

**CÁC SAI SỐ ĐẶC TRƯNG CỦA MÔ HÌNH
(SPECIFICATION ERRORS)**

GV : Đinh Công Khải – Chương trình Fulbright
Môn: Các Phương Pháp Định Lượng – MPP4

Các sai sót đặc trưng của mô hình

❑ Phương pháp “Hồi qui kinh tế trung bình” (Average Economic Regression)

Bắt đầu một mô hình với một số biến độc lập cho trước, sau đó dựa vào những chẩn đoán để thêm hoặc bớt biến giải thích trong mô hình.

→ Nhận dạng và lý giải được những sai số đặc trưng của mô hình

❑ Mục tiêu nghiên cứu

- ❖ Hiểu rõ bản chất các sai số đặc trưng
- ❖ Hậu quả của sai số đặc trưng
- ❖ Nhận dạng các sai số đặc trưng
- ❖ Phương thức sửa chữa các sai số đặc trưng

Các sai sót đặc trưng của mô hình

- **Các loại sai số đặc trưng**
- ❖ Bỏ sót biến thích hợp
- ❖ Đưa thêm biến không thích hợp vào mô hình
- ❖ Định dạng hàm sai
- ❖ Sai số đo lường

Bỏ sót biến thích hợp

- Giả sử, mô hình đúng là

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \quad (1)$$

- Mô hình chúng ta chọn là

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + v_i \quad (2)$$

- **Hậu quả của việc loại biến X_{3i} như sau:**

- ❖ $v_i = \beta_3 X_{3i} + u_i \rightarrow E(v_i) \neq 0$, vi phạm giả thiết về mô hình hồi qui tuyến tính cổ điển (CLRM).
- ❖ Phương sai của sai số ngẫu nhiên σ^2 được ước lượng không đúng

Bỏ sót biến thích hợp

$$\diamond E(\hat{\alpha}_2) = \beta_2 + \beta_3 \frac{\sum x_{3i}x_{2i}}{\sum x_{2i}^2} \quad (\text{xem Phụ lục 7A.5})$$

- Nếu X_{3i} có tương quan với X_{2i} thì $\hat{\alpha}_2$ là ước lượng bị lệch và không nhất quán
- Nếu X_{3i} không có tương quan với X_{2i} thì $\hat{\alpha}_2$ là không thiên lệch, nhưng $\hat{\alpha}_1$ là ước lượng bị lệch và giá trị dự báo sẽ bị thiên lệch

$$E(\hat{\alpha}_1) = \beta_1 + \beta_3 \left[\bar{X}_3 - \bar{X}_3 \frac{\sum x_{3i}x_{2i}}{\sum x_{2i}^2} \right]$$

Bỏ sót biến thích hợp

- ❖ Phương sai của $\hat{\alpha}_2$ cũng là một ước lượng chệch → những kết luận sai lầm khi xây dựng các khoảng tin cậy hoặc kiểm định giả thuyết thống kê.

$$\text{var}(\hat{\alpha}_2) = \frac{\sigma^2}{\sum x_{2i}^2}$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2 (1 - r_{23}^2)}$$

$$\text{var}(\hat{\alpha}_2) < \text{var}(\hat{\beta}_2)$$

- Mô hình hồi qui nên dựa trên nền tảng lý thuyết để không bỏ sót biến quan trọng

Đưa vào biến không liên quan

- Giả sử, mô hình đúng là

$$Y_i = \beta_1 + \beta_2 X_{2i} + u_i \quad (3)$$

- Mô hình chúng ta chọn là

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + v_i \quad (4)$$

- **Hậu quả của việc đưa thêm biến X_{3i} vào mô hình như sau:**
 - ❖ $E(v_i) = 0$, không vi phạm giả thiết về mô hình hồi qui tuyến tính cổ điển (CLRM).
 - ❖ Phương sai của sai số ngẫu nhiên σ^2 được ước lượng đúng

Đưa vào biến không liên quan

- ❖ Ước lượng OLS của mô hình (4) không bị thiên lệch

$$E(\hat{\alpha}_2) = \beta_2$$

$$E(\hat{\alpha}_1) = \beta_1$$

- ❖ *Ước lượng OLS sẽ không hiệu quả → kiểm định giả thuyết sẽ kém chính xác*

$$\text{var}(\hat{\alpha}_2) = \frac{\sigma^2}{\sum x_{2i}^2 (1 - r_{23}^2)}$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2}$$

$$\text{var}(\hat{\alpha}_2) > \text{var}(\hat{\beta}_2)$$

Kiểm định các sai số đặc trưng

□ Phát hiện sự có mặt các biến không cần thiết (Phương pháp Henry/LSE)

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_m X_{mi} + \beta_{m+1} X_{m+1i} + \dots + \beta_K X_{Ki} + u_i$$

- Kiểm tra dấu tiên nghiệm
- Dùng kiểm định t cho các mức ý nghĩa riêng của các β_k .
- Dùng kiểm định F cho mức ý nghĩa chung $\beta_{m+1} = \dots = \beta_K = 0$.

*** Chú ý: Cần dùng lý thuyết để định hướng cho việc xây dựng mô hình và cần tránh chiến lược “khai thác dữ liệu” (data mining)

Kiểm định các sai số đặc trưng

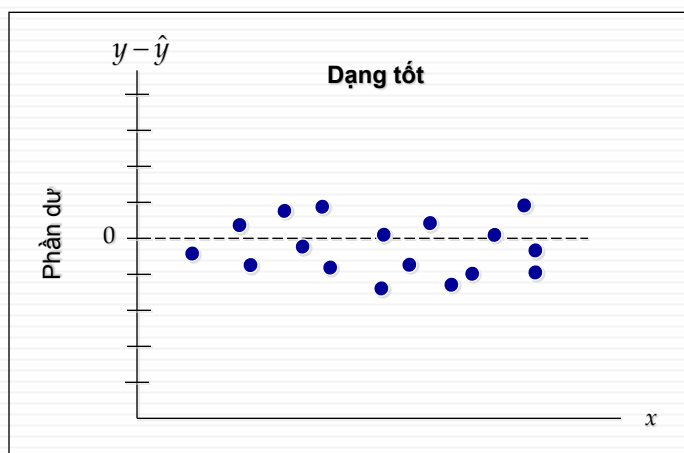
□ Kiểm định đối với việc bỏ sót các biến quan trọng và dạng hàm sai

- Dùng R^2 điều chỉnh và kiểm định t;
- So sánh dấu của các hệ số hồi qui với các dấu tiên nghiệm
- Kiểm định Durbin-Watson

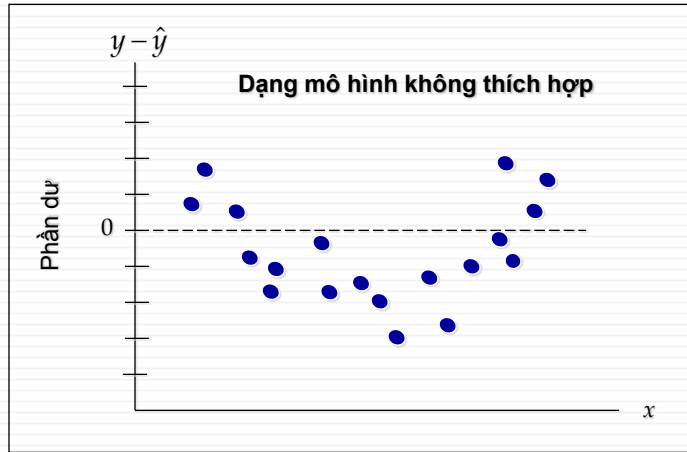
Xem xét các phần dư

- Xem xét đồ thị các phần dư để kiểm tra xem chúng ta **biến động theo một xu hướng** hay theo **một dạng đáng chú ý** không?
- Nếu có chúng ta nghi ngờ **các sai số tự tương quan** do **dạng hàm không thích hợp** hoặc do **thiếu biến quan trọng**.

ĐỒ THỊ PHẦN DƯ THEO X (Nguồn: Cao Hào Thi)



ĐỒ THỊ PHÂN DƯ THEO X (Nguồn: Cao Hào Thi)



13

Kiểm định Durbin_Watson (Nguồn: Cao Hào Thi)

$$DW = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} \approx 2(1 - \hat{\rho})$$

$$\hat{\rho} \approx \frac{\sum_{t=2}^n \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^n \hat{u}_t^2}$$

Tự tương quan dương $H_1: \rho > 0$	Không kết luận	Không tự tương quan $H_0: \rho = 0$	Không kết luận	Tự tương quan âm $H_1: \rho < 0$
0	d_L	d_U	$4 - d_U$	$4 - d_L$

Kiểm định các sai số đặc trưng

□ Kiểm định RESET của Ramsey

$$Y_i = \lambda_1 + \lambda_2 X_{2i} + u_i \quad (5)$$

- Hồi qui (5), tính ước lượng của Y, \hat{Y}_i
- Nếu \hat{u}_i quan hệ với \hat{Y}_i một cách có hệ thống thì đưa \hat{Y}_i vào vế phải (5) như là biến độc lập có thể sẽ làm tăng R^2 .
- Hồi qui (5) theo mô hình sau

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 \hat{Y}_i^2 + \beta_4 \hat{Y}_i^3 + u_i \quad (6)$$

Kiểm định các sai số đặc trưng

- Gọi R_U^2 có được từ việc hồi qui (6) và R_R^2 có được từ việc hồi qui (5)

$$F = \frac{(R_U^2 - R_R^2) / m}{(1 - R_U^2) / (n - K)}$$

(m là số biến mới đưa vào; K số biến của phương trình mới)

- Nếu F có ý nghĩa ở mức 5% thì chúng ta bác bỏ giả thuyết mô hình (5) là đúng.

Kiểm định các sai số đặc trưng

□ Kiểm định nhân tử Lagrange

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + \beta_4 X_i^3 + u_i \quad (\text{dạng đúng}) \quad (7)$$

$$Y_i = \lambda_1 + \lambda_2 X_i + u_i \quad (\text{dạng sai}) \quad (8)$$

- Hồi qui (8), tính phần dư, \hat{u}_i
- Nếu mô hình (8) đúng thì \hat{u}_i không có quan hệ với X_i^2 và X_i^3
- Nếu mô hình (8) sai thì

$$\hat{u}_i = \alpha_1 + \alpha_2 X_i + \alpha_3 X_i^2 + \alpha_4 X_i^3 + v_i \quad (9)$$

Kiểm định các sai số đặc trưng

- Với mẫu lớn, nR^2 (R^2 từ (9)) sẽ có phân phối Chi-square với tự do bằng với số biến bị giới hạn
- Nếu $nR^2 > \chi^2(df) \rightarrow$ bác bỏ giả thuyết cho rằng hàm hồi qui giới hạn là đúng