

KINH TẾ LƯỢNG ỨNG DỤNG

BÀI TẬP 3

DỮ LIỆU BẢNG, PHƯƠNG PHÁP KHÁC BIỆT TRONG KHÁC BIỆT VÀ DỰ BÁO

Ngày Phát: Thứ Hai, 20/05/2013

Ngày Nộp: Thứ Hai, 03/06/2013, lúc 13:20.

Bản in nộp tại Phòng Lab

Bản điện tử gửi vô địa chỉ <http://intranet.fetp.edu.vn:81/>

Câu 1 (30 điểm)

File Data 3.1 trình bày dữ liệu về Tỷ lệ thất nghiệp theo tháng có điều chỉnh theo mùa (UMEMP) (%) và Chỉ số giá sản xuất (Production Price Index – PPI) (X) của các nước Đức, Ý và Anh trong giai đoạn 1992-2012. Hãy xem mô hình:

$$\ln(UNEMP)_{it} = \beta_1 + \beta_2 \ln(PPI)_{it} + u_{it} \quad (1)$$

- Ước lượng mô hình đã cho trong (1) cho từng nước.
- Ước lượng mô hình, kết hợp toàn bộ các quan sát.
- Ước lượng mô hình các ảnh hưởng cố định (FEM).
- Ước lượng mô hình các thành phần sai số (ECM).
- Mô hình nào tốt hơn, FEM hay ECM? Giải thích câu trả lời của bạn.

Câu 2 (30 điểm)

File Data 3.2 trình bày số liệu 2.173 cá nhân làm công ăn lương, trong đó 1.118 người học ở miền Nam và 1.055 người học ở miền Bắc. Những người này đã tốt nghiệp trung học phổ thông nhưng không học tiếp, và được sinh ra trước năm 1975 hoặc sau năm 1976. Như vậy, 1.118 người làm công ăn lương ở miền Nam có thời gian học tập đúng 12 năm. Trong số 1.055 người hưởng lương ở miền Bắc, có 746 người được sinh ra trước năm 1975 và có thời gian học tập 11 năm, cùng với 309 người sinh ra sau năm 1976 và có 12 năm đi học.

Hãy chạy hai mô hình hồi quy theo phương pháp DID (một mô hình không có các biến kiểm soát và một mô hình có các biến kiểm soát) để ước lượng suất sinh lợi từ việc đi học.

Câu 3 (40 điểm)

Anh/Chị tiếp tục sử dụng dữ liệu chuỗi thời gian đã có ở Bài tập 2 để dự báo 3 giá trị kế tiếp trong tương lai.

Dựa vào kết quả của Câu 1 phần 7 trong Bài tập 2, Anh/Chị hãy chọn ra 2 mô hình ARIMA mà Anh/Chị nghĩ là tốt nhất. Gọi hai mô hình đó là **Mô Hình 1** và **Mô Hình 2**.

- a. Hãy dành riêng hai mươi phần trăm quan sát **CUỐI CÙNG** cho các tính toán và dự báo ngoài mẫu. Hãy sử dụng tám mươi phần trăm **ĐẦU TIÊN** của dữ liệu để làm các tính toán trong mẫu.

Sử dụng các quan sát trong mẫu, hãy ước lượng hai **Mô hình 1** và **Mô hình 2**. Từ phương trình của hai mô hình này, hãy tính các giá trị dự báo trong mẫu và ngoài mẫu.

- b. Dựa vào giá trị quan sát thực và giá trị dự báo, hãy xác định MAD (Mean Absolute Deviation – Độ lệch tuyệt đối trung bình) và RMSE (Root Mean Square Deviation – Căn bậc hai của độ lệch bình phương trung bình) trong mẫu và ngoài mẫu cho từng mô hình trong hai mô hình trên. Hãy tóm tắt các kết quả của anh chị vào bảng sau:

| | MAD | MAD | | RMSE | RMSE |
|------------------|---------------------|---------------------|--|---------------------|---------------------|
| | Trong mẫu, N = ? | Ngoài mẫu, N = ? | | Trong mẫu, N = ? | Ngoài mẫu, N = ? |
| Mô hình 1 | | | | | |
| Mô hình 2 | | | | | |

Dựa vào kết quả ở Bảng trên, Anh/Chi hãy lựa chọn mô hình tốt nhất.

- d. Đối với mô hình tốt nhất, hãy sử dụng **TOÀN BỘ** dữ liệu để ước tính lại Phương trình của mô hình tốt nhất này.
- e. Dựa vào phương trình ở câu d, hãy dự báo điểm và dự báo khoảng cho 3 giá trị kế tiếp trong tương lai.