

# Chương VIII

## Tự Tương Quan

QTKD / ĐHCN tp HCM

# I. Bản chất và nguyên nhân của tự tương quan

(1). Tự tương quan: Là sự tương quan giữa các thành phần của chuỗi các quan sát theo thời gian hay không gian.

Nếu có tự tương quan giữa các sai số ngẫu nhiên thì :

$$\text{Cov}(U_i, U_j) \neq 0 \quad (i \neq j)$$

(2). Nguyên nhân:

Quán tính dãy số liệu

Mạng nhện

Trễ

Nguyên nhân chủ quan

## II. Một số khái niệm về lược đồ tự tương quan

Xét mô hình sau đây với số liệu thời gian :

$$Y_t = \beta_1 + \beta_2 X_t + U_t$$

- Nếu  $U_t = \rho U_{t-1} + \varepsilon_t$  ( $-1 \leq \rho \leq 1$ ) (a)

Trong đó :  $\varepsilon_t$  thỏa các giả thiết của mô hình hồi qui tuyến tính cổ điển :

$$E(\varepsilon_t) = 0 \quad \forall t$$

$$\text{Var}(\varepsilon_t) = \sigma^2 \quad \forall t$$

$$\text{Cov}(\varepsilon_t, \varepsilon_{t'}) = 0 \quad (t \neq t')$$

Thì (a) được gọi là lược đồ tự tương quan bậc nhất Markov, ký hiệu AR(1) và  $\rho$  được gọi là hệ số tự tương quan bậc nhất.

- Nếu  $U_t = \rho_1 U_{t-1} + \rho_2 U_{t-2} + \dots + \rho_p U_{t-p} + \varepsilon_t$  (b)  
( $-1 \leq \rho_1, \dots, \rho_p \leq 1$ )

Trong đó :  $\varepsilon_t$  thỏa các giả thiết của mô hình hồi qui tuyến tính cổ điển .

Thì (b) được gọi là lược đồ tự tương quan bậc p Markov, ký hiệu AR(p).

### III. Ước lượng OLS khi có tự tương quan

Xét mô hình :  $Y_t = \beta_1 + \beta_2 X_t + U_t$  (1)

Với  $U_t = \rho U_{t-1} + \varepsilon_t$  ( $-1 \leq \rho \leq 1$ )

Nếu dùng OLS để ước lượng (1) thì :

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2} \quad (\text{trong đó } \sum x_i y_i = \sum X_i Y_i - n \bar{X} \bar{Y})$$

$$\sum x_i^2 = \sum X_i^2 - n(\bar{X})^2)$$

Nhưng công thức tính phương sai đã không còn như trước :



$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_t^2} + \frac{2\sigma^2}{\sum x_t^2} \left[ \rho \frac{\sum_{t=1}^{n-1} x_t x_{t+1}}{\sum x_t^2} + \right. \\ \left. + \rho^2 \frac{\sum_{t=1}^{n-2} x_t x_{t+2}}{\sum x_t^2} + \dots + \rho^{n-1} \frac{x_1 x_n}{\sum x_t^2} \right]$$

## IV. Hậu quả của việc sử dụng phương pháp OLS khi có tự tương quan

1. Các ước lượng OLS vẫn là các ước lượng tuyến tính, không chệch nhưng không còn hiệu quả nữa.
2. Ước lượng của các phương sai bị chệch (thường thấp hơn giá trị thực) nên các kiểm định t và F không còn hiệu lực nữa.
3. Thường  $R^2$  được ước lượng quá cao so với giá trị thực .
4. Sai số chuẩn của các giá trị dự báo không còn tin cậy được nữa.

# V. Cách phát hiện tự tương quan

## 1. Phương pháp đồ thị

- Hồi qui mô hình gốc  $\rightarrow$  phần dư  $e_t$ .
- Vẽ đồ thị phần dư  $e_t$  theo thời gian.
- Nếu phần dư phân bố ngẫu nhiên xung quanh trung bình của chúng, không biểu thị một kiểu mẫu nào khi thời gian tăng  $\rightarrow$  mô hình gốc không có tự tương quan.



## 2. Kiểm định Breusch-Godfrey (BG)

Xét mô hình :  $Y_t = \beta_1 + \beta_2 X_t + U_t$  (1)

với  $U_t = \rho_1 U_{t-1} + \rho_2 U_{t-2} + \dots + \rho_p U_{t-p} + \varepsilon_t$

$\varepsilon_t$  thỏa mãn các giả thiết của mô hình cổ điển Cần kiểm định  $H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0$

(không có tự tương quan)

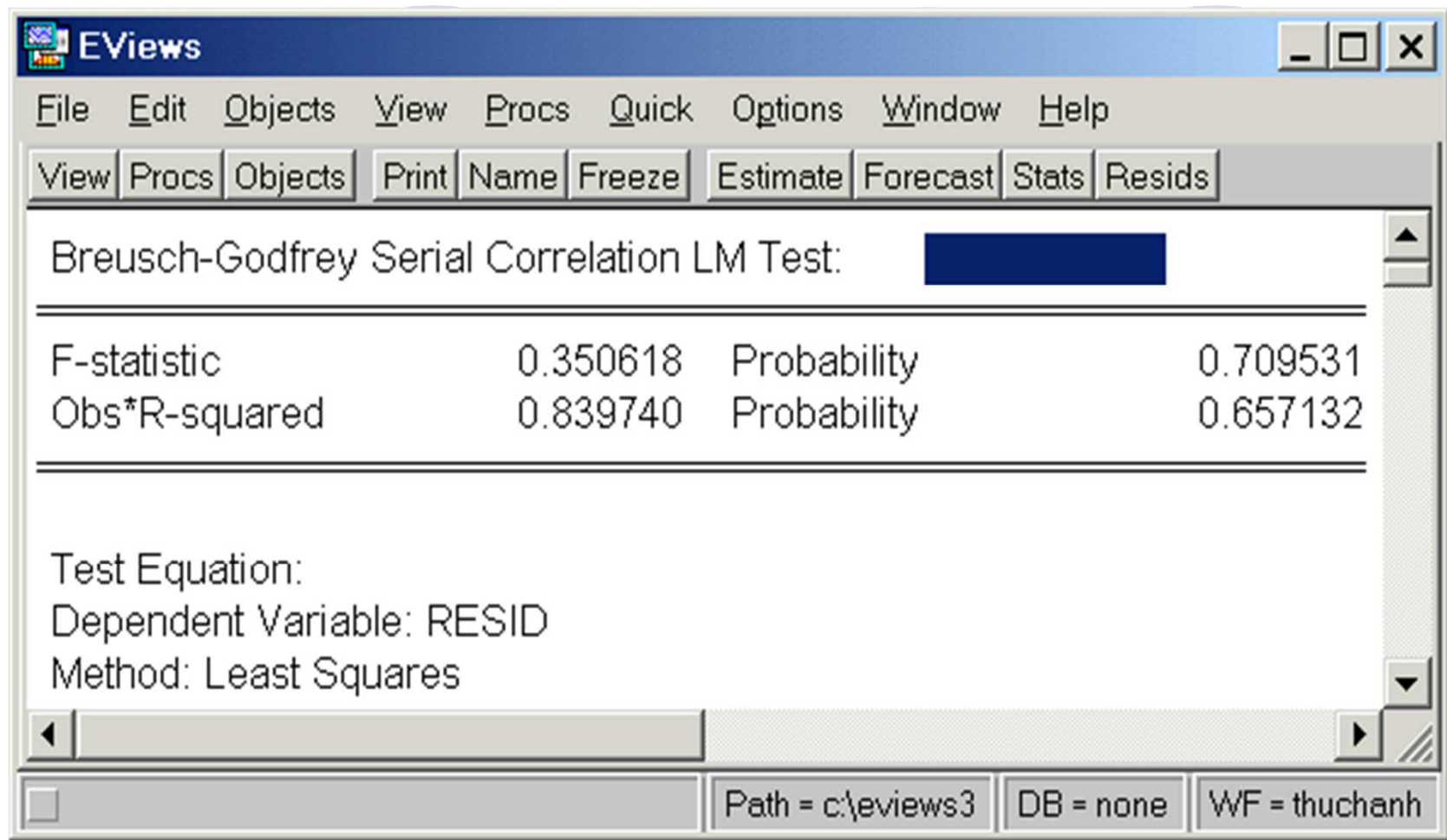
**Bước 1**: Ước lượng mô hình (1), thu  $e_t$ .

**Bước 2**: Ước lượng mô hình sau, thu  $R^2_{aux}$  :

$$e_t = \beta_1 + \beta_2 X_t + \rho_1 e_{t-1} + \rho_2 e_{t-2} + \dots + \rho_p e_{t-p} + V_t$$

**Bước 3** : Nếu  $(n-p)R^2_{aux} > \chi^2_{\alpha}(p) \rightarrow$  bác bỏ  $H_0$ ,  
nghĩa là có tự tương quan.

- Chú ý :  $(n-p)$  chính là số quan sát còn lại sau khi lấy trừ đến bậc  $p$ , nên có thể coi  $(n-p)$  là số quan sát của mẫu mới . Trong Eviews, kết quả kiểm định BG hiển thị Obs\*R-square tức là  $(n-p)R^2$ .
- Ví dụ : Hồi qui mô hình (1) rồi dùng kiểm định BG xem (1) có tự tương quan không.  
Kết quả :



Ta có :  $Obs \cdot R^2 = 0.8397$  với  $p = 0.657 > \alpha = 0.05$   
nên chấp nhận  $H_0$ , nghĩa là không có tự tương quan.

### 3. Kiểm định d của Durbin-Watson

Xét mô hình hồi qui có tự tương quan bậc nhất ( $U_t = \rho U_{t-1} + \varepsilon_t \quad (-1 \leq \rho \leq 1)$ ).

- Thống kê d. Durbin-Watson :

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \approx 2(1 - \hat{\rho})$$

$\hat{\rho}$  là ước lượng của  $\rho$  là :

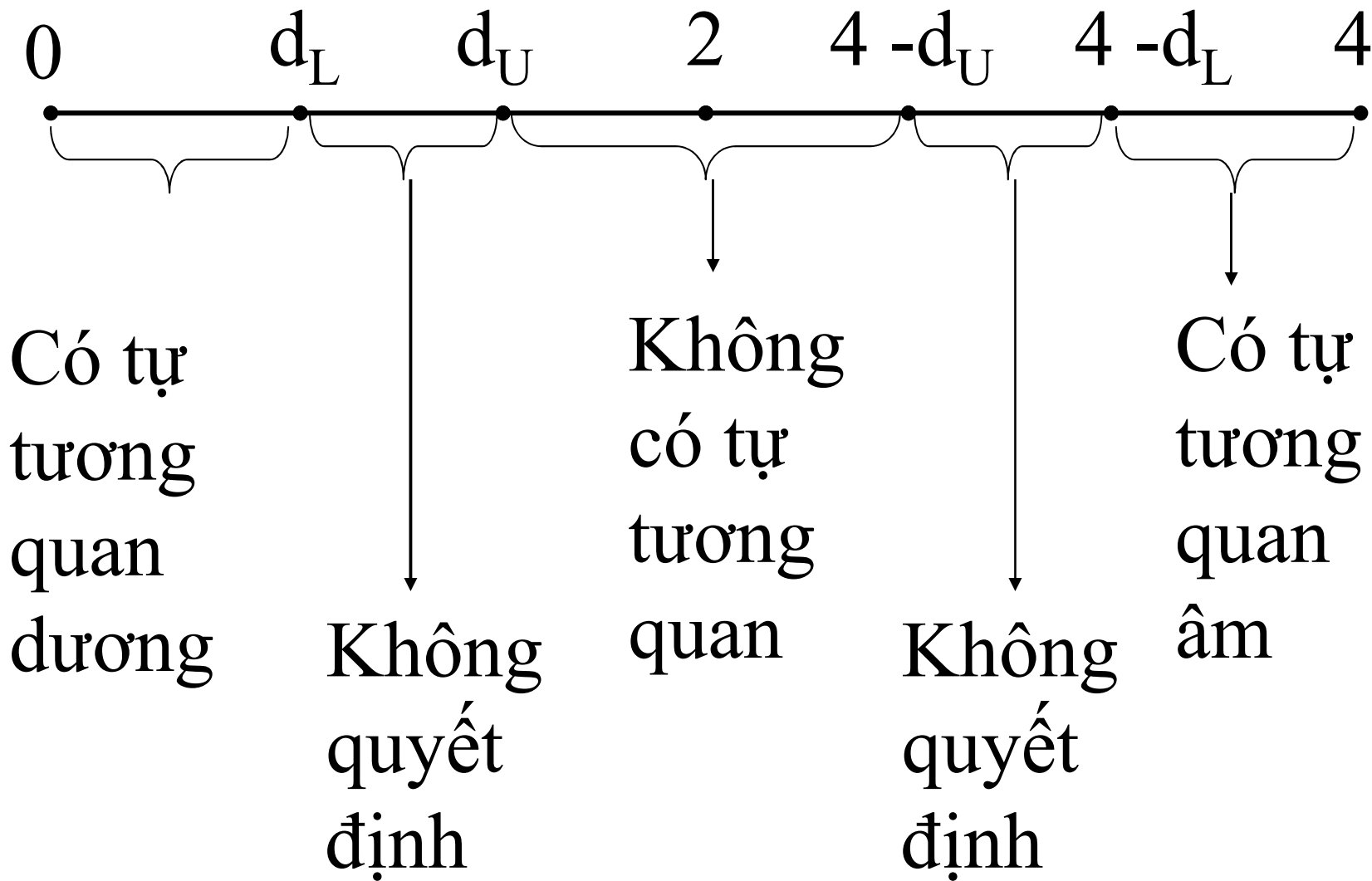
$$\hat{\rho} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2}$$

Khi  $n$  đủ lớn thì :  $d \approx 2(1 - \rho)$

Do  $-1 \leq \rho \leq 1$  nên  $0 \leq d \leq 4$

- $\rho = 0$  (không có tự tương quan)  $\rightarrow d = 2$
- $\rho = 1$  (tương quan hoàn hảo dương)  $\rightarrow d = 0$
- $\rho = -1$  (tương quan hoàn hảo âm)  $\rightarrow d = 4$

**\* Quy tắc kiểm định d của Durbin-Watson:**



Trong đó  $D_L$  và  $d_U$  là các giá trị tới hạn của thống kê Durbin-Watson dựa vào ba tham số :  $\alpha$  , số quan sát  $n$  , số biến độc lập  $k'$ .

Ví dụ : Một kết quả hồi qui được cho :

$$Y_i = 12.5 + 3.16X_i - 2.15D_i \quad (1)$$

$$n = 20 \quad d = 0.9$$

Với  $\alpha = 5\%$ ,  $n=20$ ,  $k'=2$ , ta có :

$$d_L = 1.1 \quad d_U = 1.54$$

→  $d = 0.9 \in [0, d_L]$  nên (1) có tự tương quan dương.

## Kiểm định Durbin-Watson cải biên

Với mức ý nghĩa  $2\alpha$ , ta có :

