📊 Tarea 3 — Análisis de Series de Tiempo

Curso: Análisis de Datos

Autor: Carlos Solares

Fecha: 31/10/2025

🎯 Objetivo general

Analizar el comportamiento de dos procesos estocásticos fundamentales:

1. **Ruido blanco** — proceso sin dependencia temporal.
2. **Proceso autoregresivo AR(1)** — proceso con dependencia del pasado.

Se estudian sus propiedades de **estacionariedad**, **autocorrelación** y **estimación de parámetros**.

🧩 Fundamentos teóricos

- Una **serie de tiempo** es una realización de un proceso aleatorio dependiente del tiempo.
- En un proceso estacionario:
  - $E[X_t] = \mu$
  - $\mathrm{Var}(X_t) = \sigma^2$
  - $\mathrm{Cov}(X_t, X_{t+h}) = \gamma(h)$
- La **función de autocorrelación (ACF)** mide:

$$r(h) = \frac{\mathrm{Cov}(X_t, X_{t-h})}{\mathrm{Var}(X_t)}$$

Inciso 1 — Simetría de $\gamma(h)$

Definición: $\gamma(h) = \mathrm{Cov}(X_t, X_{t+h})$.

Prueba (bajo estacionariedad débil):

1. $\gamma(-h) = \mathrm{Cov}(X_t, X_{t-h})$.
2. Reindexa con $s = t - h$: $\gamma(-h) = \mathbb{E}[(X_{s+h} - \mu)(X_s - \mu)]$.
3. Conmutatividad: $(X_{s+h} - \mu)(X_s - \mu) = (X_s - \mu)(X_{s+h} - \mu)$.
4. Estacionariedad → depende solo de h: $\gamma(-h) = \gamma(h)$.

Conclusión: la función de autocovarianza es **par** (simétrica).

Inciso 2 (a) — ¿Por qué es deseable la estacionariedad?

- **Reglas estables** en el tiempo (media, varianza, covarianzas no cambian).
- **Modelos más simples/robustos** (ARMA/ARIMA suelen asumirla).
- **Pronósticos más confiables** (no cambian “las reglas del juego”).
- **Comparabilidad temporal** entre periodos.

Inciso 2 (b) — ¿Es $Y_t = a + bt + \varepsilon_t$ estacionaria?

Supuestos: $\mathbb{E}[\varepsilon_t] = 0$, $\mathrm{Var}(\varepsilon_t) = \sigma^2$, $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t+h}) = 0$ $(h \neq 0)$.

- $\mathbb{E}[Y_t] = a + bt$ **depende de** t → **no** estacionaria si $b \neq 0$.
- $\mathrm{Var}(Y_t) = \sigma^2$ (constante).
- $\gamma_Y(h) = 0$ para $h \neq 0$.

Conclusión: no estacionaria salvo $b = 0$.

Inciso 2 (c) — ¿Cómo hacerla estacionaria?

- **Detrending** (quitar $a + bt$):
  $Z_t = Y_t - (a + bt) = \varepsilon_t$ (ruido blanco, estacionario).
  En práctica: estima $\hat{a}, \hat{b}$ por MCO y usa residuales.
- **Diferenciación**:
  $\Delta Y_t = Y_t - Y_{t-1} = b + (\varepsilon_t - \varepsilon_{t-1})$ (MA(1) + constante, estacionario).
  Centra si quieres media cero: $\Delta Y_t - b$.

Inciso 3 — Función de autocorrelación muestral

Definición:

$$r(h) = \frac{\sum_{t=h+1}^{T}(X_t - \bar{X})(X_{t-h} - \bar{X})}{\sum_{t=1}^{T}(X_t - \bar{X})^2}, \quad h = 0, 1, \ldots$$

¿Qué mide?

- **Relación lineal** entre X_t y X_{t-h}.
- $r(h) \in [-1, 1]$; cercano a 0 → poca memoria; cercano a ± 1 → fuerte dependencia.

Uso típico: detectar memoria, periodicidades, y guiar elección de modelos (p. ej., AR vs MA).

🧪 Inciso 3 — Ruido Blanco

Proceso simulado

$$X_t \sim \mathcal{N}(0, 1)$$

- No presenta memoria ni correlación temporal.
- La ACF debe ser ≈ 0 para todo $h \geq 1$.

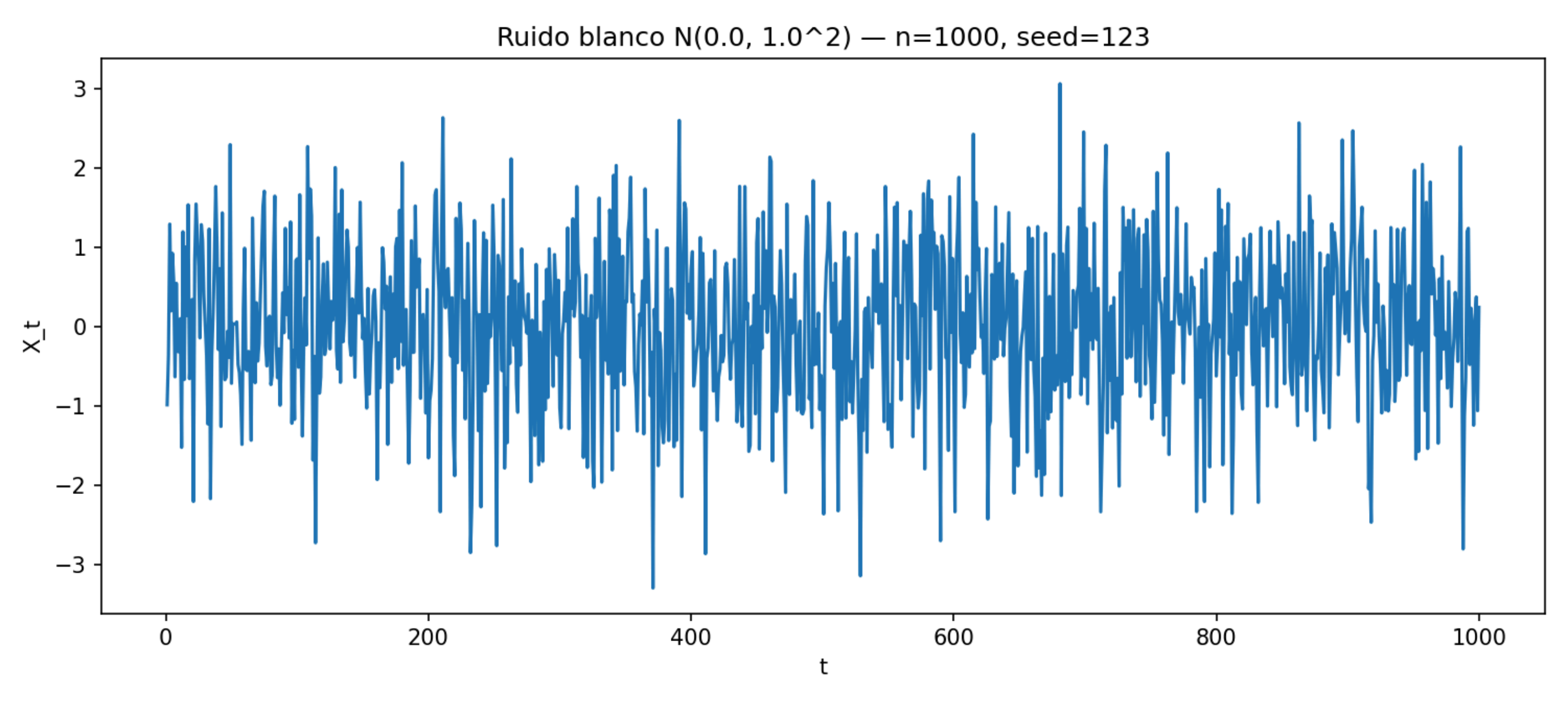
⚙ Implementación

```
cd "Inciso 3"
python -m venv .venv
.venv\Scripts\Activate.ps1     # En Windows (PowerShell)
# source .venv/bin/activate    # En macOS/Linux
pip install numpy matplotlib

python simulate_white_noise_acf.py --n 1000 --lags 40
```

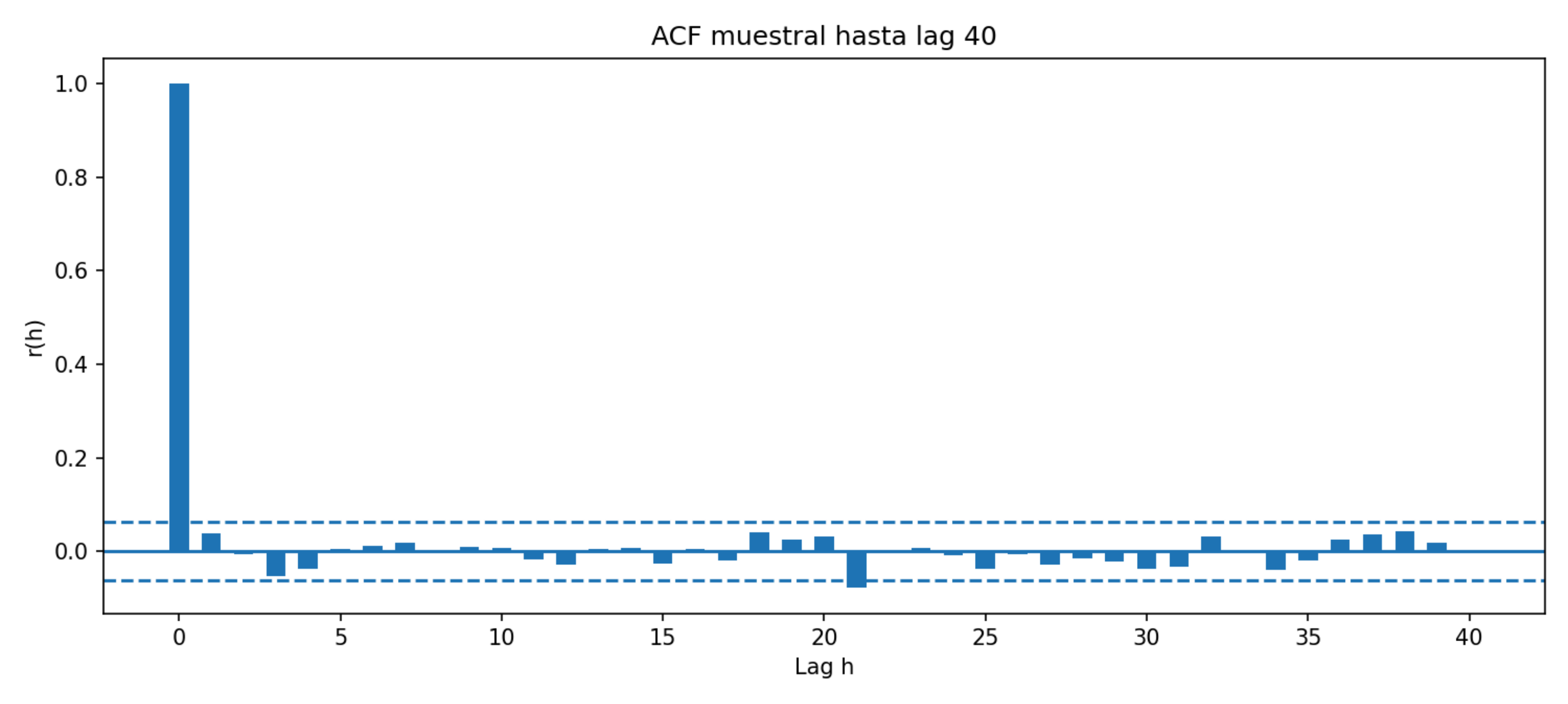
- Se generan 1000 observaciones.
- Se calcula la ACF muestral.
- Se grafican la serie y su ACF con bandas $\pm 1.96/\sqrt{T}$.

📈 Resultados — Serie simulada



- Distribución normal sin tendencia.
- Valores centrados en cero.

📉 Resultados — ACF del ruido blanco



- La ACF es ≈ 0 para todos los lags (dentro de bandas).
- Confirma independencia temporal → **ruido blanco**.

Inciso 4 — $Y = X\beta + \varepsilon$ vs $X_t = \phi X_{t-1} + \varepsilon_t$

Regresión clásica (OLS):

- $\mathbb{E}[\varepsilon] = 0, \mathrm{Var}(\varepsilon) = \sigma^2 I$,
- **Independencia** (no autocorrelación): $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ si $i \neq j$.

AR(1):

- Observaciones **dependen del pasado** → hay autocorrelación.

Si ignoras la dependencia temporal en OLS:

- **Se viola la independencia de errores**.
- Errores estándar y tests (t/F) **incorrectos**; estimadores no eficientes.

Inciso 5 (a) — AR(1): media y varianza bajo estacionariedad

Modelo: $X_t = \phi X_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$.

- $\mathbb{E}[X_t] = 0$ (para $|\phi| < 1$).
- $\mathrm{Var}(X_t) = \dfrac{\sigma^2}{1-\phi^2}$.

Inciso 5 (b) — Condición de estacionariedad

- **Si** $|\phi| < 1$ ⇒ $X_t = \sum_{j\geq 0} \phi^j \varepsilon_{t-j}$ (converge en L^2) ⇒ estacionario.
- **Si** $|\phi| \geq 1$ ⇒ varianza no acotada (p. ej., $\phi = 1$: random walk) ⇒ no estacionario.

Equivalente: AR(1) es estacionario **ssi** $|\phi| < 1$.

Inciso 5 (c) — ACF del AR(1)

Autocovarianza: $\gamma(h) = \dfrac{\sigma^2}{1-\phi^2}\phi^{|h|}$.

ACF: $\rho(h) = \dfrac{\gamma(h)}{\gamma(0)} = \phi^{|h|}$.

Interpretación: decae **geométricamente** con h.

🔁 Inciso 6 — Proceso AR(1)

Modelo

$$X_t = \phi X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

Parámetros:

$$\phi = 0.7, \quad \sigma = 1, \quad T = 200$$

- La ACF teórica es $r(h) = \phi^{|h|}$.
- Se espera un decrecimiento geométrico con h.

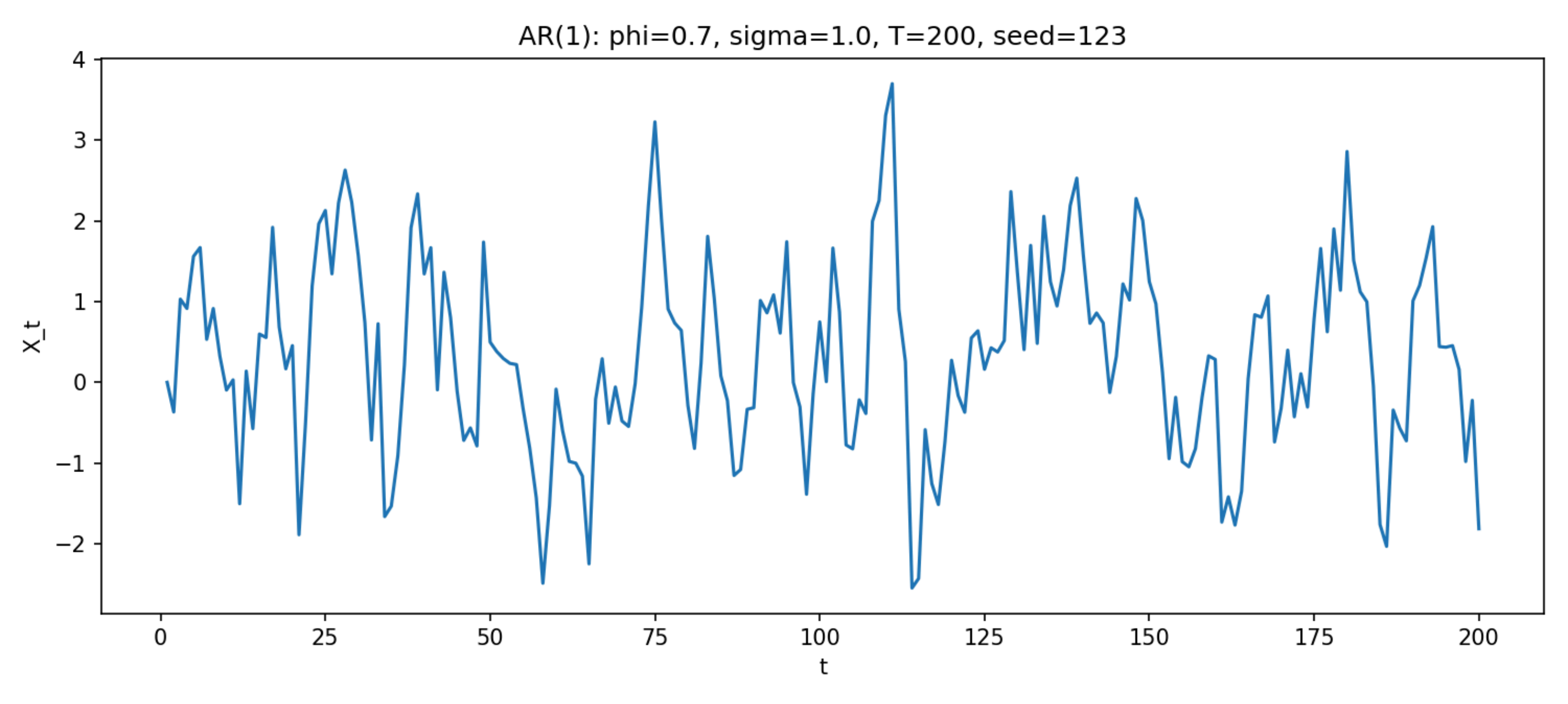
⚙ Implementación

```
cd "Inciso 6"
python -m venv .venv
.venv\Scripts\Activate.ps1     # En Windows (PowerShell)
# source .venv/bin/activate    # En macOS/Linux
pip install numpy matplotlib

python simulate_ar1_acf.py --phi 0.7 --sigma 1 --T 200 --lags 40
```

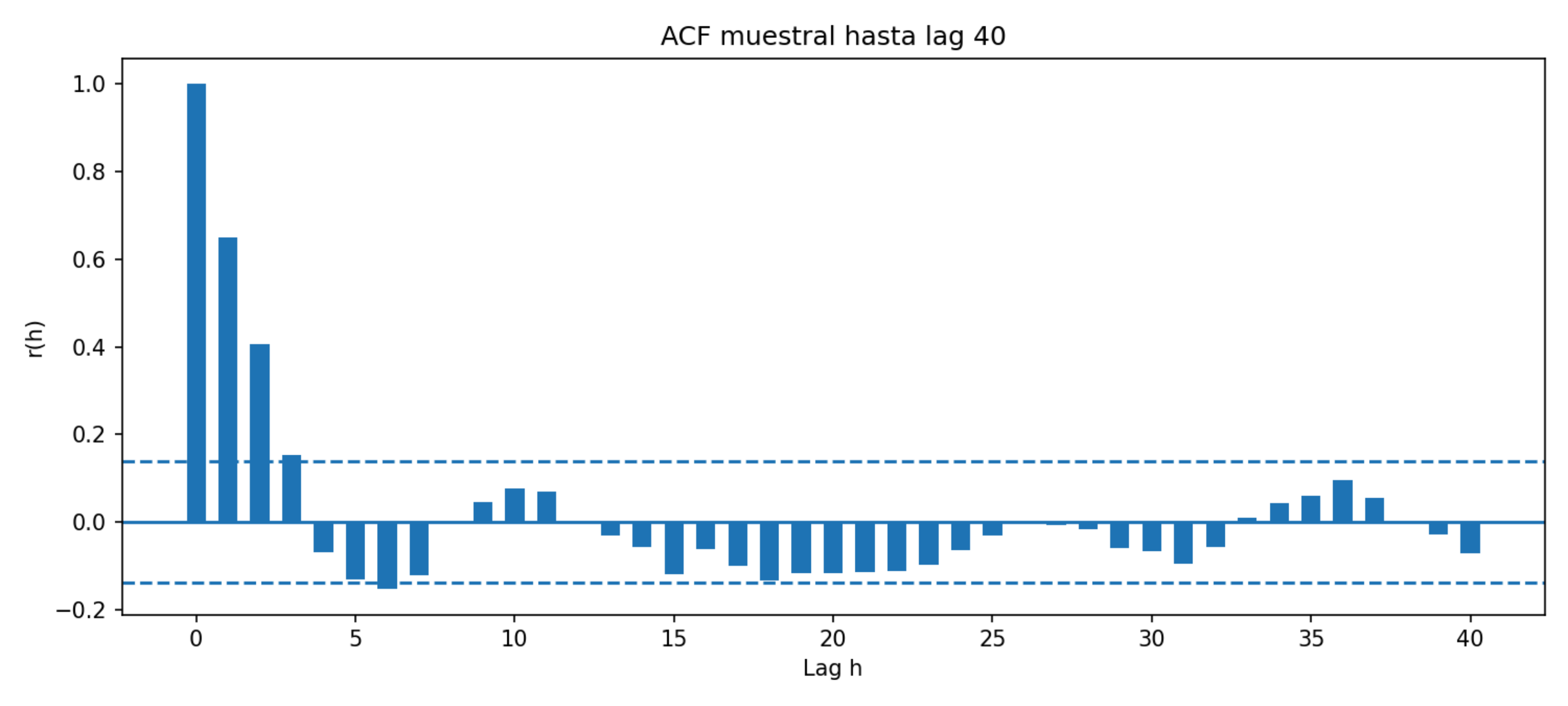
- Simula AR(1) estacionario.
- Calcula la ACF muestral.
- Estima $\hat{\phi}$ con OLS en $X_t = \beta X_{t-1} + \varepsilon_t$.

📈 Serie AR(1) simulada



- Presenta correlación positiva entre periodos consecutivos.
- Valores suavizados por el efecto de memoria.

📉 ACF del proceso AR(1)



- ACF decrece geométricamente con h.
- Forma típica de un **proceso autoregresivo**.
- Confirma dependencia temporal.

📊 Estimación de ϕ

Regresión lineal:

$$X_t = \beta X_{t-1} + \varepsilon_t$$

Resultado:

$$\hat{\phi} \approx 0.7$$

➡️ Estimación **muy cercana** al valor real → OLS válido bajo estacionariedad.

⚠️ Caso no estacionario

- Si $|\phi| \geq 1$:
  - Varianza diverge: $ \operatorname{Var}(X_t) \to \infty $
  - No se cumple estacionariedad.
  - OLS produce estimaciones **no confiables**.

Ejemplos:

- $\phi = 1$: Random Walk (camino aleatorio).
- $\phi = -1$: Alternancia con acumulación de choques.

🧠 Conclusiones

1. **Ruido blanco** → no hay dependencia temporal; ACF ≈ 0.
2. **AR(1)** → dependencia que decae como $ r(h)=\phi^h $.
3. Estimación OLS de ϕ es **consistente** si $|\phi| < 1$.
4. Si $|\phi| \geq 1$, el proceso **no es estacionario** y la varianza se vuelve infinita.

📚 Referencias

- Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (2008). *Time Series Analysis: Forecasting and Control*. Wiley.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Chatfield, C. (2003). *The Analysis of Time Series: An Introduction*. Chapman & Hall/CRC.

🏁 Fin de la presentación

Gracias por su atención.