

Федеральное государственное автономное образовательное  
учреждение высшего образования

**«Национальный исследовательский университет  
ИТМО»**

Факультет программной инженерии и компьютерной  
техники

## **Практическая работа № 6**

**«Уравнение прямой линии регрессии»**

Вариант №  $(28 \bmod 15) + 1 = 14$

Выполнил:  
Караганов П. Э.  
Группа: Р3210

Проверила:  
Селина Е. Г.

## Задание

Дана таблица распределения  $n = 100$  заводов по производственным средствам  $X$  (тыс. ден. ед.) и по суточной выборке  $Y$  (т). Известно, что между  $X$  и  $Y$  существует линейная корреляционная зависимость. Требуется:

- Найти уравнение прямой регрессии  $y$  на  $x$ .
- Построить уравнение эмпирической линии регрессии и случайные точки выборки  $(X, Y)$ .

Таблица 1: Распределение  $X$  и  $Y$

$X \setminus Y$	14	17	20	23	26	29	32	35	$m_x$
1,8	2	4	6	0	0	0	0	0	12
2,4	0	2	7	6	0	0	0	0	15
3,0	0	0	6	8	5	0	0	0	19
3,6	0	0	0	8	14	4	0	0	26
4,2	0	0	0	0	3	6	8	0	17
4,8	0	0	0	0	0	0	5	6	11
$m_y$	2	6	19	22	22	10	13	6	100

## Решение

Для подсчета числовых характеристик (выборочных средних и выборочных средних  $\bar{x}$  и  $\bar{y}$ , выборочных средних квадратичных отклонений  $s_x$  и  $s_y$  и выборочного корреляционного момента  $s_{xy}$ ) составляем расчетную таблицу (Таблица 2). При заполнении таблицы осуществляем контроль по строкам и столбцам:

$$\sum_{i=1}^6 m_{x_i} = \sum_{j=1}^8 m_{y_j} = n = 100$$

$$\sum_{i=1}^6 m_{x_i} x_i = 332.4 \quad \sum_{j=1}^8 m_{y_j} y_j = 2504$$

$$\sum_{i=1}^6 m_{x_i} x_i^2 = 1186.56 \quad \sum_{j=1}^8 m_{y_j} y_j^2 = 65308$$

$$\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} x_i y_j = 8739.6$$

Вычисляем выборочные средние:

$$\bar{X} = \frac{\sum m_{x_i} x_i}{n} = \frac{332.4}{100} = 3.324 \quad \bar{Y} = \frac{\sum m_{y_j} y_j}{n} = \frac{2504}{100} = 25.04$$

Выборочные дисперсии и корреляционный момент:

$$s_X^2 = \frac{1}{n-1} \left( \sum m_{x_i} x_i^2 - \frac{(\sum m_{x_i} x_i)^2}{n} \right) = \frac{1}{99} \left( 1186.56 - \frac{(332.4)^2}{100} \right) = 0.824873$$

$$s_Y^2 = \frac{1}{n-1} \left( \sum m_{y_i} y_i^2 - \frac{(\sum m_{y_i} y_i)^2}{n} \right) = \frac{1}{99} \left( 65308 - \frac{(2504)^2}{100} \right) = 26.341818$$

Таблица 2: Расчётная таблица для числовых характеристик

i	j	1	2	3	4	5	6	7	8	9	10	11	12	13
	$X \setminus Y$	14	17	20	23	26	29	32	35	$m_{x_i}$	$m_{x_i x_i}$	$\sum m_{ij} y_j$	$x_i^2 m_{x_i}$	$x_i \sum m_{ij} y_j$
1	1.8	2	4	6	0	0	0	0	0	12	21.6	$14 \cdot 2 + 17 \cdot 4 + 20 \cdot 6 = 258$	$1.8^2 \cdot 12 = 38.88$	$1.8 \cdot 258 = 464.4$
2	2.4	0	2	7	6	0	0	0	0	15	36.0	$17 \cdot 2 + 20 \cdot 7 + 23 \cdot 6 = 350$	$2.4^2 \cdot 15 = 86.4$	$2.4 \cdot 350 = 840$
3	3.0	0	0	6	8	5	0	0	0	19	57.0	$20 \cdot 6 + 23 \cdot 8 + 26 \cdot 5 = 454$	$3^2 \cdot 19 = 171$	$3 \cdot 454 = 1362$
4	3.6	0	0	0	8	14	4	0	0	26	93.6	$23 \cdot 8 + 26 \cdot 14 + 29 \cdot 4 = 654$	$3.6^2 \cdot 26 = 337$	$3.6 \cdot 654 = 2354.4$
5	4.2	0	0	0	0	3	6	8	0	17	71.4	$26 \cdot 3 + 29 \cdot 6 + 32 \cdot 8 = 494$	$4.2^2 \cdot 17 = 299.88$	$4.2 \cdot 494 = 2074.8$
6	4.8	0	0	0	0	0	0	5	6	11	52.8	$32 \cdot 5 + 35 \cdot 6 = 310$	$4.8^2 \cdot 11 = 253.44$	$4.8 \cdot 310 = 1488$
7	$m_{y_i}$	2	6	19	22	22	10	13	6	100	332.4	2504	1186.56	8739.6
8	$m_{y_i} y_j$	28	102	380	506	572	290	416	210	2504	—	—	—	—
9	$\sum m_{ij} x_i$	3.6	12.0	38.4	54.0	85.8	42.0	61.0	28.8	332.4	—	—	—	—
10	$y_i^2 m_{ij}$	392	1734	7600	11132	14872	8410	17216	7350	65308	—	—	—	—
11	$y_j \sum m_{ij} x_i$	50.4	204.0	768.0	1242.0	2230.8	1218	1952	1008	8739.6	—	—	—	—

Корреляционный момент вычисляем по формуле

$$s_{XY} = \frac{1}{n-1} \left( \sum m_{ij} x_i y_j - \frac{(\sum m_{xi} x_i)(\sum m_{yj} y_j)}{n} \right)$$

$$= \frac{1}{99} \left( 8739.6 - \frac{332.4 \cdot 2504}{100} \right) = 4.205091.$$

Оценкой теоретической линии регрессии является эмпирическая линия регрессии, уравнение которой имеет вид

$$y = \bar{y} + r_{XY} \frac{s_Y}{s_X} (x - \bar{x}), \text{ где}$$

$$s_X = \sqrt{0.824} \approx 0.908225 \quad s_Y = \sqrt{26.342} \approx 5.132428$$

$$r_{XY} = \frac{s_{XY}}{s_X s_Y} = \frac{4.205}{0.908 \cdot 5.132} \approx 0.902$$

Составляем уравнение эмпирической линии регрессии y на x:

$$y = 25.04 + 0.902 \cdot \frac{5.132428}{0.908225} (x - 3.324),$$

$$y = 25.04 + 5.0979 (x - 3.324),$$

$$\boxed{y = 5.0979 x + 8.0947}$$

Строим линию регрессии и случайные точки  $(x_i; y_j)$

