Madrid.