

Chương 1:  
MÔ HÌNH HỒI QUY TUYẾN TÍNH HAI BIẾN

Bộ môn Toán kinh tế  
Trường Đại học Ngân hàng TPHCM  
Blog: <https://nguyenphuongblog.wordpress.com>  
Email: [nguyenphuong0122@gmail.com](mailto:nguyenphuong0122@gmail.com)

Ngày 18 tháng 9 năm 2016

# NỘI DUNG

- 1 Mô hình và một số khái niệm
  - Mô hình hồi quy
  - Hàm hồi quy tổng thể
  - Hàm hồi quy mẫu
- 2 Phương pháp ước lượng OLS
  - Tư tưởng của phương pháp OLS
  - Công thức ước lượng hệ số chặn, hệ số góc
  - Trình bày kết quả phân tích hồi quy
- 3 Tính không chệch và độ chính xác của ước lượng OLS
  - Các giả thiết của phương pháp OLS
  - Độ chính xác của ước lượng OLS
- 4 Độ phù hợp của hàm hồi quy - hệ số xác định  $R^2$
- 5 Khoảng tin cậy cho  $\beta_1, \beta_2$  và  $\sigma^2$ 
  - Phân phối xác suất của các ước lượng
  - Khoảng tin cậy cho  $\beta_1, \beta_2$
  - Khoảng tin cậy cho phương sai sai số ngẫu nhiên
- 6 Kiểm định giả thuyết
  - Kiểm định giả thuyết về hệ số hồi quy

Bài toán quan trọng trong phân tích kinh tế: đánh giá tác động của của một biến số lên một số biến số khác.

Thí dụ: muốn đánh giá tác động của thu nhập lên chi tiêu tiêu dùng.

Suy luận thông thường: khi thu nhập tăng thì mức chi tiêu tiêu dùng sẽ gia tăng.

→ có thể biểu diễn mối quan hệ phụ thuộc hàm số giữa các biến này như sau:

$$TD = f(TN) ?$$

Mô hình hồi quy tuyến tính

Mô hình hồi quy tuyến tính hai biến thể hiện mối quan hệ phụ thuộc giữa biến Y và biến X có dạng như sau:

$$Y = \beta_1 + \beta_2 X + u$$

Biến phụ thuộc	Biến độc lập
- là biến số mà ta đang quan tâm đến giá trị của nó, thường kí hiệu là Y và nằm ở vế trái của phương trình.	- là biến số được cho là có tác động đến biến phụ thuộc, thường kí hiệu là X và nằm ở vế phải của phương trình.
- còn được gọi là biến được giải thích.	- còn được gọi là biến giải thích.

Sai số ngẫu nhiên: là yếu tố đại diện cho các yếu tố có tác động đến biến Y ngoài X.

Hồi quy nghiên cứu sự phụ thuộc của một đại lượng kinh tế này (biến phụ thuộc, biến được giải thích) vào một hay nhiều đại lượng kinh tế khác (biến độc lập, biến giải thích) dựa trên ý tưởng là ước lượng giá trị trung bình của biến phụ thuộc trên cơ sở các giá trị biết trước của các biến độc lập.

- Biến độc lập có giá trị xác định trước
- Biến phụ thuộc là đại lượng ngẫu nhiên tuân theo các quy luật phân bố xác suất.

Hàm hồi quy tổng thể là hồi quy được thực hiện trên số liệu của tổng thể và phản ánh chính xác mối quan hệ giữa biến độc lập và biến phụ thuộc.

Hàm hồi quy tổng thể – PRF:  $E(Y|X) = \beta_1 + \beta_2 X$ .

Mô hình hồi quy tổng thể – PRM:  $Y_i = \beta_1 + \beta_2 X_i + u_i, \quad i = \overline{1, N};$

hoặc:  $Y = \beta_1 + \beta_2 X + u$ .

trong đó  $E(Y|X)$  là kỳ vọng của biến  $Y$  khi biết giá trị của  $X$ , hay còn gọi là kỳ vọng của  $Y$  với điều kiện  $X$ .

Thí dụ. Hồi quy TD (tiêu dùng) theo TN (thu nhập).

Mô hình hồi quy tuyến tính như sau:  $TD = \beta_1 + \beta_2 TN + u$

Các hệ số hồi quy

- $\beta_1$  được gọi là hệ số chặn, nó chính bằng giá trị trung bình của biến phụ thuộc  $Y$  khi biến độc lập  $X$  nhận giá trị bằng 0.
- $\beta_2$  được gọi là hệ số góc cho biết: khi biến độc lập  $X$  tăng một đơn vị thì giá trị trung bình của biến phụ thuộc  $Y$  thay đổi  $\beta_2$  đơn vị.

Hàm hồi quy mẫu là hồi quy được thực hiện trên số liệu của mẫu dùng để ước lượng hàm hồi quy tổng thể.

Hàm hồi quy mẫu – SRF:  $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X.$

Mô hình hồi quy mẫu – SRM:  $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i, \quad i = \overline{1, n};$

hoặc:  $Y = \hat{\beta}_1 + \hat{\beta}_2 X + e.$

trong đó  $\hat{Y}$  là ước lượng cho  $E(Y|X_i)$ ;  $\hat{\beta}_1, \hat{\beta}_2$  là ước lượng cho  $\beta_1, \beta_2$ ;  $e_i$  là phần dư, ước lượng cho  $u_i$ .

### Ví dụ 1.1

Ước lượng hàm hồi quy tuyến tính của chi tiêu ( $Y$  – triệu đồng/tháng) theo thu nhập ( $X$  – triệu đồng/tháng), ta được:

$$\hat{Y}_i = 2,066116 + 0,831956X_i$$

Xét mô hình hồi quy tổng thể:  $Y = \beta_1 + \beta_2 X + u$

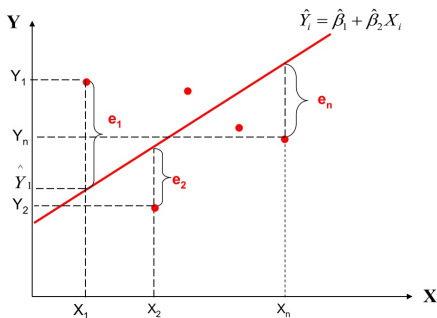
Mô hình hồi quy mẫu tại mỗi quan sát:  $Y_i = \beta_1 + \beta_2 X_i + u_i$

Với  $\widehat{\beta}_1, \widehat{\beta}_2$  là các ước lượng của  $\beta_1, \beta_2$ , ta có thể viết hàm hồi quy mẫu như sau:

$$\widehat{Y}_i = \widehat{\beta}_1 + \widehat{\beta}_2 X_i$$

Sai lệch giữa giá trị thực tế  $Y_i$  và giá trị ước lượng tương ứng từ hàm hồi quy mẫu  $\widehat{Y}_i$  là phần dư

$$e_i = Y_i - \widehat{Y}_i$$



$\widehat{\beta}_1, \widehat{\beta}_2$  được xác định dựa trên tiêu chuẩn cực tiểu tổng bình phương các phần dư, được gọi là phương pháp bình phương bé nhất.

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \widehat{Y}_i)^2 = \sum_{i=1}^n (Y_i - \widehat{\beta}_1 - \widehat{\beta}_2 X_i)^2$$

Tìm  $\widehat{\beta}_1, \widehat{\beta}_2$  sao cho:  $f(\widehat{\beta}_1, \widehat{\beta}_2) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \widehat{\beta}_1 - \widehat{\beta}_2 X_i)^2 \rightarrow \min$

$$\widehat{\beta}_1, \widehat{\beta}_2 \text{ sẽ là nghiệm của hệ sau: } \begin{cases} \frac{\partial f}{\partial \widehat{\beta}_1} = \frac{\partial \sum_{i=1}^n (Y_i - \widehat{\beta}_1 - \widehat{\beta}_2 X_i)^2}{\partial \widehat{\beta}_1} = 0 \\ \frac{\partial f}{\partial \widehat{\beta}_2} = \frac{\partial \sum_{i=1}^n (Y_i - \widehat{\beta}_1 - \widehat{\beta}_2 X_i)^2}{\partial \widehat{\beta}_2} = 0 \end{cases}$$

$$\text{Giải hệ ta được: } \begin{cases} \widehat{\beta}_1 = \overline{Y} - \widehat{\beta}_2 \overline{X} \\ \widehat{\beta}_2 = \frac{\overline{XY} - \overline{X} \cdot \overline{Y}}{X^2 - (\overline{X})^2} \end{cases} \Leftrightarrow \begin{cases} \widehat{\beta}_1 = \overline{Y} - \widehat{\beta}_2 \overline{X} \\ \widehat{\beta}_2 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \end{cases}$$

với  $x_i = X_i - \overline{X}, y_i = Y_i - \overline{Y}$  và  $\overline{X}, \overline{Y}$  là trung bình mẫu của  $X, Y$ .



## Ví dụ 2.1

Quan sát về thu nhập ( $X$  – triệu đồng/năm) và chi tiêu ( $Y$  – triệu đồng/năm) của 10 người, ta được các số liệu sau:

X	100	80	98	95	75	79	78	69	81	88
Y	90	75	78	88	62	69	65	65	60	70

Hãy ước lượng hàm hồi quy tuyến tính của  $Y$  theo  $X$ .

$$\sum Y_i = 722; \quad \sum Y_i^2 = 53108; \quad \sum X_i = 843;$$

$$\sum X_i^2 = 72045; \quad \sum X_i Y_i = 61680; \quad n = 10$$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{843}{10} = 84,3; \quad \bar{Y} = \frac{722}{10} = 72,2$$

$$\sum x_i y_i = \sum_{i=1}^n X_i Y_i - n \cdot \bar{X} \cdot \bar{Y} = 61680 - 10 \times 84,3 \times 72,2 = 815,4$$

$$\sum x_i^2 = \sum_{i=1}^n X_i^2 - n \cdot (\bar{X})^2 = 72045 - 10 \cdot (84,3)^2 = 980,1$$

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} = \frac{815,4}{980,1} = 0,831956; \quad \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} = 72,2 - 0,831956 \times 84,3 = 2,066116$$

$$\widehat{Y}_i = 2,066116 + 0,831956 X_i$$

- Giá trị  $\hat{\beta}_2 = 0,831956$  chỉ ra rằng khi thu nhập tăng 1 triệu đồng/năm thì chi tiêu trung bình của một người tăng khoảng 0,831956 triệu đồng.