

HỒI QUY TUYẾN TÍNH

Nội dung

① Hệ số Tương quan

② Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

③ Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

④ Tính toán với R

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hệ số Tương quan

Ví dụ I

Chiều cao của bố và con có liên quan gì đến nhau không?

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ví dụ I

Chiều cao của bố và con có liên quan gì đến nhau không?

Chiều cao của bố	160	154	166	175	163	172	159
Chiều cao của con	165	160	175	177	172	170	167

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

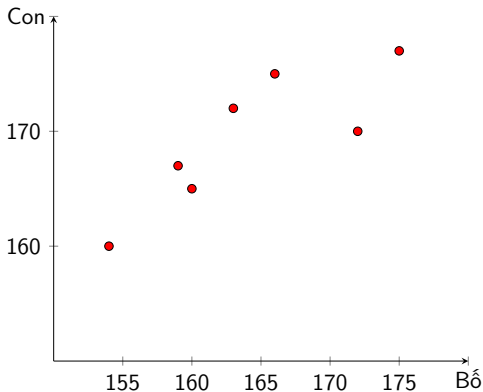
Dự đoán Quan sát
mới

Tính toán với
R

Ví dụ I

Chiều cao của bố và con có liên quan gì đến nhau không?

Chiều cao của bố	160	154	166	175	163	172	159
Chiều cao của con	165	160	175	177	172	170	167



Hình 1: Chiều cao của bố và con.

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với

R

Ví dụ II

Có mối liên hệ nào giữa độ tuổi và thời gian ngủ hay không?

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ví dụ II

Có mối liên hệ nào giữa độ tuổi và thời gian ngủ hay không?

Tuổi	4	5	28	30	50	55
Thời gian ngủ	10	9	7	7	6	5

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

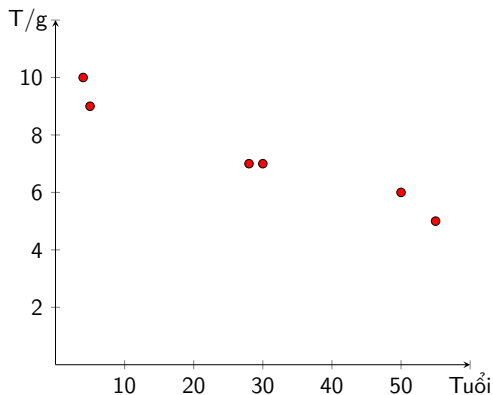
Dự đoán Quan sát
mới

Tính toán với
R

Ví dụ II

Có mối liên hệ nào giữa độ tuổi và thời gian ngủ hay không?

Tuổi	4	5	28	30	50	55
Thời gian ngủ	10	9	7	7	6	5



Hình 2: Tuổi và Thời gian ngủ.

Ví dụ III

Điểm Toán và Văn có liên quan gì đến nhau không?

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ví dụ III

Điểm Toán và Văn có liên quan gì đến nhau không?

Điểm Toán	7	8	7	9	3	5	7	10	9	8
Điểm Văn	8	9	7	5	5	7	8	7	6	8

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

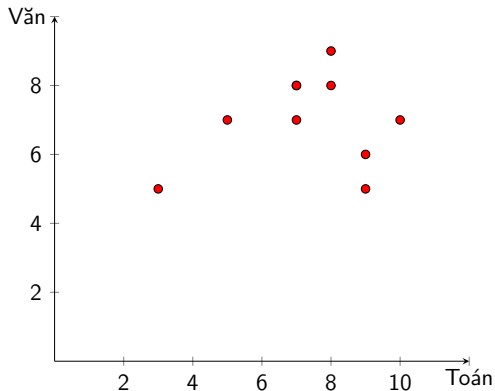
Dự đoán Quan sát
mới

Tính toán với
R

Ví dụ III

Điểm Toán và Văn có liên quan gì đến nhau không?

Điểm Toán	7	8	7	9	3	5	7	10	9	8
Điểm Văn	8	9	7	5	5	7	8	7	6	8



Hình 3: Điểm Toán và Điểm Văn.

Câu hỏi

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Trong cuộc sống, chúng ta thường gặp **hai đại lượng** và **ngghi ngờ** rằng chúng có mối liên hệ với nhau, e.g., lương cơ bản và giá nhà trọ (hoặc bất kỳ mặt hàng nào), thời gian học và điểm thi một môn học, độ tuổi và chiều cao, etc.

Câu hỏi

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Trong cuộc sống, chúng ta thường gặp **hai đại lượng** và **ngghi ngờ** rằng chúng có mối liên hệ với nhau, e.g., lương cơ bản và giá nhà trọ (hoặc bất kỳ mặt hàng nào), thời gian học và điểm thi một môn học, độ tuổi và chiều cao, etc.

Question

Làm thế nào để biết hai đại lượng có liên quan với nhau hay không?

Hệ số Tương quan

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hệ số tương quan

Cho $X = (x_1, \dots, x_n)$ và $Y = (y_1, \dots, y_n)$. Hệ số tương quan giữa X và Y được cho bởi

$$r(X, Y) = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}},$$

trong đó,

$$S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}),$$

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2,$$

$$S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2.$$

Ý nghĩa của Hệ số tương quan

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Mệnh đề 1

$$-1 \leq r(X, Y) \leq 1.$$

Ý nghĩa của Hệ số tương quan

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Mệnh đề 1

$$-1 \leq r(X, Y) \leq 1.$$

$r(X, Y)$ đo mức độ phụ thuộc tuyến tính giữa X và Y .

- Khi $r(X, Y)$ rất gần 1, thì X tăng kéo theo Y cũng tăng.
- Khi $r(X, Y)$ rất gần -1 , thì X tăng kéo theo Y giảm.
- Khi $r(X, Y)$ rất gần 0, thì sự biến thiên của X và Y không liên quan đến nhau.

Ví dụ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Trong Ví dụ I, biến x là chiều cao của bố, y là chiều cao của con

Chiều cao của bố	160	154	166	175	163	172	159
Chiều cao của con	165	160	175	177	172	170	167

$$S_{xx} = 330.857, S_{yy} = 209.714, S_{xy} = 220.571, r = 0.837.$$

Ví dụ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Trong Ví dụ I, biến x là chiều cao của bố, y là chiều cao của con

Chiều cao của bố	160	154	166	175	163	172	159
Chiều cao của con	165	160	175	177	172	170	167

$$S_{xx} = 330.857, S_{yy} = 209.714, S_{xy} = 220.571, r = 0.837.$$

- Trong Ví dụ II, biến x là Tuổi, y là Thời gian ngủ.

Tuổi	4	5	28	30	50	55
Thời gian ngủ	10	9	7	7	6	5

$$S_{xx} = 6958/3, S_{xy} = -586/3, S_{yy} = 52/3, r = -0.974.$$

Ví dụ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Trong Ví dụ I, biến x là chiều cao của bố, y là chiều cao của con

Chiều cao của bố	160	154	166	175	163	172	159
Chiều cao của con	165	160	175	177	172	170	167

$$S_{xx} = 330.857, S_{yy} = 209.714, S_{xy} = 220.571, r = 0.837.$$

- Trong Ví dụ II, biến x là Tuổi, y là Thời gian ngủ.

Tuổi	4	5	28	30	50	55
Thời gian ngủ	10	9	7	7	6	5

$$S_{xx} = 6958/3, S_{xy} = -586/3, S_{yy} = 52/3, r = -0.974.$$

- Trong Ví dụ III, x là điểm Toán, y là điểm Văn.

Điểm Toán	7	8	7	9	3	5	7	10	9	8
Điểm Văn	8	9	7	5	5	7	8	7	6	8

$$S_{xx} = 38.1, S_{xy} = 5, S_{yy} = 16, r = 0.2025.$$

Áp dụng

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Bài tập

Một công ty muốn phân tích hiệu quả của việc quảng cáo. Trong 5 tháng, công ty này thu được kết quả sau (đơn vị: trăm USD)

Tiền quảng cáo	5	8	10	15	22
Doanh thu	60	150	200	300	390

Tính hệ số tương quan.

Ứng dụng trong dự đoán

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Khi $|r(X, Y)|$ rất gần 1, biết X , ta có thể đưa ra dự đoán về Y

- Nếu biết chiều cao của bố, ta có thể dự đoán chiều cao của con.
- Nếu biết tuổi của một người, ta có thể dự đoán về thời gian ngủ một đêm của người đó.
- Biết điểm Toán không giúp ta dự đoán điểm môn Văn.

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hồi quy Tuyến tính đơn

Tại sao nên nghiên cứu hồi quy tuyến tính

Hồi quy

CMCU

Hệ số Tương quan

Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

Tính toán với R

Rất nhiều đại lượng quan sát được trong cuộc sống có **quan hệ** với nhau. Toán học mô hình hóa các mối quan hệ đó bằng một dạng gọi là hồi quy (regression). Các mô hình hồi quy giúp chúng ta

- hiểu các mối quan hệ (nhân, quả, tác động qua lại, etc.) giữa các biến.
- dự đoán các biến chưa biết dựa trên các quan sát đã có.
- xác định biến nào có ảnh hưởng rõ rệt tới biến đang quan tâm.

Tại sao nên nghiên cứu hồi quy tuyến tính

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Rất nhiều đại lượng quan sát được trong cuộc sống có **quan hệ** với nhau. Toán học mô hình hóa các mối quan hệ đó bằng một dạng gọi là hồi quy (regression). Các mô hình hồi quy giúp chúng ta

- hiểu các mối quan hệ (nhân, quả, tác động qua lại, etc.) giữa các biến.
- dự đoán các biến chưa biết dựa trên các quan sát đã có.
- xác định biến nào có ảnh hưởng rõ rệt tới biến đang quan tâm.

Mô hình hồi quy đơn giản nhất là Hồi quy tuyến tính đơn.

Mô hình

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

$$Y = \beta_0 + \beta_1 X + \varepsilon,$$

- X được gọi là biến dự đoán (predictor), Y là một biến phụ thuộc hay biến hồi đáp (response).
- Sai số ngẫu nhiên ε được giả sử là có phân bố chuẩn $\mathcal{N}(0, \sigma^2)$.
- Một mẫu ngẫu nhiên $\mathfrak{M} = \{(x_1, y_1), \dots, (x_n, y_n)\}$ gồm n quan sát từ cặp biến (X, Y) , thỏa mãn

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

với $\varepsilon_1, \dots, \varepsilon_n$ là các biến ngẫu nhiên độc lập có cùng phân bố chuẩn $\mathcal{N}(0, \sigma^2)$.

Hồi quy Tuyến tính

Hồi quy

Mô hình hồi quy tuyến tính giả thiết rằng kỳ vọng có điều kiện

CMCU

$$E[Y | X = x] = G(x) = \beta_0 + \beta_1 x$$

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

là một *hàm tuyến tính* theo x .

- β_0 được gọi là **hệ số chặn** (intercept) của đường thẳng hồi quy.
- β_1 được gọi là **độ dốc** (slope) của đường thẳng hồi quy.

Hồi quy Tuyến tính

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Mô hình hồi quy tuyến tính giả thiết rằng kỳ vọng có điều kiện

$$E[Y | X = x] = G(x) = \beta_0 + \beta_1 x$$

là một *hàm tuyến tính* theo x .

- β_0 được gọi là **hệ số chặn** (intercept) của đường thẳng hồi quy.
- β_1 được gọi là **độ dốc** (slope) của đường thẳng hồi quy.

Mục tiêu I

Chúng ta muốn tìm một ước lượng $\hat{\beta}_0, \hat{\beta}_1$ của β_0, β_1 sao cho đường thẳng

$$\hat{G}(x) = \hat{\beta}_0 + \hat{\beta}_1 x$$

gần các cặp điểm quan sát nhất, *i.e.*, tổng bình phương khoảng cách tới các điểm (x_i, y_i) 's là nhỏ nhất.

Đường thẳng hồi quy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

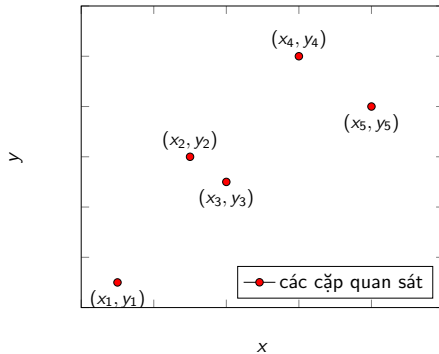
Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R



Đường thẳng hồi quy

Hồi quy

CMCU

Hệ số Tương quan

Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

Hồi quy Tuyến tính bội

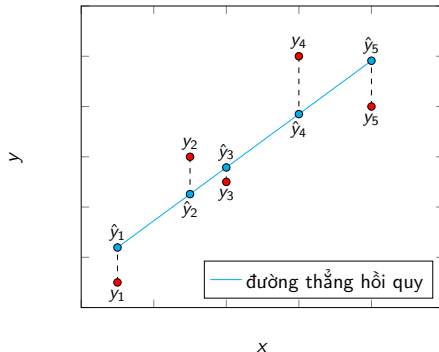
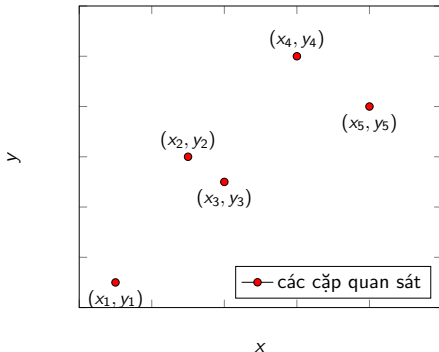
Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

Tính toán với

R



Xấp xỉ bình phương nhỏ nhất của đường thẳng hồi quy.

$$\sum_{i=1}^5 (y_i - \hat{y}_i)^2 \text{ nhỏ nhất.}$$

Phương pháp Bình phương nhỏ nhất - Least-squared Method

Hồi quy

CMCU

Hệ số Tương quan

Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

Tính toán với R



Hình 4: C.F. Gauß (1777-1855).



Hình 5: A.M. Legendre (?) (1752-1833).

Phương pháp Bình phương nhỏ nhất - Least-squared Method

Hồi quy

CMCU

Hệ số Tương quan

Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

Tính toán với R



Hình 4: C.F. Gauß (1777-1855).



Hình 5: A.M. Legendre (?) (1752-1833).

- Lần đầu được công bố bởi Legendre.¹

¹A.M. Legendre. “Nouvelles méthodes pour la détermination des orbites des comètes” (1805).

Phương pháp Bình phương nhỏ nhất - Least-squared Method

Hồi quy

CMCU

Hệ số Tương quan

Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

Tính toán với R



Hình 4: C.F. Gauß (1777-1855).



Hình 5: A.M. Legendre (?) (1752-1833).

- Lần đầu được công bố bởi Legendre.¹
- Trước đó, năm 1801, Gauß đã sử dụng phương pháp này để dự đoán vị trí của Ceres, một tiểu hành tinh.

¹A.M. Legendre. “Nouvelles méthodes pour la détermination des orbites des comètes” (1805).

Phương pháp Bình phương nhỏ nhất - Least-squared Method

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Bài toán

Cho n điểm $(x_1, y_1), \dots, (x_n, y_n)$, tìm β_0, β_1 sao cho

$$L(\beta_0, \beta_1) = (y_1 - \beta_0 - \beta_1 x_1)^2 + \dots + (y_n - \beta_0 - \beta_1 x_n)^2$$

nhỏ nhất.

Phương pháp Bình phương nhỏ nhất - Least-squared Method

Hồi quy

CMCU

Hệ số Tương quan

Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

Tính toán với R

Bài toán

Cho n điểm $(x_1, y_1), \dots, (x_n, y_n)$, tìm β_0, β_1 sao cho

$$L(\beta_0, \beta_1) = (y_1 - \beta_0 - \beta_1 x_1)^2 + \dots + (y_n - \beta_0 - \beta_1 x_n)^2$$

nhỏ nhất.

Gọi $(\hat{\beta}_0, \hat{\beta}_1)$ là một cặp số mà $L(\hat{\beta}_0, \hat{\beta}_1)$ đạt giá trị nhỏ nhất. Khi đó, $\hat{\beta}_0, \hat{\beta}_1$ thỏa mãn

$$\begin{cases} \frac{\partial L}{\partial \beta_0}(\hat{\beta}_0, \hat{\beta}_1) = 0, \\ \frac{\partial L}{\partial \beta_1}(\hat{\beta}_0, \hat{\beta}_1) = 0. \end{cases}$$

Giải hệ Phương trình

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hệ phương trình trên tương đương với

$$\frac{\partial L}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0, \quad (1)$$

$$\frac{\partial L}{\partial \beta_1} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0. \quad (2)$$

Các Phương trình 1 và 2 được gọi là các phương trình chuẩn (normal equations).

Thực nghiệm của các Phương trình Chuẩn

Hồi quy

Từ Phương trình 1, ta được

CMCU

$$\beta_0 = \frac{(y_1 + \cdots + y_n) - \beta_1(x_1 + \cdots + x_n)}{n} = \bar{y} - \beta_1\bar{x}.$$

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Nghiem của các Phương trình Chuẩn

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Từ Phương trình 1, ta được

$$\beta_0 = \frac{(y_1 + \cdots + y_n) - \beta_1(x_1 + \cdots + x_n)}{n} = \bar{y} - \beta_1\bar{x}.$$

Thay vào Phương trình 2, ta được

$$\begin{aligned}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i &= \sum_{i=1}^n x_i (y_i - \bar{y}) - \beta_1 \sum_{i=1}^n x_i (x_i - \bar{x}) \\ &= \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) - \beta_1 \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x}) \\ &= S_{xy} - \beta_1 S_{xx}.\end{aligned}$$

Nghiệm của các Phương trình Chuẩn

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Từ Phương trình 1, ta được

$$\beta_0 = \frac{(y_1 + \cdots + y_n) - \beta_1(x_1 + \cdots + x_n)}{n} = \bar{y} - \beta_1\bar{x}.$$

Thay vào Phương trình 2, ta được

$$\begin{aligned}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i &= \sum_{i=1}^n x_i (y_i - \bar{y}) - \beta_1 \sum_{i=1}^n x_i (x_i - \bar{x}) \\ &= \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) - \beta_1 \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x}) \\ &= S_{xy} - \beta_1 S_{xx}.\end{aligned}$$

Chú ý

$$\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n (y_i - \bar{y}) = 0.$$

Các hệ số của Đường thẳng Hồi quy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ta có $\partial L / \partial \beta_1 = S_{xy} - \beta_1 S_{xx} = 0$ khi và chỉ khi $\beta_1 = \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$.

Các hệ số của Đường thẳng Hồi quy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ta có $\partial L / \partial \beta_1 = S_{xy} - \beta_1 S_{xx} = 0$ khi và chỉ khi $\beta_1 = \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$.

Đường thẳng hồi quy là

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

với

$$\begin{cases} \hat{\beta}_1 &= \frac{S_{xy}}{S_{xx}}, \\ \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x}. \end{cases}$$

Các hệ số của Đường thẳng Hồi quy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ta có $\partial L / \partial \beta_1 = S_{xy} - \beta_1 S_{xx} = 0$ khi và chỉ khi $\beta_1 = \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$.

Đường thẳng hồi quy là

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

với

$$\begin{cases} \hat{\beta}_1 &= \frac{S_{xy}}{S_{xx}}, \\ \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x}. \end{cases}$$

Ý nghĩa của đường thẳng hồi quy:

Trong trường hợp $\hat{\beta}_1 \neq 0$, giá trị trung bình của Y sẽ tăng thêm $\hat{\beta}_1$ đơn vị nếu X tăng thêm 1 đơn vị.

Ví dụ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Trong Ví dụ I, $\bar{x} = 1149/7$, $\bar{y} = 1186/7$, $S_{xy} = 1544/7 = 220.571$, $S_{xx} = 2316/7 = 330.857$,

$$\hat{G}(x) = \frac{2}{3}x + 60.$$

- Trong Ví dụ II, $\bar{x} = 86/3$, $\bar{y} = 22/3$, $S_{xy} = -586/3$, $S_{xx} = 6958/3$,

$$\hat{G}(x) = -0.084x + 9.748.$$

- Trong Ví dụ III, $\bar{x} = 7.3$, $\bar{y} = 7$, $S_{xy} = 5$, $S_{xx} = 38.1$,

$$\hat{G}(x) = 0.131x + 6.042.$$

Phân tích Phương sai

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Phân tích phương sai (Analysis of Variance - ANOVA) hay biến thiên giữa các giá trị y_i 's. Một phần của biến thiên này gây ra bởi các giá trị x_i 's, phần còn lại gây ra bởi các “lỗi” ε_i 's.

Toàn bộ độ biến thiên trong các giá trị quan sát được của y_i 's được đo bởi tổng

$$SS_{\text{TOT}} = \sum_{i=1}^n (y_i - \bar{y})^2 = S_{yy} = (n - 1)s_y^2.$$

ANOVA

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Một phần của tổng trên được đóng góp bởi biến X và mô hình hồi quy được dùng. Phần biến thiên này được đo bởi tổng

$$SS_{\text{REG}} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2.$$

Đây được gọi là phần biến thiên **được giải thích bởi mô hình hồi quy**. Ta có

$$\begin{aligned} SS_{\text{REG}} &= \sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 x_i - \bar{y})^2 \\ &= \sum_{i=1}^n (\bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x_i - \bar{y})^2 \\ &= \hat{\beta}_1^2 S_{xx} = (n-1) \hat{\beta}_1^2 s_x^2. \end{aligned}$$

ANOVA (tiếp)

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Phần còn lại của tổng biến thiên gây ra bởi “lỗi,” được đo bởi tổng sau

$$SS_{\text{ERR}} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n \varepsilon_i^2.$$

SS_{ERR} được gọi là *sai số tiêu chuẩn* của đường thẳng hồi quy. Đây là phần biến thiên *không được giải thích bởi mô hình hồi quy*. Ta có,

$$SS_{\text{TOT}} = SS_{\text{REG}} + SS_{\text{ERR}}.$$

Chú ý

Sai số tiêu chuẩn SS_{ERR} được dùng để ước lượng cho σ^2 .

Hệ số xác định

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Để đo độ khớp (goodness of fit) của mô hình hồi quy, **hệ số xác định** (coefficient of determination hoặc *R*-bình phương) được sử dụng.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{SS_{\text{REG}}}{SS_{\text{TOT}}}.$$

Hệ số xác định

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Để đo độ khớp (goodness of fit) của mô hình hồi quy, **hệ số xác định** (coefficient of determination hoặc R -bình phương) được sử dụng.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{SS_{\text{REG}}}{SS_{\text{TOT}}}.$$

- R^2 đo tỉ lệ phần biến thiên trong dữ liệu được giải thích bởi mô hình hồi quy.
- R^2 càng lớn, mô hình càng giải thích tốt dữ liệu.

Hệ số xác định

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Để đo độ khớp (goodness of fit) của mô hình hồi quy, **hệ số xác định** (coefficient of determination hoặc *R*-bình phương) được sử dụng.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{SS_{\text{REG}}}{SS_{\text{TOT}}}.$$

- R^2 đo tỉ lệ phần biến thiên trong dữ liệu được giải thích bởi mô hình hồi quy.
- R^2 càng lớn, mô hình càng giải thích tốt dữ liệu.

Mệnh đề 2

Trong mô hình hồi quy tuyến tính đơn, ta có

$$R^2 = r^2(X, Y).$$

Đường thẳng hồi quy I

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

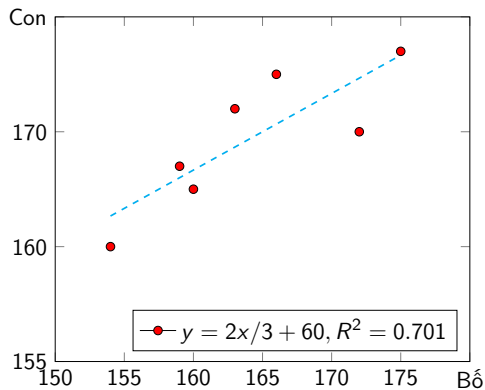
Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R



Hình 6: Chiều cao của bố và con.

Đường thẳng hồi quy II

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

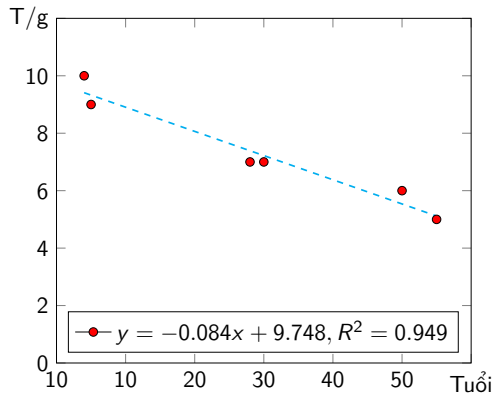
Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R



Hình 7: Tuổi và Thời gian ngủ.

Đường thẳng hồi quy III

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

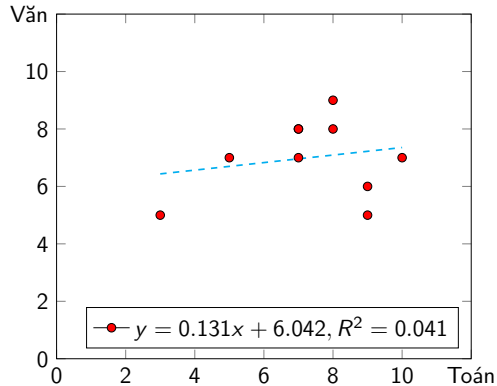
Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R



Hình 8: Điểm Toán và Điểm Văn.

Ví dụ

Hồi quy

CMCU

Hệ số Tương quan

Hồi quy Tuyến tính đơn

Mô hình hồi quy tuyến tính đơn

Phân tích Phương sai - ANOVA

Khoảng tin cậy

Dự đoán các Quan sát mới

Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát mới

Tính toán với R

Theo Cơ sở Dữ liệu Quốc tế (International Data Base), dân số thế giới tăng qua các năm được ghi lại ở bảng sau (đơn vị: nghìn người)

Năm	Dân số	Năm	Dân số	Năm	Dân số
1950	2558	1975	4089	2000	6090
1955	2782	1980	4451	2005	6474
1960	3043	1985	4855	2010	6864
1965	3350	1990	5287	2015	?
1970	3712	1995	5700	2020	?

Bảng 1: Dân số thế giới qua các năm.

Đường thẳng Hồi quy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

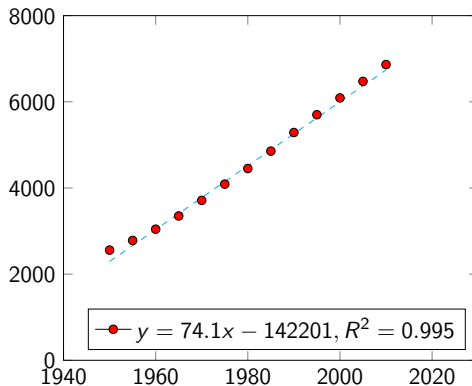
Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R



Đường thẳng Hồi quy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

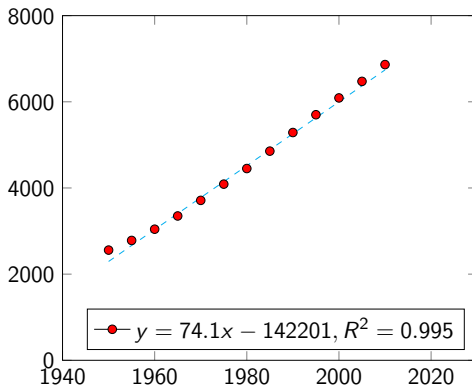
Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R



Dự đoán dân số thế giới các năm 2015, 2020, và 2024 sẽ là 7.11, 7.48 và 7.777 tỉ người (dân số thế giới tính đến 11/2024 là 8.145 tỉ người).

Áp dụng

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Dân số của Việt Nam qua các năm được cho trong bảng sau đây (đơn vị: người) (Nguồn: <https://danso.org/viet-nam/>).

Năm	Dân số	Năm	Dân số	Năm	Dân số
1955	28147443	1975	48718189	1995	74910461
1960	32670039	1980	54281846	2000	79910412
1965	37858951	1985	60896721	2005	83832661
1970	43404793	1990	67988862	2010	87967651

Bảng 2: Dân số Việt Nam qua các năm.

1. Tìm R^2 .
2. Lập đường thẳng hồi quy.
3. Dự đoán dân số Việt Nam các năm 2025, 2030, 2035.

Áp dụng (tiếp)

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Mức lương cơ bản (hay mức lương tối thiểu vùng) tại Hà Nội được cho bởi bảng sau (đơn vị: nghìn đồng) (Nguồn: <https://luatvietnam.vn/tin-phap-luat/tong-hop-muc-luong-toi-thieu-vung-qua-cac-nam-230-17768-article.html>)

Năm	2009	2010	2011	2012	2013	2014	2015	2016
Lương	800	980	1350	2000	2350	2700	3100	3500
Năm	2017	2018	2019	2020	2021	2022	2023	2024
Lương	3720	3980	4180	4420	4420	4680	4680	4960

Bảng 3: Mức lương cơ bản tại Hà Nội qua các năm.

- Tìm hệ số R^2 .
- Lập đường thẳng hồi quy.
- Dự đoán mức lương cơ bản tại Hà Nội các năm 2025, 2028.

Khoảng tin cậy cho σ^2

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ta có một mẫu gồm n quan sát $\varepsilon_i = y_i - \hat{y}_i$, $i = 1, \dots, n$. Các ε_i 's có phân bố chuẩn $\mathcal{N}(0, \sigma^2)$.

Khoảng ước lượng

Khoảng ước lượng cho σ^2 với độ tin cậy $1 - \alpha$ là

$$\left(\frac{(n-2)s^2}{c_{\alpha/2, n-2}^2}, \frac{(n-2)s^2}{c_{1-\alpha/2, n-2}^2} \right),$$

với $s^2 = \frac{SS_{ERR}}{n-2}$.

Khoảng tin cậy cho β_1

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Khoảng ước lượng

Khoảng ước lượng cho β_1 với độ tin cậy $1 - \alpha$ là

$$\left(\hat{\beta}_1 - t_{\alpha/2, n-2} \frac{s}{\sqrt{S_{xx}}}, \hat{\beta}_1 + t_{\alpha/2, n-2} \frac{s}{\sqrt{S_{xx}}} \right),$$

với $s^2 = \frac{SS_{\text{ERR}}}{n-2}$.

Khoảng tin cậy cho β_0

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Khoảng ước lượng cho hệ số β_0 được cho bởi công thức sau.

Khoảng ước lượng cho β_0

Khoảng ước lượng cho β_0 với độ tin cậy $1 - \alpha$ là

$$\left(\hat{\beta}_0 - t_{\alpha/2, n-2} s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}, \hat{\beta}_0 + t_{\alpha/2, n-2} s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}} \right).$$

Dự đoán Quan sát mới

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Với x^* là một quan sát mới của X , ta muốn dự đoán về giá trị của Y hoặc ít nhất là tìm một khoảng tin cậy cho Y . Dựa trên mô hình hồi quy

$$Y = \beta_0 + \beta_1 X + \varepsilon,$$

chúng ta có thể lấy $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x^*$ là dự đoán cho Y .

Khoảng dự đoán

Khoảng dự đoán với độ tin cậy $1 - \alpha$ cho Y là

$$\left(\hat{y} - t_{\alpha/2, n-2} s \sqrt{A(n, x^*, \bar{x})}, \hat{y} + t_{\alpha/2, n-2} s \sqrt{A(n, x^*, \bar{x})} \right),$$

với

$$A(n, x^*, \bar{x}) = 1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}.$$

Ví dụ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Trong ví dụ về dân số Thế giới, ta có $n = 13$, $\bar{x} = 1980$, $S_{xx} = 4550$ (tính từ 1950 đến 2010).

$$s = \sqrt{\frac{SS_{\text{ERR}}}{n - 2}} = 112.8.$$

Với $\alpha = 0.05$, ta có $t_{0.025,11} = 2.201$.

Ví dụ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Trong ví dụ về dân số Thế giới, ta có $n = 13$, $\bar{x} = 1980$, $S_{xx} = 4550$ (tính từ 1950 đến 2010).

$$s = \sqrt{\frac{SS_{ERR}}{n - 2}} = 112.8.$$

Với $\alpha = 0.05$, ta có $t_{0.025,11} = 2.201$.

- Với $x^* = 2015$, $\sqrt{A(13, 2015, 1980)} = 1.16$, khoảng dự đoán là

$$7110.5 \pm 288.1 = (6822.4, 7398.5) \text{ (vs. 7404).}$$

- Với $x^* = 2020$, $\sqrt{A(13, 2020, 1980)} = 1.195$, khoảng dự đoán là

$$7481 \pm 296.8 = (7184.2, 7777.8) \text{ (vs. 7821).}$$

- Với $x^* = 2024$, $\sqrt{A(13, 2024, 1980)} = 1.226$, khoảng dự đoán là

$$7777.4 \pm 304.34 = (7473.06, 8081.74).$$

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hồi quy Tuyến tính bội

Mô hình Hồi quy Bội

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Mô hình hồi quy bội với k biến độc lập

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon_k,$$

trong đó

- X_1, \dots, X_k là các biến ngẫu nhiên độc lập.
- Y là một biến phụ thuộc.
- β_0, \dots, β_k là các hệ số hồi quy.
- ε là sai số của mô hình hồi quy.

Xấp xỉ các Hệ số

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Xấp xỉ các hệ số bằng Phương pháp Bình phương nhỏ nhất.

Giả sử rằng chúng ta có $n > k$ quan sát, và x_{ij} là quan sát thứ i của biến x_j .

Các quan sát là

$$(x_{i1}, \dots, x_{ik}, y_i), \quad i = 1, \dots, n.$$

Mỗi quan sát trên thỏa mãn mô hình hồi quy bội, *i.e.*,

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i \\ &= \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i. \end{aligned}$$

Xấp xỉ (tiếp)

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hàm bình phương nhỏ nhất là

$$\begin{aligned} L &= \varepsilon_1^2 + \cdots + \varepsilon_n^2 \\ &= (y_1 - \beta_0 - \sum_{j=1}^k \beta_j x_{1j})^2 + \cdots + (y_n - \beta_0 - \sum_{j=1}^k \beta_j x_{nj})^2. \end{aligned}$$

Xấp xỉ (tiếp)

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hàm bình phương nhỏ nhất là

$$\begin{aligned} L &= \varepsilon_1^2 + \cdots + \varepsilon_n^2 \\ &= (y_1 - \beta_0 - \sum_{j=1}^k \beta_j x_{1j})^2 + \cdots + (y_n - \beta_0 - \sum_{j=1}^k \beta_j x_{nj})^2. \end{aligned}$$

Chúng ta muốn tìm các giá trị $\hat{\beta}_0, \dots, \hat{\beta}_k$ cho các ẩn số β_0, \dots, β_k sao cho $L(\hat{\beta}_0, \dots, \hat{\beta}_k)$ có giá trị nhỏ nhất. Tương tự như cho trường hợp hồi quy đơn, bộ $(\hat{\beta}_0, \dots, \hat{\beta}_k)$ phải thỏa mãn

$$\frac{\partial L}{\partial \beta_i}(\hat{\beta}_0, \dots, \hat{\beta}_k) = 0, \quad i = 0, \dots, k. \quad (3)$$

Hệ phương trình tuyến tính

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Phương trình 3 tương đương với hệ sau

$$\hat{\beta}_0 n + \hat{\beta}_1 \sum_{i=1}^n x_{i1} + \hat{\beta}_2 \sum_{i=1}^n x_{i2} \cdots + \hat{\beta}_k \sum_{i=1}^n x_{ik} = \sum_{i=1}^n y_i,$$

$$\hat{\beta}_0 \sum_{i=1}^n x_{i1} + \hat{\beta}_1 \sum_{i=1}^n x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^n x_{i1}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^n x_{i1}x_{ik} = \sum_{i=1}^n y_i x_{i1},$$

.....

$$\hat{\beta}_0 \sum_{i=1}^n x_{ik} + \hat{\beta}_1 \sum_{i=1}^n x_{ik}x_{i1} + \hat{\beta}_2 \sum_{i=1}^n x_{ik}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^n x_{ik}^2 = \sum_{i=1}^n y_i x_{ik}.$$

Hệ phương trình trên có $k + 1$ ẩn, $k + 1$ phương trình và có thể được giải bằng phương pháp khử Gauss.

Mô hình hóa Hồi quy Bội bằng Ma trận

Hồi quy

Giả sử ta có k biến dự đoán và n quan sát

CMCU

$$(x_{i1}, \dots, x_{ik}, y_i), \quad i = 1, \dots, n,$$

Hệ số Tương
quan

với mô hình hồi quy cho bởi

Hồi quy Tuyến
tính đơn

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1, \dots, n.$$

Mô hình hồi quy
tuyến tính đơn

Mô hình này được biểu diễn dưới dạng ma trận như sau

Phân tích Phương sai
- ANOVA

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

Khoảng tin cậy

Dự đoán các Quan
sát mới

trong đó

Hồi quy Tuyến
tính bội

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với

R

Tìm $\hat{\beta}$

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Chúng ta muốn tìm một véc-tơ $\hat{\beta}$ sao cho

$$L = \sum_{i=1}^n \varepsilon_i^2 = \varepsilon^T \varepsilon = (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)$$

đạt giá trị nhỏ nhất. Khi đó, $\hat{\beta}$ là nghiệm của phương trình

$$\frac{\partial L}{\partial \beta} = 0.$$

Đạo hàm theo β , ta được

$$\mathbf{X}^T \mathbf{y} - \mathbf{X}^T \mathbf{X} \beta - \beta^T \mathbf{X}^T \mathbf{X} + \mathbf{y}^T \mathbf{X} = 0 \Leftrightarrow \mathbf{X}^T \mathbf{y} = \mathbf{X}^T \mathbf{X} \beta.$$

Từ đẳng thức trên, ta thu được

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}.$$

Xấp xỉ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Các phương trình

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \cdots + \hat{\beta}_k x_{ik}, \quad i = 1, \dots, n$$

được viết lại dưới dạng

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}.$$

Ước lượng cho σ^2

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Ước lượng điểm cho phương sai σ^2 được cho bởi công thức

$$s^2 = \frac{1}{n - k - 1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{SS_{\text{ERR}}}{n - k - 1}.$$

Khoảng tin cậy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Khoảng ước lượng

Khoảng ước lượng với độ tin cậy $1 - \alpha$ cho các hệ số β_i 's được cho bởi

$$\left(\hat{\beta}_i - t_{\alpha/2, n-k-1} s \sqrt{C_i}, \hat{\beta}_i + t_{\alpha/2, n-k-1} s \sqrt{C_i} \right),$$

với C_i là hệ số ở vị trí (i, i) của ma trận $(\mathbf{X}^T \mathbf{X})^{-1}$.

Dự đoán

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Mô hình hồi quy bội được dùng để dự đoán về một quan sát mới của biến Y dựa trên các kết quả quan sát x_1^*, \dots, x_k^* của các biến độc lập X_1, \dots, X_k . Nếu $\mathbf{x}_* = (x_1^*, \dots, x_k^*)^T$, thì một ước lượng điểm của Y cho lần quan sát tiếp theo (dự đoán) tại điểm \mathbf{x}_* là

$$\hat{y}_* = \mathbf{x}_*^T \hat{\beta}.$$

Khoảng dự đoán

Khoảng dự đoán với độ tin cậy $1 - \alpha$ cho lần quan sát trong tương lai là

$$\left(\hat{y}_* - t_{\alpha/2, n-k-1} s \sqrt{A(\mathbf{x}_*, \mathbf{X})}, \hat{y}_* + t_{\alpha/2, n-k-1} s \sqrt{A(\mathbf{x}_*, \mathbf{X})} \right),$$

với $A(\mathbf{x}_*, \mathbf{X}) = 1 + \mathbf{x}_*^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_*$.

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Tính toán với R

Biểu diễn điểm

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Để biểu diễn các điểm $(x_1, y_1), \dots, (x_n, y_n)$ trên mặt phẳng tọa độ, dùng cú pháp `c()` và `plot()`.

Biểu diễn điểm

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Để biểu diễn các điểm $(x_1, y_1), \dots, (x_n, y_n)$ trên mặt phẳng tọa độ, dùng cú pháp `c()` và `plot()`.

```
> x = c(1,2,3,4,5,6)
> y = c(1,4,9,16,25,36)
> plot(x,y)
```

Các hệ số

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Tìm hệ số R^2

```
> model = lm(y~x)
> r = summary(model)$r.squared
> r
```

Các hệ số

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Tìm hệ số R^2

```
> model = lm(y~x)
> r = summary(model)$r.squared
> r
```

- Các hệ số của đường thẳng hồi quy

```
> model = lm(y~x)
> model
```


Vẽ đường thẳng hồi quy

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Để vẽ đường thẳng hồi quy, ta dùng cú pháp

```
> abline(lm(y~x))
```

Áp dụng

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Giá phòng trọ

Giá cho thuê một phòng $10 - 15m^2$ tại Hà Nội qua các năm được ghi lại ở bảng sau (đơn vị: triệu đồng/tháng)

Giá	0.7–1	1.2–1.5	1.5–2	1.8–2.4	2–2.6	1.2–2	1.6–2.5	2–3
Năm	2010	2012	2014	2016	2018	2020	2022	2024

- Vẽ đám mây điểm (năm, giá trọ).
- Tính hệ số tương quan.
- Vẽ đường thẳng hồi quy.
- Xác định các hệ số của đường thẳng hồi quy, hệ số xác định.

Ma trận

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

Hai cách để nhập một ma trận:

- Nhập theo véc-tơ hàng. Giả sử các véc-tơ hàng của ma trận A là r_1, \dots, r_m . Sau khi nhập các dòng này, dùng cú pháp

```
> A = rbind(r1,r2,...,rm)
```

- Nhập theo véc-tơ cột. Giả sử các véc-tơ cột của ma trận A là c_1, \dots, c_n . Sau khi nhập các véc-tơ này, dùng cú pháp

```
> A = cbind(c1,c2,...,cn)
```

Một số phép toán với ma trận

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Phép cộng/trừ hai ma trận cùng cỡ A và B

> $A + B$

> $A - B$

- Phép nhân ma trận A với vô hướng a

> $A * a$

- Phép nhân ma trận A với ma trận B

> $A \%*\% B$

- Định thức của một ma trận vuông

> $\det(A)$

- Ma trận chuyển vị của ma trận A được tìm bởi cú pháp

> $t(A)$

Phép toán với ma trận (tiếp) & véc-tơ

Hồi quy

CMCU

Hệ số Tương
quan

Hồi quy Tuyến
tính đơn

Mô hình hồi quy
tuyến tính đơn

Phân tích Phương sai
- ANOVA

Khoảng tin cậy

Dự đoán các Quan
sát mới

Hồi quy Tuyến
tính bội

Mô hình Hồi quy Bội

Khoảng tin cậy

Dự đoán Quan sát
mới

Tính toán với
R

- Để tìm nghịch đảo của ma trận (vuông) A , ta dùng thư viện `matlib` và cú pháp sau

```
> library(matlib)
> inv(A)
```

- Tích vô hướng của hai véc-tơ x, y có thể được tìm bởi câu lệnh sau

```
> sum(x * y)
```

Tích vô hướng của hai véc-tơ

Tích vô hướng của hai véc-tơ $\mathbf{x} = (x_1, \dots, x_n)$ và $\mathbf{y} = (y_1, \dots, y_n)$ là giá trị

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \dots + x_n y_n.$$

End of Chapter 6