Series temporales

1. Introducción.

2. Análisis de la serie.

3. Modelos ARMA y ARIMA.

4. Predicción.

Introducción

- **Serie temporal:** Conjunto de observaciones z_t, cada una recogida en un tiempo específico t.

  1. Tiempo discreto (habitualmente a intervalos de tiempo fijos).
  2. Tiempo continuo.

- **Objetivo:** Explicar la evolución de la serie a lo largo del tiempo y prever sus valores futuros.

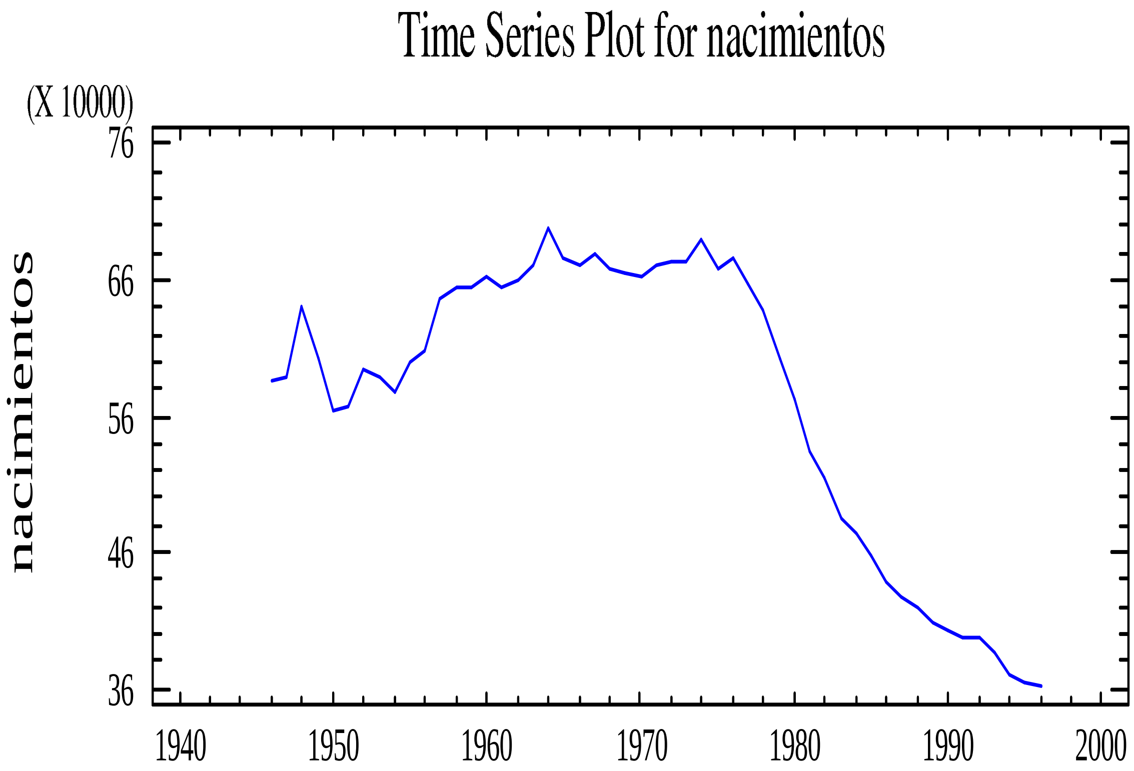
- La representación principal es un gráfico temporal.

Ejemplos:

- Serie de nacimientos.
- Serie de temperaturas.
- Serie de ventas de una empresa.
- Serie de la demanda de energía eléctrica.
- Serie de cotizaciones en bolsa.

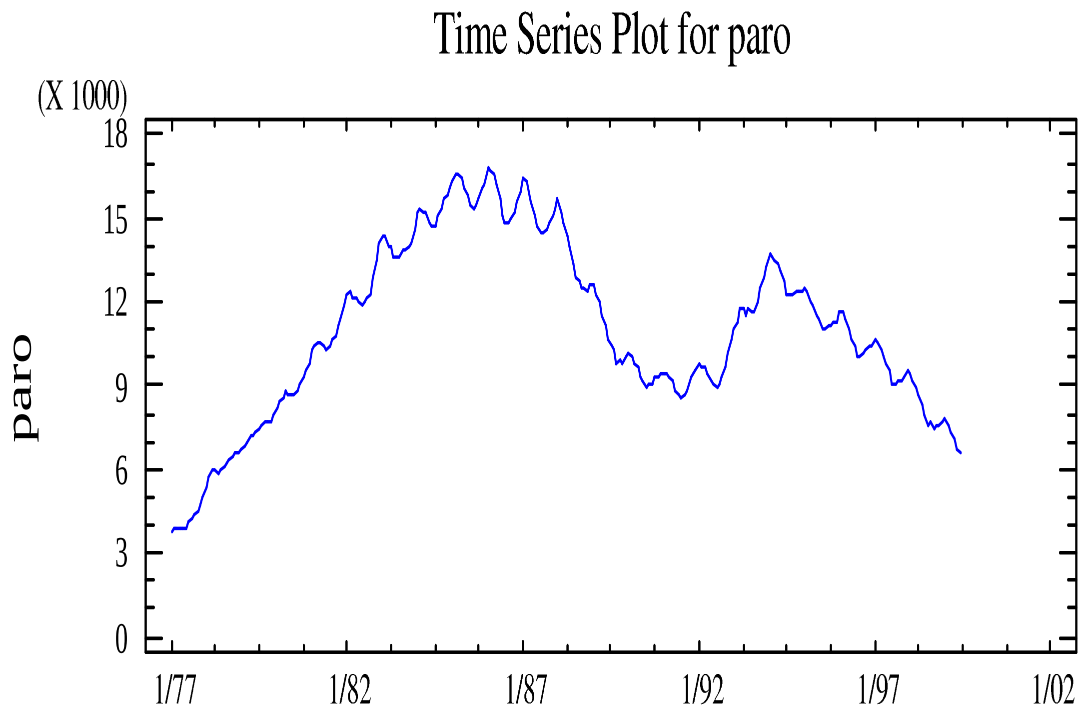
Ejemplos

Número de nacimientos en España desde 1946 a 1996 (serie anual).



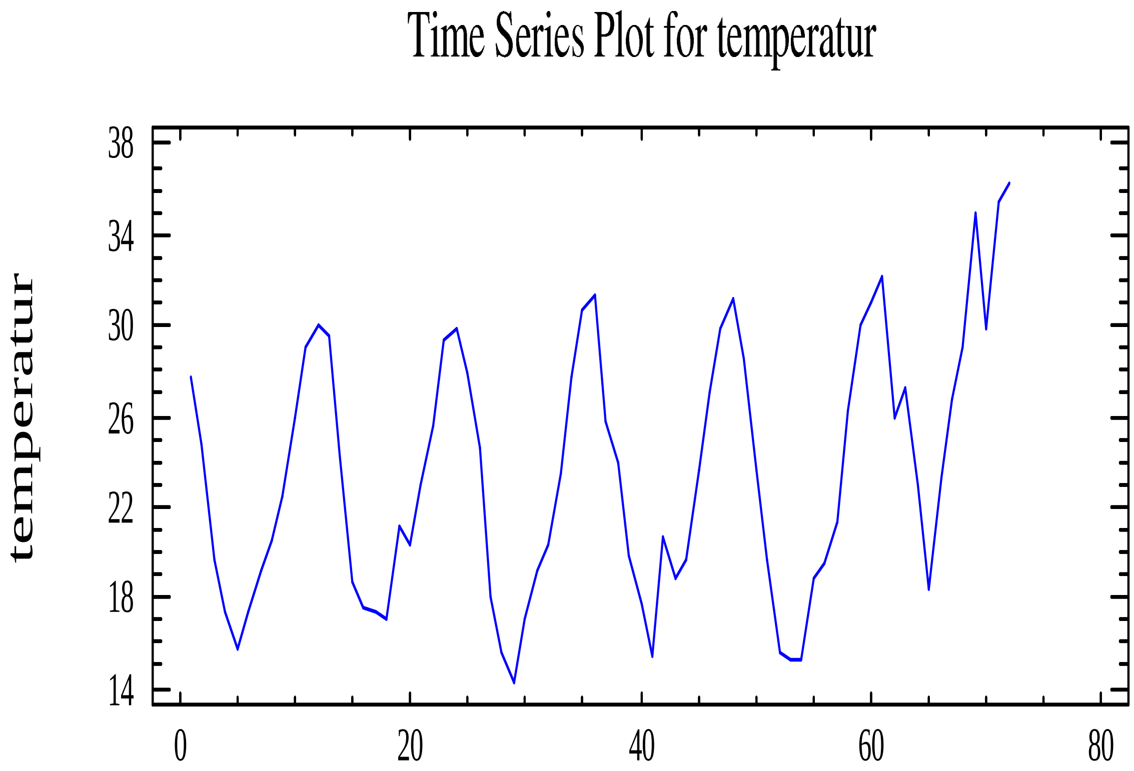
Ejemplos

Número de inscripciones mensuales en el INEM (España) desde enero de 1977 a julio de 1999 (serie mensual).



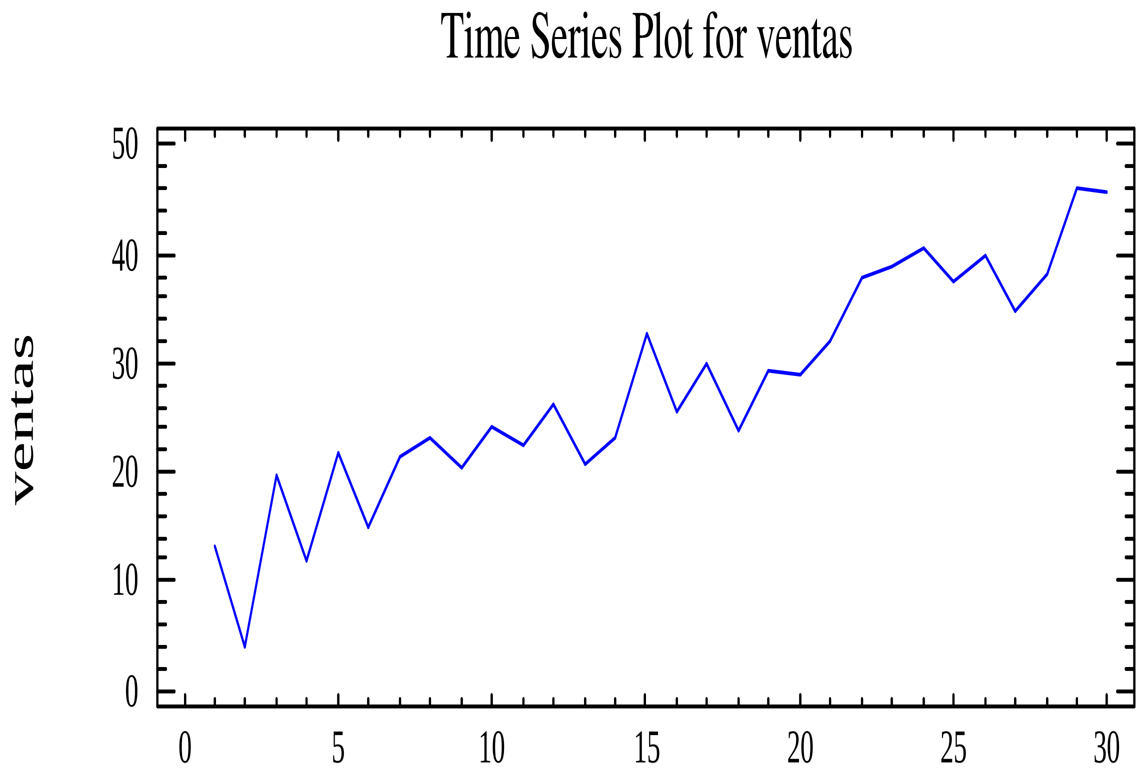
Ejemplos

Temperaturas medias en Valencia (serie mensual).



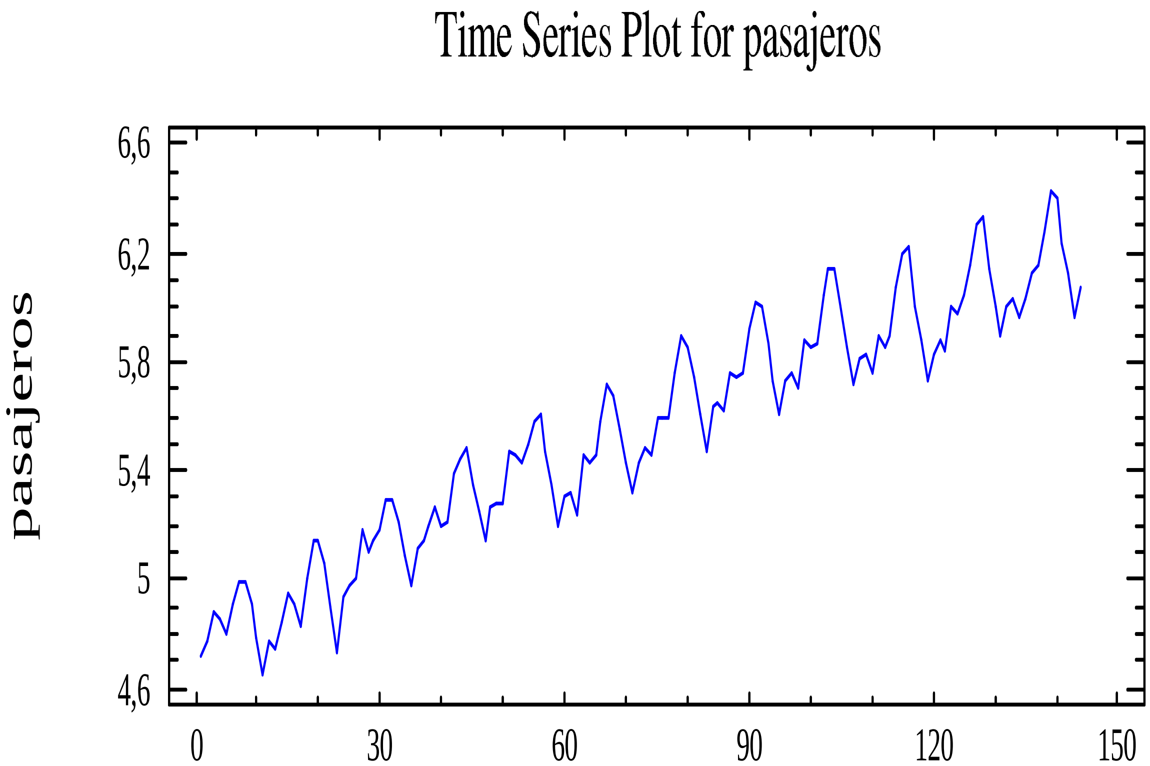
Ejemplos

Ventas de una empresa (beneficios mensuales en MPta).



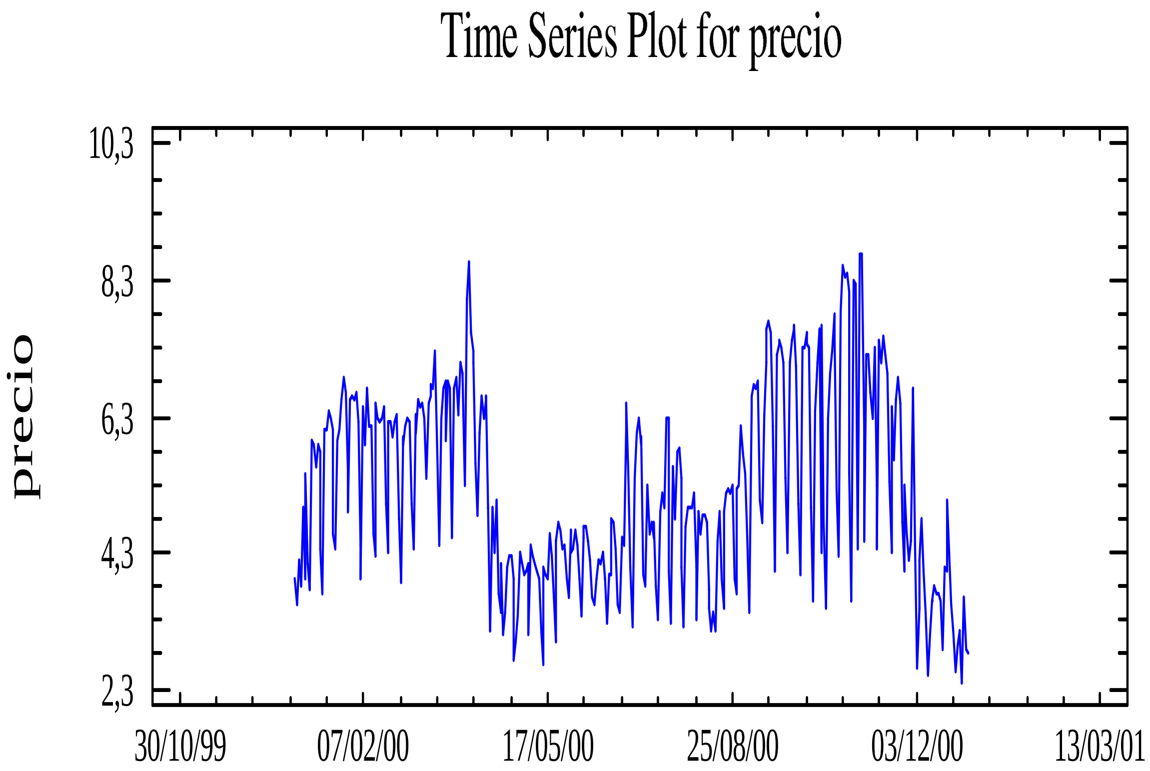
Ejemplos

Número de pasajeros de una compañía aérea (serie mensual).



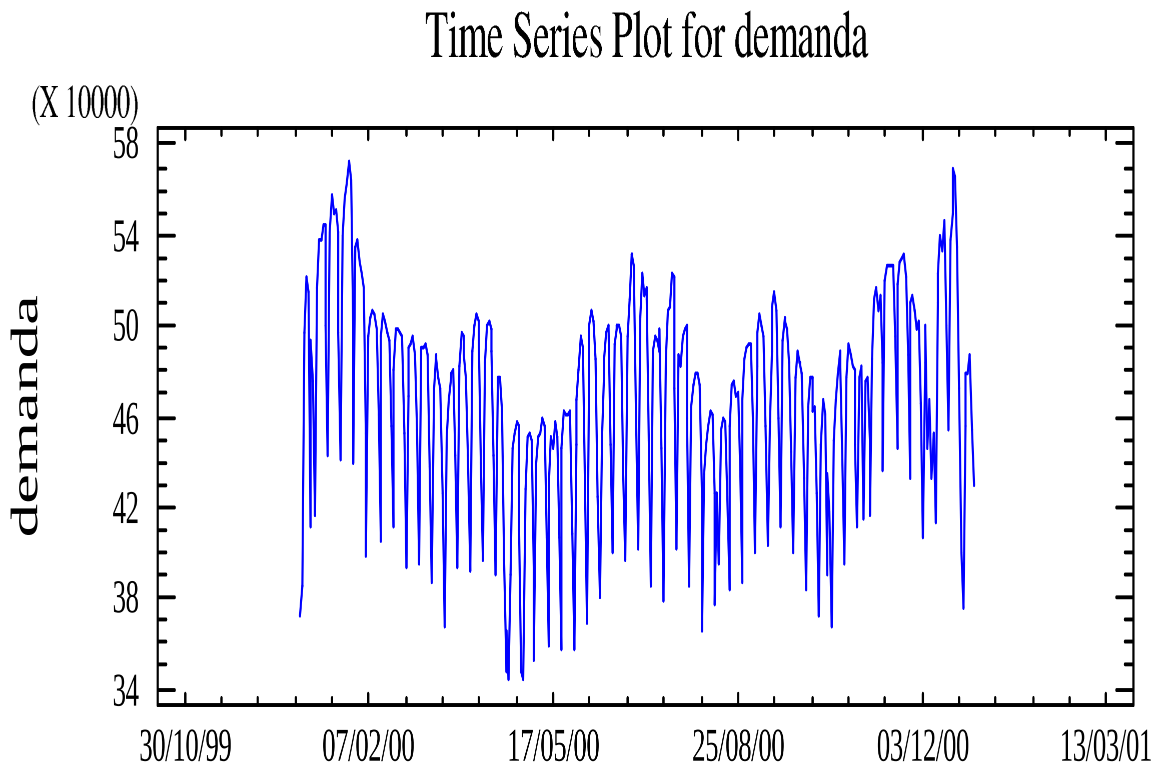
Ejemplos

Precios medios de energía (Pta/KWh) en España durante 2000 (serie diaria).



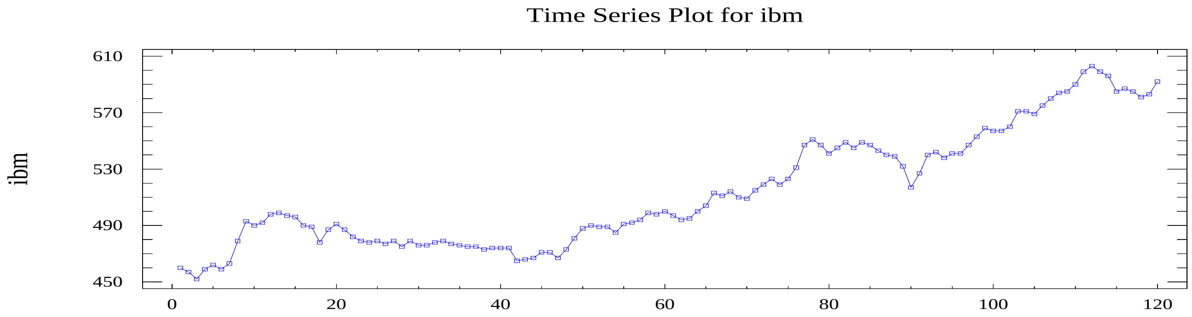
Ejemplos

Demandas medias de energía (KWh) en España durante 2000 (serie diaria).



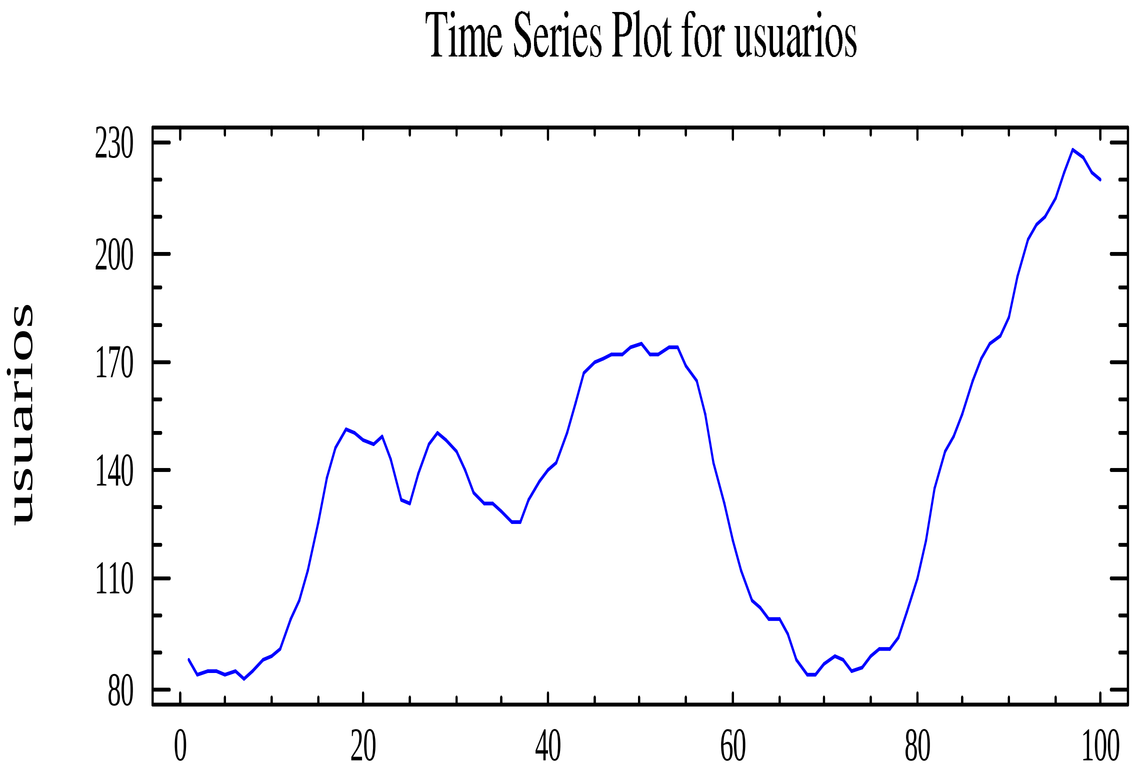
Ejemplos

Cierre de las acciones de IBM (serie diaria).



Ejemplos

Número de usuarios conectados a un servidor de Internet por minuto.



Introducción: Periodicidad

- Anual
- Mensual
- Semanal
- Diaria

Claramente, existen otros tipos de periodicidad: semestral, trimestral, horaria.

El tipo de periodicidad es una característica muy importante en una serie y aparecerán determinadas pautas debida a ella.

Tipos de series

No tiene sentido aplicar los métodos descriptivos de una variable, ya que no se tiene en cuenta la dependencia temporal. Por esta razón, ahora ya no disponemos de una muestra de tamaño n de una v.a. sino que disponemos de n muestras de tamaño 1 de n v.a.

- **Estacionarias:** Media y varianza constantes.
- **No Estacionarias:** Media y/o varianza cambian a lo largo del tiempo.

Además,

- **Estacionales:** Pauta de evolución que se repite.

Análisis: etapas

1. Análisis descriptivo: tendencia, estacionalidad, cambios estructurales, atípicos.

2. Análisis estadístico:

   a) Conseguir una serie estacionaria (residuos): eliminar la tendencia y componentes estacionales y aplicar transformaciones para estabilizar la variabilidad.
   b) Elegir un modelo que ajuste bien los residuos (se hace uso de las correlaciones).

3. A partir del modelo, predecir los residuos, r_t, y a partir de éstos predecir la serie original, z_t.

Análisis descriptivo de Series Temporales

1. **Tendencia:**
   - **Constante**
   - **Variable:** La tendencia cambia con el tiempo, tanto en magnitud como en dirección.

2. **Estacionalidad:** Unos meses están sistemáticamente por encima o por debajo de la media del año.

3. **Componente irregular:** Variaciones aleatorias alrededor de las componentes anteriores.
   - **Constante**
   - **Variable:** La variabilidad varía con el tiempo o con la magnitud de la serie.

Análisis descriptivo de una serie

- **Objetivo:** Identificar si la serie tiene tendencia, estacionalidad, cambios estructurales, heterocedasticidad, atípicos,...

- **Herramientas:**
  - Gráfico de la serie.
  - Gráfico de la descomposición estacional.

Gráfico de la serie de pasajeros

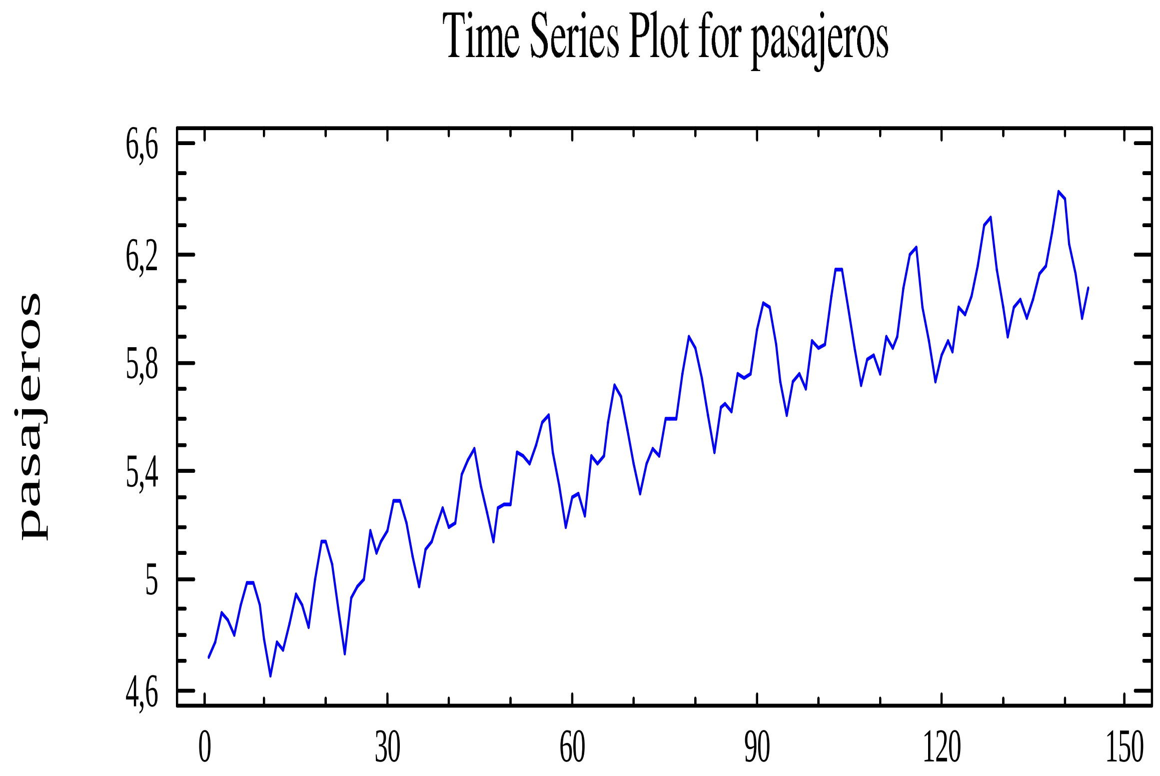


Gráfico del índice estacional

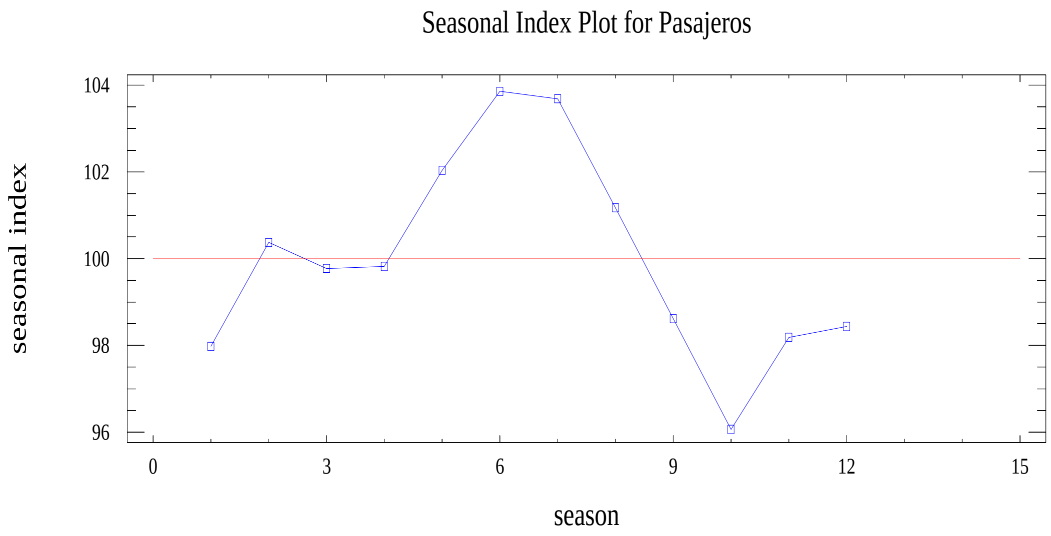
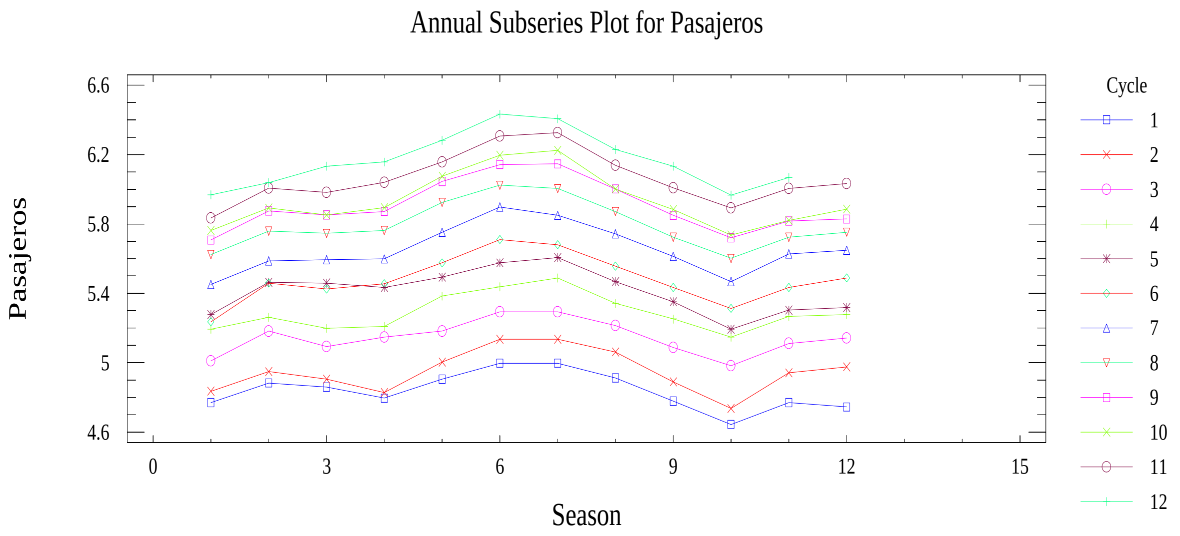


Gráfico de las subseries anuales



Análisis estadístico de la serie

- Para analizar y modelar la serie es necesario identificar la estructura que la genera, es decir, cómo influyen las observaciones del pasado en las futuras.

- Herramientas:
  - Función de Autocorrelación Simple **(FAS)**
  - Función de Autocorrelación Parcial **(FAP)**

Función de autocorrelación simple

- Proporciona la estructura de dependencia lineal de la serie.
- Los valores que se observan en la serie son:

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

- z_{t+1} representa el valor de la serie para un periodo próximo, es decir, un valor futuro.

Función de autocorrelación simple

- **Objetivo de la FAS:** Estudiar cómo se influyen las diversas observaciones.

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

- **Idea:** Dar el coeficiente de correlación entre las observaciones separadas un número determinado de periodos.

$$\rho_1, \rho_2, \ldots, \rho_k, \ldots \quad \in [-1, 1]$$

$\rho_1 \longrightarrow$ Cómo influye una observación en la siguiente, $z_i \longrightarrow z_{i+1}$.

$\rho_2 \longrightarrow$ Cómo influye una observación sobre la que está dos periodos hacia adelante, $z_i \longrightarrow z_{i+2}$.

$\rho_3 \longrightarrow$ Cómo influye una observación sobre la que está tres periodos hacia adelante, $z_i \longrightarrow z_{i+3}$.

Función de autocorrelación simple

Conclusión:

- La **FAS** proporciona información sobre cómo una observación influye en las siguientes.

Gráfico de la FAS de la serie de ventas

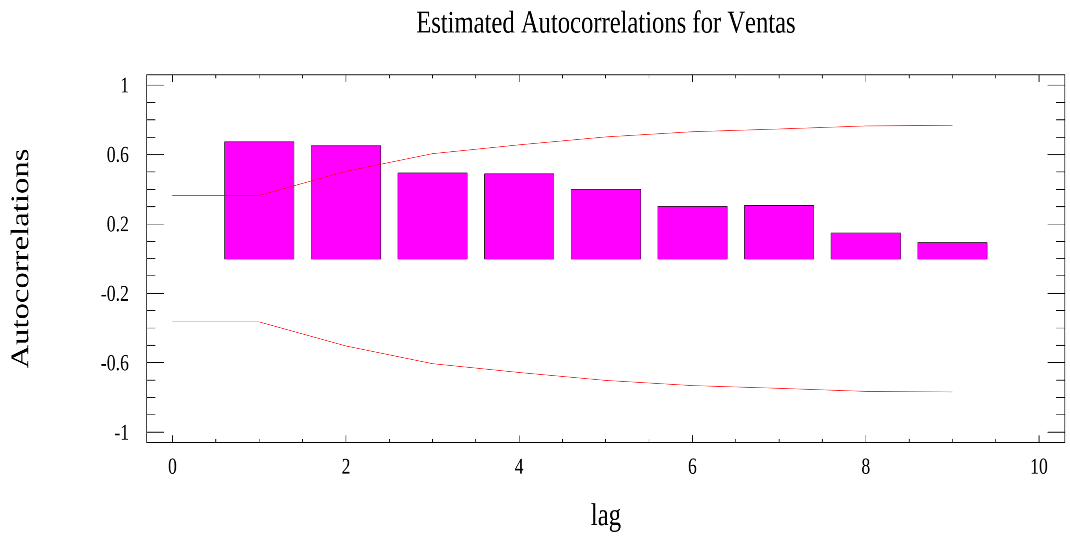


Gráfico de la FAS de la serie de pasajeros

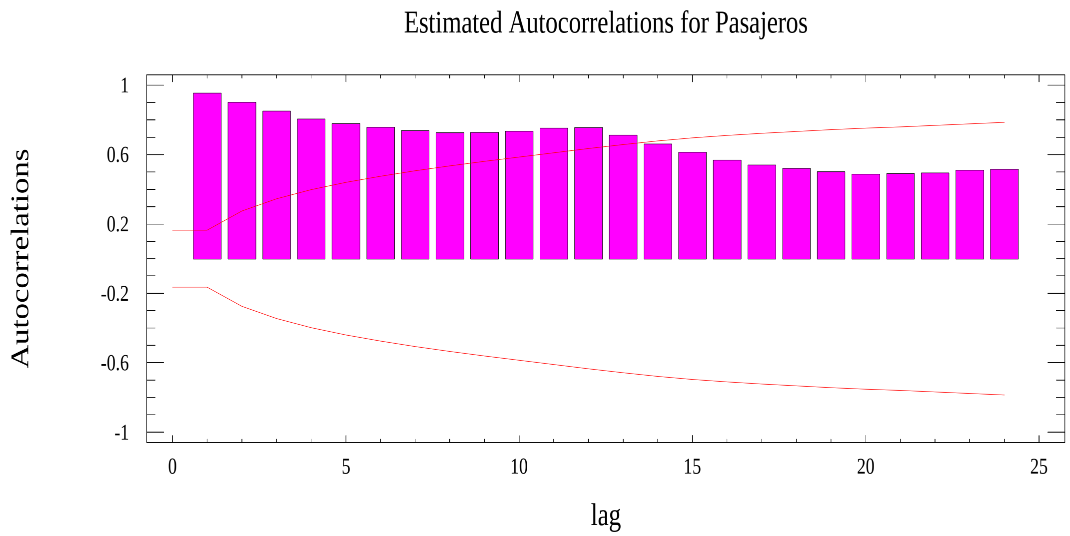


Gráfico de la FAS de la serie de acciones de IBM

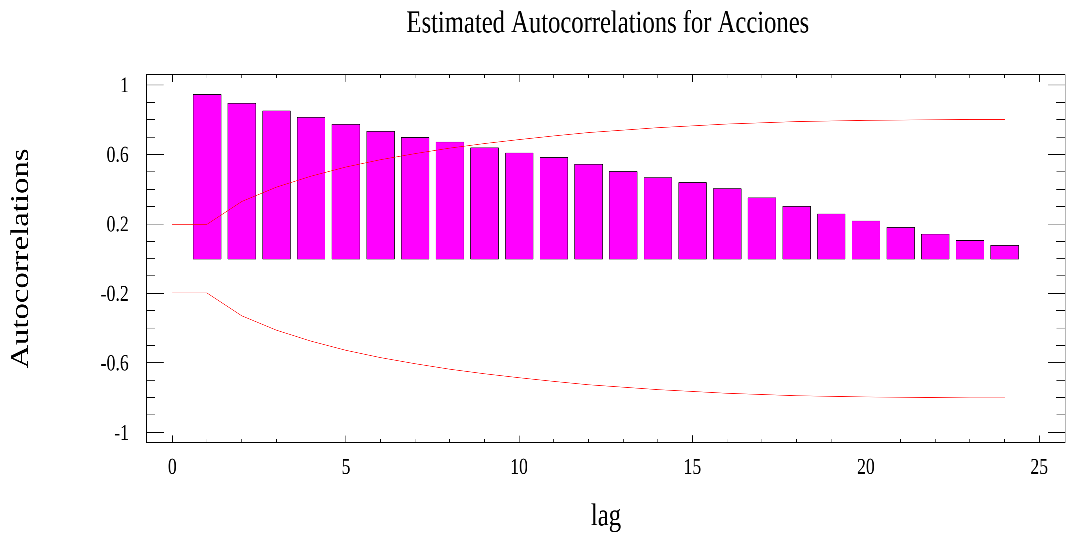
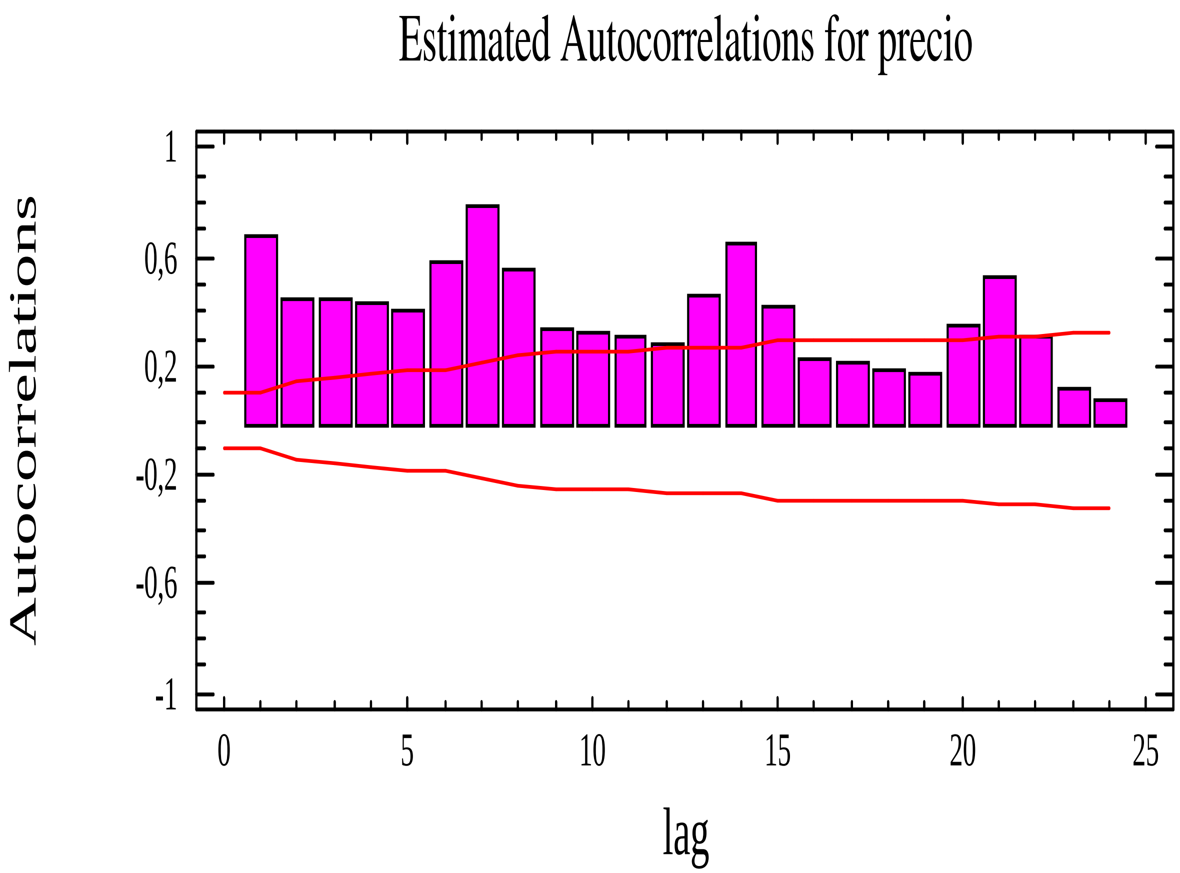


Gráfico de la FAS de la serie de precios



Función de autocorrelación simple

- **Problema de la FAS:** Si $\rho_1 \neq 0$, 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.

- Si $z_1 \longrightarrow z_2$ y $z_2 \longrightarrow z_3$, entonces $z_1 \longrightarrow z_3$.

- Hay que distinguir entre varias cadenas de influencia:
  - Cadena de influencia general a través de ρ_1.
  - Cadena de influencia directa. Cómo influye z_1 sobre z_3 directamente, sin pasar por z_2.

Función de autocorrelación parcial

- Proporciona la relación **directa** existente entre observaciones separadas por k retardos.
- Elimina el problema que presentaba la FAS.
- En la FAS, si ρ_1 es significativo, también lo será ρ_2.
- En la FAP, esto no tiene por qué ocurrir.

Gráfico de la FAP de la serie de ventas

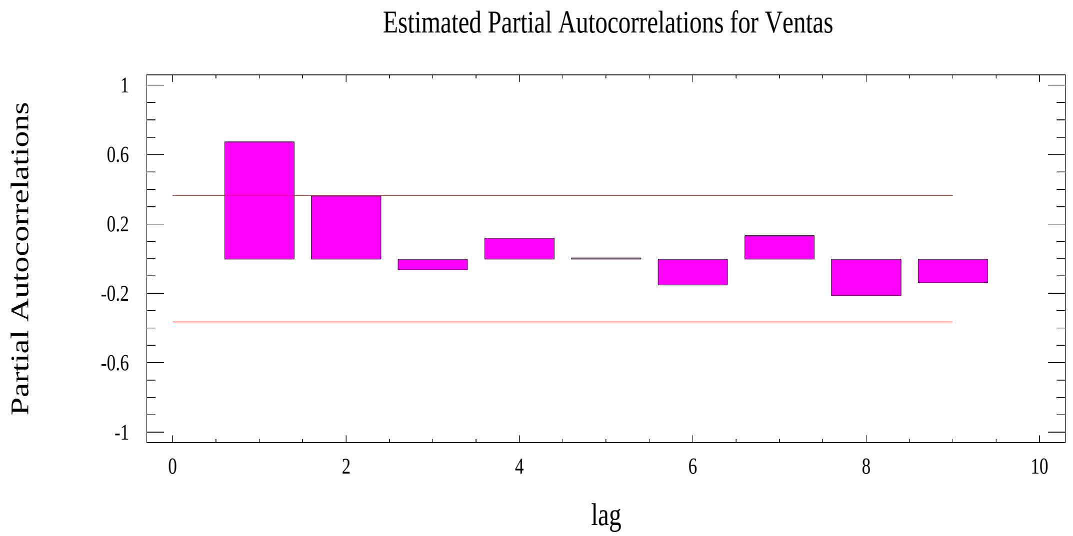


Gráfico de la FAP de la serie de pasajeros

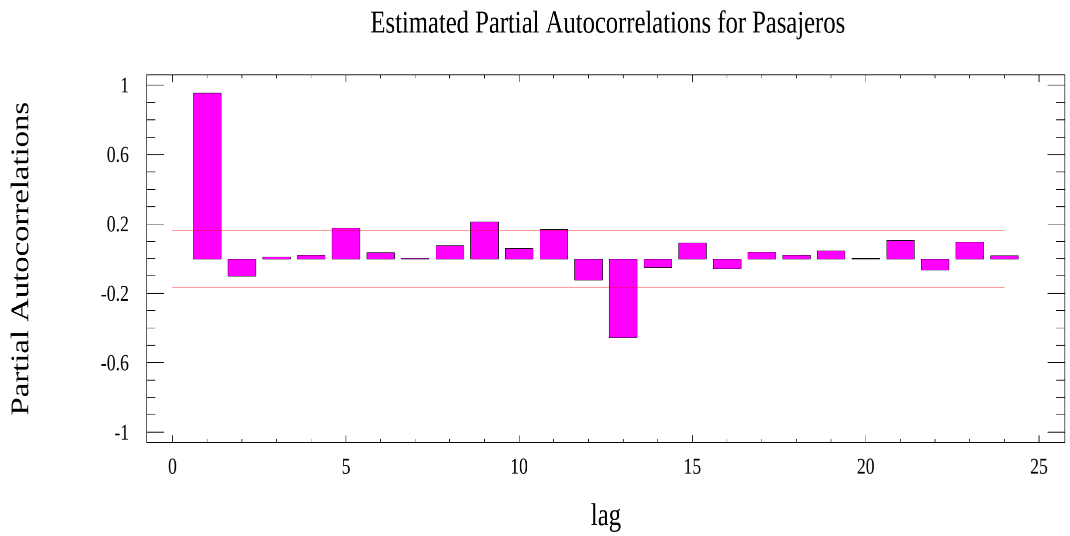


Gráfico de la FAP de la serie de acciones de IBM

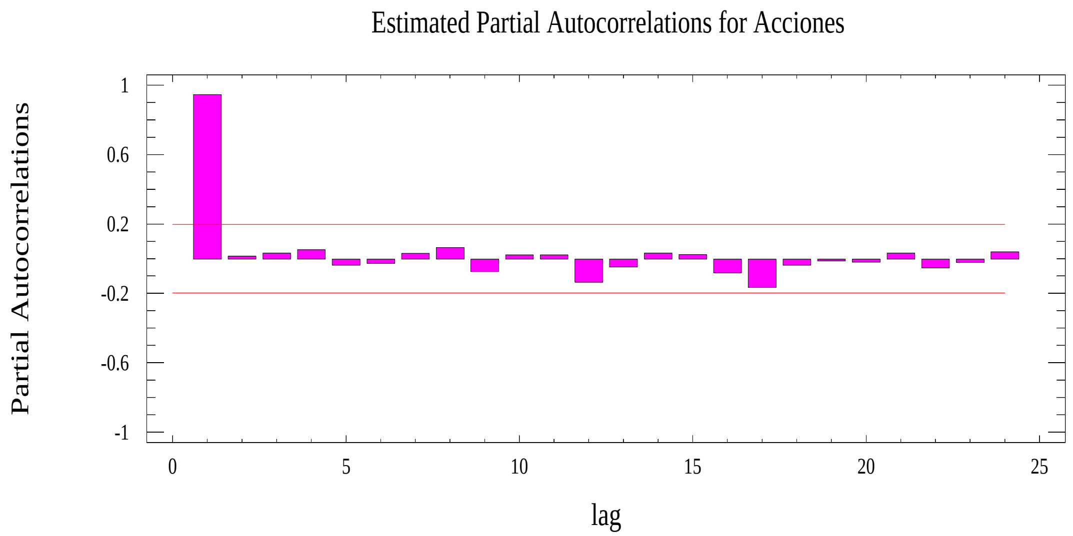
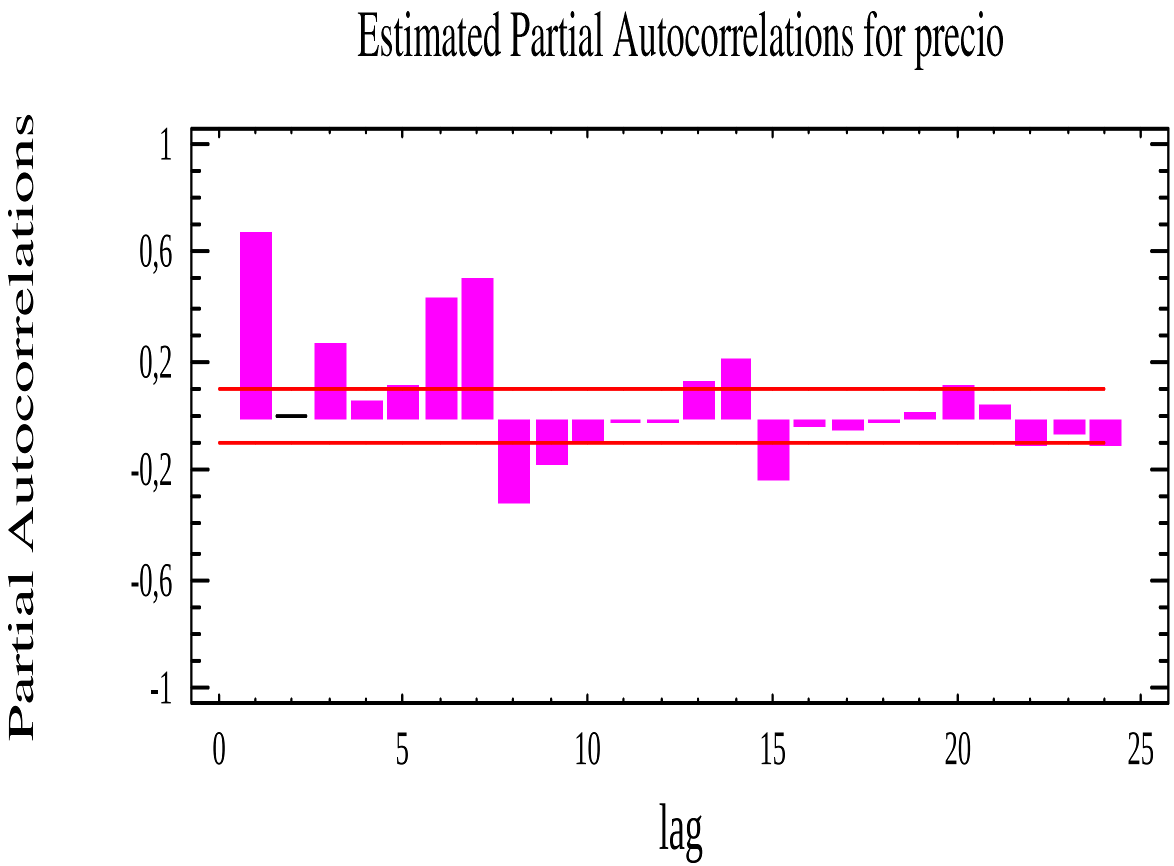


Gráfico de la FAP de la serie de precios



Series estacionarias

Características:

- No tiene tendencia.
- Homocedástica.
- No tiene ciclos estacionales.
- La estructura de dependencia se mantiene constante. Condición importante para modelizar la serie y poder prever su evolución.
- La influencia de las observaciones sobre las posteriores decrece con el tiempo.

Tipo especial de serie estacionaria: **Ruido Blanco.** Es una serie estacionaria en la que ninguna observación influye en las siguientes. Tanto la FAS como la FAP de un RB son una sucesión de palos no significativos.

Gráfico de la FAS de un ruido blanco

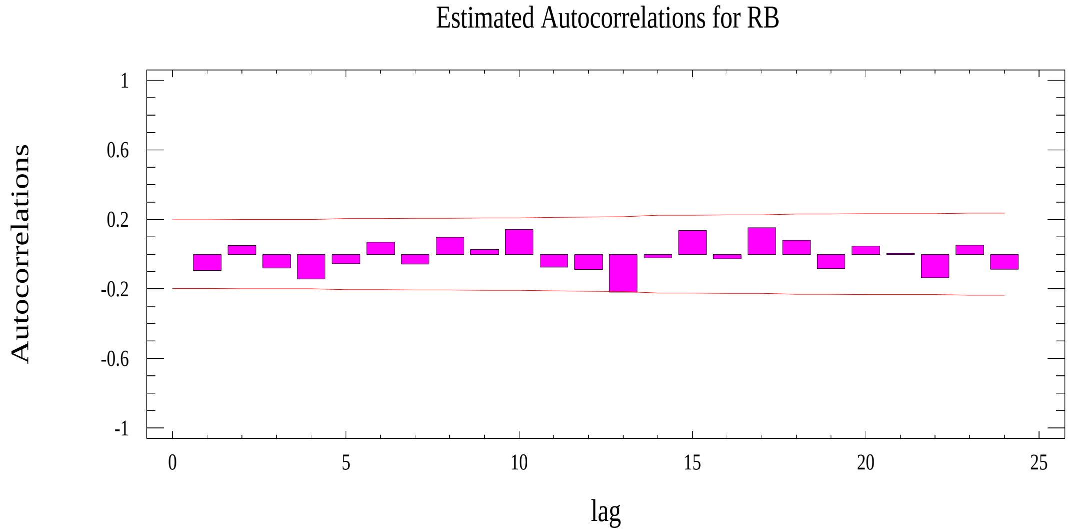
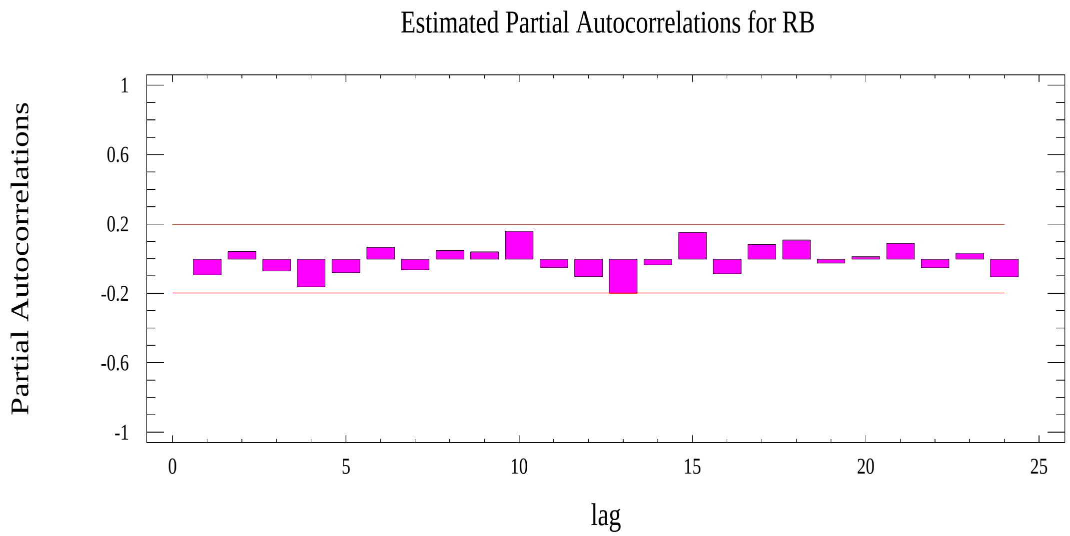


Gráfico de la FAP de un ruido blanco



Objetivo

Hay que ir consiguiendo componentes irregulares (residuos) cada vez más parecidos a un ruido blanco.

Gráfico de la serie de acciones de IBM

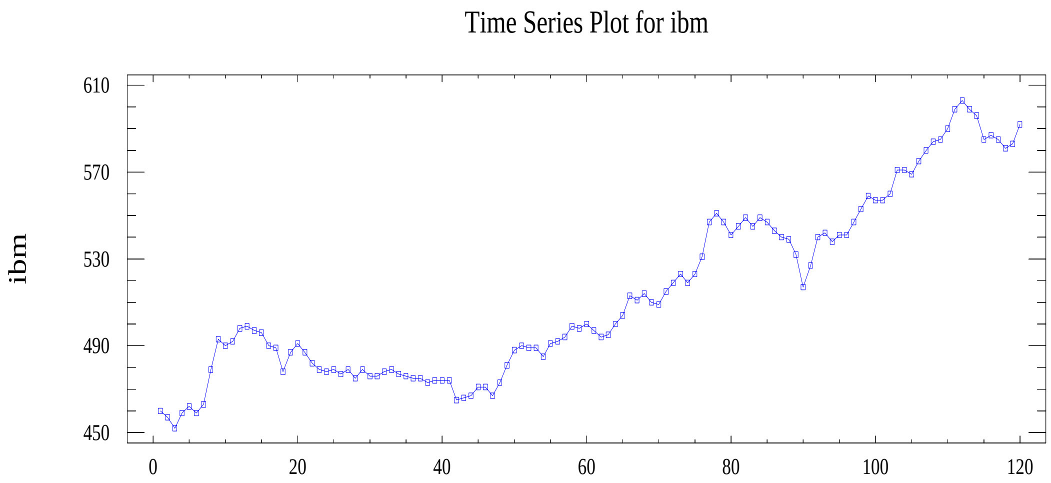


Gráfico de la FAS de IBM

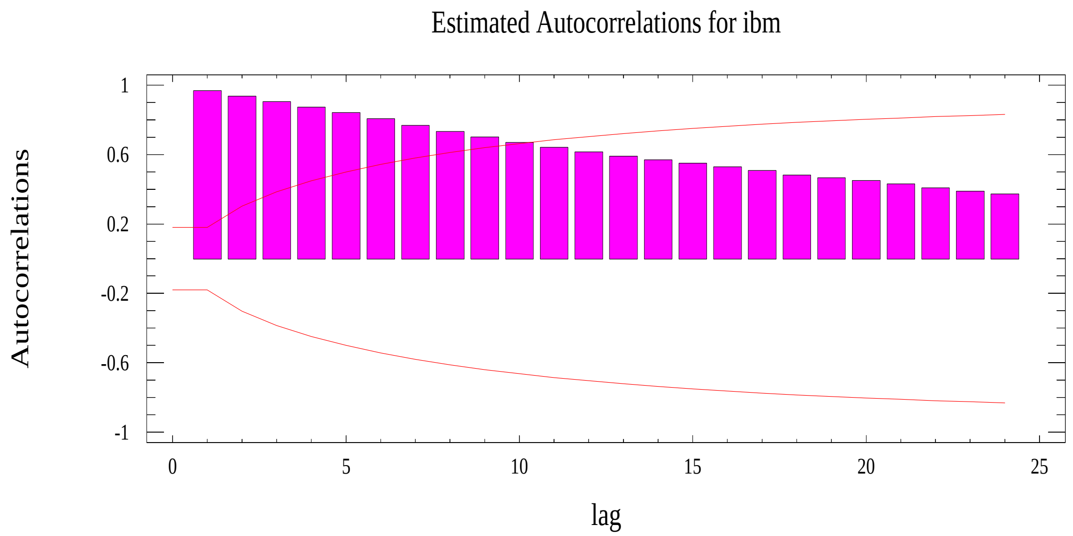


Gráfico de la serie de acciones de IBM en diferencias

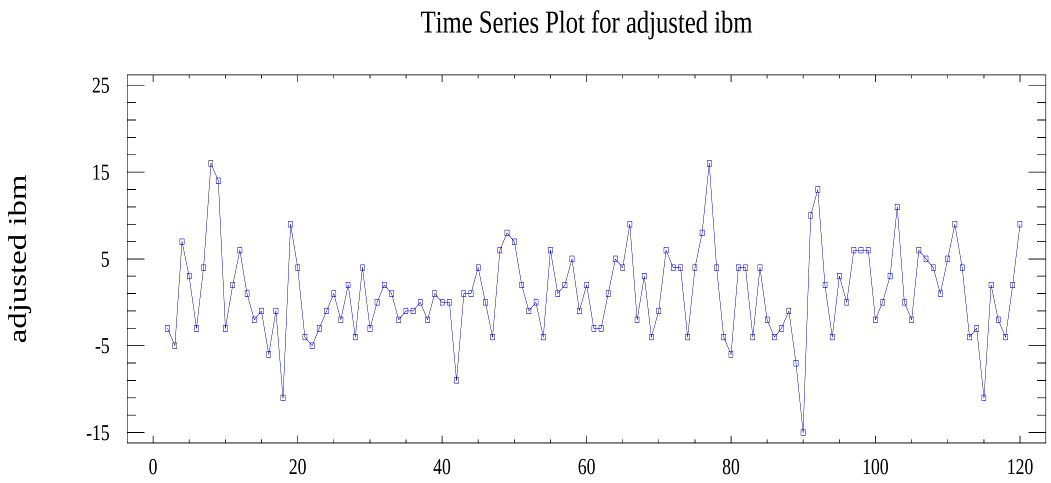
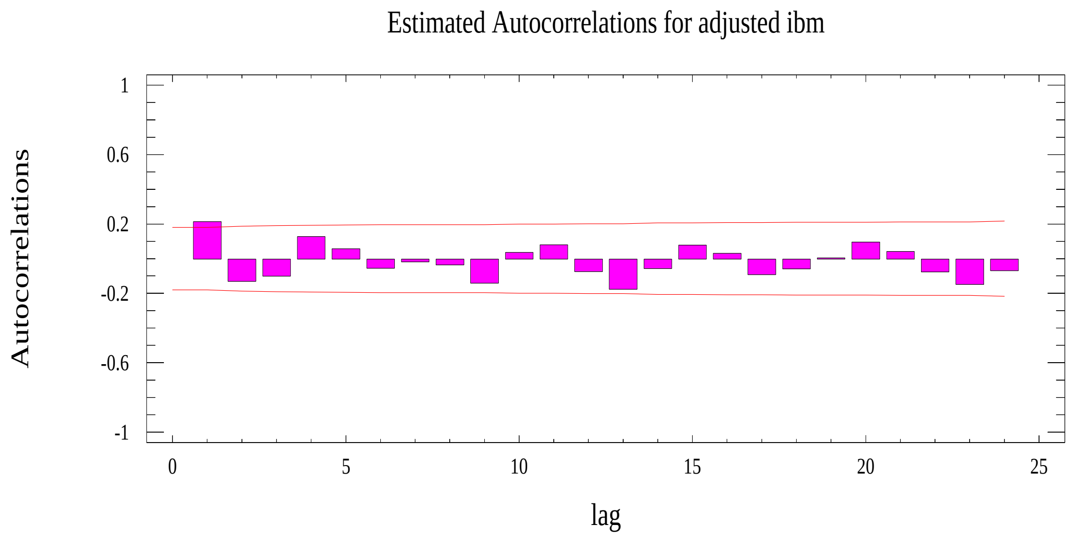


Gráfico de la FAS de IBM en diferencias



Modelos ARMA y ARIMA

- Recordemos que el objetivo principal de series temporales es predecir una serie de datos no determinista (contiene un componente aleatorio).
- Si el componente aleatorio es estacionario, se pueden desarrollar técnicas eficientes para predecir valores futuros de la serie.
- Si el componente aleatorio no es estacionario, primero se debe convertir dicho componente en estacionario, se predice dicho componente y, finalmente, se deshace la conversión para recuperar la predicción de interés.

Modelos ARMA y ARIMA

- Formalmente, un proceso z_t es estacionario si

  1. $\mu_t = E(z_t) = \mu \;\; \forall t$ (tendencia constante).
  2. $\sigma_t^2 =$ Var$(z_t) = \sigma^2 \;\; \forall t$ (varianza constante).
  3. Cov$(z_t, z_{t+k}) =$ Cov$(z_t, z_{t-k}) = \gamma_k \;\; \forall t, \; k = 0, \pm 1, \pm 2, \ldots$ (dependencia temporal constante).

- La FAS de un proceso estacionario se define como

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \; k = 0, 1, 2, \ldots \quad (\rho_0 = 1).$$

- Para modelar un proceso estacionario, nos centraremos en técnicas basadas en modelos ARMA y ARIMA: “AutoRegressive Moving Average” y “AutoRegressive Integrated Moving Average”.

Procesos autorregresivos

- Una observación depende (linealmente) de las anteriores.
- Se denominan procesos AutoRegresivos (AR) y se caracterizan por su orden (de dependencia).
- Se supone que la serie proviene de un proceso estacionario.

Procesos autorregresivo de primer orden AR(1)

- Es el más sencillo.
- Sigue la siguiente ecuación:

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

- Cada observación se construye a partir de la anterior más una perturbación aleatoria:

$$a_t \sim \mathcal{N}(0, \sigma_a^2) \text{ independientes } \forall t \;\equiv\; \text{ ruido blanco}$$

Ejemplo: Si $\phi = 0.6$ y $z_0 = 0$, entonces $z_1 = 0.6 \times 0 + a_1 = a_1$.
$z_2 = 0.6 \times a_1 + a_2$ y así sucesivamente.

Procesos autorregresivos de primer orden AR(1)

- El proceso AR(1) es estacionario si $|\phi| < 1$ (ejercicio).
- Si $|\phi| > 1$ el proceso es explosivo.
- La media del proceso es (ejercicio): $\mu = \frac{c}{1-\phi}$.
- La varianza del proceso es (ejercicio): $\sigma^2 = \frac{\sigma_a^2}{1-\phi^2}$.

Procesos autorregresivos de primer orden AR(1)

- Además, $\rho_k = \phi^k$ (ejercicio) $\rightarrow$ decrecimiento exponencial:
  - Si ϕ es positivo, la FAS decrece positivamente.
  - Si ϕ es negativo, la FAS decrece de forma alternada.
- Sólo existe influencia directa de primer orden.
  - Si ϕ es positivo, la FAP tendrá un único palo significativo y será positivo.
  - Si ϕ es negativo, la FAP tendrá un único palo significativo y será negativo.

Gráfico de un AR(1) con $\phi > 0$

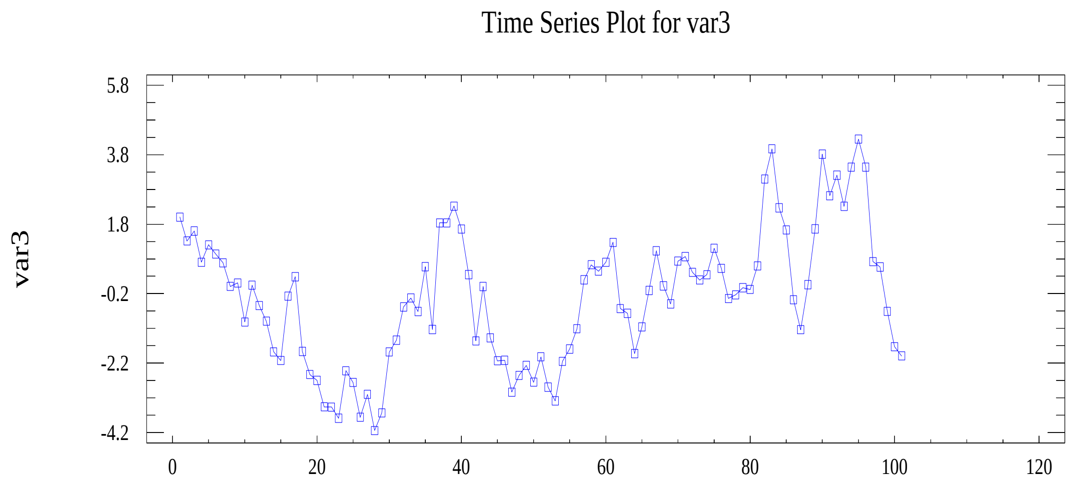


Gráfico de la FAS de un AR(1) con $\phi > 0$

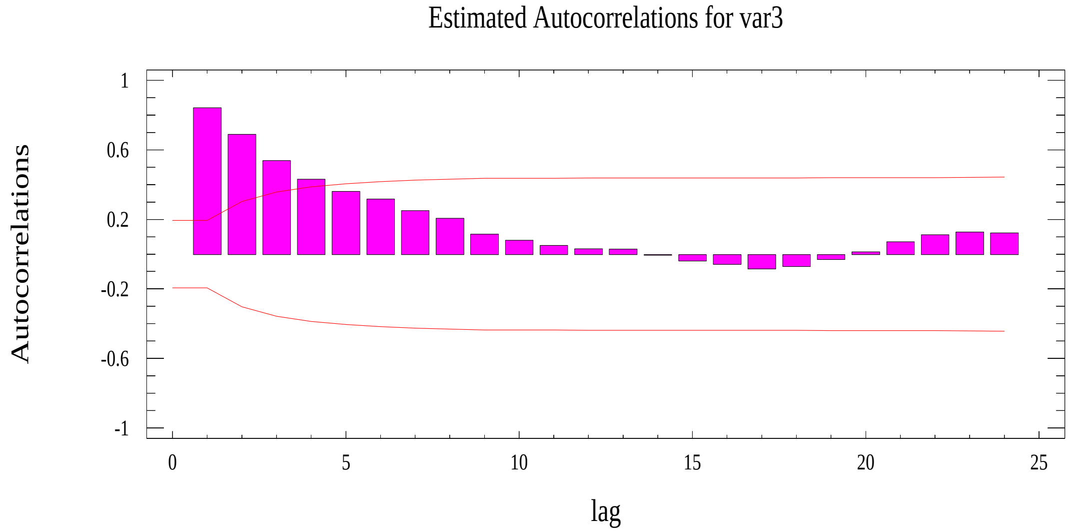


Gráfico de la FAP de un AR(1) con $\phi > 0$

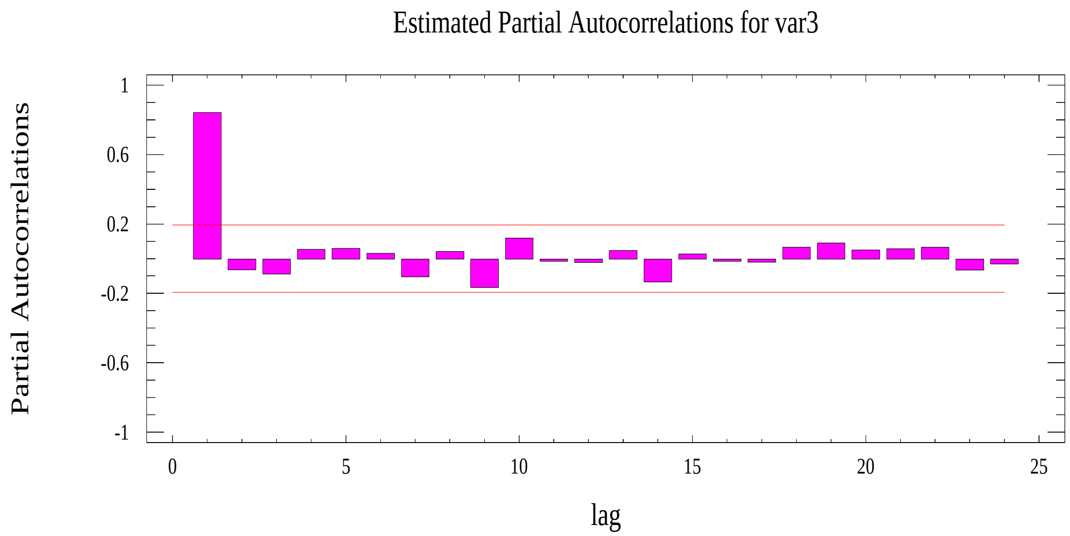


Gráfico de un AR(1) con $\phi < 0$

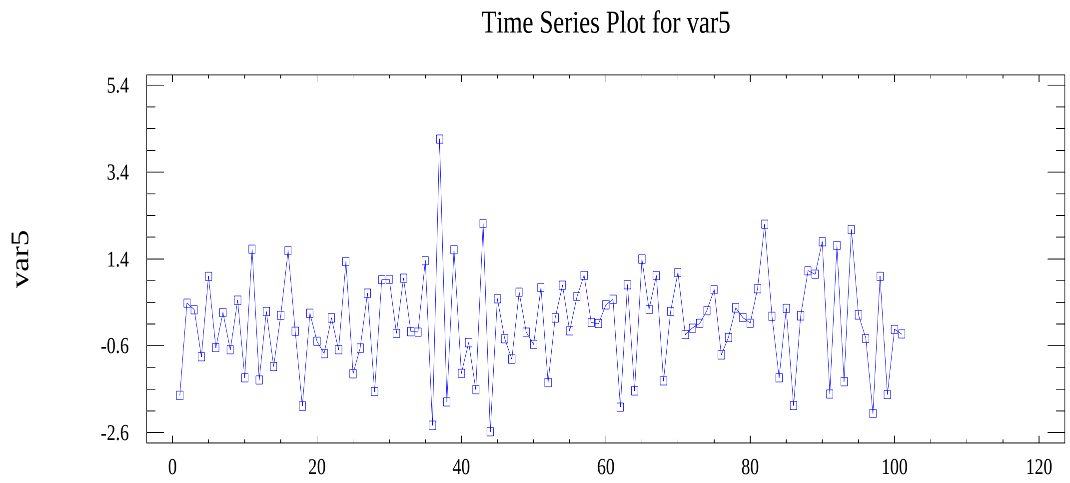


Gráfico de la FAS de un AR(1) con $\phi < 0$

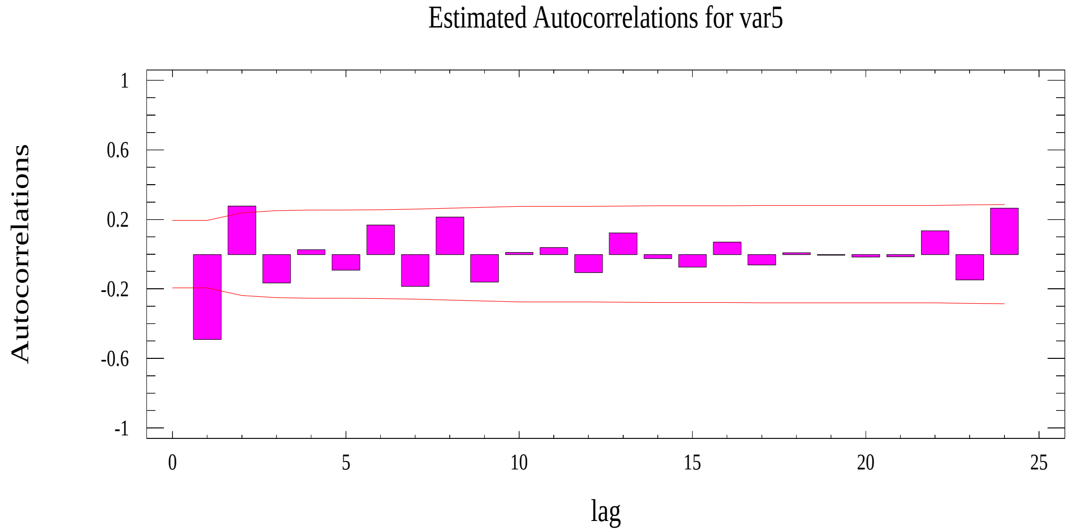
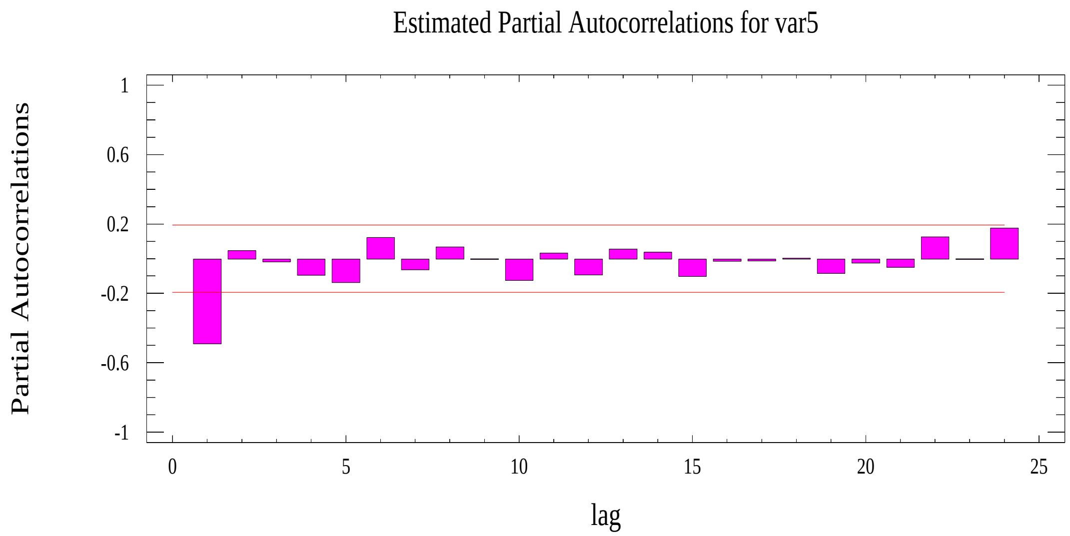


Gráfico de la FAP de un AR(1) con $\phi < 0$



Procesos autorregresivos de segundo orden AR(2)

- La ecuación de un proceso AR(2) es:

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

- El polinomio característico de un AR(2) se define como:

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

- Para que el proceso sea estacionario, las raíces del polinomio característico deben estar fuera del círculo unidad.
- La FAS es más compleja y admite varias posibilidades, en función de cómo sean las raíces del polinomio característico.
- La FAP tendrá únicamente dos palos significativos.

Gráfico de un AR(2)

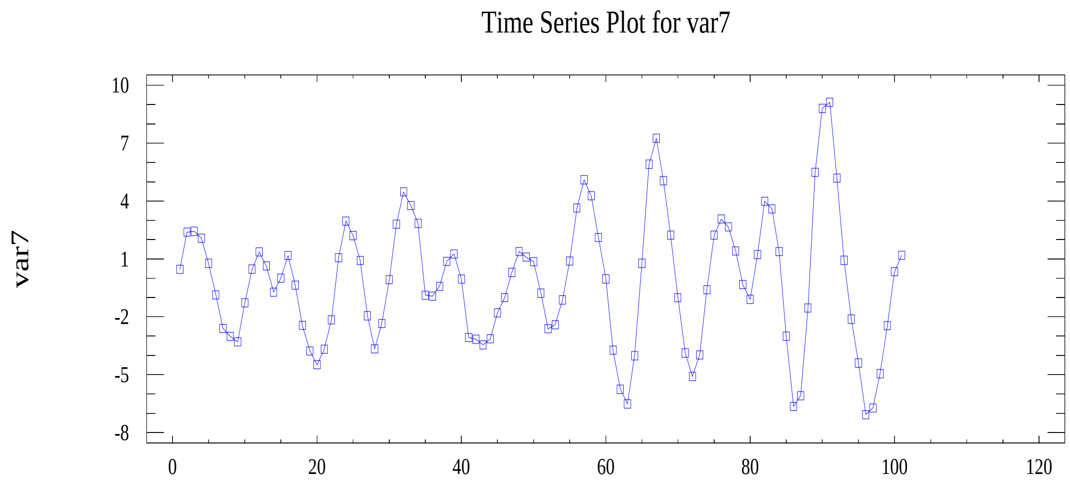


Gráfico de la FAS de un AR(2)

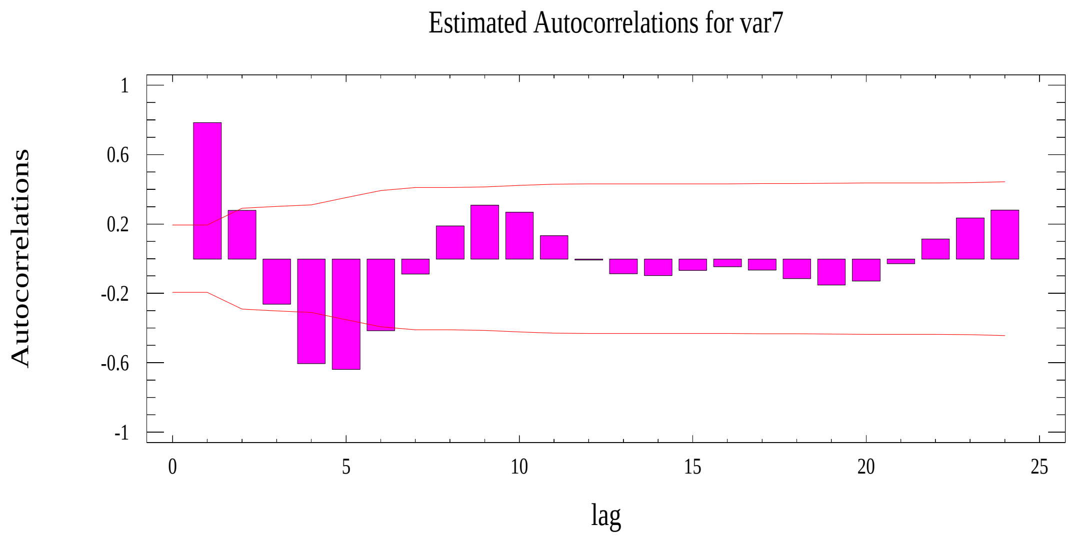
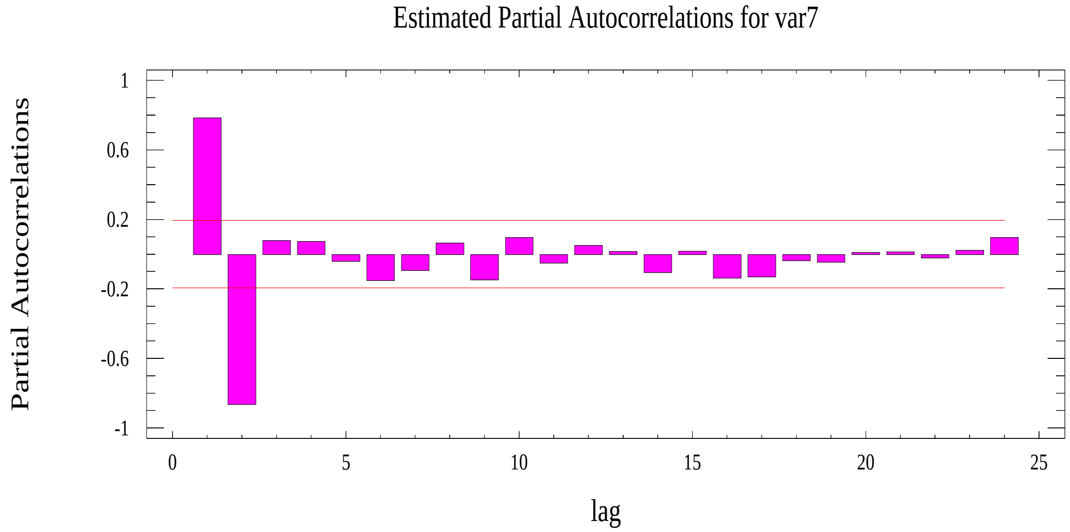


Gráfico de la FAP de un AR(2)



Procesos autorregresivos de orden superior, AR(p)

- La ecuación de un proceso AR(p) es:

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

- Una observación está influida por las p observaciones anteriores.
- El polinomio característico es:

$$1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p$$

- Para que el proceso sea estacionario, las raíces del polinomio característico deben estar fuera del círculo unidad.
- La FAS es más compleja.
- La FAP tendrá únicamente p palos significativos.

Procesos de media móvil, MA

- Característica de los procesos AR: tienen memoria larga, es decir, los palos de la FAS decrecen lentamente.
- Un proceso AR tarda bastante tiempo en absorber los impactos externos.
- La realidad demuestra que existen series que absorben rápidamente los impactos.
- Se introduce otra familia de procesos de memoria corta: procesos de media móvil (“moving average”).

Procesos de media móvil de primer orden, MA(1)

- Su ecuación es (suponiendo que a z_t le hemos restado la media):

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

- A a_t también se le denomina innovación y representa los efectos externos a la serie.
- En un MA(1), los valores vendrán dados por los efectos externos muy recientes, el actual a_t y el anterior a_{t-1}.
- Tiene memoria corta y absorbe rápidamente los impactos.
- La FAS de un MA(1) es como la FAP de un AR(1), ídem para la FAP:
  - La FAS tiene un único palo significativo.
  - LA FAP tiene varios palos que van decreciendo.

Gráfico de un MA(1)

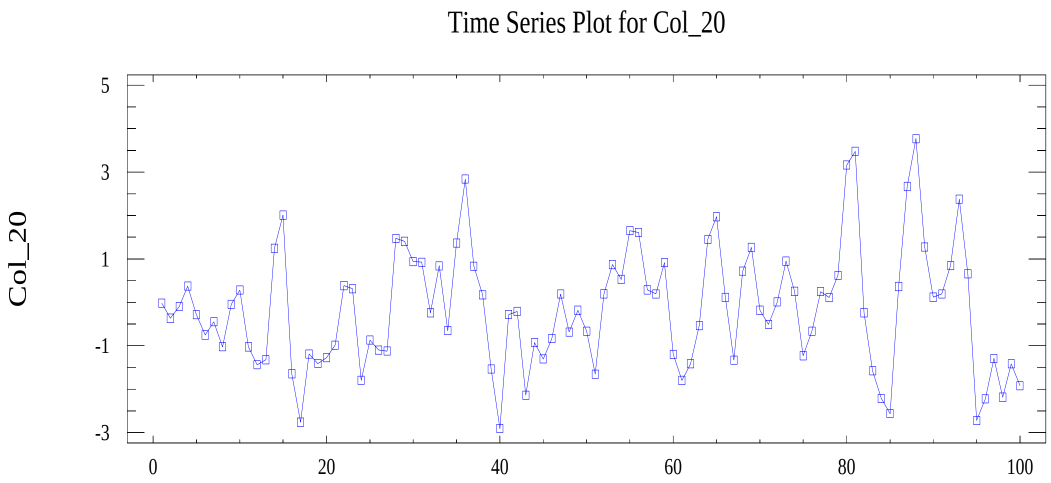


Gráfico de la FAS de un MA(1)

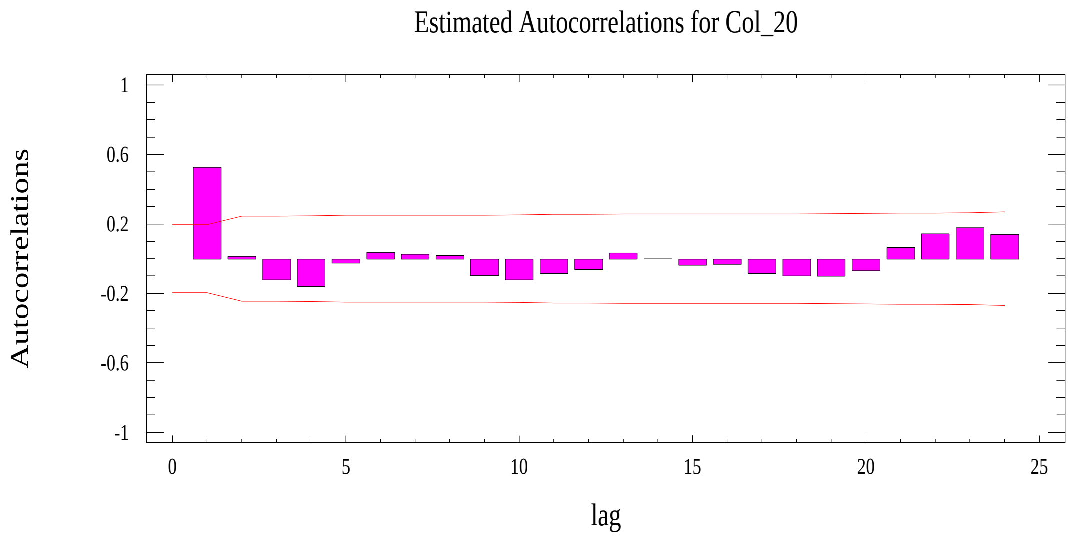
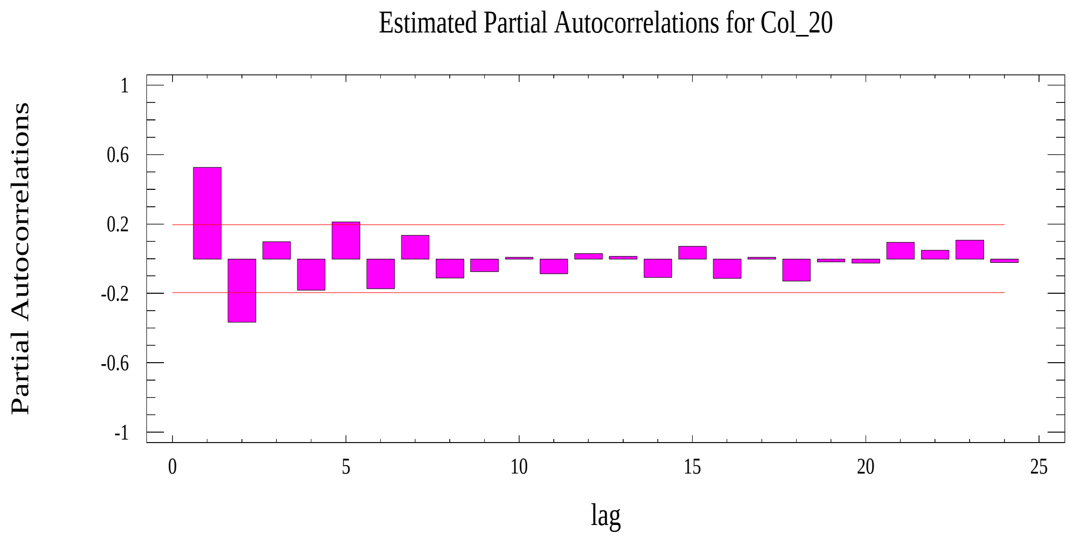


Gráfico de la FAP de un MA(1)



Procesos de media móvil de orden superior, MA(q)

- La ecuación de un proceso MA(q) (de media 0) es:

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

- El proceso tarda q periodos en absorber los impactos.
- Su FAS tendrá q palos significativos.
- Su FAP presentará un decrecimiento.

Procesos ARMA(p,q)

- En la práctica, las series presentan parte AR y parte MA.
- Los procesos ARMA(p,q) son una combinación de estructuras autorregresivas y de media móvil y tienen una parte AR(p) y una parte MA(q).
- La FAS y la FAP serán combinación de ambos procesos:
  - FAS: Los primeros q palos vendrán dados por la parte MA. A partir de ahí, se producirá un decrecimiento, que viene dado por la parte AR.
  - FAP: Los primeros p palos vendrán dados por la parte AR. A partir de ahí, se producirá un decrecimiento, que viene dado por la parte MA.

Gráfico de un ARMA(1,1)

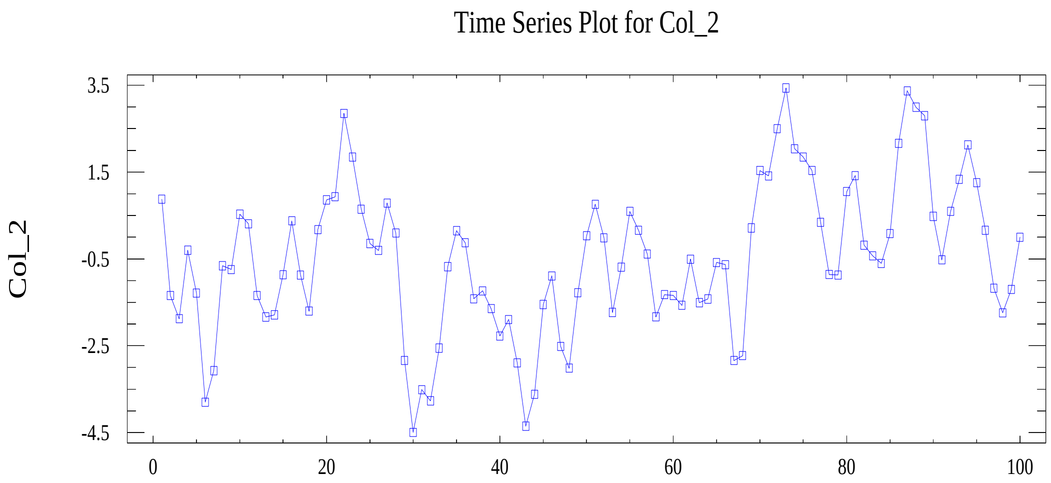


Gráfico de la FAS de un ARMA(1,1)

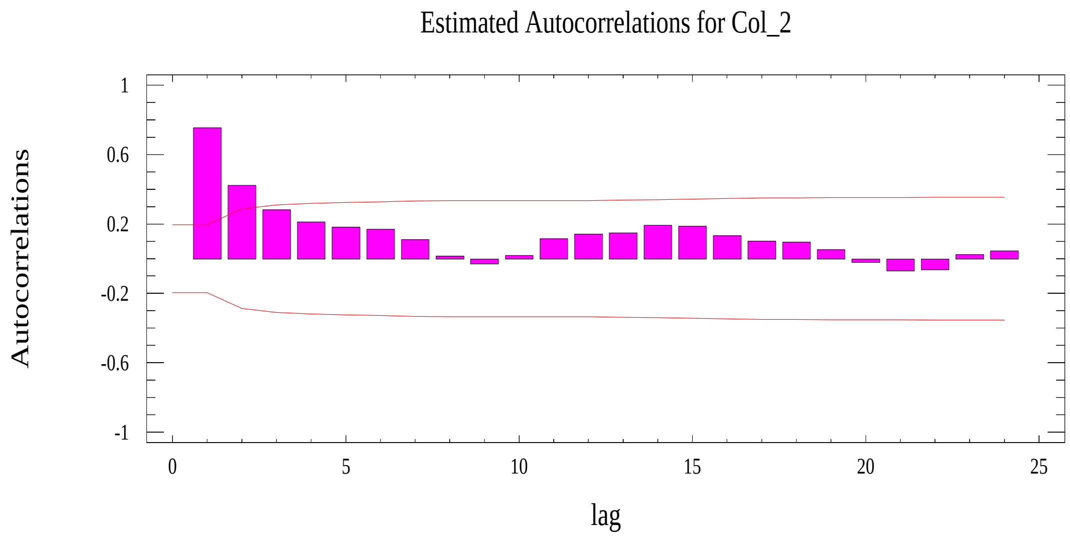
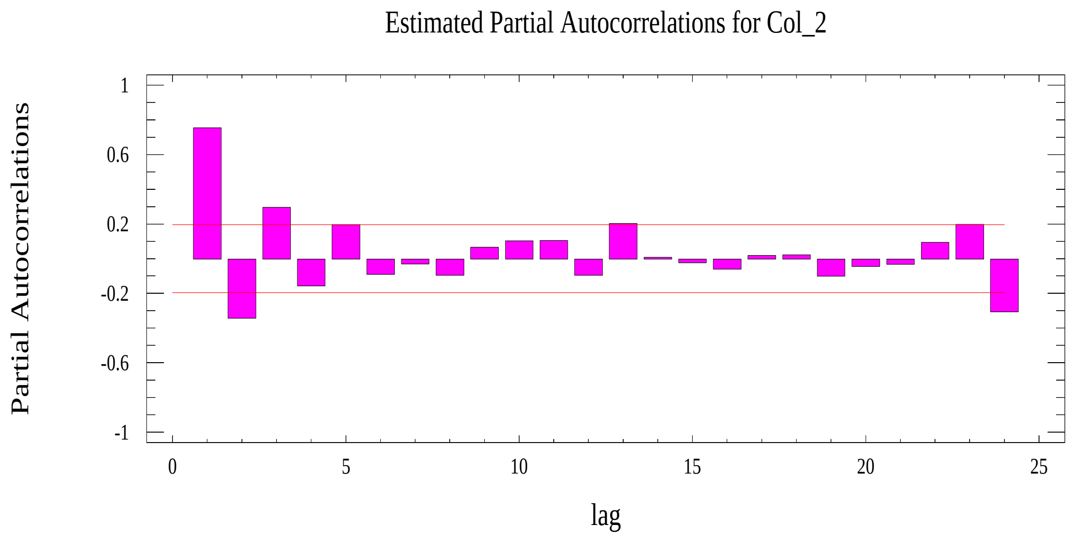


Gráfico de la FAP de un ARMA(1,1)



Series no estacionarias. Procesos ARIMA(p,d,q)

- Los modelos ARMA sirven para ajustar series estacionarias.
- Si una serie no lo es, hay que transformarla como se indicó en el primer tema: eliminar tendencia, heterocedasticidad,...
- Los procesos ARIMA eliminan la tendencia de una serie diferenciándola d veces. Típicamente, $d = 1, 2$. Por tanto, un ARIMA(2,1,1) es un ARMA(2,1) diferenciado 1 vez.
- Notación:

$$\Delta z_t = z_t - z_{t-1},\ \Delta^2 z_t = \Delta\{z_t - z_{t-1}\} = z_t - 2z_{t-1} + z_{t-2},\ \text{etc.}$$

Gráfico de la serie de acciones de IBM

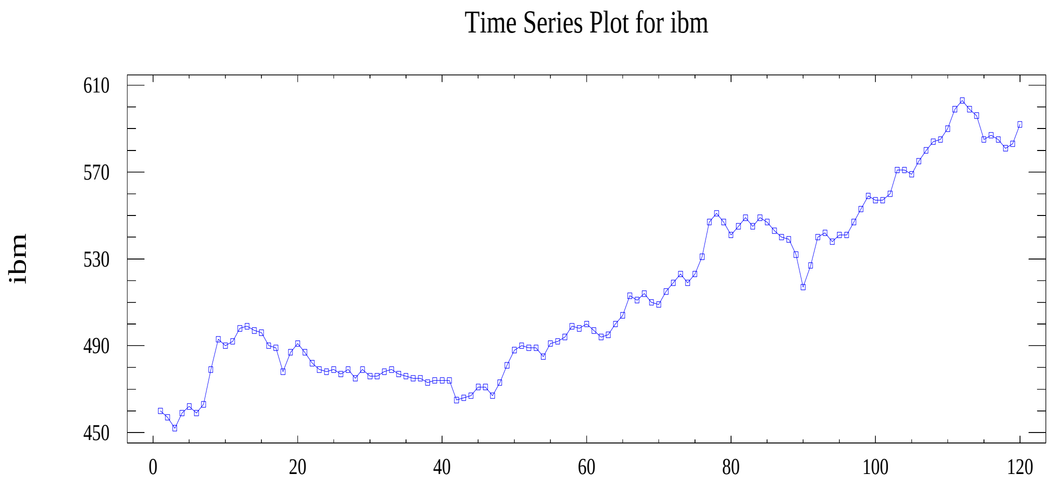


Gráfico de la FAS de IBM

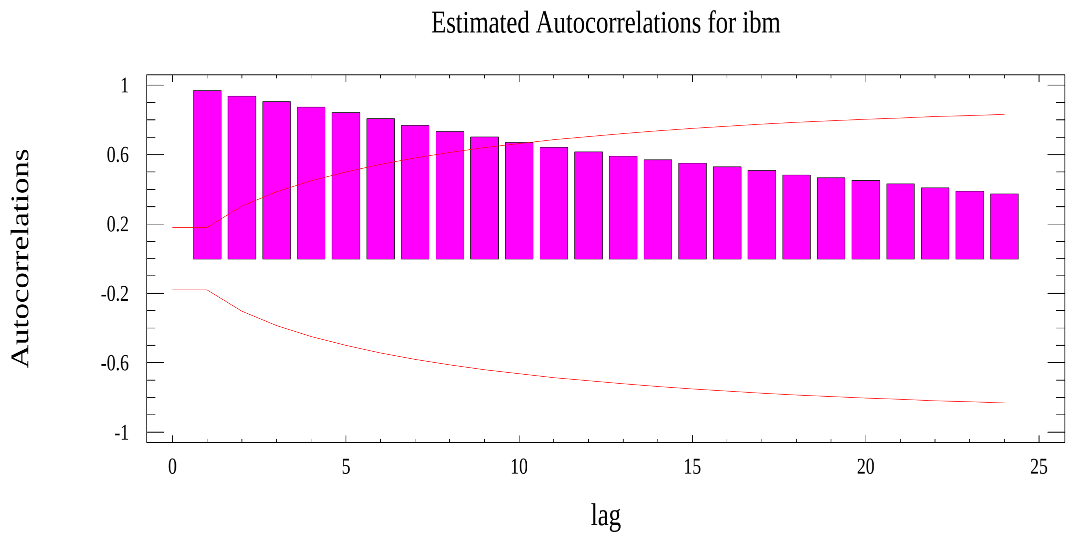


Gráfico de la serie de acciones de IBM diferenciada

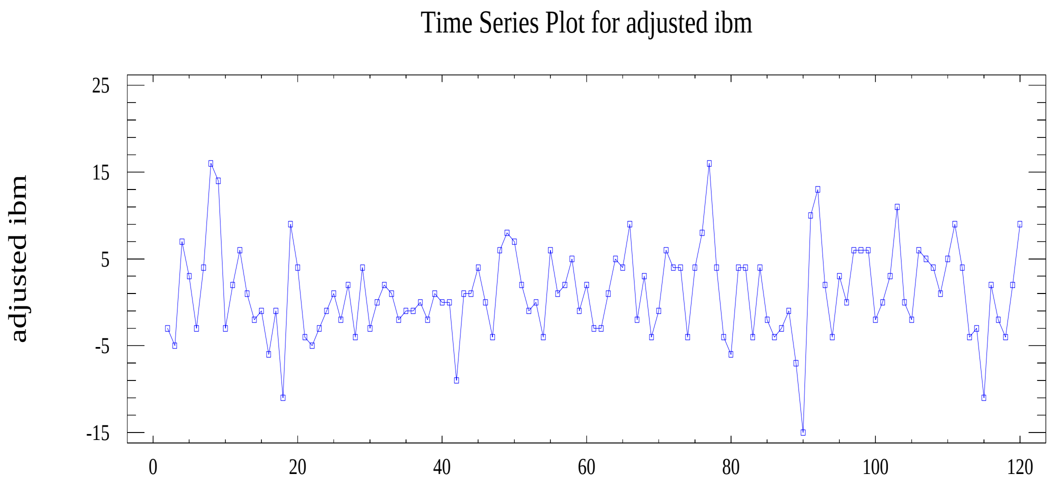


Gráfico de la FAS de IBM diferenciada

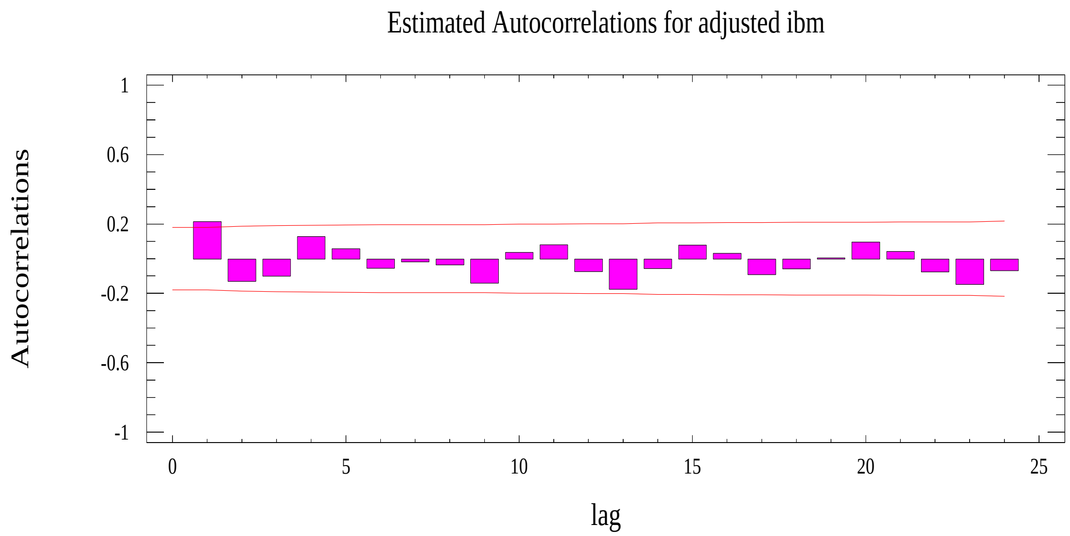
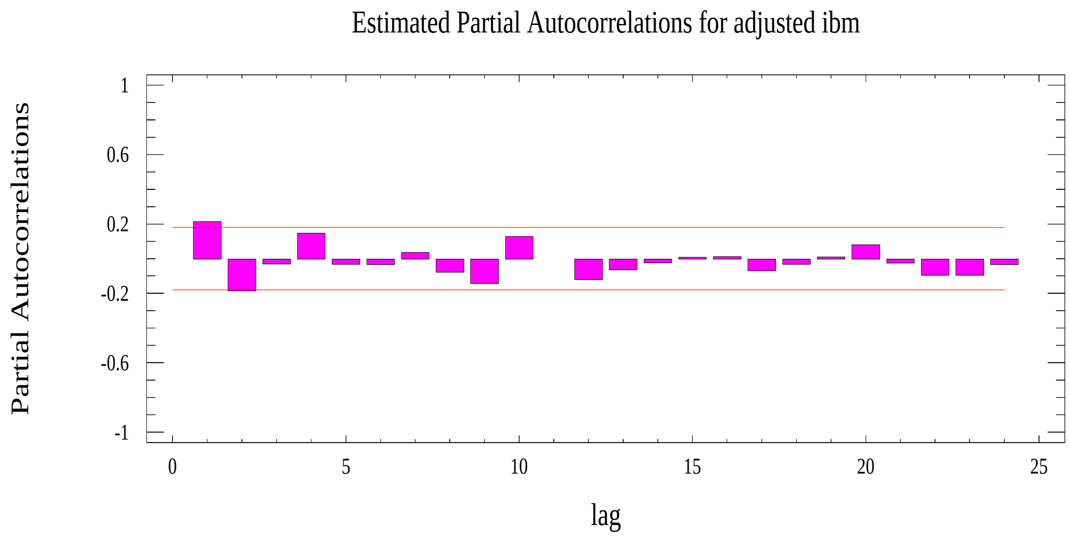


Gráfico de la FAP de IBM diferenciada



Modelo para la serie de IBM diferenciada

- Tiene pinta de que ΔIBM sigue un modelo ARMA(0,1), esto es, la serie de acciones de IBM sigue un modelo ARIMA(0,1,1).

- Este modelo es:

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

- Hay que estimar μ, θ_1, σ_a^2.

Estimación de modelos ARIMA(p,d,q)

- Una vez identificada la serie, es necesario estimar los parámetros: μ, ϕ_i, θ_j, σ_a^2.

- La estimación es compleja (sobre todo en presencia de MA) y requiere algoritmos de optimización sofisticados. Básicamente, la estimación se basa en métodos de mínimos cuadrados o en métodos de máxima verosimilitud.

- El resultado es una tabla similar a la de regresión: figuran los valores estimados, los errores estándar y los estadísticos t.

Estimación de modelos ARIMA(p,d,q)

- Ejemplo de la serie de IBM. Si se ajusta un MA(1) a la serie diferenciada:

```
                            ARIMA Model Summary
Parameter            Estimate        Stnd. Error     t               P-value
----------------------------------------------------------------------------
MA(1)                -0,270708       0,0908496       -2,97974        0,003509
Mean                 1,12023         0,583698        1,91919         0,057396
Constant             1,12023
----------------------------------------------------------------------------
Backforecasting: yes
Estimated white noise variance = 25,9011 with 117 degrees of freedom
Estimated white noise standard deviation = 5,08931
```

Estimación de modelos ARIMA(p,d,q)

- Ejemplo de la serie de IBM. Si se ajusta un MA(2) a la serie diferenciada:

```
                           ARIMA Model Summary
Parameter             Estimate         Stnd. Error      t               P-value
----------------------------------------------------------------------------
MA(1)                 -0,229556        0,0931328        -2,46482        0,015173
MA(2)                 0,069806         0,0918814        0,759741        0,448951
Mean                  1,12009          0,528482         2,11945         0,036185
Constant              1,12009
----------------------------------------------------------------------------
Backforecasting: yes
Estimated white noise variance = 26,0423 with 116 degrees of freedom
Estimated white noise standard deviation = 5,10317
```

Diagnosis

- Si una serie está bien identificada y se le ajusta un modelo correcto, los residuos deben carecer de estructura, es decir, deben ser ruido blanco.

- La diagnosis está basada en:
  - FAS y FAP de los residuos. Hay que comprobar que los palos no son significativos.
  - El test de Box-Pierce proporciona información sobre si los primeros palos de la función de autocorrelación de los residuos son cero. Este test indica problemas cuando el p-valor es bajo, por ejemplo menor que 0.05. Cuanto mayor sea, hay más evidencia a favor de que los residuos son ruido blanco.

En la serie ajustada de IBM, el p-valor del estadístico de Box-Pierce (aplicado a los 24 primeros residuos de la FAS) es 0.72, luego no hay evidencia para rechazar la hipótesis nula, es decir, no se rechaza que los residuos sean ruido blanco.

Gráfico de la FAS de los residuos de la serie IBM

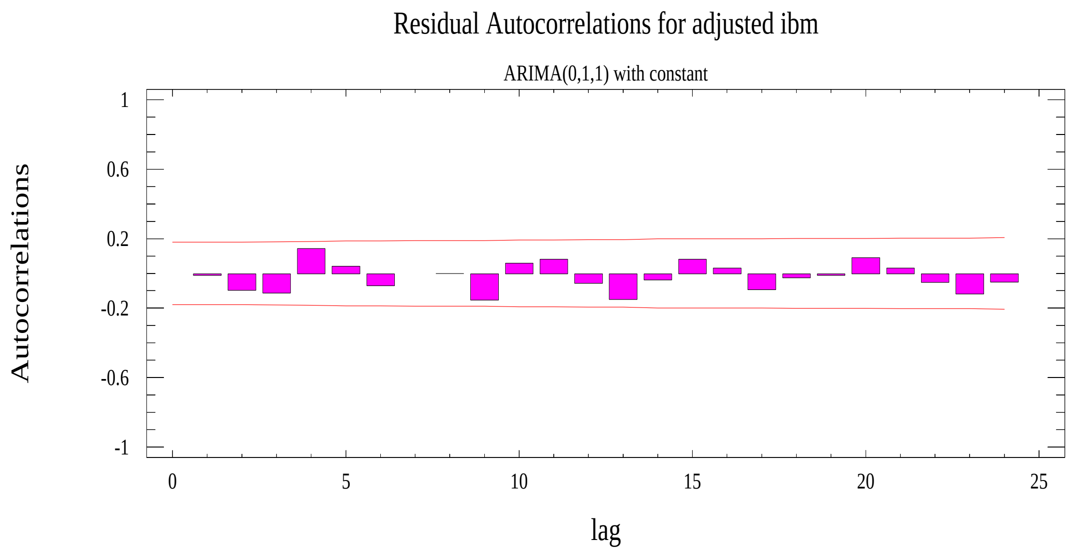
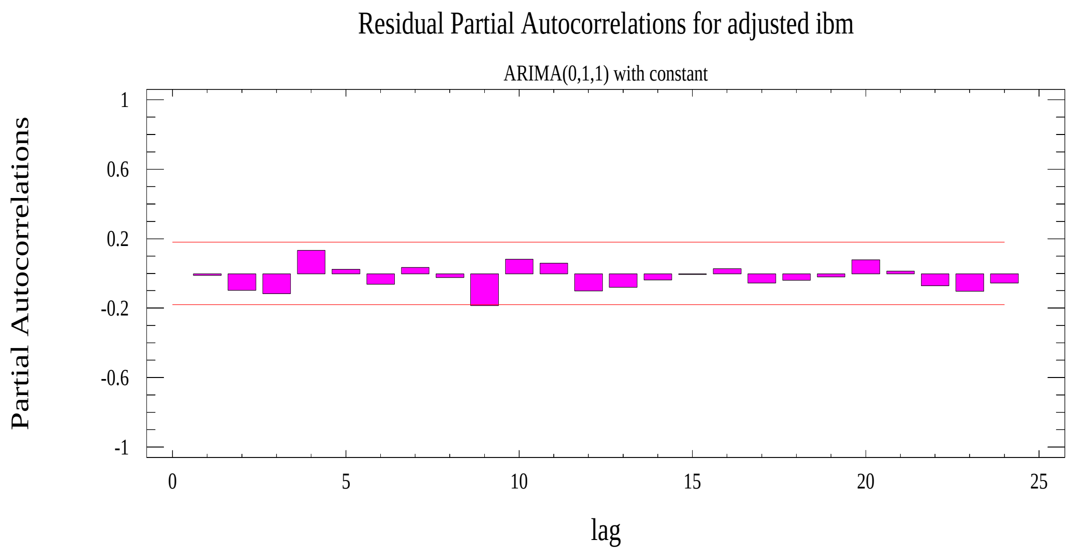


Gráfico de la FAP de los residuos de la serie IBM



Modelos estacionales

- Muchas series con periodicidad menor que la anual tienen estacionalidad.
- La estacionalidad aparece en forma de ciclo, es decir, la serie tiene una pauta que se repite.
- El orden de la estacionalidad indica cada cuántos periodos se repite el ciclo.
- Por ejemplo, en series mensuales, en general, se encontrará estacionalidad de orden 12.

Modelos estacionales

- Las series estacionales tienen dos tipos de estructura de dependencia.
  - Estructura regular. Se ajustan modelos ARIMA(p,d,q).

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

  - Estructura estacional. Se ajustan modelos ARIMA estacionales.
    Por ejemplo, si la serie es mensual

$$z_1 \longrightarrow z_{13} \longrightarrow \cdots \longrightarrow z_{t-12} \longrightarrow z_t \longrightarrow z_{t+12} \longrightarrow \cdots$$

Modelos estacionales

- Las series estacionales pueden tener una relación autorregresiva o de medias móviles.
- Si la serie es mensual, hay que fijarse en los palos de la FAS y de la FAP separados por 12 retardos.
- La identificación se realiza estudiando la FAS y la FAP en los retardos estacionales.
- Cuando una serie tiene parte regular y estacional, puede haber interacción entre ellas, tanto en la FAS como en la FAP.
- Las series estacionales tienen ciclo estacional. La metodología ARIMA permite eliminar el ciclo tomando diferencias de orden estacional: $\Delta_d z_t = z_t - z_{t-d}$.

Gráfico de la serie del IPI de Francia: IPI

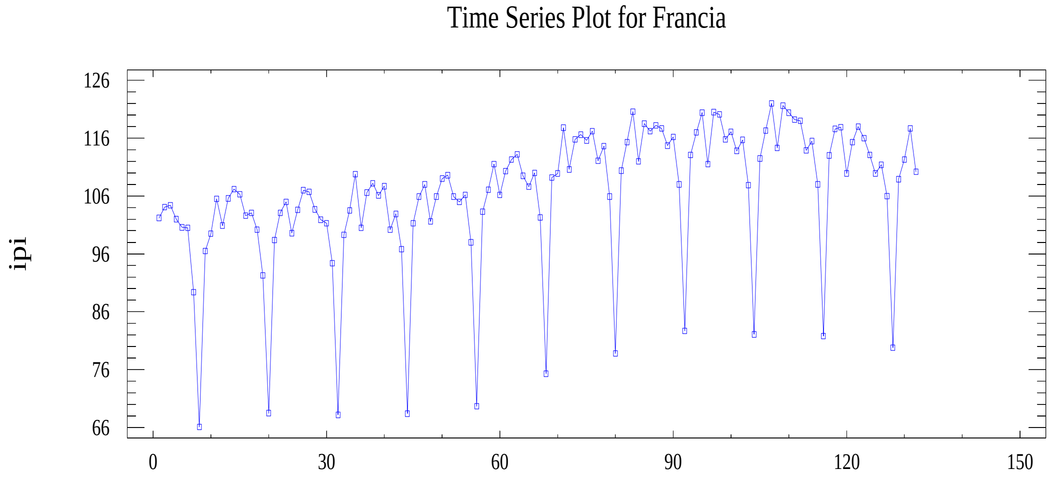


Gráfico de la FAS del IPI de Francia

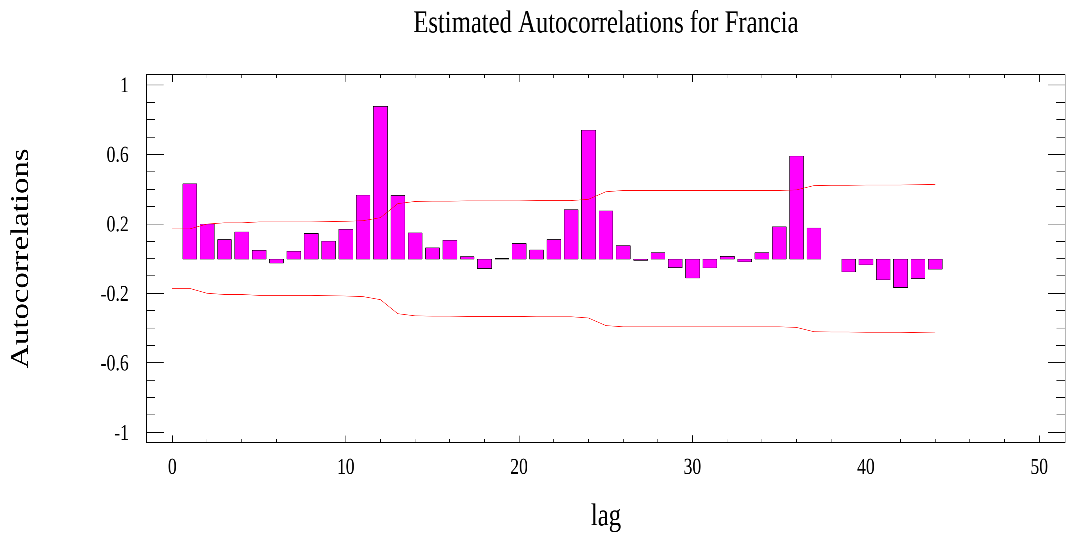


Gráfico de la FAP del IPI de Francia

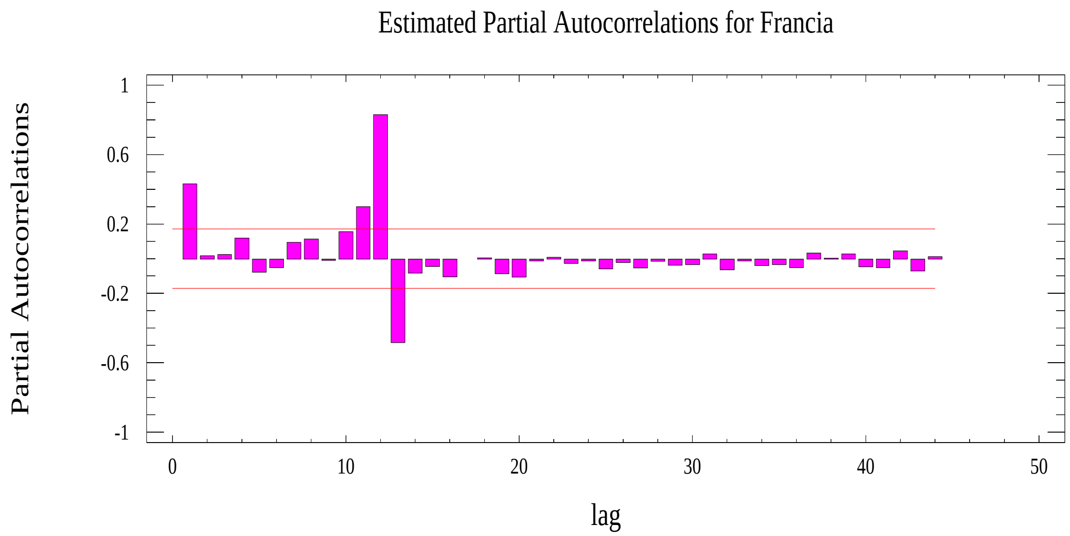


Gráfico del IPI de Francia con una diferencia regular: $\triangle$ IPI

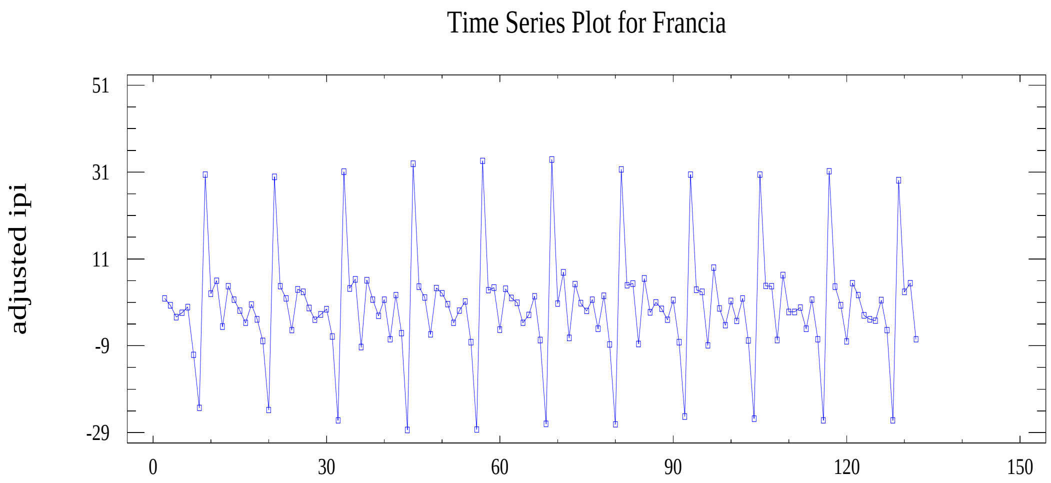


Gráfico de la FAS de $\triangle$ IPI

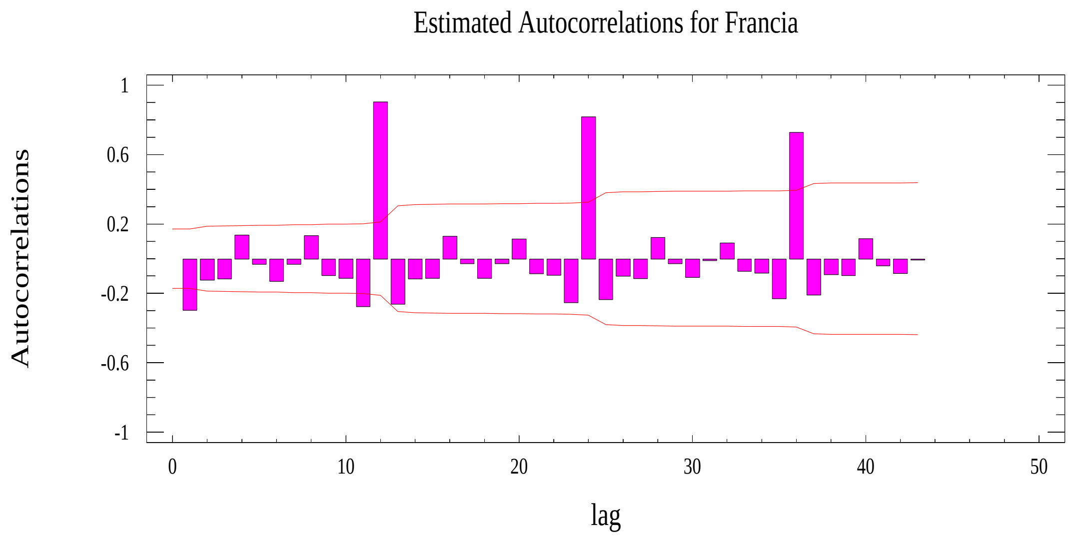


Gráfico de la FAP de $\triangle$ IPI

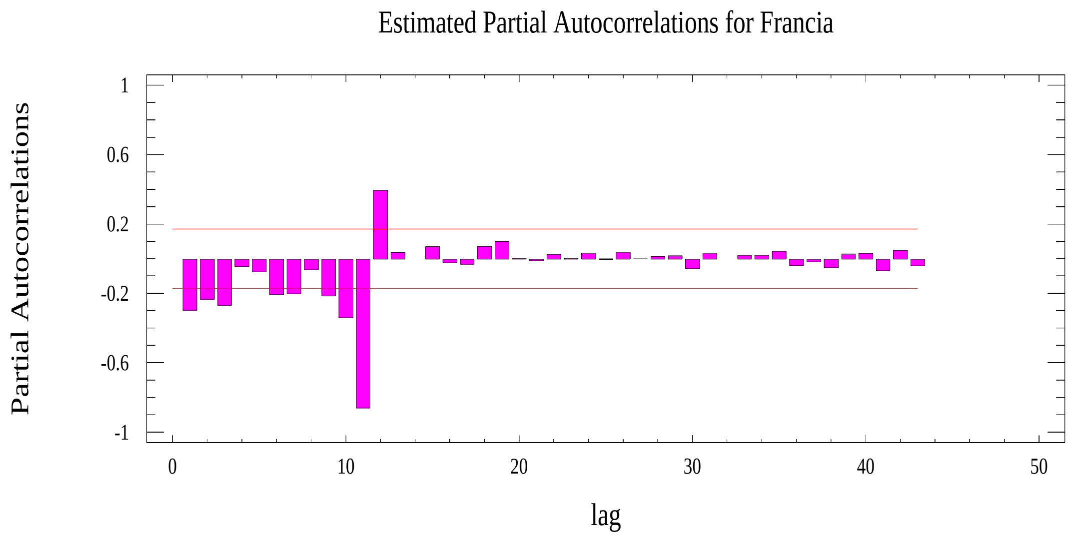


Gráfico del IPI con diferencia estacional: Δ_{12} IPI

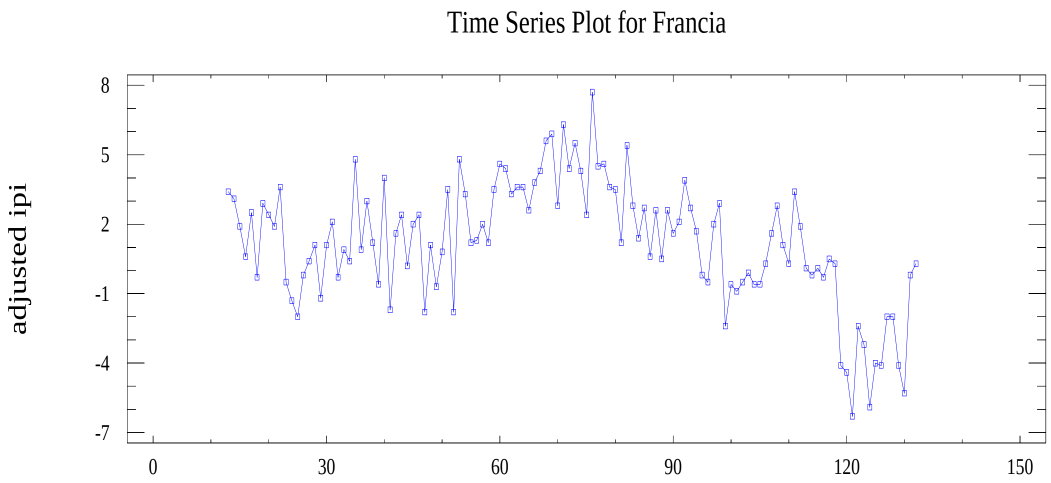


Gráfico de la FAS de Δ_{12} IPI

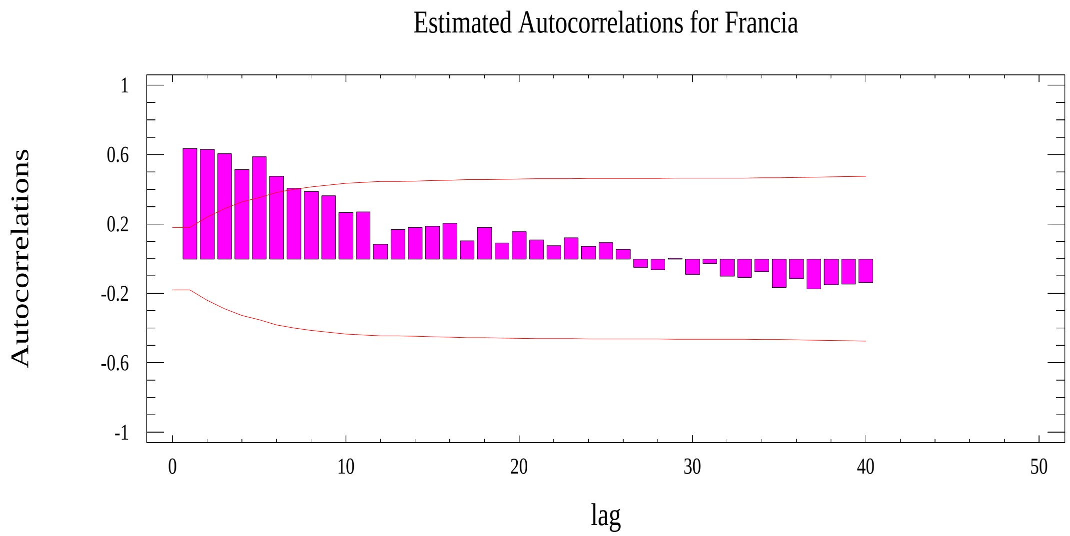


Gráfico de la FAP de Δ_{12} IPI

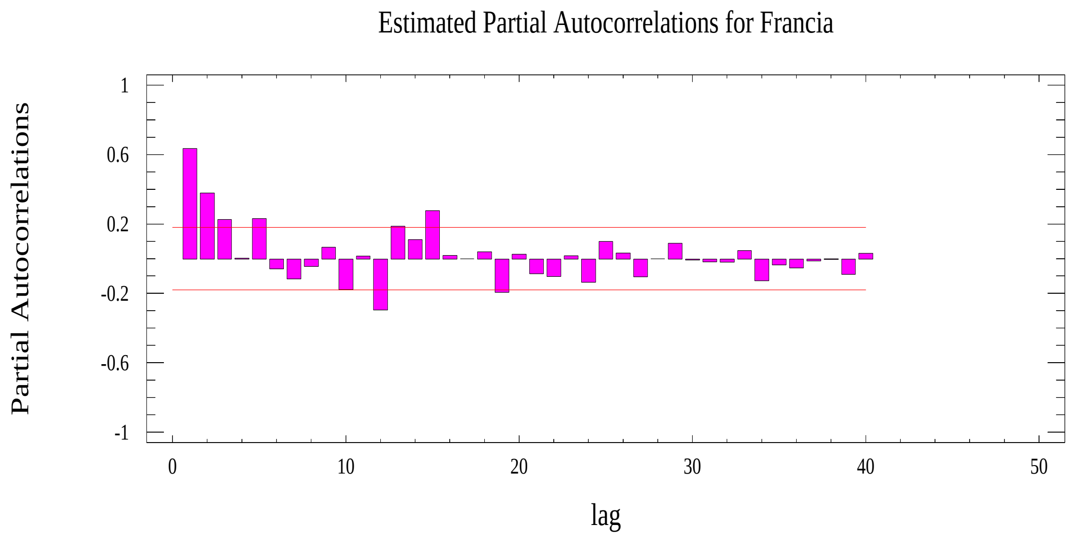


Gráfico de $\Delta\Delta_{12}$ IPI

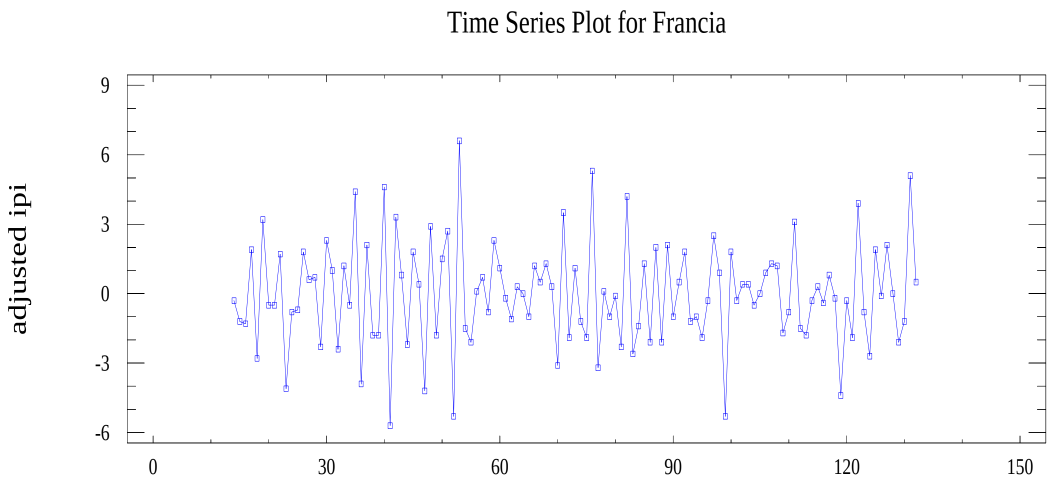


Gráfico de la FAS de $\Delta\Delta_{12}$ IPI

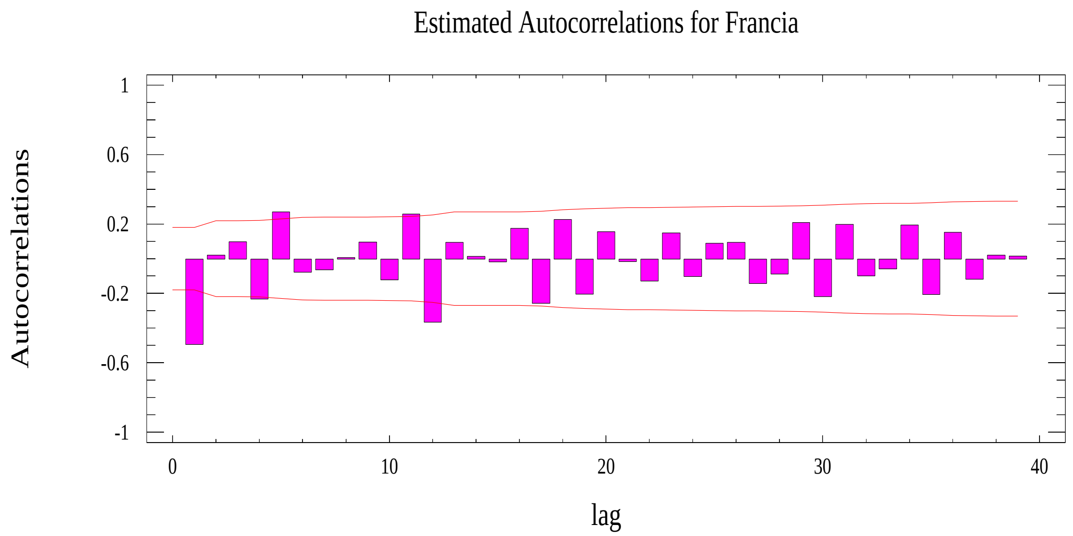
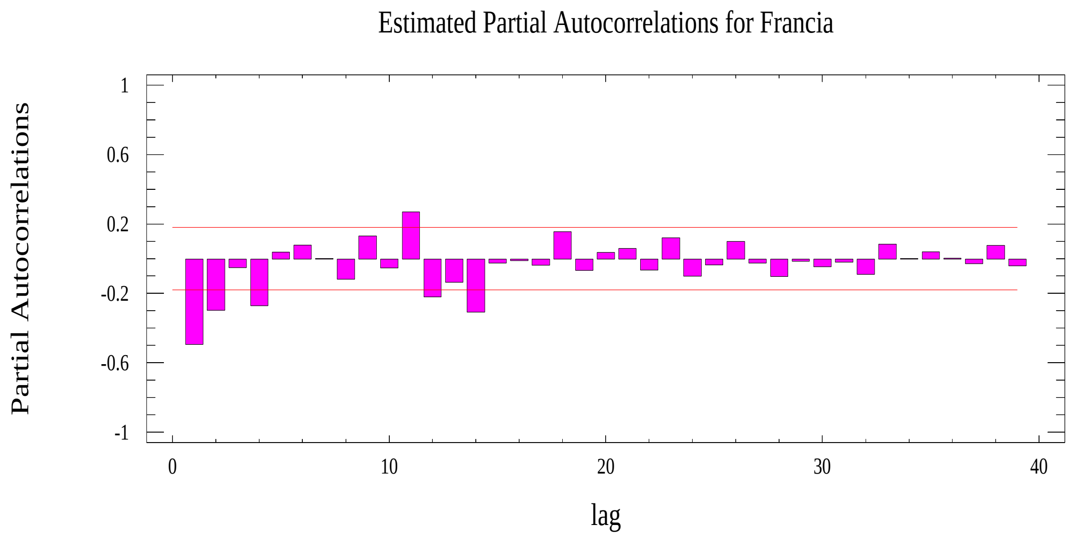


Gráfico de la FAP de $\Delta\Delta_{12}$ IPI



Análisis de la FAS y FAP de la serie $\Delta\Delta_{12}$ IPI

- **Parte regular:** Hay que fijarse en los primeros retardos. Un palo grande en la FAS y un descenso en la FAP. Podría ser un MA(1).
- **Parte estacional:** Los palos estacionales son grandes en el retardo 12 y no se ven palos de orden superior. Podría ser un MA(1)$_{12}$.
- Se ajusta un modelo ARIMA(0,1,1)$\times$(0,1,1)$_{12}$.
- Los parámetros ajustados son significativos.
- La diagnosis es correcta. El test de Box-Pierce da un p-valor de 0.60.

Gráfico de la FAS de los residuos del IPI de Francia

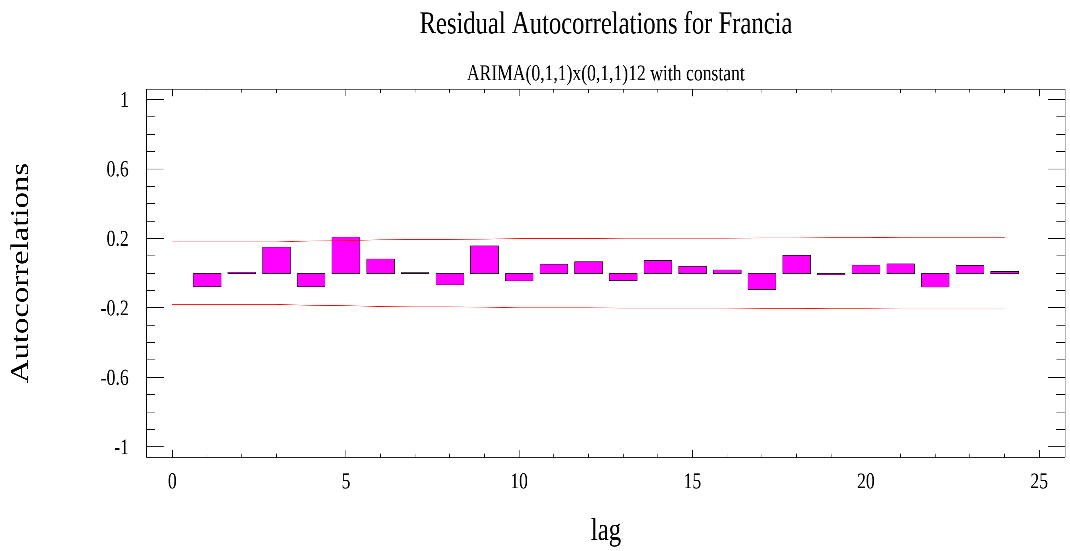
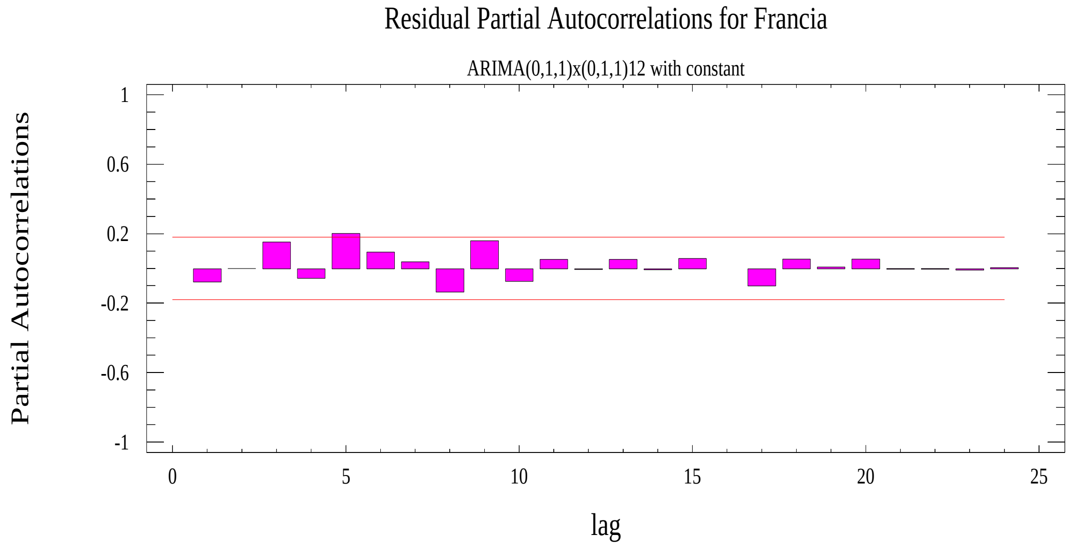


Gráfico de la FAP de los residuos del IPI de Francia



Resumen: Metodología Box-Jenkins

• Etapas en la construcción de un modelo de series temporales (basado en la metodología Box y Jenkins, (1970)):

1. Gráfico de la serie, fas y fap.
2. Identificación de un modelo para los datos (residuos).
3. Estimación de los parámetros.
4. Diagnosis (gráfico de los residuos, fas y fap). ¿Es el modelo adecuado?

$$\begin{cases} \text{Sí} \rightarrow & \text{Predicción.} \\ \text{No} \rightarrow & \text{Ir a 2.} \end{cases}$$

Resumen: Metodología Box-Jenkins

Estacionaridad: Es necesario que las propiedades sean estables a lo largo del tiempo.

La dependencia entre observaciones debe tender a cero a medida que aumentan los retardos.

Procesos integrados: La gran mayoría de los procesos no son estacionarios y su nivel medio varía con el tiempo. Sin embargo, es frecuente que el proceso se convierta en estacionario al diferenciarlo.

La primera diferencia del proceso $\{z_t\}$ es un nuevo proceso $\{w_t\}$ obtenido mediante $w_t = z_t - z_{t-1}$. Análogamente, llamaremos proceso segunda diferencia del original a $y_t = w_t - w_{t-1} = z_t - 2z_{t-1} + z_{t-2}$.

Proceso integrado de orden h si al diferenciarlo h veces se obtiene un proceso estacionario.

Resumen: Metodología Box-Jenkins

• Ejemplo: Un proceso que sea suma de una tendencia polinómica y un proceso estacionario será un proceso integrado: $z_t = a + bt + u_t$,

$$w_t = z_t - z_{t-1} = a + bt + u_t - (a + b(t-1) + u_{t-1}) = b + u_t - u_{t-1} = b + \alpha_t$$

donde α_t es otro proceso estacionario.

AR(p): $z_t = \phi_1 z_{t-1} + \ldots + \phi_p z_{t-p} + a_t$.

• Para facilitar el manejo de estos procesos se define el operador de retardo B (equivalente a Δ):

$Bz_t = z_{t-1} \qquad B^k z_t = B \cdot Bz_t = z_{t-k}$

• Con esta notación, la ecuación de un AR(p) es:

$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) z_t = a_t$, y llamando $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$, entonces se obtiene la ecuación general de un proceso autorregresivo: $\phi_p(B) z_t = a_t$.

Resumen: Metodología Box-Jenkins

MA(q): $z_t = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$

Los procesos AR son casos particulares de los procesos MA.

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

$(1 - \phi_1 B) z_t = a_t \quad |\phi_1| < 1$

$(1 - \phi_1 B)^{-1} = 1 + \phi_1 B + \phi_1^2 B^2 + \ldots$

ARMA(p,q):

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

En forma compacta: $\phi_p(B) z_t = \theta_q(B) a_t$

• FAS: q-p+1 valores iniciales con cualquier estructura. Decrecimiento a partir del coeficiente $q - p$ como una mezcla de exponenciales y sinusoides determinada por la parte autorregresiva.

• FAP: Estructura similar.

Resumen: Metodología Box-Jenkins

Procesos no estacionarios. Ejemplo: $z_t = \phi_1 z_{t-1} + a_t$

Si $|\phi_1| > 1$ el proceso es explosivo. Si $|\phi_1| = 1$ es un proceso integrado de orden 1, ya que su primera diferencia es un proceso estacionario. Operador diferencia $\nabla = 1 - B$.

ARIMA(p,d,q):

$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)(1 - B)^d z_t = (1 - \theta_1 B - \ldots - \theta_q B)a_t$.

Equivalentemente: $\phi_p(B)\nabla^d z_t = \theta_q(B)a_t$.

Resumen: Metodología Box-Jenkins

- **Estacionalidad:** Pauta regular de comportamiento periódico en la serie.

$\nabla_s = 1 - B^s$ convierte en general un proceso estacional en estacionario.

ARIMA(p,d,q)$\times$ARIMA(P,D,Q)$_s$:

$$\Phi_P(B)\phi_p(B)\nabla^d\nabla_s^D z_t = \Theta_Q(B)\theta_q(B)a_t,$$

donde

$$\Phi_p(B) = (1 - \Phi_s B^s - \Phi_{2s} B^{2s} - \cdots - \Phi_{Ps} B^{Ps}),$$
$$\Theta_q(B) = (1 - \Theta_s B^s - \cdots - \theta_{Qs} B^{Qs}).$$

Resumen: Metodología Box-Jenkins

FAS y FAP:

- Retardos bajos: Se observará únicamente la parte regular.
- Retardos estacionales: Se observará básicamente la parte estacional.
- Alrededor de los retardos estacionales observaremos la interacción entre la parte regular y estacional.

Resumen: Metodología Box-Jenkins

Identificación de la estructura no estacionaria

- Transformar la serie para que tenga varianza constante.
- Determinar el número de diferencias para que la media sea constante.
- Si es estacional con periodo s, aplicar una diferencia estacional para convertirla en estacionaria.

Resumen: Metodología Box-Jenkins

- **Diagnosis:** Residuos. Contrates.
- **Predicción con modelos ARIMA:** último tema.

Nota: El modelo ARIMA estacional más frecuente es el llamado modelo de pasajeros de avión o líneas aéreas:

$$\nabla\nabla_{12}\, z_t = (1-\theta B)(1-\Theta B^{12})a_t$$

Predicción

- Una vez ajustado el modelo es posible utilizarlo para predecir valores futuros.

- Esto implica obtener la ecuación siguiente:

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

- Sustituyendo los valores se obtienen los valores futuros: $\hat{z}_1$, $\hat{z}_2$,...

- Las predicciones siempre tienen que ir acompañadas de los correspondientes intervalos de predicción (complicados de calcular).

Predicción

- Medidas del comportamiento en la predicción:

$$\text{MSE} = \sum_t (z_t - \hat{z}_t)^2$$

$$\text{MAE} = \sum_t |z_t - \hat{z}_t|$$

- Estas medidas se obtienen de la etapa de estimación y también se pueden obtener de datos de validación (externos a la estimación).

Previsiones para la serie del IPI con el modelo ajustado