



Business AI Lab

Phân tích Chuỗi thời gian: Từ Tính dừng đến Mô hình ARMA

Lý thuyết, Kiểm định và Quy trình Xây dựng Mô hình

Giảng viên: Trong-Nghia Nguyen
Khoa KHDL & TTNT (Faculty of DS & AI)
Môn học: Chuỗi thời gian (Time-series)
Học kỳ: Mùa xuân 2026

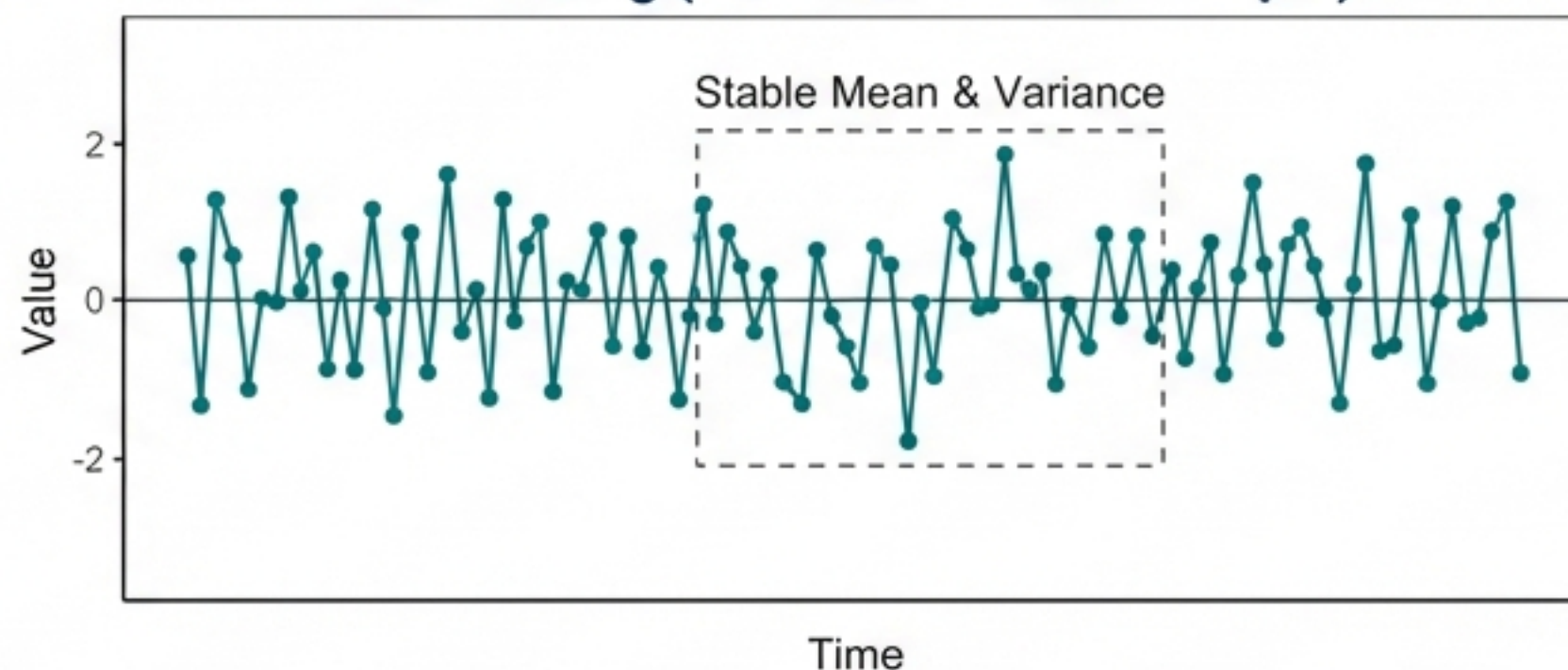
Định nghĩa về Tính dừng (Stationarity)

Tính dừng đề cập đến sự ổn định thống kê của một chuỗi theo thời gian. Đây là giả định nền tảng cho các mô hình AR, MA, và ARMA.

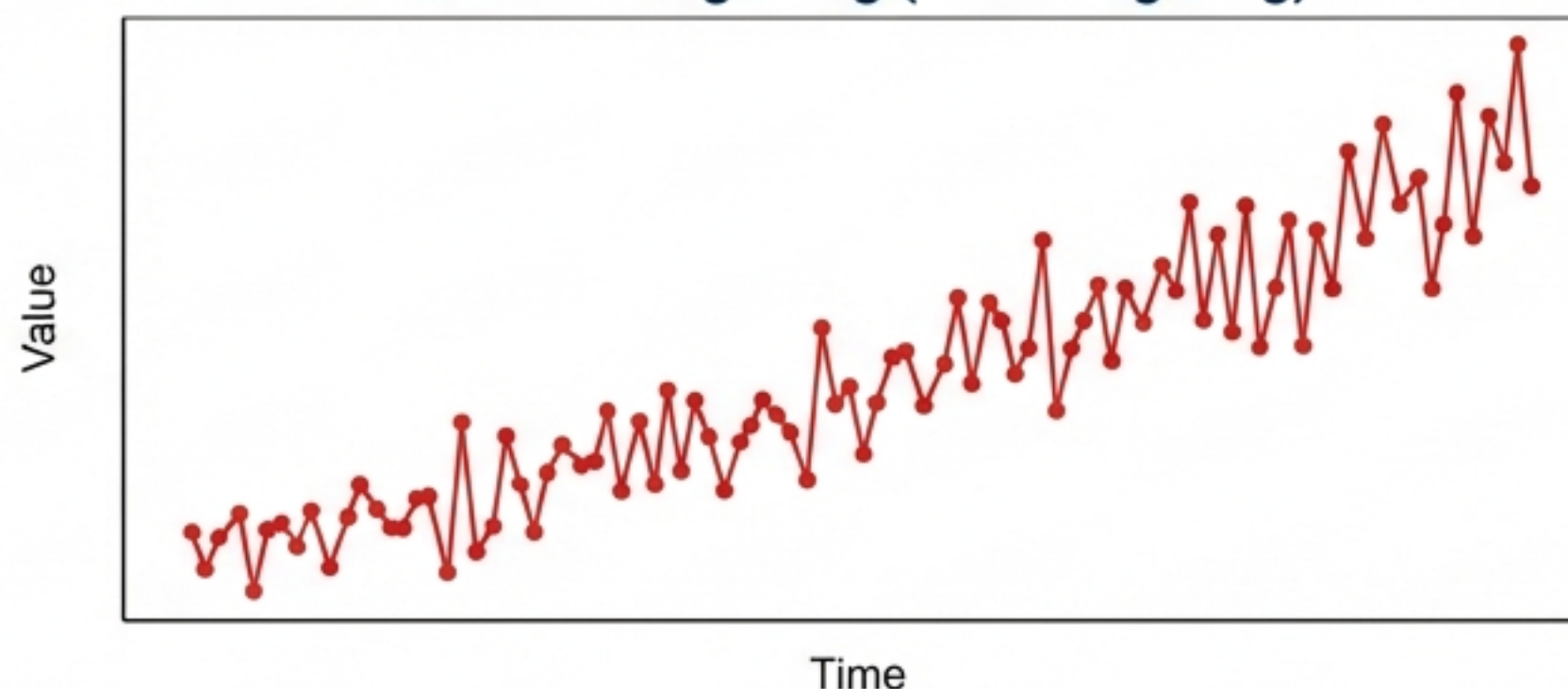
Hai loại tính dừng:

1. **Tính dừng chặt** (Strict Stationarity): Toàn bộ phân phối xác suất không đổi theo thời gian. (Quá nghiêm ngặt cho thực tế).
2. **Tính dừng yếu** (Weak Stationarity):
 - Kỳ vọng (Mean) là hằng số: $\mathbb{E}[X_t] = \mu$
 - Phương sai (Variance) là hằng số: $\text{Var}(X_t) = \sigma^2$
 - Hiệp phương sai chỉ phụ thuộc vào độ trễ (lag), không phụ thuộc thời gian t .

Chuỗi Dừng (Mean & Variance ổn định)

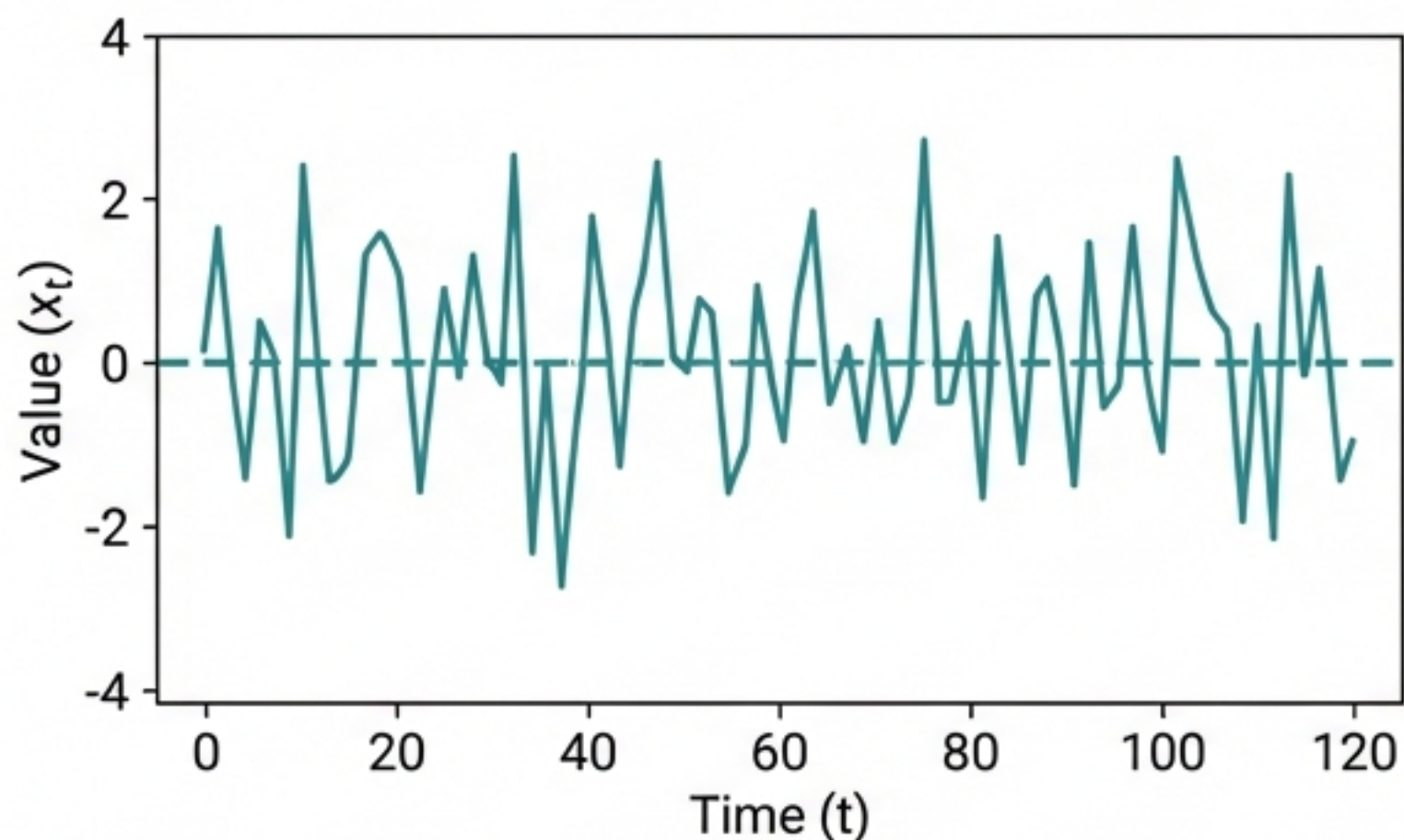


Chuỗi Không Dừng (Xu hướng tăng)



Trực quan hóa: Phân biệt Chuỗi Dừng và Không Dừng

Chuỗi A: Dừng (Stationary)



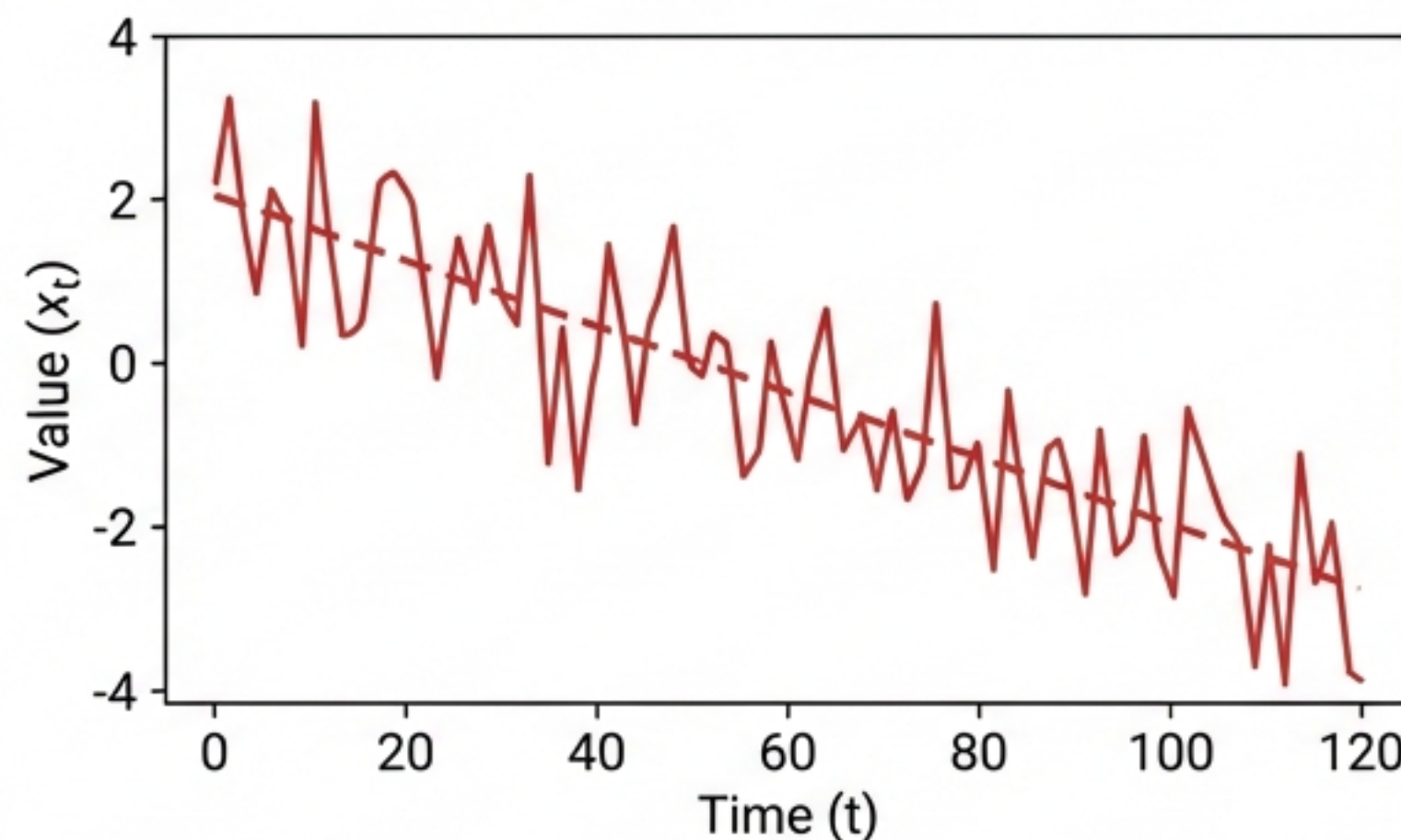
Mô hình: $x_t = \epsilon_t$

Mean: Hằng số (≈ 0)

Variance: Ổn định

Kết luận: Thỏa mãn giả định mô hình.

Chuỗi B: Không Dừng (Non-Stationary)



Mô hình: $x_t = -0.03t + \epsilon_t$

Mean: Thay đổi theo thời gian ($\mathbb{E}[x_t]$ phụ thuộc t)

Variance: Tương đối ổn định

Kết luận: Vi phạm giả định. Cần xử lý.

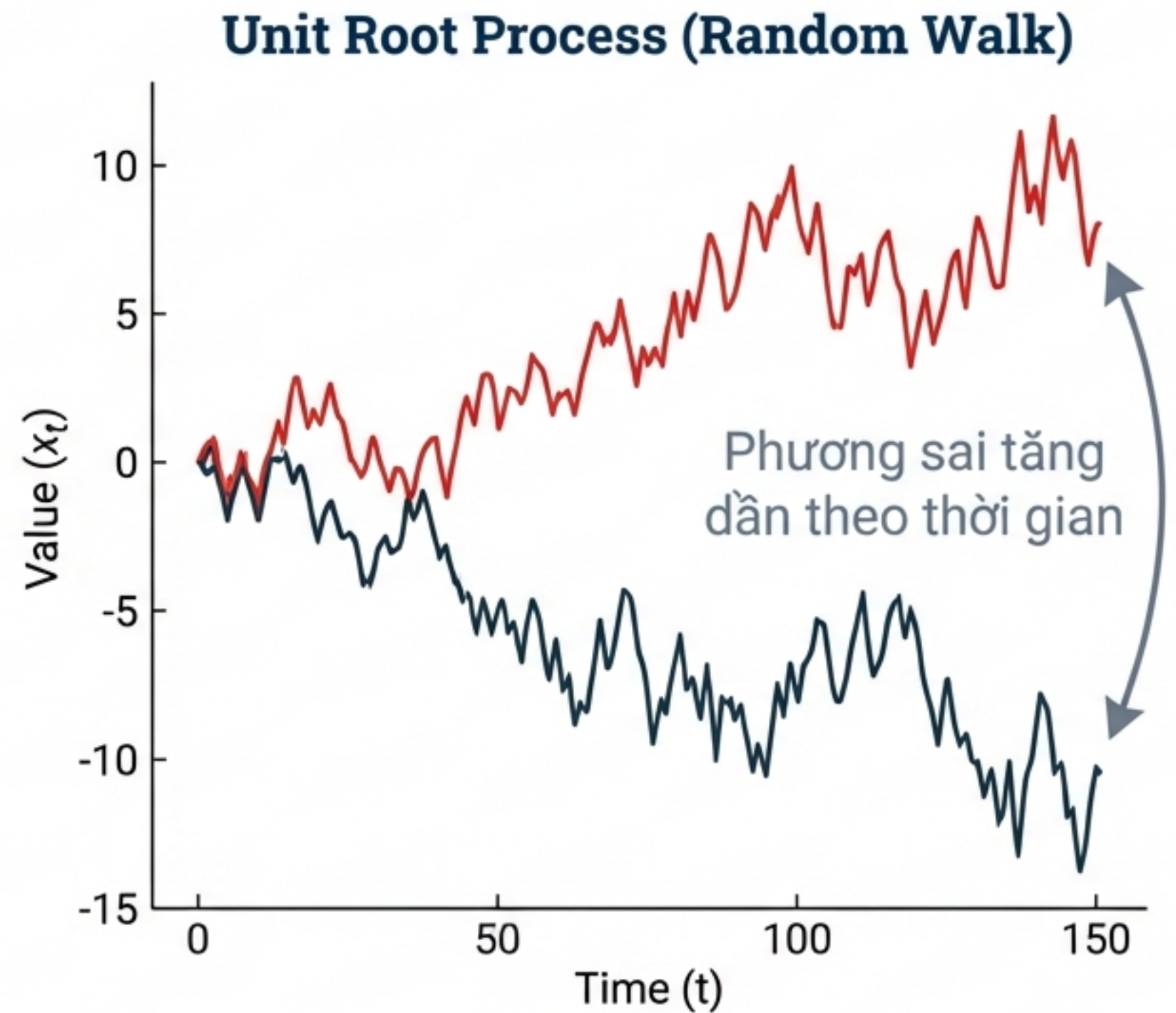
Cạm bẫy của Nghiệm đơn vị (The Unit Root Trap)

Một số chuỗi dữ liệu trông có vẻ ổn định trong ngắn hạn nhưng thực chất là Bước ngẫu nhiên (Random Walk). Đây là trường hợp đặc biệt của AR(1) khi hệ số $\phi = 1$.

Mô hình: $x_t = x_{t-1} + \epsilon_t$

Hậu quả:

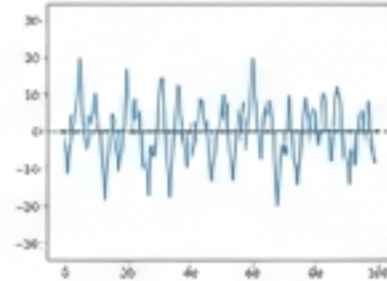
1. **Tác động vĩnh viễn (Permanent Effect):** Các cú sốc không suy giảm về 0 mà tích tụ mãi mãi.
2. **Phương sai tăng dần:** Biến động của chuỗi lớn dần theo thời gian.
3. **Kết luận:** Kiểm tra bằng mắt là không đủ. Cần kiểm định thống kê.



Kiểm định Nghiệm đơn vị (Unit Root Tests)

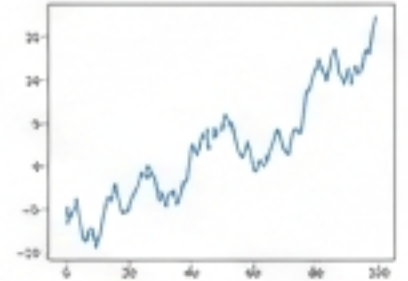
H1: Stationary (Dừng)

Chuỗi dao động quanh giá trị trung bình (Mean-Reverting).



H0: Unit Root (Có nghiệm đơn vị)

Chuỗi có xu hướng ngẫu nhiên (Stochastic Trend).



Công cụ chính: Augmented Dickey–Fuller (ADF) Test

Quy tắc Quyết định (Decision Rule):

Kiểm tra p-value hoặc ADF Statistic.

- Nếu $p\text{-value} < 0.05$ → Bác bỏ H_0 → Kết luận: **Chuỗi Dừng.**
- Nếu $p\text{-value} > 0.05$ → Chấp nhận H_0 → Kết luận: **Chuỗi Không Dừng.**

Giải pháp: Kỹ thuật Sai phân (Differencing)

Nếu chuỗi không dừng, ta sử dụng sai phân để loại bỏ xu hướng.

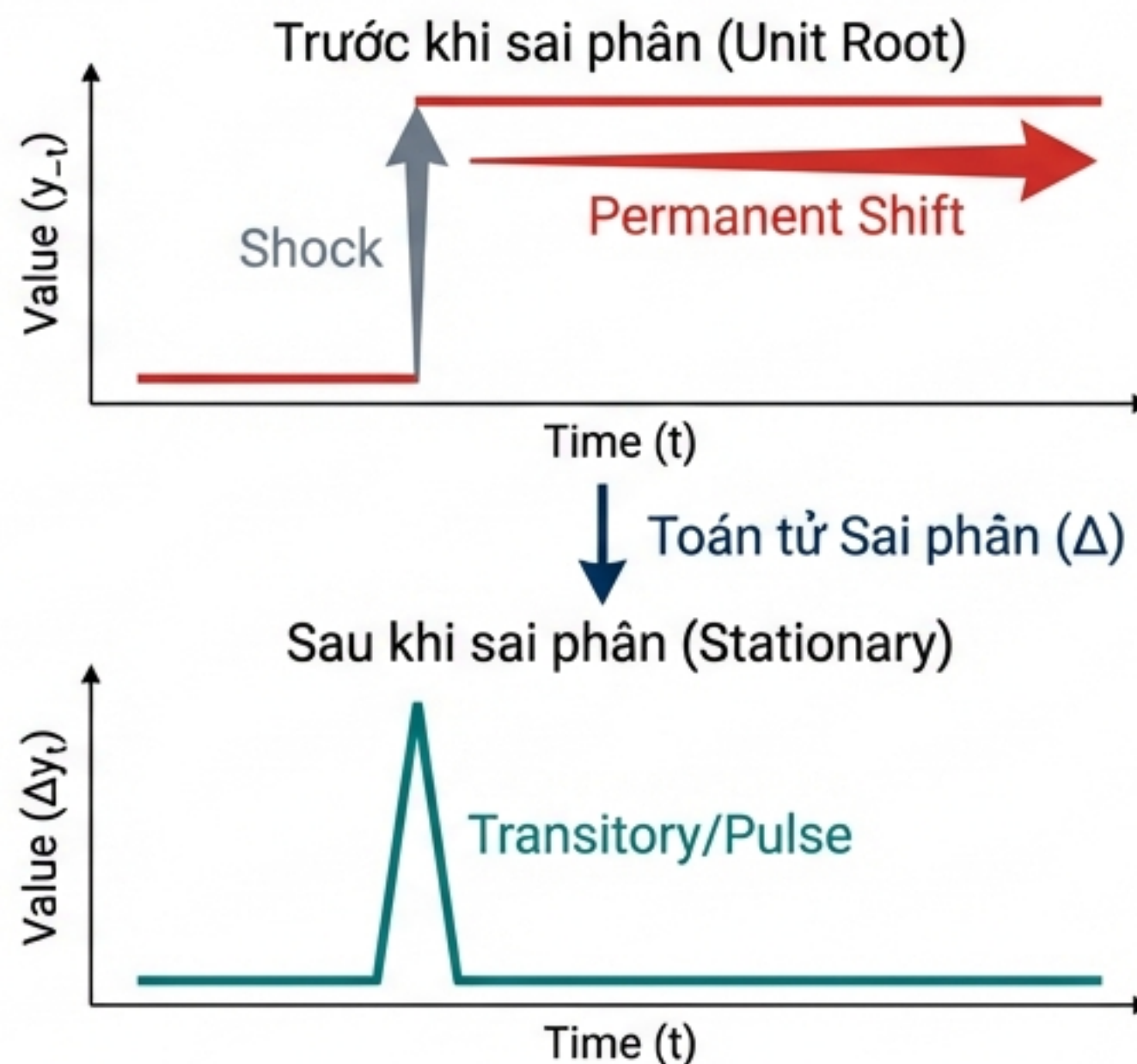
Định nghĩa Sai phân bậc 1:

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

Tác dụng:

- Loại bỏ xu hướng ngẫu nhiên (stochastic trend).
- Ổn định giá trị trung bình về 0.
- Biến đổi Random Walk thành White Noise.

Visual Intuition



Tổng quan Mô hình ARMA(p, q)

Kết hợp giữa AR (AutoRegressive) và MA (Moving Average)

$$x_t = c + \underbrace{\sum_{i=1}^p \phi_i x_{t-i}}_{\text{Phần AR (Memory)}} + \underbrace{\sum_{j=1}^q \theta_j \epsilon_{t-j}}_{\text{Phần MA (Shock)}} + \epsilon_t$$

AR (Bậc p)

- Tự hồi quy.
- Giá trị hiện tại phụ thuộc vào các giá trị trong quá khứ.
- Đại diện cho “trí nhớ” (memory) của chuỗi.

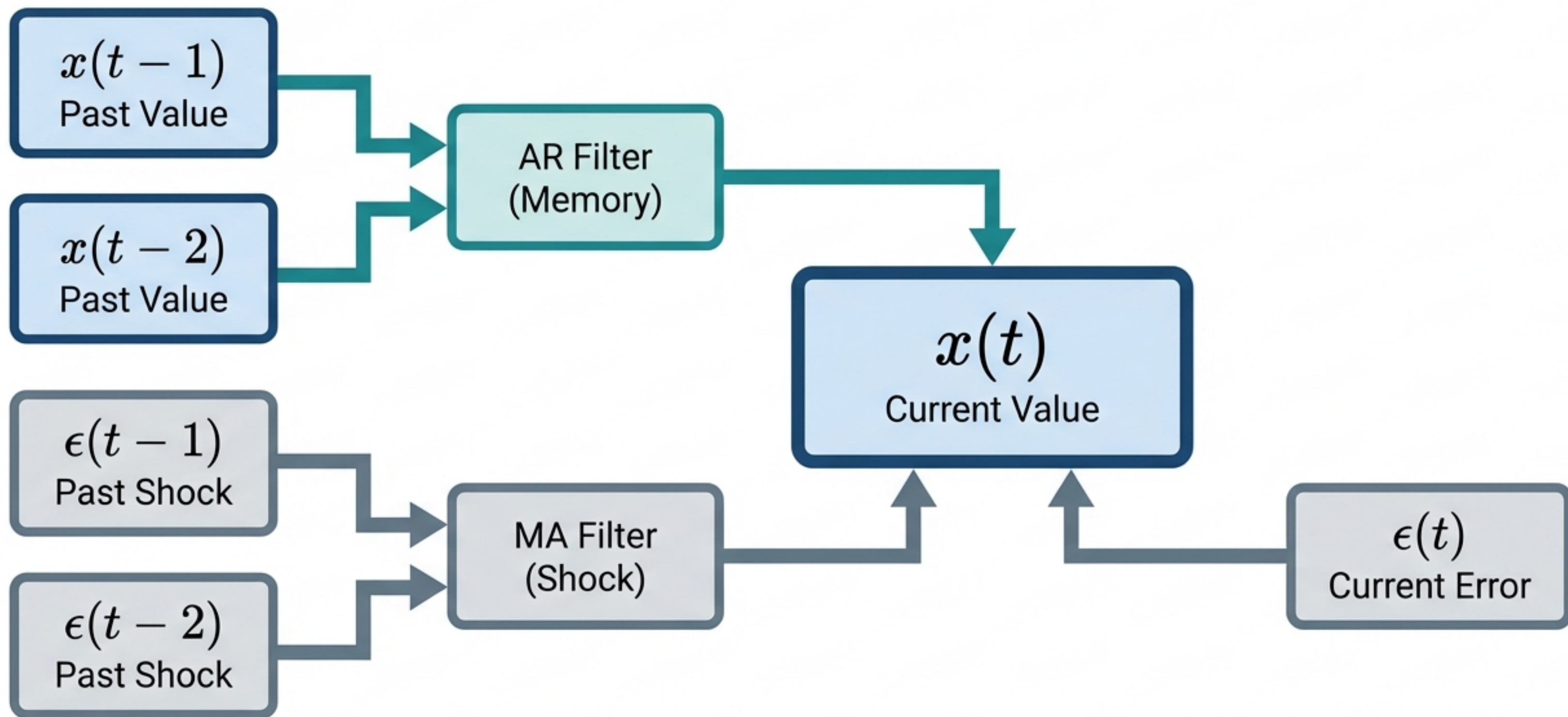
MA (Bậc q)

- Trung bình trượt.
- Giá trị hiện tại phụ thuộc vào các sai số/cú sốc trong quá khứ.
- Đại diện cho tác động ngắn hạn.

Điều kiện

- Chỉ áp dụng cho chuỗi dừng (Stationary Series).
- Nhiễu ϵ_t là White Noise.

Cấu trúc hoạt động của ARMA



Sơ đồ thể hiện cách ARMA tổng hợp thông tin từ quá khứ (AR) và các cú sốc (MA) để hình thành giá trị hiện tại.

Đọc kết quả thực nghiệm (Python Output)

SARIMAX Results						
	coef	std err	z	P> z	[0.025	0.975]
const	-0.0057	2.716	-0.002	0.998	-5.328	5.317
ar.L1	0.4078	0.074	5.527	0.000	0.263	0.552
ma.L1	0.2042	0.081	2.515	0.012	0.045	0.363
sigma2	395.7234	24.175	16.369	0.000	348.342	443.105

Hệ số AR: 0.4078. Giá trị hiện tại chịu ảnh hưởng ~40% từ giá trị $t - 1$.

P-value < 0.05. Hệ số có ý nghĩa thống kê (Significant).

P-value = 0.012 (< 0.05). Thành phần MA có ý nghĩa.

Tiêu chuẩn lựa chọn mô hình (Information Criteria)

Cân bằng giữa Độ khớp (Fit) và Độ phức tạp (Complexity)

AIC (Akaike)

- Phạt nhẹ độ phức tạp.
- Ưu tiên: Hiệu suất dự báo.
- Công thức:

$$-2 \log(L) + 2k$$

HQIC (Hannan-Quinn)

- Mức phạt trung bình.
- Cân bằng giữa AIC và BIC.

BIC (Bayesian)

- Phạt nặng độ phức tạp.
- Ưu tiên: Mô hình đơn giản.
- Công thức:

$$-2 \log(L) + k \log(n)$$

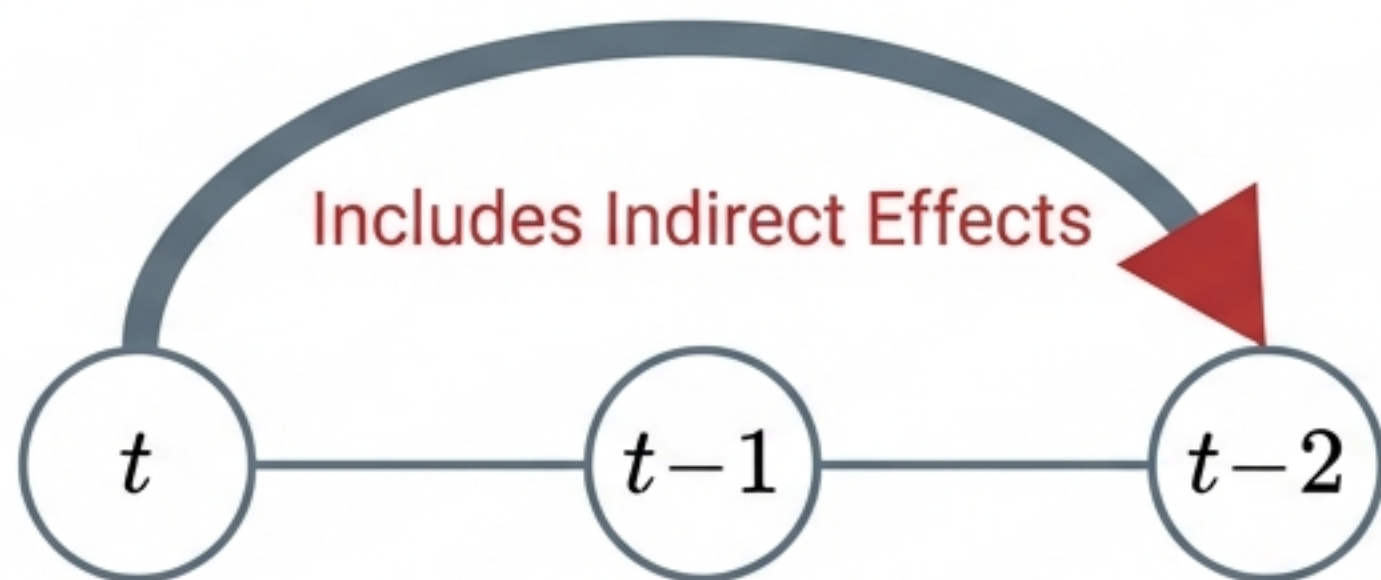
Quy tắc vàng: Giá trị càng thấp = Mô hình càng tốt (Lower is Better)

Công cụ chẩn đoán: ACF & PACF

Xác định cấu trúc tương quan của chuỗi

ACF (Autocorrelation Function)

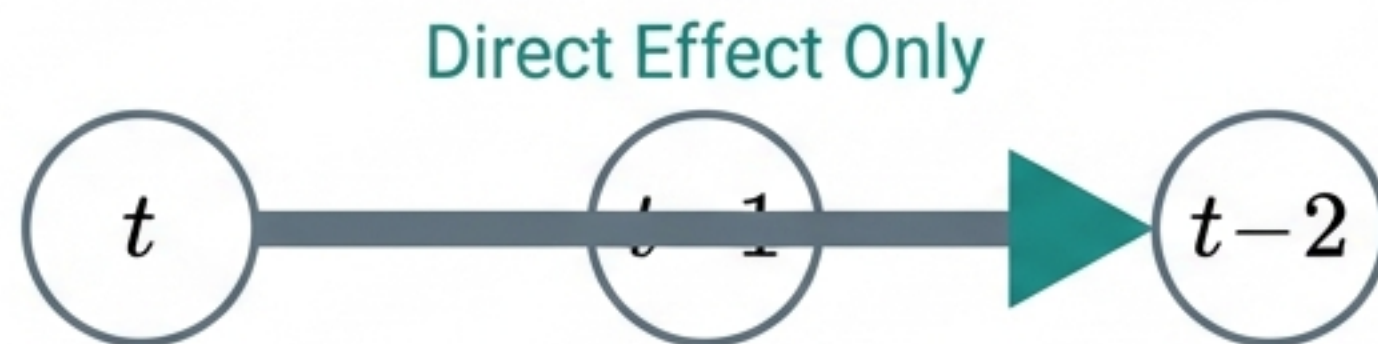
Tương quan giữa x_t và x_{t-k} .



Bao gồm cả tác động gián tiếp qua các lag trung gian.

PACF (Partial Autocorrelation Function)

Tương quan riêng phần.



Chỉ đo lường tác động trực tiếp, đã loại bỏ ảnh hưởng trung gian.

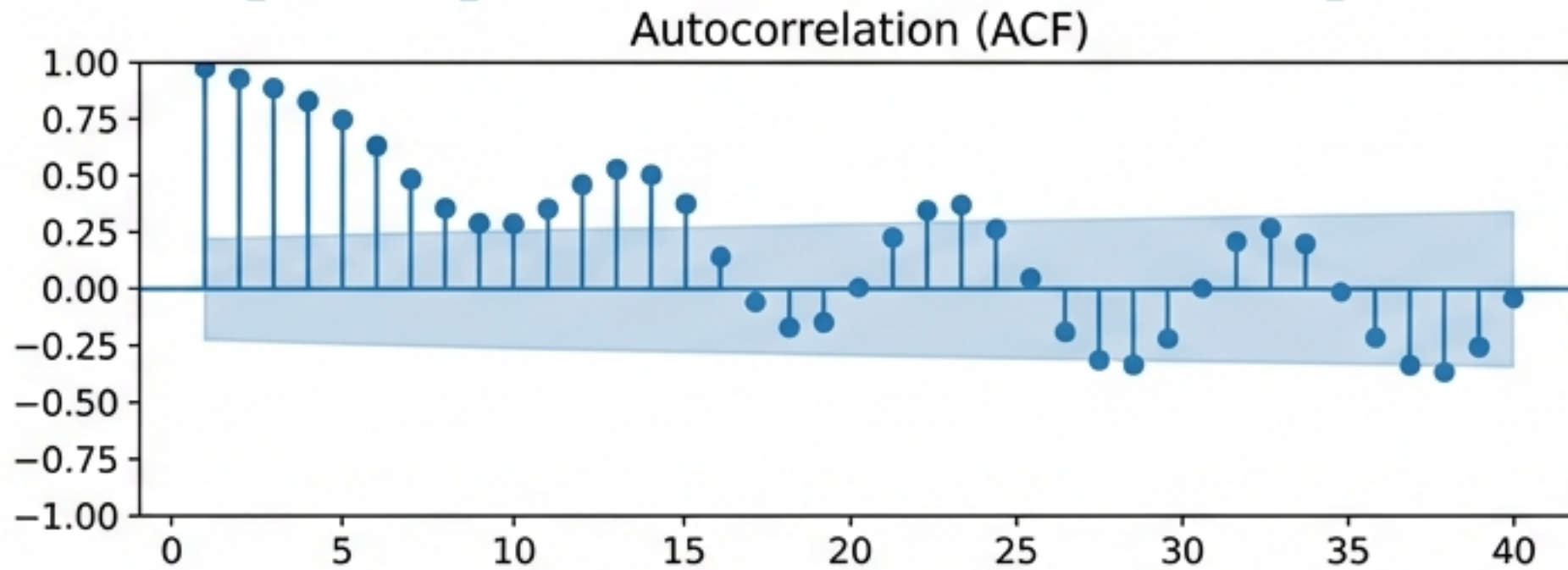
Mục tiêu: Sử dụng hình dạng biểu đồ để xác định bậc p (AR) và q (MA).

Quy tắc xác định bậc p và q (Rules of Thumb)

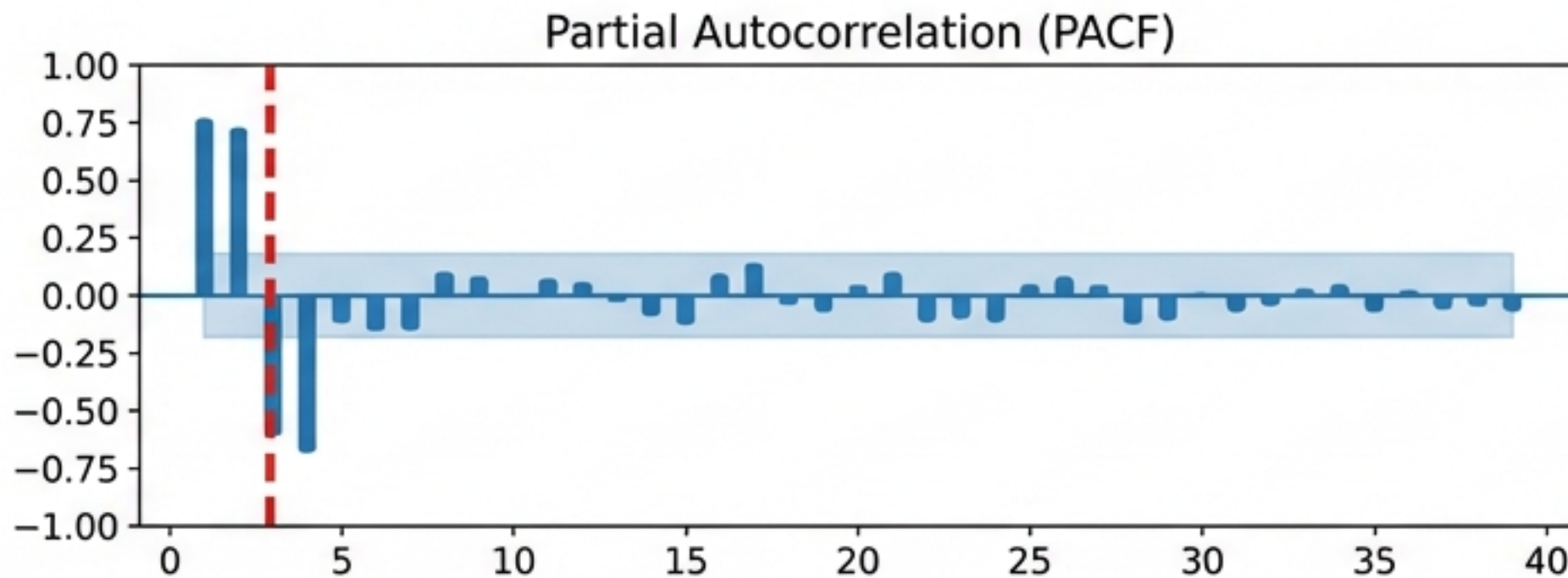
Mô hình	ACF Pattern	PACF Pattern
AR(p) (AutoRegressive)	Giảm dần (Decay) - dạng sóng hoặc mũ. 	Cắt (Cut-off) tại lag p . 
MA(q) (Moving Average)	Cắt (Cut-off) tại lag q . 	Giảm dần (Decay). 
ARMA(p, q)	Giảm dần (Decay). 	Giảm dần (Decay). 

Mẹo thực hành: Dùng PACF để tìm p. Dùng ACF để tìm q.

Ví dụ thực tế: Vết đen mặt trời (Sunspots Data)



ACF giảm dần chậm
→ Gợi ý thành phần AR.



Cut-off tại Lag 2

Các lag sau Lag 2 nằm trong
khoảng tin cậy (không có ý nghĩa).

Kết luận: Dữ liệu phù hợp với mô hình **AR(2)**.

Tổng kết Quy trình Phân tích (The Analyst's Checklist)

- **1. Kiểm tra dữ liệu (Visual Check)**
Vẽ biểu đồ chuỗi thời gian. Quan sát xu hướng và phương sai.
- **2. Chẩn đoán (Diagnosis)**
Sử dụng kiểm định ADF. Nếu **p-value > 0.05** → Có nghiệm đơn vị (Unit Root).
- **3. Xử lý (Preprocessing)**
Sử dụng sai phân (Δy_t) để đưa về chuỗi dừng.
- **4. Nhận dạng (Identification)**
Quan sát ACF (tìm q) và PACF (tìm p).
- **5. Ước lượng (Estimation)**
Chạy mô hình. So sánh chỉ số AIC/BIC (càng thấp càng tốt).
- **6. Kiểm định (Validation)**
Kiểm tra ý nghĩa thống kê của hệ số và đảm bảo phần dư (residual) là **nhieu trắng**.

Cảm ơn (Thank You)

Q&A Session

Trong-Nghia Nguyen
Khoa KHDL & TTNT | Spring 2026