Series Temporales

Teresa Villagarcía

1. Introducción.

El estudio de series temporales tiene por objeto analizar la evolución de una variable a través del tiempo. La diferencia esencial entre las series temporales y los análisis no temporales (Estadística descriptiva, Diseño de experimentos o Regresión) es que en los análisis previos no importaba el orden en que estaban tomadas las observaciones y éste se podía variar sin problemas. En series temporales el orden es muy importante y variarlo supone cambiar la información contenida en la serie.

Es muy importante conocer la periodicidad de los datos de las series que se están analizando. La periodicidad puede ser:

- Anual: Se toma un dato cada año.
- Mensual: Se toma un dato cada mes.
- Semanal: Se toma un dato cada semana.
- Diaria: Se toma un dato cada día.

Indudablemente existen muchos más tipos de periodicidad, como semestral o trimestral. El tipo de periodicidad va a ser algo importante en el análisis de la serie, y aparecerán determinadas pautas debidas a ella.

Existen numerosos ejemplos de series temporales:

1. **Serie del IPC en España**. Esta serie puede ser anual o mensual. Por ejemplo la serie de ipc anual desde 1975-1997 tiene 17 datos. La serie mensual desde Enero de 1975 hasta Diciembre de 1997 tiene 17x12 datos.

2. **Serie de Temperaturas en Madrid**. Esta serie suele ser mensual. Si fuera anual perderíamos mucha información, pues un invierno extremadamente frío puede compensarse con un verano muy cálido, de modo que la temperatura media del año sea templada.

3. **Serie de ventas de una empresa**. Este tipo de series puede ser anual, mensual o semanal.

4. **Demanda de Energía eléctrica**. Esta serie suele obtenerse con periodicidad horaria.

5. **Series de cotizaciones de bolsa**. Este tipo de series se obtienen con la periodicidad que se quiera. Así, por ejemplo podemos estudiar el índice Nikkei diario, o con cualquier otra periodicidad.

2. Descripción de series temporales.

Las series suelen representarse mediante un gráfico que muestra su evolución con el tiempo. Cuando se representa una serie se suele prestar atención a una serie de características.

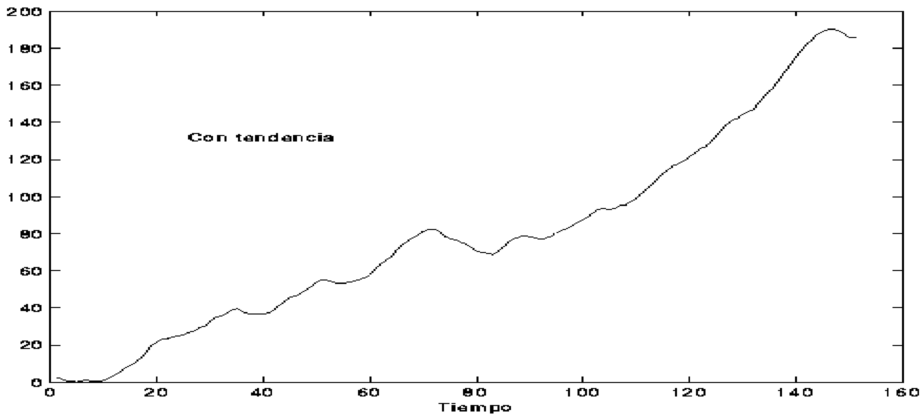


Figura 1: Serie con tendencia

La figura 1 muestra una serie con **TENDENCIA**. La tendencia implica que la serie tiende a crecer o a decrecer a largo plazo. Cuando una serie permanece más o menos constante, oscilando en torno a un valor, decimos que la serie no tiene tendencia. La figura 2 muestra una serie que no tiene tendencia.

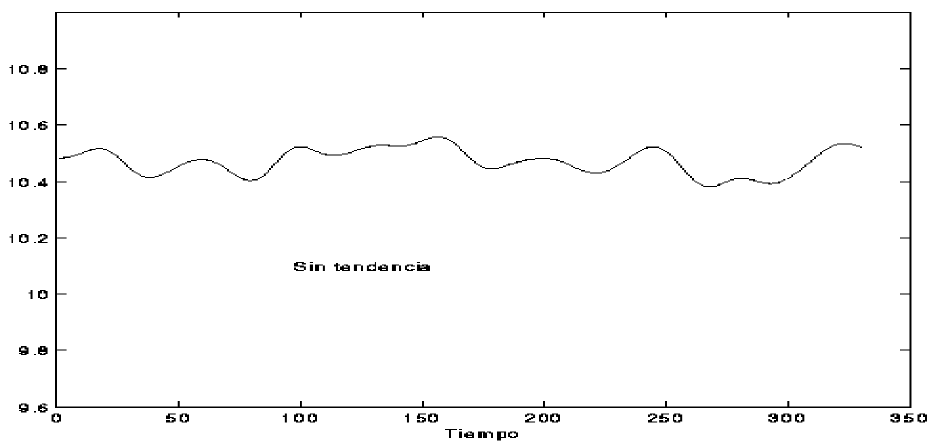


Figura 2: Serie sin tendencia

Otra característica de las series es su variabilidad. Decimos que una serie es **HOMOCEDÁSTICA**, si su variabilidad se mantiene constante a lo largo de la serie. La serie de la figura 2 es homocedástica pues su variabilidad no aumenta con el tiempo. Cuando la variabilidad de la serie aumenta o disminuye a lo largo del tiempo, decimos que la serie es **HETEROCEDÁSTICA**. La figura 3 muestra una serie heterocedástica en la que la varianza va aumentando con el tiempo.

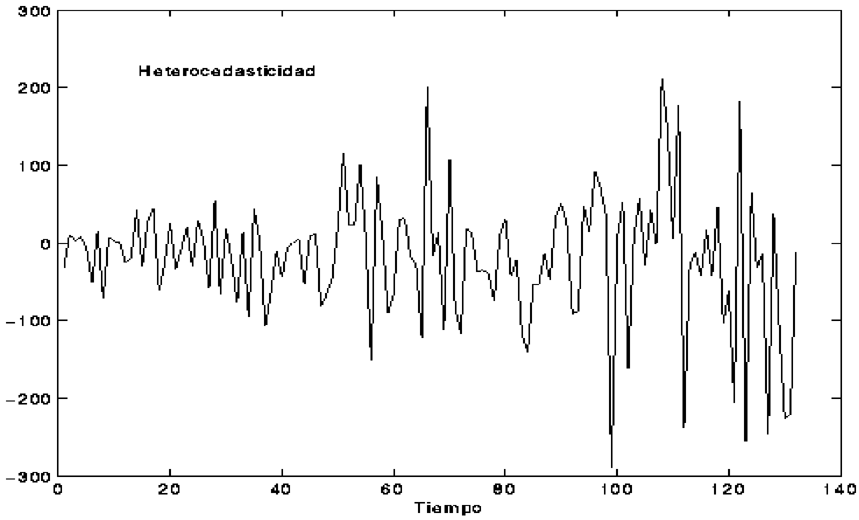


Figura 3: Serie con heterocedasticidad

Indudablemente una serie puede tener tendencia y ser heterocedástica. La figura 4 muestra un caso así.

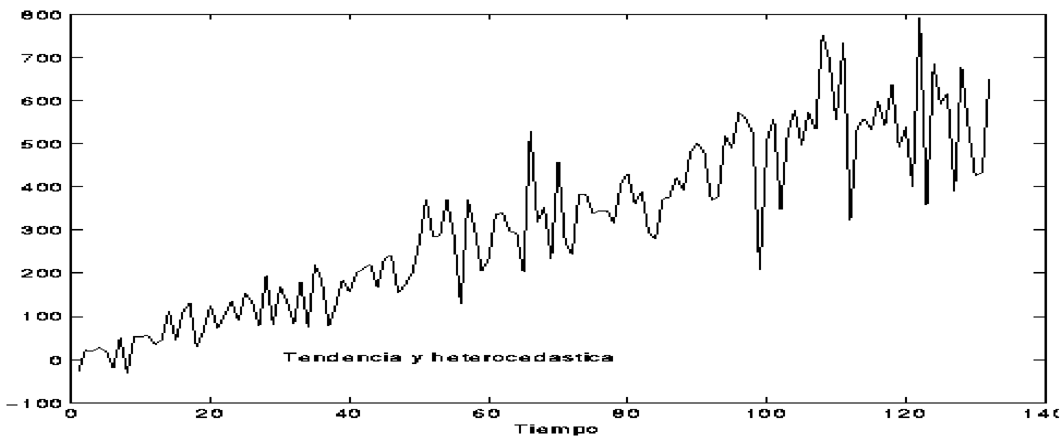


Figura 4: Serie con tendencia y Heterocedasticidad

2.1 Estacionalidad.

Las series que hemos visto en las figuras anteriores, no presentan efectos cíclicos. Sin embargo, si se estudian series de periodicidad menor que la anual suelen aparecer determinados ciclos característicos. Así, la serie de temperaturas máximas medias por meses en ciudades tiene el aspecto de la figura 5, que muestra las temperaturas de Madrid. Esta serie presenta un ciclo estacional, es decir tiene una estructura que se repite año tras año. El fenómeno del ciclo estacional es muy corriente en la mayoría de series socioeconómicas y biológicas.

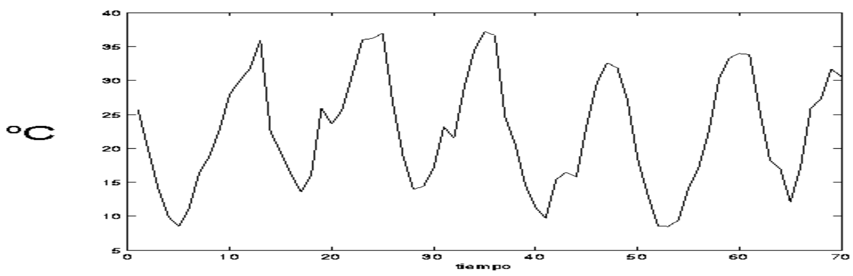


Figura 5: Temperatura en Madrid. Meses.

2.2 Análisis descriptivo de la serie.

Cuando se aborda el análisis de series temporales, utilizamos determinadas técnicas descriptivas que no son habituales en otros métodos estadísticos.

El análisis descriptivo previo tiene por objetivo identificar si la serie tiene tendencia, heterocedasticidad y estacionalidad. Para ello se utiliza:

- El gráfico de la serie
- Gráfico de la descomposición estacional.

El gráfico de la serie permite detectar la existencia de tendencia fácilmente. El gráfico de descomposición estacional, que muestra la figura 6, proporciona para cada uno de los meses el valor medio de las temperaturas. Ese valor medio se representa mediante una recta horizontal. Así, en la figura 6 (que comienza en el mes de Septiembre) puede observarse que la temperatura máxima media de septiembre es de unos 30 grados centígrados, mientras que en Enero es de 9 grados. La observación de las líneas horizontales permite por tanto reconstruir el ciclo estacional con claridad.

Las líneas verticales que salen de cada horizontal son las observaciones individuales. Así, el primer septiembre fue especialmente frío.

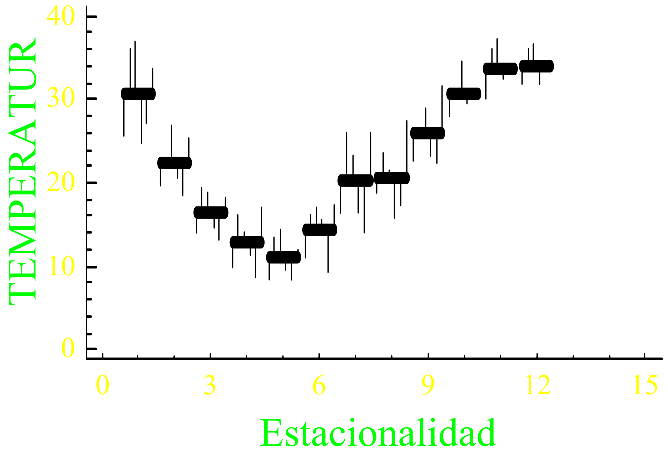


Figura 6: Descomposición estacional de la serie de Temperaturas de Madrid.

3. Funciones de autocorrelación.

Hasta ahora hemos descrito el aspecto de la serie. Sin embargo, cuando se quiere analizar la serie es necesario identificar la estructura que la genera, es decir cómo influyen las observaciones del pasado en las observaciones del futuro. Para identificar esta dependencia utilizamos dos fuentes de información, la Función de Autocorrelación Simple (FAS) y la Función de Autocorrelación Parcial (FAP)

3.1 Función de autocorrelación simple.

La función de autocorrelación simple se una serie proporciona la estructura de dependencia lineal de la misma.

Si denominamos z_t a la serie temporal, los valores que se observan van a ser:

$$z_1, z_2, \cdots z_{t-2}, z_{t-1}, z_t, z_{t+1}, \cdots$$

donde z_1 representa el primer valor de la serie, z_2 el segundo, y z_t será el valor actual de la serie. De este modo z_{t+1} representa el valor de la serie para próximo periodo, es decir un valor futuro.

Si denominamos z_1 influye sobre z_2 como $z_1 \rightarrow z_2$, la función de autocorrelación simple tiene por objetivo estudiar cómo se influyen las diversas observaciones:

$$z_1 \rightarrow z_2 \rightarrow \cdots \rightarrow z_{t-1} \rightarrow z_t \rightarrow z_{t+1} \rightarrow \cdots$$

La idea de la función de autocorrelación es proporcionar el coeficiente de correlación entre las observaciones separadas un número determinado de periodos. Así la FAS, va a ser una sucesión de números

$$\rho_1, \rho_2, \cdots \rho_k, \cdots$$

que representan cómo influye una observación sobre la siguiente (ρ_1), sobre la segunda posterior (ρ_2) o sobre la *k* periodos posterior (ρ_k).

$\rho_1 \rightarrow$	Cómo influye una observación sobre la siguiente. $z_i \rightarrow z_{i+1}$
$\rho_2 \rightarrow$	Cómo influye una observación sobre la dos periodos posterior. $z_i \rightarrow z_{i+2}$
$\rho_3 \rightarrow$	Cómo influye una observación sobre la tres periodos posterior. $z_i \rightarrow z_{i+3}$

Y así sucesivamente

Los coeficientes de la FAS, $\rho_1, \rho_2, \cdots \rho_k, \cdots$ están acotados entre $[-1, +1]$. Cuando un ρ_i vale cero, quiere decir que no existe efecto entre una observación y la *i* posiciones posterior.

Si ρ_i es próximo a 1 *indica que hay mucha relación entre una observación y la i posiciones posterior*, *y que esa relación es positiva*.

Si ρ_i es próximo a -1 *indica que hay mucha relación entre una observación y la i posiciones posterior*, *y que esa relación es negativa*.

La FAS proporciona cómo una observación influye sobre las posteriores.

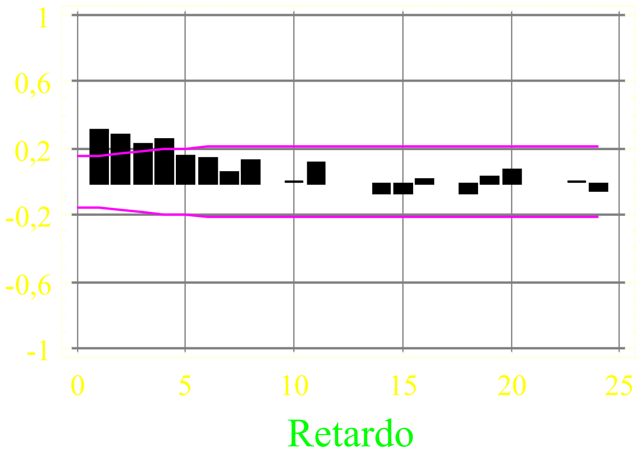


Figura 7: Función de Autocorrelación Simple muestral. FAS

La figura 7 muestra una FAS. Como puede observarse la función de autocorrelación proporciona los coeficientes de correlación de la serie consigo misma -de ahí el nombre de Autocorrelación- para distintos retardos. Es el caso de la figura, se ve que los coeficientes (o palos de la función) son significativos para retardos bajos. Las bandas horizontales que se observan en la figura proporcionan los límites para considerar significativo un retardo. Es decir si un palo está dentro de las bandas lo consideraremos no significativo en general. En el caso de la figura 7, los palos de retardos superiores no son significativos. Esto indica que una observación no influye excesivamente sobre las que están muy alejadas de ella.

La FAS tiene un problema, y es que si por ejemplo ρ_1 es distinto de cero, entonces

$$z_1 \longrightarrow z_2 \longrightarrow \cdots \longrightarrow z_{t-1} \longrightarrow z_t \longrightarrow z_{t+1} \longrightarrow \cdots$$

es decir existe una cadena de influencia separada por un retardo. Pero si $z_1 \longrightarrow z_2$ y $z_2 \longrightarrow z_3$, entonces $z_1 \longrightarrow z_3$. Por tanto, la FAS en general, si ρ_1 es distinto de cero, encontrará que ρ_2, ρ_3 ... etc, serán distintos de cero.

Sin embargo es necesario distinguir entre varias cadenas de influencia posibles:

- La cadena de influencia general, a través de ρ_1.

- Las cadenas de influencia directa. Es decir cómo influye z_1 sobre z_3 directamente, es decir **SIN PASAR A TRAVÉS DE** z_2.

Para resolver este problema se construye la Función de Autocorrelación Parcial.

3.2 Función de Autocorrelación Parcial (FAP).

La función de autocorrelación parcial proporciona la relación directa que existe entre observaciones separadas por *k* retardos. Esta es una información muy valiosa sobre la estructura de la serie, ya que elimina el problema que presentaba la función de autocorrelación simple de que si $z_1 \rightarrow z_2$ y $z_2 \rightarrow z_3$, entonces $z_1 \rightarrow z_3$. En la función de autocorrelación simple, el primer palo (relación entre z_1 y z_2 o z_2 y z_3) será significativo. Y el segundo también, ya que si $z_1 \rightarrow z_2$ y $z_2 \rightarrow z_3$, entonces $z_1 \rightarrow z_3$.

En la FAP, esto no ocurre. el primer palo, será significativo, y el segundo no lo será. La figura 8 muestra una función de autocorrelación parcial.

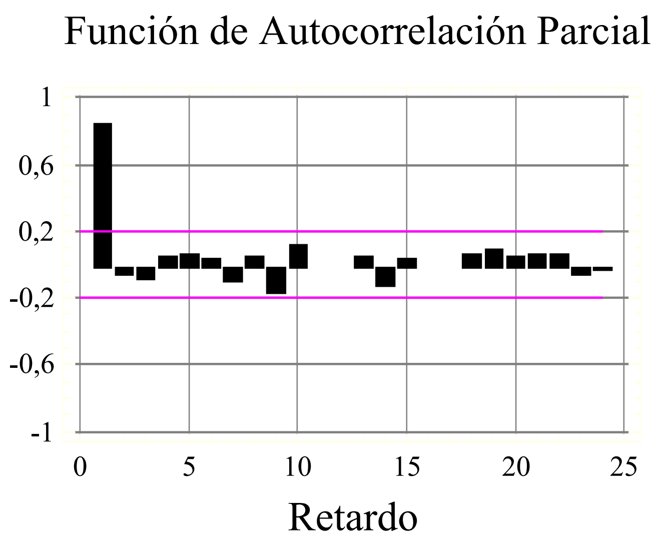


Figura 8: Función de Autocorrelación Parcial (FAP) muestral.

Como puede observarse, el primer retardo es significativo, mientras que ninguno de los demás lo es. Esto implica que la serie presenta relación directa entre una observación y la siguiente, pero no existe ninguna otra relación directa.

4. Series estacionarias.

Se define una serie como estacionaria cuando cumple las siguientes características:

- No tiene tendencia.
- Es homocedástica.

- No tiene ciclos estacionales

- La estructura de dependencia se mantiene constante, es decir si una observación influye sobre la posterior, ésto ocurre SIEMPRE y no únicamente entre las observaciones 23 y 24. Esta condición es importante para modelizar la serie, pues si el fenómeno que genera la serie cambia, es imposible que podamos prever la evolución de la serie.

- La influencia de las observaciones sobre las posteriores decrece con el tiempo.

Un tipo especial de serie estacionaria es la serie denominada ruido blanco. Un ruido blanco es una serie estacionaria tal que ninguna observación influye sobre las siguientes.

La FAS y la FAP del ruido blanco es una sucesión de palos no significativos. Las figuras 9 y 10 muestran respectivamente la FAS y FAP de un ruido blanco.

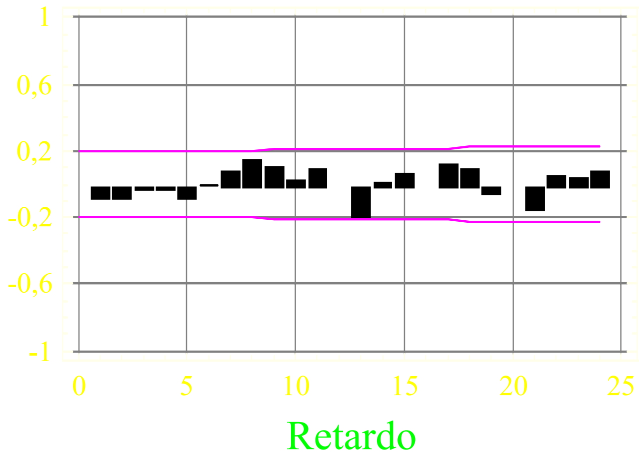


Figura 9: Función de Autocorrelación Simple de un ruido blanco.

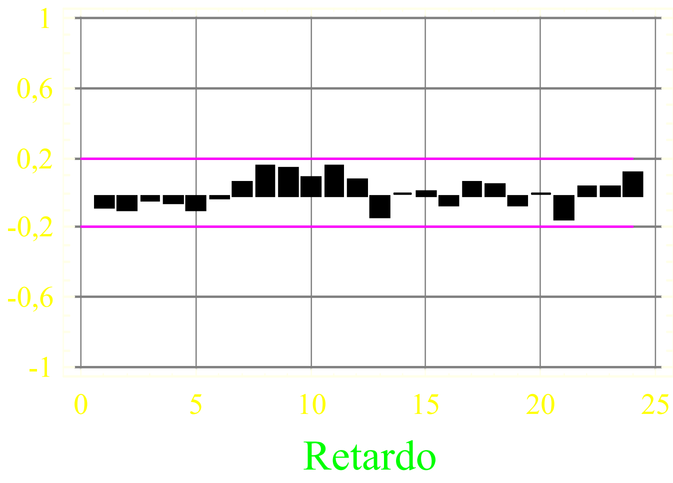


Figura 10: Función de Autocorrelación Parcial de un ruido blanco.

Las técnicas que vamos a presentar se aplican únicamente a series estacionarias. Si la serie no es estacionaria, es necesario transformarla hasta conseguir que lo sea.

4.1 Transformaciones para conseguir estacionariedad.

En la sección anterior se ha explicado cómo es una serie estacionaria. Cuando la serie no lo es, es preciso transformarla. Este proceso es muy sencillo. Para eliminar la tendencia se toman una o varias diferencias en la serie. Una serie se diferencia restando a cada observación la observación anterior:

$$w_t = z_t - z_{t-1}$$

Es decir, la serie w_t se obtiene restando de cada observación de z_t la observación anterior. Evidentemente la serie diferenciada w tiene una observación menos que la serie original z, ya que la primera observación se pierde. Si el gráfico de una serie muestra tendencia, se diferencia la serie y se comprueba si ha perdido la tendencia. En caso de que no la haya perdido se diferencia una segunda vez. Es muy raro necesitar más de una o dos diferencias para eliminar la tendencia de una serie.

Normalmente se utiliza la nomenclatura $\nabla z_t = w_t$ para representar la serie con una diferencia. Vamos a ver un ejemplo:

Las figuras 11 y 12 muestran respectivamente la evolución de las cotizaciones de IBM en la bolsa de Nueva York y su FAS (Datos tomados del programa Statgraphics). Puede observarse que la serie tiene tendencia. Y en la FAS se observa una pauta que aparece cuando la serie tiene tendencia y por tanto no es estacionaria: los palos disminuyen muy lentamente.

Las figuras 13 y 14 muestran la serie ∇*IBM*. Como puede observarse, la serie ya no tiene tendencia, y su FAS decrece rápidamente. Consecuentemente, podemos ver si una serie es estacionaria en su gráfico y en su FAS. En caso de que no lo sea, tomaremos tantas diferencias como sea preciso para que la serie resulte estacionaria.

Figura 11: Acciones de IBM. Tiene tendencia

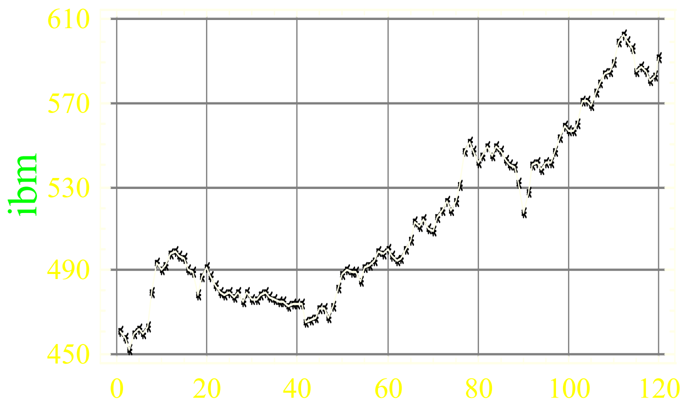


Figura 12: FAS para IBM con tendencia

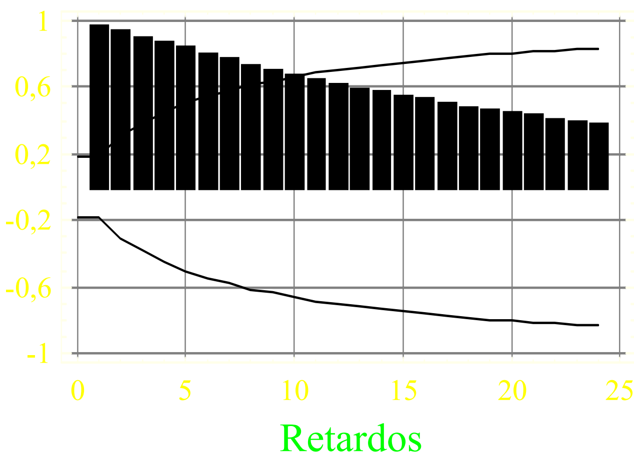


Figura 13: Acciones de IBM en diferencias.

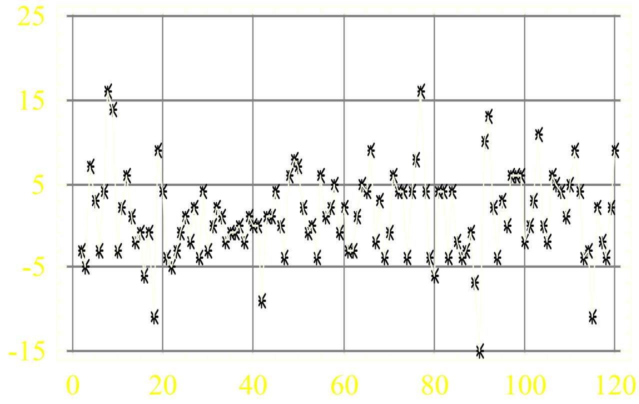
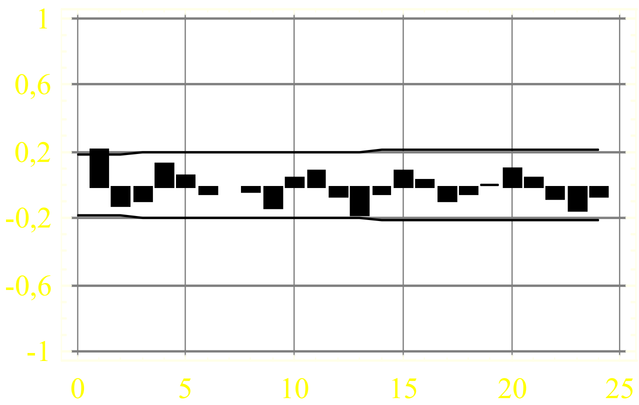


Figura 14: FAS para IBM en diferencias



5. Procesos autorregresivos. AR.

Los procesos autorregresivos forman una familia de procesos tales que una observación depende de las observaciones anteriores. Se denominan procesos AR y se caracterizan por su orden.

5.1 Proceso autorregresivo de primer orden. AR(1).

El proceso autorregresivo de primer orden es el más sencillo de la familia de procesos autorregresivos. Se dice que un proceso o serie es autorregresivo de primer orden si sigue la siguiente ecuación:

$$z_t = \phi z_{t-1} + a_t$$

Si una serie sigue un proceso AR(1), cada observación se construye a partir de la anterior, más una perturbación aleatoria a_t. Por ejemplo, si $\phi = 0.6$ y $z_7 = 50,$ podemos decir $z_8 = 0.6\ 50 + a_t = 30 + a_t$.

El papel de a_t es precisamente permitir que z_8 valga ”algo” en torno a 30, y no exactamente 30. Si no existiera el término a_t, la serie sería determinista, y conocido un valor de z ya conoceríamos todos los demás. Como en la práctica ésto no ocurre, es necesario introducir un término aleatorio que permita la introducción de errores.

El proceso AR(1) para ser estacionario debe tener un valor absoluto de $\phi < 1$. Si el valor de ϕ fuera mayor que 1, el proceso se hace explosivo y rápidamente tiende a crecer desmesuradamente.

5.1.1 FAS y FAP de los procesos AR(1)

La FAS y la FAP del proceso AR(1) tienen un aspecto muy característico. Se puede

demostrar que la FAS del proceso AR(1) tiene la expresión:

$$\rho_k = \phi^k$$

es decir que el palo k, de la FAS es igual al coeficiente ϕ elevado a la potencia k. Esto implica que la FAS de un proceso AR(1) pueda tener el siguiente aspecto:

- ϕ positivo: La FAS será una función positiva y decreciente.
- ϕ negativo: La FAS será una función alternada, y tendrá palos pares positivos, y palos impares negativos.

En cuanto a la FAP, como indica la expresión (ref: ar1), sólo existe influencia de primer orden, ya que si z_t depende de z_{t-2}, es a través de z_{t-1}. Las FAP será por tanto:

- ϕ positivo: La FAP tendrá un único palo, el primero. Este palo será positivo.
- ϕ negativo: La FAP tendrá un único palo y será negativo.

En la figura 15 pueden apreciarse las FAP de procesos AR(1) con $\phi = 0.8$ y $\phi = -0.8$.

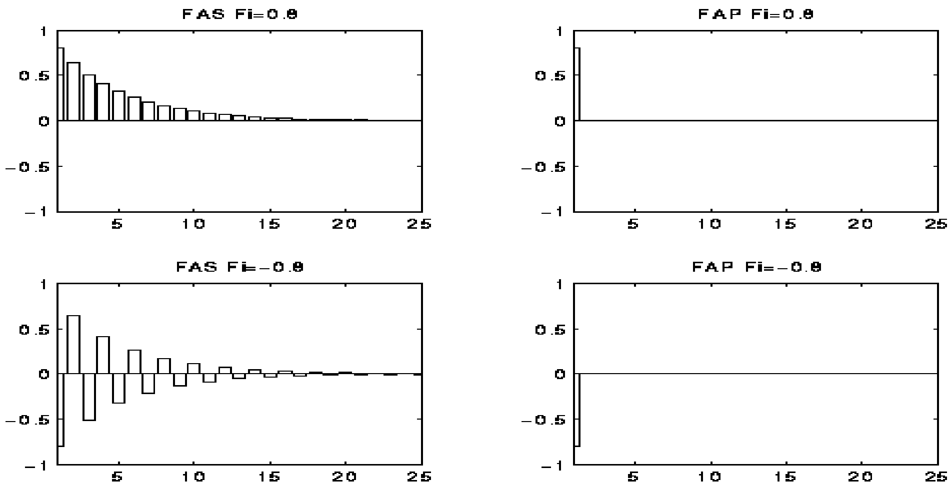


Figura 15: FAS y FAP de AR(1)

La figura 16 presenta las FAS y FAP de procesos AR(1) con $\phi = 0.6$ y $\phi = -0.6$.

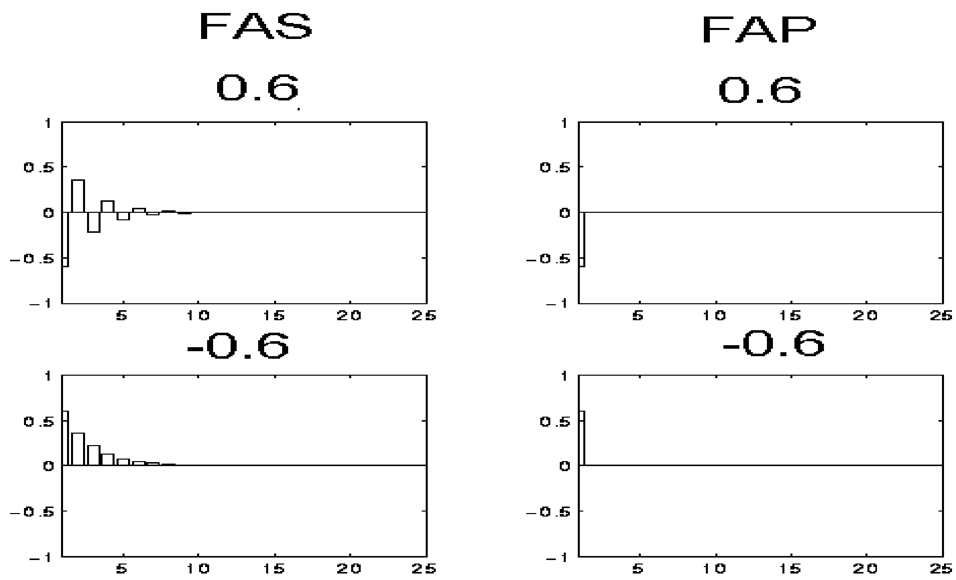


Figura 16: FAS y FAP de AR(1)

Una serie podrá ser un proceso AR(1) si su FAS y su FAP muestral se parecen a las de un proceso AR(1). Si ésto ocurre, será interesante ajustarle a la serie un proceso AR(1).

Ejemplo:

Las figuras 17a, 17b y 17c muestran el gráfico de la serie z_t. Puede observarse que la serie es estacionaria, pues no tiene ni tendencia, ni heterocedasticidad. La FAS y la FAP lo confirman, ya que la FAS tiende rápidamente a cero.

La FAS corresponde a la de un proceso AR(1) con ϕ positivo, pues se produce un decrecimiento exponencial del tamaño de los palos. Si fuera ciertamente un AR(1), la FAP debería tener un único palo positivo. Efectivamente, la FAP de la serie tiene un solo palo positivo.

Figura 17a

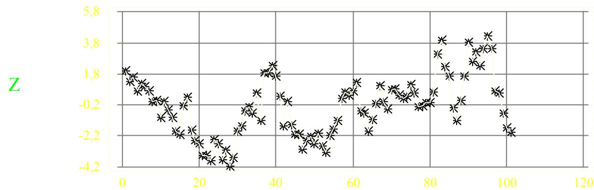


Figura 17B: FAS de Z

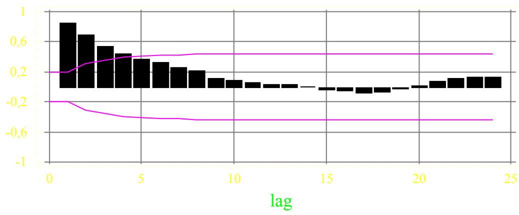
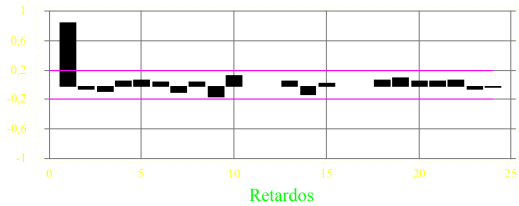


Figura 17c: FAP de Z



5.2 Proceso autorregresivo de segundo orden. AR(2).

El proceso autorregresivo de segundo orden tiene por ecuación

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

Se puede demostrar que el proceso AR(2) será estacionario siempre que las raíces del denominado polinomio característico, que se define como:

$$1 - \phi_1 x - \phi_2 x^2 = 0$$

estén fuera del círculo unidad. Como es bien sabido las raíces de una ecuación de segundo grado, pueden ser raíces reales o complejas conjugadas. Para que el proceso sea estable, si las raíces son reales deben ser mayores que la unidad y si son imaginarias tener módulo mayor que uno.

La FAS de un proceso AR(2) es algo más compleja que la de un AR(1), y admite las siguientes posibilidades:

- Dos raíces reales positivas. En este caso la FAS es la superposición de dos exponenciales decrecientes como las mostradas en la figura 15 y 16 para el caso de raíces positivas. La figura 18 muestra dos posibles casos de Fas con raíces positivas.

- Raíces reales negativas. La FAS será la superposición de dos exponenciales decrecientes con los palos de signos alternados como las que se mostraron en las figuras 15 y 16. La figura 19 muestra dos posibles casos de Fas con raíces negativas.

- Raíces imaginarias. En este caso la FAS y la serie tienen un aspecto cíclico.. La figura 20 muestra diversas posibilidades de FAS de procesos AR(2) ciclicos.

La FAP del AR(2), tendrá únicamente dos palos. Las figuras 18, 19 y 20 muestran las FAP de los procesos AR(2) con diversos tipos de raíces.

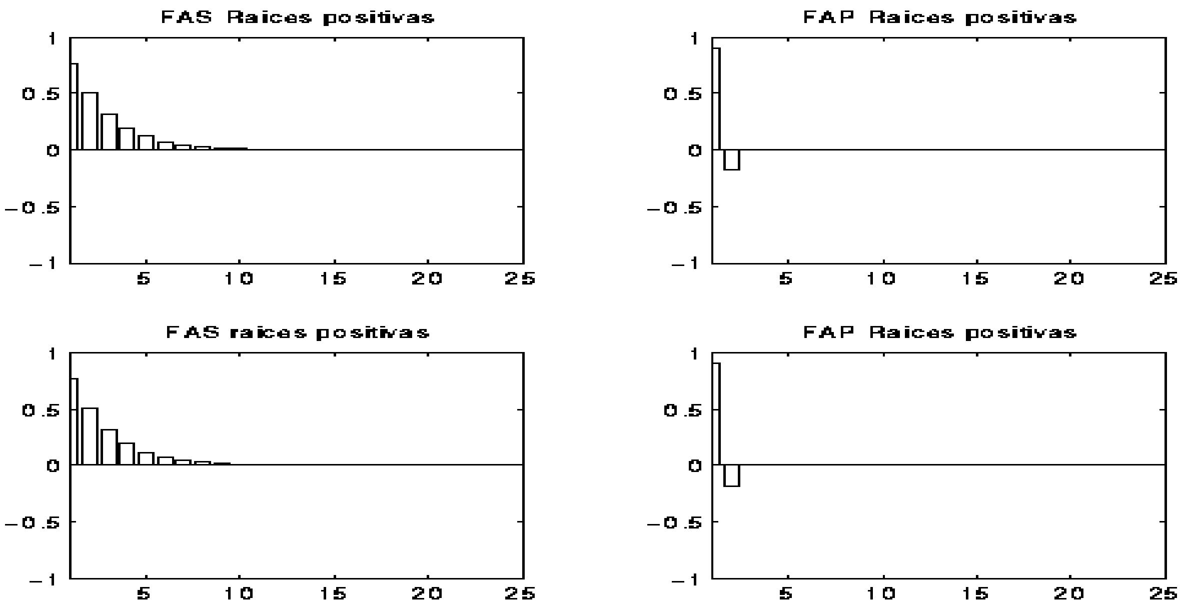


Figura 18: FAS y FAP de AR(2). Raíces positivas.

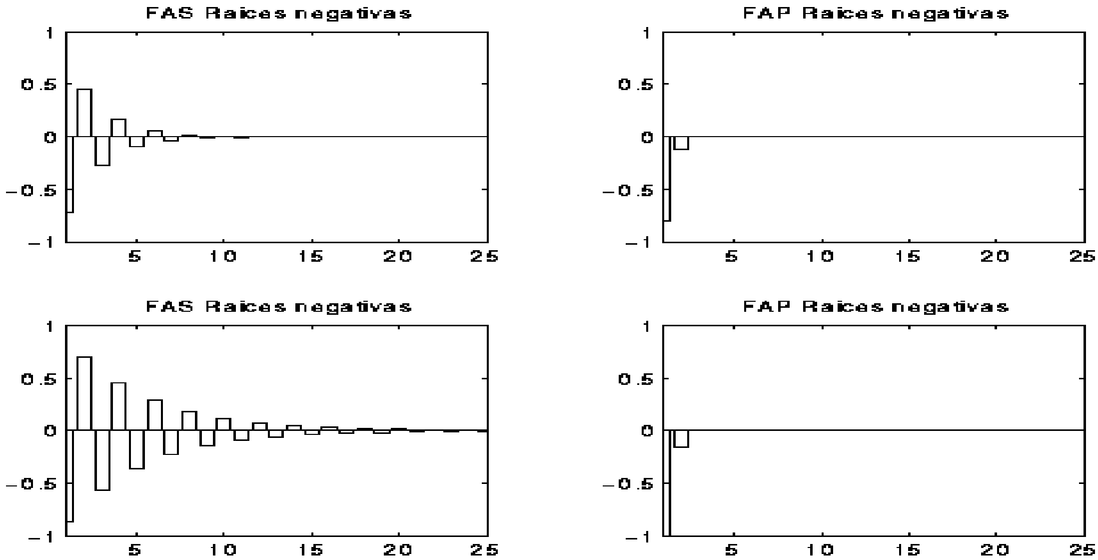


Figura 19: FAS y FAP de AR(2)
Raices negativas.

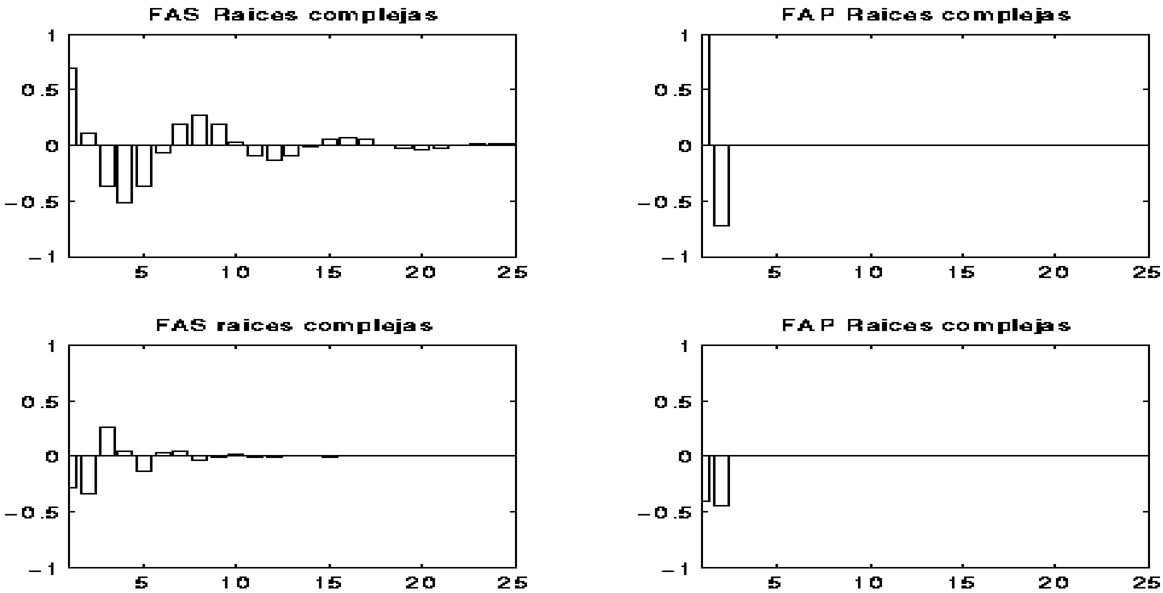


Figura 20: FAS y FAP de AR(2)
Raices complejas.
Estructura cíclica.

5.3 Proceso autorregresivo de orden superior. AR(p).

El proceso autorregresivo de orden superior, AR(p) sigue la ecuación:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + a_t$$

es decir que un observación estará influida por las *p* observaciones anteriores de forma directa. La FAS y FAP son combinación de las del proceso AR(1). Así, cada raíz positiva del polinomio característico

$$x^p - \phi_1 x^{p-1} - \cdots - \phi_p x^p = 0$$

contribuirá con una exponencial decreciente. Cada raíz negativa contribuirá con una exponencial decreciente con palos positivos y negativos alternados.

Finalmente cada pareja de raíces complejas conjugadas aporta una componente cíclica que aparecerá también en la serie.

La FAP del proceso AR(p) será más sencilla, pues tendrá *p* palos significativos y el resto de los palos no lo serán.

6. Procesos de Media Móvil. MA.

Los procesos autorregresivos que se han introducido tienen una característica común: todos ellos tienen memoria larga. La memoria larga aparece en que los palos de la FAS van decreciendo lentamente. Así, un proceso autorregresivo tarda bastante tiempo en absorber los impactos externos. Por ejemplo, si una serie es autorregresiva y se produce una perturbación externa (Por ejemplo, tráfico aéreo y comienza la guerra del golfo) la serie tarda mucho tiempo en absorber el impacto, y el tráfico aéreo se resiente durante muchos periodos debido al efecto de la guerra del golfo.

La realidad nos muestra que existen series, como la de tráfico aéreo precisamente, que absorben rápidamente los impactos. Este tipo de fenómenos no se pueden modelizar exclusivamente mediante modelo autorregresivos, y es necesario introducir una nueva familia de modelos para poder representarlos.

La familia de procesos matemáticos que representa los procesos de memoria corta se denomina genéricamente procesos de media móvil.

6.1 Procesos de Media Móvil de primer orden. MA(1)

El proceso de media móvil de primer orden tiene una ecuación

$$z_t = a_t - \theta a_{t-1}$$

donde a_t se denomina la innovación y representa los efectos externos a la serie. Si un proceso es MA(1), sus valores estarán formados por los efectos externos muy recientes, el actual a_t y el inmediatamente anterior a_{t-1}. El proceso MA tiene memoria corta, y absorbe rápidamente los impactos. La figura 21 muestra una serie AR(1) con $\rho = 0.9$ y un proceso MA(1) con $\theta = 0.9$ a la que en el periodo 50 le viene un impacto externo muy grande.

El impacto puede verse en el salto que da la serie. Sin embargo la serie AR, tarda en volver a su nivel, mientras que la serie MA vuelve inmediatamente, en dos observaciones.

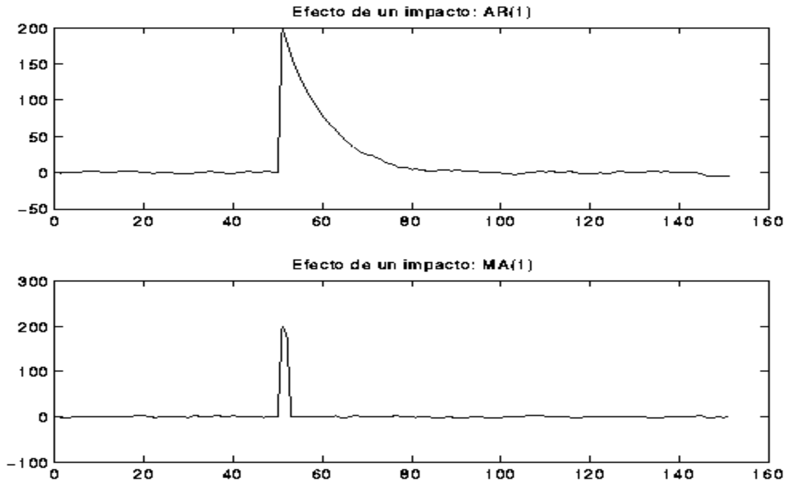


Figura 21: Efecto de impactos en modelos AR y MA

Se puede demostrar que las FAS y FAP de los procesos MA son muy similares a las del AR(1), pero cambiadas. Es decir la FAS del MA(1) tiene un único palo (De ahí la corta memoria). Mientras que la FAP tiene muchos palos que van decreciendo lentamente del mismo modo que lo hacían los palos de la FAP del AR(1). La figura 22 muestra la FAS y FAP de un proceso MA(1)

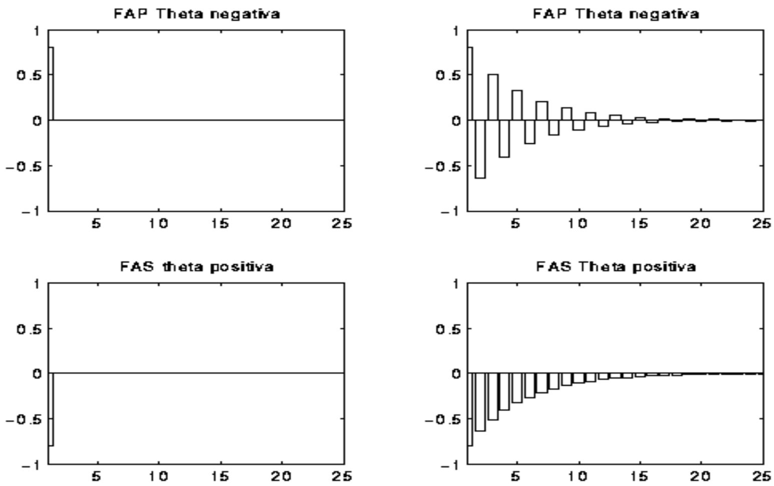


Figura 22: FAS y FAP de MA(1)

6.2 Procesos de Media Móvil de orden superior. MA(q)

El proceso de media móvil de orden superior MA(q) tiene la siguiente ecuación:

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

y tarda q periodos en absorber los impactos.

Su FAS tendrá q palos significativos, y su FAP tendrá un decrecimiento.

7. Procesos ARMA(p,q)

En la práctica se observa que las series no son puras AR o MA, sino que presentan parte AR y parte MA.

Los procesos ARMA(p,q) son combinación de estructuras autorregresivas y de media móvil que tienen una parte AR(p) y una parte MA(q). Su FAS y su FAP serán combinación de ambos procesos y tendrán la siguiente estructura:

- **FAS**: Los primeros *q* palos de la FAS vendrán establecidos por la parte MA. Estos palos le van a permitir a la serie absorber rápidamente los impactos externos, tal como corresponde a la parte MA. A partir del retardo *q* se producirá un decrecimiento de los palos que vendrá dado por la estructura AR.

- **FAP**: Los primeros *p* palos de la FAP vendrán establecidos por la parte AR. A partir del retardo *p* se producirá un decrecimiento de los palos que vendrá dado por la estructura MA.

La figura 23 muestra la FAS y FAP de dos series ARMA(1,1)

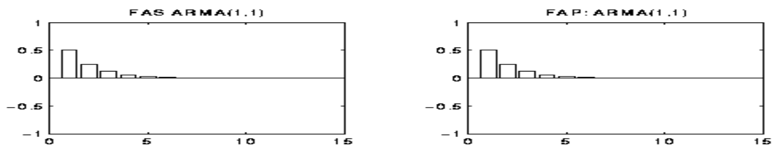
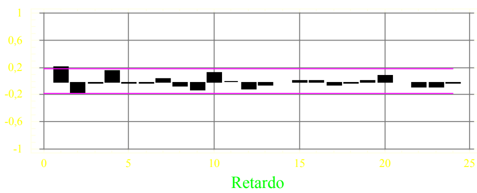


Figura 23: FAS y FAP de ARMA(1,1)

8. Series No Estacionarias. Procesos ARIMA(p,d,q)

Como se ha explicado en la sección 4, los modelos ARMA únicamente sirven para ajustar series estacionarias. Si la serie no lo es habrá que transformarla como ya se ha indicado. Si una serie como la de IBM de la figura 11 tiene tendencia, se toma una diferencia y se obtiene la serie de la figura 13 que ya no tiene tendencia. Su FAP se muestra a continuación, en la figura 24. Entre la FAS de la figura 14 y la FAP de la 24 se observa que el modelo de esta serie ∇IBM puede ser un MA(1), ya que la FAS tiene un palo y en la FAP parece haber más estructura. La serie ∇IBM sigue un modelo ARMA(0,1)=MA(1).

Figura 24: FAP de IBM en diferencias



El modelo que sigue IBM sin diferenciar se denomina modelo ARIMA(p,d,q), donde:

- p representa el orden de la parte autorregresiva de la serie estacionaria. En el caso de IBM el orden será 0.

- q representa el orden de la media móvil de la serie estacionaria. En el caso de IBM será 1.

- d representa el número de diferencias que ha habido que tomar para que la serie que, inicialmente no era estacionaria, sea finalmente estacionaria. En el caso de IBM será 1.

Podemos concluir por tanto que la representación general de una serie será el modelo ARIMA(p,d,q). Así, si una serie requiere 2 diferencias para ser estacionaria en tendencia, y hay que transformarla a Logaritmos para estabilizar la varianza, y se le identifica un modelo AR(2) una vez que se ha vuelto estacional, decimos que la serie $\log(z_t)$ sigue un modelo ARIMA(2,2,0). En el caso de IBM, utilizando esta nomenclatura, el modelo será ARIMA(0,1,1).

9.Operadores de retardo y diferencias:

Habitualmente en series temporales se utiliza una nomenclatura específica: Los operadores de retardos (B) y de diferencias (∇).

En esta sección se va a introducir la nomenclatura y a proporcionar algunos ejemplos.

9.1 Operador de retardo:

El operador de retardo, B, retrasa una posición la observación. Así por ejemplo

$$Bz_t = z_{t-1}$$

Es decir B aplicada a z_t retrasa ésta y la convierte en z_{t-1}.

Aplicando el operador dos veces obtenemos:

$$B^2 z_t = B(Bz_t) = Bz_{t-1} = z_{t-2}$$

$$B^2 z_t = z_{t-2}$$

Y aplicándolo k veces, obtenemos

$$B^k z_t = z_{t-k}$$

Veamos ahora cómo se escriben los modelos en forma de operador de retardos:

1. **Modelo AR(1)**:

$$z_t = \phi_1 z_{t-1} + a_t$$

Aplicando el Operador de Retardo, B, obtenemos:

$$z_t = \phi_1 B z_t + a_t$$

y pasando los miembros que contienen z_t al lado izquierdo de la ecuación,

$$z_t - \phi_1 B z_t = a_t$$

escribiendo el resultado en forma de polinomio,

$$(1 - \phi_1 B) z_t = a_t$$

que es la manera alternativa de escribir el modelo autorregresivo de primer orden en forma de polinomio.

2. **Modelo AR(2)**

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

Aplicando el Operador de Retardo, B, obtenemos:

$$z_t = \phi_1 B z_t + \phi_2 B^2 z_t + a_t$$

y pasando los miembros que contienen z_t al lado izquierdo de la ecuación,

$$z_t - \phi_1 B z_t - \phi_2 B^2 z_t = a_t$$

escribiendo el resultado en forma de polinomio,

$$(1 - \phi_1 B - \phi_2 B^2) z_t = a_t$$

que es la manera alternativa de escribir el modelo autorregresivo de primer orden en forma de polinomio. Nótese que para que el proceso AR(2) sea estacionario, las dos raices del polinomio en B, $(1 - \phi_1 B \phi_2 - B^2)$ deben estar fuera del círculo unidad.

3. **Modelo AR(p)**

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + a_t$$

Aplicando el Operador de Retardo, B, obtenemos:

$$z_t = \phi_1 B z_t + \phi_2 B^2 z_t + \cdots + \phi_p B^p z_t + a_t$$

y pasando los miembros que contienen z_t al lado izquierdo de la ecuación,

$$z_t - \phi_1 B z_t - \phi_2 B^2 z_t - \cdots - \phi_p B^p z_t = a_t$$

escribiendo el resultado en forma de polinomio,

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) z_t = a_t$$

que es la manera alternativa de escribir el modelo autorregresivo de orden p en forma de polinomio. Nótese que para que el proceso AR(p) sea estacionario, las *p* raices del polinomio en B, $(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$ deben estar fuera del círculo unidad.

4. **Modelo MA(1)**

$$z_t = a_t - \theta_1 a_{t-1}$$

Aplicando el Operador de Retardo, B, obtenemos:

$$z_t = a_t - \theta_1 B a_t$$

escribiendo el resultado en forma de polinomio,

$$z_t = (1 - \theta_1 B) a_t$$

que es la manera alternativa de escribir el modelo de media móvil de primer orden en forma de polinomio.

5. **Modelo MA(2)**

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

Aplicando el Operador de Retardo, B, obtenemos:

$$z_t = a_t - \theta_1 B a_t - \theta_2 B^2 a_t$$

escribiendo el resultado en forma de polinomio,

$$z_t = (1 - \theta_1 B - \theta_2 B^2) a_t$$

que es la manera alternativa de escribir el modelo de media móvil de segundo orden en forma de polinomio.

6. **Modelo MA(q)**

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

Aplicando el Operador de Retardo, B, obtenemos:

$$z_t = a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \cdots - \theta_q B^q a_t$$

escribiendo el resultado en forma de polinomio,

$$z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t$$

que es la manera alternativa de escribir el modelo de media móvil de orden q en forma de polinomio.

7. **Modelo ARMA(1,1)**:

El modelo ARMA(1,1) contiene una parte AR(1) y un parte MA(1), se escribe muy sencillamente en forma de polinomio, sin más que poner los polinomios correspondientes a un AR(1) en el lado de la z_t y el polinomio correspondiente a un MA(1) en lado del a_t.

$$(1 - \phi_1 B) z_t = (1 - \theta_1 B) a_t$$

que es la manera alternativa de escribir el modelo de ARMA(1,1) en forma de polinomio.

8. **Modelo general ARMA(p,q)**.

El modelo ARMA(p,q) contiene una parte AR(p) y un parte MA(q), se escribe muy sencillamente en forma de polinomio, sin más que poner los polinomios correspondientes a un AR(p) en el lado de la z_t y el polinomio correspondiente a un MA(q) en lado del a_t.

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t$$

que es la manera alternativa de escribir el modelo de ARMA(1,1) en forma de polinomio.

9.2 Operador diferencias

Cuando una serie tiene tendencia es preciso quitarla mediante una diferencia regular. Una diferencia regular le resta a cada elemento de la serie la observación anterior:

$$w_t = z_t - z_{t-1}$$

La expresión puede escribirse :

$$w_t = z_t - z_{t-1} = z_t - Bz_t = (1 - B) z_t$$

es decir que tomar una diferencia equivale a premultiplicar z_t por $(1 - B)$. A esta expresión se le denomina

$$z_t - z_{t-1} = \nabla z_t$$

Cuando escribimos ∇z_t se indica que se ha tomado una diferencia d a la serie. Si se toman dos diferencias de orden 1, se representa por $\nabla^2 z_t$. Y eso ocurre cuando la serie sigue teniendo tendencia después de tomar la primera diferencia.

Veamos cómo se escriben los diferentes modelos ARIMA(p,d,q):

1. **AR(1) con una diferencia**. **ARIMA(1,1,0)**: El modelo AR(1) tiene una expresión, en forma polinómica

$$(1 - \phi_1 B) z_t = a_t$$

Si hay que tomar una diferencia, el modelo será el mismo pero aplicado a ∇z_t.

$$(1 - \phi_1 B) \nabla z_t = a_t$$

que equivale a

$$(1 - \phi_1 B)(1 - B) z_t = a_t$$

y multiplicando los términos de los dos paréntesis

$$(1 - (1 + \phi_1) B - \phi_1 B^2) z_t = a_t$$

y quitando el operador

$$z_t = (1 + \phi_1) z_{t-1} + \phi_1 z_{t-2} + a_t$$

2. **MA(1) con una diferencia**. **ARIMA(0,1,1)**. El modelo MA(1) tiene una expresión, en forma polinómica

$$z_t = (1 - \theta_1 B)a_t$$

Si z precisa de una diferencia, en lugar de z se operará con ∇z

$$\nabla z_t = (1 - \theta_1 B)a_t$$

y ésto equivale a

$$(1 - B)z_t = (1 - \theta_1 B)a_t$$

$$z_t = z_{t-1} + a_t + \theta_1 a_{t-1}$$

3. **AR(1) con dos diferencias**. **ARIMA(1,2,0)**. El modelo AR(1) tiene una expresión, en forma polinómica

$$(1 - \phi_1 B)z_t = a_t$$

Si hay que tomar una diferencia, el modelo será el mismo pero aplicado a $\nabla^2 z_t$.

$$(1 - \phi_1 B)\nabla^2 z_t = a_t$$

que equivale a

$$(1 - \phi_1 B)(1 - B)^2 z_t = a_t$$

y multiplicando al cuadrado el segundo paréntesis

$$(1 - \phi_1 B)(1 + B^2 - 2B)z_t = a_t$$

$$(1 - \phi_1 B + B^2 - \phi_1 B^3 - 2B + 2\phi_1 B^2)z_t = a_t$$

$$(1 - (\phi_1 + 2) + (1 + 2\phi_1)B^2 - \phi_1 B^3)z_t = a_t$$

$$z_t = (\phi_1 + 2)z_{t-1} - (1 + 2\phi_1)z_{t-2} + \phi_1 z_{t-3} + a_t$$

10. Estimación de modelos ARIMA.

Una vez identificada la serie, es decir cuando se conoce el modelo ARIMA(p,d,q) que puede seguir la serie, es preciso estimar los parámetros.

La estimación de modelos ARIMA es compleja y utiliza algoritmos de optimización no lineales para calcular los valores de los parámetros utilizando el método de máxima verosimilitud.

El resultado de la estimación es una tabla análoga a la de regresión. Esta tabla proporciona los valores estimados de los parámetros y sus errores estándar y estadísticos *t*.

Vamos a ajustar un modelo a la serie de IBM. Como se ha visto en las figuras 11, 12, 13 y 14 la serie IBM no es estacionaria debido a su tendencia. Tomamos por tanto una diferencia regular ∇IBM. Las figuras 14 y 24 presentan la FAS y FAP de ∇IBM. En ellas puede observarse: La FAS presenta un solo palo positivo significativo. La FAP presenta dos palos. El primero positivo y el segundo negativo. La estructura podría corresponder a la de un modelo MA(1) con $\theta < 0$. Ajustando un MA(1) se obtiene:

$$
\begin{array}{llll}
\nabla IBM_t = & a_t - & 0.27a_{t-1} \\
\text{Std. Error} & & (0.09) \\
\text{t} & & (-2.97)
\end{array}
$$

El ajuste indica que θ es significativamente distinta de 0, ya que la t es mayor que 2 en valor absoluto.

Si hubiésemos estimado un MA(2), la ecuación obtenida habría sido:

$$
\begin{array}{llll}
\nabla IBM_t = & a_t - & 0.23a_{t-1} + & 0.07a_{t-2} \\
\text{Std. Error} & & (0.09) & (0.09) \\
\text{t} & & (-2.46) & (0.75)
\end{array}
$$

que indica que θ_2 no es significativa.

11. Diagnosis.

El ajuste de un modelo ARIMA(p,d,q) implica cumplir satisfactoriamente la etapa de identificación del modelo. Si una serie está bien identificada, y se le ajusta el modelo correcto, los residuos de la serie deben carecer completamente de estructura. Como se ha visto anteriormente, una serie sin estructura de dependencia es precisamente el ruido blanco.

El ruido blanco es una serie estacionaria producida por observaciones independientes. Su FAS y su FAP teóricas son cero, es decir ningún palo debe ser significativo.

La diagnosis de modelos ARIMA se basa precisamente en este hecho: Si la serie está bien ajustada, sus residuos deben ser ruido blanco. Así, tras ajustar un modelo y comprobar la significatividad de los coeficientes realizaremos la diagnosis del mismo basada en:

1. FAS y FAP de los residuos. Si la serie está bien ajustada, la FAS y la FAP de los residuos debe ser nula. En la práctica debemos comprobar que los palos de las funciones no son significativos.

2. Test de no significatividad de los palos de la FAS y la FAP. El test de Box-Pierce proporciona información sobre si los primeros palos de la función de autocorrelación simple de los residuos son cero o no.

3. Test de rachas y estructuras extrañas en los residuos.

Las figuras 25 y 26 muestran la FAS y la FAP de los residuos del modelo ARIMA(0,1,1) para IBM. En ellas puede apreciarse que los palos no son significativos, por lo que podemos aceptar que los residuos son ruido blanco.

Figura 25. FAS de los residuos de ARIMA(0,1,1) para IBM

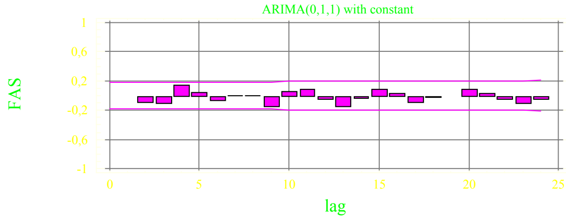
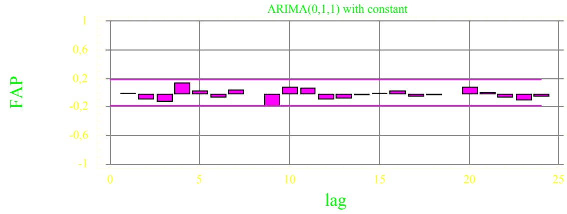


Figura 26: FAP de los residuos de ARIMA(0,1,1) para IBM



El test de Box-Pierce aplicado a los 24 primeros palos de la FAS de los residuos da un valor de 18.63 con un p-valor de 0.72. Este resultado debe interpretarse como que no hay evidencia para rechazar que los residuos sean ruido blanco. El test de Box-Pierce indica problemas cuando el p-valor se hace menor de 0.05. Cuanto mayor sea, hay más evidencia a favor de que los residuos son ruido blanco.

Finalmente los test de rachas no indican nada anormal en los residuos. Podemos por tanto asumir que el modelo ARIMA(0,1,1) es adecuado para modelizar los datos de IBM.

12. Modelos Estacionales

Muchas series con periodicidad menor que la anual tienen estacionalidad. La estacionalidad aparece en forma de ciclo, de modo que la serie tiene una pauta que se repite año tras año. La serie de temperaturas en Madrid de la figura 5 presenta un clara estacionalidad.

El orden de la estacionalidad indica cada cuantos periodos se repite el ciclo. Así, en datos mensuales es normal encontrar estacionalidad de orden 12. En datos trimestrales la estacionalidad será de orden 4, en datos cuatrimestrales la estacionalidad será de orden 3 y así sucesivamente.

Las series estacionales tienen además de la estructura de dependencia de las series no estacionales, es decir

$$z_1 \longrightarrow z_2 \longrightarrow \cdots \longrightarrow z_{t-1} \longrightarrow z_t \longrightarrow z_{t+1} \longrightarrow \cdots \tag{1}$$

una estructura adicional. Si por ejemplo la serie es mensual, es decir que tiene estacionalidad de orden 12, existirá una estructura de dependencia estacional en la que un mes de marzo influye sobre el siguiente mes de marzo *directamente*.

$$z_1 \rightarrow z_{13} \rightarrow \cdots \rightarrow z_{t-12} \rightarrow z_t \rightarrow z_{t+12} \rightarrow \cdots \quad (2)$$

Se distinguen entonces dos tipos de estructuras:

1. Estructura Regular: se refiere a la relación establecida en (1). Corresponde a los modelos ya estudiados. Ajustaremos la parte regular mediante un modelo ARIMA(p,d,q)

2. Estructura estacional: se refiere a la relación (2). Ajustaremos modelos ARIMA estacionales.

12.1 Modelos ARIMA estacionales.

Una serie estacional puede tener una relación autorregresiva o de media móvil. Un serie estacional autorregresiva de primer orden tendrá una FAS y una FAP como las de las figuras 27. Es decir es la misma estructura que un modelo AR(1) pero los palos están separados por 12 retardos.

Un modelo MA(1) presentará una FAS como las de las figuras 27. Las FAP de estos modelos se presenta en la figura 28.

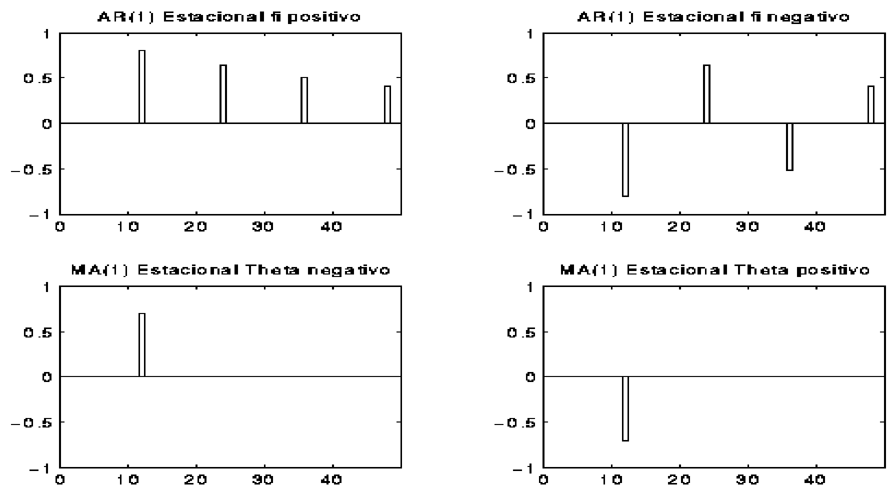


Figura 27: FAS de Modelos Estacionales

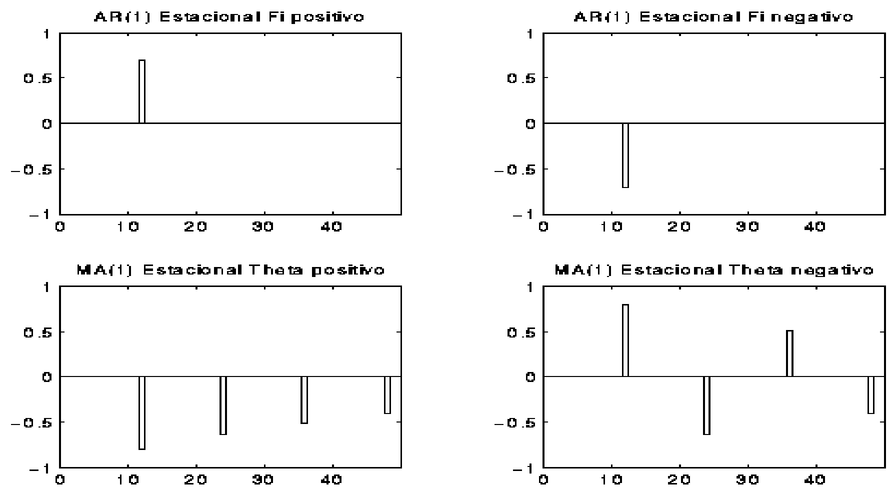


Figura 28: FAP de Modelos estacionales

12.2 Interacción entre la parte regular y la estacional.

Los modelos estacionales pueden ser como en los casos anteriores $AR(1)_{12}$ o $MA(1)_{12}$, o presentar estructuras más complejas como modelos ARMA. En principio la parte estacional se puede modelizar de la misma forma que la parte regular. La identificación se realiza estudiando la FAS y la FAP en los retardos estacionales. Cuando una serie tiene parte regular y parte estacional, la FAS presenta una interrelación entre ellas. La figura 29 muestra ejemplos de esta interacción.

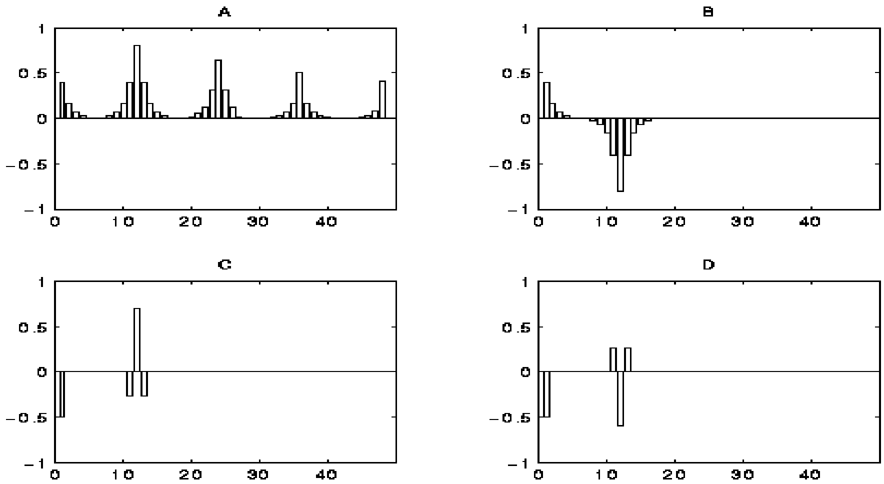


Figura 29: Modelos Estacionales FAS

Figura29: A: Modelo $AR(1)\text{x}AR(1)_{12}$ con ambas $\phi > 0$.

B: Modelo $AR(1)\text{x}MA(1)_{12}$ con $\phi > 0$ y $\Theta > 0$.
C: Modelo $MA(1)\text{x}MA(1)_{12}$ con $\theta > 0$ y $\Theta < 0$.
D: Modelo $MA(1)\text{x}MA(1)_{12}$ con $\theta > 0$ y $\Theta > 0$.

Básicamente las FAS y FAP de modelos estacionales tienen las siguientes características:

- **FAS**:

1. En los retardos bajos se observa únicamente la parte regular.
2. En los retardos estacionales (12, 24,...) se observa básicamente la parte estacional.
3. En los retardos cercanos a los estacionales se observa la interacción entre las dos partes. A ambos lados del retardo estacional aparece la FAS regular.

- **FAP**:

1. En los retardos bajos se observa únicamente la FAP de la parte regular.

2. A la derecha de cada retardo estacional aparece la FAP regular. Si el coeficiente estacional es positivo la FAP aparece invertida. Si es negativo, aparece con su signo.

3. A la izquierda de los coeficientes estacionales aparece la FAP con su signo.

12.3 Series estacionales no estacionarias.

Tal como hemos definido las series estacionarias, en la sección 4, ya se vió que las series estacionales no eran estacionarias. El ajuste de modelos a series estacionales requiere por tanto como paso previo el estacionarizar la serie, y para ello es necesario eliminar el ciclo estacional.

La eliminación del ciclo estacional es inmediata: Basta con tomar diferencias del orden estacional. Así, si la serie z_t es estacional de orden 12, y presenta un ciclo, definiremos la serie

$$w_t = z_t - z_{t-12} \tag{3}$$

que es posible que sea estacionaria. Nótese que (3) puede escribirse utilizando el operador diferencias B.

Las propiedades del operador diferencia son:

- $Bz_t = z_{t-1}$
- $B^2 z_t = B(Bz_t) = Bz_{t-1} = z_{t-2}$
- $B^k z_t = z_{t-k}$

La expresión (3) puede escribirse :

$$w_t = z_t - z_{t-12} = z_t - B^{12} z_t = (1 - B^{12}) z_t = \nabla_{12}\, z_t$$

Cuando escribimos $\nabla_{12}\, z_t$ se indica que se ha tomado una diferencia de orden 12 a la serie. Si se toma una diferencia de orden 1, es decir una diferencia regular,

$$w_t = z_t - z_{t-1} = z_t - Bz_t = (1 - B) z_t = \nabla z_t$$

La expresión ∇z_t indica que se ha tomado una diferencia de orden 1. si se toman dos diferencias de orden 1, se representa por $\nabla^2\, z_t$. Cuando una serie requiere una diferencia estacional y una regular lo representamos mediante $\nabla\,\nabla_{12}\, z_t$.

Una serie no estacionaria en la parte regular produce en la FAS palos que tienden a decrecer muy lentamente. Una serie no estacionaria en la parte estacional produce palos estacionales que no terminan de decrecer.

Ejemplo:

Las figura 30, 31 y 32 presentan la evolución del Índice de Producción Industrial (IPI) en Francia, su FAS y su FAP. Como puede observarse, la serie tiene tendencia y ciclo estacional. Por tanto la serie no es estacionaria.

Puede observarse en la FAS que los retardos estacionales no terminan de descender.

Figura 30: Serie IPI en Francia

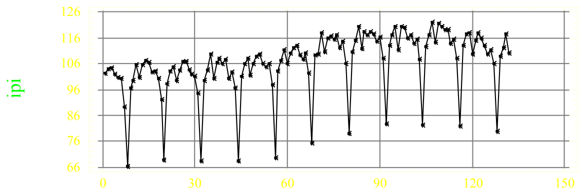


Figura 31: Fas de IPI

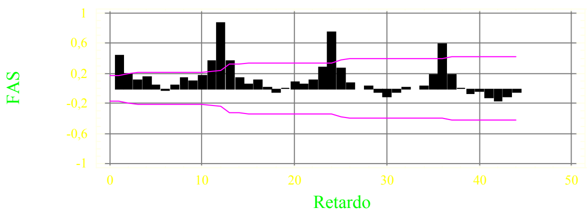
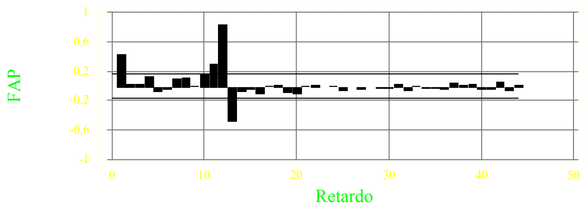


Figura 32: FAP de IPI



Si se toma una diferencia regular para eliminar la tendencia, se obtienen las figuras 33, 34 y 35. En estas figuras se ve que la diferencia regular elimina la tendencia, pero el ciclo estacional permanece en la serie.

Figura 33: Serie IPI. Diferencia regular

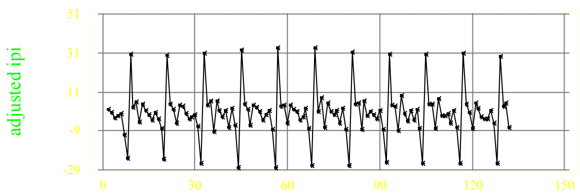


Figura 34: Fas de IPI. Diferencia regular.

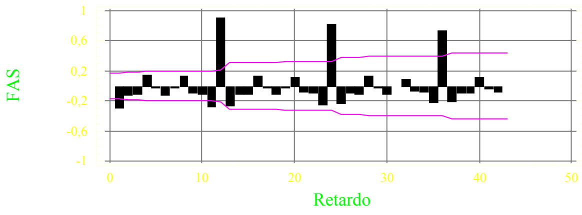
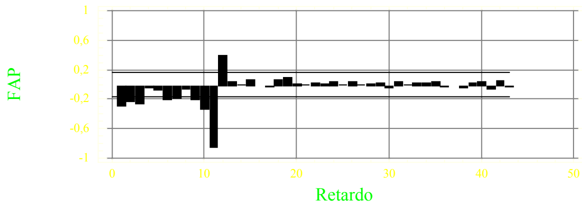


Figura 35: FAP de IPI. Diferencia regular.



En la serie ∇IPI se observa con claridad la falta de estacionariedad de la parte estacional. La FAS no termina de decrecer.

Si hubiéramos tomado una diferencia estacional pero no la regular, es decir $\nabla_{12} IPI$ que quiere decir que a la serie *IPI* se le ha tomado una diferencia de orden 12, se obtienen las figuras 36, 37 y 38. En ellas se aprecia claramente que los palos estacionales ya han decrecido debido a la diferencia de orden 12, pero los regulares denotan falta de estacionalidad.

Figura 36: Serie IPI. Diferencia estacional

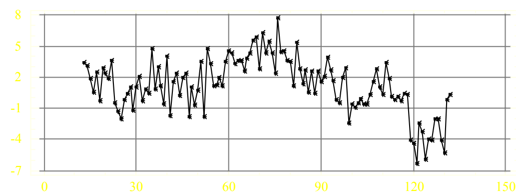


Figura 37: Fas de IPI. Diferencia estacional.

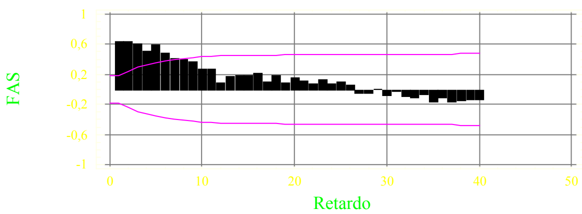
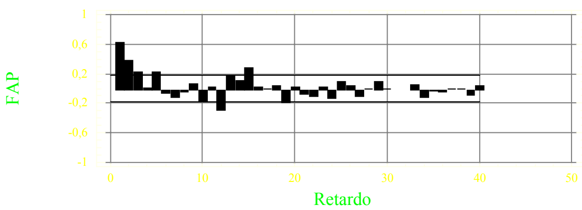
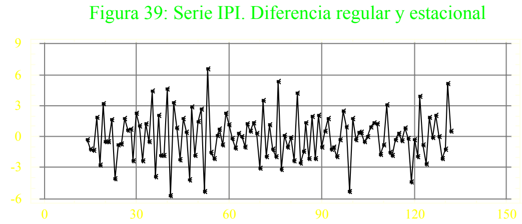


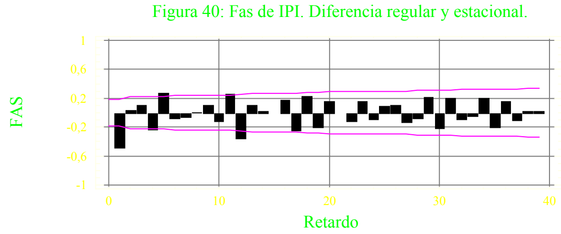
Figura 38: FAP de IPI. Diferencia estacional.

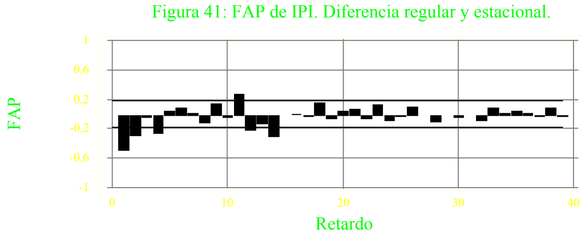


Es necesario por tanto tomar una diferencia estacional y una diferencia regular.

La figura 39 muestra la serie *IPI* con una diferencia estacional y una regular. Este serie se denomina $\nabla\nabla_{12}IPI$. El orden en que se tomen las diferencias es irrelevante. Las figuras 40 y 41 muestran la FAS y la FAP de la serie $\nabla\nabla_{12}IPI$. En ellas se aprecia que la serie ya es estacionaria.







El análisis de la FAS y la FAP de la serie $\nabla\nabla_{12}IPI$ indica:

- **Parte regular**: Los primeros retardos contienen la información de la parte regular. Vemos un palo grande en la FAS y un descenso en la FAP. Podemos pensar en un modelo MA(1).

- **Parte estacional**: Los palos estacionales son claros en el retardo 12 en ambas Funciones de Autocorrelación. No se ven palos de orden superior. Podemos pensar en un modelo $MA(1)_{12}$

Ajustando un modelo $ARIMA(0,1,1)x(0,1,1)_{12}$ se obtiene la siguiente estimación:

$$\hat{\theta} = 0.63 \quad \hat{\Theta} = 0.88$$
$$t = 8.93 \quad t = 25.1$$

que indica que ambos parámetros son significativos. La FAS y FAP de los residuos es adecuada, y el test de Box-Pierce proporciona una valor de 19.62 con un p-valor de 60. Podemos concluir por tanto que el modelo es adecuado.

13.Previsión.

Una vez ajustado el modelo es posible utilizarlo para prever los valores futuros de la serie.

Matemáticamente esto implica obtener la ecuación:

$$z_t = f(z_{t-1}, z_{t-2}, \cdots, a_t, a_{t-1}, \cdots)$$

y substituyendo los valores se obtienen los valores futuros. Los intervalos de predicción para estos valores futuros son relativamente complejos de calcular, y en la práctica los proporcionan los ordenadores.

En la figura 42 se muestran las previsiones para la serie IPI con el modelo ARIMA(0,1,1)x$(0,1,1)_{12}$ que hemos ajustado.

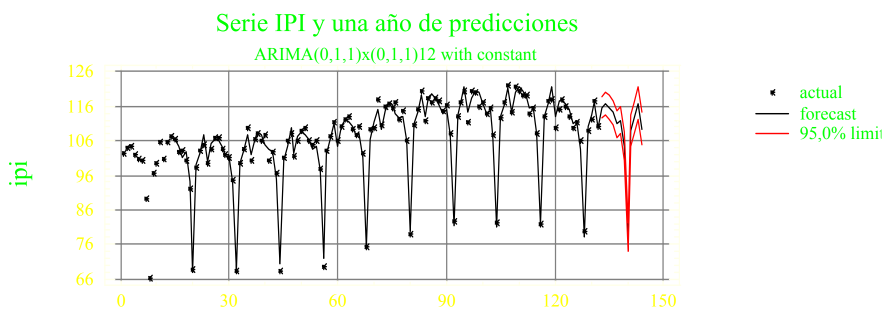


Figura 42: Serie IPI y sus predicciones