

HỒI QUI ĐA BIẾN: KIỂM ĐỊNH GIẢ THUYẾT VÀ LỰA CHỌN MÔ HÌNH

GV : Đinh Công Khải – Chương trình Fulbright
Môn: Các Phương Pháp Định Lượng – MPP4

Giả thiết về qui luật chuẩn

- Giả thiết $u_i \sim N(0, \sigma^2)$
- Các tính chất của ước lượng OLS trong hồi qui đa biến theo giả thiết phân phối chuẩn $\hat{\beta}_k \sim N(\beta_k, \sigma_{\hat{\beta}_k}^2)$
- Ước lượng $\sigma_{\hat{\beta}_k}^2$ trong hàm hồi qui với 2 biến độc lập

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

$$\text{var}(\hat{\beta}_2) = \frac{\sum x_{3i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i}x_{3i})^2} \sigma^2$$

$$\text{var}(\hat{\beta}_3) = \frac{\sum x_{2i}^2}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i}x_{3i})^2} \sigma^2$$

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-3}$$

Kiểm định hệ số hồi qui riêng

□ Phương pháp kiểm định ý nghĩa: Kiểm định t

• Kiểm định 2 phía

$$H_0: \beta_k = a$$

$$H_a: \beta_k \neq a$$

Trị kiểm định thống kê

$$t = \frac{\hat{\beta}_k - \beta_k}{S_{\hat{\beta}_k}}$$

Kiểm định hệ số hồi qui riêng

Qui tắc bác bỏ

- ❖ Bác bỏ nếu $|t| > t_{\alpha/2}$ với $t_{\alpha/2}$ dựa trên phân phối t với **bậc tự do là (n-K)**
- ❖ Hoặc $p_{value} < \alpha$.

• Kiểm định 1 phía

$$H_0: \beta_k \geq a$$

$$H_0: \beta_k \leq a$$

$$H_a: \beta_k < a$$

$$H_a: \beta_k > a$$

Qui tắc bác bỏ

- ❖ Bác bỏ nếu $t < -t_\alpha$ $t > t_\alpha$
- ❖ Hoặc $p_{value} < \alpha$ $p_{value} < \alpha$

Kiểm định hệ số hồi qui riêng

- Phương pháp kiểm định dựa trên khoảng tin cậy $(1-\alpha)100\%$

$$\hat{\beta}_k \pm t_{\alpha/2} s_{\hat{\beta}_k}$$

Qui tắc bác bỏ

Bác bỏ H_0 nếu 0 không nằm trong khoảng tin cậy $(1-\alpha)100\%$ của β_k

Kiểm định ý nghĩa thống kê của các hệ số hồi qui

- Phương pháp kiểm định ý nghĩa: Kiểm định F (Kiểm định Wald)

Giả thuyết

$$H_0: \beta_2 = \beta_3 = \dots = \beta_K = 0$$

H_a : Ít nhất có một tham số β_k khác 0

Trị kiểm định F:

$$F = \frac{MSE}{MSR} = \frac{ESS / (K - 1)}{RSS / (n - K)} \sim F_{(K-1, n-K, \alpha)}$$

Qui tắc bác bỏ: Bác bỏ H_0 nếu $F \geq F_{(K-1, n-K, \alpha)}$ hoặc $p_{\text{value}} \leq \alpha$

Kiểm định ý nghĩa thống kê của các hệ số hồi qui

- **Mối quan hệ giữa R^2 và F**

$$F = \frac{R^2 / (K - 1)}{(1 - R^2) / (n - K)}$$

- Khi R^2 càng lớn thì F càng lớn.
- Kiểm định F là thước đo *ý nghĩa chung của mô hình hồi qui* và cũng là *kiểm định ý nghĩa của R^2* .
- Kiểm định $H_0: \beta_2 = \beta_3 = \dots = \beta_K = 0$ tương đương kiểm định $H_0: R^2 = 0$

Lựa chọn mô hình

- **Phương pháp “từ tổng quát đến đơn giản” (Hendry/LSE)**

Sử dụng các kiểm định để loại bỏ biến

- Kiểm tra xem dấu của các hệ số hồi qui ước lượng có đúng kỳ vọng không
- Sử dụng kiểm định t và kiểm định Wald
- Sử dụng R^2 điều chỉnh

Lựa chọn mô hình

- Phương pháp “từ đơn giản đến tổng quát”
- Liệu đưa thêm 1 hay nhiều biến giải thích có làm tăng mức ý nghĩa chung của mô hình hay không?
- Giả sử chúng ta có một mô hình với m biến (mô hình cũ)

$$(R): Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_m X_{mi} + u_i$$

Sau đó chúng ta bổ sung thêm $(K - m)$ biến giải thích (mô hình mới)

$$(U): Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_m X_{mi} + \beta_{m+1} X_{m+1i} + \dots + \beta_K X_{Ki} + v_i$$

Lựa chọn mô hình

- Dùng kiểm định Wald

$$H_0: \beta_{m+1} = \beta_{m+2} = \dots = \beta_K = 0$$

H_a : Ít nhất có một tham số β_k ở trên khác 0

Trị kiểm định

$$F = \frac{[ESS_{new} - ESS_{old}]/(K - m)}{RSS_{new}/(n - K)} = \frac{(R_{new}^2 - R_{old}^2)/(K - m)}{(1 - R_{new}^2)/(n - K)}$$

Quy luật bác bỏ H_0 : $F > F(K-m, n-K, \alpha)$ hoặc $p_{value} < \alpha \rightarrow$ bổ sung các biến vào mô hình làm tăng một cách ý nghĩa ESS và R^2 .

Lựa chọn mô hình

□ Kiểm định nhân tử Lagrange

$$(R): Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_m X_{mi} + u_i$$

$$(U): Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_m X_{mi} + \beta_{m+1} X_{m+1i} + \dots + \beta_K X_{Ki} + v_i$$

Kiểm định giả thuyết

$$H_0: \beta_{m+1} = \beta_{m+2} = \dots = \beta_K = 0$$

H_a : Ít nhất có một tham số β_k ở trên khác 0

Lựa chọn mô hình

• Bước 1: Ước lượng mô hình (R)

• Bước 2: Tính phần dư, \hat{u}_R

• Bước 3: Ước lượng mô hình

$$\hat{u}_{Ri} = \alpha_1 + \alpha_2 X_{2i} + \dots + \alpha_m X_{mi} + \alpha_{m+1} X_{m+1i} + \dots + \alpha_K X_{Ki} + \varepsilon_i \quad (*)$$

• Bước 4: Với mẫu lớn, nR^2 (R^2 từ *) sẽ có phân phối Chi-square với tự do bậc bằng với số biến bị giới hạn (K-m).

• Nếu $nR^2 > \chi^2$ (df=K-m) \rightarrow bác bỏ giả thuyết H_0 .

Lựa chọn dạng hàm hồi qui (phép thử MWD)

□ Các giả thuyết

$$H_0: Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + u_i \quad (1)$$

$$H_a: \ln Y_i = \beta_1 + \beta_2 \ln X_{2i} + \dots + \beta_K \ln X_{Ki} + v_i \quad (2)$$

□ Quy trình kiểm định

- ❖ Ước lượng mô hình tuyến tính (1); tính \hat{Y} ; tính $\ln \hat{Y}$
- ❖ Ước lượng mô hình tuyến tính logarit (2) và tính $\ln \hat{Y}$
- ❖ Tạo biến mới $Z_1 = (\ln \hat{Y} - \ln Y)$
- ❖ *Hồi qui Y theo Xs và Z_1 , bác bỏ H_0 nếu hệ số hồi qui của Z_1 có ý nghĩa thống kê*

Lựa chọn dạng hàm hồi qui (phép thử MWD)

- ❖ Tạo biến mới $Z_2 = (\text{anti log of } \ln Y - \hat{Y})$

- ❖ *Hồi qui $\ln Y$ theo $\ln X_s$ và Z_2 , bác bỏ H_a nếu hệ số hồi qui của Z_2 có ý nghĩa thống kê theo kiểm định t thông thường.*

Các tiêu chuẩn chọn mô hình khác

□ Kiểm định AIC (Akaike Info Criterion)

$$\left(\frac{RSS}{n}\right)e^{2k/n}$$

- Mô hình nào có giá trị của tiêu chuẩn này thấp hơn sẽ được chọn
- Thích hợp trong phân tích chuỗi thời gian

□ Kiểm định Schwarz

$$\left(\frac{RSS}{n}\right)n^{k/n}$$

- Mô hình nào có giá trị của tiêu chuẩn này thấp hơn sẽ được chọn
- Thích hợp đối với những mô hình đơn giản

Các tiêu chuẩn chọn mô hình khác

□ Kiểm định Hannan – Quinn (HQ Criterion)

$$\left(\frac{RSS}{n}\right)(\ln n)^{2k/n}$$

- Mô hình nào có giá trị của tiêu chuẩn này thấp hơn sẽ được chọn