Series de tiempo

Patricia Kisbye - Georgina Flesia

Series de tiempo financieras

¿Qué es una serie de tiempo?

- Un conjunto de datos indexados con un orden temporal:

$$x_1, x_2, x_3, x_4, \ldots., x_T$$

  donde además interesan aquellos datos que muestran cierta aleatoriedad.

- -3, 2.5, 3.8, 4.8, 7.1, -0.2, ….

Motivación

Las series de tiempo aparecen en diferentes contextos:

- Ciencia
- Tecnología
- Economía
- Finanzas
- Política

Series de tiempo en finanzas

- Ejemplos

Los datos que cambian con el tiempo: precios de activos, ventas, índices, bonos, etc.

- Ejemplo de precios de activos

	Open	High	Low	Close	Adj Close	Volume
Date						
2022-11-23	0.028874	0.071673	0.037266	0.075310	0.075310	0.333769
2022-11-25	0.064099	0.008568	0.046054	-0.001858	-0.001858	-0.770872
2022-11-28	-0.027945	0.017662	-0.009065	0.000328	0.000328	0.606192
2022-11-29	0.027567	-0.011310	-0.001398	-0.011492	-0.011492	-0.108446
2022-11-30	-0.013935	0.043980	0.010463	0.073903	0.073903	0.269923

Características de una serie de tiempo

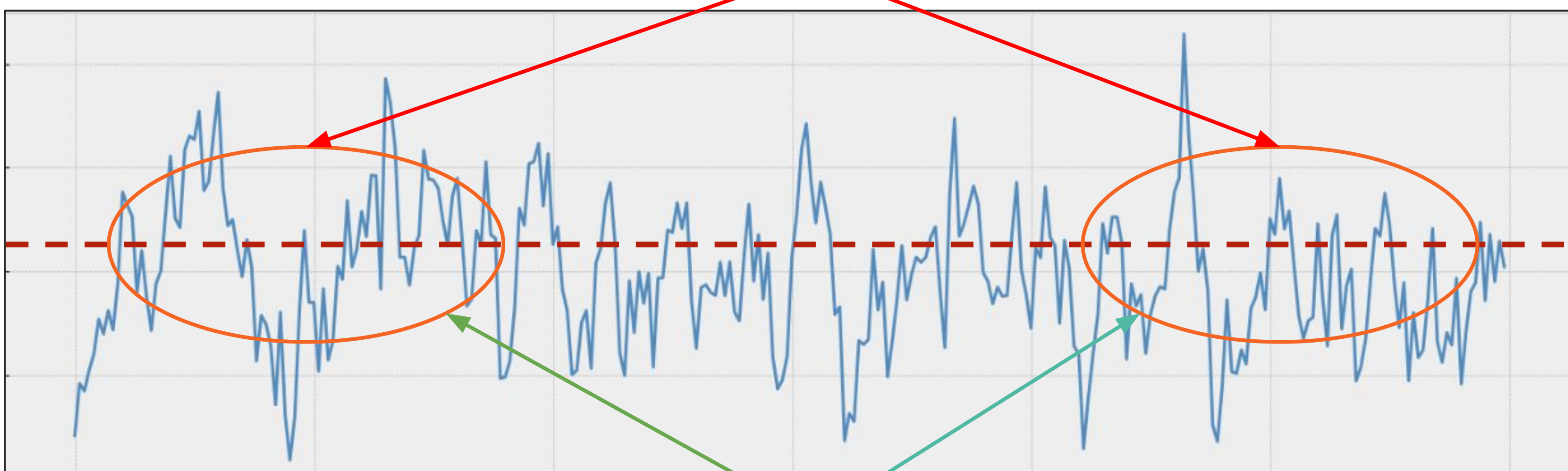
- La serie de tiempo puede verse como:
  - una muestra de tamaño 1 de un proceso estocástico
  - T observaciones de T variables aleatorias, no independientes.

- 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.

Estacionariedad

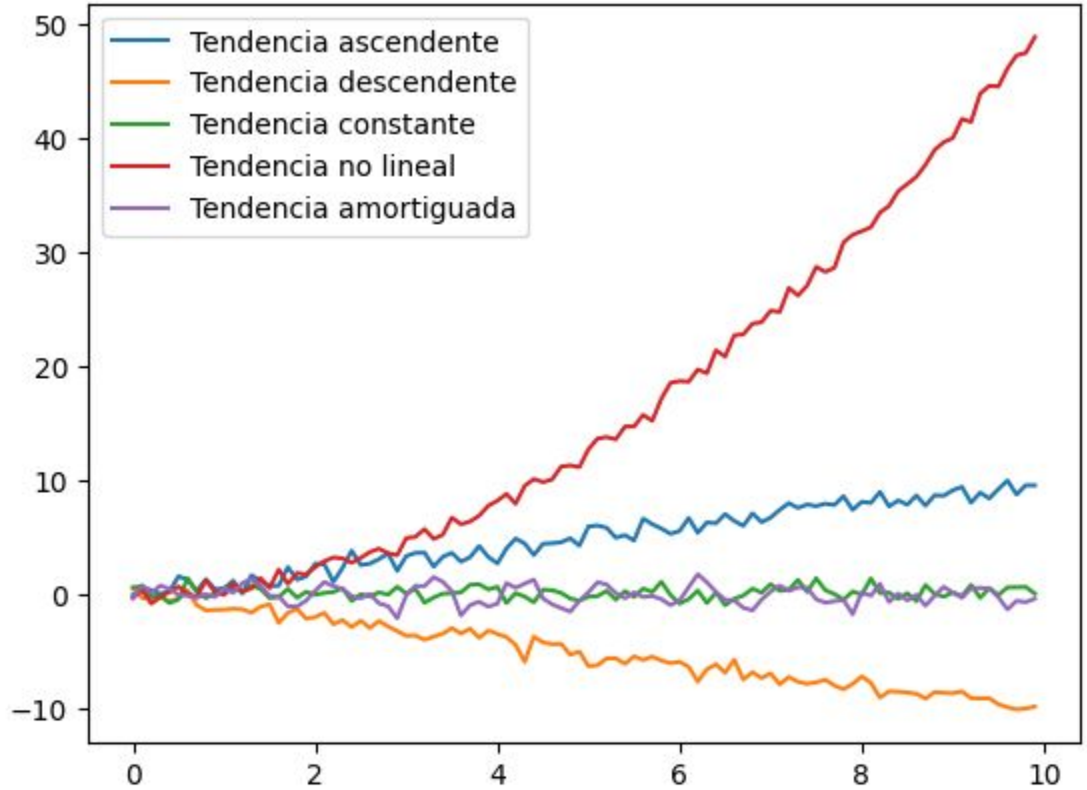
Estacionariedad estricta: igual distribución de probabilidad



Estacionariedad débil: Media, varianza y covarianza invariantes.

Tendencia

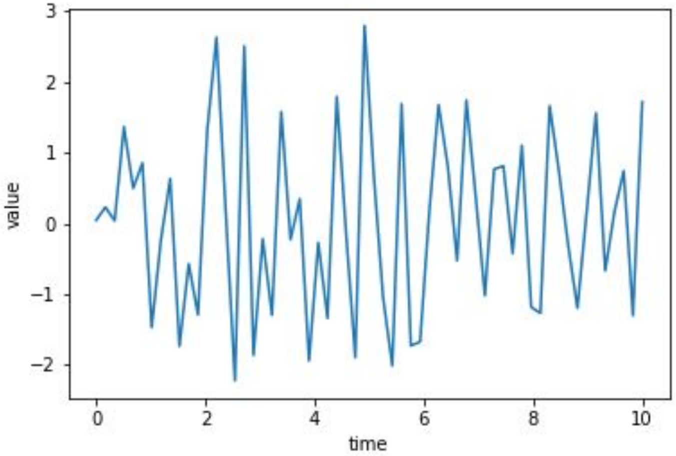
- Cambio a largo plazo en el nivel de la serie



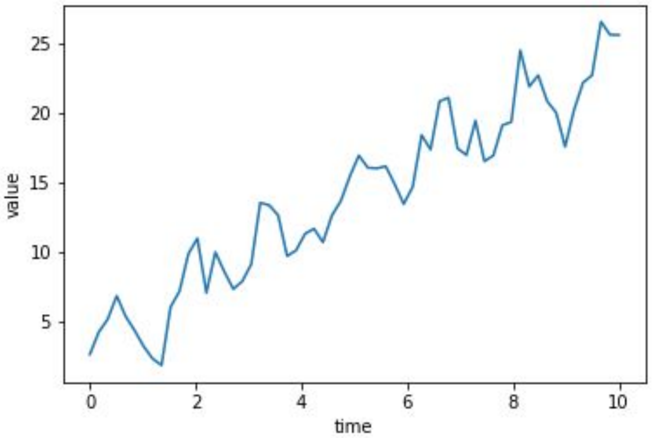
Estacionariedad

Serie estacionaria

- Tendencia constante



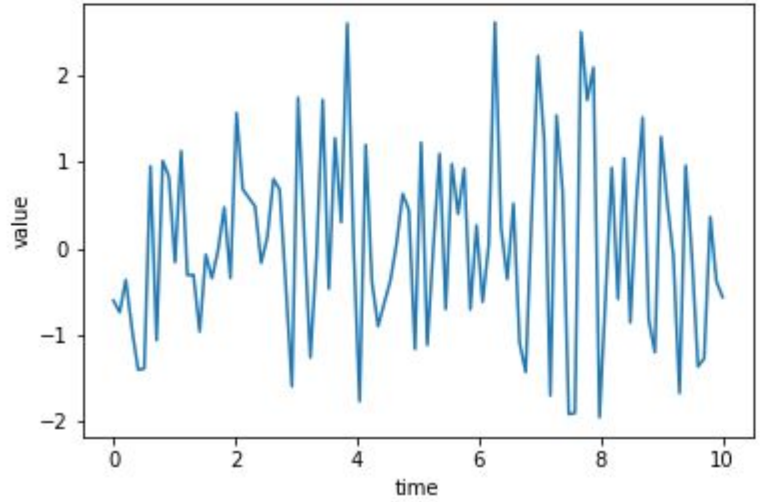
Serie no estacionaria



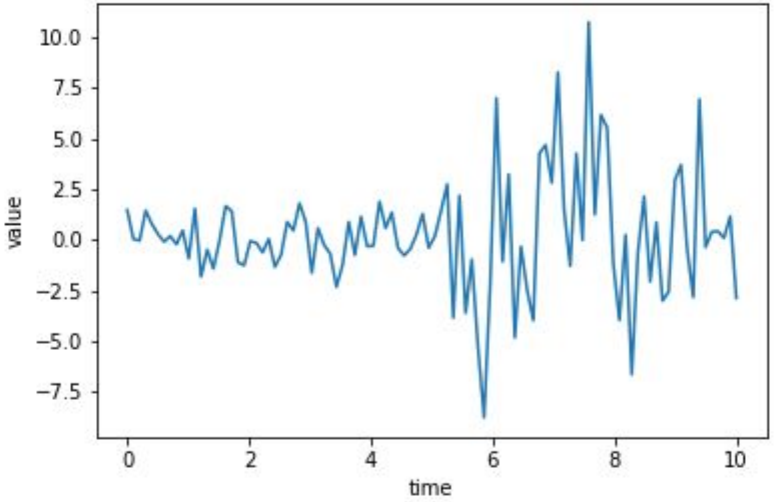
Estacionariedad

Serie estacionaria

- La media es constante
- La varianza es constante



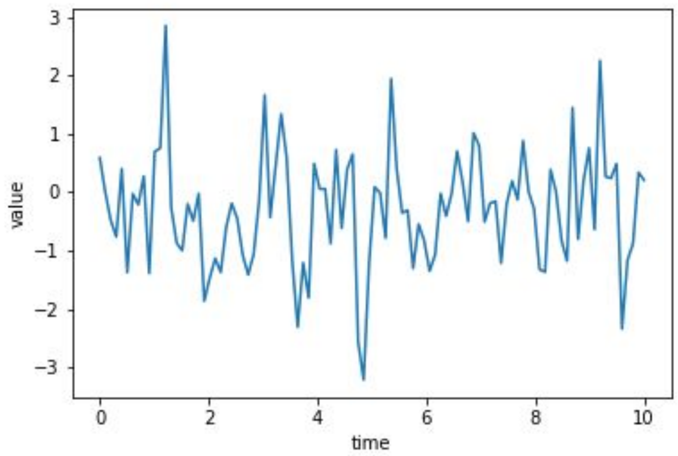
Serie no estacionaria



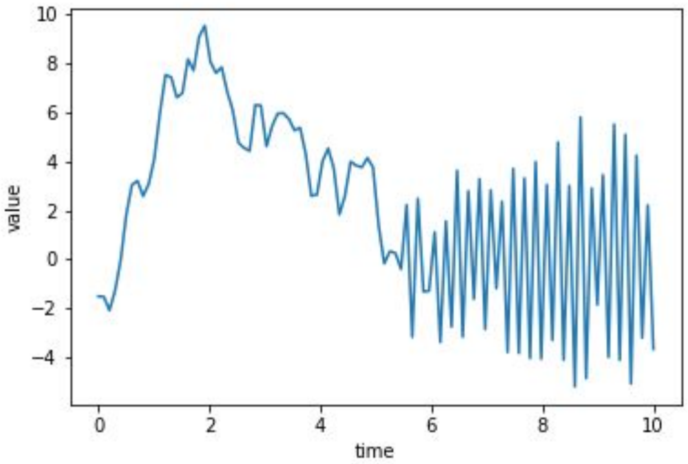
Estacionariedad

Serie estacionaria

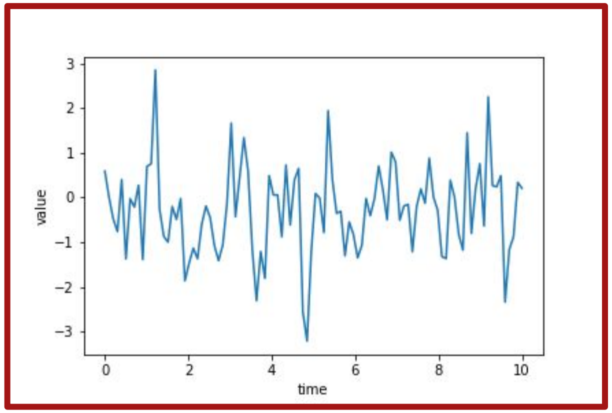
- La media es constante
- La varianza es constante
- La autocorrelación es constante



Serie no estacionaria



Definiciones



- Media de la serie

$$\mu(t) = E(x_t)$$

- Varianza de la serie:

$$\sigma^2(t) = E[(x_t - \mu(t))^2]$$

- Covarianza de la serie:

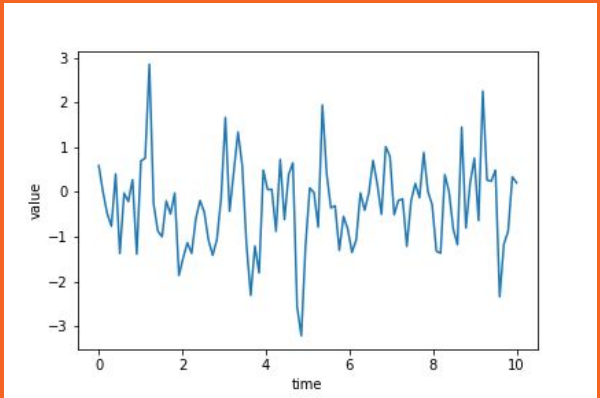
$$\gamma_h(t) = E[(x_t - \mu(t))(x_{t+h} - \mu(t+h))]$$

Estacionariedad

- La media, la varianza y la covarianza no dependen de t.

- Autocorrelación de retardo h

$$\rho_h = \frac{cov(x_t, x_{t+h})}{\sigma^2}$$



- $\mu(t) = \mu,$
- $\sigma^2(t) = \sigma^2$
- $\gamma_h(t) = \gamma_h.$

Ruido blanco

- Sin tendencia o tendencia constante.
- Media = 0
- Varianza constante
- Sus términos no están autocorrelacionados

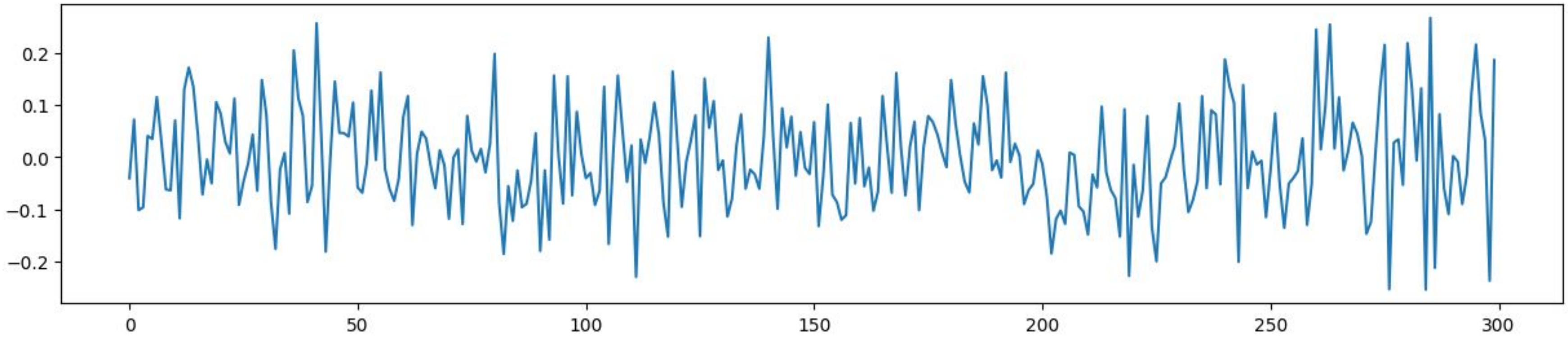
  0.1, -0.3, 0.8, 0.4, -0.5, 0.9, …

- Ruido blanco gaussiano: los valores provienen de una distribución normal $N(0, \sigma)$.

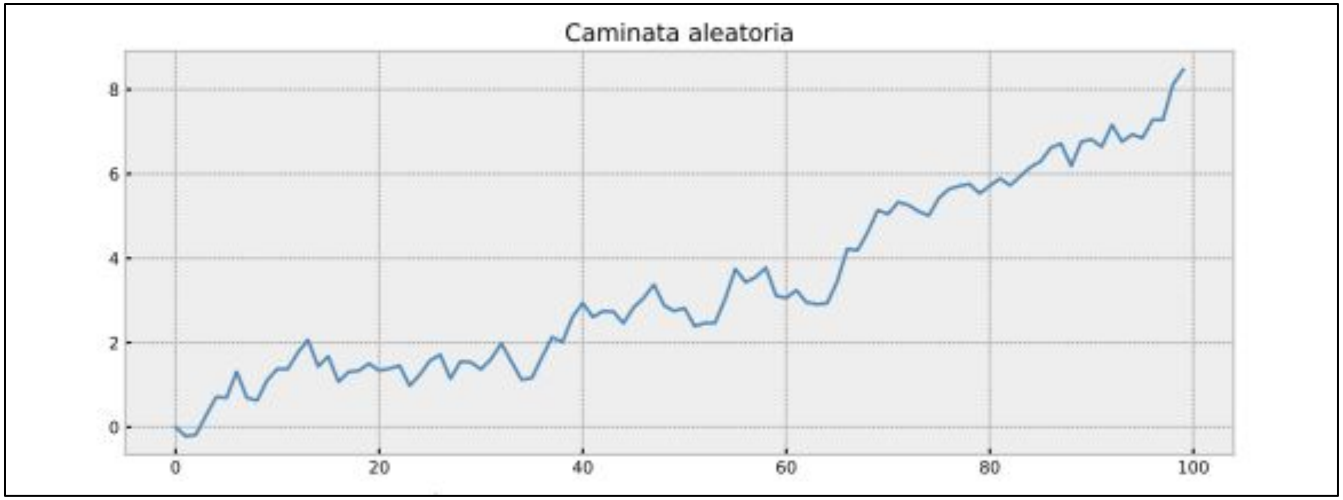
Construcción de un ruido blanco

```
import numpy as np
wn = np.random.normal(0, 0.1, 300)

plt.plot(wn)
```



Caminata aleatoria



- Es un proceso x_t que cumple:

$$x_t = x_{t-1} + a_t \quad \text{o también} \quad x_t = x_o + x_{t-1} + a_t$$

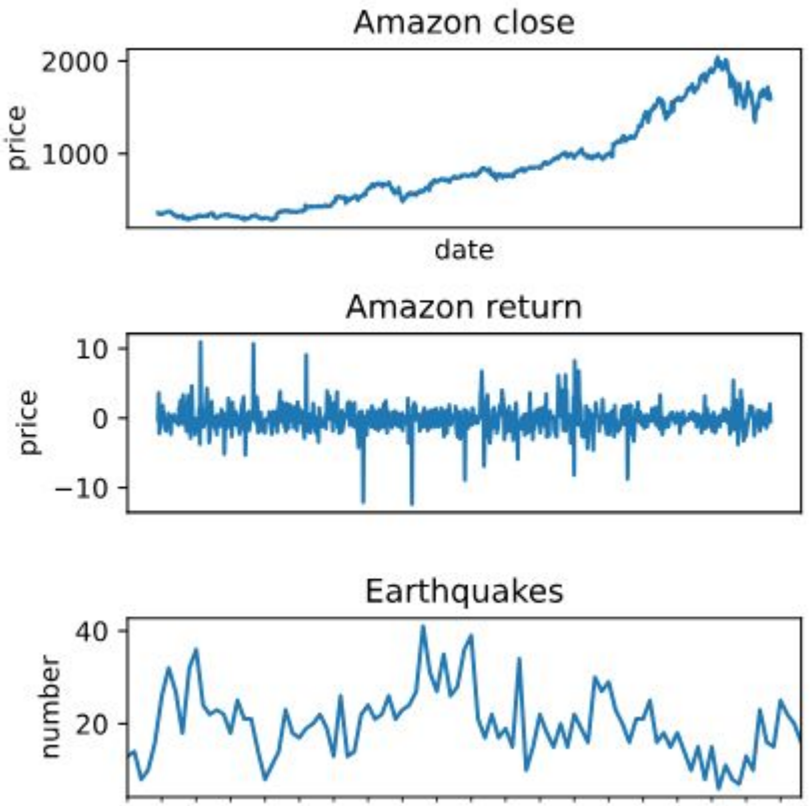
donde a_t es un ruido blanco de varianza σ.

- La varianza no es constante: $Var(x_t) = \sigma t^2$
- La tendencia es estocástica.

Actividades

Test de estacionariedad

Inspeccionando el gráfico de cada serie, ¿cuál respuesta sería adecuada a la siguiente pregunta?



La serie ¿es estacionaria?

- No, porque la varianza no es constante.
- Sí, parece ser estacionaria
- No, porque muestra una tendencia.
- No, porque muestra ciclos.

Tendencia

Trabajamos en la Actividad 1 de la Notebook.

```
spy = pd.read_csv('SPY.csv', index_col = 'Date', parse_dates = True)

eurusd = pd.read_csv('EURUSD.csv', index_col = 'Date', parse_dates = True)

vix = pd.read_csv('VIX.csv', index_col = 'Date', parse_dates = True)
```

continuamos …

Series no estacionarias

Para analizar estacionariedad de una serie, se puede aplicar el test estadístico ADF

- ADF: Augmented Dickey Fuller

- Testea si existe raíz unitaria: estacionariedad en tendencia

- Hipótesis nula: La serie no es estacionaria. Existe una raíz unitaria.

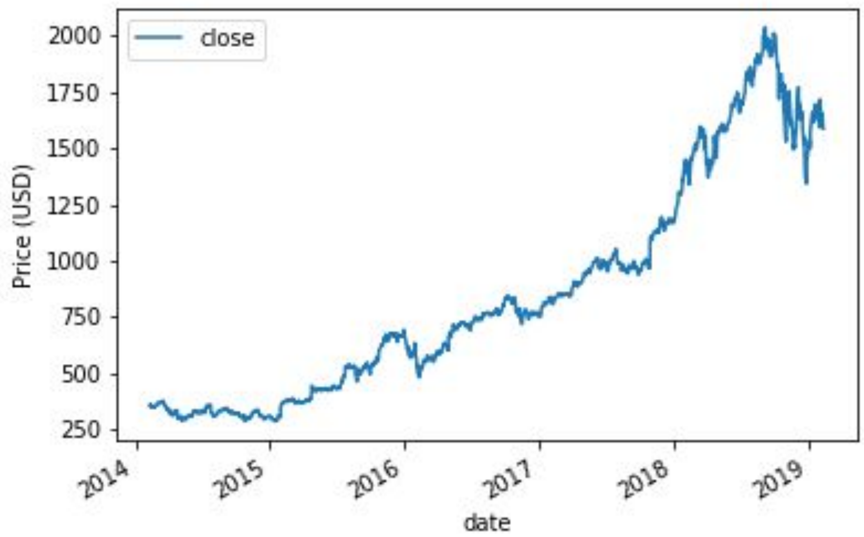
- Ejemplo de raíz unitaria: Caminata aleatoria

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

- En términos del operador de retardo $L\,x_t = x_{t-1}$:

$$\boxed{(1 - L)}\,\mathbf{x_t = a_t}$$

Test ADF




```
from statsmodels.tsa.stattools import adfuller
results = adfuller(prices["Sales"])
```

Interpretación del resultado

```
print(results)
```

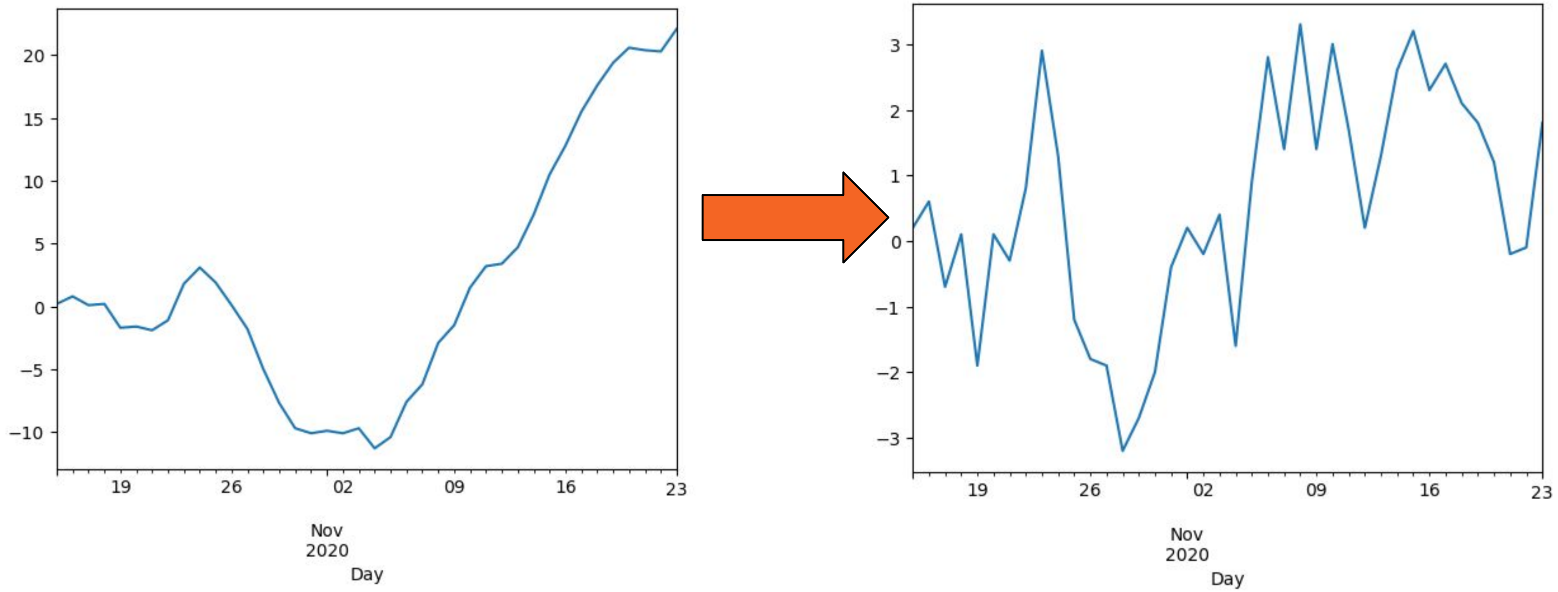
```
(-2.566067912904697, 0.10021507466490315, 0, 39, {'1%':
-3.610399601308181, '5%': -2.939108945868946, '10%':
-2.6080629651545038}, 95.7660871995601)
```

- Elemento 0: Valor del estadístico observado
- Elemento 1: p-valor de la prueba:
  - Si es pequeño se rechaza H0. Se rechaza la no estacionariedad.
- Elemento 4: valores críticos del test

https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html

Estacionarizar una serie

Existen distintas transformaciones para lograr la estacionariedad de una serie.



Diferenciar la serie

```
stationary_prices = prices.diff().dropna()
```

$$\Delta y_t = y_t - y_{t-1}$$

18.1 -12.3 -13.3 5.6

25.1, 43.2, 30.9, 17.6, 23.2 , ...

Otras transformaciones

- Logaritmo: `np.log(prices)`
- Cambio proporcional: `prices.shift(1) / prices`
- Cambio porcentual: `prices.pct_change()`

Actividades

Actividad

- Actividad 2 de la notebook

- Aplicamos Test ADF a algunas series.
- Aplicamos transformaciones a algunas series.

continuamos

ACF y PACF

- Tienen sentido para series (débilmente) estacionarias

  ACF: Función de autocorrelación

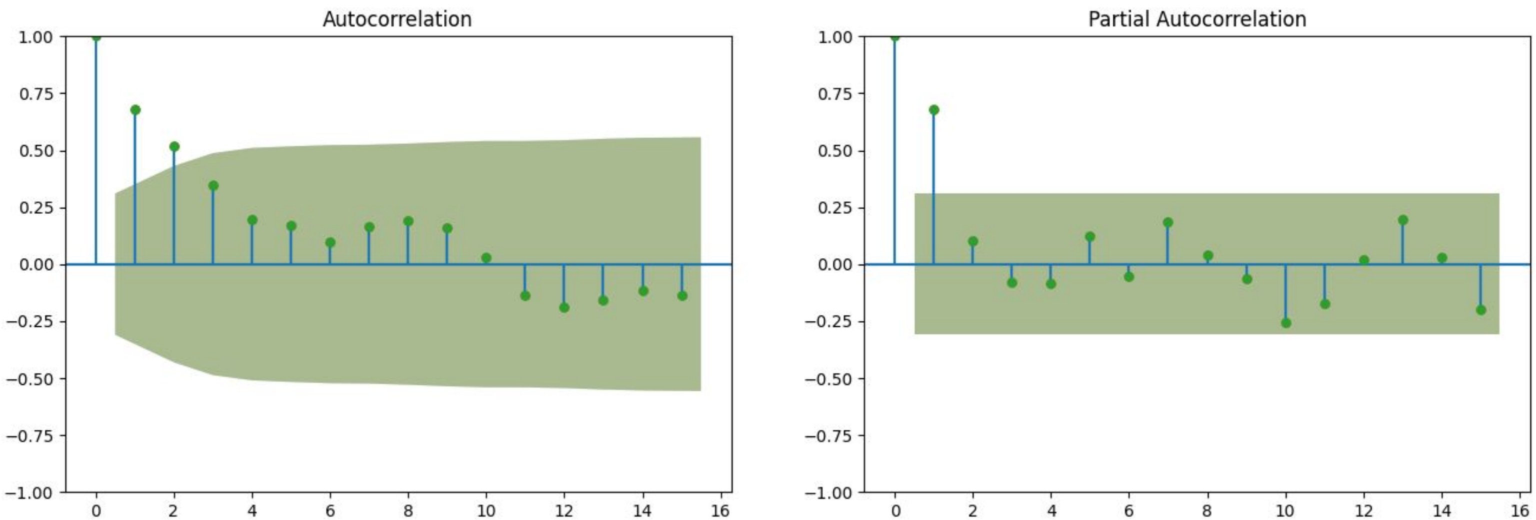
  - Calcula la correlación entre x_t y x_{t+h}

$$\boxed{\rho_h = \frac{cov(x_t, x_{t+h})}{\sigma^2}}$$

  PACF: Función de autocorrelación parcial.

  - Descarta la influencia de los valores intermedios: x_{t+1}, x_{t+2}, hasta $x_{t+(h-1)}$

ACF y PACF



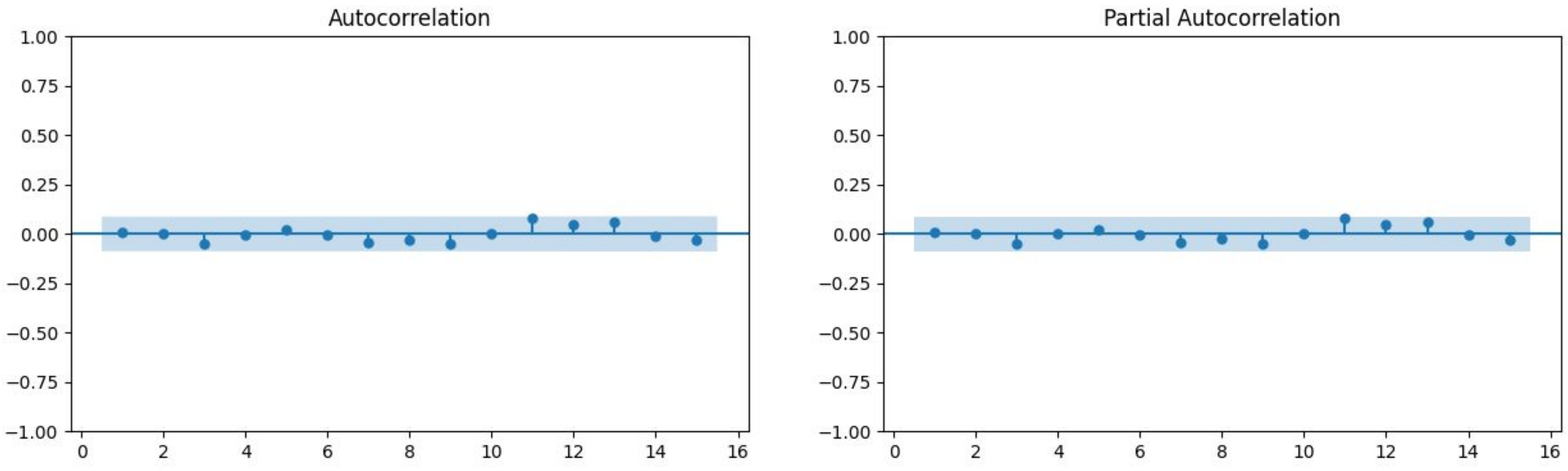

```
import statsmodels.api as sm

sm.graphics.tsa.plot_acf(prices.values.squeeze(), lags=15, ax = acf_ax, alpha = 0.05)
sm.graphics.tsa.plot_pacf(prices.values.squeeze(), lags=15, ax = pacf_ax, alpha = 0.05)
```

ACF y PACF de un ruido blanco

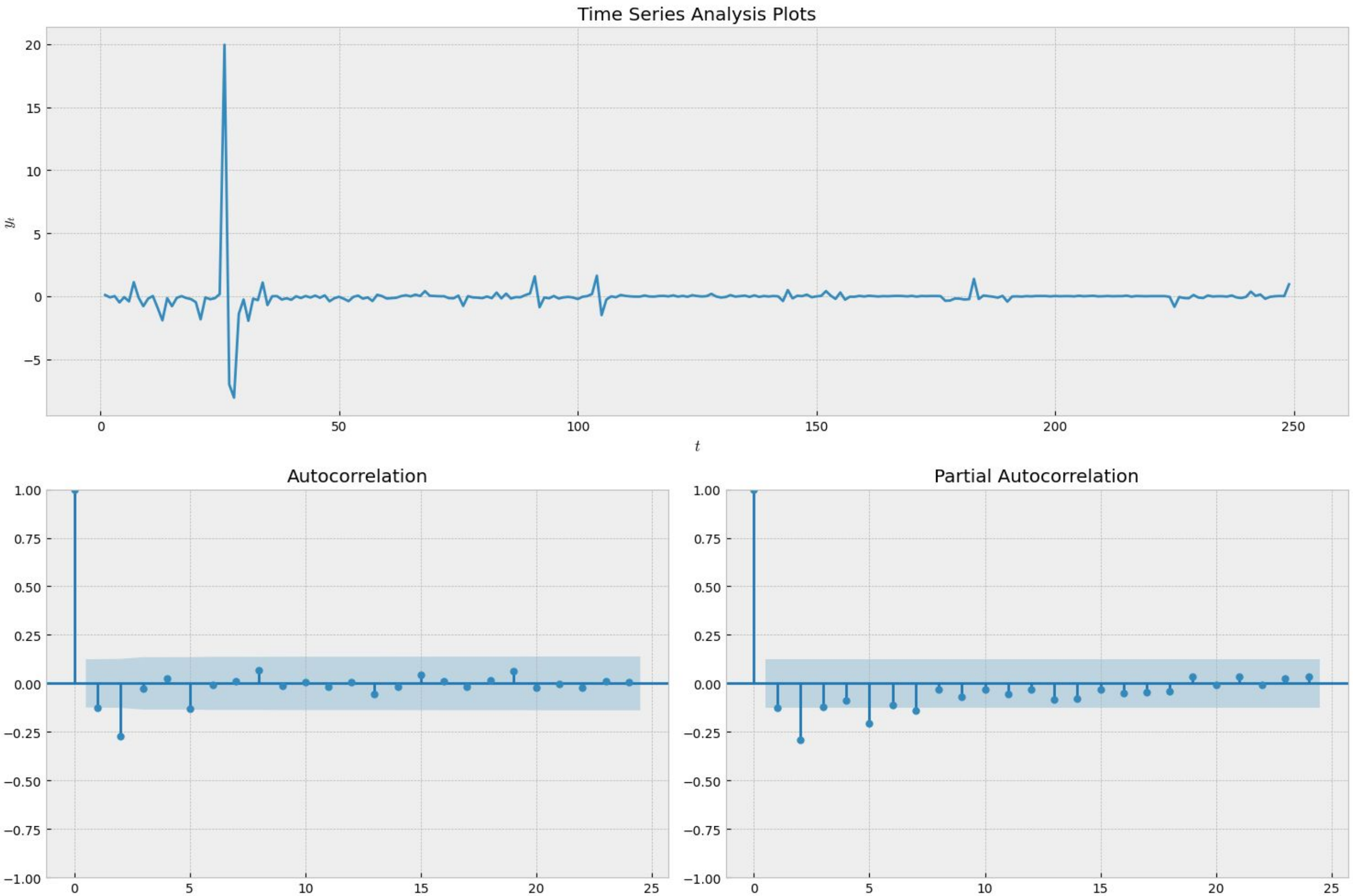
```
wn = np.random.normal(0,1, 500)

sm.graphics.tsa.plot_acf(wn, lags = 15, zero = False, ax = ax_acf);
sm.graphics.tsa.plot_pacf(wn,lags = 15, zero = False, ax = ax_pacf)
```



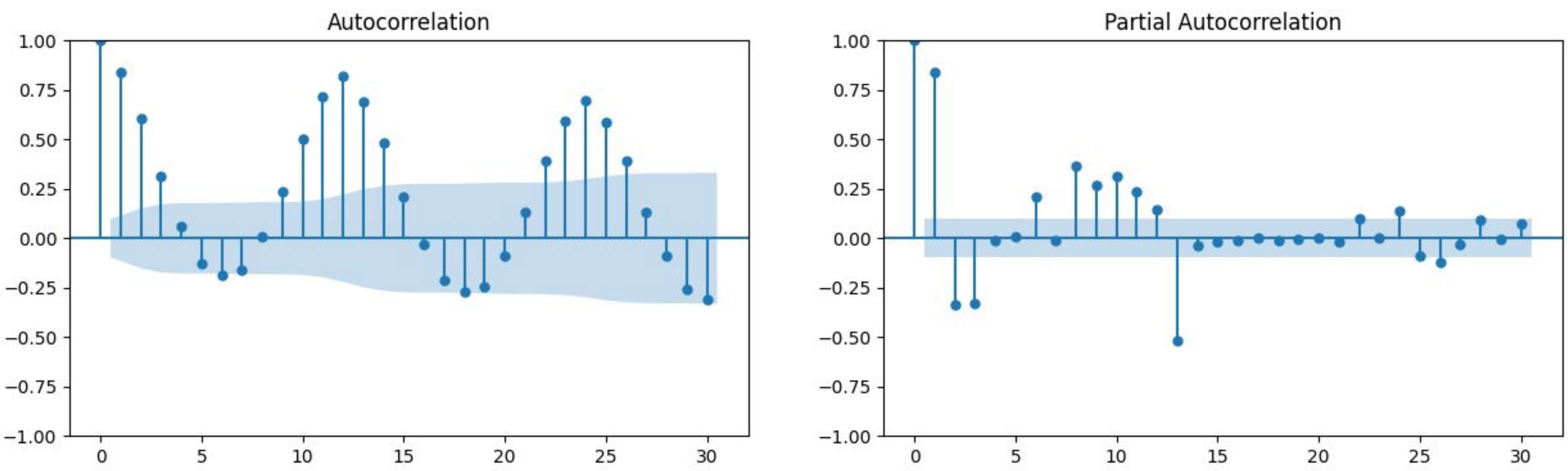
ACF y PACF de una serie de precios

AMTD Digital Inc. (HKD). Serie diferenciada.



ACF y PACF de una serie estacional

```
sm.graphics.tsa.plot_acf(candy["IPG3113N_20231017"], lags = 30,   ax = ax_acf)
sm.graphics.tsa.plot_pacf(candy["IPG3113N_20231017"],  lags = 30,  ax = ax_pacf)
```



Modelos autorregresivos

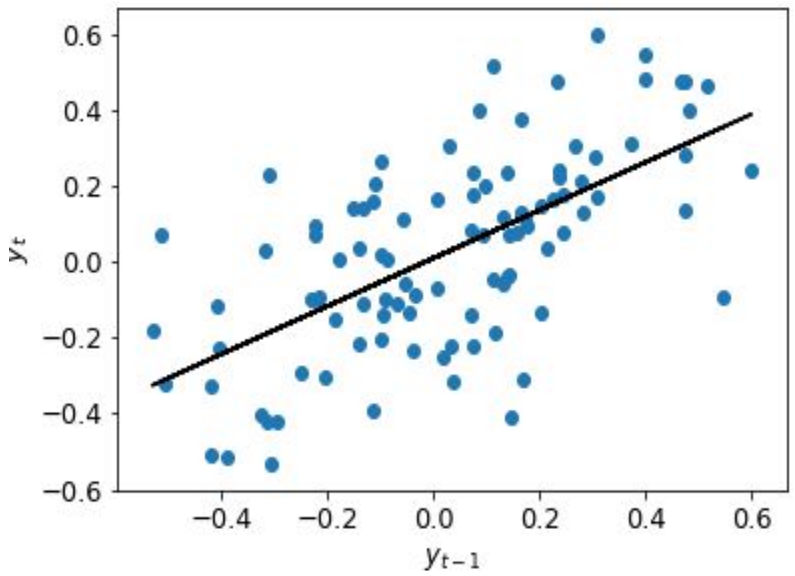
Modelos Autorregresivos

- Son modelos de series de tiempo
- Asumen estacionariedad de la serie
- Intentan capturar la dependencia con los retardos y con las innovaciones anteriores.

Modelos AR

Modelo Autorregresivo AR(1)

$$y_t = a_0 + a_1\, y_{t-1} + \varepsilon_t$$



Término constante

- Serie con término constante o intercepto:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

- La media de esta serie NO es a_0.

$$E[y_t] = a_0 + a_1 E[y_{t-1}] + E[\varepsilon_t]$$

$$E[y_t] = a_0 / (1 - a_1) = \mu$$

Término constante

- Se puede transformar en una serie sin término constante restando la media:

$$y_t - \mu = a_0 + a_1 (y_{t-1} - \mu) - (1 - a_1)\mu + \varepsilon_t$$

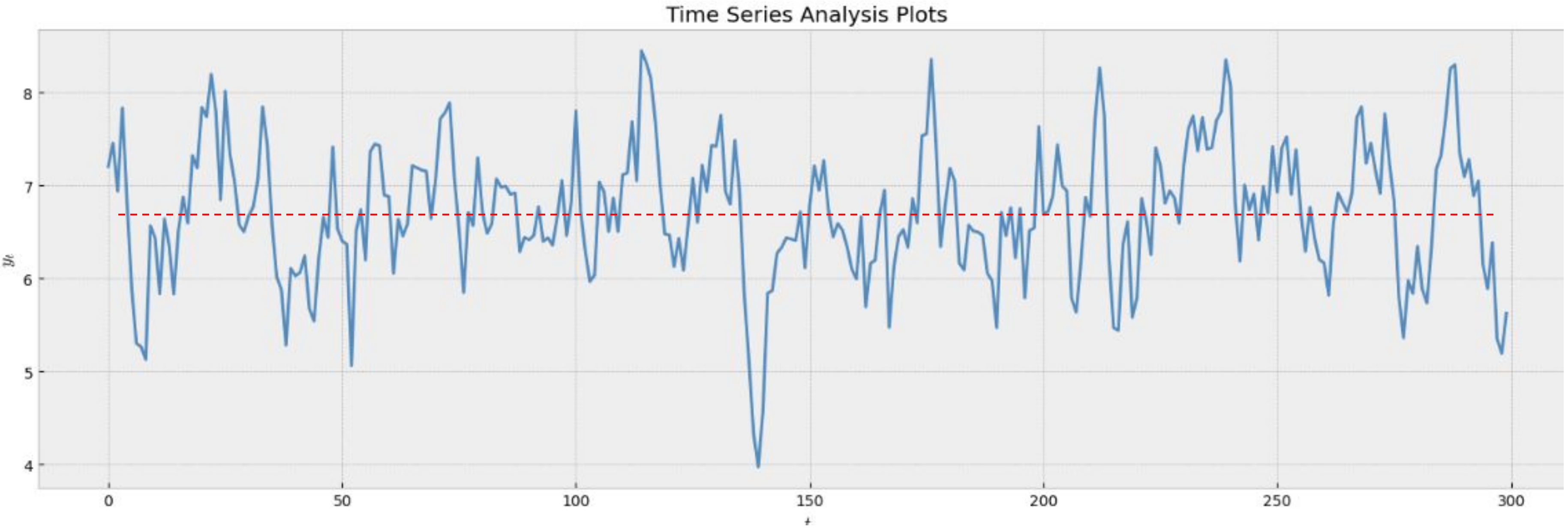
- La serie transformada es $x_t = y_t - \mu$

$$x_t = a_1 x_{t-1} + \varepsilon_t$$

Término constante

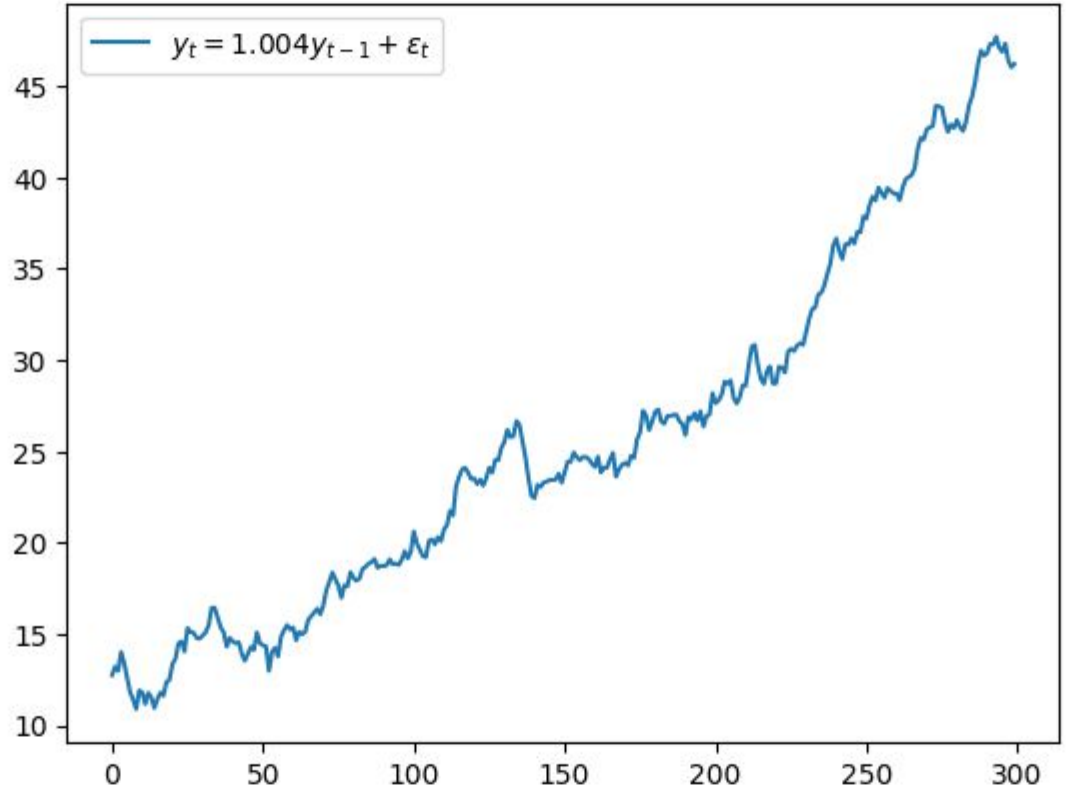
AR(1): $y_t = 2 + 0.7\, y_{t-1} + \varepsilon_t$

Media ≃ 6.67



Coeficiente del AR(1)

Modelo autorregresivo: $y_t = 1.004\, y_{t-1} + \varepsilon_t$



AR(1)

- El coeficiente a_1 interviene en la varianza de una serie AR(1) .
- Para que la serie sea estacionaria se requiere $|a_1| < 1$

$$y_t = a_1\, \mathbf{y_{t-1}} + \varepsilon_t = a_1\, \mathbf{(a_1\, y_{t-2} + \varepsilon_{t-1})} + \varepsilon_t = \ldots = \mathbf{(a_1)^t\, y_0} + \mathbf{(a_1)^{t-1}}\varepsilon_1 + \ldots + \mathbf{a_1}\varepsilon_{t-1} + \varepsilon_t$$

- La varianza de la serie AR(1) estacionaria es:

$$\boxed{\mathrm{Var}\,(y_t) = \mathrm{Var}(\varepsilon_t)\,/\,(1 - a_1^{\,2})}$$

Operador de retardo

- Si escribimos la ecuación para un AR(1) en términos del operador de retardo:

$$y_t - a_1 y_{t-1} = \varepsilon_t$$

$$(1 - a_1 L)\, y_t = \varepsilon_t$$

- La ecuación $1 - a_1 x = 0$ tiene solución con valor absoluto > 1 ($x = 1 / a_1$)

Modelos AR(p)

- Modelo AR(1)

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

- Modelo AR(2)

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$$

- Modelo AR(p)

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \ldots + a_p y_{t-p} + \varepsilon_t$$

Estacionariedad AR(p)

$$(1 - a_1 L - a_2 L^2 - a_3 L^3 - \ldots + a_p L^p)\, y_t = \varepsilon_t$$

- Estacionariedad: Las soluciones de $1 - a_1 x - a_2 x^2 - a_3 x^3 - \ldots - a_p x^p = 0$ deben tener módulo > 1.

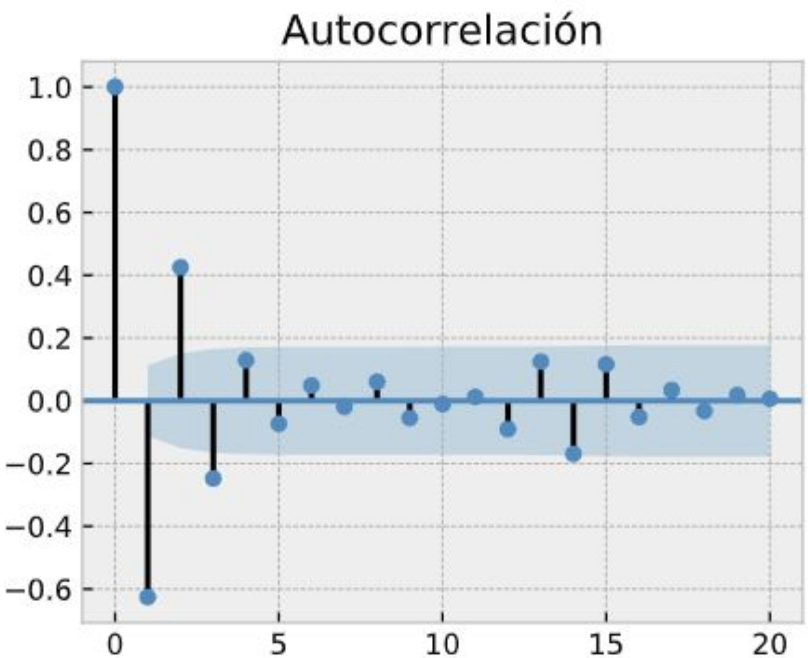
- AR(2):

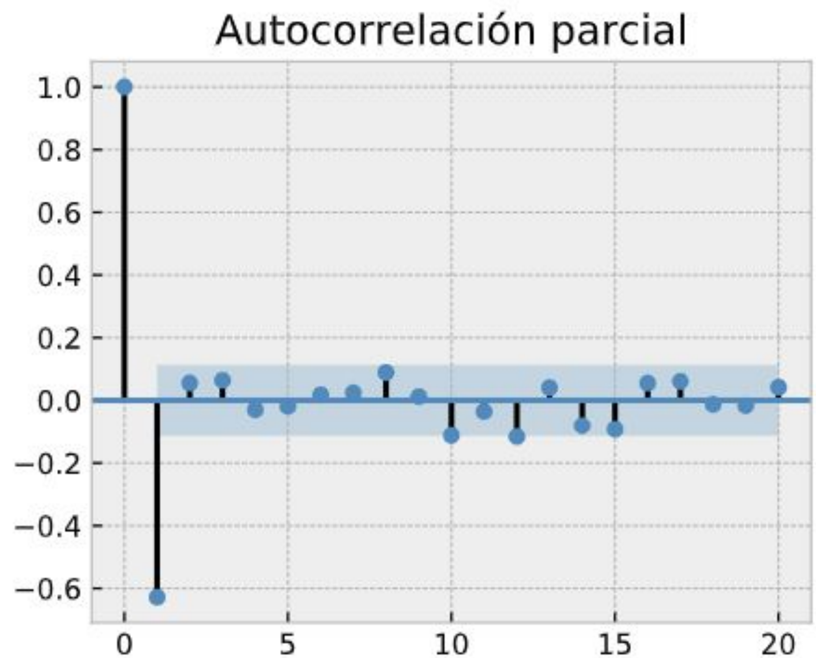
$$(1 - z_1 L)(1 - z_2 L)\, y_t = \varepsilon_t$$

ACF y PACF

En un modelo AR(p)

- El ACF no da mayor información
- Las autocorrelaciones parciales de orden p+1 en adelante se anulan.

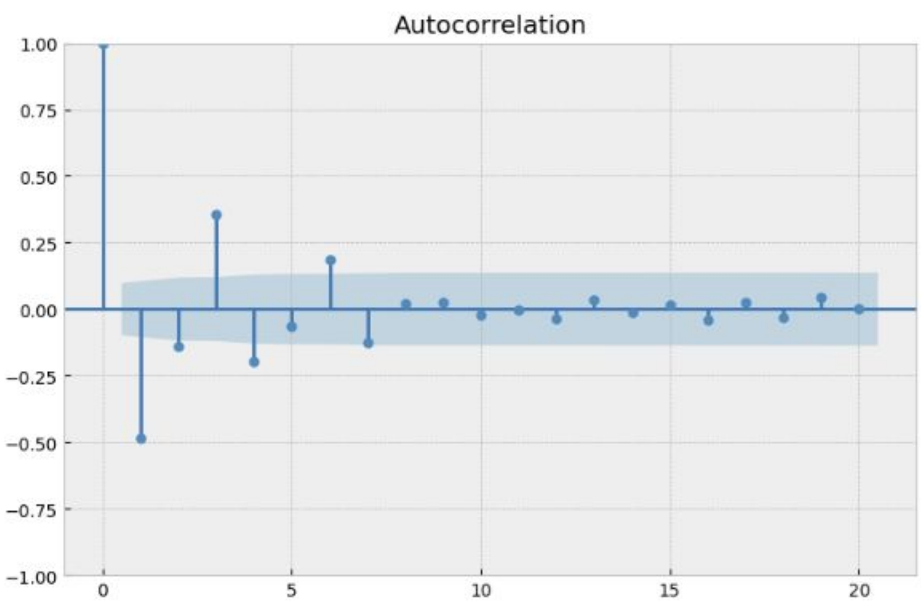


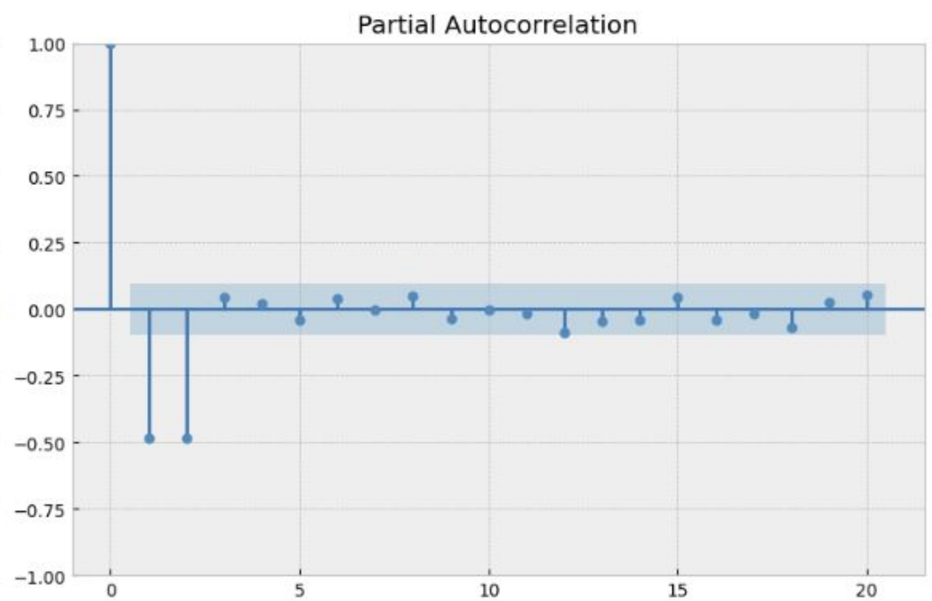


ACF y PACF

Modelo AR(2)

- El ACF no da mayor información
- Las <u>autocorrelaciones parciales</u> de orden 3 en adelante se anulan.

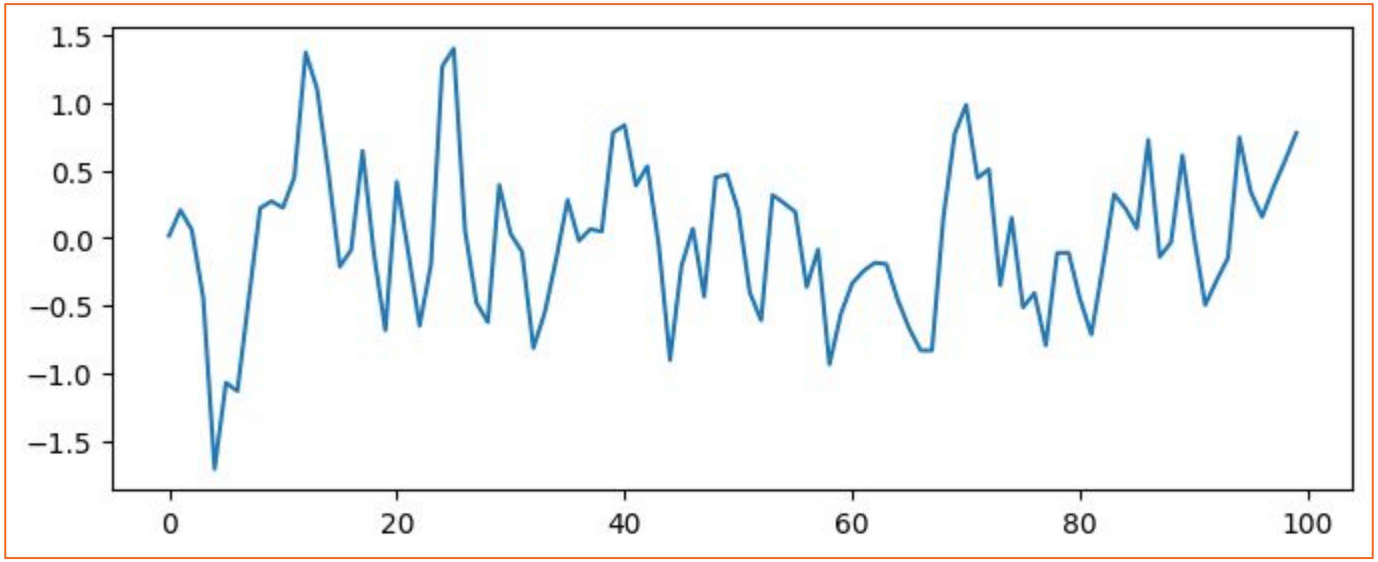




Creación de un AR

$$y_t = 0.5\, y_{t-1} + \varepsilon_t \longrightarrow y_t - 0.5\, y_{t-1} = \varepsilon_t$$

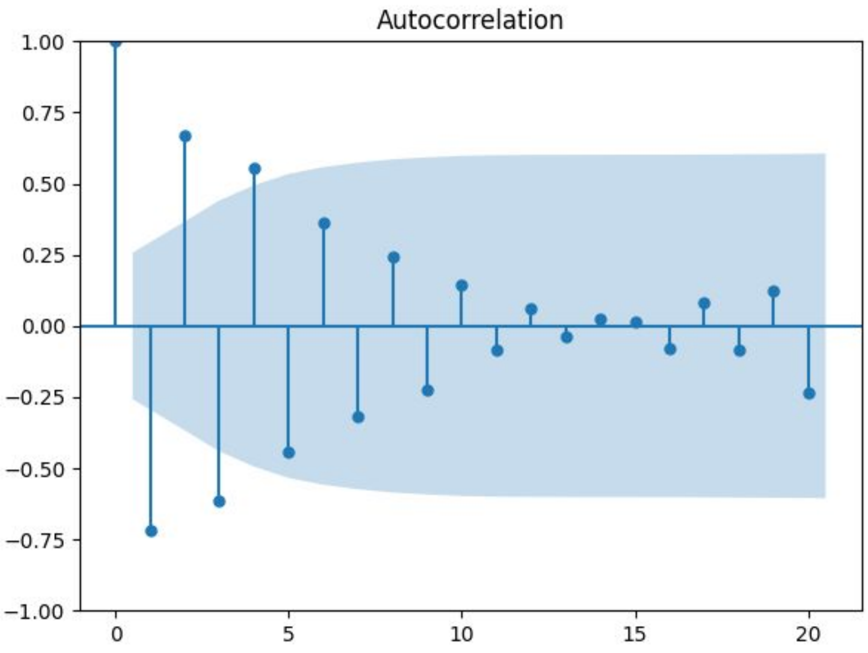
```
from statsmodels.tsa.arima_process import arma_generate_sample
ar_coefs = [1, -0.5]  ## Atención a los signos.
ma_coefs = [1]
serie = arma_generate_sample(ar_coefs, ma_coefs, nsample = 100, scale = 0.5)
```

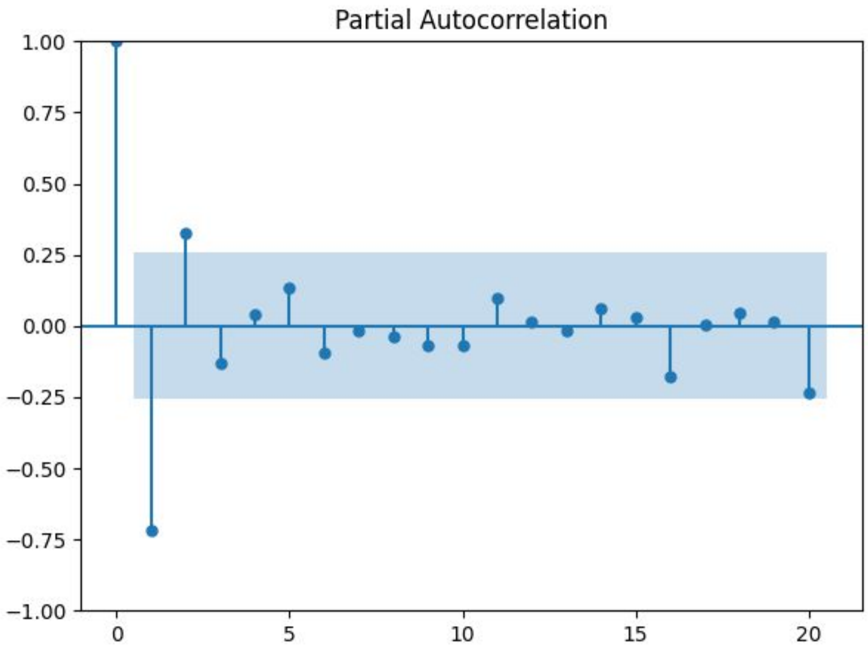


Creación de un AR(2)

$$y_t = 0.5\, y_{t-1} - 0.4\, y_{t-2} + \varepsilon_t$$

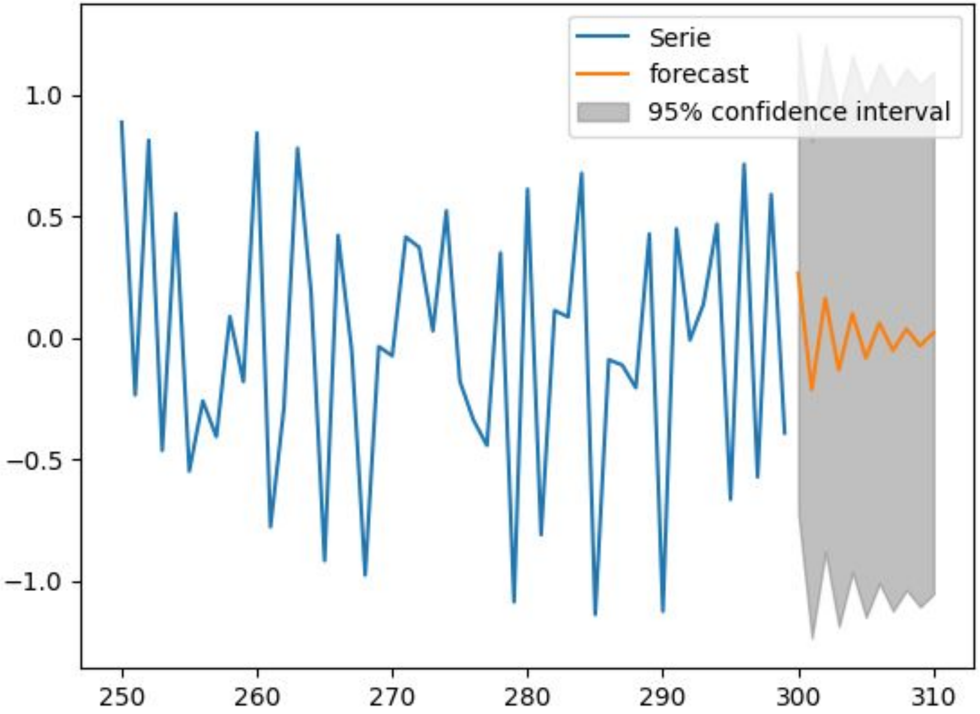
```
ar_coefs = [1, -0.5, 0.4]  ## Atención a los signos.
ma_coefs = [1]
serie = arma_generate_sample(ar_coefs, ma_coefs, nsample=100)
```





Predicción

$$\hat{y}_{t+1} = E[y_{t+1} \mid I(t)\,] = a_1\, y_t + a_2\, y_{t-1} + \ldots + a_p\, y_{t+1-p}$$



Modelos MA

- Moving Average: Modelos de media móvil

- Modelo MA(1)

$$y_t = \varepsilon_t + m_1 \varepsilon_{t-1}$$

- Modelo MA(2)

$$y_t = \varepsilon_t + m_1 \varepsilon_{t-1} + m_2 \varepsilon_{t-2}$$

- Modelo MA(q)

$$y_t = \varepsilon_t + m_1 \varepsilon_{t-1} + m_2 \varepsilon_{t-2} + m_3 \varepsilon_{t-3} + \ldots + m_q \varepsilon_{t-q}$$

Modelos MA(q)

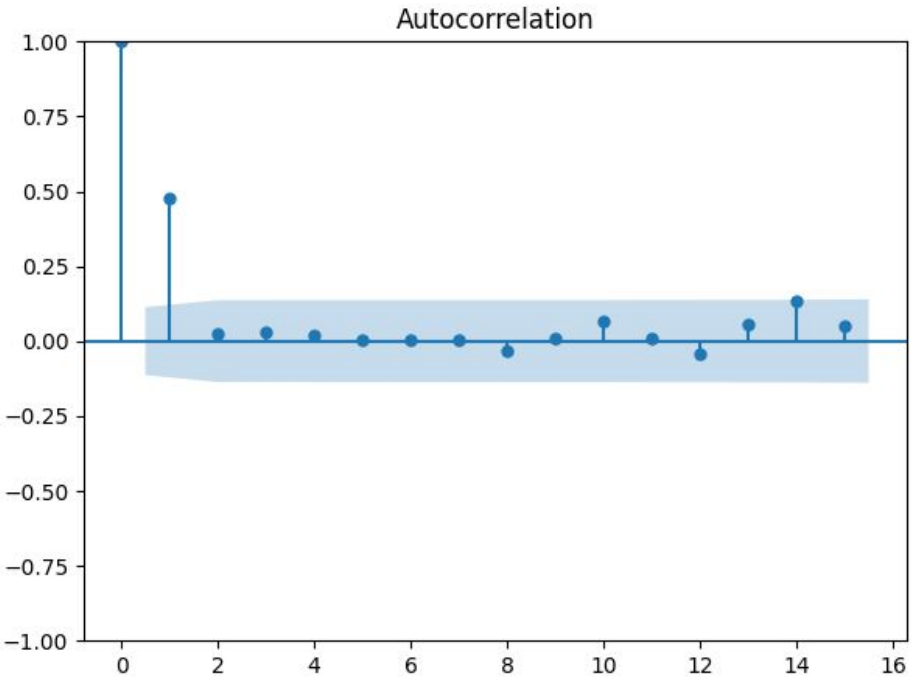
$$y_t = \varepsilon_t + m_1 \varepsilon_{t-1} + m_2 \varepsilon_{t-2} + m_3 \varepsilon_{t-3} + \ldots + m_q \varepsilon_{t-q}$$

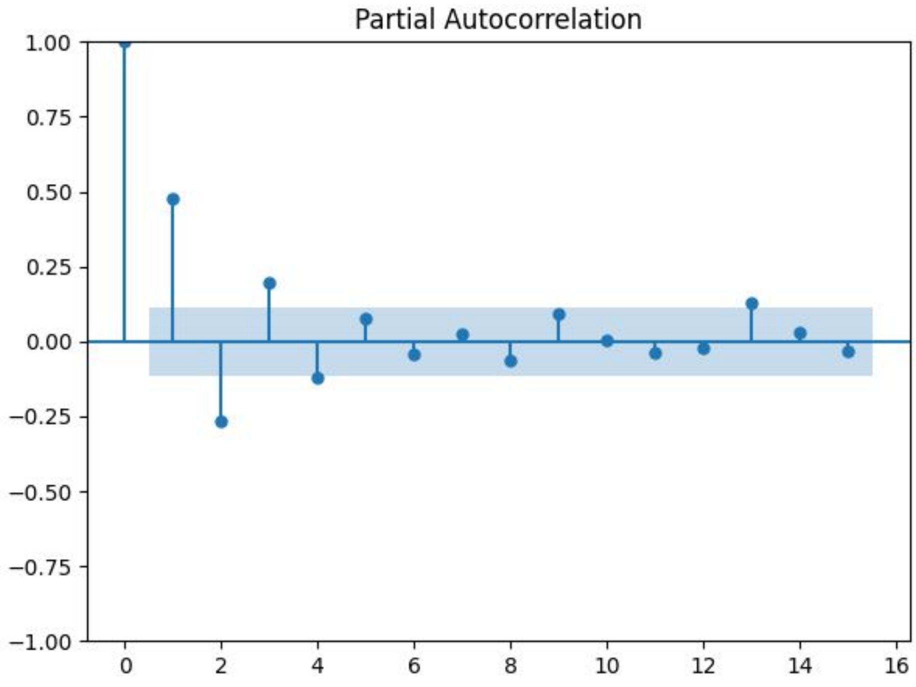
- **Estacionariedad**: Siempre son estacionarias
- **PACF:** Las autocorrelaciones parciales no brindan mayor información.
- **ACF:** $\rho_h = 0$ para $h > q$.

Creación de un MA(1)

$$y_t = \varepsilon_t + 0.7\,\varepsilon_{t-1}$$

```
ar_coefs = [1]
ma_coefs = [1, 0.7]
serie = arma_generate_sample(ar_coefs, ma_coefs, nsample=300)
```





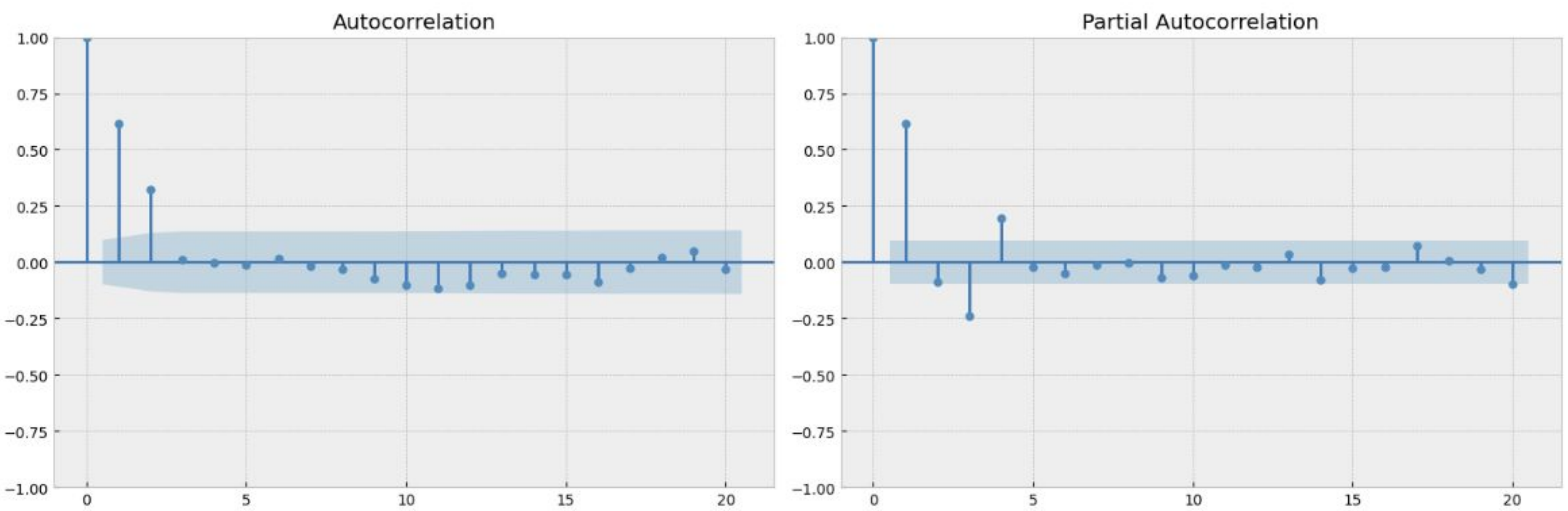
MA(2)

$$y_t = \varepsilon_t + 0.7\,\varepsilon_{t-1} + 0.5\,\varepsilon_{t-2}$$

MA(2):

ACF: se anula a partir del tercer retardo

PACF: no da mayor información.



Modelos ARMA

- Modelo ARMA(1, 1)

$$y_t = a_1 y_{t-1} + m_1 \varepsilon_{t-1} + \varepsilon_t$$

- Modelo ARMA(2,1)

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + m_1 \varepsilon_{t-1} + \varepsilon_t$$

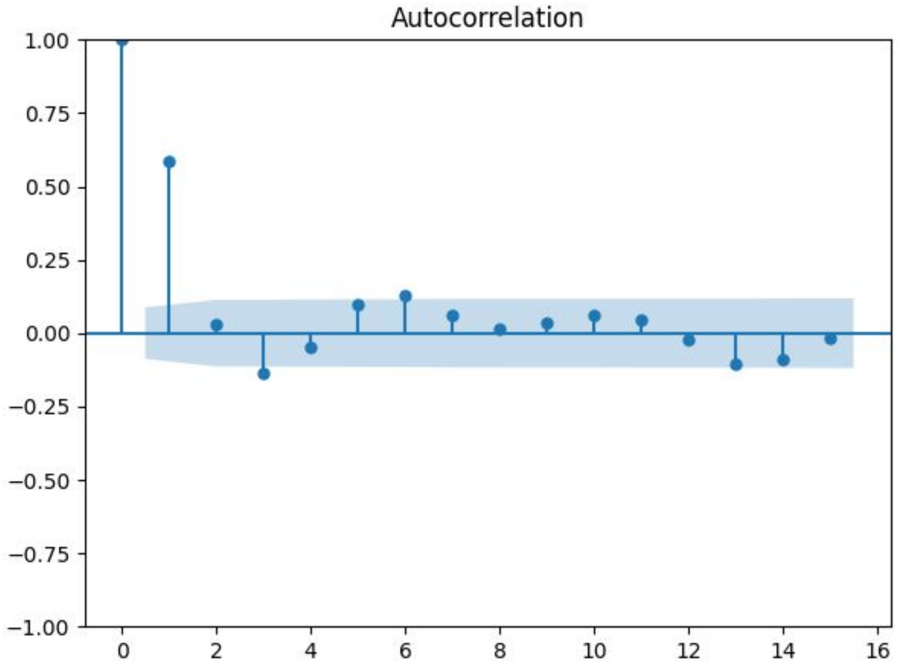
- Modelo ARMA (p,q)

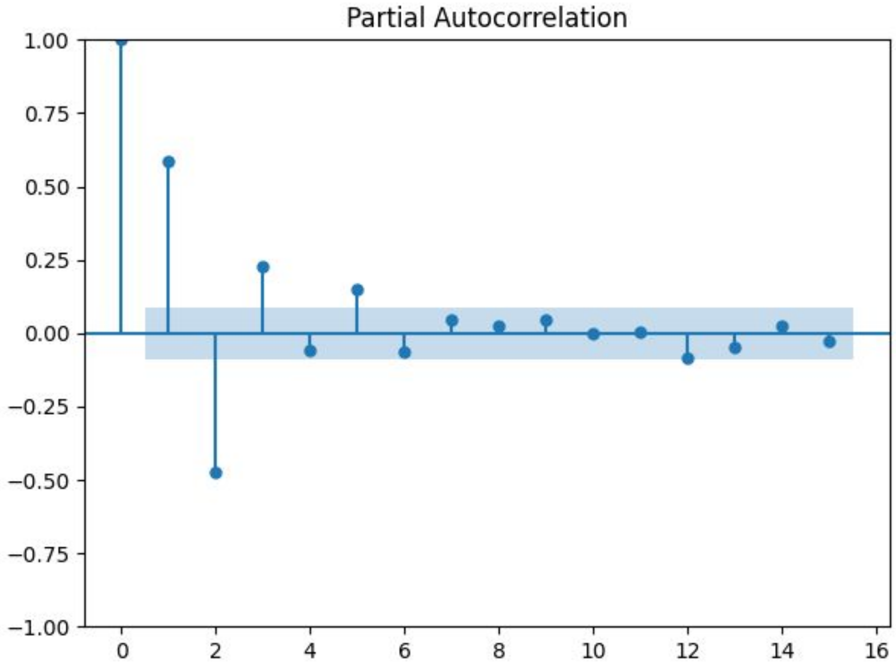
$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + m_1 \varepsilon_{t-1} + m_2 \varepsilon_{t-2} + \ldots + m_q \varepsilon_{t-q} + \varepsilon_t$$

Creación de un ARMA(2, 1)

$$y_t = 0.3\, y_{t-1} - 0.1\, y_{t-2} + 0.7\, \varepsilon_{t-1} + \varepsilon_t$$

```
ar_coefs = [1, -0.3, 0.1]
ma_coefs = [1, 0.7]
serie = arma_generate_sample(ar_coefs, ma_coefs, nsample=300)
```





Ajuste de un modelo ARMA

```
from statsmodels.tsa.arima.model import ARIMA

model = ARIMA(y, order=(2,0,1))
# Fit model
results = model.fit()


print(results.summary())
```

Resumen

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  500
Model:                 ARIMA(2, 0, 1)   Log Likelihood                -704.334
Date:                Wed, 15 Nov 2023   AIC                           1418.669
Time:                        12:32:38   BIC                           1439.742
Sample:                             0   HQIC                          1426.938
                                - 500
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1189      0.091     -1.308      0.191      -0.297       0.059
ar.L1          0.4456      0.062      7.235      0.000       0.325       0.566
ar.L2         -0.2246      0.055     -4.099      0.000      -0.332      -0.117
ma.L1          0.5603      0.055     10.146      0.000       0.452       0.669
sigma2         0.9774      0.058     16.785      0.000       0.863       1.091
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):                 6.05
Prob(Q):                              0.93   Prob(JB):                         0.05
Heteroskedasticity (H):               0.66   Skew:                             0.19
Prob(H) (two-sided):                  0.01   Kurtosis:                         3.38
===================================================================================
```

Resumen estadístico

- Datos generales de la serie
- Criterios de información: AIC, BIC, …
- Parámetros y el grado de "no significancia" estadística.
- Tests sobre los residuos:
    - Ljung Box + Prob(Q): La hipótesis nula es el ruido blanco.
    - Prob(Q) pequeño: rechaza que sea un ruido blanco.

    - Heteroscedasticidad + Prob(H): La hipótesis nula es varianza constante.
    - Prob(H) pequeño: rechaza que la varianza sea constante.

    - Jarque Bera: testea la normalidad de los residuos

https://analyzingalpha.com/interpret-arima-results

Actividades

Actividades

¿Qué modelo se está generando con las siguientes instrucciones?

```
ar_coefs = [1, 0.4, -0.1]
ma_coefs = [1, 0.2]
serie = arma_generate_sample(ar_coefs, ma_coefs, nsample = 100,
scale = 0.5)
```

1. ARMA(3,2)
2. ARMA(2,3)
3. ARMA(1,2)
4. ARMA(2,1)

Actividades

¿Qué modelo se está generando con las siguientes instrucciones?

```
ar_coefs = [1, -0.4]
ma_coefs = [1]
serie = arma_generate_sample(ar_coefs, ma_coefs, nsample = 100)
```

1. $y_t = \varepsilon_t + 0.4\,\varepsilon_{t-1}$
2. $y_t = 0.4\,y_{t-1}$
3. $y_t = 0.4\,y_{t-1} + \varepsilon_t$
4. $y_t = -0.4\,y_{t-1} + \varepsilon_t$

Actividades

¿Cuál de las siguientes es equivalente a un modelo AR(1)?

1. ARMA(1,0)
2. ARMA(0,1)
3. MA(1)
4. ARMA(1,1)

Continuamos…

—

Series integradas

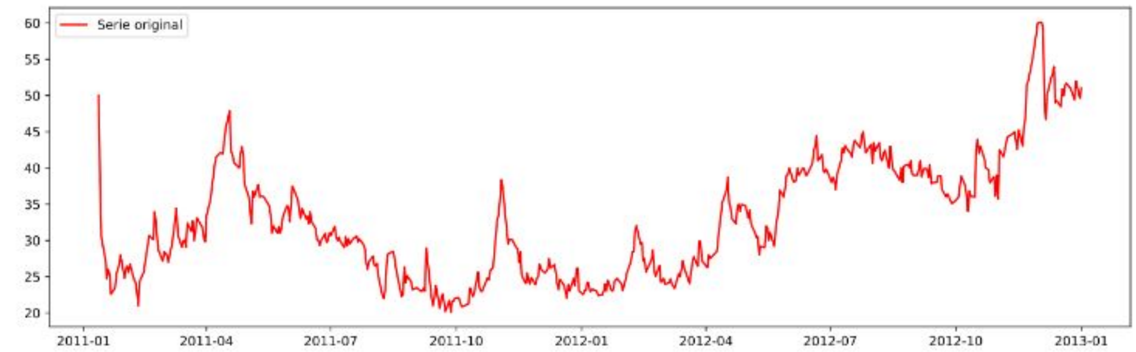
- El gráfico muestra una tendencia a largo plazo.

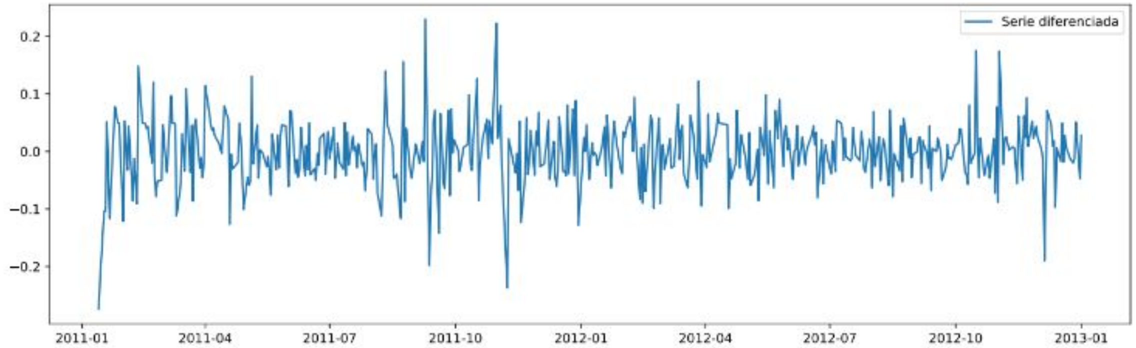
- Al diferenciar una vez la serie resulta estacionaria.

$$\Delta y_t = y_t - y_{t-1}$$

- Si Δy_t es ARMA(p, q) entonces

Modelo ARIMA(p,1,q)





Doble diferenciación

$$(x_t) = [5, 4, 6, 7, 9, 12]$$

- Primera diferenciación:

$$(\Delta x_t)' = [-1, 2, 1, 2, 3]$$

- Segunda diferenciación:

$$(\Delta(\Delta x_t))' = [3, -1, 1, 1]$$

- No es lo mismo que diferenciar con el segundo retardo: $y_t - y_{t-2}$

$$(x_t - L^2 x_t)' = [1, 3, 3, 5]$$

Construcción de una serie integrada

$$(x_t) = [5, 4, 6, 7, 9, 12]$$

- Serie integrada:

$$[5,\ 5 + 4,\ 5 + 4 + 6, 5 + 4 + 6 + 7, ...]$$

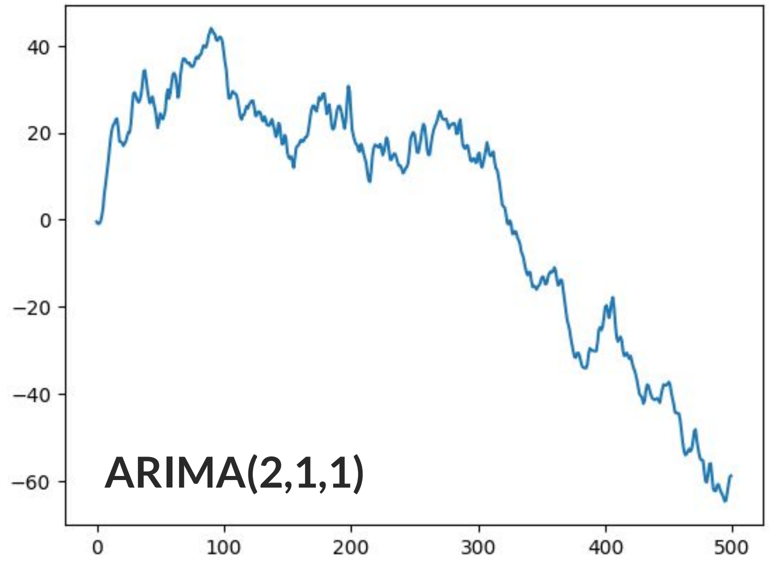
$$[5, 9, 15, 22, 31, 43]$$

- Diferenciamos:

$$[4, 6, 7, 9, 12]$$

Construcción de una serie integrada

```
serie_arima_1 = serie_arma.cumsum()
plt.plot(integrada)
```



Pequeña síntesis

ARIMA(0,0,0)+c	Modelo con media constante
ARIMA(0,1,0):	Caminata aleatoria
ARIMA(0,1,0) + c:	Caminata aleatoria con tendencia. (MGB)
ARIMA(1,0,0) + c:	Regresión con el retardo de primer orden
ARIMA(2,0,0) + c:	Regresión con el retardo de primer orden y segundo orden
ARIMA(1,1,0) + c:	Regresión de Δy_t con el retardo de primer orden
ARIMA(2, 1,0) + c:	Regresión de Δy_t con el retardo de primer orden y el de segundo orden

Pequeña síntesis

ARIMA(0,1,1)	Suavizado simple exponencial
ARIMA(0,1,1) + C	Suavizado simple exponencial más tendencia
ARIMA(1,1,2)	Suavizado lineal exponencial con tendencia amortiguada
ARIMA(0,2,2)	Suavizado lineal exponencial

Ejemplo: ARIMA (0,1,1):

$$y_t - y_{t-1} = \Delta y_t = \varepsilon_t + m_1 \varepsilon_{t-1}$$

$$\hat{y}_t = y_{t-1} + m_1 \varepsilon_{t-1} = (1 - m_1)\, y_{t-1} + m_1 \hat{y}_{t-1}$$

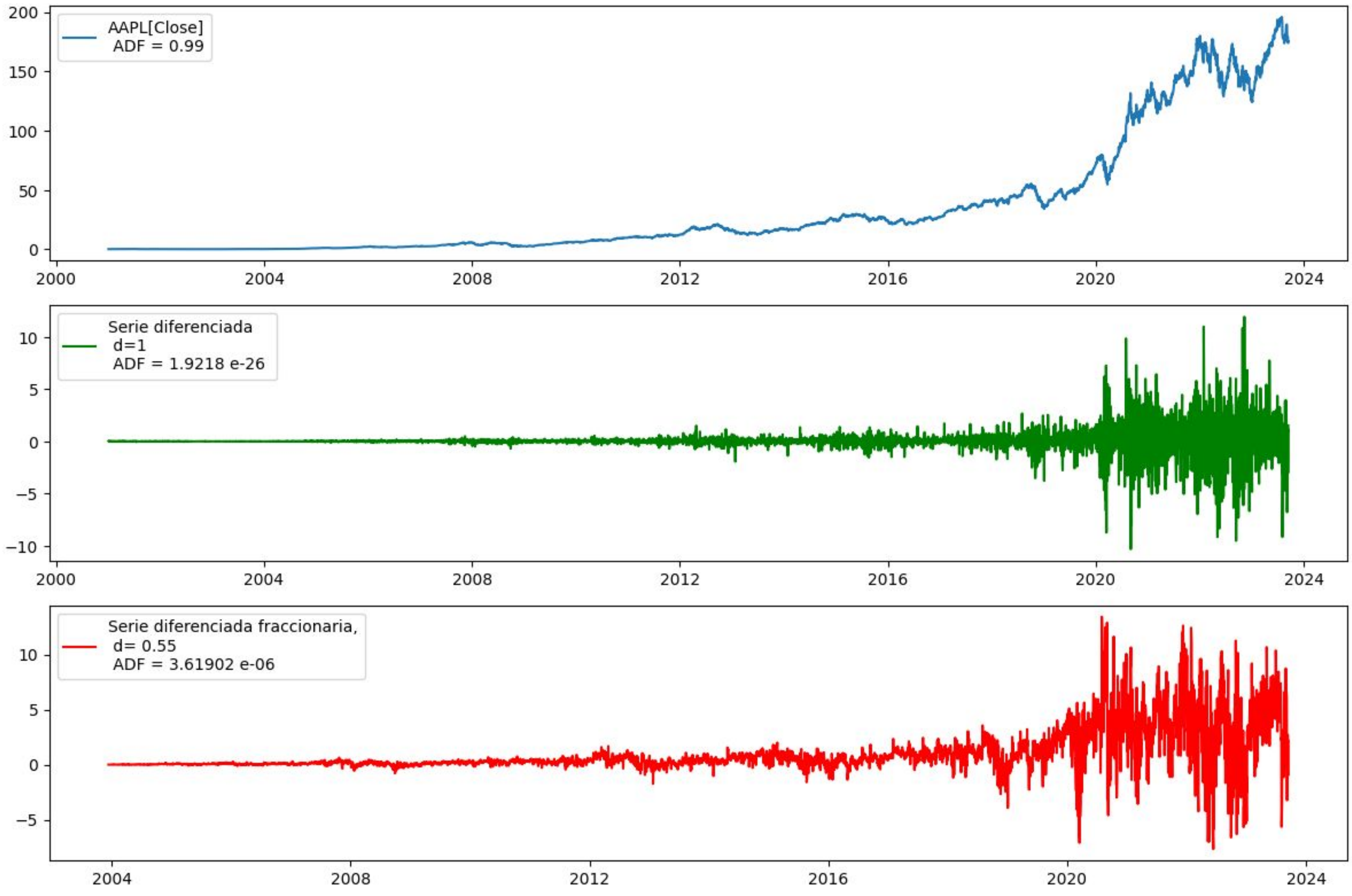
https://people.duke.edu/~rnau/411arim.htm

Modelo ARFIMA

- Series de tiempo que rechazan ADF pero deja dudas.
- Hay una alta persistencia en las correlaciones.
- Puede ocurrir que al diferenciarlas, o aplicarles alguna transformación, se pierdan propiedades.
- Nuevo modelo: ARFIMA

FI: Fraccionario integrado

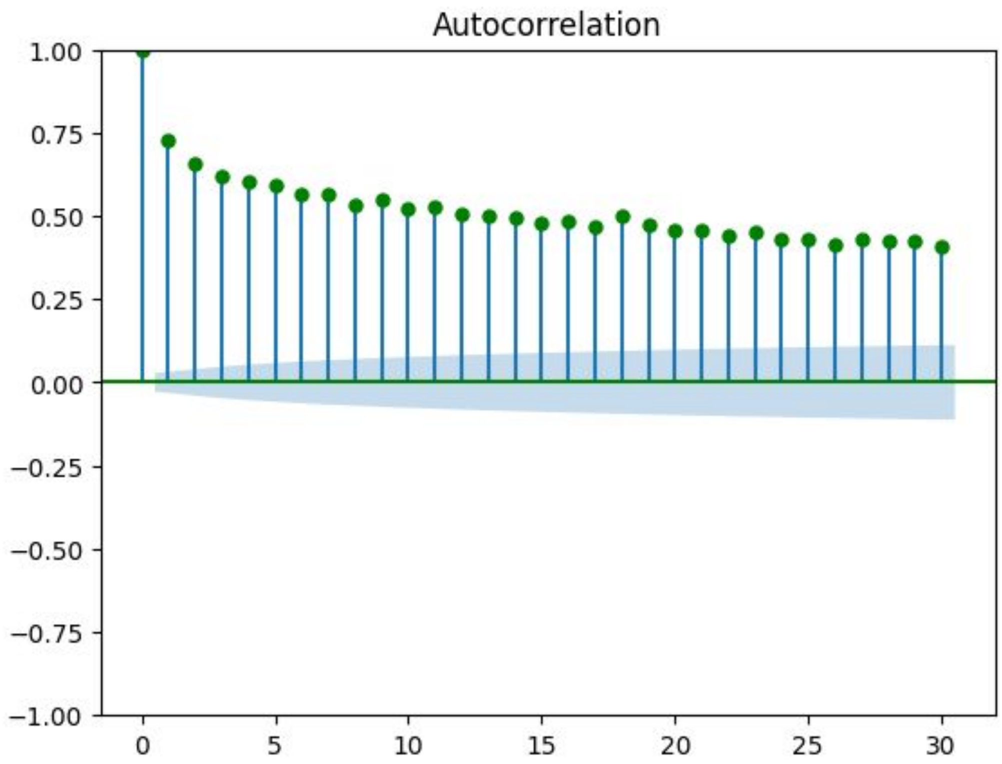
ARFIMA



AAPL[Close]
ADF = 0.99
200
150
100
50
0
2000
2004
2008
2012
2016
2020
2024
Serie diferenciada
d=1
ADF = 1.9218 e-26
10
5
0
−5
−10
Serie diferenciada fraccionaria,
d= 0.55
ADF = 3.61902 e-06

ARFIMA

La función de autocorrelación ACF muestra alta persistencia.



Modelo ARFIMA

- Diferenciación de orden d:

$$(1 - L)^d y_t = \varepsilon_t$$

- Si d no es entero:

$$(1 - L)^d y_t = y_t - d\, y_{t-1} + d(d-1)/2!\, y_{t-2} - d(d-1)(d-2)/3!\, y_{t-3} + \ldots$$

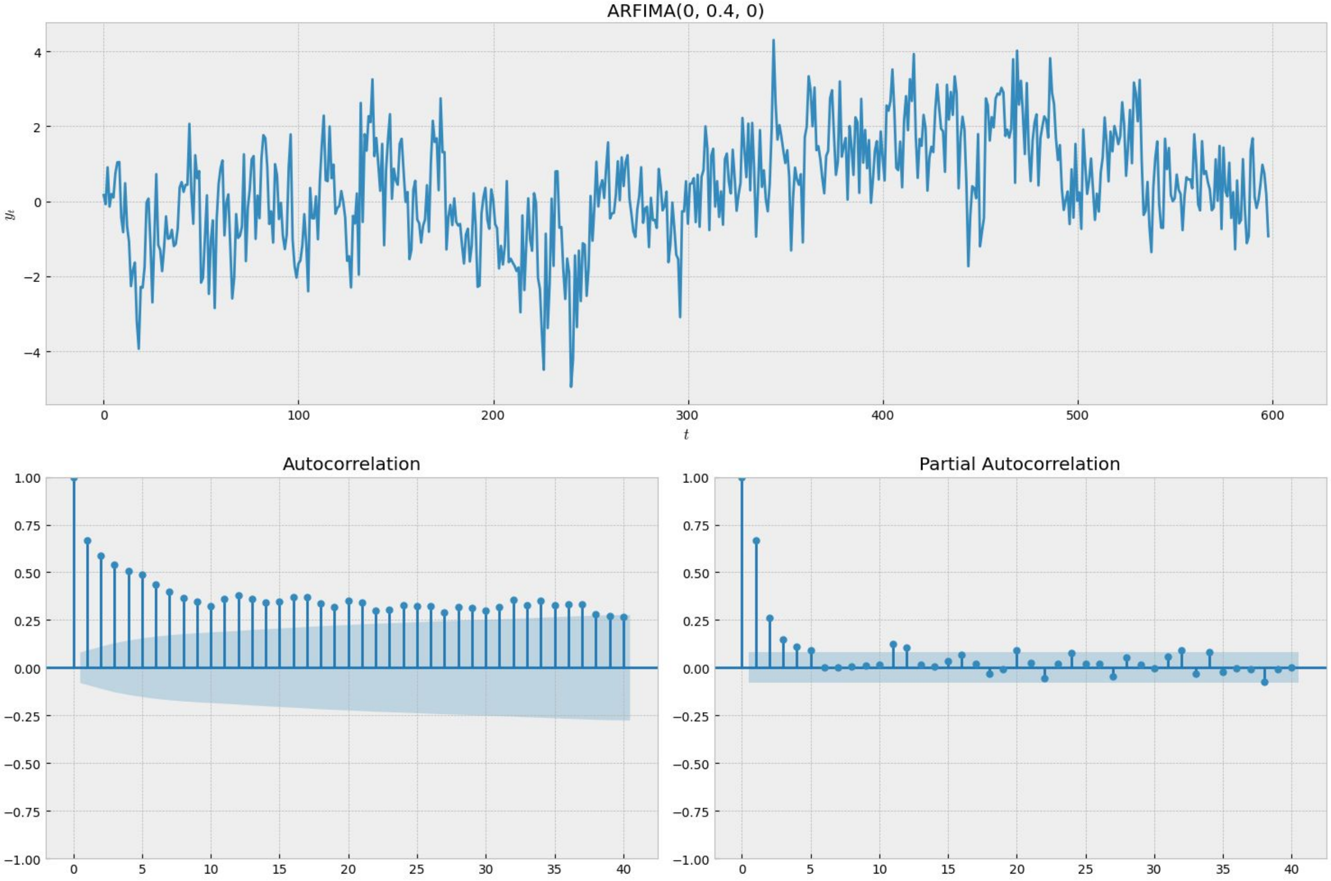
- Si d es entero (por ejemplo d= 2):

$$(1 - L)^d y_t = y_t - 2\, y_{t-1} + y_{t-2}$$

- El proceso y_t es **estacionario** si $|d| < 0.5$.

Modelo ARFIMA

d < 1: La serie es estacionaria



Modelos ARCH y GARCH

Origen

- Una vez hecho el ajuste de una serie financiera a un modelo ARIMA, ¿cómo es el comportamiento de los residuos?
- En algunos casos “parecen” un ruido blanco, pero no lo son. Su varianza no es constante.
- Las series financieras no suelen tener una volatilidad constante.

$$log(P_t / P_{t-1}) = ARIMA + \varepsilon_t \quad \Longleftrightarrow \quad (P_t - P_{t-1}) / P_{t-1} = ARIMA + \varepsilon_t$$

Volatilidad

¿Qué es la volatilidad?

- Describe la dispersión de los retornos de un activo a lo largo del tiempo.
- Se suele calcular a través del desvío estándar o la varianza de los retornos de un activo.
- Volatilidad histórica. Volatilidad implícita.
- A mayor volatilidad, mayor riesgo del activo.
- A movimientos bruscos de precios le siguen períodos de alta volatilidad.

Cálculo de la volatilidad histórica

1. Calculamos el retorno de los activos:

$$Retorno_i = \frac{P_{i+1} - P_i}{P_i}$$

2. Calculamos la media muestral de los retornos

$$Media\ \mu = \frac{\sum_{i=1}^{T} Retorno_i}{T}$$

3. Calculamos el desvío estándar muestral de los retornos.

$$Volatilidad = \sqrt{\frac{1}{T-1}\sum_{i=1}^{T}(Retorno_i - \mu)^2}$$

Cálculo de la volatilidad histórica

```
rets = data["Adj Close"].pct_change()

print(rets.dropna())

volatility_aapl = rets.std()
```

308.233124 x 1.000510 = 308.390228

Date		Date	
2022-01-03	328.727600		
2022-01-04	323.090942	2022-01-04	-0.017147
2022-01-05	310.688171	2022-01-05	-0.038388
2022-01-06	**308.233124**	2022-01-06	-0.007902
2022-01-07	**308.390228**	2022-01-07	**0.000510**

Volatilidad

- Si los datos son diarios, obtenemos retornos diarios.

- Para pasar de volatilidad diaria a mensual usamos:

$$\sigma_{mensual} = \sqrt{21}\,\sigma_{diario}$$

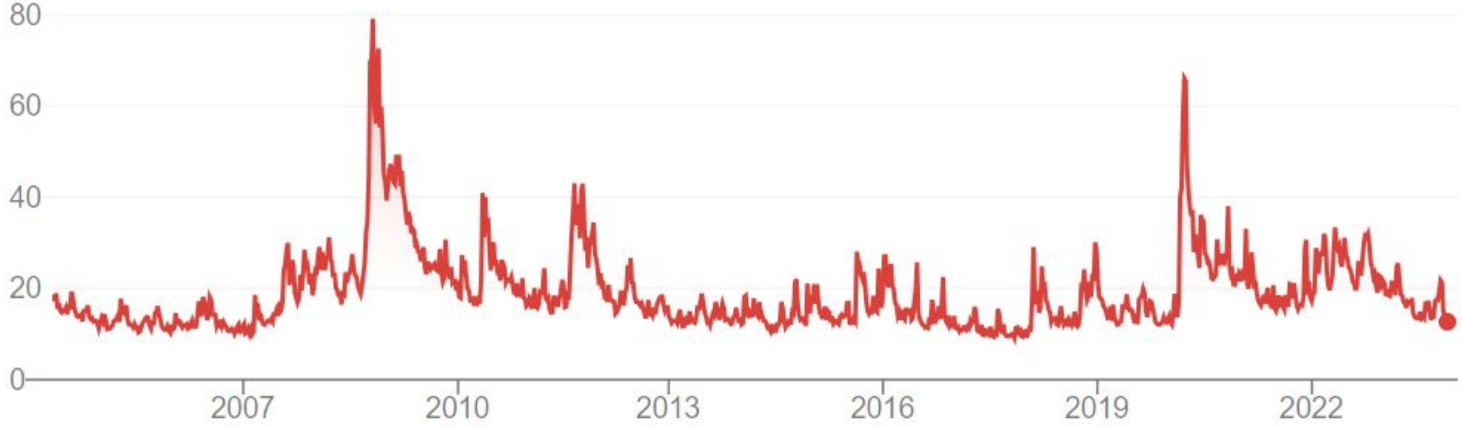
- Para pasar de la volatilidad diaria a la anual usamos:

$$\sigma_{anual} = \sqrt{252}\,\sigma_{diario}$$

Modelos de volatilidad

- Heterocedasticidad
  - Hetero (distinto)
  - Skedasis (dispersión)
- Las series de tiempo pueden mostrar cambios en la volatilidad a lo largo del tiempo.

Clusters de Volatilidad



Actividades

Modelos ARCH y GARCH

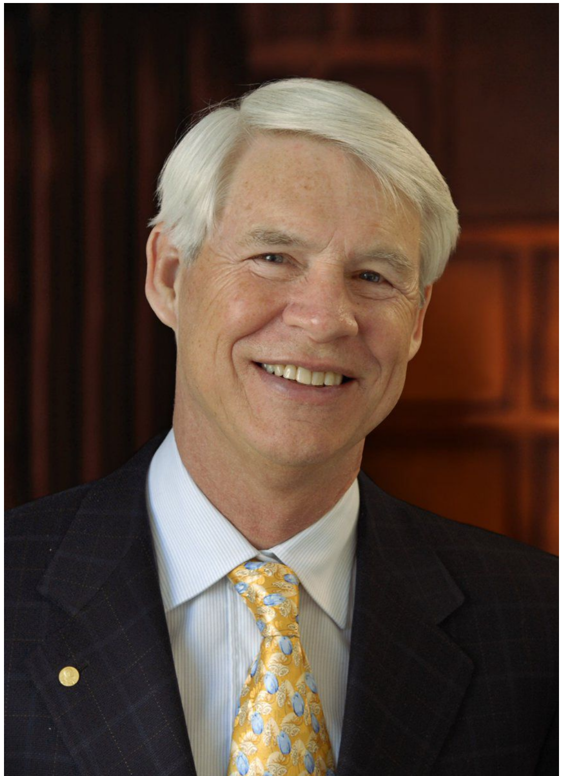
En la industria financiera, los modelos GARCH son una aproximación a un modelo de volatilidad.

- La volatilidad se mide frecuentemente como el desvío estándar o la varianza del retorno de precios.
- En general, mayor volatilidad indica un mayor nivel de riesgo.
- Un shock en el precio de un activo puede conducir a movimientos de precios significativos que son persistentes en el tiempo. Este fenómeno se denomina “cluster de volatilidad”.

Modelo ARCH

Modelo Autoregresivo de Heterocedasticidad Condicional

Desarrollado por Robert Engle. Premio Nobel de Economía en 2003.



Modelo GARCH

Modelo ARCH generalizado

Desarrollado por Tim Bollerslev, discípulo de Engle

Términos estadísticos relacionados

- Ruido blanco: valores no correlacionados, con media cero y varianza finita.
- Residuo = Predicción - Observación

$$\hat{y}_t = y_t + (\hat{y}_t - y_t)$$

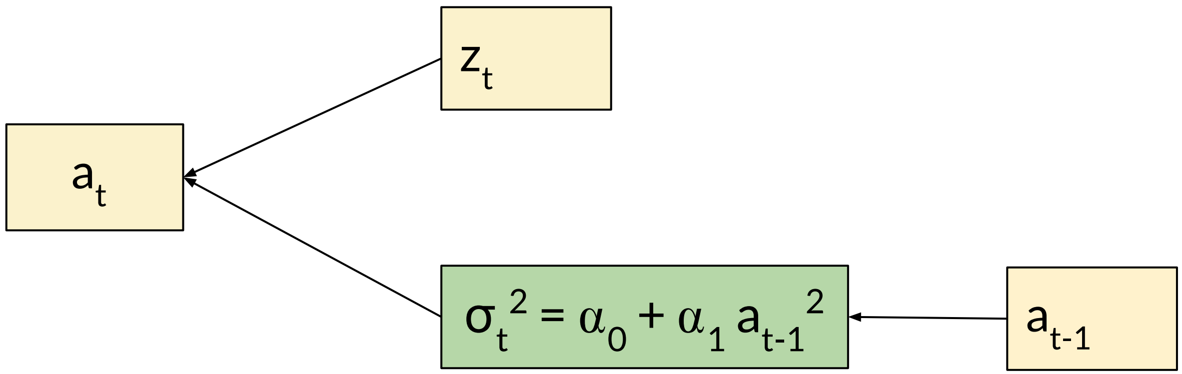
a_t = residuo

ARCH(1)

- La serie es un ruido blanco z_t multiplicado por una volatilidad σ_t :

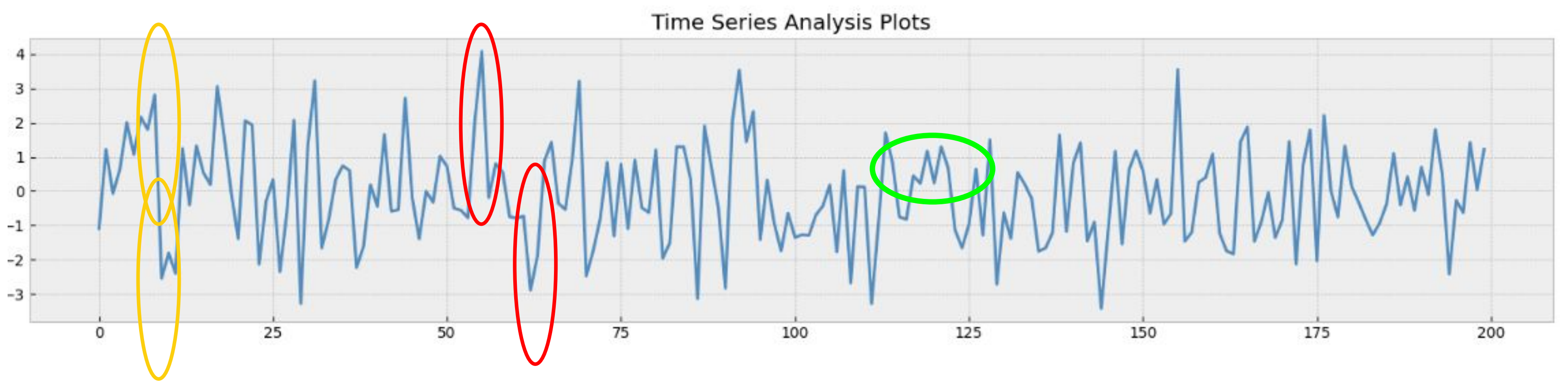
$$a_t = \sigma_t * z_t$$

- La volatilidad depende del valor del residuo en el paso anterior
- Un cambio brusco en at produce alta volatilidad.



ARCH(1)

Problema: Cambios rápidos



Se busca que cuando hay un cambio brusco, la serie se mantenga en el nivel alcanzado.

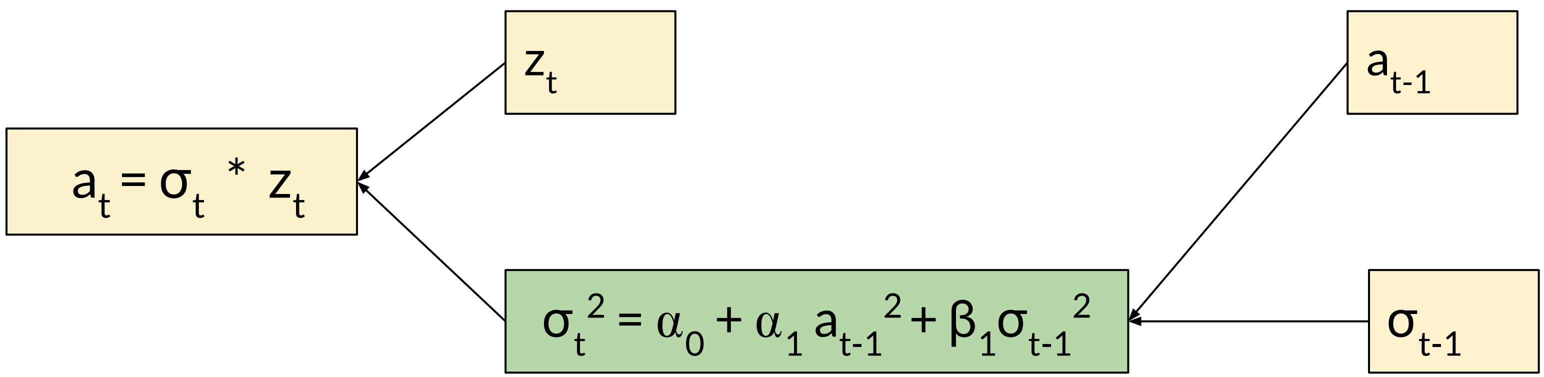
Solución: La volatilidad depende de la volatilidad pasada.

GARCH(1,1)

- La serie es un ruido blanco z_t multiplicado por una volatilidad σ_t :

$$a_t = \sigma_t * z_t$$

- La volatilidad depende del valor de la serie y de la volatilidad en el paso anterior.



Restricciones de los parámetros

GARCH(1, 1)

- Todos los coeficientes no negativos

$$\alpha_0, \alpha_1, \beta_1 \geq 0$$

- Es un modelo con reversión a la media de la varianza a largo plazo

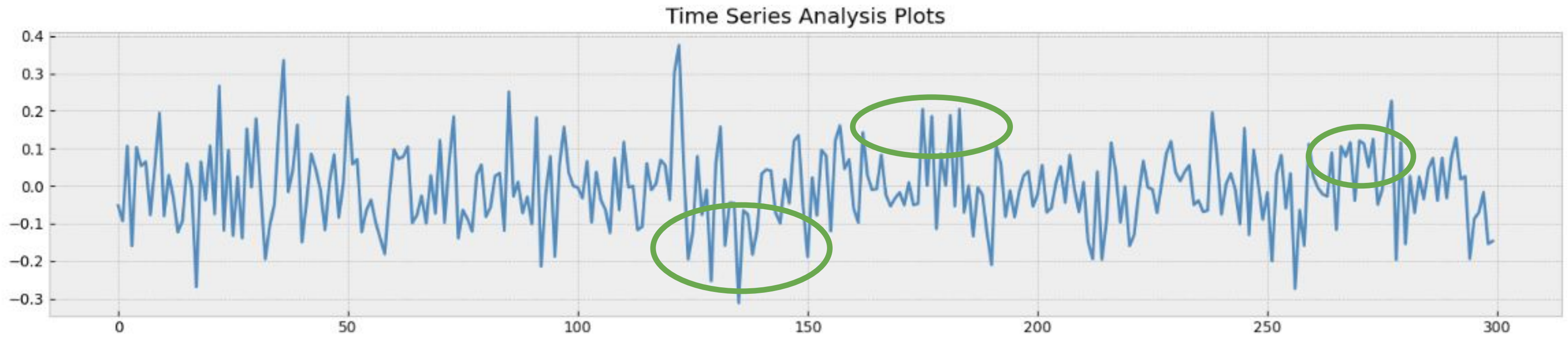
$$\alpha_1 + \beta_1 < 1$$

- La varianza a largo plazo está dada por

$$\alpha_0 / (1 - \alpha_1 - \beta_1)$$

GARCH(1,1)

Es factible modelar la persistencia de valores altos (bajos) ante un cambio brusco de la serie.



- A mayor α_1, mayor es el impacto inmediato de las innovaciones anteriores.
- A mayor β_1, mayor duración del impacto

Problema: No modela la asimetría de las respuestas a movimientos bruscos.

Modelos ARCH

ARCH(1)

$$\sigma_t^{\,2} = \alpha_0 + \alpha_1\, a_{t-1}^{\,2}$$

ARCH(p)

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

- Modelos autorregresivos
- Dependencia de la volatilidad de un promedio ponderado de valores pasados de la serie.

Modelos GARCH

GARCH(1,1)

$$\sigma_t^2 = \alpha_0 + \alpha_1\, a_{t-1}^2 + \beta_1\, \sigma_{t-1}^2$$

GARCH(p, q)

$$\sigma_t^2 = \alpha_0 + \alpha_1\, a_{t-1} + \alpha_2\, a_{t-2} + \ldots. + \alpha_p\, a_{t-p} + \beta_1\, \sigma_{t-1}^2 + \beta_2\, \sigma_{t-2}^2 + \ldots. + \beta_q\, \sigma_{t-q}^2$$

Restricciones de los parámetros

GARCH(1, 1)

- Todos los coeficientes no negativos

$$\alpha,\ \beta \geq 0$$

- Es un modelo con reversión a la media de la varianza a largo plazo

$$\alpha + \beta < 1$$

- La varianza a largo plazo está dada por

$$\omega / (1 - \alpha - \beta)$$

Continuamos en jupyter