Capítulo 5

Procesos Integrados

Un proceso no estacionario puede no ser estable en la media, en la varianza o en las autocorrelaciones. Por ejemplo, las series 3, 5-13, 19, 29-31, 35-37, y 39 del Capítulo 0 no son estables en la media.

Procesos integrados: Procesos no estacionarios comunes son los procesos integrados, que se convierten en estacionarios simplemente tomando diferencias.

- Si (Y_t) no es estacionario pero la serie (Z_t) de las primeras diferencias

$$Z_t = \nabla Y_t = Y_t - Y_{t-1},$$

  sí lo es, entonces (Y_t) es un proceso *"integrado de orden uno"*, $I(1)$.

- Si (∇Y_t) no es estacionario, pero la serie (Z_t) de las segundas diferencias

$$Z_t = \nabla^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-1},$$

  sí lo es, entonces (Y_t) es un proceso *"integrado de orden dos"*, $I(2)$.

- Un proceso es *"integrado de orden d"*, $I(d)$, $d \geq 0$, si al diferenciarlo d veces (y no menos) resulta un proceso estacionario.

- Un proceso estacionario es integrado de orden 0 y se denota por $I(0)$.

El término *integrado* indica que el proceso es suma de los elementos hasta t de un proceso estacionario. Si (Y_t) es integrado de orden d, llamando $Z_t = \nabla^d Y_t$ al proceso estacionario, entonces (Y_t) es suma (integración) de (Z_t). Por ejemplo, sea el proceso de primeras diferencias $Z_t = \nabla Y_t = Y_t - Y_{t-1}$.

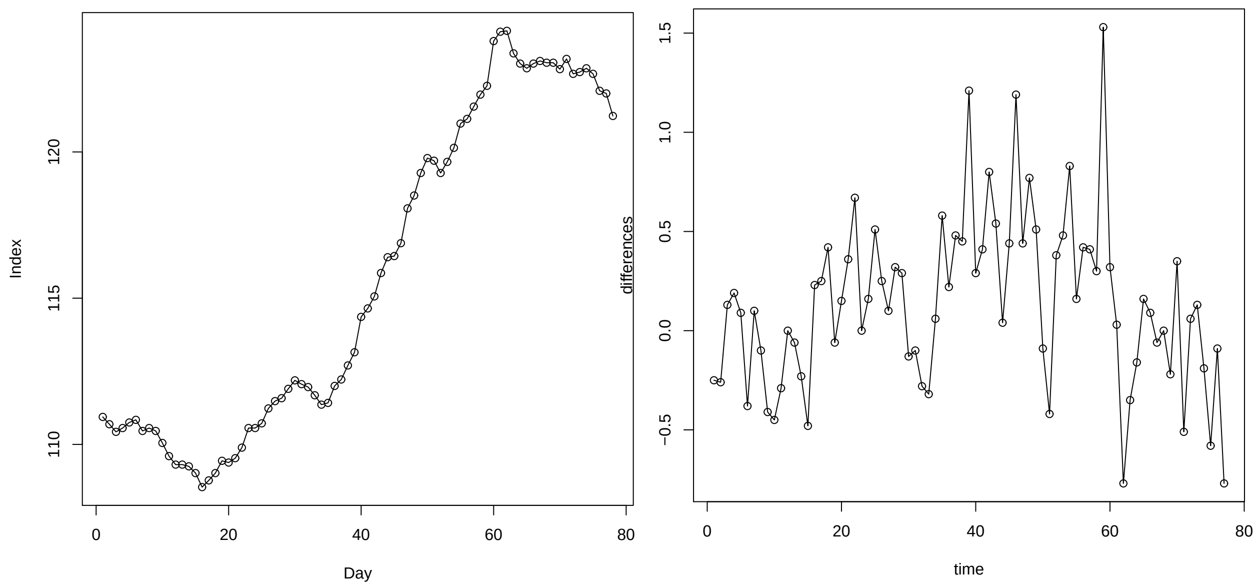
Despejamos Y_t en función de los valores de Z_t:

$$\begin{aligned} Y_t &= \ldots \\ &= \ldots \\ \\ &= \ldots \end{aligned}$$

En la práctica se suelen considerar procesos integrados de orden $d \leq 3$.

Figura 5.1: Índice de Dow Jones, 28.08.–18.12.1972, y sus primeras diferencias

Índice Crecimientos del índice



5.1. Procesos integrados de orden 1, $I(1)$

5.1.1. El paseo aleatorio con deriva

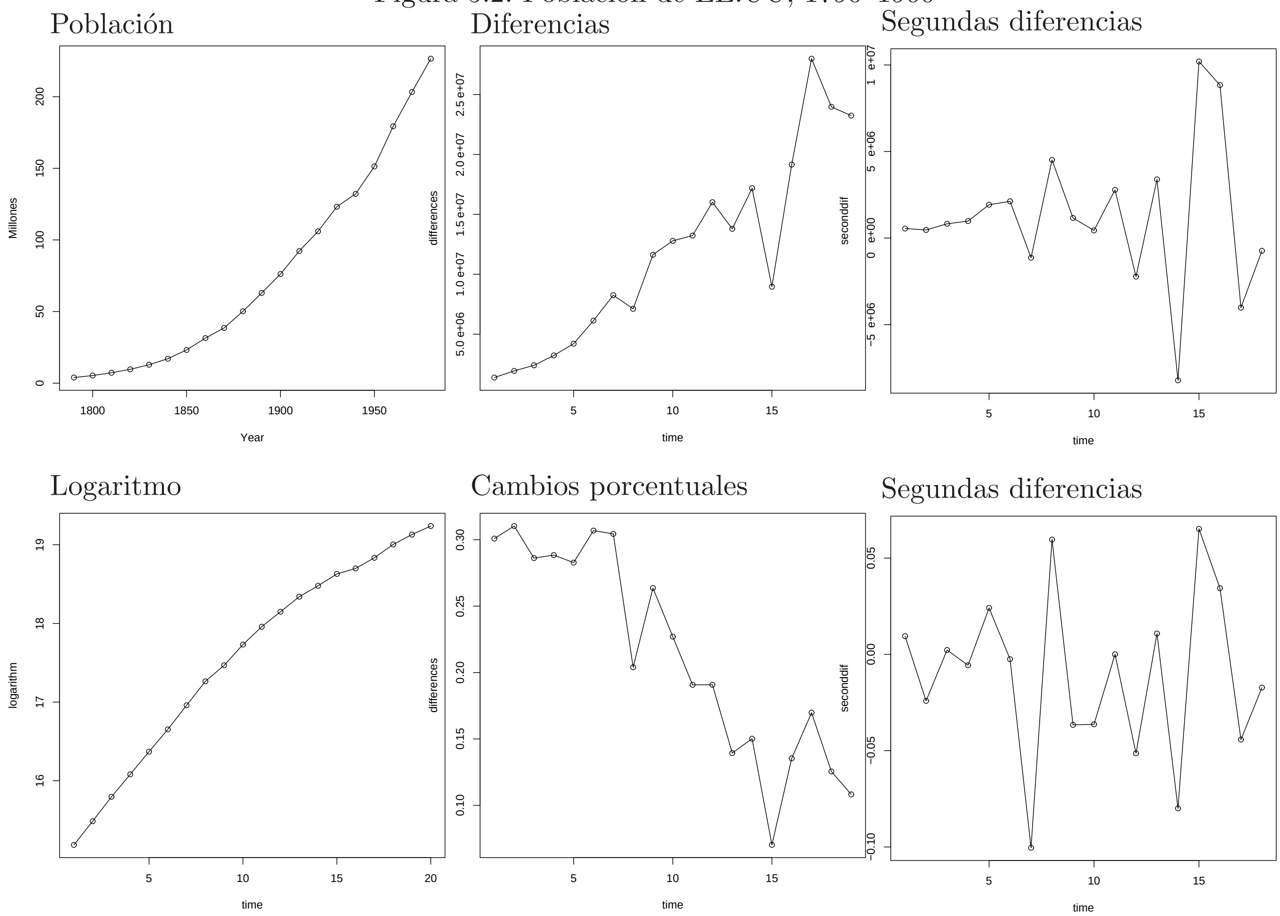
El proceso AR(1) dado por

$$Y_t = c + \phi Y_{t-1} + a_t$$

es estacionario si $|\phi| < 1$ y explosivo si $|\phi| > 1$. Para $\phi = 1$ es $I(1)$, ya que verifica

$$Y_t - Y_{t-1} = c + a_t.$$

Figura 5.2: Población de EE.UU, 1790–1900



Paseo aleatorio con deriva c**:** Es un proceso AR(1) con $\phi = 1$, es decir

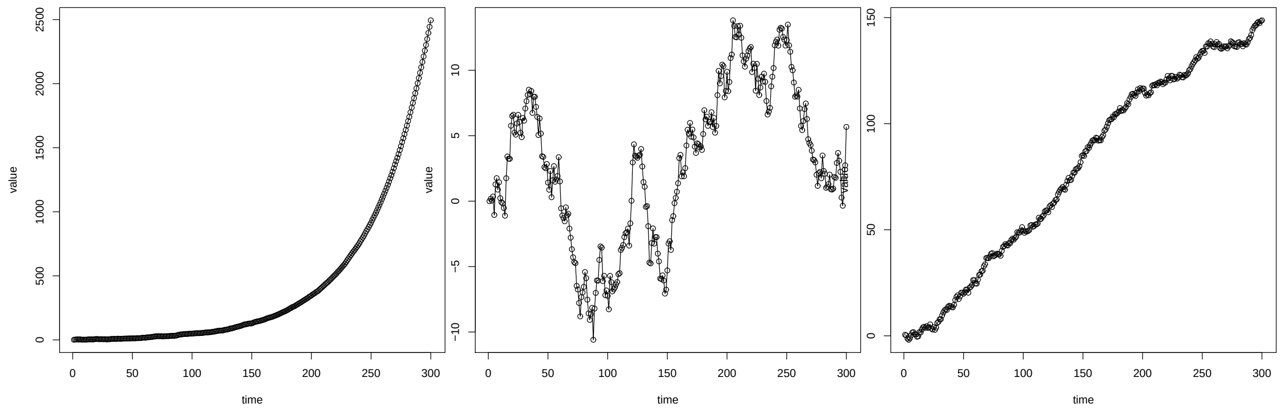
$$Y_t = c + Y_{t-1} + a_t \Leftrightarrow \nabla Y_t = c + a_t$$

Nota:

- En un proceso estacionario o $I(0)$, la constante c regula la media constante. Restando la media podemos trabajar con un proceso sin constante.
- En un proceso $I(1)$, la constante c regula la tendencia. Se dice que el proceso tiene "deriva" c.

Figura 5.3: Procesos AR(1) con $\phi \geq 1$

$\phi = 1{,}02$ y $c = 0$ $\qquad$ $\phi = 1$ y $c = 0$ $\qquad$ $\phi = 1$ y $c = 0{,}5$



Solución de la ecuación: Calculamos la solución de la ecuación del paseo aleatorio de forma recursiva suponiendo $Y_0 = 0$,

$$
\begin{aligned}
Y_1 &= \ldots \\
Y_2 &= \ldots \\
Y_3 &= \ldots \\
&\vdots \\
Y_t &= \ldots
\end{aligned}
$$

Esperanza:

$$E(Y_t) = \ldots$$

Varianza:

$$Var(Y_t) = \ldots$$

Autocovarianzas:

$$Cov(Y_t, Y_{t+h}) = \ldots$$

$$Cov(Y_t, Y_{t-h}) = \ldots$$

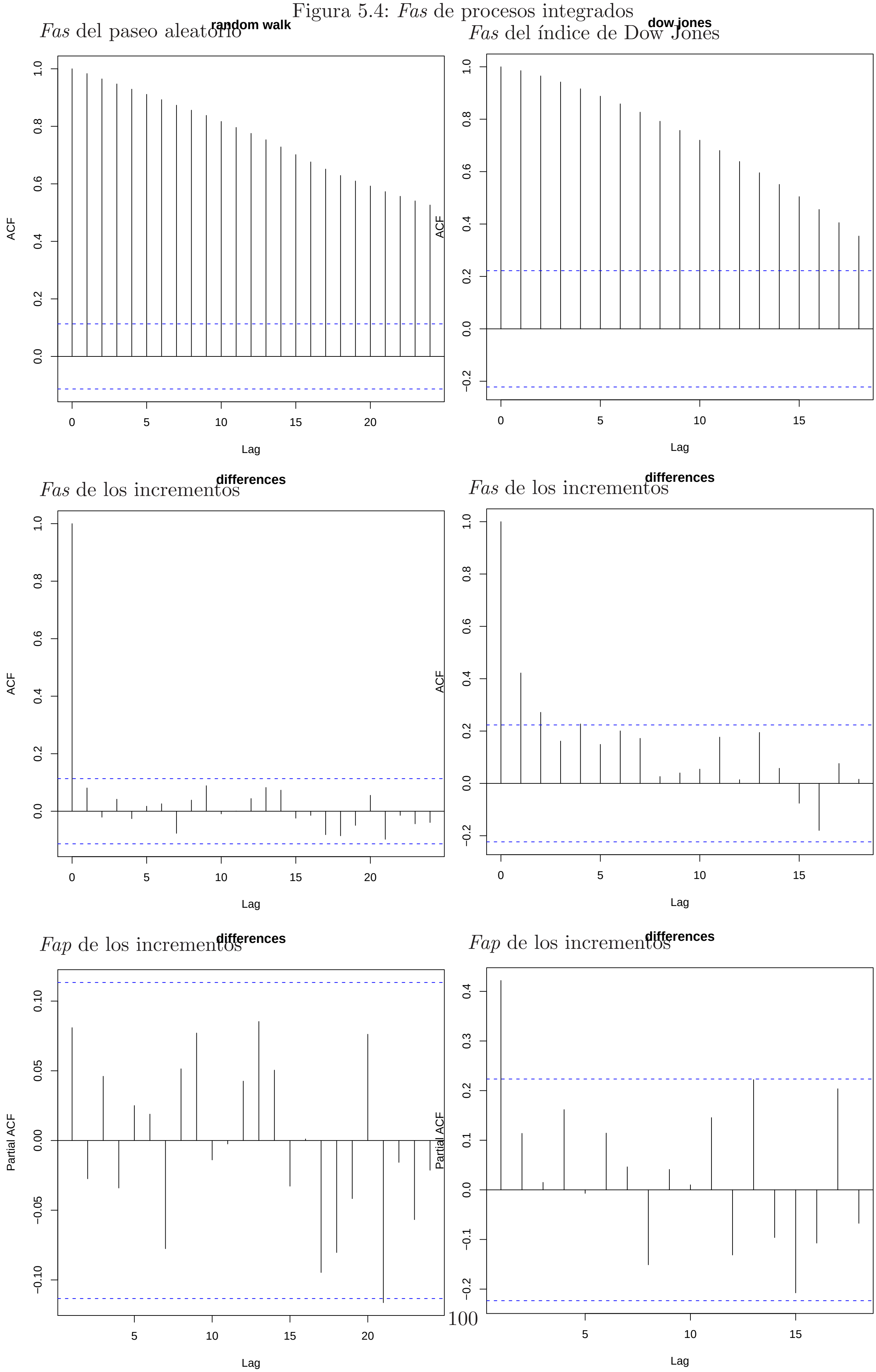
Autocorrelaciones:

$$\rho(Y_t, Y_{t+h}) = \ldots$$

✓ El paseo aleatorio no es estable en media ni en varianza y las autocovarianzas no dependen solo del retardo;

✓ Para un t fijo, las autocorrelaciones decrecen con h de forma lineal.

Figura 5.4: *Fas* de procesos integrados



5.1.2. El proceso de alisado exponencial simple o IMA$(1,1)$

Modelo: El IMA$(1,1)$ es el caso límite del ARMA$(1,1)$ cuando $\phi_1 \to 1$, es decir,

$$Y_t = c + Y_{t-1} + a_t - \theta_1 a_{t-1} \Leftrightarrow \nabla Y_t = c + (1 - \theta_1 B) a_t$$

Este modelo coincide con el paseo aleatorio si $\theta_1 = 0$. Es un modelo popular para la predicción.

Solución: Sustituyendo de forma recursiva, obtenemos

$$\begin{aligned} Y_t &= ... \\ &= ... \\ &= ... \\ &= ... \end{aligned}$$

Si suponemos que la serie comienza en $Y_0 = a_0 = 0$, entonces

$$Y_t = ...$$

Media: Tomando esperanza, obtenemos

$$E(Y_t) = ...$$

Varianza: Tomando varianza, se obtiene

$$Var(Y_t) = ...$$

Autocovarianzas: Tomamos $h > 0$,

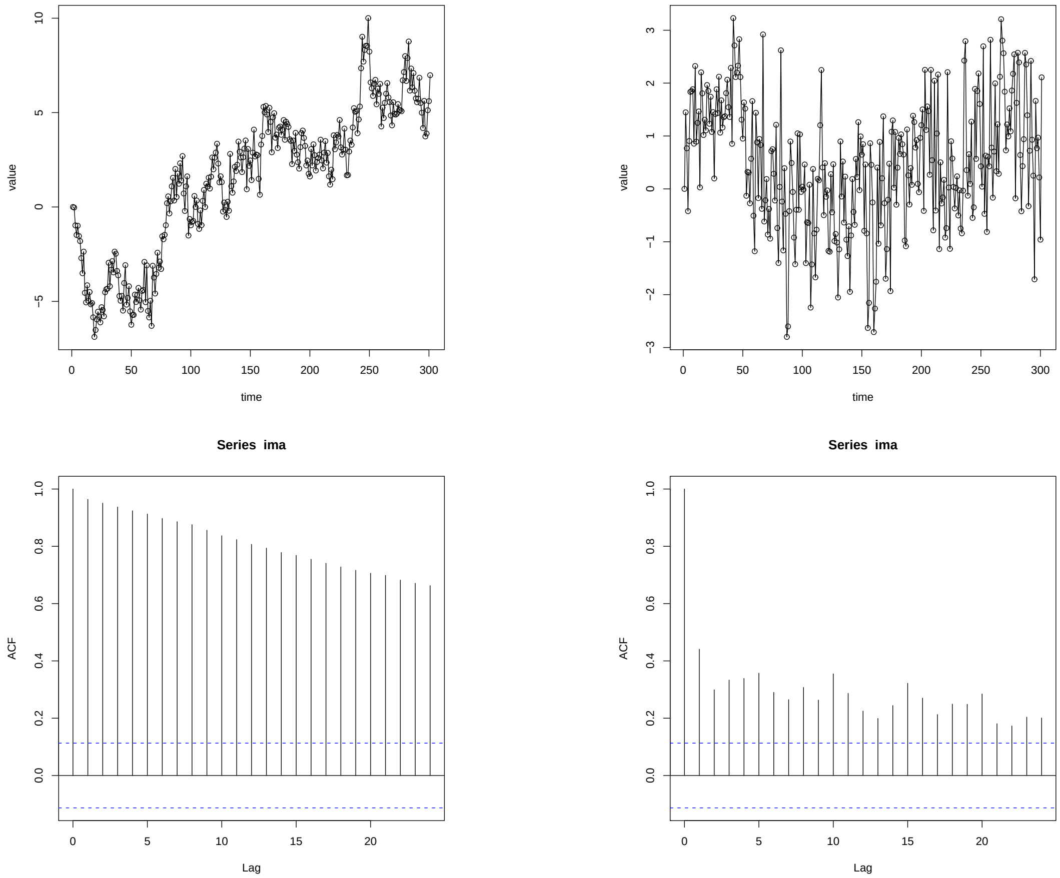
$$\begin{aligned} Cov(Y_t, Y_{t+h}) &= ... \\ &= ... \\ &= ... \end{aligned}$$

Autocorrelaciones: Para $h > 0$, se tiene

$$\rho(Y_t, Y_{t+h}) = ...$$

- Para t grande, si $\theta_1 \approx 0$, las autocorrelaciones se pueden approximar por $\rho(Y_t, Y_{t+h}) \cong (1 + h/t)^{-1/2}$, como para el paseo aleatorio.

Figura 5.5: Series IMA(1,1) y sus correlogramas, $\theta_1 = 0{,}4$ (izq.) y $\theta_1 = 0{,}9$ (der.)



- Si $\theta_1 \cong 1$, las autocorrelaciones no tienen por qué ser grandes, y van decreciendo de forma lineal.

Representación AR(∞): Invertimos el proceso IMA$(1,1)$ con $c = 0$, dado por

$$\nabla Y_t = (1 - \theta_1 B)a_t,$$

donde $|\theta_1| < 1$. ...

¿Cuánto suman los pesos de las observaciones pasadas?

...

✓ El valor de la serie en t es una media ponderada de los valores de la serie en los instantes anteriores, con pesos que decrecen al alejarse en el pasado.

✓ Si $\theta_1 \cong 0$, se tienen solo en cuenta las observaciones más recientes.

✓ Si $\theta_1 \cong 1$, se hace una media ponderada de muchos datos históricos.

5.2. Procesos integrados de orden 2, $I(2)$

Un modelo $I(2)$ que aparece en muchas aplicaciones es el IMA$(2,1)$ dado por la ecuación

$$\nabla^2 Y_t = (1-\theta B)a_t \Leftrightarrow (1-B)^2 Y_t = (1-\theta B)a_t.$$

Consideremos un paseo aleatorio con deriva no constante, dado por la ecuación

$$\nabla Y_t = c_t + u_t, \tag{5.1}$$

donde u_t es una ruido blanco y c_t sigue el modelo

$$c_t = c_{t-1} + \epsilon_t,$$

siendo ϵ_t otro ruido blanco independiente de u_t. Entonces al tomar otra diferencia en (5.1) obtenemos

$$\begin{aligned} \nabla^2 Y_t &= \ldots \\ &= \ldots \\ &= \ldots \\ &= \epsilon_t + Z_t, \end{aligned}$$

donde $Z_t = u_t - u_{t-1}$ es un MA(1) no invertible. Pero se verifica

Ruido blanco + MA(1) no invertible = MA(1) invertible.

Por tanto, $\nabla^2 Y_t$ es una MA(1) invertible, y entonces Y_t es un IMA$(2,1)$.

✓ Los procesos integrados de orden 2 son generalizaciones de los $I(1)$ con pendiente variable.

5.3. Procesos integrados ARIMA(p, d, q)

Modelo: Un proceso ARIMA(p, d, q) ("autoregressive integrated moving average") sigue la ecuación

...

donde

p es el orden de la parte autorregresiva estacionaria;

d es el número de raíces unitarias u orden de integración;

q es el orden de la parte media móvil.

En función del operador retardo, el modelo es

$$\Phi_p(B)\nabla^d Y_t = c + \Theta_q(B)a_t$$

El paseo aleatorio es un ARIMA(..., ..., ..) y el alisado exponencial simple es un ARIMA(..., ..., ...).

✓ Estos procesos se caracterizan porque su *fas* decrece lentamente.

Efectivamente, tomemos un proceso no integrado ARMA(p, q) dado por

$$\Phi_p(B)Y_t = \Theta_q(B)a_t.$$

Sabemos que la *fas* es solución de la ecuación en diferencias

$$\Phi_p(B)\rho_h = 0.$$

La solución a esta ecuación es

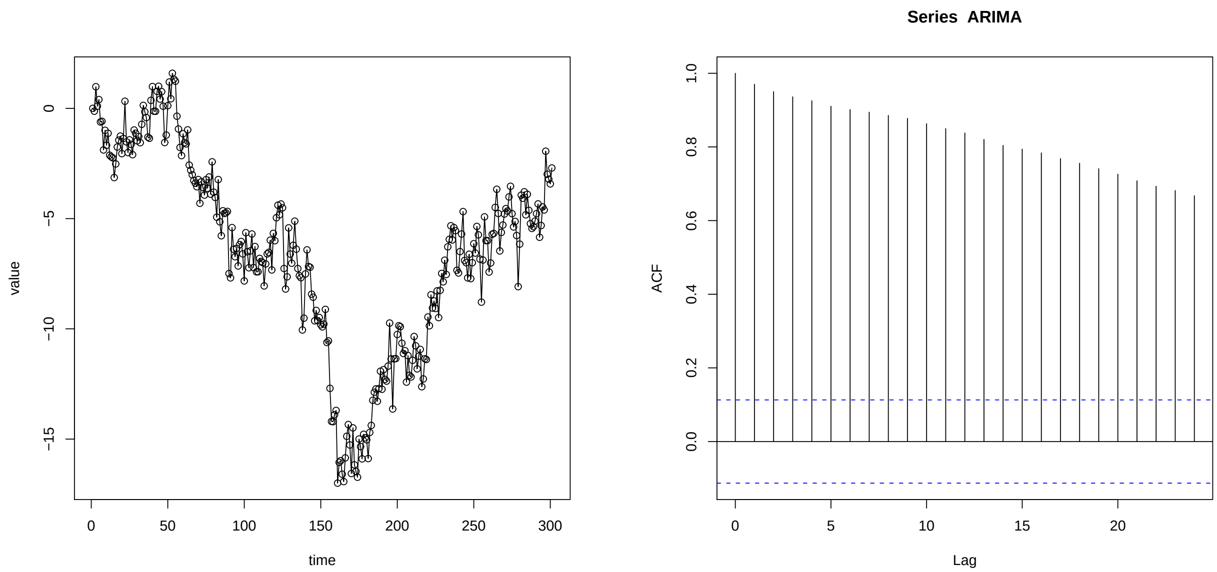
$$\rho_h = \sum_{i=1}^{p} A_i G_i^h$$

donde $G_1^{-1}, \ldots, G_p^{-1}$ son las raíces de la ecuación característica $\Phi_p(B) = 0$ y $A_1, \ldots, A_p$ son constantes. Supongamos que los G_i están todos dentro del círculo unidad, pero uno de ellos, G_j está cerca de 1; es decir, el proceso es "casi" integrado. Tomamos $G_j = 1 - \epsilon$, $\epsilon > 0$ pequeño. Entonces, para h grande,

$$\rho_h = \sum_{i=1;i\neq j}^{p} A_i G_i^h + A_j G_j^h \cong A_j G_j^h = A_j(1-\epsilon)^h,$$

y como ϵ es pequeño, esto tiende a cero muy lentamente.

Figura 5.6: Serie ARIMA(1,1,1) con $\phi_1 = 0{,}4$ y $\theta_1 = 0{,}6$ (izq.) y su correlograma (der.)



5.4. Procesos integrados y tendencias

✓ Los procesos ARIMA incluyen muchos procesos no estacionarios que aparecen en la práctica.

✓ En particular, incluyen como casos límite a procesos con tendencia tanto determinista como estocástica.

Se verifica que

Tendencia polinómica + Proceso estacionario = Proceso integrado.

Ejemplo 37 *Consideremos el proceso con tendencia determinista lineal,*

$$Y_t = b + ct + u_t, \quad u_t \quad \textit{ruido blanco.}$$

Tomando la primera diferencia, obtenemos

$$\begin{aligned} \nabla Y_t &= ... \\ &= ... \end{aligned}$$

donde ∇u_t es estacionario (es diferencia de procesos estacionarios) pero no es invertible. Por tanto, ∇Y_t es integrado de orden 1, $I(1)$.

Ejemplo 38 *Consideremos el proceso con tendencia determinista polinómica*

$$\begin{aligned} Y_t &= \mu_t + u_t \\ \mu_t &= \beta_0 + \beta_1 t + \cdots + \beta_d t^d \end{aligned}$$

Diferenciando d veces obtenemos un proceso estacionario no invertible,

$$\nabla^d Y_t = d\beta_d + \nabla^d u_t,$$

donde $\nabla^d u_t$ es un proceso estacionario al ser diferencia entre procesos estacionarios. Por tanto, Y_t es un proceso integrado de orden d, $I(d)$.

Ejemplo 39 *Consideremos el proceso anterior con $d = 2$, es decir,*

$$Y_t = \mu_t + u_t, \quad \mu_t = \beta_0 + \beta_1 t + \beta_2 t^2,$$

pero donde ahora u_t es un $AR(1)$, es decir,

$$(1 - \phi_1 B)u_t = a_t \Leftrightarrow u_t = (1 - \phi_1 B)^{-1} a_t.$$

Tomando diferencia de orden 2, obtenemos

$$\nabla^2 Y_t = ...$$

donde

$$\begin{aligned} & \nabla^2 u_t = ... \\ \Leftrightarrow \quad & (1 - \phi_1 B)\nabla^2 u_t = ... \end{aligned}$$

Es decir, el proceso $\nabla^2 u_t$ es un $ARMA(1, 2)$ no invertible. Entonces, $\nabla^2 Y_t$ es un $ARIMA(1, 1, 2)$.

Nota 13 *Modelizamos unos datos con un proceso ARIMA. Entonces*

- *Si tenemos una serie con tendencia polinómica de orden d y la diferenciamos d veces, obtenemos componentes media móvil no invertibles.*
- *La serie diferenciada tendrá media distinta de 0, y esta media es función del coeficiente de orden mayor de la tendencia.*

Ejemplo 40 *Consideremos el proceso con nivel estocástico pero pendiente determinista definido por*

$$Y_t = \mu_t + u_t$$

donde el nivel en t se obtiene sumando una constante c (la pendiente) al nivel en $t-1$ de la forma

$$\mu_t = \mu_{t-1} + c + \epsilon_t,$$

y donde (u_t) y (ϵ_t) son ruidos blancos independientes. Tomando primera diferencia, obtenemos

$$\nabla Y_t = ...$$

donde

$$\nabla \mu_t = ... \qquad \Rightarrow \nabla Y_t = ...$$

Es decir, ∇Y_t es suma de un ruido blanco y un $MA(1)$ no invertible, que es un $MA(1)$ invertible. Por tanto, Y_t es un $IMA(1,1)$.

Ejemplo 41 *En la práctica es difícil encontrar unos datos en los que la tendencia siga exactamente una función concreta. Los procesos con tendencia estocástica son bastante más realistas. Consideremos el proceso definido por*

$$\begin{aligned} Y_t &= \mu_t + u_t, \\ \mu_t &= \mu_{t-1} + \beta_t + \epsilon_t, \\ \beta_t &= \beta_{t-1} + \delta_t, \end{aligned}$$

donde (u_t), (ϵ_t) y (δ_t) son ruidos blancos independientes. En este proceso, la media crece de forma lineal, pero con pendiente variable.

Tomando una diferencia, obtenemos

$$\nabla Y_t = ...$$

donde

$$\nabla \mu_t = ... \qquad \Rightarrow \nabla Y_t = ...$$

Tomando ahora una segunda diferencia, obtenemos

$$\nabla^2 Y_t = ...$$

El término de la derecha es un $MA(2)$ invertible, ya que es suma de un $MA(2)$, un $MA(1)$ y ruido blanco; por tanto, Y_t es un $IMA(2,2)$.

✓ *Si las varianzas de ϵ_t y δ_t son cero, obtenemos un modelo con tendencia determinista.*