Part III

Modelos Box-Jenkins

Bibliografía

Bibliografía

- Brockwell, P.J. y Davis, R.A. (2002). *Introduction to Time Series and Forecasting.* 2ª edición. Springer.
- Cowpertwait, P.S.P. y Metcalfe, A.V. (2009). *Introductory Time Series with R.* Springer.
- González, M. y del Puerto, I.M. (2009). *Series Temporales.* Colección manuales uex-60.

Bibliografía

Bibliografía

- Makridakis, S., Wheelwright, S.C. y Hyndman, R.J. (1998).
  Forecasting. Methods and Applications. 3ª edición. Wiley.
- Peña, D. (2005).
  Análisis de Series Temporales. Alianza Editorial.
- Shumway, R.H. y Stoffer, D.S. (2006).
  Time Series Analysis and Its Applications. With R Examples. 2ª edición. Springer.

Introducción

Comenzamos recordando la notación general:

- $x_1, x_2, \ldots, x_T$: serie de tiempo.
- $\{X_t\}_t$: proceso generador de la serie de tiempo.
- $\{a_t\}_t$: proceso de ruido blanco (innovaciones).

Además, supondremos que:

- a_t es independiente de $X_{t-1}, X_{t-2}, \ldots$

Esto implica que el conocimiento de los valores de $X_{t-1}, X_{t-2}, \ldots$ no aporta información acerca de a_t.

Modelos Box-Jenkins

Introducción

El primer objetivo de este tema es la construcción de un modelo estocástico sencillo que, de forma razonable, haya podido generar a la serie de tiempo de que disponemos.

Entonces, basándonos en dicho modelo, podremos:

- Comprender la dinámica de la serie de tiempo.
- Efectuar predicciones de valores futuros de la serie de tiempo.

Modelos Box-Jenkins

Introducción

La planificación del presente tema puede resumirse en las siguientes etapas:

1. Presentación de una clase de modelos estocásticos paramétricos pero flexibles (han demostrado su utilidad como posibles generadores de series de tiempo reales).
2. Exposición de técnicas que nos permitan seleccionar alguno de dichos modelos como generador de nuestra serie de tiempo.
3. Construcción de estimadores para los parámetros del modelo.
4. Predicción de futuros valores de la serie en base al modelo estimado.

Procesos autorregresivos: AR

Imaginemos que deseamos construir un modelo para la cantidad de agua (X_t) que hay al final del mes t en un embalse. Puesto que dicha cantidad es aleatoria, el modelo ha de ser un modelo estocástico. Haremos las siguientes suposiciones:

- Durante el mes t llega al embalse una cantidad de agua $c + a_t$, siendo c la cantidad media de agua que llega y a_t una v.a. de media cero y varianza constante que hace que la entrada de agua varíe de un mes a otro.
- Cada mes se gasta una proporción fija $1 - \phi_1$ de las existencias iniciales.

Se tiene que el modelo buscado es:

$$X_t = c + \phi_1 X_{t-1} + a_t.$$

Procesos autorregresivos: AR

Un proceso estacionario $\{X_t\}_t$ que admite la representación

$$X_t = c + \phi_1 X_{t-1} + a_t,$$

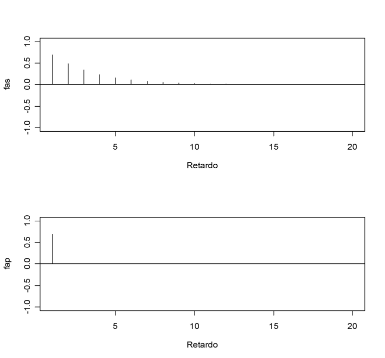
donde c y ϕ_1 son constantes, se conoce como un "proceso autorregresivo de orden 1", y se denota por AR(1).

Se verifica que:

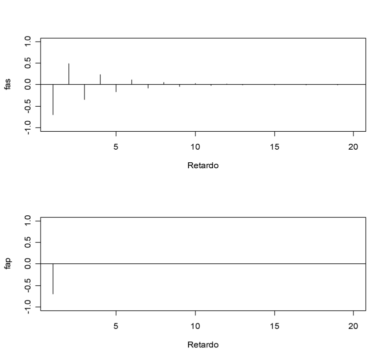
- El proceso AR(1) explica el valor actual (X_t) como una función lineal de 1 valor pasado (X_{t-1}).
- El proceso AR(1) siempre es invertible.

Ejemplos de la fas y la fap de procesos AR(1)

AR(1): $\phi_1 > 0$



AR(1): $\phi_1 < 0$



Series de Tiempo

Germán Aneiros Pérez

Introducción

Procesos ARMA: Construcción e identificación

Procesos ARIMA: Construcción e identificación

Procesos ARIMA estacionales: Construcción e identificación

Estimación

Diagnosis

Procesos autorregresivos: AR

Un proceso estacionario $\{X_t\}_t$ que admite la representación

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t,$$

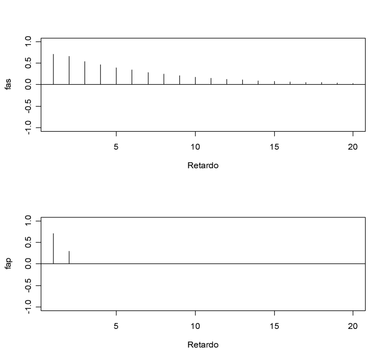
donde c, ϕ_1 y ϕ_2 son constantes, se conoce como un "proceso autorregresivo de orden 2", y se denota por AR(2).

Se verifica que:

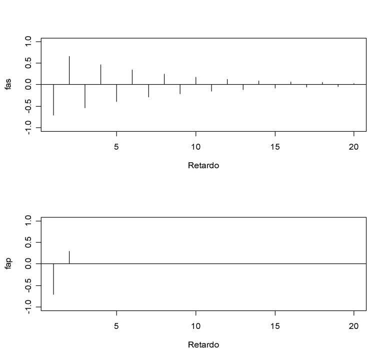
- El proceso AR(2) explica el valor actual (X_t) como una función lineal de 2 valores pasados $(X_{t-1}$ y $X_{t-2})$.
- El proceso AR(2) siempre es invertible.

Ejemplos de la fas y la fap de procesos AR(2)

AR(2): $\phi_1 > 0, \phi_2 > 0$

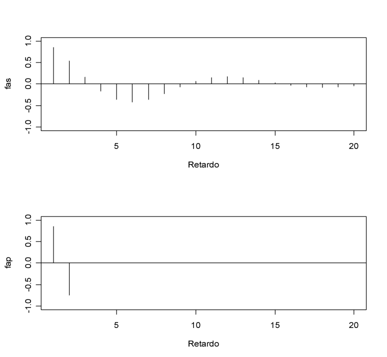


AR(2): $\phi_1 < 0, \phi_2 > 0$

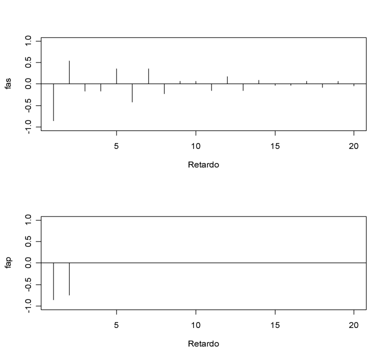


Ejemplos de la fas y la fap de procesos AR(2)

AR(2): $\phi_1 > 0, \phi_2 < 0$



AR(2): $\phi_1 < 0, \phi_2 < 0$



Modelos Box-Jenkins

Procesos autorregresivos: AR

Un proceso estacionario $\{X_t\}_t$ que admite la representación

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + a_t,$$

donde $c, \phi_1, \ldots, \phi_p$ son constantes, se conoce como un "proceso autorregresivo de orden p" (AR(p)).

Se verifica que:

- Estacionario $\Leftrightarrow 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p \neq 0$ $\forall z$ con $|z| = 1$.
- Causal $\Leftrightarrow 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p \neq 0$ $\forall z$ con $|z| \leq 1$.
- Siempre es invertible.
- La fap se anula para todo retardo mayor que p.

Procesos de medias móviles: MA

Un proceso $\{X_t\}_t$ que admite la representación

$$X_t = c + a_t + \theta_1 a_{t-1},$$

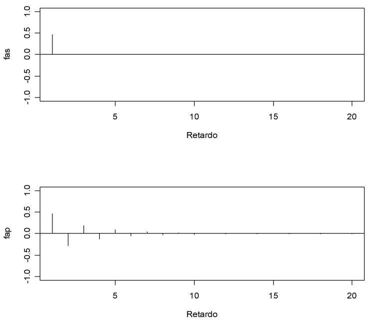
donde c y θ_1 son constantes, se conoce como un "proceso de medias móviles de orden 1", y se denota por MA(1).

Se verifica que:

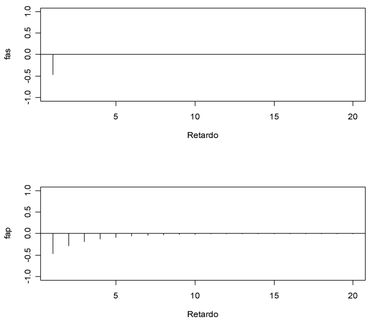
- El proceso MA(1) explica el valor actual (X_t) como una función lineal de 1 valor pasado de un proceso de ruido blanco (a_{t-1}).
- El proceso MA(1) siempre es estacionario y causal.

Ejemplos de la fas y la fap de procesos MA(1)

MA(1): $\theta_1 > 0$



MA(1): $\theta_1 < 0$



Procesos de medias móviles: MA

Un proceso $\{X_t\}_t$ que admite la representación

$$X_t = c + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2},$$

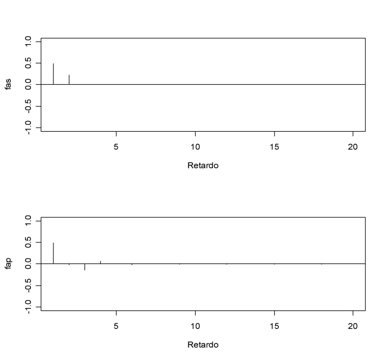
donde c, θ_1 y θ_2 son constantes, se conoce como un "proceso de medias móviles de orden 2", y se denota por MA(2).

Se verifica que:

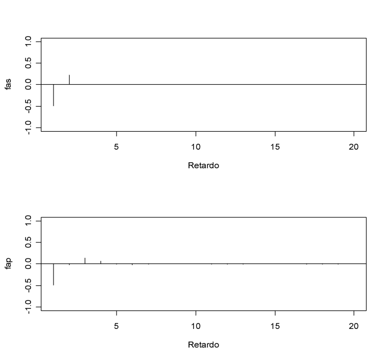
- El proceso MA(2) explica el valor actual (X_t) como una función lineal de 2 valores pasados de un proceso de ruido blanco (a_{t-1} y a_{t-2}).
- El proceso MA(2) siempre es estacionario y causal.

Ejemplos de la fas y la fap de procesos MA(2)

MA(2): $\theta_1 > 0, \theta_2 > 0$

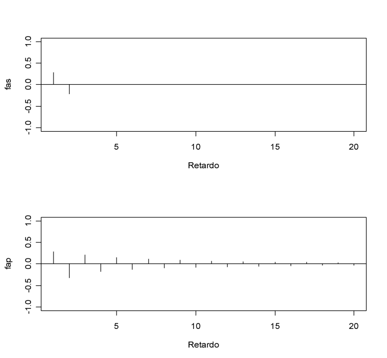


MA(2): $\theta_1 < 0, \theta_2 > 0$

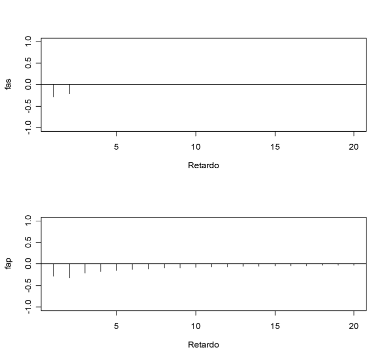


Ejemplos de la fas y la fap de procesos MA(2)

MA(2): $\theta_1 > 0, \theta_2 < 0$



MA(2): $\theta_1 < 0, \theta_2 < 0$



Procesos de medias móviles: MA

Un proceso $\{X_t\}_t$ que admite la representación

$$X_t = c + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q},$$

donde $c, \theta_1, \ldots, \theta_q$ son constantes, se conoce como un "proceso de medias móviles de orden q" (MA(q)).

Se verifica que:

- Siempre es estacionario y causal.
- Invertible $\Leftrightarrow 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_p z^q \neq 0$ $\forall z$ con $|z| \leq 1$.
- La fas se anula para todo retardo mayor que q.

Procesos ARMA

La introducción en un mismo proceso estacionario de estructura autorregresiva (AR) y de medias móviles (MA) da lugar a los procesos ARMA.

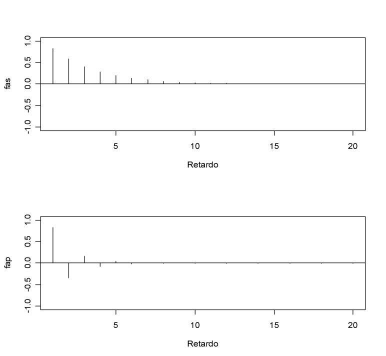
Así, por ejemplo, se tienen los procesos:

- ARMA(1,1): $X_t = c + \phi_1 X_{t-1} + a_t + \theta_1 a_{t-1}$.
- ARMA(2,1): $X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t + \theta_1 a_{t-1}$.
- ARMA(1,2): $X_t = c + \phi_1 X_{t-1} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$.
- ARMA(2,2):

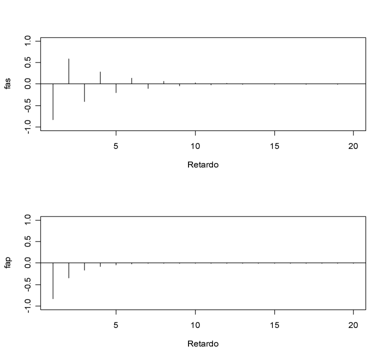
$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}.$$

Ejemplos de la fas y la fap de procesos ARMA(1,1)

ARMA(1,1): $\phi_1 > 0, \theta_1 > 0$



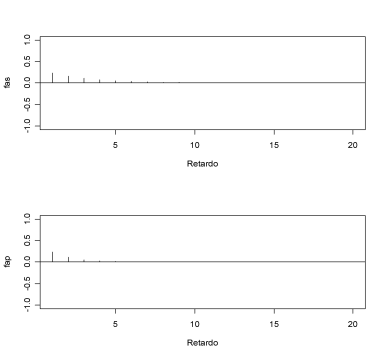
ARMA(1,1): $\phi_1 < 0, \theta_1 < 0$



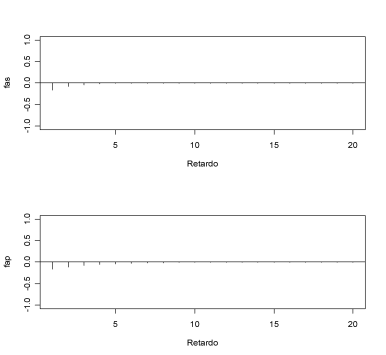
Series de Tiempo
Germán Aneiros Pérez
Introducción
Procesos ARMA: Construcción e identificación
Procesos ARIMA: Construcción e identificación
Procesos ARIMA estacionales: Construcción e identificación
Estimación
Diagnosis

Ejemplos de la fas y la fap de procesos ARMA(1,1)

ARMA(1,1): $\phi_1 > 0, \theta_1 < 0$

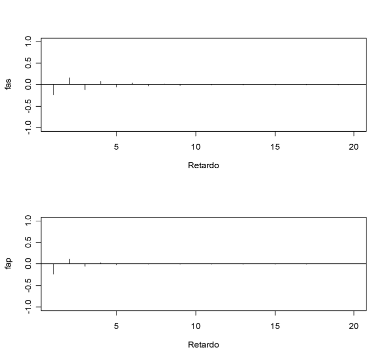


ARMA(1,1): $\phi_1 > 0, \theta_1 < 0$

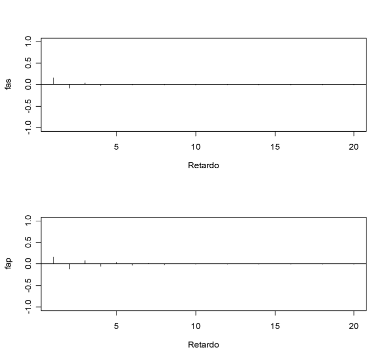


Ejemplos de la fas y la fap de procesos ARMA(1,1)

ARMA(1,1): $\phi_1 < 0, \theta_1 > 0$



ARMA(1,1): $\phi_1 < 0, \theta_1 > 0$



Procesos ARMA

Un proceso estacionario $\{X_t\}_t$ que admite la representación

$$\begin{aligned} X_t = \; & c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} \\ & + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}, \end{aligned}$$

donde $c, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ son constantes, se conoce como un proceso ARMA(p,q).

Se verifica que:

- $c = \mu(1 - \phi_1 - \cdots - \phi_p)$.
- ARMA(p,0) $\Leftrightarrow$ AR(p).
- ARMA(0,q) $\Leftrightarrow$ MA(q).

Procesos ARMA

- Estacionario $\Leftrightarrow 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p \neq 0$
$$\forall z \text{ con } |z| = 1.$$
- Causal $\Leftrightarrow 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p \neq 0 \ \forall z \text{ con } |z| \leq 1.$
- Invertible $\Leftrightarrow 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_p z^q \neq 0$
$$\forall z \text{ con } |z| \leq 1.$$
- La clase de procesos ARMA es muy flexible:

  Si $\{Y_t\}$ es un proceso estacionario tal que $\gamma_{Y,h} \to 0$ cuando $h \to \infty$, entonces, dado cualquier número entero $k > 0$, existe un proceso ARMA $\{X_t\}$ tal que $\gamma_{X,h} = \gamma_{Y,h}$ $\forall h = 0, 1, \ldots, k$.

Procesos ARMA

La ecuación que define al proceso ARMA(p,q)

$$\begin{aligned} X_t = \quad & c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} \\ & + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}, \end{aligned}$$

se puede escribir en la forma compacta

$$\phi(B) X_t = c + \theta(B) a_t,$$

donde

$$\begin{aligned} \phi(B) &= \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right), \\ \theta(B) &= \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) \end{aligned}$$

y B denota al operador retardo, definido por $BX_t = X_{t-1}$.

Modelos Box-Jenkins

Procesos ARMA

	fas	fap
AR(p)	Muchos coeficientes no nulos*	Se anula para todo retardo mayor que p
MA(q)	Se anula para todo retardo mayor que q	Muchos coeficientes no nulos*
ARMA(p,q)	Muchos coeficientes no nulos*	Muchos coeficientes no nulos*

*A partir de los primeros retardos, convergen muy rápidamente a cero, como suma de funciones exponenciales y/o sinusoidales.

Procesos ARMA

La clase de procesos ARMA que acabamos de estudiar:

- Es una familia de procesos estacionarios.
- Es muy flexible: puede modelizar gran variedad de series generadas por procesos estacionarios.

Dada una "serie de tiempo estacionaria", trataremos de modelizarla a través de un proceso ARMA(p,q).

Las primeras cuestiones a resolver son:

1. ¿Cómo asesorarnos acerca de si una serie de tiempo ha sido generada o no por un proceso estacionario (ARMA)?
2. Si efectivamente ha sido generada por un proceso ARMA, ¿cómo identificar los valores de p y q?

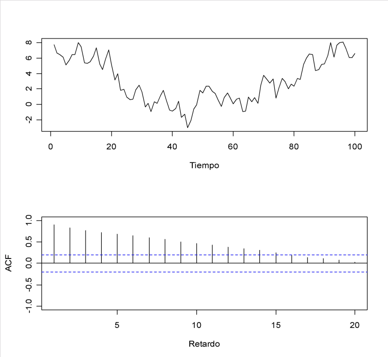
Modelos Box-Jenkins

La serie de tiempo. . . ,
¿ha sido generada por un proceso estacionario (ARMA)?

Serie estacionaria (ARMA)

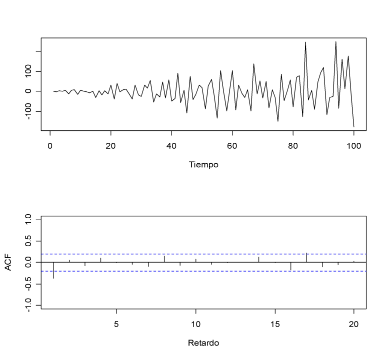
- El gráfico de la serie frente al tiempo debe mostrar:
  - Nivel constante.
  - Variabilidad constante.
- La fas muestral $\widehat{\rho}_k$ debe converger a cero muy rápidamente a medida que el retardo k crece.

Nivel no constante (tendencia)

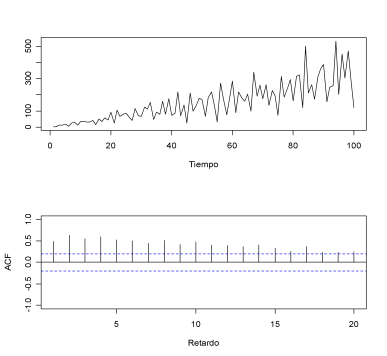


Modelos Box-Jenkins

Varianza no constante

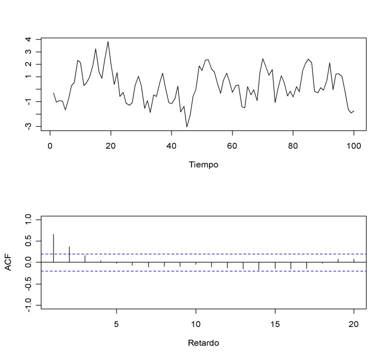


Nivel y varianza no constantes

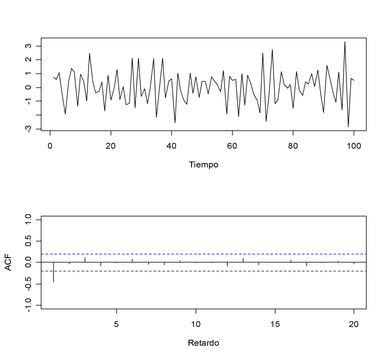


Modelos Box-Jenkins

Serie estacionaria



Serie estacionaria



Procesos ARMA

Dada una serie de tiempo generada por un proceso ARMA. . . , ¿cuáles son los órdenes p y q correspondientes? (Identificación)

Para responder a esta pregunta nos basaremos, en un principio, en la información que nos suministran la fas y la fap muestrales ($\widehat{\rho}_k$ y $\widehat{\alpha}_k$, respectivamente).

Tanto $\widehat{\rho}_k$ como $\widehat{\alpha}_k$ dependen de la serie de tiempo observada, y sus valores cambiarán con ella (son aleatorios).

Por ello, para obtener información a partir de $\widehat{\rho}_k$ y $\widehat{\alpha}_k$ necesitamos conocer su distribución muestral.

Procesos ARMA

Distribución muestral de $\widehat{\rho}_k$ y de $\widehat{\alpha}_k$. Inferencia

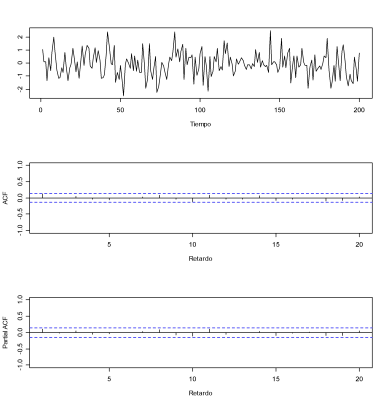
Si el tamaño de la serie (T) es "grande" se verifica que:

$$\text{Ruido blanco i.i.d.} \Rightarrow \begin{cases} \widehat{\rho}_k \approx N\left(0, \dfrac{1}{\sqrt{T}}\right), & \forall\ k \\ \widehat{\alpha}_k \approx N\left(0, \dfrac{1}{\sqrt{T}}\right), & \forall\ k \end{cases}$$

Conclusión: Si la serie ha sido generada por un proceso de ruido blanco i.i.d., debería cumplirse para cada $k = 1, 2, \ldots$ (con un nivel de significación individual aproximado del 5%) que:

$$\widehat{\rho}_k \in \left(-\frac{1.96}{\sqrt{T}}, \frac{1.96}{\sqrt{T}}\right) \text{ y } \widehat{\alpha}_k \in \left(-\frac{1.96}{\sqrt{T}}, \frac{1.96}{\sqrt{T}}\right)$$

Serie, fas y fap



Conclusión

Los gráficos de la izquierda sugieren que la serie:

1. Es estacionaria.
2. Ha sido generada por un proceso de ruido blanco.

Procesos ARMA

Distribución muestral de $\widehat{\alpha}_k$. Inferencia

Bajo condiciones generales (incluyendo T "grande") se verifica que:

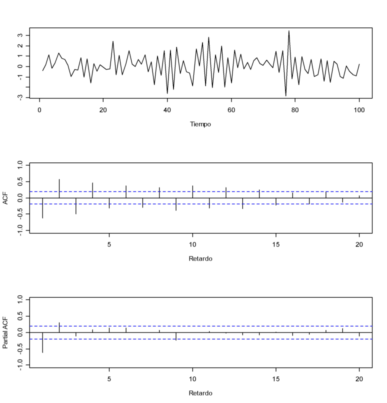
$$\mathrm{AR(p)} \Rightarrow \widehat{\alpha}_k \approx N\left(0, \frac{1}{\sqrt{T}}\right), \ \forall\ k > p.$$

Conclusión: Si la serie ha sido generada por un proceso AR(p), debería cumplirse para cada $k > p$ (con un nivel de significación individual aproximado del 5%) que:

$$\widehat{\alpha}_k \in \left(-\frac{1.96}{\sqrt{T}}, \frac{1.96}{\sqrt{T}}\right)$$

Modelos Box-Jenkins

Serie, fas y fap



Conclusión

Los gráficos de la izquierda sugieren que la serie:

1. Es estacionaria.
2. Ha sido generada por un proceso AR(2).

Procesos ARMA

Distribución muestral de $\widehat{\rho}_k$. Inferencia

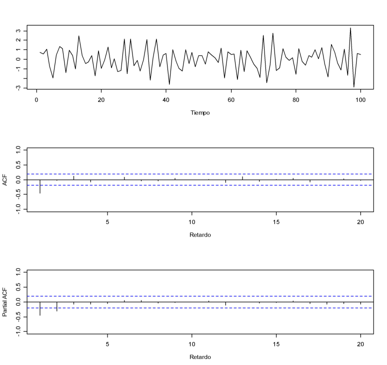
Bajo condiciones generales (incluyendo T "grande") se verifica que:

$$\text{MA(q)} \Rightarrow \widehat{\rho}_k \approx N\left(0, \frac{\sqrt{1+2\left(\rho_1^2+\cdots+\rho_q^2\right)}}{\sqrt{T}}\right), \ \forall\ k>q.$$

Conclusión: Si la serie ha sido generada por un proceso MA(q), debería cumplirse para cada $k>q$ (con un nivel de significación individual aproximado del 5%) que:

$$\widehat{\rho}_k \in \left(\pm 1.96\sqrt{\frac{1+2\left(\widehat{\rho}_1^2+\cdots+\widehat{\rho}_q^2\right)}{T}}\right)$$

Serie, fas y fap



Conclusión

Los gráficos de la izquierda sugieren que la serie:

1. Es estacionaria.
2. Ha sido generada por un proceso MA(1) o por un AR(2).

Procesos ARMA

La identificación de los órdenes p y q de un proceso ARMA mixto ($p \neq 0 \neq q$) a partir del estudio de la fas ($\widehat{\rho}_k$) y la fap ($\widehat{\alpha}_k$) muestrales es una tarea muy complicada.

De momento, nos limitamos a identificar procesos AR o MA a partir de $\widehat{\rho}_k$ y $\widehat{\alpha}_k$.

Más tarde, veremos un método (semi-automático) para identificar procesos ARMA generales.