

Trường Chính sách công và Quản lý Fulbright

Học kỳ Xuân năm 2020

THỐNG KÊ VÀ PHÂN TÍCH DỮ LIỆU

BÀI TẬP 4

Ngày Phát: Thứ năm 12/03/2020

Ngày Nộp: Chủ nhật 15/03/2020

Học viên nộp bản điện tử bao gồm file Word và Eviews trước 13h20 trên ứng dụng Microsoft Teams

Học viên cần trình bày rõ ràng trên Word

Câu 1 (40 điểm)

Anh/chị hãy tải dữ liệu về chi tiêu cho giáo dục của các bang ở Mỹ trong file Eduspend-Income-DataPS4.xls. Trong bộ dữ liệu này:

- Eduspend: là chi tiêu bình quân đầu người cho giáo dục ở mỗi bang (USD)
- Income: là thu nhập bình quân đầu người của mỗi bang (USD)
- a. Anh/Chị hãy sử dụng các lý thuyết đã học và/hoặc kinh nghiệm của anh/chị để xác định các biến độc lập và biến phụ thuộc trong mô hình hồi quy của mình.
- b. Anh/chị kỳ vọng như thế nào về dấu của các hệ số ước lượng. Hãy giải thích?
- c. Sử dụng phần mềm Excel, chạy mô hình hồi quy theo các biến đã xác định bằng các phương pháp mà các anh/chị đã được hướng dẫn và thực hành. Trình bày ngắn gọn cách thực hiện.
- d. Từ kết quả của câu 3, anh/chị hãy viết phương trình hồi quy. Dấu của các hệ số ước lượng quan sát được có giống như sự kỳ vọng của các anh/chị không? Nếu có sự khác biệt, anh/chị hãy giải thích vì sao?
- e. Giải thích ý nghĩa của các hệ số ước lượng?
- f. Hãy dùng kiểm định t để kiểm định mối quan hệ giữa hai biến độc lập và phụ thuộc có ý nghĩa về mặt thống kê ở mức ý nghĩa 5% hay không?
- g. Xác định và giải thích ý nghĩa của hệ số xác định R^2 và R^2 hiệu chỉnh.
- h. Dựa vào kết quả của mô hình hồi quy, hãy đề xuất giải pháp cải tiến biến phụ thuộc mà anh/chị đã chọn.

Câu 2 (60 điểm)

Anh/Chị hãy sử dụng file dữ liệu Employee-DataPS4.xls. File này có các biến sau:

id	số thứ tự
age	tuổi
Gender	1: Nam, 2: nữ
Jobcat (vị trí công việc)	1: nhân viên, 2: trưởng nhóm/tổ trưởng/Quản đốc

	3: quản lý
educ	số năm đi học
prevexp	số tháng đã đi làm
salary	tiền lương hiện tại trong một năm (USD)
salbegin	tiền lương khởi điểm trong một năm (USD)

- a. Anh/Chị hãy ước lượng mô hình sau đây và cho biết những nhận xét của bạn từ kết quả tính toán:

$$\text{Salary} = \beta_1 + \beta_2 \text{educ} + u$$

- b. Anh/Chị hãy ước lượng mô hình sau đây và cho biết những nhận xét của bạn từ kết quả tính toán:

$$\text{Salary} = \beta_1 + \beta_2 \text{educ} + \beta_3 \text{age} + u$$

- c. Anh/Chị hãy tạo biến giả male: với quy ước 1: nam, 0 nữ; và biến female: 1 nữ, 0: nam; từ biến jobcat, hãy tạo ra các biến giả phản ánh các vị trí công việc: clerical (1: nhân viên, 0 không phải nhân viên); teamleader (1: trưởng nhóm, 0: không phải là trưởng nhóm); manager (1: quản lý; 0: không phải là quản lý); educmale là tích của 2 biến educ và male; Từ đó, ước lượng mô hình sau:

$$\text{Salary} = \beta_1 + \beta_2 \text{educ} + \beta_3 \text{male} + \beta_4 \text{educmale} + \beta_5 \text{teamleader} + \beta_6 \text{manager} + u$$

- d. Anh/Chị hãy cho biết ý nghĩa kinh tế của hệ số hồi quy đứng trước biến manager, biến teamleader.

- e. Theo Anh/Chị, hệ số hồi quy đứng trước biến educmale cho bạn biết điều gì?

- f. Anh/Chị hãy viết phương trình để ước tính tiền lương cho nam, phương trình để ước tính tiền lương cho nữ, và cho biết trong điều kiện cố định các yếu tố khác, trung bình mức chênh lệch tiền lương của nam so với nữ sẽ như thế nào?

- g. Với mô hình ở câu a, Anh/Chị hãy dự đoán (dự báo điểm) mức lương trung bình của một người bằng bao nhiêu nếu biết người này có 16 năm đi học, ở chức vụ quản lý, là nam, có tuổi bằng 34, đã từng làm việc trong nghề là 5 năm, có một vợ hai con.

- h. Anh/Chị hãy ước lượng mô hình sau đây và cho biết những nhận xét của bạn từ kết quả tính toán:

$$\text{Salary} = \beta_1 + \beta_2 \text{educ} + \beta_3 \text{female} + \beta_4 \text{clerical} + \beta_5 \text{manager} + u$$