

Trường Chính sách công và Quản lý Fulbright

Học kỳ Xuân 2019

Các phương pháp định lượng 2

BÀI TẬP 2

Ngày Phát: Thứ năm 07/03/2019

Ngày Nộp: Thứ ba 19/03/2019

Bản in nộp trước 8h20 tại Hộp nộp bài tập | Bản điện tử nộp trước 8h20 trên học liệu mở

Lưu ý:

- ✓ Học viên được yêu cầu nộp cả phần bài làm và các file dữ liệu (data, do-file, log-file, ...) từ phần mềm để Ban giảng viên kiểm tra, đối chiếu khi chấm điểm.

Câu 1:

Học viên được cung cấp file dữ liệu *lbw.dta* gồm 189 quan sát về trọng lượng của trẻ khi vừa sinh ra.¹ Học viên được yêu cầu sử dụng các mô hình, công cụ cần thiết để ước lượng, kiểm định mô hình, dự báo và phân tích các yếu tố ảnh hưởng đến xác suất trẻ sinh ra bị nhẹ cân (dưới 2500gr).

Thông tin các biến có trong file dữ liệu như bảng sau:

| Biến | Giải thích |
|-------|--|
| low | = 1 nếu trọng lượng khi vừa sinh ra < 2500gr = 0 nếu trọng lượng khi vừa sinh ra >= 2500gr |
| bwt | Trọng lượng khi vừa sinh ra (gr) |
| age | Tuổi của bà mẹ tại thời điểm sinh (tuổi) |
| lwt | Trọng lượng của bà mẹ tại thời điểm khi vừa bắt đầu mang thai (lbs) |
| race | = 1 nếu bà mẹ là người da trắng = 2 nếu bà mẹ là người da đen = 3 nếu bà mẹ là các màu da khác |
| smoke | = 1 nếu bà mẹ có hút thuốc (đặc biệt trong giai đoạn đang mang thai) = 0 nếu không |
| ptl | Số lần sinh non trong quá khứ của bà mẹ |
| ht | = 1 nếu bà mẹ có tiền sử các bệnh về huyết áp = 0 nếu không |
| ui | = 1 nếu bà mẹ có hiện tượng tử cung dễ bị kích thích trong giai đoạn mang thai = 0 nếu không |
| ftv | Số lần phải đi khám bác sĩ trong 3 tháng đầu của thời kỳ mang thai |

Học viên thực hành các yêu cầu cụ thể sau đây:

- Thực hiện các thống kê mô tả cần thiết để trình bày về tình trạng sinh nhẹ cân của trẻ và trọng lượng của trẻ khi vừa sinh ra. Mô tả về mối quan hệ giữa các biến quan sát trong tập dữ liệu với tình trạng sinh nhẹ cân và trọng lượng của trẻ khi vừa sinh ra.

¹ Dữ liệu thực tế được Hosmer và Lemeshow thu thập vào năm 2000. Tham khảo từ Dataset của STAT12.

b. Giả sử cơ sở lý thuyết về y tế nói chung và sức khỏe sinh sản nói riêng ủng hộ mối quan hệ giữa các biến quan sát có trong tập dữ liệu. Hãy đề xuất, ước lượng, thực hiện các kiểm định cần thiết để chọn ra mô hình phù hợp nhất và đánh giá khả năng dự báo của các mô hình lựa chọn, bao gồm:

- 01 mô hình các yếu tố ảnh hưởng đến trọng lượng khi vừa sinh ra của trẻ (OLS)
- 01 mô hình các yếu tố ảnh hưởng đến khả năng trẻ sinh ra bị nhẹ cân (được lựa chọn từ một trong các mô hình lpm/logit/probit)

Lưu ý:

- Việc biến đổi biến (hoặc tạo biến mới từ các biến hiện hữu) có thể được xem xét trước khi thực hiện hồi quy.
 - Để chọn lựa mô hình, anh/chị cần giải thích ngắn gọn phương thức các anh/chị đã làm để chọn ra mô hình mà anh/chị cho là phù hợp nhất.
 - Với mỗi mô hình anh/chị lựa chọn, anh/chị cần kiểm định, kết luận và giải thích ý nghĩa của tất cả các hệ số hồi quy có trong mô hình.
- c. Anh/chị hãy viết một kết luận ngắn gọn (tối đa 01 trang A4) để nói về tác động của yếu tố lên trọng lượng cũng như tình trạng nhẹ cân của trẻ khi sinh. Đồng thời đưa ra các khuyến nghị phù hợp cho các bà mẹ chuẩn bị hoặc đang mang thai.

Câu 2:

Học viên được cung cấp file dữ liệu *ebt.xls* ghi nhận lợi nhuận trước thuế (*ebt*) và chi phí thực hiện các chương trình khuyến mãi (*promotion*) trong giai đoạn từ 1959 - 2018 của một doanh nghiệp.² Anh/chị hãy sử dụng thông tin được cung cấp và ước lượng mô hình sau đây:

$$ebt_t = \beta_0 + \beta_1 * promotion_t + u_t$$

- a. Bằng các phép thử và các kiểm định cần thiết, theo anh/chị mô hình ở trên có dấu hiệu của hiện tượng hồi quy giả hay không? Giải thích ngắn gọn?
- b. Bằng các phép thử và các kiểm định cần thiết, theo anh/chị 2 chuỗi dữ liệu cung cấp ở mô hình trên có phải là 2 chuỗi đồng kết hợp hay không? Giải thích ngắn gọn?
- c. Giả sử *ebt_t* và *promotion_t* ở trên là 2 chuỗi đồng kết hợp, hãy dùng cơ chế hiệu chỉnh sai số ECM để nghiên cứu tác động ngắn hạn và dài hạn của chi phí thực hiện các chương trình khuyến mãi lên lợi nhuận trước thuế của doanh nghiệp. Anh/chị hãy giải thích ý nghĩa của các hệ số hồi quy ước lượng từ mô hình hiệu chỉnh sai số ECM.

² Dữ liệu mô phỏng để thực hành

Câu 3:

Học viên được cung cấp file dữ liệu *unem_gdp.xls* ghi nhận tỷ lệ thất nghiệp (*unem*) và tốc độ tăng trưởng GDP (*gdp*) của Hoa Kỳ trong giai đoạn từ 1980 – 2018.³ Anh/chị hãy sử dụng file dữ liệu cung cấp để xem mối quan hệ giữa tăng trưởng và tỷ lệ thất nghiệp qua mô hình đơn giản như sau:

$$unem_t = \alpha + \beta_0 * gdp_t + \beta_1 * gdp_{t-1} + \dots + \beta_k * gdp_{t-k} + u_t$$

- Ước lượng mô hình hồi quy trong đó biến phụ thuộc là *unem_t* và biến độc lập là *gdp_t* với các độ trễ thích hợp. Dựa vào kết quả ước lượng và các kiểm định, so sánh cần thiết, hãy đề xuất mô hình phù hợp nhất theo anh/chị. Hãy kiểm tra và xử lý hiện tượng tương quan chuỗi (tự tương quan) trong mô hình lựa chọn của anh/chị (nếu có). Giải thích tác động ngắn hạn và dài hạn từ kết quả mô hình lựa chọn.
- Giả định, anh/chị muốn xem xét tác động của tăng trưởng GDP kỳ vọng (*gdp_t^{*}*) lên tỷ lệ thất nghiệp (*unem_t*) thông qua mô hình điều chỉnh kỳ vọng (Adaptive Expectation Model) như sau đây:

$$unem_t = \beta_1 + \beta_2 * gdp_t^* + u_t$$

và

$$gdp_t^* = \gamma * gdp_{t-1} + (1 - \gamma) * gdp_{t-1}^* + e_t$$

Hãy thực hiện các phép biến đổi cần thiết để có thể ước lượng mô hình điều chỉnh kỳ vọng dựa vào dữ liệu đã cung cấp. Hãy kiểm tra và xử lý hiện tượng tương quan chuỗi (nếu có) trước khi tính toán các tham số chưa biết trong mô hình kỳ vọng ở trên. Giải thích ý nghĩa của các tham số vừa tính toán và bình luận về tác động của số nhân dài hạn trong tình huống này?

- Một giả định khác, để xem xét ảnh hưởng phân phối trễ của tốc độ tăng trưởng GDP lên tỷ lệ thất nghiệp, anh/chị hãy sử dụng kỹ thuật phân phối trễ Almon. Hãy xác định mô hình phù hợp nhất để giải thích tác động phân phối trễ của tốc độ tăng trưởng GDP lên tỷ lệ thất nghiệp tại Hoa Kỳ.
- Hãy thực hiện kiểm định nhân quả Granger để xác định xem mối quan hệ giữa tỷ lệ thất nghiệp và tốc độ tăng trưởng GDP tại Hoa Kỳ là có mối quan hệ một chiều, hai chiều hay độc lập? Hãy đưa ra câu trả lời dựa trên mô hình với độ trễ phù hợp nhất từ dữ liệu cung cấp.

---HẾT---

³ Dữ liệu được thu thập từ World Economic Outlook Database của IMF, cập nhật đến tháng 10/2018.