

# VỀ MỘT PHƯƠNG PHÁP KIỂM ĐỊNH HIỆN TƯỢNG PHƯƠNG SAI CỦA SAI SỐ THAY ĐỔI TRONG CÁC MÔ HÌNH HỒI QUI

TS. Nguyễn Cao Văn  
ĐH Kinh tế Quốc dân

## *1. Đặt vấn đề*

Phương pháp bình phương nhỏ nhất để ước lượng các hệ số của mô hình hồi qui luôn đòi hỏi phải thỏa mãn hàng loạt giả thiết, trong đó có giả thiết là phương sai của các sai số ngẫu nhiên là đồng đều (homoscedasticity) để đảm bảo thu được các ước lượng có các tính chất tốt nhất. Tuy nhiên trong thực tế nhiều giả thiết thường xuyên bị vi phạm làm ảnh hưởng

đến chất lượng của các ước lượng, trong đó có hiện tượng phương sai của sai số ngẫu nhiên thay đổi. Nhiều tác giả đã đề xuất các phương pháp khác nhau để phát hiện sự tồn tại của hiện tượng này trong các mô hình hồi qui nhằm đưa ra các biện pháp khắc phục song chưa có phương pháp nào thực sự hoàn thiện để phát hiện ra hiện tượng đó. Hậu quả là khi sử dụng các phương pháp khác nhau lại cho các

kết luận mâu thuẫn nhau. Phần lớn các kiểm định thông dụng như kiểm định PARK, GLEJSER, SPEARMAN, GOLDFELD - QUAND, BREUSCH - PAGAN, WHITE... đều gặp tình huống như vậy, nhất là khi áp dụng chúng cho các mô hình hồi qui bội. Vì vậy giải pháp thực tế cho vấn đề này là tiến hành càng nhiều kiểm định càng tốt và nếu phần lớn các kiểm định đó cho cùng một kết luận như nhau thì kết luận đó mới có thể xem là đáng tin cậy. Chính vì lý do đó mà chúng tôi đề xuất thêm một phương pháp nữa để kiểm định hiện tượng này trong các mô hình hồi qui.

## 2. Mô tả phương pháp

Xét mô hình hồi qui bội k biến sau:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_{ki} + U_i \quad (1)$$

và cần kiểm định cặp giả thuyết:

$H_0$ : Phương sai của các sai số ngẫu nhiên  $U_i$  là đồng đều

$H_1$ : Phương sai của các sai số ngẫu nhiên  $U_i$  thay đổi.

Đối với các sai số ngẫu nhiên  $U_i$  chỉ cần thỏa mãn hai giả thiết sau:

+ Các sai số ngẫu nhiên  $U_i$  cùng phân phối chuẩn

+ Các sai số ngẫu nhiên  $U_i$  không tương quan với nhau.

Lúc đó trên cơ sở một mẫu điều tra, thủ tục kiểm định được tiến hành theo các bước sau:

**Trường hợp 1:** Nếu có thể chia mẫu ra thành k nhóm ( $k \geq 2$ ) bằng nhau, mỗi nhóm phải có ít nhất hai quan sát ( $n_1 = n_2 = \dots = n_k = n$ ).

**Bước 1:** Dùng phương pháp bình phương nhỏ nhất OLS hồi quy mô hình (1) lần lượt với mỗi nhóm n quan sát và thu được các tổng bình phương các phần dư  $RSS_1, RSS_2, \dots, RSS_k$  tương ứng của mỗi nhóm.

**Bước 2:** Chọn tiêu chuẩn kiểm định giả thuyết là thống kê sau:

$$G = \frac{RSS_{MAX}}{RSS_1 + RSS_2 + \dots + RSS_k}$$

Cochran đã chứng minh được rằng (xem 1) phân phối xác suất của thống kê G chỉ phụ thuộc vào số bậc tự do  $(n-1)$  và số lượng nhóm k và tuân theo qui luật phân bố Cochran. Lúc đó miền bác bỏ giả thuyết ở mức ý nghĩa  $\alpha$  được xác định bằng biểu thức:

$$W_\alpha = \left\{ G = \frac{RSS_{MAX}}{RSS_1 + RSS_2 + \dots + RSS_k}; G > G_\alpha^{(n-1, k)} \right\}$$

Các giá trị tới hạn của thống kê G được tìm trong bảng phân phối COCHRAN.

**Trường hợp 2.** Nếu không thể chia các quan sát của mẫu thành các nhóm có số quan sát như nhau (chẳng hạn do đặc điểm của mỗi giai đoạn quan sát). Giả sử số quan sát của các nhóm tương ứng là  $n_1, n_2, \dots, n_k$ .

Lúc đó trước tiên cũng dùng OLS hồi quy mô hình (1) riêng cho từng nhóm quan sát và nếu ký hiệu RSS là trung bình số học của các tổng bình phương các phần dư:

$$RSS = \frac{\sum_{i=1}^k h_i RSS_i}{h}$$

trong đó  $h = \sum h_i$  với  $h_i = n_i - 1$  tức là số bậc tự do của  $RSS_i$  thì tiêu chuẩn kiểm định là thống kê sau:

$$B = \frac{V}{C}$$

trong đó  $V = 2,303 (h \cdot \lg RSS - \sum h_i \lg RSS_i)$

$$C = 1 + \frac{1}{3(k-1)} \cdot \left[ \sum_{i=1}^k \frac{1}{h_i} - \frac{1}{h} \right]$$

Bartlett [xem (2) (3)] đã chứng minh rằng nếu giả thuyết  $H_0$  là đúng thì thống kê  $B$  phân phối xấp xỉ khi bình phương với  $k - 1$  bậc tự do nếu mọi  $n_i > 3$ .

Do đó miền bác bỏ giả thuyết ở mức ý nghĩa  $\alpha$  có dạng:

$$W_\alpha = \left\{ B = \frac{C}{V}; B > \chi_{\alpha}^{2(k-1)} \right\}$$

### 3. Thí dụ áp dụng và so sánh với các phương pháp kiểm định khác.

Sử dụng số liệu thống kê của Đài Loan từ 1958 đến 1985 theo hai chỉ tiêu là tổng vốn đầu tư trực tiếp ra nước ngoài FDI và tổng sản phẩm quốc nội GDP ta hồi quy mô hình sau:

$$FDI_t = \beta_1 + \beta_2 GDP_t + U_t \quad (2)$$

Sử dụng trường hợp 1, giả sử chia các quan sát của mẫu thành 4 nhóm, mỗi nhóm có  $n=8$  quan sát và dùng OLS hồi quy mô hình (2) cho từng nhóm ta thu được kết quả sau:

$$RSS_1 = 78,393 \quad RSS_2 = 613,3936 \\ RSS_3 = 2787,4 \quad RSS_4 = 12127,6$$

Từ đó giá trị quan sát của  $G = 0,777$

Theo bảng giá trị tới hạn COCHRAN thì  $G_{0,05}^{(7,4)} = 0,6129$

Vậy  $G_{qs}$  thuộc miền bác bỏ, do đó bác bỏ  $H_0$  và thừa nhận phương sai của sai số ngẫu nhiên trong mô hình

(2) là thay đổi.

Nếu sử dụng trường hợp 2, giả sử chia các quan sát của mẫu thành 3 nhóm với số quan sát tương ứng là  $n_1 = 10$ ,  $n_2 = 10$  và  $n_3 = 8$  ta tìm được

$$RSS_1 = 293,5072 \quad RSS_2 = 2733,1 \\ RSS_3 = 12876,4$$

Từ đó  $RSS = 5309,49088$  và  $V = 27,665 \quad C = 1,0542$

Như vậy giá trị quan sát của  $B = 26,243$

Mà giá trị tới hạn  $\chi_{0,05}^{2(3)} = 7,815$

Vậy  $B_{qs}$  cũng thuộc miền bác bỏ do đó cũng thu được kết luận giống như trường hợp trên là phương sai của sai số trong mô hình (2) là thay đổi.

Cùng với tập số liệu trên, các kiểm định thông dụng khác về hiện tượng phương sai của sai số thay đổi đối với mô hình (2) cho kết quả như sau:

+ Kiểm định PARK cho hàm hồi quy mẫu sau :

$$\ln E^2 = -4,0404 + 0,92295 \ln GDP$$

và  $T_{qs} = 3,051$  (P-value = 0,005)

+ Kiểm định GLEJSER cho hàm hồi quy mẫu sau:

$$|E| = 24,484 - 26944,6/GDP$$

và  $T_{qs} = -2,5102$  (P-value = 0,019)

+ Kiểm định SPEARMAN cho

$$R_s = 0,58128 \quad \text{và} \quad T_{qs} = 3,6426$$

mà  $T_{0,05}^{(26)} = 1,706$

+ Kiểm định GOLDFELD - QUAND cho

$$RSS_1 = 239,5072 \quad RSS_2 = 13123,6$$

Từ đó  $F_{qs} = 54,794$  mà  $F_{0,05}^{(8,8)} = 3,44$

+ Kiểm định WHITE cho hàm hồi quy mẫu sau:

$$E^2 = 351,8679 - 0,039483GDP + 0,01168GDP^2$$

và  $R^2 = 0,55106$  từ đó  $\chi^2 = 15,42968$

mà  $\chi_{0,05}^{2(3)} = 7,815$

+ Kiểm định dựa trên biến phụ thuộc với mô hình

$$E^2 = \beta_1 + \beta_2 \text{FDIMU}^2 + U$$

cho  $\chi^2(1) = 14,3819$  ( P-value = 0)

và  $F(1,26) = 27,4582$  ( P-value = 0)

Như vậy so sánh kết quả kiểm định của phương pháp do chúng tôi đề nghị với kết quả của một số phương pháp thông dụng khác thì các kết luận thu được là tương tự. Tất nhiên với các mô hình khác, khi hiện tượng phương sai của sai số thay đổi không thật sự hiển nhiên thì có thể xảy

ra tình huống là các kết luận của các phương pháp khác nhau sẽ là khác nhau.

#### ***Tài liệu tham khảo:***

1. Snedecor G.W and W.G. Cochran. 1989. Statistical methods 8<sup>th</sup> ed. Ames IA. Iowa state University Preess

2. Bartlett M.S. 1937. Properties of sufficiency and statistical tests. Proceeding of the royal society. Senes A 160: 268 - 282.

3. Judge et al. 1985. The theory and practice of econometrics 2d ed. New York John Wiley and Sons.