Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2013

Các Phương Pháp Phân Tích Định Lượng

Bài tập 8

HỒI QUY ĐA BIẾN

Ngày Phát: **Thứ ba 3/12/2013**
Ngày Nộp: **Thứ ba 7/12/2013**
Bản in nộp lúc **8h20 sáng**, tại Hộp nộp bài tập trong phòng Lab
Bản điện tử gửi theo hướng dẫn của trợ giảng.

Bài 1 (50 điểm)

Bạn hãy sử dụng file WOODY's DATA ở bài tập 6, thực hiện các thao tác cần thiết trên Eviews để trả lời các câu hỏi sau:

a. Nếu xem Y là biến đo lường lượng cầu về dịch vụ nhà hàng, bạn kỳ vọng như thế nào về mối quan hệ giữa Y và biến N, P, I và vì sao?

b. Vẽ ba đồ thị phân tán giữa Y và N, giữa Y và P, giữa Y và I và cho biết nhận xét của bạn.

c. Tính hệ số tương quan cặp giữa Y và N, giữa Y và P, giữa Y và I, đồng thời kiểm định xem từng hệ số tương quan trên có ý nghĩa thống kê ở độ tin cậy 90% hay không? Kết quả tính toán của bạn có ủng hộ những kỳ vọng ban đầu của bạn ở câu a hay không?

d. Tính thống kê mô tả (trung bình, độ lệch, min, max) của các biến Y, N, P, I

e. Bạn hãy ước lượng hàm hồi quy mẫu sau:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 N_i + \hat{\beta}_3 P_i + \hat{\beta}_4 I_i + \hat{u}_i$$

f. Từ hàm hồi quy ở câu e, hãy tính tác động biên của từng biến giải thích lên biến Y (về mặt trung bình) và cho biết những nhận xét của bạn.

g. Nếu xét tại giá trị trung bình của các biến giải thích, hãy tính hệ số co giãn của Y theo từng biến giải thích, và cho biết những nhận xét của bạn.

h. Hãy ước lượng hàm hồi quy sau và cho biết ý nghĩa kinh tế của các hệ số hồi quy đứng trước các biến giải thích trong mô hình:

$$\ln(Y_i) = \alpha_1 + \alpha_2 \ln(N_i) + \alpha_3 \ln(P_i) + \alpha_4 \ln(I_i) + \hat{u}_i$$

i. Giả sử bạn chọn mô hình ở câu h để dự báo trung bình lượt khách đến một nhà hàng mà bạn quan tâm. Theo bạn giá trị dự báo điểm sẽ là bao nhiêu nếu biết nhà hàng này ở khu vực có N=5 đối thủ, P= 60000 người, I= 25000 USD *(có thể tính gần đúng)*

j. Nếu chỉ sử dụng R^2 đã hiệu chỉnh để so sánh, theo bạn có thể chọn được mô hình ở câu e, và mô hình ở câu h mô hình nào tốt hơn hay không, và vì sao?

Bài 2 (25 điểm)

Một nghiên cứu của tác giả Khương Ninh đăng trên tạp chí Công Nghệ Ngân Hàng số tháng 8/2010 có kết quả hồi quy sau:

Mô hình 1: INV=0.9762+0.0768PRO + $\hat{u}_i$

(thống kê t) (13.0593) (6.7066)

n=810, $\overline{R_2}$=0.074, P-value(F)=0.000

Mô hình 2: INV=0.9426+0.4941PRO -0.0667PROxASS+ $\hat{u}_i$

(thống kê t) (13.0248) (7.5750) (-6.4683)

n=810, $\overline{R_2}$=0.138, P-value(F)=0.000

Mô hình 3:

INV=0.66+0.50***PRO -0.07***PROxASS+0.0033**DASL+0.04*EDU-0.0095*AGE
+0.0159VOCA-0.04INP-0.10ENL+0.05COM + $\hat{u}_i$

n=810, $\overline{R_2}$=0.314, P-value(F)=0.000

Ghi chú: dấu *, **, *** sau các hệ số hồi quy cho biết hệ số hồi quy có ý nghĩa thống kê tương ứng ở mức ý nghĩa 10%, 5%, và 1%

Trong đó, thông tin của một số biến như sau: INV là tỷ số giữa giá trị đầu tư vào máy móc thiết bị, nhà xưởng … và tổng giá trị tài sản cố định của doanh nghiệp; PRO là tỷ số giữa lợi nhuận của năm trước và giá trị tài sản cố định của doanh nghiệp, PROxASS là tích số giữa PRO và ln(giá trị tài sản cố định), DASL là tốc độ tăng trưởng của doanh số (%), EDU cho biết trình độ học vấn của người quản lý doanh nghiệp, AGE là tuổi của doanh nghiệp tính từ khi thành lập đến năm điều tra, VOCA là số lần dự các lớp tập huấn chuyên môn của người quản lý doanh nghiệp.

Hãy tạm thời bỏ qua những sai sót của tác giả (nếu có) trong việc xây dựng và ước lượng mô hình, cũng như giả sử rằng mô hình này không vi phạm các giả định của của phương pháp OLS, từ kết quả trên, bạn hãy trả lời các câu hỏi sau:

a. Tính hệ số tương quan của biến INV và PRO

b. Trong mô hình 1, theo bạn biến PRO có ý nghĩa thống kê ở độ tin cậy 95% hay không, hãy cho biết ý nghĩa kinh tế của hệ số này?

c. Theo bạn thống kê F của mô hình 1 sẽ bằng bao nhiêu?

d. Từ các mô hình trên, theo bạn tác động biên của PRO lên INV có phụ thuộc vào biến ASS hay không?

e.Theo bạn, hệ số 0.0033 đứng trước biến DASL (trong mô hình 3) cho bạn biết điều gì? đây có phải là hệ số co giãn của INV theo DASL hay không, và vì sao?

Bài 3 (25 điểm)

Từ dữ liệu VHLSS2008, sử dụng phần mềm Stata; (a) bạn hãy ước lượng hàm hồi quy tuyến tính bội thể hiện mối quan hệ giữa chi tiêu của hộ và các biến giải thích sau đây: thu nhập của hộ, tuổi của chủ hộ, số người trong hộ (có sử dụng trọng số trong hồi quy). Sau đó, bạn hãy cho biết những nhận xét của bạn từ kết quả tính toán. (b) Trong các biến giải thích trên, theo bạn biến nào ảnh hưởng mạnh nhất (quan trọng nhất) đến chi tiêu của hộ. (c) Nếu sử dụng VIF và hệ số tương quan cặp giữa các biến giải thích để chẩn đoán hiện tượng đa cộng tuyến, theo bạn mô hình có dấu hiệu bị vi phạm đa cộng tuyến hay không?