

# Hồi Quy Tuyến Tính Đơn Biến

Nguyễn Thị Hồng Nhung<sup>1</sup>

<sup>1</sup>Trường Đại học Khoa học Tự nhiên, ĐHQG-Tp HCM.

[nthnhung@hcmus.edu.vn](mailto:nthnhung@hcmus.edu.vn)

## Nội dung

## Giới thiệu

## Mô hình hồi quy tuyến tính đơn biến

## Ước lượng phương trình hồi quy

## Tính phù hợp của mô hình

### Hệ số xác định

## Bài tập

## Giới thiệu

Độ sạch của oxy và tỷ lệ phần trăm hydrocarbon tại một số điểm đo thu được kết quả như bảng sau:

## Giới thiệu

Độ sạch của oxy và tỷ lệ phần trăm hydrocarbon tại một số điểm đo thu được kết quả như bảng sau:

TT	x(%)	y(%)	TT	x(%)	y(%)	TT	x(%)	y(%)
1	0.99	90.01	8	1.23	91.77	15	1.11	89.85
2	1.02	89.05	9	1.55	99.42	16	1.2	90.39
3	1.15	91.43	10	1.4	93.65	17	1.26	93.25
4	1.29	93.74	11	1.19	93.54	18	1.32	93.41
5	1.46	96.73	12	1.15	92.52	19	1.43	94.98
6	1.36	94.45	13	0.98	90.56	20	0.95	87.33
7	0.87	87.59	14	1.01	89.54	21	1.32	94.01

## Giới thiệu

Độ sạch của oxy và tỷ lệ phần trăm hydrocarbon tại một số điểm đo thu được kết quả như bảng sau:

TT	x(%)	y(%)	TT	x(%)	y(%)	TT	x(%)	y(%)
1	0.99	90.01	8	1.23	91.77	15	1.11	89.85
2	1.02	89.05	9	1.55	99.42	16	1.2	90.39
3	1.15	91.43	10	1.4	93.65	17	1.26	93.25
4	1.29	93.74	11	1.19	93.54	18	1.32	93.41
5	1.46	96.73	12	1.15	92.52	19	1.43	94.98
6	1.36	94.45	13	0.98	90.56	20	0.95	87.33
7	0.87	87.59	14	1.01	89.54	21	1.32	94.01

- Mỗi tương quan giữa độ sạch của oxy và tỷ lệ phần trăm hydrocarbon.

## Giới thiệu

Độ sạch của oxy và tỷ lệ phần trăm hydrocarbon tại một số điểm đo thu được kết quả như bảng sau:

TT	x(%)	y(%)	TT	x(%)	y(%)	TT	x(%)	y(%)
1	0.99	90.01	8	1.23	91.77	15	1.11	89.85
2	1.02	89.05	9	1.55	99.42	16	1.2	90.39
3	1.15	91.43	10	1.4	93.65	17	1.26	93.25
4	1.29	93.74	11	1.19	93.54	18	1.32	93.41
5	1.46	96.73	12	1.15	92.52	19	1.43	94.98
6	1.36	94.45	13	0.98	90.56	20	0.95	87.33
7	0.87	87.59	14	1.01	89.54	21	1.32	94.01

- ▶ Mỗi tương quan giữa độ sạch của oxy và tỷ lệ phần trăm hydrocarbon.
- ▶ Mức độ tương quan.

## Giới thiệu

Độ sạch của oxy và tỷ lệ phần trăm hydrocarbon tại một số điểm đo thu được kết quả như bảng sau:

TT	x(%)	y(%)	TT	x(%)	y(%)	TT	x(%)	y(%)
1	0.99	90.01	8	1.23	91.77	15	1.11	89.85
2	1.02	89.05	9	1.55	99.42	16	1.2	90.39
3	1.15	91.43	10	1.4	93.65	17	1.26	93.25
4	1.29	93.74	11	1.19	93.54	18	1.32	93.41
5	1.46	96.73	12	1.15	92.52	19	1.43	94.98
6	1.36	94.45	13	0.98	90.56	20	0.95	87.33
7	0.87	87.59	14	1.01	89.54	21	1.32	94.01

- ▶ Mỗi tương quan giữa độ sạch của oxy và tỷ lệ phần trăm hydrocarbon.
- ▶ Mức độ tương quan.
- ▶ Tiên đoán tỷ lệ phần trăm hydrocarbon ứng với độ sạch oxy.

## Giới thiệu

Độ sạch của oxy và tỷ lệ phần trăm hydrocarbon tại một số điểm đo thu được kết quả như bảng sau:

TT	x(%)	y(%)	TT	x(%)	y(%)	TT	x(%)	y(%)
1	0.99	90.01	8	1.23	91.77	15	1.11	89.85
2	1.02	89.05	9	1.55	99.42	16	1.2	90.39
3	1.15	91.43	10	1.4	93.65	17	1.26	93.25
4	1.29	93.74	11	1.19	93.54	18	1.32	93.41
5	1.46	96.73	12	1.15	92.52	19	1.43	94.98
6	1.36	94.45	13	0.98	90.56	20	0.95	87.33
7	0.87	87.59	14	1.01	89.54	21	1.32	94.01

- ▶ Mỗi tương quan giữa độ sạch của oxy và tỷ lệ phần trăm hydrocarbon.
- ▶ Mức độ tương quan.
- ▶ Tiên đoán tỷ lệ phần trăm hydrocarbon ứng với độ sạch oxy.



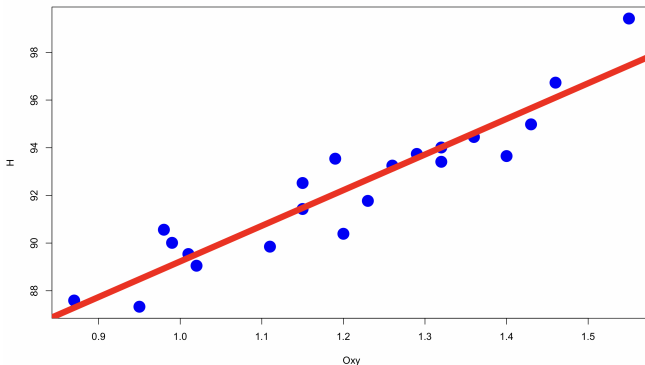
## Giới thiệu

Mối tương quan giữa độ sạch của oxy và tỷ lệ phần trăm hydrocarbon.



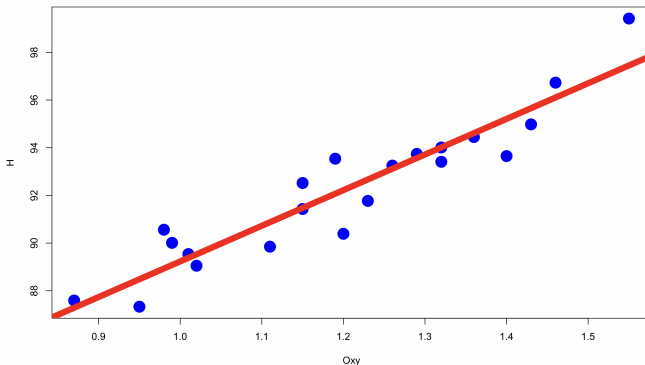
## Giới thiệu

Mối tương quan giữa độ sạch của oxy và tỷ lệ phần trăm hydrocarbon.



## Giới thiệu

Mối tương quan giữa độ sạch của oxy và tỷ lệ phần trăm hydrocarbon.



## Phương sai và hiệp phương sai

- # Gọi  $X$  và  $Y$  là hai biến ngẫu nhiên rút ra từ một mẫu quan sát  $n$  đối tượng.

## Phương sai và hiệp phương sai

- # Gọi  $X$  và  $Y$  là hai biến ngẫu nhiên rút ra từ một mẫu quan sát  $n$  đối tượng.
- # Đo lường độ dao động giữa  $X$  và  $Y$ : phương sai.

## Phương sai và hiệp phương sai

- # Gọi  $X$  và  $Y$  là hai biến ngẫu nhiên rút ra từ một mẫu quan sát  $n$  đối tượng.
- # Đo lường độ dao động giữa  $X$  và  $Y$ : phương sai.
- # Hiệp phương sai giữa  $X$  và  $Y$ : đo lường cường độ tương quan.

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

## Phương sai và hiệp phương sai

- # Gọi  $X$  và  $Y$  là hai biến ngẫu nhiên rút ra từ một mẫu quan sát  $n$  đối tượng.
- # Đo lường độ dao động giữa  $X$  và  $Y$ : phương sai.
- # Hiệp phương sai giữa  $X$  và  $Y$ : đo lường cường độ tương quan.

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- # Nếu hiệp phương sai bằng 0 thì  $X$  và  $Y$  độc lập với nhau.

## Phương sai và hiệp phương sai

- # Gọi  $X$  và  $Y$  là hai biến ngẫu nhiên rút ra từ một mẫu quan sát  $n$  đối tượng.
- # Đo lường độ dao động giữa  $X$  và  $Y$ : phương sai.
- # Hiệp phương sai giữa  $X$  và  $Y$ : đo lường cường độ tương quan.

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- # Nếu hiệp phương sai bằng 0 thì  $X$  và  $Y$  độc lập với nhau.
- # Hiệp phương sai âm: lệch pha giữa hai phân phối theo hướng ngược chiều nhau.



## Phương sai và hiệp phương sai

- # Gọi  $X$  và  $Y$  là hai biến ngẫu nhiên rút ra từ một mẫu quan sát  $n$  đối tượng.
- # Đo lường độ dao động giữa  $X$  và  $Y$ : phương sai.
- # Hiệp phương sai giữa  $X$  và  $Y$ : đo lường cường độ tương quan.

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- # Nếu hiệp phương sai bằng 0 thì  $X$  và  $Y$  độc lập với nhau.
- # Hiệp phương sai âm: lệch pha giữa hai phân phối theo hướng ngược chiều nhau.
- # Hiệp phương sai dương: lệch pha giữa hai phân phối theo hướng cùng chiều nhau.

## Phương sai và hiệp phương sai

- # Gọi  $X$  và  $Y$  là hai biến ngẫu nhiên rút ra từ một mẫu quan sát  $n$  đối tượng.
- # Đo lường độ dao động giữa  $X$  và  $Y$ : phương sai.
- # Hiệp phương sai giữa  $X$  và  $Y$ : đo lường cường độ tương quan.

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- # Nếu hiệp phương sai bằng 0 thì  $X$  và  $Y$  độc lập với nhau.
- # Hiệp phương sai âm: lệch pha giữa hai phân phối theo hướng ngược chiều nhau.
- # Hiệp phương sai dương: lệch pha giữa hai phân phối theo hướng cùng chiều nhau.

## Hệ số tương quan

- # Hệ số tương quan giữa  $X$  và  $Y$  là một hiệp phương sai được chuẩn hóa.

## Hệ số tương quan

- # Hệ số tương quan giữa  $X$  và  $Y$  là một hiệp phương sai được chuẩn hóa.
- #  $r_{XY}$  được xác định bằng:

$$r_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

## Hệ số tương quan

- # Hệ số tương quan giữa  $X$  và  $Y$  là một hiệp phương sai được chuẩn hóa.
- #  $r_{XY}$  được xác định bằng:

$$r_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$



## Mô hình hồi quy tuyến tính đơn biến

♣ Phương trình mô tả  $Y$  liên hệ với  $x$  và một sai số được gọi là **mô hình hồi quy**.

## Mô hình hồi quy tuyến tính đơn biến

- ♣ Phương trình mô tả  $Y$  liên hệ với  $x$  và một sai số được gọi là **mô hình hồi quy**.
- ♣ Mô hình **hồi quy tuyến tính đơn biến** là:

$$y = \beta_0 + \beta_1 x + \epsilon.$$



## Mô hình hồi quy tuyến tính đơn biến

- ♣ Phương trình mô tả  $Y$  liên hệ với  $x$  và một sai số được gọi là **mô hình hồi quy**.
- ♣ Mô hình **hồi quy tuyến tính đơn biến** là:

$$y = \beta_0 + \beta_1 x + \epsilon.$$

Trong đó:

$\beta_0$  và  $\beta_1$  được gọi là **các tham số của mô hình**,  
 $\epsilon$  là biến ngẫu nhiên, được gọi là **sai số**.

## Phương trình hồi quy tuyến tính đơn biến

♣ Phương trình hồi quy tuyến đơn:

$$E(y) = \beta_0 + \beta_1 x$$

# Phương trình hồi quy tuyến tính đơn biến

♣ Phương trình hồi quy tuyến tính đơn:

$$E(y) = \beta_0 + \beta_1 x$$

♠ Đồ thị của phương trình hồi quy tuyến tính đơn là đường thẳng.

# Phương trình hồi quy tuyến tính đơn biến

♣ Phương trình hồi quy tuyến tính đơn:

$$E(y) = \beta_0 + \beta_1 x$$

- ♠ Đồ thị của phương trình hồi quy tuyến tính đơn là đường thẳng.
- ♠  $\beta_0$  là tung độ gốc của đường hồi quy.

# Phương trình hồi quy tuyến tính đơn biến

## ♣ Phương trình hồi quy tuyến tính đơn:

$$E(y) = \beta_0 + \beta_1 x$$

- ♠ Đồ thị của phương trình hồi quy tuyến tính đơn là đường thẳng.
- ♠  $\beta_0$  là tung độ gốc của đường hồi quy.
- ♠  $\beta_1$  là độ dốc của đường hồi quy.

# Phương trình hồi quy tuyến tính đơn biến

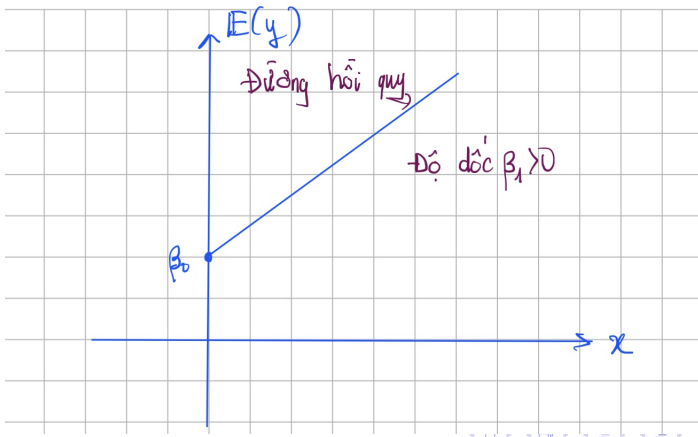
## ♣ Phương trình hồi quy tuyến tính đơn:

$$E(y) = \beta_0 + \beta_1 x$$

- ♠ Đồ thị của phương trình hồi quy tuyến tính đơn là đường thẳng.
- ♠  $\beta_0$  là tung độ gốc của đường hồi quy.
- ♠  $\beta_1$  là độ dốc của đường hồi quy.

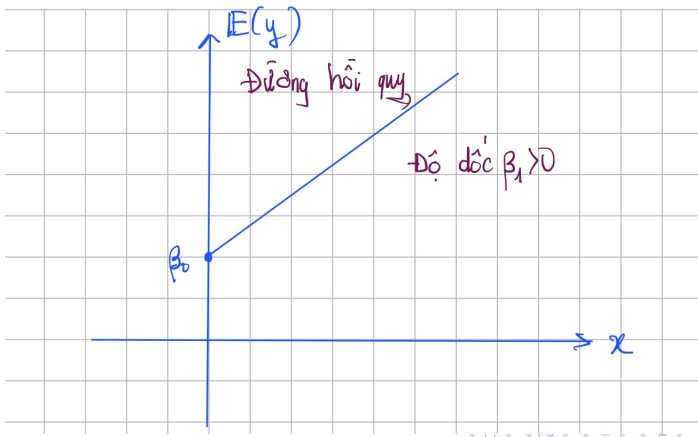
# Phương trình hồi quy tuyến tính đơn biến

## ♣ Quan hệ tuyến tính đồng biến



# Phương trình hồi quy tuyến tính đơn biến

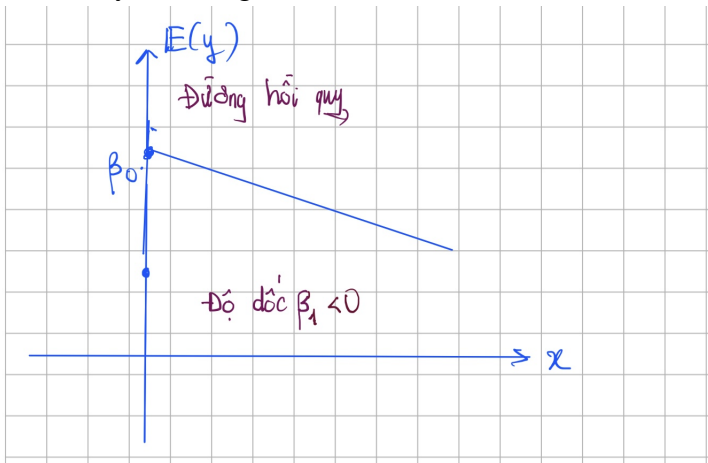
## ♣ Quan hệ tuyến tính đồng biến





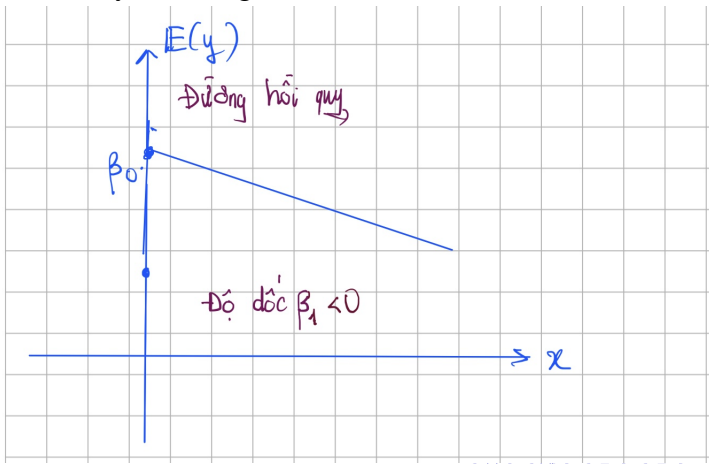
## Phương trình hồi quy tuyến tính đơn biến

### ♣ Quan hệ tuyến tính nghịch biến



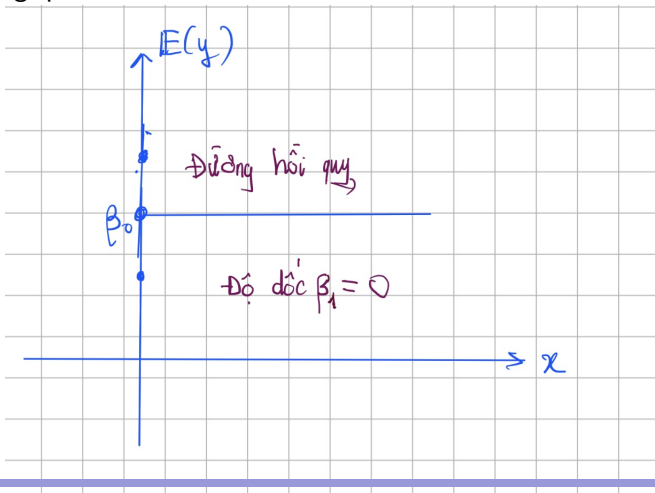
## Phương trình hồi quy tuyến tính đơn biến

### ♣ Quan hệ tuyến tính nghịch biến



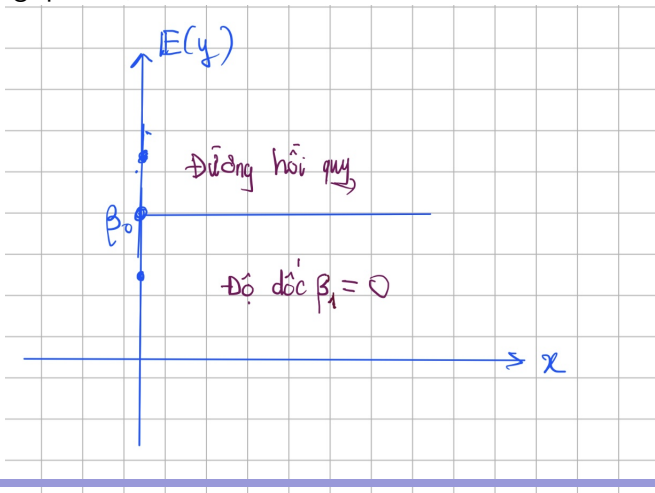
## Phương trình hồi quy tuyến tính đơn biến

♣ không quan hệ



## Phương trình hồi quy tuyến tính đơn biến

♣ không quan hệ



## Nội dung

## Giới thiệu

## Mô hình hồi quy tuyến tính đơn biến

## Ước lượng phương trình hồi quy

## Tính phù hợp của mô hình

### Hệ số xác định

## Bài tập

# Ước lượng phương trình hồi quy tuyến tính đơn biến

♣ Phương trình hồi quy tuyến đơn:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$



## Ước lượng phương trình hồi quy tuyến tính đơn biến

♣ Phương trình hồi quy tuyến đơn:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

♠ Đồ thị được gọi là đường hồi quy ước lượng.

♠  $\hat{\beta}_0$  là tung độ gốc của đường hồi quy.





# Ước lượng phương trình hồi quy tuyến tính đơn biến

## ♣ Phương trình hồi quy tuyến đơn:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

- ♠ Đồ thị được gọi là đường hồi quy ước lượng.
- ♠  $\hat{\beta}_0$  là tung độ gốc của đường hồi quy.
- ♠  $\hat{\beta}_1$  là độ dốc của đường hồi quy.
- ♠  $\hat{y}$  là giá trị ước lượng của  $y$  đối với giá trị  $x$  cho trước.

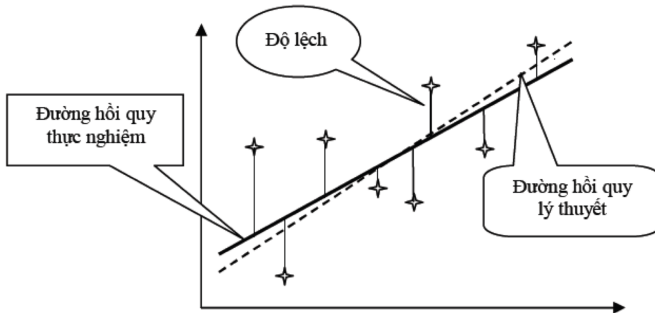
# Ước lượng phương trình hồi quy tuyến tính đơn biến

## ♣ Phương trình hồi quy tuyến đơn:

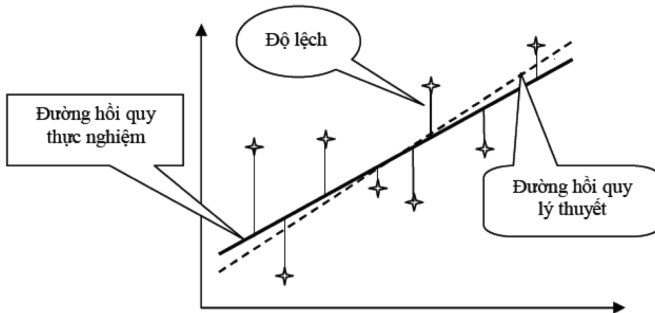
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

- ♠ Đồ thị được gọi là đường hồi quy ước lượng.
- ♠  $\hat{\beta}_0$  là tung độ gốc của đường hồi quy.
- ♠  $\hat{\beta}_1$  là độ dốc của đường hồi quy.
- ♠  $\hat{y}$  là giá trị ước lượng của  $y$  đối với giá trị  $x$  cho trước.

## Phương pháp bình phương tối thiểu



## Phương pháp bình phương tối thiểu



## Phương pháp bình phương tối thiểu

- ◇ Tổng bình phương của các độ lệch

$$L(\beta_0, \beta_1) = \sum (y_i - \hat{y}_i)^2$$

trong đó

- ▶  $y_i$  giá trị quan sát của biến phụ thuộc đối với quan sát thứ  $i$ .
  - ▶  $\hat{y}_i$  : giá trị ước lượng của biến phụ thuộc đối với quan sát thứ  $i$ .
  - ▶ Tìm  $\beta_0, \beta_1$  để  $L(\beta_0, \beta_1)$  đạt cực tiểu.
- ◇ Điều kiện cần của cực tiểu

$$\frac{\partial L}{\partial \beta_0} = \frac{\partial L}{\partial \beta_1} = 0$$

## Phương pháp bình phương tối thiểu

◇ Độ dốc của phương trình hồi quy ước lượng

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum (x_i y_i) - n\bar{x} \cdot \bar{y}}{\sum (x_i - \bar{x})^2}$$

## Phương pháp bình phương tối thiểu

- ◇ Độ dốc của phương trình hồi quy ước lượng

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum (x_i y_i) - n\bar{x} \cdot \bar{y}}{\sum (x_i - \bar{x})^2}$$

- ◇ Tung độ dốc của phương trình hồi quy ước lượng

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$



## Phương pháp bình phương tối thiểu

- ◇ Độ dốc của phương trình hồi quy ước lượng

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sum (x_i y_i) - n\bar{x} \cdot \bar{y}}{\sum (x_i - \bar{x})^2}$$

- ◇ Tung độ dốc của phương trình hồi quy ước lượng

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$



# Hệ số xác định

- Mỗi liên hệ giữa SST, SSR và SSE

$$SST = SSR + SSE$$

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

Với

◇  $SST = \sum (y_i - \bar{y})^2$ : Tổng bình phương toàn phần

# Hệ số xác định

- Mối liên hệ giữa SST, SSR và SSE

$$SST = SSR + SSE$$

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

Với

- ◇  $SST = \sum (y_i - \bar{y})^2$ : Tổng bình phương toàn phần
- ◇  $SSR = \sum (\hat{y}_i - \bar{y})^2$ : Tổng bình phương hồi quy.

## Hệ số xác định

- Mối liên hệ giữa SST, SSR và SSE

$$SST = SSR + SSE$$

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

Với

- ◇  $SST = \sum (y_i - \bar{y})^2$ : Tổng bình phương toàn phần
- ◇  $SSR = \sum (\hat{y}_i - \bar{y})^2$ : Tổng bình phương hồi quy.
- ◇  $SSE = \sum (y_i - \hat{y}_i)^2$ : Tổng bình phương sai số

## Hệ số xác định

◇ Hệ số xác định là:

$$R^2 = \frac{SSR}{SST}$$

◇ Hệ số xác định cho biết mức độ biến thiên trong Y có thể giải thích bằng mức độ biến thiên trong nhóm X.

◇  $0 \leq R^2 \leq 1$ .



## Bài tập

SV làm bài trong sách: Giáo trình Bài tập XSTK