

Chủ đề 3: MÔ HÌNH HỒI QUY BỘỊ

Lê Kim Long

Phạm Thành Thái

Khoa Kinh tế - Đại học Nha Trang

MÔ HÌNH HỒI QUI BA BIẾN

- Mô hình ba biến được viết như sau:

$$E(Y / X_2, X_3) = \beta_1 + \beta_2 X_2 + \beta_3 X_3$$

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

- Ý nghĩa hệ số hồi qui riêng?

Mô hình hồi qui k biến

- Mô hình k biến được viết như sau:

$$E(Y/ X_2, X_3, \dots, X_k) = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$$

$$\text{Hay: } Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + U_i$$

- Ý nghĩa các hệ số hồi qui riêng?

Ứng dụng trên Eviews

Equation: UNTITLED Workfile: DATA-01::Untitled\				
View Proc Object Print Name Freeze Estimate Forecast Stats Resids				
Dependent Variable: Y				
Method: Least Squares				
Date: 03/02/10 Time: 16:01				
Sample: 1960 1982				
Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X2	0.004889	0.004962	0.985370	0.3383
X3	-0.651888	0.174400	-3.737889	0.0016
X4	0.243242	0.089544	2.716443	0.0147
X5	0.104318	0.070644	1.476674	0.1580
X6	-0.071110	0.098381	-0.722805	0.4796
C	38.59691	4.214488	9.158150	0.0000
R-squared	0.944292	Mean dependent var	39.66957	
Adjusted R-squared	0.927908	S.D. dependent var	7.372950	
S.E. of regression	1.979635	Akaike info criterion	4.423160	
Sum squared resid	66.62224	Schwarz criterion	4.719376	
Log likelihood	-44.86635	F-statistic	57.63303	
Durbin-Watson stat	1.100559	Prob(F-statistic)	0.000000	

Hệ số xác định bội (R^2) và hệ số xác định bội đã hiệu chỉnh (\bar{R}^2):

Hệ số xác định bội R^2 có thể được tính bằng một trong hai công thức sau:

$$+ R^2 = \frac{ESS}{TSS} = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$+ R^2 = \frac{ESS}{TSS} = \frac{\beta_2 \sum_{i=1}^n y_i x_{2i} + \beta_3 \sum_{i=1}^n y_i x_{3i} + \dots + \beta_k \sum_{i=1}^n y_i x_{ki}}{\sum_{i=1}^n y_i^2}$$

Hệ số xác định bội (R^2) và hệ số xác định bội đã hiệu chỉnh (\bar{R}^2):

✓ Ý nghĩa của hệ số xác định?

✓ $0 \leq R^2 \leq 1$. Nếu $R^2 = 1$, có nghĩa là đường hồi qui giải thích 100% sự thay đổi của Y . Nếu $R^2 = 0$, có nghĩa là mô hình không giải thích sự thay đổi nào của Y .

Hệ số xác định bội (R^2) và hệ số xác định bội đã hiệu chỉnh (\bar{R}^2):

👉 Hệ số xác định bội hiệu chỉnh \bar{R}^2 được tính theo công thức sau:

$$+ \quad \bar{R}^2 = 1 - \frac{\frac{RSS}{n-k}}{\frac{TSS}{n-1}} = 1 - (1 - R^2) \frac{n-1}{n-k}$$

👉 \bar{R}^2 có các tính chất sau:

✓ Nếu $k > 1$ thì $\bar{R}^2 \leq R^2 \leq 1$, điều này có nghĩa là nếu số biến giải thích tăng lên thì \bar{R}^2 tăng chậm hơn R^2 .

Hệ số xác định bội (R^2) và hệ số xác định bội đã hiệu chỉnh (\bar{R}^2):

👉 \bar{R}^2 có các tính chất sau:

✓ $R^2 \square 0$, nhưng \bar{R}^2 có thể âm. Khi \bar{R}^2 âm thì chúng ta coi như nó bằng 0.

Người ta thường dùng hệ số \bar{R}^2 để cân nhắc khi xem xét việc đưa thêm biến mới vào mô hình. Chúng ta chỉ đưa thêm biến mới vào mô hình nếu thoả mãn hai điều kiện sau đây:

+ Chừng nào \bar{R}^2 còn tăng.

+ Hệ số hồi qui của biến mới đưa vào mô hình phải khác không có ý nghĩa về mặt thống kê.

Hệ số xác định bội (R^2) và hệ số xác định bội đã hiệu chỉnh (\bar{R}^2):

Khi nào thì biết được hệ số hồi qui của biến mới đưa vào khác không có ý nghĩa thống kê? Khi mà giả thiết:

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0$$

bị bác bỏ, trong đó β_k là hệ số của biến X_k mà chúng ta định đưa vào mô hình.

Ma trận tương quan

Giả sử chúng ta có mô hình hồi qui bội sau đây:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + U_i$$

Gọi r_{tj} là hệ số tương quan giữa biến thứ t và biến thứ j . Nếu $t=1$ thì r_{1j} là hệ số tương quan giữa biến Y và biến X_j .

Khi đó:

$$r_{1j}^2 = \frac{\left(\sum_{i=1}^n y_i x_{ji} \right)^2}{\sum_{i=1}^n y_i^2 \sum_{i=1}^n x_{ji}^2} ; \quad r_{tj}^2 = \frac{\left(\sum_{i=1}^n x_{ti} x_{ji} \right)^2}{\sum_{i=1}^n x_{ti}^2 \sum_{i=1}^n x_{ji}^2}$$

Ma trận tương quan

Trong đó: $x_{ji} = X_{ji} - \bar{X}_j$

Dễ dàng thấy rằng: $r_{tj} = r_{jt}$; $r_{jj} = 1$

Ma trận tương quan có dạng sau:

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{kk} \end{pmatrix} = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1k} \\ r_{21} & 1 & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & 1 \end{pmatrix}$$

Kiểm định giả thiết

- Kiểm định giả thiết về các hệ số hồi qui riêng:
 - + Xây dựng khoảng tin cậy.
 - + Kiểm định mức ý nghĩa.
 - + Giá trị xác suất P ($P_{\text{-value}}$).

Kiểm định giả thiết đồng thời

- Ý nghĩa?
- Giả thiết:

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

H_1 : Không phải tất cả các tham số đồng thời bằng không

- Trị thống kê kiểm định đối với giả thiết này là :

$$F_c = \frac{ESS / (k - 1)}{RSS / (n - k)} \sim F(k-1, n-k)$$

Kiểm định giả thiết đồng thời

-Mặc khác:

$$F_c = \frac{ESS / (k - 1)}{RSS / (n - k)} = \frac{R^2 (n - k)}{(1 - R^2)(k - 1)} \square F(k - 1, n - k)$$

Nguyên tắc ra quyết định: Bác bỏ giả thiết không khi

$$F_c > F_\alpha(k - 1, n - k)$$

Hoặc giá trị p-value của thống kê F nhỏ hơn mức ý nghĩa cho trước (thường nhỏ hơn 10%)