

Chương 6: TỰ TƯƠNG QUAN

1. KHÁI NIỆM TỰ TƯƠNG QUAN
2. HẬU QUẢ HIỆN TƯỢNG TỰ TƯƠNG QUAN
3. PHÁT HIỆN TỰ TƯƠNG QUAN
4. KHẮC PHỤC HIỆN TƯỢNG TỰ TƯƠNG QUAN

6.1 KHÁI NIỆM TỰ TƯƠNG QUAN (AUTOCORRELATION)

6.1.1 Khái niệm.

Tự tương quan (TTQ) là sự tương quan giữa các sai số ngẫu nhiên được sắp xếp theo thứ tự thời gian (trong dữ liệu chuỗi thời gian) hoặc không gian (trong dữ liệu chéo).

Tức là: $\text{Corr}(u_i, u_j) \neq 0$ với $i \neq j$

Hiện tượng TTQ thường xảy ra với dữ liệu chuỗi thời gian. Vì vậy, phương trình hồi quy chương này được viết dưới dạng:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + U_t$$

6.1.1 Khái niệm.

Các tên gọi khác của TTQ trong dữ liệu chuỗi thời gian:

- Serial Correlation – tương quan chuỗi
- Autocorrelation – tự tương quan
- AutoRegression – tự hồi quy

❖ **TTQ bậc 1 (AR (1))**: Sai số u_t tương quan với u_{t-1} .

Phương trình AR(1): $U_t = \rho U_{t-1} + \varepsilon_t$

Với hệ số tự tương quan ρ thỏa $|\rho| < 1$ và ε_t không có TTQ.

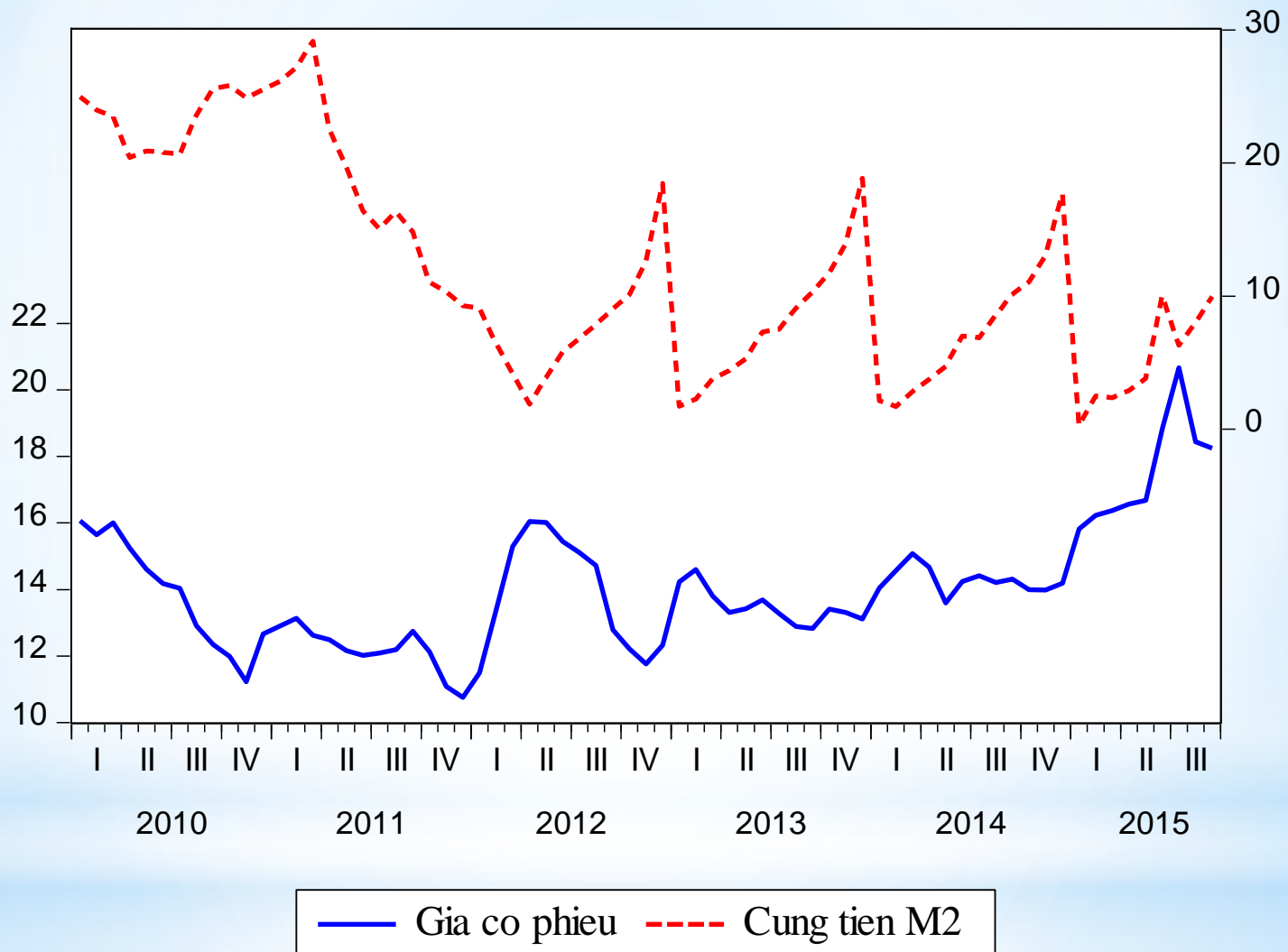
❖ **TTQ bậc m (AR(m))**: Sai số u_t tương quan với các sai số ở m kỳ trước đó. ($u_{t-1}, u_{t-2}, \dots, u_{t-m}$)

Phương trình AR(m): $U_t = \rho_1 + \rho_2 U_{t-1} + \dots + \rho_{m+1} U_{t-m} + \varepsilon_t$

6.1.2 Nguyên nhân hiện tượng TTQ.

Nguyên nhân khách quan:

- *Yếu tố mùa vụ*: thường xuất hiện với các số liệu có tần suất nhỏ hơn 1 năm, ví dụ: tần suất tháng, quý...
- *Yếu tố xu thế*: thường xuất hiện với các số liệu có chiều dài chuỗi thời gian khá lớn.
- *Hiện tượng mạng nhện*: biến phụ thuộc bị ảnh hưởng bởi giá trị của biến độc lập ở kỳ trước đó.
- *Các độ trễ*: Biến phụ thuộc ở kỳ t phụ thuộc vào chính nó ở các kỳ trước.



6.1.2 Nguyên nhân hiện tượng TTQ.

Nguyên nhân chủ quan:

- *Xử lý số liệu*: phép lấy trung bình, phép nội suy và ngoại suy
- *Chọn mô hình không phù hợp*:
 - + Thiếu biến quan trọng.
 - + Sai dạng hàm

6.2 HẬU QUẢ HIỆN TƯỢNG TỰ TƯƠNG QUAN

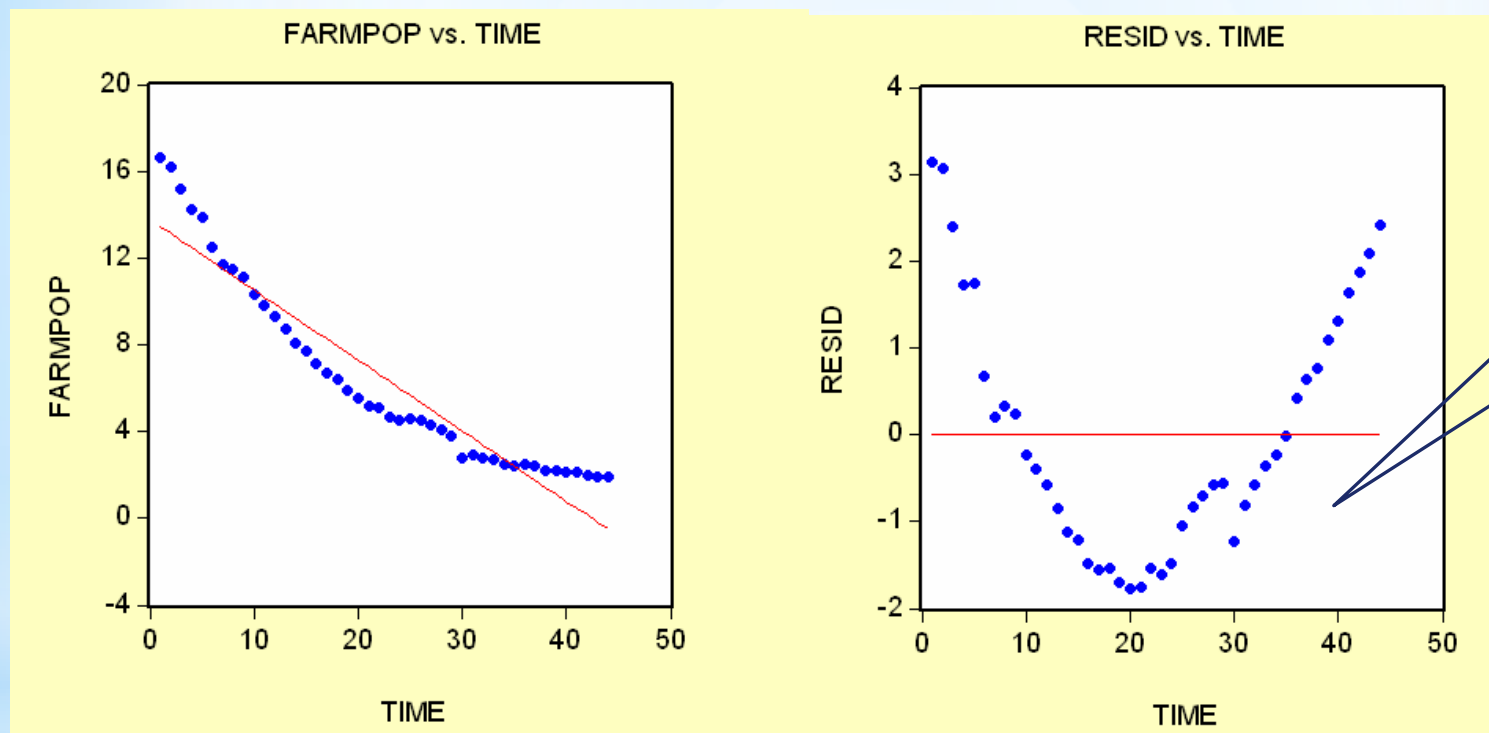
Giả sử các giả thiết GT2 – GT5 thỏa mãn. Khi đó:

- Các ước lượng OLS vẫn không chệch và vững.
- Phương sai của các hệ số ước lượng OLS bị chệch nên:
 - ✓ Bài toán tìm khoảng tin cậy cho kết quả không chính xác.
 - ✓ Bài toán kiểm định giả thuyết thống kê về các hệ số là không đáng tin cậy

6.3 PHÁT HIỆN TỰ TƯƠNG QUAN

6.3.1 Sử dụng đồ thị

- Vẽ đồ thị phân tán giữa (e_t, e_{t-1}) hoặc (e_t, t) .



Xét mô hình hồi quy: $FARMPOP = \beta_1 + \beta_2 TIME + u$,
Trong đó $TIME$ là biến xu thế.

6.3.2 Kiểm định tự tương quan bậc 1

Xét phương trình hồi quy tổng thể:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + U_t \quad (*)$$

Biến ngoại sinh chặt:

X_j được gọi là biến ngoại sinh chặt nếu:

$\text{Cov}(X_j, u_s) = 0$ với mọi s .

6.3.2 Kiểm định tự tương quan bậc 1

Xét phương trình hồi quy tổng thể:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + U_t \quad (*)$$

Và phương trình AR(1): $U_t = \rho U_{t-1} + \varepsilon_t$

i) Kiểm định t

B1: Ước lượng mô hình (*), thu được các phần dư e_t .

B2: Ước lượng e_t theo e_{t-1} với $t = 2, 3, \dots, n$

$$e_t = \rho e_{t-1} + v_t \quad (\text{có thể thêm hệ số chặn})$$

B3: Sử dụng thống kê t thông thường để kiểm định cặp giả thuyết: $H_0: \rho = 0$, $H_1: \rho \neq 0$.

Điều kiện áp dụng:
Các biến độc lập là ngoại sinh chặt và cỡ mẫu lớn

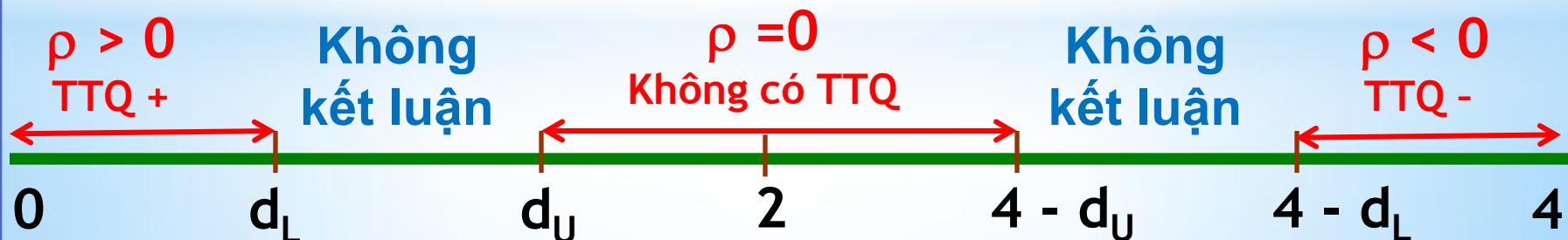
6.3.2 Kiểm định tương quan chuỗi bậc 1

ii) Kiểm định Durbin – Watson (DW) (1950)

B1: Ước lượng mô hình (*), thu được các phần dư e_t .

B2: Tính giá trị thống kê
$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad \approx 2(1 - \hat{\rho})$$

B3: Tra bảng *thống kê Durbin – Watson* với mức ý nghĩa α , số quan sát n và số biến độc lập $k' = k - 1$ để tìm d_U và d_L và ra quyết định theo nguyên tắc.



**Điều
kiện áp
dụng
kiểm
định
Durbin –
Watson**

Các giả thiết GT2 – GT5 phải thỏa mãn (trừ giả thiết về TTQ).

Chỉ kiểm định TTQ bậc 1.

Chuỗi dữ liệu liên tục.

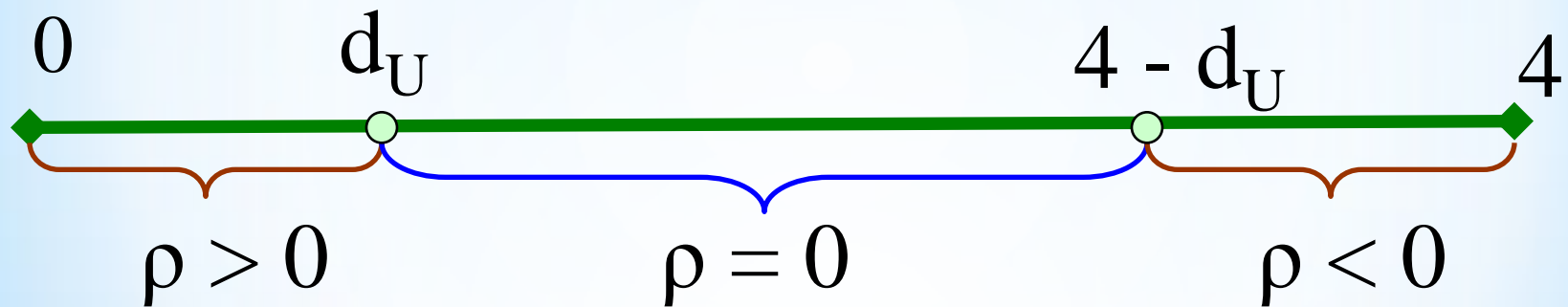
Mô hình hồi quy gốc có hệ số chặn.

Các biến giải thích là *biến ngoại sinh chặt*,

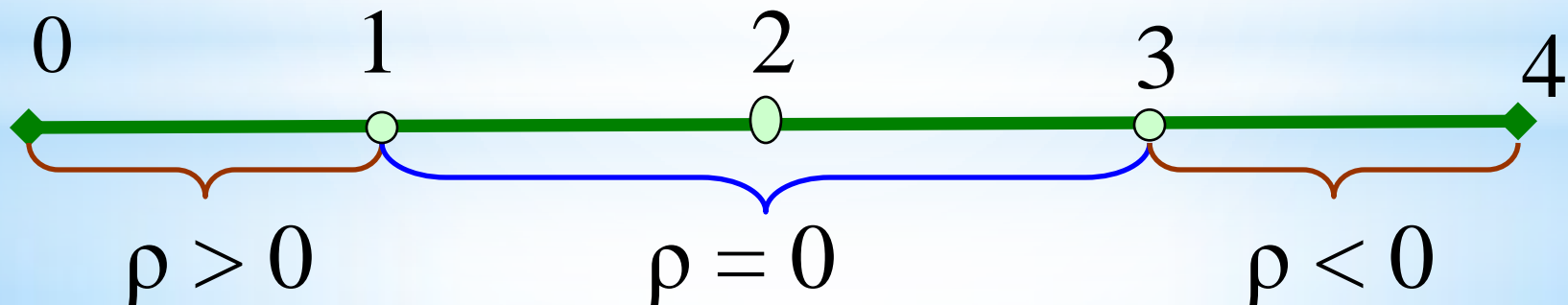
Khuyết điểm: có 2 vùng không quyết định được, và không có bảng tra khi n lớn.

* Kiểm định Durbin – Watson (DW) cải biên

Tra bảng mức ý nghĩa 2α , số quan sát n và số biến độc lập k' , ta có d_U và d_L :



* Kiểm định Durbin – Watson theo kinh nghiệm



6.3.3 Kiểm định tự tương quan bậc bất kỳ

i) Kiểm định F

B1: Ước lượng mô hình $Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + U_t$ (*)
thu được các phần dư e_t .

B2: Ước lượng mô hình hồi quy phụ:

$$e_t = \alpha_1 + \alpha_2 X_{2t} + \dots + \alpha_k X_{kt} + \rho_1 e_{t-1} + \dots + \rho_m e_{t-m} + v_t$$

B3: Sử dụng thống kê F thông thường để kiểm định
cặp giả thuyết:

H_0 : tất cả ρ_1, \dots, ρ_m bằng 0.

H_1 : ít nhất một trong các ρ_j ($j = 1, \dots, m$) khác không.

Nếu
biến
độc lập
là **ngoại
sinh**
chặt thì
không
cần đưa
vào mô
hình hồi
quy phụ

6.3.3 Kiểm định tự tương quan bậc bất kỳ

ii) Kiểm định Breusch – Godfrey (1978) (kiểm định nhân tử Lagrange – LM Test)

B1: Ước lượng mô hình $Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + U_t$ thu được các phần dư e_t .

B2: Ước lượng mô hình hồi quy phụ:

$$e_t = \alpha_1 + \alpha_2 X_{2t} + \dots + \alpha_k X_{kt} + \rho_1 e_{t-1} + \dots + \rho_m e_{t-m} + v_t$$

Và thu được hệ số xác định R^2 .

B3: Thực hiện kiểm định cặp giả thuyết

H_0 : tất cả ρ_1, \dots, ρ_m bằng 0.

H_1 : ít nhất một trong các ρ_j ($j = 1, \dots, m$) khác không.

Nếu $(n - m) R^2 > \chi^2_{\alpha}(m)$ thì bác bỏ H_0 .

Nếu biến độc lập là **ngoại sinh chặt** thì không cần đưa vào mô hình hồi quy phụ

6.4 KHẮC PHỤC HIỆN TƯỢNG TỰ TƯƠNG QUAN

6.4.1 Phương pháp bình phương bé nhất tổng quát GLS

Trường hợp TTQ bậc 1: $U_t = \rho U_{t-1} + \varepsilon_t$

Giả sử đã biết ρ , $|\rho| < 1$ và ε_t thỏa mãn các giả thiết OLS.

Xét hồi quy hai biến: $Y_t = \beta_1 + \beta_2 X_t + U_t$ (a)

Quan sát kỳ trước (t-1) $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + U_{t-1}$ (b)

Suy ra: $Y_t - \rho Y_{t-1} = \beta_1(1 - \rho) + \beta_2(X_t - \rho X_{t-1}) + (U_t - \rho U_{t-1})$ (c)

Đặt:
$$\begin{cases} \beta_1^* = \beta_1(1 - \rho); \beta_2^* = \beta_2 \\ Y_t^* = Y_t - \rho Y_{t-1}; X_t^* = X_t - \rho X_{t-1} \end{cases}$$

Khi đó (d) trở thành $Y_t^* = \beta_1^* + \beta_2^* X_t^* + \varepsilon_t$ (d)

Đây là phương trình hồi quy tuyến tính thông thường

6.4.2 Phương pháp bình phương bé nhất tổng quát FGLS.

(Feasible generalized least squares)

Trường hợp TTQ bậc 1: $U_t = \rho U_{t-1} + \varepsilon_t$

Giả sử chưa biết ρ .

B1: Ước lượng mô hình hồi quy gốc, thu được các phần dư e_t .

B2: Ước lượng mô hình hồi quy phụ e_t theo e_{t-1}

$e_t = \rho e_{t-1} + v_t$ thu được ước lượng của ρ là $\hat{\rho}$.

B3: Sử dụng $\hat{\rho}$ để khắc phục tự tương quan như trường hợp đã biết ρ .

Các ước lượng FGLS là chệch và không vững nếu có biến độc lập không là biến ngoại sinh chặt

6.4.3 Phương pháp sử dụng phương sai hiệu chỉnh

- Tương tự trường hợp phương sai sai số thay đổi, phương pháp sai số chuẩn vững áp dụng cho trường hợp TTQ được đề xuất bởi Newey – West (1987), với ý tưởng:
- Vẫn sử dụng các hệ số ước lượng OLS.
- Ma trận hiệp phương sai các hệ số ước lượng được tính toán không dựa trên giả thiết về TTQ và phương sai sai số thay đổi.