

Chương 4

PHÂN TÍCH MỐI TƯƠNG QUAN

NỘI DUNG

Các loại quan hệ

Quan hệ tuyến tính

Các dạng quan hệ tuyến tính

**Mô hình tuyến tính đơn các đặc trưng
định lượng**

CÁC LOẠI QUAN HỆ

Một thí nghiệm khảo sát ảnh hưởng của đạm lên năng suất lúa có 4 nghiệm thức và 3 lần lặp lại như sau:

NT	I	II	III	Trung bình
0	4253	4467	3970	4230
50	5360	5437	5529	5442
100	6580	6651	6752	6661
150	7065	7170	7215	7150

Bảng 4.2: Kết quả theo dõi 7 chỉ tiêu trên 36 giống lúa

Giống	Cao cây (cm)	TS nhánh	Nhánh HH	Dài bông (cm)	Số hạt /bông	M100 hạt (g)	NSTT (g/bụi)	NS hồi quy (g/bụi)
1	109,5	46,1	41,9	20,3	52,5	3,9	85,7	86,2
2	119,1	36,2	33,4	20,0	63,2	4,0	83,9	85,4
3	114,4	47,6	44,6	21,7	48,6	4,4	96,4	94,3
4	119,7	57,2	55,2	20,0	52,0	4,1	117,0	114,1
5	117,0	44,0	42,1	21,2	64,1	4,3	115,8	108,4
6	121,5	35,9	34,0	22,1	47,6	4,3	68,7	69,8
7	120,9	53,1	50,4	22,1	57,1	4,8	108,5	122,4
8	119,5	33,3	30,8	20,0	40,9	4,0	50,7	50,7
9	126,8	46,7	43,8	19,8	61,1	3,7	98,3	98,2
10	111,7	43,7	42,3	20,2	77,4	3,3	108,1	111,8
11	120,1	41,5	40,5	21,7	51,1	4,3	88,7	87,8
12	121,2	52,5	50,7	22,0	51,2	3,9	101,4	101,5
13	127,8	30,6	28,3	22,5	66,8	3,8	72,3	78,2
14	127,7	42,9	40,1	21,9	52,6	4,4	92,7	90,7
15	131,8	49,7	47,5	20,1	41,5	4,3	84,4	88,4
16	111,8	42,1	37,3	18,7	43,2	4,6	74,1	75,6
17	117,5	40,2	38,1	19,4	56,7	4,7	101,7	97,1
18	109,8	47,6	45,9	20,1	45,5	4,4	91,3	92,1
19	122,4	57,0	54,8	19,0	43,5	4,2	100,2	103,7
20	110,2	50,0	47,4	23,1	52,8	4,4	110,0	105,2
21	129,3	55,7	52,8	20,3	52,8	4,2	116,6	112,4
22	126,1	45,4	41,2	20,3	52,2	4,3	91,6	90,5

1. Tương quan và hồi quy tuyến tính (đường thẳng), gồm:

- Tương quan và hồi quy và tuyến tính một biến

$$y = f(x) = a + bx$$

- Tương quan và hồi quy và tuyến tính đa biến

$$(y = f(x_i) = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n)$$

2. Tương quan và hồi quy phi tuyến tính (đường cong), gồm:

- Tương quan và hồi quy và phi tuyến tính một biến
- Tương quan và hồi quy và phi tuyến tính đa biến.

QUAN HỆ TUYẾN TÍNH

Các dạng quan hệ tuyến tính

Phương trình biểu thị mối quan hệ tuyến tính một biến giữa X và Y có dạng:

$$y = f(x) = a + bx$$

Phương trình trên gọi là phương trình hồi quy tuyến tính một biến, trong đó

y là hàm số (số phụ thuộc),

x là đối số (số độc lập);

b là hệ số góc (còn gọi là hệ số hồi quy)

a là hằng số.

Các bước phân tích tương quan tuyến tính đơn (tìm hệ số a và b)

1. Tính trung bình X và Y (biến độc lập và biến phụ thuộc).

$$\bar{X} = \frac{\sum X}{n} \quad \bar{Y} = \frac{\sum Y}{n}$$

2. Tính phương sai X và Y

$$\sum x^2 = \sum_{i=1}^n (X_i - \bar{X})^2$$

$$\sum y^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2$$

3 . Tính tổng tích số của sai lệch X và Y

$$\sum xy = \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})$$

4 . Tính hệ số góc b

$$b = \frac{\sum xy}{\sum x^2}$$

5 . Tính hệ hằng số a

$$a = \bar{Y} - b\bar{X}$$

6 . Tính hệ số tương quan

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

$|r| \simeq 0$: X và Y độc lập hoặc có quan hệ phi tuyến tính

$0 < |r| \leq 0,3$: X và Y có quan hệ yếu

$0,3 < |r| \leq 0,5$: X và Y có quan hệ vừa

$0,5 < |r| \leq 0,7$: X và Y có quan hệ tương đối chặt

$0,7 < |r| \leq 0,9$: X và Y có quan hệ chặt

$0,9 < |r| \leq 1$: X và Y có quan hệ rất chặt

7 . Tính khoảng tin cậy khi biến X thay đổi

$$s_{y \cdot x}^2 = \frac{\sum y^2 - \frac{(\sum xy)^2}{\sum x^2}}{n - 2}$$

$$C.I. = b \pm t_{\alpha} \sqrt{\frac{s_{y \cdot x}^2}{\sum x^2}}$$

t_{α} tra bảng với $df = n - 2$

Đánh giá sự tồn tại của hệ số tương quan

Hệ số tương quan lý thuyết được tính sẵn trong bảng r với độ tự do $n - 2$ ở các mức tin cậy khác nhau.

- Nếu $r > r_{\alpha}$ với $n - 2$ bậc tự do $\Rightarrow r$ tồn tại với độ tin cậy $1 - \alpha$.
- Nếu $r < r_{\alpha}$ với $n - 2$ bậc tự do $\Rightarrow r$ không tồn tại với độ tin cậy $1 - \alpha$.

Ví dụ: Một tương quan đơn thực hiện với với $n = 20$ quan sát \Rightarrow độ tự do là 18,

Nếu $r > 0,444$ thì r tồn tại với độ tin cậy 95%

Nếu $r > 0,561$ thì r tồn tại với độ tin cậy 99%

Nếu $r_{xy} > 0,679$ thì r tồn tại với độ tin cậy 99,9%.