

Índice

1. Identificación	2
1.1. Instrumentos de identificación	3
2. Estimación por máxima verosimilitud	4
2.1. Propiedades de los estimadores de máxima verosimilitud (MV)	4
3. Diagnosis	5
3.1. Instrumentos de diagnosis	5
3.1.1. Reformulación	5
3.1.2. Componentes deterministas	5
3.1.3. Otras herramientas estadísticas a usar, una vez comprobada la ausencia de correlación serial en los residuos	6
4. Raíces unitarias	6
4.1. Notación: operadores retardo y diferencia y modelos ARIMA	6
4.1.1. Notación: ARIMA	6
4.2. Raíces unitarias en los polinomios AR y MA	6
4.3. Paseos aleatorios	7
5. Modelos ARIMA estacionales (SARIMA)	8
5.1. MA(1) estacional con raíz positiva	8
5.2. AR(1) estacional con raíz positiva	9
5.3. ARIMA $(0,0,1)x(0,0,1)$	11
5.4. ARIMA $(1,0,0)x(0,0,1)$	11
5.5. ARIMA $(1,0,0)x(1,0,0)$	12
5.6. ARIMA $(0,0,1)x(1,0,0)$	12
6. Resumen del análisis univariante de series temporales	13
6.1. Ideas principales respecto a la modelización univariante	13
6.2. Metodología	14
7. Introducción intuitiva e informal al concepto de ergodicidad	14
7.1. Ejemplo de un proceso no ergódico	15

Lección 7. Modelos ARIMA y SARIMA. Identificación y diagnosis

Marcos Bujosa

24 de noviembre de 2025

Resumen

Repasaremos los instrumentos de identificación y diagnosis del análisis univariante. Extenderemos la notación para incorporar modelos con raíces unitarias y modelos estacionales. Finalmente resumiremos las ideas principales del análisis univariante.

- ([slides](#)) — ([html](#)) — ([pdf](#)) — ([mybinder](#))

Carga de algunos módulos de python y creación de directorios auxiliares

```
# Para trabajar con los datos y dibujarlos necesitamos cargar algunos módulos de python
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib as mpl
# definimos parámetros para mejorar los gráficos
mpl.rcParams['text', usetex=False]
import matplotlib.pyplot as plt # data visualization
import datafram_image as dfi # export tables as .png
```

- Directorio auxiliar para albergar las figuras de la lección: para publicar la lección como pdf o página web, necesito los gráficos como ficheros .png alojados algún directorio específico:

```
imagenes_leccion = "./img/lecc07" # directorio para las imágenes de la lección
import os
os.makedirs(imagenes_leccion, exist_ok=True) # crea el directorio si no existe
```

Gráficos para las ACF, PACF y densidades espectrales teóricas

Cargamos las funciones auxiliares (véase la carpeta `src/`)

```
import warnings
warnings.filterwarnings("ignore", category=UserWarning)
%run -i ./src/analisis_armas.py
```

1. Identificación

- Combinando herramientas gráficas y estadísticas, se puede inferir un modelo para los datos.
- Este proceso de especificación empírica de un modelo es conocido como "*identificación*".

El proceso de identificación puede estructurarse como una secuencia de preguntas:

1. ¿Es la serie estacionaria?
 2. ¿Tiene una media significativa?
 3. ¿Es persistente la ACF? ¿sigue alguna pauta reconocible?
 4. ¿Es persistente la PACF? ¿sigue alguna pauta reconocible?
- La identificación se apoya en estadísticos muestrales (media, autocorrelaciones, etc.) cuya representatividad respecto del proceso estocástico subyacente depende de la estacionariedad y la ergodicidad¹.
 - Tras inducir la estacionariedad, especificamos un modelo tentativo decidiendo cuál de las funciones ACF o PACF es finita y cuál es persistente

	ACF finita	ACF persistente
PACF finita	Ruido blanco: retardos conjuntamente NO significativos	<u>AR</u> : orden indicado por la PACF
PACF persistente	<u>MA</u> : orden indicado por la ACF	<u>ARMA</u>

La parametrización de mayor orden en modelos ARMA con series económicas suele ser ARMA(2, 1)

1.1. Instrumentos de identificación

	Instrumento	Objetivo y observaciones
Transf. logarítmica	Gráficos rango-media y serie temporal.	Conseguir independizar la variabilidad de los datos de su nivel. Las series económicas necesitan esta transformación frecuentemente.
d , orden de diferenciación	Gráfico de la serie temporal. ACF (caída lenta y lineal). Contrastes de raíz unitaria (DF o ADF y KPSS).	Conseguir que los datos fluctúen en torno a una media estable. En series económicas, d suele ser 0, 1 ó 2.
Constante	Media de la serie transformada. Desviación típica de la media.	Si la media de la serie transformada es significativa, el modelo debe incluir un término constante.
p , orden AR.	Si PACF cae abruptamente en el retardo p y la ACF decae lentamente.	En series económicas p suele ser ≤ 2 .
q , orden MA.	Si ACF cae abruptamente en el retardo q y PACF decae lentamente.	En series económicas q suele ser ≤ 1 .

¹Véase la sección [Introducción intuitiva e informal al concepto de ergodicidad](#).

2. Estimación por máxima verosimilitud

Los parámetros de los modelos ARMA se estiman por máxima verosimilitud exacta (resolviendo un sistema de ecuaciones no lineales de forma iterativa).

2.1. Propiedades de los estimadores de máxima verosimilitud (MV)

Invarianza Si $\hat{\beta}^{MV}$ es el estimador de máxima verosimilitud de β y $g(\beta)$ es una función continua, entonces el estimador de máxima verosimilitud de $g(\beta)$ es $g(\hat{\beta}^{MV})$.

Ejemplo. Si $\hat{\sigma}_{MV}^2$ es el estimador MV de σ^2 , entonces el estimador MV de σ es $\sqrt{\hat{\sigma}_{MV}^2}$.

Consistencia A medida que el tamaño muestral n crece, el estimador $\hat{\beta}^{MV}$ se aproxima en probabilidad al valor verdadero β .

$$\operatorname{plim}_{n \rightarrow \infty} \hat{\beta}^{MV} = \beta$$

Es decir, a medida que crece la muestra, la probabilidad de estimar valores alejados del verdadero valor tiende a cero; formalmente y de manera más explícita:

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|\hat{\beta}^{MV} - \beta| > \epsilon) = 0.$$

Insesgadez asintótica El estimador puede ser sesgado en muestras pequeñas, pero su sesgo desaparece conforme $n \rightarrow \infty$. Es decir, el valor esperado converge al parámetro verdadero.

$$\lim_{n \rightarrow \infty} E(\hat{\beta}_n^{MV}) = \beta$$

Diferencia con insesgadez Un estimador **insesgado** cumple $E(\hat{\beta}) = \beta$ para todo n , mientras que un estimador **asintóticamente insesgado** lo cumple solo en el límite.

Eficiencia asintótica El estimador se vuelve más preciso a medida que crece n . Su varianza tiende a cero, concentrándose alrededor del valor verdadero.

$$\lim_{n \rightarrow \infty} \operatorname{Var}(\hat{\beta}_n^{MV}) = 0$$

Nota Bajo condiciones regulares, el MLE es **asintóticamente eficiente**, es decir, alcanza la menor varianza posible entre los estimadores consistentes (cota de Cramér–Rao).

Propiedades	Significado
Invarianza	Si $\hat{\beta}^{MV}$ es el estimador de máxima verosimilitud de β , entonces el estimador de máxima verosimilitud de $g(\beta)$ es $g(\hat{\beta}^{MV})$.
Consistencia	$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(\hat{\beta}^{MV} - \beta > \epsilon) = 0$
Insesgadez asintótica	$\lim_{n \rightarrow \infty} E(\hat{\beta}_n^{MV}) = \beta$
Eficiencia asintótica	$\lim_{n \rightarrow \infty} \operatorname{Var}(\hat{\beta}_n^{MV}) = 0$

3. Diagnosis

3.1. Instrumentos de diagnosis

	Instrumento	Possible diagnóstico
d , orden de diferenciación	Proximidad a 1 de alguna raíz de los polinomios AR o MA.	Conviene diferenciar si la raíz es AR; o quitar una diferencia si es MA (<i>salvo si hay tendencia determinista</i>).
d , orden de diferenciación	Gráfico de los residuos.	Si muestra rachas largas de residuos positivos o negativos, puede ser necesaria una diferencia adicional.
Constante	Media de los residuos.	Si es significativa: añadir una constante.
Constante	Constante estimada.	Si NO es significativa: el modelo mejorará quitando el término constante.
p y q ,	Contrastes de significación de los parámetros estimados.	Pueden sugerir eliminar parámetros irrelevantes.
p y q ,	ACF/PACF residuos. Test Q de Ljung-Box para la ACF.	Indican posibles pautas de autocorrelación no modelizadas.
p y q ,	Correlaciones elevadas entre los parámetros estimados.	Puede ser síntoma de sobreparametrización.

3.1.1. Reformulación

Siempre debemos analizar las características de los residuos.

- Si los residuos presentan correlación serial debemos reformular el modelo.

Ejemplo Supongamos que hemos estimado un modelo AR(1) y el correlograma de los residuos presenta estructura MA(1): debemos estimar un modelo ARMA(1,1) y analizar los nuevos residuos y la significatividad de los parámetros

3.1.2. Componentes deterministas

Un MA(1) con parámetro próximo a 1 puede indicar sobrediferenciación, pero también la presencia de una componente determinista.

Ejemplo El modelo $(1 - B)Y_t = \beta + (1 - \theta B)U_t$ es equivalente al modelo $Y_t = \beta t + U_t$ cuando $\theta = 1$.

(fíjese que el primer modelo no es invertible; y el segundo no es estacionario).

Una vez superadas las pruebas de diagnostico, aún se puede aplicar un análisis exploratorio; consistente en añadir parámetros AR o MA para comprobar si resultan significativos y mejoran el modelo.

- no aumentar los órdenes autorregresivos y de medias móviles simultáneamente.

3.1.3. Otras herramientas estadísticas a usar, una vez comprobada la ausencia de correlación serial en los residuos

1. Término constante. Cuando el modelo ni tiene término constante: contrastar si la media de los residuos es significativa ($H_0 : \mu = 0$) para decidir si incorporarlo

$$\frac{\widehat{\mu}}{\widehat{dt(\mu)}} \underset{H_0}{\sim} t_{n-1}; \quad \widehat{dt(\mu)} = \frac{\widehat{\sigma}}{\sqrt{n}}.$$

2. Verificar visualmente que la varianza de los residuos es constante.
3. Contrastar la normalidad de los residuos con el test Jarque-Bera. [Jarque-Bera test \(Wikipedia\)](#)

La transformación logarítmica suele inducir normalidad si no hay datos negativos.

```
from scipy import stats
rng = np.random.default_rng()
x = rng.normal(0, 1, 100000)
jarque_bera_test = stats.jarque_bera(x)
jarque_bera_test
# jarque_bera_test.statistic
# jarque_bera_test.pvalue
```

4. Raíces unitarias

4.1. Notación: operadores retardo y diferencia y modelos ARIMA

El operador diferencia ∇ se define a partir del operador retardo como $\nabla = (1 - \mathbb{B})$:

$$\nabla Y_t = (1 - \mathbb{B})Y_t = Y_t - Y_{t-1}.$$

El operador diferencia estacional es $\nabla_S = (1 - \mathbb{B}^S)$:

$$\nabla_S Y_t = (1 - \mathbb{B}^S)Y_t = Y_t - Y_{t-S}.$$

4.1.1. Notación: ARIMA

Extendemos la notación a procesos con raíces autorregresivas unitarias con “ARIMA(p, d, q)”; donde d indica el número de diferencias que la serie necesita para ser $I(0)$,

$$\phi_p * \nabla^d * \mathbf{Y} = \theta_q * \mathbf{U};$$

es decir

$$\phi_p(\mathbb{B}) \nabla^d Y_t = \theta_q(\mathbb{B}) U_t; \quad t \in \mathbb{Z}.$$

4.2. Raíces unitarias en los polinomios AR y MA

Cuando un polinomio tiene alguna raíz igual a uno se dice que tiene “raíces unitarias”.

Si el polinomio AR estimado tiene alguna raíz “próxima a uno” es síntoma de infradiferenciación.

Si el polinomio MA estimado tiene alguna raíz “próxima a uno” es síntoma de sobrediferenciación.

Ejemplos:

Modelo expresado con raíces unitarias en ϕ o θ	Modelo equivalente sin raíces unitarias en ϕ o θ
$(1 - 1,5B + ,5B^2)Y_t = U_t$	$(1 - 0,5B)\nabla Y_t = U_t$
$(1 - ,5B + 0,7B^2)\nabla^2 Y_t = (1 - B)U_t$	$(1 - ,5B + 0,7B^2)\nabla Y_t = U_t$
$\nabla Y_t = \beta + (1 - B)U_t$	$Y_t = \beta t + U_t$ (¡no estacionario!)

4.3. Paseos aleatorios

Un paseo aleatorio representa una variable cuyos incrementos son ruido blanco:

$$Y_t = \mu + Y_{t-1} + U_t.$$

Cuando $\mu \neq 0$ se denomina *paseo aleatorio con deriva*: $\nabla Y_t = \mu + U_t$.

```
fig = plot_paseo_aleatorio_analysis(trend="t", pendiente=0.25, n=400, semilla=2025)
fig.savefig('./img/lecc07/ACF-RWcd.png', dpi=300, bbox_inches='tight')
```



El proceso tiene mayor inercia cuanto mayor es $|\mu|$. El signo de μ determina el signo de la pendiente global.

Cuando $\mu = 0$ se denomina sencillamente *paseo aleatorio*: $\nabla Y_t = U_t$

```
fig = plot_paseo_aleatorio_analysis(n=400, semilla=2026)
fig.savefig('./img/lecc07/ACF-RW.png', dpi=300, bbox_inches='tight')
```



5. Modelos ARIMA estacionales (SARIMA)

El período estacional S es el número mínimo de observaciones necesarias para recorrer un ciclo estacional completo. Por ejemplo, $S = 12$ para datos mensuales, $S = 4$ para datos trimestrales, $S = 24$ para datos horarios, etc.

Captamos la estacionalidad con modelos ARIMA(p, d, q) \times (P, D, Q) $_S$

$$\phi_p(\mathbf{B}) \Phi_P(\mathbf{B}^S) \nabla^d \nabla_S^D Y_t = \theta_q(\mathbf{B}) \Theta_q(\mathbf{B}^S) U_t; \quad t \in \mathbb{Z}$$

donde

$$\begin{aligned} \Phi_P(\mathbf{B}^S) &= 1 - \Phi_1 \mathbf{B}^{1 \cdot S} - \Phi_2 \mathbf{B}^{2 \cdot S} - \dots - \Phi_P \mathbf{B}^{P \cdot S} \\ \Theta_Q(\mathbf{B}^S) &= 1 - \Theta_1 \mathbf{B}^{1 \cdot S} - \Theta_2 \mathbf{B}^{2 \cdot S} - \dots - \Theta_Q \mathbf{B}^{Q \cdot S} \\ \nabla_S^D &= (1 - \mathbf{B}^S)^D \end{aligned}$$

Es decir, el modelo consta de polinomios autorregresivos y de media móvil tanto regulares (en minúsculas) como estacionales (en mayúsculas).

Veamos un ejemplo de un modelo MA(1) estacional y otro de un modelo AR(1) estacional...

5.1. MA(1) estacional con raíz positiva

$$\text{MA}(1) \text{ estacional: } \Theta = 1 - 0.9z^{12} \Rightarrow X_t = (1 - 0.9\mathbf{B}^{12})U_t$$

```
SMA1 = [1, -0.9]
fig = plot_arma_seasonal_parametric_diagnostics(seasonal_ma_params=SMA1, s=12)
fig.savefig('img/lecc07/ACF-SMA1p.png', dpi=300, bbox_inches='tight')
```



	P. real	P. imag.	Módulo	Frecuencia	Periodo	Multip.
11	1.008819	0.000000	1.008819	0.000000	Inf	1
9	0.873663	-0.504409	1.008819	0.083333	12.0	1
10	0.873663	0.504409	1.008819	0.083333	12.0	1
7	0.504409	-0.873663	1.008819	0.166667	6.0	1
8	0.504409	0.873663	1.008819	0.166667	6.0	1
5	0.000000	-1.008819	1.008819	0.250000	4.0	1
6	0.000000	1.008819	1.008819	0.250000	4.0	1
3	-0.504409	-0.873663	1.008819	0.333333	3.0	1
4	-0.504409	0.873663	1.008819	0.333333	3.0	1
1	-0.873663	-0.504409	1.008819	0.416667	2.4	1
2	-0.873663	0.504409	1.008819	0.416667	2.4	1
0	-1.008819	0.000000	1.008819	0.500000	2.0	1

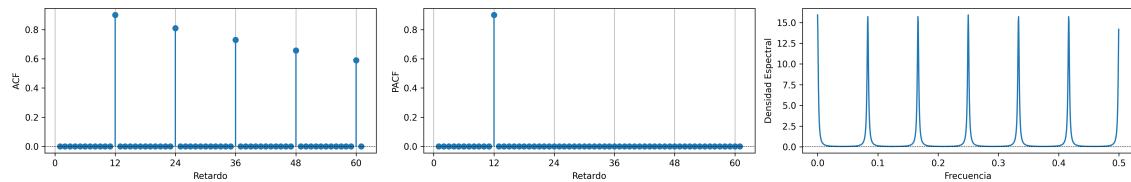
```
fig = plot_sarima_analysis(seasonal_ma_params=SMA1, s=12, seed=2025)
fig.savefig('./img/lecc07/Sim-SMA1p.png', dpi=300, bbox_inches='tight')
```



5.2. AR(1) estacional con raíz positiva

$$\text{AR}(1) \text{ estacional: } \Phi = 1 - 0.9z^{12} \Rightarrow (1 - 0.9B^{12})X_t = U_t$$

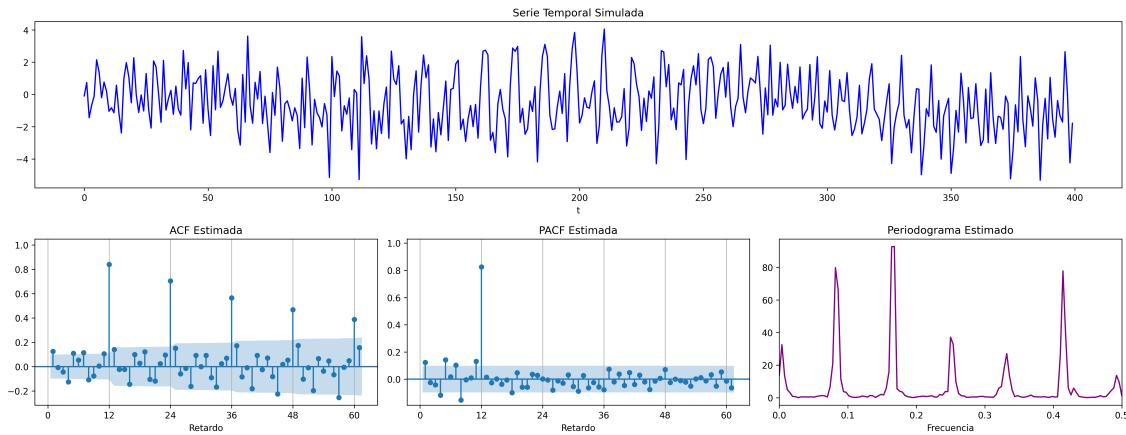
```
SAR1 = [1, -0.9]
fig = plot_arma_seasonal_parametric_diagnostics(seasonal_ar_params=SAR1, s=12)
fig.savefig('./img/lecc07/ACF-SAR1p.png', dpi=300, bbox_inches='tight')
```



	P. real	P. imag.	Módulo	Frecuencia	Periodo	Multip.
11	1.008819	0.000000	1.008819	0.000000	Inf	1
9	0.873663	-0.504409	1.008819	0.083333	12.0	1
10	0.873663	0.504409	1.008819	0.083333	12.0	1
7	0.504409	-0.873663	1.008819	0.166667	6.0	1
8	0.504409	0.873663	1.008819	0.166667	6.0	1
5	0.000000	-1.008819	1.008819	0.250000	4.0	1
6	0.000000	1.008819	1.008819	0.250000	4.0	1
3	-0.504409	-0.873663	1.008819	0.333333	3.0	1
4	-0.504409	0.873663	1.008819	0.333333	3.0	1
1	-0.873663	-0.504409	1.008819	0.416667	2.4	1
2	-0.873663	0.504409	1.008819	0.416667	2.4	1
0	-1.008819	0.000000	1.008819	0.500000	2.0	1

Evidentemente las raíces son iguales a las del caso anterior (aunque ahora corresponden al polinomio autorregresivo).

```
fig = plot_sarima_analysis(seasonal_ar_params=SAR1, s=12, seed=2025)
fig.savefig('./img/lecc07/Sim-SAR1p.png', dpi=300, bbox_inches='tight')
```



Con estos dos ejemplos hemos podido apreciar que:

- las pautas de autocorrelación son análogas a las de los MA(1) y AR(2), pero ahora los retardos significativos corresponden a los retardos estacionales, es decir, a múltiplos del período estacional S .
- En estos ejemplos, en los que $S = 12$, los retardos estacionales son: 12, 24, 36, 48, 60,...
- las correlaciones correspondientes a los “retardos regulares” (es decir, todos menos los estacionales) son no significativas en general.

Veamos ahora un par de ejemplos de modelos estacionales multiplicativos (i.e., con parte regular y parte estacional).

5.3. ARIMA (0,0,1)x(0,0,1)

$$\text{ARIMA}(0, 0, 1) \times (0, 0, 1)_{12}: \quad X_t = (1 - 0.9B)(1 - 0.9B^{12})U_t$$

```
MA1 = [1, -0.8]
SMA1 = [1, -0.9]
fig = plot_arma_seasonal_parametric_diagnostics(ma_params=MA1, seasonal_ma_params=SMA1, s=12)
fig.savefig('./img/lecc07/ACF-MA1SMA1.png', dpi=300, bbox_inches='tight')
```



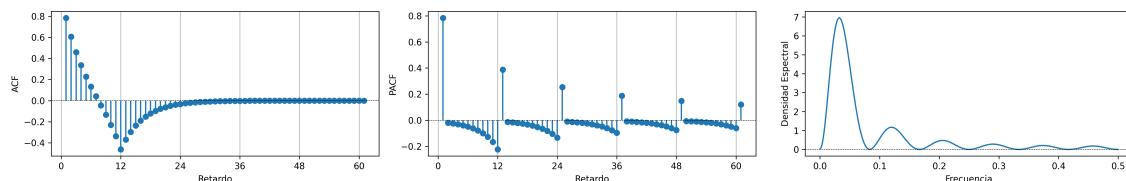
```
fig = plot_sarima_analysis(ma_params=MA1, seasonal_ma_params=SMA1, s=12, seed=2025)
fig.savefig('./img/lecc07/Sim-MA1SMA1.png', dpi=300, bbox_inches='tight')
```



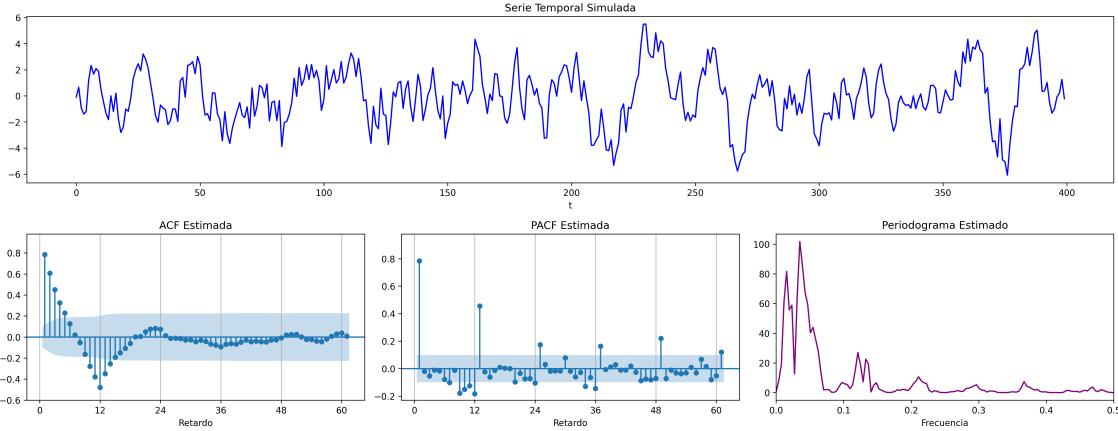
5.4. ARIMA (1,0,0)x(0,0,1)

$$\text{ARIMA}(1, 0, 0) \times (0, 0, 1)_{12}: \quad (1 - 0.9B)X_t = (1 - 0.9B^{12})U_t$$

```
AR1 = [1, -0.8]
SMA1 = [1, -0.9]
fig = plot_arma_seasonal_parametric_diagnostics(ar_params=AR1, seasonal_ma_params=SMA1, s=12)
fig.savefig('./img/lecc07/ACF-AR1SMA1.png', dpi=300, bbox_inches='tight')
```



```
fig = plot_sarima_analysis(ar_params=AR1, seasonal_ma_params=SMA1, s=12, seed=2025)
fig.savefig('./img/lecc07/Sim-AR1SMA1.png', dpi=300, bbox_inches='tight')
```



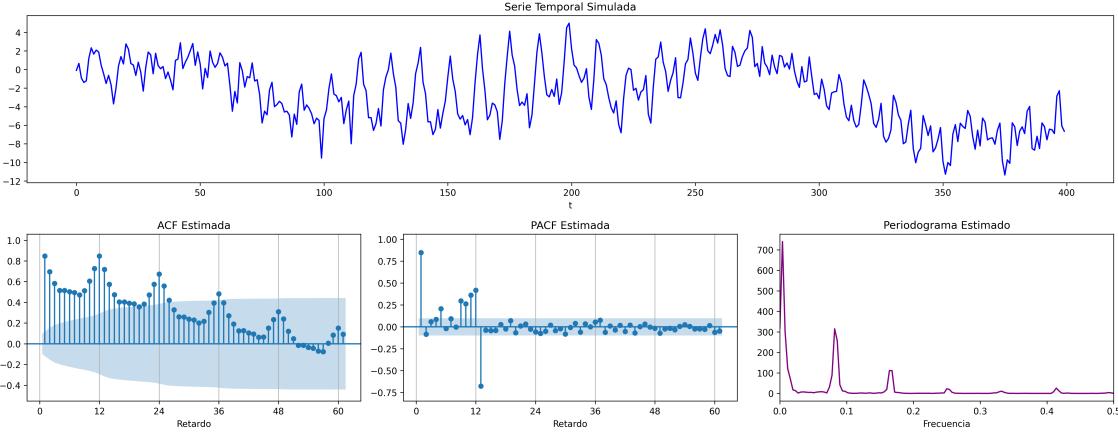
5.5. ARIMA (1,0,0)x(1,0,0)

$$\text{ARIMA}(1,0,0) \times (1,0,0)_{12}: \quad (1 - 0.9B)(1 - 0.9B^{12})X_t = U_t$$

```
AR1 = [1, -0.8]
SAR1 = [1, -0.9]
fig = plot_arima_seasonal_parametric_diagnostics(ar_params=AR1, seasonal_ar_params=SAR1, s=12, logs=True)
fig.savefig('./img/lecc07/ACF-AR1SAR1.png', dpi=300, bbox_inches='tight')
```



```
fig = plot_sarima_analysis(ar_params=AR1, seasonal_ar_params=SAR1, s=12, seed=2025)
fig.savefig('./img/lecc07/Sim-AR1SAR1.png', dpi=300, bbox_inches='tight')
```



5.6. ARIMA (0,0,1)x(1,0,0)

$$\text{ARIMA}(0,0,1) \times (1,0,0)_{12}: \quad (1 - 0.9B^{12})X_t = (1 - 0.9B)U_t$$

```

MA1 = [1, -0.8]
SAR1 = [1, -0.9]
fig = plot_arima_seasonal_parametric_diagnostics(ma_params=MA1, seasonal_ar_params=SAR1, s=12)
fig.savefig('./img/lecc07/ACF-MA1SAR1.png', dpi=300, bbox_inches='tight')

```



```

fig = plot_sarima_analysis(ma_params=MA1, seasonal_ar_params=SAR1, s=12, seed=2025)
fig.savefig('./img/lecc07/Sim-MA1SAR1.png', dpi=300, bbox_inches='tight')

```



En estos cuatro ejemplos hemos podido apreciar que

- en el entorno de los retardos estacionales surgen una serie de coeficientes significativos (“satélites”) que proceden de la interacción entre las estructuras regular y estacional
- Estos satélites son útiles para identificar en qué retardos estacionales hay autocorrelaciones no nulas, pero no requieren una parametrización especial.

6. Resumen del análisis univariante de series temporales

6.1. Ideas principales respecto a la modelización univariante

- Es una modelización sin variables exógenas
- Modelizan la interdependencia temporal con polinomios de órdenes reducidos.
- Está especialmente indicada para hacer predicción.
- Parte de dos supuestos sobre el proceso estocástico subyacente:
 1. es débilmente estacionario

2. tiene representación como proceso lineal: $Y_t = \mu + \sum_{j=0}^{\infty} a_j U_{t-j}$; con $\mu \in \mathbb{R}$, $a \in \ell^2$ y $U \sim WN(0, \sigma^2)$
- Utiliza múltiples instrumentos: (a) gráficos (b) función de autocorrelación (c) función de autocorrelación parcial, (d) estadístico Q de Ljung-Box, etc...
 - Si la serie original no "parece" débilmente estacionaria, se induce esta propiedad mediante las transformaciones adecuadas

	ACF finita	ACF persistente
PACF finita	Ruido blanco: retardos conjuntamente NO significativos	<u>AR</u> : orden indicado por la PACF
PACF persistente	<u>MA</u> : orden indicado por la ACF	<u>ARMA</u>

6.2. Metodología

Tres fases:

Identificación Se elige una especificación provisional para el proceso estocástico subyacente en función de las características medibles de los datos ("dejar que los datos hablen")

Estimación suele requerir métodos iterativos (*Gretl se encarga de esto*)

Diagnosis de la calidad estadística del modelo ajustado. Algunos controles estándar son:

- Significatividad de los parámetros estimados
- Estacionariedad y homocedasticidad de los residuos
- ¿Existe un patrón de autocorrelación residual que podría ser modelado? ¿O hemos logrado que los residuos sean **ruido blanco**?

Si la diagnosis no es satisfactoria, se vuelve a la primera fase.

Si la diagnosis es satisfactoria... ¡hemos logrado un modelo aceptable!... que podremos usar para realizar pronósticos.

7. Introducción intuitiva e informal al concepto de ergodicidad

La **ergodicidad** de un proceso estocástico es importante porque conecta el comportamiento en el tiempo de una sola realización del proceso con las propiedades estadísticas del conjunto de todas las posibles realizaciones. En otras palabras, nos dice cuándo "promediar en la muestra" puede arrojar un buen estimador de algún momento teórico. Es decir, si el proceso es **ergódico**:

- estimadores como la media muestral, la varianza muestral o las autocorrelaciones son **consistentes y útiles**, porque en el límite nos dan el verdadero parámetro poblacional. Por ejemplo, cuando el proceso es ergódico, el **promedio** de una sola realización (la media muestral de una serie larga) **converge** al valor esperado teórico (el promedio de la distribución incondicional del proceso).

- En muchas aplicaciones, eso también se traduce en eficiencia asintótica (dependiendo del tipo de proceso y del estimador).

En cambio, si el proceso **no es ergódico** los promedios en el tiempo dependen de la trayectoria concreta observada, y pueden **no coincidir** con los momentos teóricos.

- En tal caso, un estimador basado en una única realización puede estar **sesgado** o ser **inconsistente**: no converge al parámetro teórico, sino a algo dependiente de la trayectoria particular. Lo cual significa que los métodos estadísticos usuales de series temporales (media muestral, autocorrelaciones, regresiones dinámicas) dejan de ser fiables.

En términos prácticos: la ergodicidad es lo que “salva” la inferencia empírica. Sin ergodicidad, trabajar con datos de una sola realización en el tiempo (que es lo que tenemos en economía, física o finanzas) no nos dice realmente cómo es la distribución del proceso estocástico.

- Si un proceso es ergódico, entonces una única trayectoria suficientemente larga contiene información estadística suficiente para realizar inferencia estadística sobre el proceso estocástico.
- Esto significa que podemos estimar valores esperados (media, varianza, correlaciones) observando un solo experimento durante mucho tiempo, sin necesidad de repetirlo muchas veces.
- En la práctica, cuando medimos fenómenos físicos, económicos o sociales, solo observamos **una** trayectoria del proceso (ejemplo: una serie temporal de precios).
- La ergodicidad garantiza que los cálculos hechos sobre esa única serie son representativos de la distribución subyacente.
- La ergodicidad suele estudiarse en procesos estacionarios. No todos los procesos estacionarios son ergódicos, pero la ergodicidad asegura que la “estabilidad en el tiempo” realmente refleja la “estabilidad en probabilidad”.
- Implicaciones en teoría y aplicaciones:
 - En estadística de series temporales: justifica el uso de una muestra temporal.
 - En economía: permite usar datos históricos de una variable para inferir expectativas a largo plazo.

En resumen: la ergodicidad es crucial porque es lo que hace posible inferir la estructura probabilística de un proceso observando solo una realización larga en el tiempo, algo que es lo habitual en el mundo real.

7.1. Ejemplo de un proceso no ergódico

Consideremos un proceso de Bernoulli con un *parámetro aleatorio fijo* p :

1. Primero se elige al azar una probabilidad p con distribución uniforme en el intervalo $[0, 1]$.
- 2 Luego se genera una secuencia infinita de lanzamientos de moneda con esa probabilidad p .

En dicho experimento, como el valor esperado para el parámetro p es $0,5$ (es el valor esperado de una distribución uniforme $[0, 1]$), la esperanza de un lanzamiento es $E(X) = p = 0,5$; y si prome-

diáramos los promedios de muchas realizaciones de esas secuencias infinitas, nos aproximariamos al verdadero valor esperado $E(X) = 0,5$.

Pero si tan solo tenemos una realización del proceso, a largo plazo el promedio temporal de la secuencia tenderá casi seguro al valor particular de p que nos haya salido al principio, que no es necesariamente a 0,5.

Por tanto, este proceso no es ergódico: los promedios temporales dependen de la trayectoria inicial (dependen del valor de p que nos “tocó” inicialmente).