GIỚI THIỆU Ý NGHĨA CỦA CÁC THAM SỐ TRONG CÁC MÔ HÌNH HỒI QUI

Lê Dân(*)

Các mô hình hồi quy hiện đang được ứng dụng nhiều trong thực tế. Một trong những vấn đề được nhiều độc giả quan tâm là ý nghĩa của các tham số. Bài viết giới thiệu ý nghĩa các tham số của một số mô hình hồi quy thường gặp

1. Ý nghĩa của hệ số hồi qui tuyến tính dạng tổng quát

Theo dạng ngẫu nhiên

$$Y_i = \beta_1+\beta_2X_{2i}+\ldots+\beta_kX_{ki} + u_i \;\; \forall i=\overline{1,n} \qquad (1)$$

Hay theo dạng kỳ vọng

$$E(Y_i) = \beta_1+\beta_2X_{2i}+\ldots+\beta_kX_{ki} \;\; \forall i=\overline{1,n} \qquad (2)$$

Trong đó:

Y là biến phụ thuộc và X_j là biến giải thích hay biến độc lập.

β_1 gọi là hệ số chặn và β_j ($\forall j=\overline{2,k}$) là các hệ số góc hay còn gọi các hệ số hồi qui riêng.

u_i là các sai số ngẫu nhiên có kỳ vọng bằng 0 phương sai hữu hạn.

Xét mô hình (2), chúng ta nhận thấy $E(Y_i)= \beta_1$ khi X_{ji}=0 và:

$$\beta_j = \frac{\partial E(Y_i)}{\partial X_{ji}}$$

Trong kinh tế, chúng ta có thể tính xấp xỉ như sau:

$$\beta_j = \frac{\partial E(Y_i)}{\partial X_{ji}} \approx \frac{\Delta E(Y_i)}{\Delta X_{ji}}$$

Với Δ thể hiện mức tăng của từng chỉ tiêu. Khi ΔX_{ji}=1, thì $\beta_j = \Delta E(Y_i)$

Với biểu thức này có thể giải thích ý nghĩa của β_j ($\forall j=\overline{2,k}$) như sau: trong điều kiện các nhân tố khác không đổi, khi X_j tăng lên một đơn vị (theo đơn vị của X_j) thì E(Y) sẽ tăng bình quân β_j đơn vị (theo đơn vị của Y).

2. Ý nghĩa của hệ số hồi qui Log-Log tổng quát

Theo dạng ngẫu nhiên

$$LnY_i = \beta_1+\beta_2LnX_{2i}+\ldots+\beta_kLnX_{ki} + u \;\; \forall i=\overline{1,n} \qquad (3)$$

Hay theo dạng kỳ vọng

$$LnE(Y_i) = \beta_1+\beta_2LnX_{2i}+\ldots+\beta_kLnX_{ki} \;\; \forall i=\overline{1,n} \qquad (4)$$

Với Ln ký hiệu của logarit theo cơ số tự nhiên.

Dạng (3) và (4) chính là dạng hàm sản xuất Cobb-Douglas đã được tuyến tính hoá.

Ý nghĩa kinh tế của các hệ số trong hàm hồi qui (3) và (4) có khác trong hàm hồi qui (1) và (2) không?

Đối với mô hình (4), chúng ta có thể thực hiện đạo hàm riêng như sau:

$$\beta_j = \frac{\partial LnE(Y_i)}{\partial LnX_{ji}} = \frac{\dfrac{\partial E(Y_i)}{Y_i}}{\dfrac{\partial X_i}{X_i}}$$

(*) Tiến sĩ, Khoa Thống kê, tin học, Trường Đại học Kinh tế Đà Nẵng

Trong kinh tế, chúng ta có thể tính xấp xỉ như sau:

$$\beta_j = \frac{\dfrac{\partial E(Y_i)}{Y_i}}{\dfrac{\partial X_i}{X_i}} \approx \frac{\dfrac{\Delta E(Y_i)}{Y_i}}{\dfrac{\Delta X_i}{X_i}}$$

Với $\dfrac{\Delta E(Y_i)}{Y_i}$ và $\dfrac{\Delta X_i}{X_i}$ thể hiện tốc độ tăng của từng chỉ tiêu. Khi $\dfrac{\Delta X_i}{X_i} = 1$, thì $\beta_j = \dfrac{\Delta E(Y_i)}{Y_i}$

Như vậy, có thể nói β_j chính là hệ số co giãn của $E(Y_i)$ theo X_j.

Với biểu thức này có thể giải thích ý nghĩa của β_j ($\forall j = \overline{2,k}$) như sau: trong điều kiện các nhân tố khác không đổi, khi X_j tăng lên 1% thì E(Y) sẽ tăng bình quân β_j%.

3. Ý nghĩa của hệ số hồi qui Tuyến tính - Log tổng quát

Theo dạng ngẫu nhiên

$$Y_i = \beta_1+\beta_2 LnX_{2i}+\ldots+\beta_k LnX_{ki} + u_i \quad \forall i = \overline{1,n} \quad (5)$$

Hay theo dạng kỳ vọng

$$E(Y_i) = \beta_1+\beta_2 LnX_{2i}+\ldots+\beta_k LnX_{ki} \quad \forall i = \overline{1,n} \quad (6)$$

Ý nghĩa của các hệ số trong hàm này được giải thích như thế nào?

Thực hiện đạo hàm riêng trong mô hình (6) như sau:

$$\beta_j = \frac{\partial E(Y_i)}{\partial LnX_{ji}} = \frac{\partial E(Y_i)}{\dfrac{\partial X_i}{X_i}}$$

Trong kinh tế, chúng ta có thể tính xấp xỉ như sau:

$$\beta_j = \frac{\partial E(Y_i)}{\dfrac{\partial X_i}{X_i}} \approx \frac{\Delta E(Y_i)}{\dfrac{\Delta X_i}{X_i}}$$

Với $\Delta E(Y_i)$ thể hiện mức tăng của $E(Y_i)$ và $\dfrac{\Delta X_{ji}}{X_{ji}}$ thể hiện tốc độ tăng của X_j. Khi $\dfrac{\Delta X_{ji}}{X_{ji}} = 1$, thì $\beta_j = \Delta E(Y_i)$

Với biểu thức này có thể giải thích ý nghĩa của β_j ($\forall j = \overline{2,k}$) như sau: trong điều kiện các nhân tố khác không đổi, khi X_j tăng lên 1% thì E(Y) sẽ tăng bình quân β_j đơn vị (theo đơn vị tính của Y).

4. Ý nghĩa của hệ số hồi qui Log-Tuyến tính tổng quát

Theo dạng ngẫu nhiên

$$LnY_i = \beta_1+\beta_2 X_{2i}+\ldots+\beta_k X_{ki} + u_i \quad \forall i = \overline{1,n} \quad (7)$$

Hay theo dạng kỳ vọng

$$E(LnY_i) = \beta_1+\beta_2 X_{2i}+\ldots+\beta_k X_{ki} \quad \forall i = \overline{1,n} \quad (8)$$

Ý nghĩa của các hệ số trong hàm này được giải thích như thế nào?

Thực hiện đạo hàm riêng theo biến X_j trong mô hình (8) như sau:

$$\beta_j = \frac{\partial LnE(Y_i)}{\partial X_{ji}} = \frac{\dfrac{\partial E(Y_i)}{Y_i}}{\partial X_{ji}}$$

Trong kinh tế, chúng ta có thể tính xấp xỉ như sau:

$$\beta_j = \frac{\partial LnE(Y_i)}{\partial X_i} \approx \frac{\dfrac{\Delta E(Y_i)}{Y_i}}{\Delta X_i}$$

Với $\Delta(X_{ji})$ thể hiện mức tăng của X_j và $\frac{\Delta Y_i}{Y_i}$ thể hiện tốc độ tăng của Y. Khi $\Delta X_{ji} = 1$, thì $\beta_j = \frac{\Delta E(Y_i)}{Y_i}$

Với biểu thức này có thể giải thích ý nghĩa của β_j ($\forall j = \overline{2,k}$) như sau: trong điều kiện các nhân tố khác không đổi, khi X_j tăng lên 1 đơn vị (theo đơn vị tính của X_j) thì E(Y) sẽ tăng bình quân β_j%.

5. Ý nghĩa của hệ số hồi qui tương ứng với biến giả trong mô hình Log-Tuyến tính tổng quát

Xét mô hình hồi qui log tuyến tính như sau:

$$Ln(Y_i) = \alpha + \sum_{j=1}^{k} \beta_j X_{ji} + \lambda D_i + u_i \qquad (9)$$

Với X_j là các biến liên tục có hệ số hồi qui là β_j và D là biến giả có hệ số hồi qui là λ, u_i là các sai số ngẫu nhiên có kỳ vọng bằng 0 phương sai hữu hạn.

Theo Halvorsen và Palmquist (1980) và Kennedy (1981), để xác định phần biến động của Y khi D_i=1(g) được tính là $Y_{D=1}=(1+g)Y_{D=0}$

Lấy logaric hai vế, chúng ta có:

$Ln(Y_{D=1}) = Ln((1+g)Y_{D=0})$.

Như vậy, khi biến đổi thêm sẽ có:

$$Ln(Y_{D=1}) - Ln(Y_{D=0}) = Ln(1+g) \qquad (10)$$

Hơn nữa, từ mô hình (9), chúng ta có:

$$Ln(Y_{D=1}) - Ln(Y_{D=0}) = \lambda \qquad (11)$$

Kết hợp (10) và (11), chúng ta được $g=Exp(\lambda)-1$.

Như vậy, để xác định ảnh hưởng của biến giả đến biến động của biến phụ thuộc trong hàm log tuyến tính, cần tính antilog của hệ số hồi qui của biến giả đã được ước lượng và trừ cho 1.

Ví dụ: có tài liệu giả định về tiền lương, tuổi nghề và giới tính như trên bảng sau:

TUỔI NGHỀ, GIỚI TÍNH VÀ TIỀN LƯƠNG CỦA CÔNG NHÂN

(số liệu giả định)

Tiền lương (Y)	Tuổi nghề (TN)	Giới tính (D)
25	1	1
21.5	1	0
26	2	1
23	2	1
27	3	0
24	3	0
28.5	3	0
25.1	4	1
27	5	1
30	5	1
28	6	0
29.5	6	1
33.5	7	1
31	7	0

Thực hiện hồi qui theo mô hình sau:

$Ln(Y_i)=\alpha+\beta TN_i+\lambda D_i+u_i$

Kết quả hồi qui bằng công cụ Regression trong Data Analysis của chương trình Microsoft Excel như sau

Regression Statistics	
Multiple R	0.845212
R Square	0.714384
Adjusted R Square	0.662454
Standard Error	0.07074
Observations	14

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.13768	0.06884	13.75663	0.001016
Residual	11	0.055045	0.005004		
Total	13	0.192725			

	Coefficients	*Std Error*	*t Stat*	*P-value*
Intercept	3.089413	0.046129	66.97383	1.02E-15
TN	0.048764	0.009383	5.19677	0.000296
D	0.019071	0.038236	0.498769	0.627765

Với kết quả này, chúng ta giải thích ý nghĩa của các hệ số hồi qui là:

- Trong điều kiện các nhân tố khác không đổi, tuổi nghề tăng 1 năm thì tiền lương tăng 4,8%.

- Để giải thích hệ số của biến giả trước tiên cần tính $e^{0.019071}$=1.019 như vậy, trong điều kiện các nhân tố khác không đổi, thì tiền lương của nam lớn hơn của nữ 1,9%.

Hiện nay, trong phân tích kinh tế, các nhà kinh tế sử dụng rất nhiều mô hình khác nhau. Tính đa dạng của các mô hình tạo nên nội dung phân tích phong phú nhưng cũng làm cho việc giải thích ý nghĩa của các mô hình trở nên khó khăn hơn. Bài viết này cũng chỉ trình bày cách tiếp cận toán học trong việc giải thích ý nghĩa của các tham số trong một số mô hình. Hy vọng với cách tiếp cận này sẽ là ý tưởng cho việc giải thích ý nghĩa các mô hình hồi qui khác■

Tài liệu tham khảo:

1. S.Charles Maurice, Charles W.Smithson (1990), *Kinh tế quản lý*, Trung tâm tài liệu Thông tin ĐHKT Quốc dân, Hà Nội.

2. Jan Kmenta (1986), *Elements of Econometrics*, Second Edition, Macmillan, NewYork.

3. Guijarati (1988), *Basic Econometrics*, Mc Graw Hill Publishing, NewYork.

4. Maddala (1992), *Introduction to Econometrics*, Macmillan Publishing Company.

5. William H.Greene (1991), *Econometric Analysis*, Macmillan publishing company, NewYork.

6. Paul Newbold (1995), *Statistics for Business& Economics*, Fourth Edition, Prentice-Hall International, Inc.