

Лабораторная работа №2

Оценка коэффициентов простой линейной регрессии

и построение доверительных интервалов

с уровнем значимости $\alpha = 0,05$

Пусть имеется множество из " n " точек $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$, которые представляют результаты измерений некоторого эксперимента. Предположим, что функция регрессии для Y от фактора X является линейной функцией, т. е. имеет вид:

$$Y = \beta_0 + \beta_1 \cdot X.$$

С учетом этого предположения требуется найти оценки неизвестных параметров β_0 и β_1 .

Для получения оценок параметров модели β_0 и β_1 используем метод наименьших квадратов

Вариант (методический пример)

X	0	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5
Y	1,35	1,83	-0,10	0,33	-0,24	-0,75	-1,61	-0,75	-4,27	-2,52	-2,60

Формулировка задания и расчетные формулы

Для построения модели необходимо оценить:

а) математическое ожидание X и Y по формулам:

$$\bar{X} = \sum_{i=1}^n X_i / n, \quad \bar{Y} = \sum_{i=1}^n Y_i / n;$$

$$\bar{X} = \frac{27,5}{11} = 2,5,$$

$$\bar{Y} = \frac{-9,33}{11} = -0,8482.$$

б) параметры моделей β_0 и β_1 по формулам:

$$b_1 = \left[n \cdot \sum_{i=1}^n X_i \cdot Y_i - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right) \right] / \left[n \cdot \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 \right],$$

$$b_1 = \frac{11 \cdot (-49,92) - 27,5 \cdot (-9,33)}{11 \cdot 96,25 - 27,5^2} = -0,9671$$

$$b_0 = \left[\sum_{i=1}^n X_i^2 \sum_{i=1}^n Y_i - \sum_{i=1}^n X_i \cdot \sum_{i=1}^n X_i \cdot Y_i \right] / \left[n \cdot \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 \right]$$

$$b_0 = \frac{[96,25 \cdot (-9,33) - 27,5 \cdot (-49,92)]}{[11 \cdot 96,25 - 27,5^2]} = 1,5695$$

в) стандартное отклонение в модели линейной регрессии:

$$s = \left[\sum_{i=1}^n (Y_i - b_0 - b_1 \cdot X_i)^2 / (n - 2) \right]^{1/2}$$

$$s = (6,77495/9)^{1/2} = 0,8676$$

г) коэффициент множественной детерминации R^2 :

$$R^2 = \left[\sum_{i=1}^n (b_0 + b_1 \cdot X_i - \bar{Y})^2 \right] / \sum_{i=1}^n (Y_i - \bar{Y})^2$$

$$R^2 = 25,72027/32,49477 = 0,7915$$

а также выписать уравнения для модели регрессии и для границ доверительных интервалов с уровнем значимости $1-a$, где $a = 0.05$,

$t(n-2, 1-a/2)$ - значение статистики Стьюдента равно $t(9; 0.975) = 2.26$:

– для модели регрессии:

$$Y_0 = b_0 + b_1 \cdot X;$$

X_i	Y_0
0	1,5695
0,5	1,08595
1	0,6024
1,5	0,11885
2	-0,3647
2,5	-0,84825
3	-1,3318
3,5	-1,81535
4	-2,2989
4,5	-2,78245
5	-3,266

– для верхней границы доверительного интервала

$$Y1 = b_0 + b_1 x + t(n-2, 1-a/2) \cdot s \left[1/n + (x - \bar{X})^2 / \sum_{i=1}^n (X_i - \bar{X})^2 \right]^{1/2}$$

X_i	Y_1
0	2,673
0,5	2,037
1	1,416
1,5	0,815
2	-0,171
2,5	-0,260
3	-1,138
3,5	-1,119
4	-1,486
4,5	-1,831
5	-2,162

– для нижней границы доверительного интервала

$$Y2 = b_0 + b_1 x - t(n-2, 1-a/2) \cdot s \left[1/n + (x - \bar{X})^2 / \sum_{i=1}^n (X_i - \bar{X})^2 \right]^{1/2}$$

X_i	Y_2
0	0,466
0,5	0,135
1	-0,211
1,5	-0,577
2	-0,559
2,5	-1,436
3	-1,526
3,5	-2,511
4	-3,112
4,5	-3,733
5	-4,367

е) построить графики линий Y_0 , Y_1 , Y_2 .

Варианты индивидуальных ~~заданий~~ заданий для лабораторной работы №2.

Вариант 1.

X	0	0.2	1	0.3	2	0.5	3	0.7	4	0.9	5
Y	1.4	1.53	2.1	1.1	2.3	3.1	2.1	3.6	4.1	3.8	2.9

Вариант 2.

X	0	3.8	1	4.1	2	3.6	3	1.1	4	2.9	5
Y	1.1	2.3	3.1	2.5	3.3	4.2	1.9	2.5	3.7	4.2	3.9

Вариант 3.

X	0	5.1	1	4.9	2	5.2	3	4.1	4	3.8	5
Y	2.2	3.4	4.2	1.1	2.5	4.8	3.4	2.7	2.9	3.4	4.8

Вариант 4.

X	0	2.1	1	3.4	2	2.5	3	3.1	4	4.6	5
Y	1.1	1.7	2.3	2.2	3.4	4.4	4.2	3.4	2.8	3.7	4.5

Вариант 5.

X	0	3.1	1	3.2	2	2.5	3	4.1	4	4.2	5
Y	2.1	2.8	3.6	4.4	3.9	5.2	4.2	3.7	3.8	5.4	4.3

Вариант 6.

X	0	3	1	2.1	2	3.2	3	3.4	4	4.1	5
Y	1.3	2.4	3.2	4.1	2.5	3.4	2.8	4.1	2.5	3.2	4.4

Вариант 7.

X	0	2.1	1	2.2	2	4.1	3	3.6	4	4.1	5
Y	1.1	2.3	3.3	3.6	3.8	2.5	4.8	6.1	6.4	5.8	6.1

Вариант 8.

X	0	2.1	1	3.1	2	4.1	3	3.3	4	4.2	5
Y	2.2	3.4	2.7	2.9	4.1	4.7	5.2	4.2	5.1	3.4	2.8

Вариант 9.

X	0	3.1	1	2.3	2	4.1	3	1.4	4	5.1	5
Y	1.1	2.1	3.2	1.2	4.3	3.4	4.8	3.8	5.2	3.4	4.7

Вариант 10.

X	0	2.1	1	1.2	2	3.1	3	3.8	4	4.2	5
Y	3.1	1.3	2.5	4.8	5.1	4.8	6.3	3.9	5.7	6.1	3.4

Вариант 11.

X	0	1.1	1	2.3	2	3.2	3	2.8	4	2.9	5
Y	3.3	1.2	4.1	5.2	6.1	5.5	5.8	4.3	3.8	2.1	4.3

Вариант 12.

X	0	2.1	1	2.8	2	3.4	3	4.3	4	5.2	5
Y	3.1	4.1	4.3	4.6	5.1	1.6	6.3	5.4	6.2	3.4	2.8

Вариант 13

X	0	3.1	1	4.2	2	3.5	3	5.3	4	6.1	5
Y	1.1	2.3	4.1	2.4	3.6	4.2	4.8	6.4	5.6	1.8	2.7

Вариант 14

X	0	1.2	1	2.4	2	3.1	3	4.2	4	3.1	5
Y	1.1	3.2	4.3	4.4	5.1	1.5	6.5	4.6	3.5	4.8	2.9

Вариант 15

X	0	3.1	1	2.1	2	4.1	3	3.2	4	4.1	5
Y	1.1	2.1	3.2	2.3	4.1	2.4	5.1	1.6	5.6	3.7	2.8

Вариант 16.

X	0	3.1	1	2.3	2	4.1	3	3.8	4	5.1	5
Y	2.4	5.3	6.1	1.6	5.2	2.8	6.1	4.7	2.8	1.5	4.3

Вариант 17.

X	0	3.1	1	2.2	2	4.1	3	3.2	4	4.3	5
Y	2.1	3.4	4.1	5.1	2.5	3.6	4.7	6.1	5.2	2.8	3.6

Вариант 18.

X	0	4.1	1	3.4	2	5.1	3	6.2	4	3.8	5
Y	1.3	2.4	5.6	4.1	1.1	4.8	2.7	3.4	4.6	5.2	4.1