

Теория вероятностей

Домашнее задание №2

Федоров Егор, Р3215, вариант 19

1 ИДЗ 19.1

В результате эксперимента получены данные, записанные в виде статистического ряда. Требуется:

1. Записать значения результатов эксперимента в виде вариационного ряда;
2. Найти размах варьирования и разбить его на 9 интервалов;
3. Построить полигон частот, гистограмму относительных частот и график эмпирической функции распределения;
4. Найти числовые характеристики выборки \bar{x} и D_B ;
5. Приняв в качестве нулевой гипотезы H_0 : генеральная совокупность, из которой извлечена выборка, имеет нормальное распределение, проверить ее, пользуясь критерием Пирсона при уровне значимости $\alpha = 0.025$;
6. Найти доверительные интервалы для математического ожидания и среднего квадратичного отклонения при надежности $\gamma = 0.9$;

1.1 Решение

1. Вариационный ряд для данной выборки представлен на табл. 2.
2. Размах варьирования $\omega = x_{\max} - x_{\min} = 80.8 - 10.8 = 70$. Длина частотного интервала $h = \omega/l$, где $l = 9$. Тогда $h = 70/9 \approx 7.78$. Границы интервалов и их середины представлены на табл. 3.
3. Полигон частот, гистограмма относительных частот и график эмпирической функции распределения представлены на рисунках 1, 2, 3.
- 4.

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = 46.243$$

$$D_B = \frac{1}{n} \sum_{i=1}^n (x_i^2 - \bar{x}^2) = 364.7415 \quad \sigma_B = \sqrt{D_B} \approx 19.0982$$

$$S^2 = \frac{n}{n-1} D_B = 368.4257$$

5. Перейдем к случайной величине $z = (x - \bar{x})/\sigma_B$. Вычислим концы интервалов $z_i = (x_i - \bar{x})/\sigma_B$, причем z_1 положим стремящимся к $-\infty$, а $z_{m+1} \rightarrow +\infty$. Вычисления теоретических вероятностей представлены на таблице 4 и 5. Наблюдаемое значение критерия Пирсона:

$$\begin{aligned} \chi_{\text{набл}}^2 &= \sum_{i=1}^9 \frac{(n_i - n'_i)^2}{n'_i} \approx \frac{(6 - 6.81)^2}{46.38} + \frac{(13 - 8.1)^2}{65.61} + \frac{(11 - 11.5)^2}{132.2} + \frac{(11 - 14.87)^2}{221.1} + \frac{(15 - 15.83)^2}{250.6} + \\ &+ \frac{(14 - 14.76)^2}{217.9} + \frac{(8 - 11.98)^2}{143.5} + \frac{(13 - 8.03)^2}{64.48} + \frac{(9 - 8.076)^2}{65.22} = \\ &= \frac{0.6561}{46.38} + \frac{24.01}{65.61} + \frac{0.25}{132.2} + \frac{14.98}{221.1} + \frac{0.6889}{250.6} + \frac{0.5776}{217.9} + \frac{15.84}{143.5} + \frac{24.7}{64.48} + \frac{0.8538}{65.22} = \\ &= 0.01415 + 0.366 + 0.00189 + 0.06773 + 0.002749 + 0.002651 + 0.1104 + 0.3831 + 0.01309 \\ &= 0.9617 \end{aligned}$$

Уровню значимости $\alpha = 0.025$ и числу степеней свободы $k = l - 3 = 6$ соответствует $\chi_{\text{крит}}^2 = 14.4$.

Таким образом, $\chi_{\text{набл}}^2 < \chi_{\text{крит}}^2$, а значит гипотеза H_0 принимается.

6. Так как X распределена нормально, то с надежностью γ можно утверждать, что математическое ожидание a покрывается доверительным интервалом

$$\left(\bar{x} - \frac{\tilde{\sigma}_B}{\sqrt{n}} t_\gamma; \bar{x} + \frac{\tilde{\sigma}_B}{\sqrt{n}} t_\gamma \right)$$

$$t_\gamma = 1.65 \quad \tilde{\sigma}_B = \sqrt{S^2} = \sqrt{368.4257} \quad n = 10$$

$$\frac{\tilde{\sigma}_B}{\sqrt{n}} t_\gamma = \frac{\sqrt{368.4257}}{10} 1.65 \approx 3.167 \quad \bar{x} = 46.243$$

Таким образом, искомый доверительный интервал:

$$(43.075; 49.41)$$

Для оценки σ с уверенностью γ используем доверительный интервал $(\tilde{\sigma}_B(1-q); \tilde{\sigma}_B(1+q))$ При $n = 100$ и $\gamma = 0.9$, $q = 0.143$. Тогда искомый доверительный интервал:

$$(16.4496; 21.9392)$$

19.3	44.5	49.9	26.9	50.2	51.1	18.6	72.7	35.4	25.4
42.7	17.5	51.7	49.3	26.2	47.1	71.4	27.1	75.7	43.2
25.5	27.2	80.4	50.4	70.2	14.9	52.4	62.3	41.7	49.5
40.6	14.5	62.8	34.5	53.4	26.1	69.3	52.5	27.3	80.3
25.3	43.1	27.4	80.1	68.4	63.3	13.4	55.4	39.5	33.1
38.4	19.7	63.8	40.4	80.8	56.4	66.1	27.5	79.1	24.6
28.6	47.9	78.4	57.4	66.5	37.3	23.4	67.6	11.1	64.3
22.7	64.8	36.2	58.7	10.8	47.7	58.4	29.2	46.7	77.2
51.9	31.3	44.7	66.3	20.1	65.3	45.5	76.3	67.8	35.1
66.9	18.9	42.9	50.7	34.9	43.5	32.5	48.4	53.1	65.8

Таблица 1: Данные, полученные в результате эксперимента

10.8	11.1	13.4	14.5	14.9	17.5	18.6	18.9	19.3	19.7
20.1	22.7	23.4	24.6	25.3	25.4	25.5	26.1	26.2	26.9
27.1	27.2	27.3	27.4	27.5	28.6	29.2	31.3	32.5	33.1
34.5	34.9	35.1	35.4	36.2	37.3	38.4	39.5	40.4	40.6
41.7	42.7	42.9	43.1	43.2	43.5	44.5	44.7	45.5	46.7
47.1	47.7	47.9	48.4	49.3	49.5	49.9	50.2	50.4	50.7
51.1	51.7	51.9	52.4	52.5	53.1	53.4	55.4	56.4	57.4
58.4	58.7	62.3	62.8	63.3	63.8	64.3	64.8	65.3	65.8
66.1	66.3	66.5	66.9	67.6	67.8	68.4	69.3	70.2	71.4
72.7	75.7	76.3	77.2	78.4	79.1	80.1	80.3	80.4	80.8

Таблица 2: Вариационный ряд

i	x_i	x_{i+1}	$x'_i = (x_i + x_{i+1})/2$	n_i	$W_i = n_i/n$	W_i/h
1	10.8000	18.5778	14.6889	6	0.0600	0.0077
2	18.5778	26.3556	22.4667	13	0.1300	0.0167
3	26.3556	34.1333	30.2444	11	0.1100	0.0141
4	34.1333	41.9111	38.0222	11	0.1100	0.0141
5	41.9111	49.6889	45.8000	15	0.1500	0.0193
6	49.6889	57.4667	53.5778	14	0.1400	0.0180
7	57.4667	65.2444	61.3556	8	0.0800	0.0103
8	65.2444	73.0222	69.1333	13	0.1300	0.0167
9	73.0222	80.8000	76.9111	9	0.0900	0.0116

Таблица 3: Распределение данных по интервалам

i	x_i	x_{i+1}	$x_i - \bar{x}$	$x_{i+1} - \bar{x}$	$z_i = (x_i - \bar{x})/\sigma$	$z_{i+1} = (x_{i+1} - \bar{x})/\sigma$
1	10.8000	18.5778	—	-27.6652	—	-1.4486
2	18.5778	26.3556	-27.6652	-19.8874	-1.4486	-1.0413
3	26.3556	34.1333	-19.8874	-12.1097	-1.0413	-0.6341
4	34.1333	41.9111	-12.1097	-4.3319	-0.6341	-0.2268
5	41.9111	49.6889	-4.3319	3.4459	-0.2268	0.1804
6	49.6889	57.4667	3.4459	11.2237	0.1804	0.5877
7	57.4667	65.2444	11.2237	19.0014	0.5877	0.9949
8	65.2444	73.0222	19.0014	26.7792	0.9949	1.4022
9	73.0222	80.8000	26.7792	—	1.4022	—

Таблица 4: Вычисление теоретических частот

i	z_i	z_{i+1}	$\Phi(z_i)$	$\Phi(z_{i+1})$	$P_i = \Phi(z_{i+1}) - \Phi(z_i)$	$n'_i = P_i \cdot n$
1	—	-1.4486	-0.5	-0.4319	0.0681	6.81
2	-1.4486	-1.0413	-0.4319	-0.3508	0.081	8.1
3	-1.0413	-0.6341	-0.3508	-0.2357	0.1151	11.51
4	-0.6341	-0.2268	-0.2357	-0.0870	0.1487	14.87
5	-0.2268	0.1804	-0.0870	0.0714	0.1583	15.83
6	0.1804	0.5877	0.0714	0.21904	0.1476	14.76
7	0.5877	0.9949	0.21904	0.33891	0.1198	11.98
8	0.9949	1.4022	0.33891	0.41924	0.0803	8.03
9	1.4022	—	0.41924	0.5	0.08076	8.076

Таблица 5: Вычисление теоретических частот

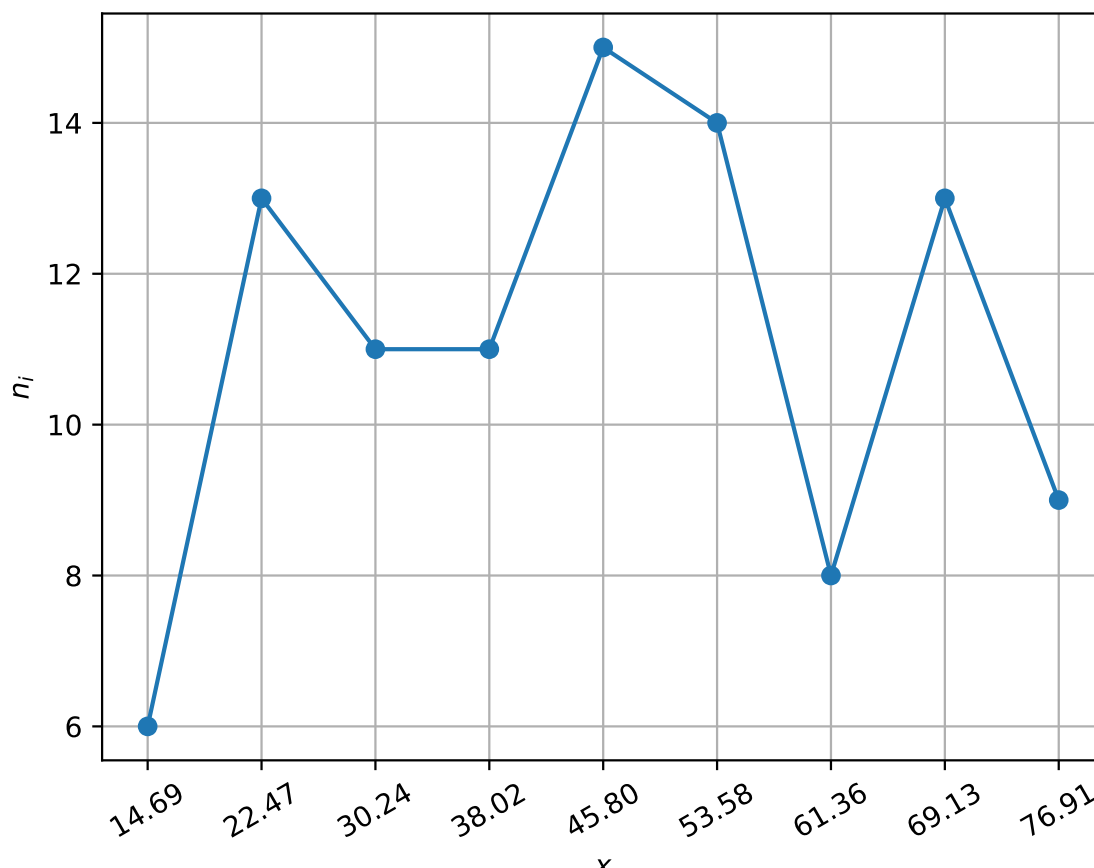


Рис. 1: Полигон частот

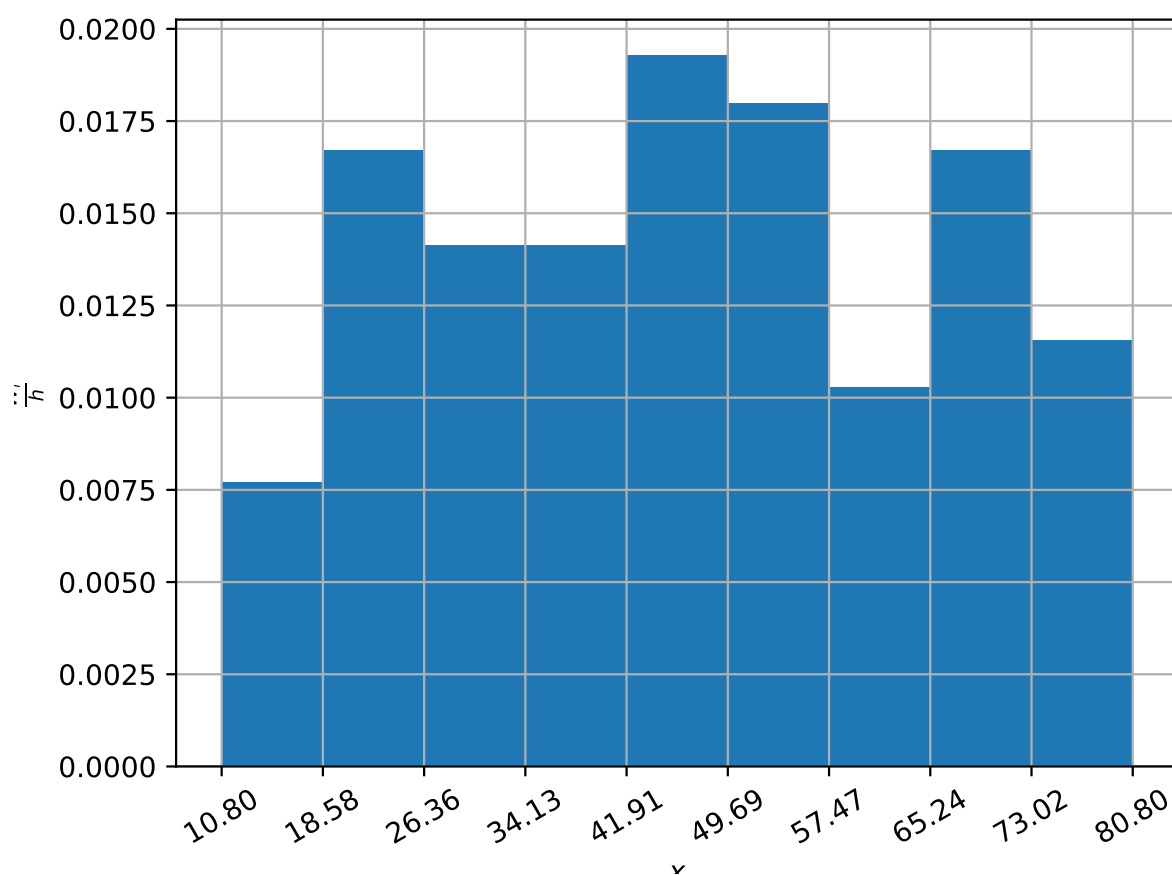


Рис. 2: Гистограмма относительных частот

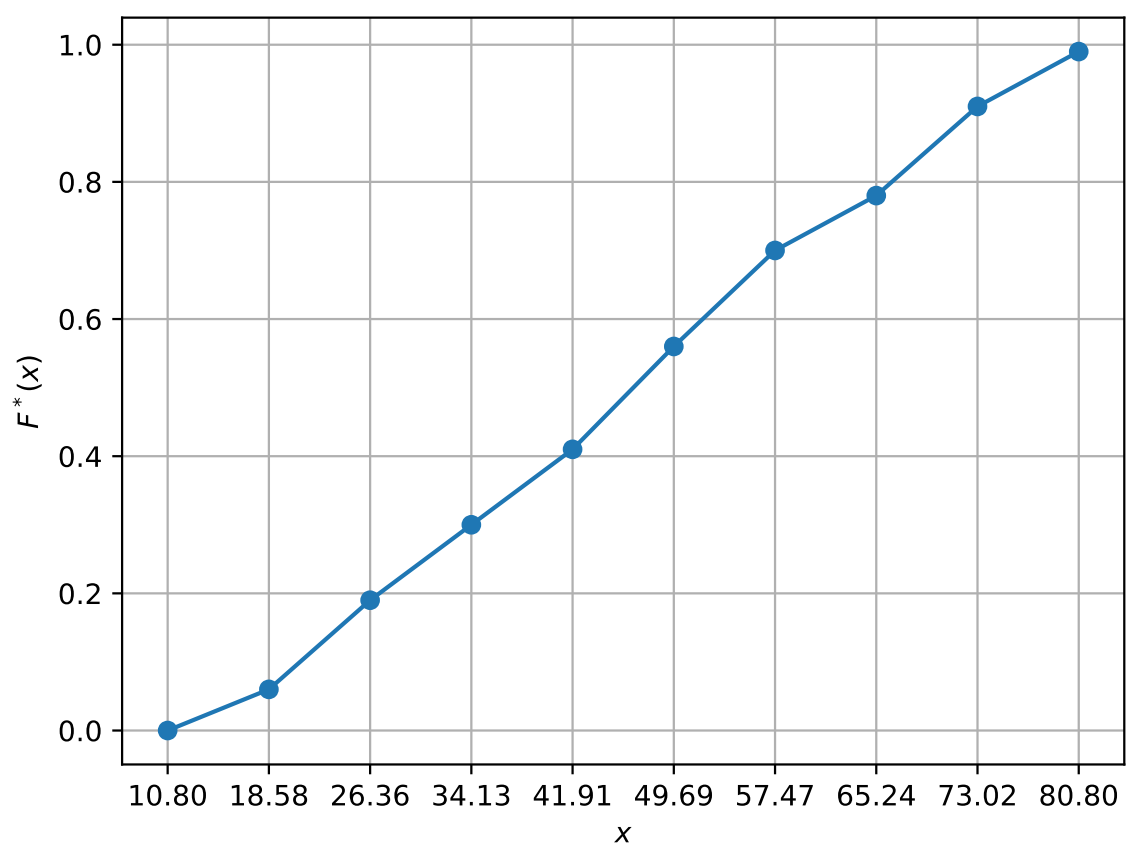


Рис. 3: График эмпирической функции распределения

2 ИДЗ 19.2

Дана таблица распределения 100 заводов по производственным средствам X (тыс. ден. ед.) и по суточной выработке 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) .

2.1 Решение

Найдем выборочные средние \bar{x} и \bar{y} .

$$\bar{x} = \frac{1}{n} \sum_{i=1}^6 m_{xi} x_i = \frac{1}{100} (11 \cdot 120 + 13 \cdot 130 + 15 \cdot 140 + 26 \cdot 150 + 160 \cdot 20 + 15 \cdot 170) = \frac{14760}{100} = 147.6$$

$$\bar{y} = \frac{1}{n} \sum_{j=1}^8 m_{yj} y_j = \frac{1}{100} (5 \cdot 8.0 + 9 \cdot 8.8 + 9.6 \cdot 8 + 17 \cdot 10.4 + 19 \cdot 11.2 + 13 \cdot 12.0 + 16 \cdot 12.8 + 13 \cdot 13.6) = \frac{1123.2}{100} = 11.232$$

Выборочная дисперсия s_x^2 :

$$\begin{aligned} s_x^2 &= \frac{1}{n-1} \left(\sum m_{xi} x_i^2 - \frac{1}{100} \left(\sum m_{xi} x_i \right)^2 \right) = \frac{1}{n-1} \left(\sum m_{xi} x_i^2 - \frac{1}{100} (100\bar{x})^2 \right) = \\ &= \frac{1}{99} \left[(11 \cdot 120^2 + 13 \cdot 130^2 + 15 \cdot 140^2 + 26 \cdot 150^2 + 20 \cdot 160^2 + 15 \cdot 170^2) - \frac{1}{100} (100\bar{x})^2 \right] = \\ &= \frac{1}{99} \left[2202600 - \frac{1}{100} 14760^2 \right] = \\ &= \frac{728}{3} = 242.66 \end{aligned}$$

Для вычисления корреляционного момента сначала вычислим сумму $\sum \sum m_{ij} x_i y_j$:

$$\begin{aligned} \sum \sum m_{ij} x_i y_j &= 5 \cdot 120 \cdot 8.0 + 6 \cdot 120 \cdot 8.8 + 3 \cdot 130 \cdot 8.8 + 4 \cdot 130 \cdot 9.6 + 6 \cdot 130 \cdot 10.4 + \\ &+ 4 \cdot 140 \cdot 9.6 + 5 \cdot 140 \cdot 10.4 + 6 \cdot 140 \cdot 11.2 + 6 \cdot 150 \cdot 10.4 + 13 \cdot 150 \cdot 11.2 + 7 \cdot 150 \cdot 12.0 + \\ &+ 6 \cdot 160 \cdot 12.0 + 9 \cdot 160 \cdot 12.8 + 5 \cdot 160 \cdot 13.6 + 7 \cdot 170 \cdot 12.8 + 8 \cdot 170 \cdot 13.6 = \\ &= 168096 \end{aligned}$$

Корреляционный момент s_{xy} :

$$\begin{aligned} s_{xy} &= \frac{1}{n-1} \left(\sum \sum m_{ij} x_i y_j - \frac{1}{n} \left(\sum m_{xi} x_i \right) \left(\sum m_{yj} y_j \right) \right) = \\ &= \frac{1}{n-1} \left(\sum \sum m_{ij} x_i y_j - \frac{1}{n} (100\bar{x})(100\bar{y}) \right) = \\ &= \frac{1}{99} \left(168096 - \frac{1}{100} (14760)(1123.2) \right) = \\ &= 23.3503 \end{aligned}$$

Тогда зависимость y от x выражается формулой:

$$y = \bar{y} + r_{xy} \frac{s_y}{s_x} (x - \bar{x}) = \bar{y} + \frac{s_{xy}}{s_x s_y} \frac{s_y}{s_x} (x - \bar{x}) = \bar{y} + \frac{s_{xy}}{s_x^2} (x - \bar{x}) =$$

Подставляя полученные ранее значения, получаем:

$$y = 11.232 + \frac{23.3503}{242.66} (x - 147.6) = 0.096x - 2.971$$

График линии регрессии представлен на рисунке 4

$X \backslash Y$	8.0	8.8	9.6	10.4	11.2	12.0	12.8	13.6	m_x
120	5	6	-	-	-	-	-	-	11
130	-	3	4	6	-	-	-	-	13
140	-	-	4	5	6	-	-	-	15
150	-	-	-	6	13	7	-	-	26
160	-	-	-	-	-	6	9	5	20
170	-	-	-	-	-	-	7	8	15
m_y	5	9	8	17	19	13	16	13	100

Таблица 6: Таблица распределения

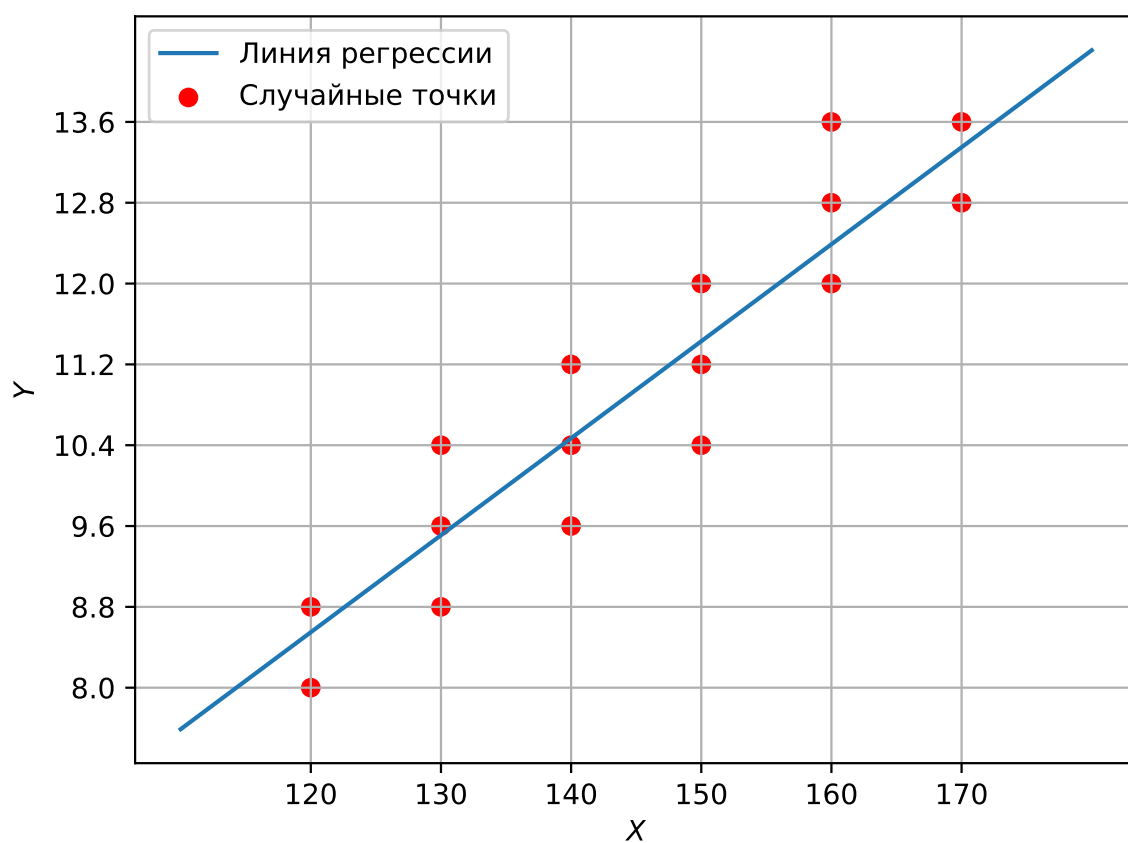


Рис. 4: Линия регрессии и случайные точки