

## ИД3-19.2 (вариант 9)

Дано:

Дана таблица распределения 100 автомашин по затратам на перевозки X (ден. ед.) и по протяженности маршрутов перевозок Y (км). Известно, что между X и Y существует линейная корреляционная зависимость. Требуется:

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

X\Y	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	m <sub>x</sub>
250	3	4	5	-	-	-	-	-	12
450	-	6	2	8	-	-	-	-	16
650	-	-	-	5	14	9	-	-	28
850	-	-	-	6	8	6	-	-	20
1050	-	-	-	-	5	7	4	-	16
1250	-	-	-	-	-	-	5	3	8
m <sub>y</sub>	3	10	7	19	27	22	9	3	100

Решение:

Для подсчета числовых характеристик (выборочных средних X' и Y', выборочных средних квадратичных отклонений S<sub>x</sub> и S<sub>y</sub> и выборочного корреляционного момента S<sub>xy</sub>) составляем расчетную таблицу. При заполнении таблицы осуществляем контроль по строкам и столбцам:

$$\sum_{i=1}^6 m_{xi} = \sum_{j=1}^8 m_{yj} = n = 100$$

$$\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} x_i = \sum_{i=1}^6 m_{xi} x_i = 72200$$

$$\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} y_j = \sum_{j=1}^8 m_{yj} y_j = 94.8$$

$$\sum_{i=1}^6 \left( x_i \sum_{j=1}^8 m_{ij} y_j \right) = \sum_{j=1}^8 \left( y_j \sum_{i=1}^6 m_{ij} x_i \right) = 76220$$

Вычисляем выборочные средние X' и Y', i = {1, 6}; j = {1, 8}:

$$X' = \frac{\sum \sum m_{ij} x_i}{n} = \frac{\sum m_{xi} x_i}{n} = \frac{72200}{100} = 722$$

$$Y' = \frac{\sum m_{yj} y_j}{n} = \frac{94.8}{100} = 0.948$$

Выборочные дисперсии находим по формулам:

$$S_x^2 = \frac{1}{n-1} \left( \sum m_{xi} x_i^2 - \frac{1}{n} \left( \sum m_{xi} x_i \right)^2 \right) = \frac{1}{99} \left( 60410000 - \frac{1}{100} (72200)^2 \right) = 83652.53$$

$$S_y^2 = \frac{1}{n-1} \left( \sum m_{yj} y_j^2 - \frac{1}{n} \left( \sum m_{yj} y_j \right)^2 \right) = \frac{1}{99} \left( 100.4 - \frac{1}{100} (94.8)^2 \right) = 0.10636$$

	j	1	2	3	4	5	6	7	8	9	10	11	12	13
i	X\Y	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	$m_x$	$m_{xi}$	$\sum_{j=1}^k m_{yj} y_j$	$m_{xi} x_i^2$	$x_i \sum_{j=1}^k m_{ij} y_j$
1	250	3	4	5	-	-	-	-	-	12	3000	5.2	750000	1300
2	450	-	6	2	8	-	-	-	-	16	7200	10.0	3240000	4500
3	650	-	-	-	5	14	9	-	-	28	18200	28.8	11830000	18720
4	850	-	-	-	6	8	6	-	-	20	17000	20.0	14450000	17000
5	1050	-	-	-	-	5	7	4	-	16	16800	19.0	17640000	19950
6	1250	-	-	-	-	-	-	5	3	8	10000	11.8	12500000	14750
7	$m_y$	3	10	7	19	27	22	9	3	100	72200	94.8	60410000	76220
8	$m_y y_j$	0.6	4.0	4.2	15.2	27.0	26.4	12.6	4.8	94,8	-	-	-	-
9	$\sum_{i=1}^m m_{xi} x_i$	750	3700	2150	11950	21150	18300	10450	3750	72200	-	-	-	-
10	$m_{ij} y_j^2$	0.12	1.60	2.52	12.16	27.00	31.68	17.64	7.68	100.4	-	-	-	-
11	$y_j \sum_{i=1}^m m_{ij} x_i$	150	1480	1290	9560	21150	21960	14630	6000	76220	-	-	-	-

Корреляционный момент вычисляем по формуле:

$$S_{xy} = \frac{1}{n-1} \left( \sum \sum m_{ij} x_i y_j - \frac{1}{n} \left( \sum m_{xi} x_i \right) \left( \sum m_{yj} y_j \right) \right) = \frac{1}{99} \left( 76220 - \frac{1}{100} (72200 * 94.8) \right) = 78.53$$

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

$$y = y' + r_{xy} \frac{s_y}{s_x} (x - x'), \text{ где } S_x = \sqrt{83652.53} \approx 289.23; S_y = \sqrt{0.10636} \approx 0.32613$$

$$r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{78.53}{289.23 * 0.32613} \approx 0.8325$$

Составляем уравнение эмпирической линии регрессии y на x:

$$y = 0.948 + 0.8325 * \frac{0.32613}{289.23} (x - 722)$$

$$y = 0.00093871x + 0.270251$$

Строим линию регрессии и случайные точки  $(x_i, y_i)$ :

