Introducción a Series temporales

Modelos AR, MA, ARMA, ARIMA, GARCH

Georgina Flesia, Patricia Kisbye
FaMAF.

26 de noviembre de 2022

Qué es una serie de tiempo

Series de tiempo:

Una **serie de tiempo** es un conjunto de datos indexados con un orden temporal:

$$(x_t)_{t \in T}, \qquad T \subset \mathbb{R}$$

- T discreto: serie temporal discreta.
- T un intervalo real: serie temporal continua.

Ejemplos de series de tiempo

- Precios de productos financieros, índices, tasas de interés, inflación, PBI, etc.
- Producción anual de trigo, maíz, exportaciones, etc.
- Población, tasas de natalidad, mortalidad, accidentes, crímenes,
- Medicina: electrocardiograma, encefalogramas, etc.
- FAMAF: cantidad de inscriptos, alumnos, egresados, etc. por año.

Estudiar series

Objetivos:

- Modelos probabilísticos que permitan entender y describir los procesos aleatorios que generan la serie temporal.
- Predicción de valores futuros.
- Control de procesos.

Características de una serie de tiempo

- La serie $x_1, x_2, x_3, \ldots, x_T$:
  - una muestra de tamaño 1 de un proceso estocástico, o
  - T observaciones de T variables aleatorias, no independientes.
- La dependencia entre los valores de la serie no permite aplicar el análisis estadístico usual.
- Para definir algunos modelos estadísticos, una primera aproximación es suponer estacionariedad de la serie, es decir, que la serie se *comporte siempre igual*.
- ¿Qué significa que sea **estacionaria**?

Serie de retornos

La mayoría de los estudios financieros involucran retornos de activos, en lugar de precios:

- El retorno del activo es una síntesis completa y libre de escalas de la oportunidad de inversión.
- La serie de retornos es más manejable que la de precios, por tener mejores propiedades.

Definición de retorno

- Retorno simple:

$$R_t = \frac{P_t - P_{t-1}}{P_t} = \frac{P_t}{P_{t-1}} - 1$$

- Retorno compuesto:

$$1 + R_t[k] = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}}$$

- Retorno continuo:

$$e^{r_t} = \frac{P_t}{P_{t-1}}$$

$$1 + R_t = \frac{105}{100} \qquad \color{red}{R_t = 0{,}05}$$

$$e^{r_t} = \frac{105}{100} \qquad \color{blue}{r_t = 0{,}048} = \ln(1 + 0{,}05)$$

La distribución de los retornos

Algunas hipótesis que se suelen proponer en relación a los retornos:

- Retornos simples independientes e igualmente distribuidos, con media y varianza constante.
- Log retornos independientes e igualmente distribuidos con media μ y varianza σ^2. Modelo de Black Scholes.
- Retornos con mezcla de distribuciones.

Series estacionarias

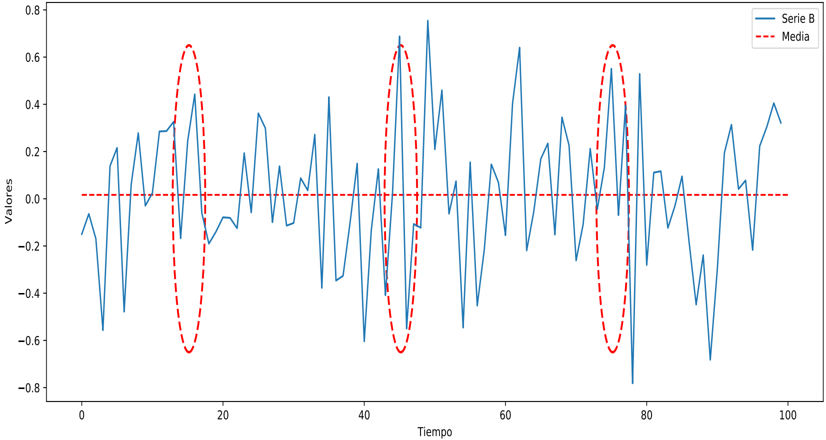
Serie de tiempo: $x_1, x_2, x_3, \ldots, x_T$

Proceso estrictamente estacionario

La distribución conjunta de $(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$ es la misma que la distribución conjunta de $(x_{t_1+h}, x_{t_2+h}, \ldots, x_{t_k+h})$ para cualquier subconjunto $(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$. Sólo puede depender de h.

En la práctica son procesos difíciles de encontrar.

Estacionariedad



Proceso (débilmente) estacionario

$E[x_t] < \infty$, $Var(x_t) < \infty$, y:

- $E[x_t] = \mu$, es una constante independiente de t.
- $Cov(x_t, x_{t+h}) = \gamma_h$, depende de h, no de t.

- γ_h: Autocovarianza del retardo h.
- ρ_h: Autocorrelación del retardo h

$$
\begin{aligned}
\rho_h &= \frac{Cov(x_t, x_{t+h})}{\sqrt{Var(x_t)}\sqrt{Var(x_{t+h})}} = \frac{\gamma_h}{\gamma_0} \qquad -1 \leq \rho_h \leq 1. \\
\rho_0 &= 1.
\end{aligned}
$$

```
import statsmodels.tsa.api as stm
...
stm.graphics.plot_acf(serie, lags = 20)
```

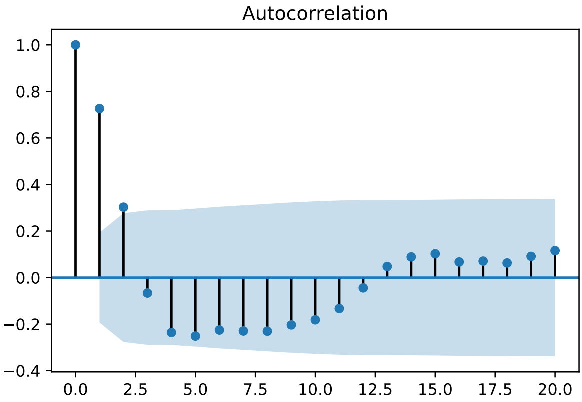


Figura: Función de autocorrelación

Estimaciones

Si la serie es estacionaria y $\lim_{h\to\infty} \rho_h = 0$ (condición suficiente), entonces existen estimadores consistentes:

- $\hat{\mu}$ para la media:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} x_t$$

- $\hat{\gamma}_h$ para las autocovarianzas:

$$\frac{1}{T}\sum_{t=1}^{T-h}(x_t - \hat{\mu})(x_{t+h} - \hat{\mu}).$$

```
import quandl
quandl.get("NSE/SPYL", start_date="2011-01-12", end_date="2013-01-01")
```

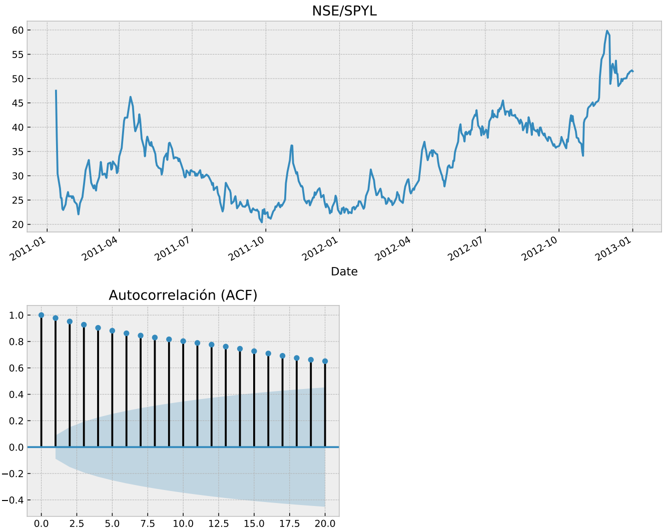


Figura: Función de autocorrelación

Ruido blanco

Ruido blanco

Es un proceso a_t con las siguientes propiedades:

- Media 0: $E[a_t] = 0$.
- Varianza finita constante: $Var(a_t) = \sigma_a^2 < \infty$.
- Autocovarianzas nulas: $Cov(a_t, a_{t+h}) = 0$ para $h \neq 0$.

- Si cada a_t tiene distribución normal: $a_t \sim N(0, \sigma_a)$, se trata de un Ruido blanco Gaussiano.

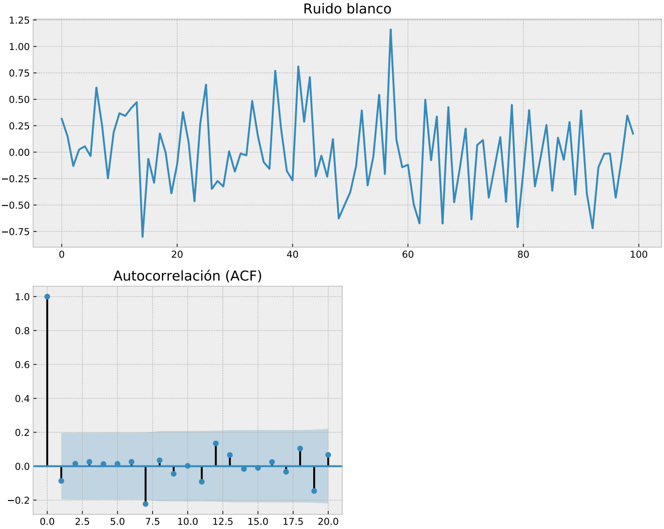


Figura: Ruido blanco

Series de tiempo lineales

Series lineales

Son las series que pueden escribirse como

$$x_t = \mu_t + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots,$$

donde μ_t es la media y $\{a_t\}$ es un ruido blanco.

- a_t: es la **innovación** del proceso en tiempo t.
- Para las series lineales débilmente estacionarias $\mu_t = \mu$.

Teorema de Wold: Toda serie estacionaria con media 0 y sin componente determinística puede escribirse de esa forma, con la condición:

$$\sum_{j=1}^{\infty} |\psi_j| < \infty.$$

Caminata aleatoria

Caminata aleatoria

Es un proceso estocástico x_t que satisface:

$$x_t = x_{t-1} + a_t,$$

donde a_t es un ruido blanco.

Notamos que:

$$x_t = x_{t-1} + a_t = x_{t-2} + a_{t-1} + a_t = x_{t-3} + a_{t-2} + a_{t-1} + a_t = \dots.$$

$$\boxed{x_t = x_0 + a_t + a_{t-1} + \cdots + a_1 = x_0 + \sum_{j=0}^{t-1} a_{t-j}}$$

- No es estacionaria. Tiene una tendencia estocástica: $Var(x_t) = t\sigma_a^2$.

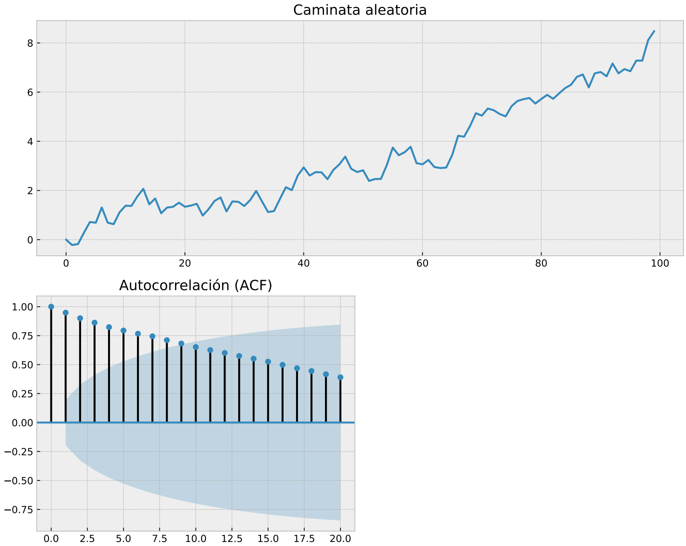


Figura: Caminata aleatoria

Ejemplo

Ejemplo

$$x_t = rx_{t-1} + a_t$$

a_t un ruido blanco.

$$\begin{aligned} x_t &= rx_{t-1} + a_t = r(rx_{t-2} + a_{t-1}) + a_t \\ &= r^2 x_{t-2} + ra_{t-1} + a_t = r^2(rx_{t-3} + a_{t-2}) + ra_{t-1} + a_t \\ &= r^3 x_{t-3} + r^2 a_{t-2} + ra_{t-1} + a_t \end{aligned}$$

$$x_t = r^t x_0 + \sum_{j=0}^{t-1} r^j a_{t-j}$$

- Si $|r| < 1$ la suma $\sum_j r^j$ es convergente.
- Si $|r| < 1$ la serie es estacionaria.

Modelos Autorregresivos

- $AR(p)$. El valor de x_t depende de observaciones anteriores: x_{t-1}, x_{t-2} . . . .
- $MA(q)$. Las observaciones dependen linealmente de innovaciones anteriores: a_{t-1}, a_{t-2}, ...
- $ARMA(p, q)$. Combinaciones de ambas
- **Series integradas** $I(p)$. No estacionarias, pero sus diferencias $x_t - x_{t-1}$ (o de orden mayor) sí.
- $ARIMA(p, I, q)$. Combinaciones de las tres.
- **ARCH(p), GARCH(p,q)**. Las innovaciones a_t condicionales no tienen varianza constante.
- Muchos otros.

Modelos AR(p)

$$AR(p)$$

- Son modelos **A**uto**R**regresivos, de orden p.
- Estos modelos intentan explicar la i-ésima observación a partir de las observaciones anteriores.
- Según el número de observaciones anteriores consideradas es el orden p.

$$\begin{aligned} x_t &= \phi_0 + \phi_1 x_{t-1} + a_t & AR(1) \\ x_t &= \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t & AR(2) \end{aligned}$$

- a_t: innovación, es un ruido blanco.

Modelo $AR(1)$

$$x_t = \phi_0 + \phi_1 x_{t-1} + a_t$$

- Predicción: Dado x_{t-1}, el valor **esperado** de x_t es $\phi_0 + \phi_1 x_{t-1}$:

$$\hat{x}_t = E[x_t \mid \text{tiempo pasado}] = \phi_0 + \phi_1 x_{t-1}$$

- Si la serie es estacionaria:

$$\mu = \frac{\phi_0}{1-\phi_1} \qquad\qquad x_t - \mu = \phi_1(x_{t-1} - \mu) + a_t$$

- Restando la media μ se trabaja directamente con:

$$y_t = \phi_1 y_{t-1} + a_t.$$

- Es estacionario si y sólo si $|\phi_1| < 1|$.

Modelo $AR(p)$ - Autocorrelaciones

- La ACF de un AR(1) es
$$\rho_h = \phi_1^h, \quad h > 0.$$
- Para un AR(2): $\rho_h = \phi_1 \rho_{h-1} + \phi_2 \rho_{h-2}, \quad h > 2.$
- La ACF no da demasiada información para distinguir un AR(p).

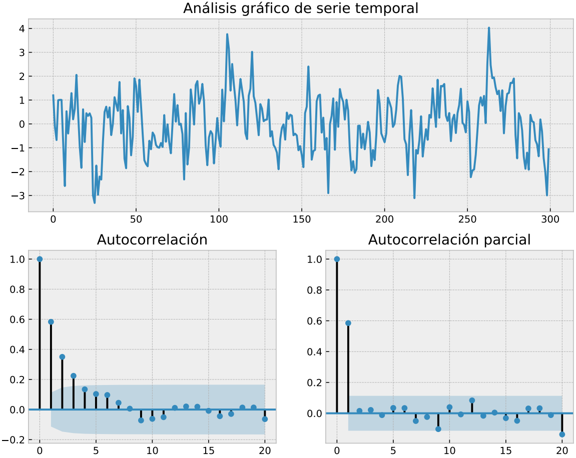
Autocorrelación parcial (PACF)

Es la correlación entre x_t y x_{t+l} quitando la dependencia sobre $x_{t+1}, x_{t+1}, \ldots, x_{t+l-1}$.

- Si es un AR(p), se anulan todas las autocorrelaciones parciales desde $p+1$ en adelante.

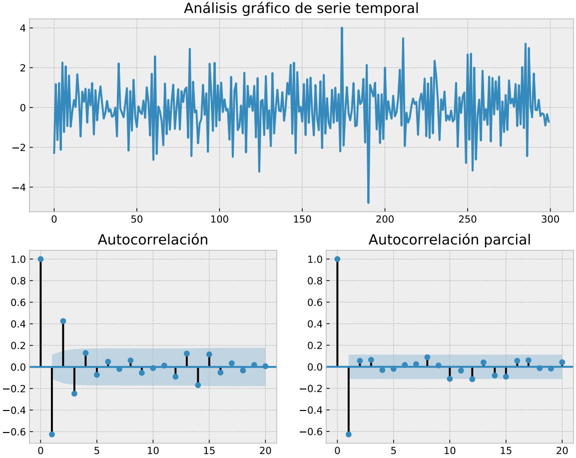
Modelo AR(1)

```
phi = 0.6
ar = np.r_[1, -phi], ma = np.r[1,0]
ar1 = smt.arma_generate_sample(ar=ar, ma=ma, nsample=n)
```



Modelo AR(1)

```
phi = -0.6
ar = np.r_[1, -phi], ma = np.r[1,0]
ar1 = smt.arma_generate_sample(ar=ar, ma=ma, nsample=n)
```



Propiedades $AR(2)$

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t$$

- Si es débilmente estacionario: $\mu = \frac{\phi_0}{1-\phi_1-\phi_2}$
- Función de autocorrelación ACF: $\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2}$.
- Para que sea débilmente estacionaria, las raíces de

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

  deben tener valor absoluto mayor que 1.
- Estas propiedades se generalizan a un $AR(p)$.

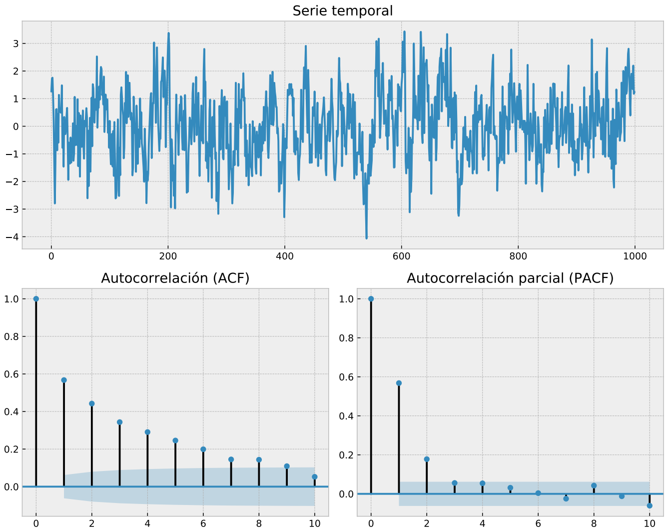


Figura: Modelo AR(2)

El operador Lag o Retardo

Operador Retardo

$$Lx_t = x_{t-1}, \qquad L^2 x_t = x_{t-2}, \; \dots L^k x_t = x_{t-k}$$

- Un proceso $AR(p)$:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots \phi_p x_{t-p} + a_t$$

- lo podemos escribir como:

$$\underbrace{(I - (\phi_1 L + \phi_2 L^2 + \cdots + \phi_p L^p))}_{=\Phi(L)} x_t = a_t.$$

- Si las raíces del polinomio $\Phi(z) = (I - (\phi_1 z + \phi_2 z^2 + \cdots + \phi_p z^p))$ tienen módulo mayor que 1: invertibilidad.

$$x_t = \Psi(L) a_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

Modelos MA(q)

$$MA(q)$$

- Son modelos de **media móvil**: **M**oving**A**verage, de orden q.
- Estos modelos intentan explicar la i-ésima observación a partir de las innovaciones anteriores.
- Según el número de observaciones anteriores consideradas es el orden q.

$$\begin{aligned} x_t &= \mu + a_t - \theta_1 a_{t-1} && MA(1) \\ x_t &= \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} && MA(2) \end{aligned}$$

- a_t: innovación, es un ruido blanco.
- μ: es la media del proceso.

Propiedades de $MA(q)$

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

- Son siempre (débilmente) estacionarias.
- Para un $MA(1)$,

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2}, \qquad \rho_h = 0,\ h > 1.$$

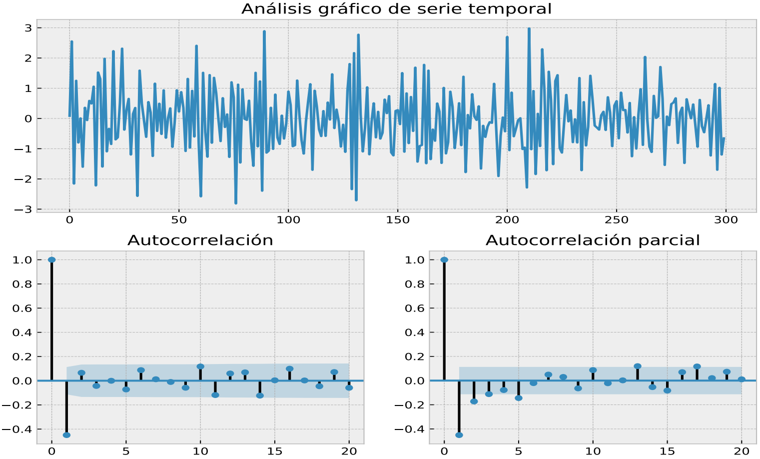
- Para un $MA(q)$:

$$\rho_h = 0 \qquad h > q$$

- La Función de Autocorrelación Parcial (PACF) no da mayor información.

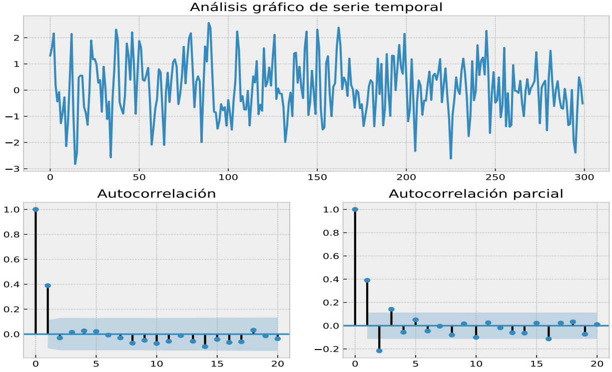
Modelo MA(1)

```
theta = 0.6
ar = np.r_[1, .0], ma = np.r[1,-theta]
ar1 = smt.arma_generate_sample(ar=ar, ma=ma, nsample=n)
```



Modelo MA(1)

```
theta = -0.6
ar = np.r_[1, .0], ma = np.r[1,-theta]
ar1 = smt.arma_generate_sample(ar=ar, ma=ma, nsample=n)
```



Modelos ARMA(p,q)

$$ARMA(p, q)$$

- **A**uto **R**egresive **M**oving **A**verage.
- Son modelos que conjugan a los autorregresivos con los de media móvil.
- a_t: innovación, es un ruido blanco.

ARMA(1,2)

$$\underbrace{x_t - \phi_1 x_{t-1}}_{AR(1)} = \phi_0 + \underbrace{a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}}_{MA(2)}$$

Series integradas

Si tenemos una serie de tiempo x_t, podemos:

- **diferenciarla:**

$$\Delta x_t = x_t - x_{t-1}$$

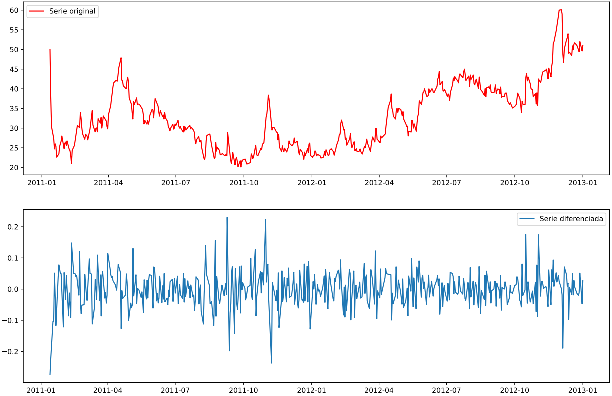
- y luego **integrarla:**

$$x_t = x_{t-1} + \Delta x_t.$$

Series $I(1)$

Una serie es **integrada de orden 1** si al diferenciarla **una vez** resulta estacionaria.

Ejemplos: precios de activos, índices, tasas de interés.


Serie original
60
55
50
45
40
35
30
25
20
2011-01
2011-04
2011-07
2011-10
2012-01
2012-04
2012-07
2012-10
2013-01
Serie diferenciada
0.2
0.1
0.0
−0.1
−0.2
2011-01
2011-04
2011-07
2011-10
2012-01
2012-04
2012-07
2012-10
2013-01

Raíces unitarias

- Si en un modelo AR(p) ocurre:

$$1 - \phi_1 - \phi_2 - \cdots - \phi_p = 0$$

  se dice que el proceso tiene un **raíz unitaria.**

- Los procesos con raiz unitaria no son estacionarios. Se pueden transformar en un proceso estacionarios diferenciando:

una vez	$x_t - x_{t-1} = \Delta x_t = (I - L)x_t$
dos veces	$(x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = (I - L)^2 x_t$
...	...

- Para testear la presencia de raíces unitarias: **Test ADF**: Augmented Dickey Fuller Test.

Modelos ARIMA

Modelo $ARIMA(p, d, q)$

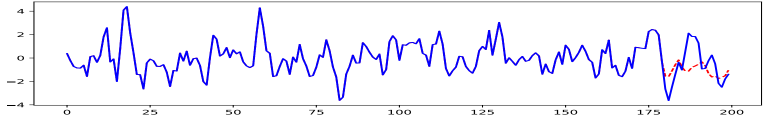
Un proceso es un ARIMA(p, d, q) si al diferenciarlo d veces resulta un ARMA(p, q).

- Predicción en un ARIMA$(2, 1, 1)$:

$$x_t - x_{t-1} = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} - a_t - \theta_1 a_{t-1}$$

- Dadas las observaciones x_1, x_2, ..., x_h:

$$\begin{aligned} \hat{x}_{h+1} &= E[x_{h+1} \mid \text{observaciones anteriores}] \\ &= x_h + \phi_0 + \phi_1 x_h + \phi_2 x_{h-1} - \theta_1 a_h \\ \text{Error} &= a_{h+1} \end{aligned}$$



La volatilidad como medida de riesgo

Los modelos presentados hasta el momento asumen que la volatilidad condicional es constante:

$$Var[a_t \mid F_{t-1}] = \sigma_a^2.$$

En la práctica se observan otras características:

- Clusters de volatilidad: períodos de alta volatilidad y otros de baja volatilidad.
- La volatilidad evoluciona de manera continua, los saltos son raros.
- La volatilidad no diverge a infinito, se mantiene en un rango: es frecuentemente **estacionaria**.
- Reacciona de manera diferente a una suba fuerte de precios a una baja fuerte de precios.

Modelos de heteroscedasticidad condicional

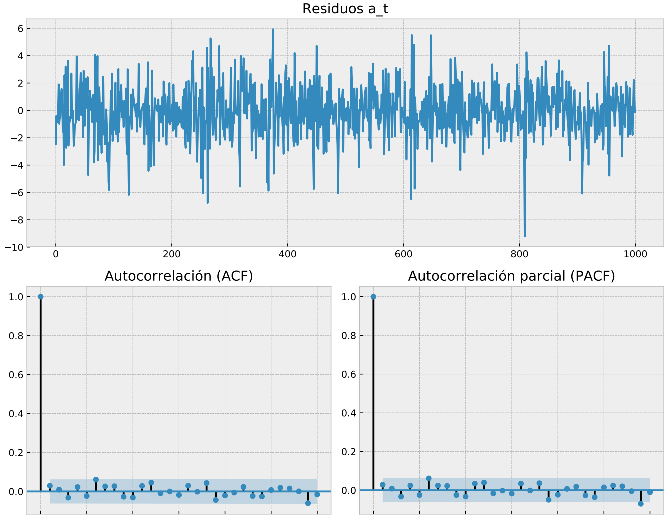
Supongamos que hemos ajustado una serie de tiempo de retornos de un activo a un determinado modelo ARIMA μ_t:

$$x_t - \mu_t = a_t,$$

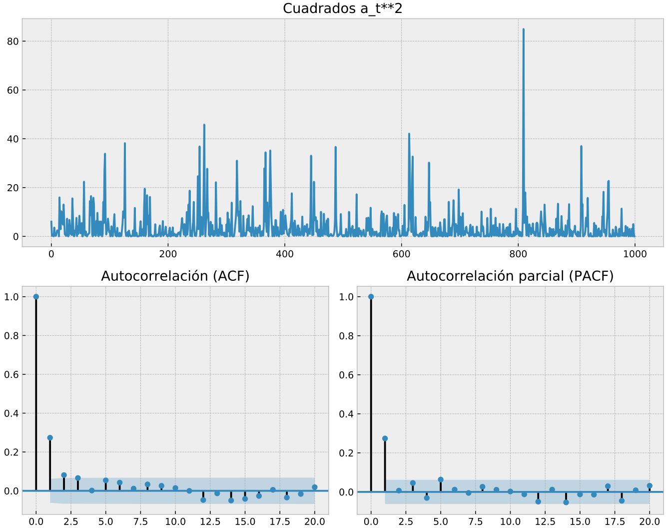
donde a_t son los errores.

- Los errores a_t no están serialmente correlacionados.
- Al analizar sus cuadrados a_t^2 se observa una dependencia lineal.

Residuos: ¿ruido blanco?



Residuos al cuadrado: dependencia lineal



Construcción del modelo

- Especificar un modelo para la media de los retornos de modo de eliminar la dependencia lineal: $x_t - \mu_t = a_t$, x_t un ARIMA(p,d,q).
- Usar los residuos a_t para testear un modelo autorregresivo de heteroscedasticidad condicional.
- Ajustar a un modelo de volatilidad: ARCH, GARCH.
- Existen otros modelos: EGARCH, TGARCH.

Modelos ARCH(p) y GARCH(p, q)

$$a_t = \sigma_t \epsilon_t,$$

- σ_t es la **volatilidad**.
- Se modela σ_t^2.

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \cdots + \alpha_p a_{t-p}^2$$

- La volatilidad depende del cuadrado de las innovaciones anteriores. Luego modela igual respuesta para innovaciones positivas que negativas.
- Responde muy lentamente a shocks aislados de los retornos.
- No provee una explicación de la fuente de aleatoriedad de la serie financiera.

Construcción de un ARCH(1)

```
a0 = 2
a1 = .5

y = w = np.random.normal(size=1000)
Y = np.empty_like(y)

for t in range(len(y)):
    y[t] = w[t] * np.sqrt((a0 + a1*y[t-1]**2))
```

Modelo GARCH(p, q)

$$x_t - \mu_t = a_t, \qquad a_t = \sigma_t a_t,$$

- Se modela σ_t^2:

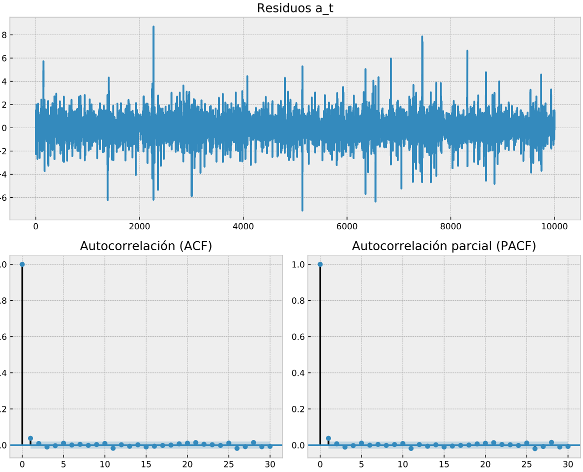
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

- Permite modelar clusters de volatilidad.
- No resuelve el problema de respuesta a innovaciones positivas y negativas.

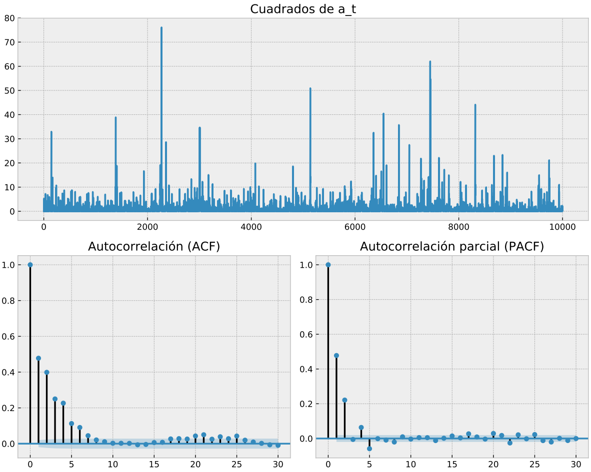
Generación de un GARCH(1,1)

```
a0 = 0.2
a1 = 0.5
b1 = 0.3

n = 10000
w = np.random.normal(size=n)
eps = np.zeros_like(w)
sigsq = np.zeros_like(w)

for i in range(1, n):
    sigsq[i] = a0 + a1*(eps[i-1]**2) + b1*sigsq[i-1]
    eps[i] = w[i] * np.sqrt(sigsq[i])



am = arch_model(eps).fit()
am.summary()
```

GARCH(1,1)


Residuos a_t
8
6
4
2
0
−2
−4
−6
0
2000
4000
6000
8000
10000
Autocorrelación (ACF)
Autocorrelación parcial (PACF)
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30

GARCH(1,1) - cuadrados



Resumen estadístico: GARCH(1,1)

```
                     Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                      y   R-squared:                      -0.000
Mean Model:             Constant Mean   Adj. R-squared:                 -0.000
Vol Model:                      GARCH   Log-Likelihood:               -12237.3
Distribution:                  Normal   AIC:                           24482.6
Method:            Maximum Likelihood   BIC:                           24511.4
                                        No. Observations:                10000
Date:                Sun, Sep 09 2018   Df Residuals:                     9996
Time:                        18:56:37   Df Model:                            4
                                  Mean Model
==============================================================================
                  coef    std err          t      P>|t|      95.0% Conf. Int.
------------------------------------------------------------------------------
mu         -6.7225e-03  6.735e-03     -0.998      0.318 [-1.992e-02,6.478e-03]
                            Volatility Model
========================================================================
                 coef    std err          t      P>|t|  95.0% Conf. Int.
------------------------------------------------------------------------
omega          0.2021  1.043e-02     19.383  1.084e-83 [  0.182,  0.223]
alpha[1]       0.5162  2.016e-02     25.611 1.144e-144 [  0.477,  0.556]
beta[1]        0.2879  1.870e-02     15.395  1.781e-53 [  0.251,  0.325]
========================================================================
```