



# Hàm hồi quy tổng thể và hồi quy mẫu

Bởi:

Phạm Trí Cao

## Hàm hồi quy tổng thể và hồi quy mẫu

### Hàm hồi quy tổng thể (PRF)

Ví dụ 3.1. Hồi quy tiêu dùng  $Y$  theo thu nhập  $X$ .

Theo Keynes thì hàm tiêu dùng như sau

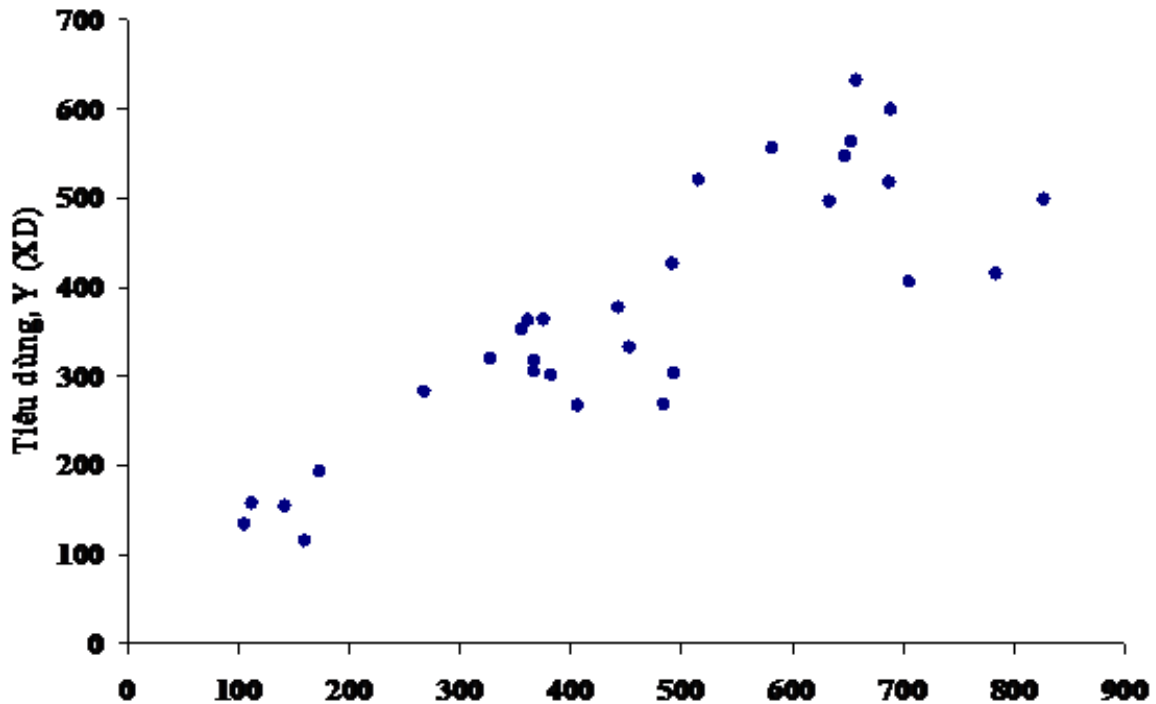
Damodar N Gujarati, Basic Economics-3<sup>rd</sup> Edition, p4.

:

$Y = \beta_1 + \beta_2 X$ , với  $\beta_2$  là xu hướng tiêu dùng biên,  $0 < \beta_2 < 1$ . (3.1)

Chúng ta kiểm chứng giả thiết trên với số liệu từ một nước giả định  $Z$  có dân số 30 người với số liệu tiêu dùng và thu nhập của từng người như đồ thị phân tán sau.

Số liệu ở phụ lục 3.1.PL cuối chương 3.



Thu nhập X (XD)

Hình 3.1. Đồ thị phân tán quan hệ giữa tiêu dùng và thu nhập khả dụng.

Đồ thị 3.1. cho thấy có mối quan hệ đồng biến giữa tiêu dùng và thu nhập khả dụng, hay là thu nhập tăng sẽ làm tiêu dùng tăng. Tuy quan hệ giữa Y và X không chính xác như hàm bậc nhất (3.1).

Trong phân tích hồi quy chúng ta xem biến độc lập X có giá trị xác định trong khi biến phụ thuộc Y là biến ngẫu nhiên. Điều này tưởng như bất hợp lý. Khi chúng ta chọn ngẫu nhiên người thứ i thì chúng ta thu được đồng thời hai giá trị:  $X_i$  là thu nhập và  $Y_i$  là tiêu dùng của người đó. Vậy tại sao lại xem  $Y_i$  là ngẫu nhiên? Câu trả lời như sau: Xét một mức thu nhập  $X_i$  xác định, cách lấy mẫu của chúng ta là chọn ngẫu nhiên trong số những người có thu nhập là  $X_i$ . Thu nhập góp phần chính yếu quyết định tiêu dùng như thể hiện ở hàm số (1.3), tuy nhiên còn nhiều yếu tố khác cũng tác động lên tiêu dùng nên ứng với một cách lấy mẫu thì với nhiều lần lấy mẫu với tiêu chí  $X = X_i$  ta nhận được các giá trị  $Y_i$  khác nhau. Vậy chính xác hơn biến phụ thuộc Y là một biến ngẫu nhiên có điều kiện theo biến độc lập X. Ước lượng tốt nhất cho Y trong trường hợp này là giá trị kỳ vọng của Y ứng với điều kiện X nhận giá trị  $X_i$  xác định.

Hàm hồi quy tổng thể (PRF):

$$E(Y/X=X_i) = \beta_1 + \beta_2 X \quad (3.2)$$

Hàm hồi quy tổng thể và hồi quy mẫu

Đối với một quan sát cụ thể thì giá trị biến phụ thuộc lệch khỏi kỳ vọng toán, vậy:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_i \quad (3.3)$$

$\beta_1$  và  $\beta_2$  : các tham số của mô hình

$\beta_1$  : tung độ gốc

$\beta_2$ : độ dốc

Giá trị ước lượng của  $Y_i$

$$\hat{Y}_i = \beta_1 + \beta_2 X_i$$

$\beta_i$  : Sai số của hồi quy hay còn được gọi là nhiễu ngẫu nhiên

Nhiễu ngẫu nhiên hình thành từ nhiều nguyên nhân:

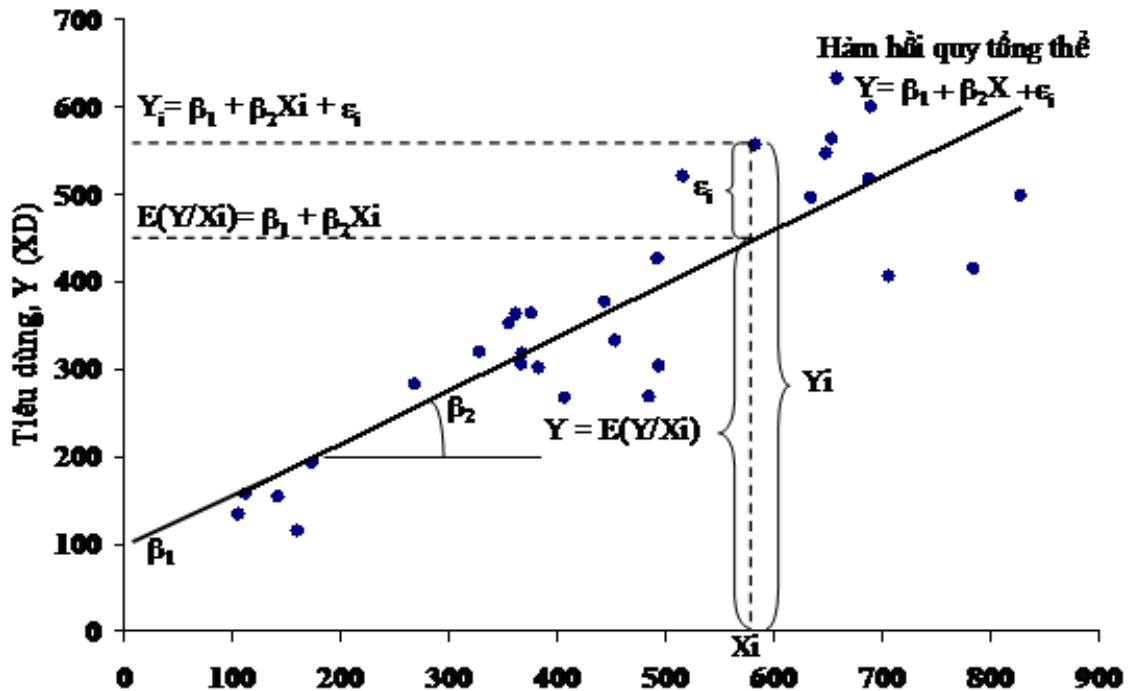
Bỏ sót biến giải thích.

Sai số khi đo lường biến phụ thuộc.

Các tác động không tiên đoán được.

Dạng hàm hồi quy không phù hợp.

Dạng hàm hồi quy (3.2) được gọi là hồi quy tổng thể tuyến tính. Chúng ta sẽ thảo luận chi tiết về thuật ngữ hồi quy tuyến tính ở cuối chương. Hình 3.2 cho ta cái nhìn trực quan về hồi quy tổng thể tuyến tính và sai số của hồi quy.



Thu nhập X (XD)

Hình 3.2. Hàm hồi quy tổng thể tuyến tính

### Hàm hồi quy mẫu (SRF)

Trong thực tế hiếm khi chúng có số liệu của tổng thể mà chỉ có số liệu mẫu. Chúng ta phải sử dụng dữ liệu mẫu để ước lượng hàm hồi quy tổng thể.

Hàm hồi quy mẫu:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i \quad (3.4)$$

Trong đó

$\hat{\beta}_1$ : ước lượng cho  $\beta_1$ .

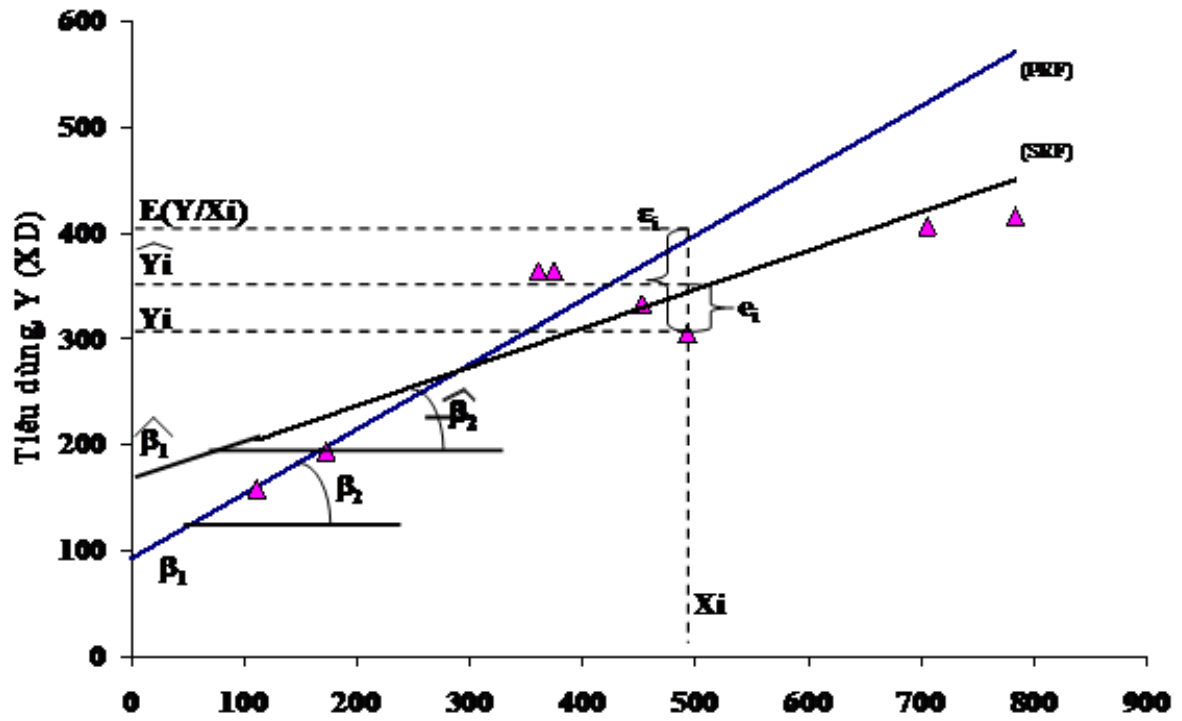
$\hat{\beta}_2$ : Ước lượng cho  $\beta_2$ .

Đối với quan sát thứ  $i$ :

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i \quad (3.5)$$

## Hàm hồi quy tổng thể và hồi quy mẫu

Hình 3.3 cho thấy sự xấp xỉ của hàm hồi quy mẫu (SRF) và hàm hồi quy tổng thể (PRF).



Thu nhập X (XD)

Hình 3.3. Hồi quy mẫu và hồi quy tổng thể