

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:

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

$\begin{matrix} Y \\ X \end{matrix}$	1200	2700	4200	6700	8200	9700	11200	12700	m_x
20	4	2	5	-	-	-	-	-	11
520	-	-	7	5	2	-	-	-	14
1020	-	-	-	9	14	6	-	-	29
1520	-	-	-	7	8	6	-	-	21
2020	-	-	-	-	4	5	7	-	16
2520	-	-	-	-	-	3	2	4	9
m_y	4	2	12	21	28	20	9	4	100

Решение:

Для подсчета числовых характеристик (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.

Вычисляем выборочные средние \bar{X} и \bar{Y} :

$$\bar{X} = \frac{\sum \sum m_{ij} x_i}{n} = \frac{\sum m_{x_i} x_i}{n} = 1240$$

$$\bar{Y} = \frac{\sum m_{y_j} y_j}{n} = 7765$$

Выборочные дисперсии находим по формулам:

$$s_x^2 = \frac{1}{n-1} \left(\sum m_{x_i} x_i^2 - \frac{1}{n} \left(\sum m_{x_i} x_i \right)^2 \right) = 511600$$

$$s_y^2 = \frac{1}{n-1} \left(\sum m_{y_i} y_i^2 - \frac{1}{n} \left(\sum m_{y_i} y_i \right)^2 \right) = 6838275$$

СКО

$$S_x = 715,26$$

$$S_y = 2615,01$$

X	Y							
	1200	2700	4200	6700	8200	9700	11200	12700
20	96000	108000	420000	0	0	0	0	0
			152880	17420				
520	0	0	00	000	8528000	0	0	0
				61506	1170960			
1020	0	0	0	000	00	59364000	0	0
				71288	9971200			
1520	0	0	0	000	0	88464000	0	0
					6625600			
2020	0	0	0	0	0	97970000	158368000	0
2520	0	0	0	0	0	73332000	56448000	128016000

Корреляционный момент вычисляем по формуле:

$$S_{xy} = \frac{1}{n-1} \left(\sum \sum m_{ij} x_i y_j - \frac{1}{n} \left(\sum m_{x_i} x_i \right) \left(\sum m_{y_j} y_j \right) \right) = 11196800$$

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

$$y = \bar{y} + r_{xy} \frac{S_y}{S_x} (x - \bar{x}),$$

$$r_{xy} = \frac{S_{xy}}{S_x S_y} = 0,84$$

Составляем уравнение эмпирической линии регрессии y на x:

$$y = 3,07 * (x - 1240) + 7765 = 3,07 * x + 3964,05$$

