

Университет ИТМО

Факультет программной инженерии и компьютерной техники

Индивидуальное домашнее задание №6

по «Математической статистике»

Вариант 9

Выполнили:

Студенты группы Р3233

Хасаншин Марат

Шикунов Максим

Номер команд: 9

Санкт-Петербург

2024

Цель работы

Оценить характер зависимости между параметрами двумерной случайной величины (обработка двумерной выборки)

Исходные данные

$x_i^* \backslash y_j^*$	31	41	46	51
26	0	0	45	10
36	20	30	80	0
46	15	0	0	0

Ход выполнения

Высчитаем n_i и n_j :

$x_i^* \backslash y_j^*$	31	41	46	51	n_i
26	0	0	45	10	55
36	20	30	80	0	130
46	15	0	0	0	15
n_j	35	30	125	10	$n=200$

По формулам считаем:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^k n_i x_i^*$$

$$\bar{Y}_n = \frac{1}{n} \sum_{j=1}^m n_j y_j^*$$

$$\sigma_n(X) = \sqrt{\frac{1}{n} \sum_{i=1}^k n_i (x_i^*)^2 - \bar{X}_n^2}$$

$$\sigma_n(Y) = \sqrt{\frac{1}{n} \sum_{j=1}^m n_j (y_j^*)^2 - \bar{Y}_n^2}$$

$$r_n(X, Y) = \frac{\sum_{i=1}^k \sum_{j=1}^m n_{ij} x_i^* y_j^* - n \cdot \bar{X}_n \bar{Y}_n}{n \cdot \sigma_n(X) \cdot \sigma_n(Y)}$$

$$y(x) = \bar{Y}_n + r_n(X, Y) \frac{\sigma_n(Y)}{\sigma_n(X)} (x - \bar{X}_n)$$

Получается:

$$\bar{X}_{200}=34$$

$$\bar{Y}_{200} = 42.875$$

$$\sigma_n(X) = 5.567764$$

$$\sigma_n(Y) = 5.882973$$

$$r_n(X, Y) = -0.61059$$

$$y(x) = 42.875 - 0.61059 \cdot \left(\frac{5.882973}{5.567764}\right)(x - 34)$$

$$y(x) = 64.805 - 0.645x$$

Сравним оценки условных математических ожиданий, вычисленные:

а) по данным таблицы, полагая, как и ранее, $P(y_j^*) = p_j^* = \frac{n_{ij}}{n_i}$

б) на основе последнего уравнения

Например, при $x^* = 36$ имеем:

$$\text{а) } M(Y|X = 36) = \frac{(20 \cdot 31 + 30 \cdot 41 + 80 \cdot 46)}{130} = 42.538$$

$$\text{б) } M(Y|X = 36) = 64.805 - 0.645 \cdot 36 = 41.585$$

Ответы примерно одинаковы

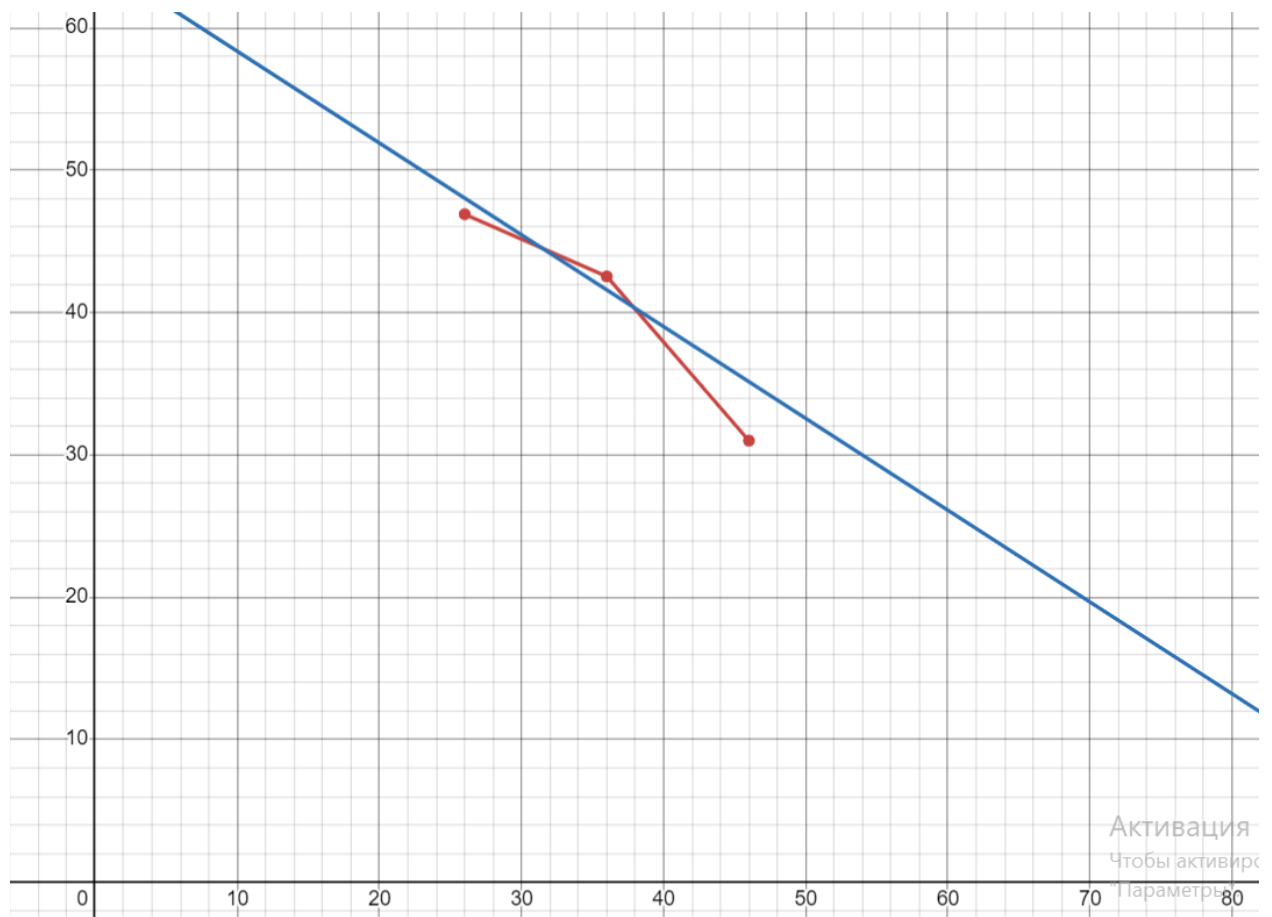
Построим зависимость условного математического ожидания компонента у от значений компонента х и график функции регрессии линейного вида.

а)

x_i	26	36	46
$M(Y X = x_i)$	46.91	42.538	31

б)

$$y(x) = 64.805 - 0.645x$$



Вывод

По нашим данным нормализовали и линейно выровняли характер зависимости двумерной выборки компонента y от компонента x .