

Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2013

Các Phương Pháp Phân Tích Định Lượng

Bài tập 10

LỰA CHỌN DẠNG HÀM & KIỂM ĐỊNH ĐẶC TRƯNG MÔ HÌNH

Ngày Phát: Thứ Ba, 17/12/2013

Ngày Nộp: Thứ Năm, 26/12/2013

Bản in nộp lúc 8:20 sáng, tại Hộp nộp bài tập trong phòng Lab

Bản điện tử gửi theo hướng dẫn của trợ giảng.

Bài 1

Dữ liệu ở file **data 4-8.wf1** của Ramanathan về các hệ thống cáp của một mẫu gồm 40 thị trường truyền hình, có các biến sau:

Sub: số người đăng ký thuê bao trên mỗi hệ thống (ngàn người)

Home: số lượng nhà mà hệ thống đi qua

Inst: chi phí lắp đặt (USD)

Svc: chi phí dịch vụ hàng tháng của mỗi hệ thống (USD)

Tv: số lượng tín hiệu truyền tải bởi mỗi hệ thống cáp

Age: tuổi của mỗi hệ thống (năm)

Air: số lượng tín hiệu truyền hình miễn phí nhận được

Y: thu nhập bình quân đầu người của mỗi thị trường cáp truyền hình (USD)

a. Bạn hãy ước lượng mô hình không giới hạn sau và giải thích kết quả

$$\text{Sub} = \beta_1 + \beta_2 \text{age} + \beta_3 \text{age}^2 + \beta_4 \text{air} + \beta_5 \text{air}^2 + \beta_6 \text{y} + \beta_7 \text{y}^2 + \beta_8 \text{home} + \beta_9 \text{svc} + \beta_{10} \text{tv} + \beta_{11} \text{inst} + u \quad (1)$$

b. Theo bạn, Mô hình (1) có bị vi phạm sai số đặc trưng: thừa biến không cần thiết hay không? Nếu có, hãy khắc phục sai số đặc trưng này (giả sử bạn quan tâm đến độ tin cậy ít nhất là 90%) bằng cách loại lần lượt các biến không có ý nghĩa thống kê ra khỏi mô hình và chỉ ra mô hình mà bạn chọn. Nếu không khắc phục, theo bạn, mô hình sẽ có những trục trặc gì?

c. Về mặt thống kê, từ mô hình (1), bằng kiểm định Wald bạn hãy đưa ra kết luận xem có nên loại đồng thời cả ba biến inst, tv, svc ra khỏi mô hình (1) hay không?

d. Về mặt thống kê, bằng kiểm định nhân tử Lagrange, bạn hãy đưa ra kết luận xem mô hình mà bạn chọn ở câu b có bị thiếu ba biến quan trọng sau đây hay không: inst, tv, svc.

e. Hãy xem mô hình mà bạn lựa chọn ở câu b có bị sai số đặc trưng không bằng thống kê Durbin-Watson.

f. Hãy sử dụng kiểm định RESET của Ramsey để xem xét mô hình mà bạn đã chọn ở câu b có bị sai số đặc trưng hay không, và cho biết nhận xét của bạn từ kết quả tính toán.

g. Với mô hình mà bạn đã chọn ở câu b, hãy cho biết tác động biên của Y theo các biến có ở trong mô hình (nếu tác động biên phụ thuộc vào một biến giải thích khác, hãy tính tác động biên này ở giá trị trung bình của các biến giải thích)?

h. Anh Tèo đã ước lượng mô hình sau:

$$\ln(\text{sub}) = \alpha_1 + \alpha_2 \text{age} + \alpha_3 \text{age}^2 + \alpha_4 \text{home} + v \quad (2)$$

Trong mô hình trên, nếu biến home có ý nghĩa thống kê ở mức ý nghĩa 10%, và ít nhất một trong 2 biến age, age^2 có ý nghĩa thống kê ở mức ý nghĩa 10% thì Tèo có thể sử dụng. Theo bạn, từ thông tin của mẫu, hệ số ước lượng của α_4 cho Tèo biết điều gì?

i. Với mô hình 2, từ thông tin của mẫu, xét tại giá trị trung bình của home, age, theo bạn tác động biên, và hệ số co giãn của sub theo age sẽ là bao nhiêu?

j. Với Stata, bạn hãy cho biết những câu lệnh cần thiết trong do-file để tính toán được các kết quả cần thiết nhằm trả lời cho câu a, b, c, h. Các kết quả ấy là gì?

Bài 2.

Bạn hãy thu thập dữ liệu về GDP theo giá cố định của Việt Nam từ năm 1995 đến 2010.

a. Tạo ra biến T là thứ tự thời gian, với T=1 cho năm 1995, T=2 cho năm 1996 ... và T=16 cho năm 2010 và vẽ biểu đồ biến động của GDP theo T. Dựa vào đồ thị, theo bạn dạng hàm GDP=f(T) nào có thể phù hợp cho chuỗi dữ liệu GDP trên?

b. Giả sử bạn đang phân vân một trong các mô hình sau:

$$\widehat{\text{GDP}} = a_0 + a_1 T$$

$$\widehat{\text{GDP}} = b_0 + b_1 T + b_2 T^2$$

$$\widehat{\text{GDP}} = c_0 + c_1 \ln(T)$$

$$\widehat{\text{GDP}} = e^{\lambda_0 + \lambda_1 T}$$

Hãy ước lượng các mô hình trên và chọn ra 1 mô hình mà bạn cho là phù hợp nhất.

c. Với mô hình mà bạn đã chọn ở câu b, nếu quy luật biến động của dữ liệu theo thời gian vẫn còn đúng đến năm 2012; bạn hãy cho biết giá trị dự báo điểm, khoảng tin cậy 95% của giá trị dự báo GDP (giá trị cá biệt) sẽ là bao nhiêu?