

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАУЧНО-ОБРАЗОВАТЕЛЬНАЯ КОРПОРАЦИЯ ИТМО»

ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И КОМПЬЮТЕРНОЙ ТЕХНИКИ



Индивидуальное домашнее задание №2

Вариант 12

Выполнил:

Степанов Арсений Алексеевич

Группа:

ТеорВер 2.4

Преподаватель:

Селина Елена Георгиевна

Санкт-Петербург, 2023г.

Задание №1

Исходные данные

32	105	48	80	144	128	64	112	18	81
66	129	113	17	94	78	90	51	104	34
110	149	36	103	82	53	93	130	68	150
114	84	55	131	70	38	102	77	16	135
41	19	142	61	85	159	115	57	72	101
56	100	86	146	73	40	141	25	87	126
151	71	94	15	125	76	54	99	39	140
17	124	52	98	139	37	147	88	69	109
35	158	67	30	93	123	50	138	21	97
96	121	49	137	89	145	91	65	92	33

Вариационный ряд

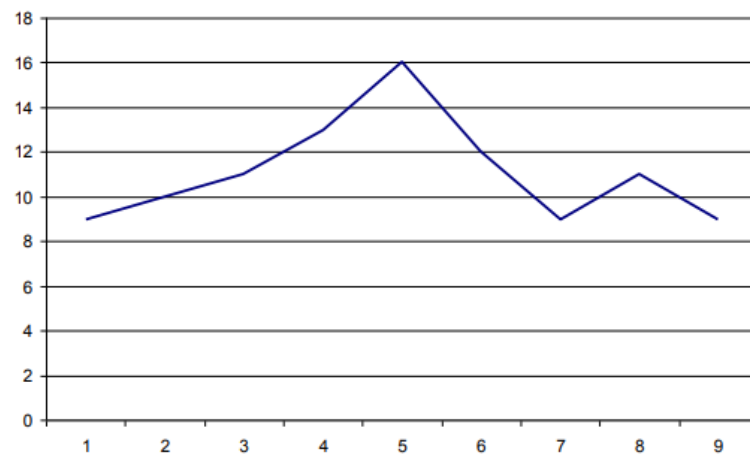
15	16	17	17	18	19	21	25	30	32
33	34	35	36	37	38	39	40	41	48
49	50	51	52	53	54	55	56	57	61
64	65	66	67	68	69	70	71	72	73
76	77	78	80	81	82	84	85	86	87
88	89	90	91	92	93	93	94	94	96
97	98	99	100	101	102	103	104	105	109
110	112	113	114	115	121	123	124	125	126
128	129	130	131	135	137	138	139	140	141
142	144	145	146	147	149	150	151	158	159

Размах выборки $R = x_{\max} - x_{\min} = 159 - 15 = 144$

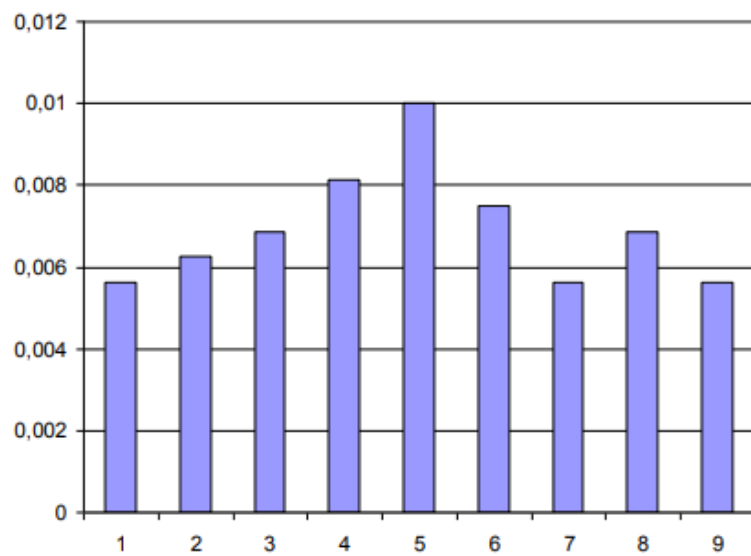
Величина интервала находится по формуле: $h = \frac{144}{9} = 16$

Интервал $[x_i, x_{i+1})$	n_i	$w_i, w_i = \frac{n_i}{nh}$	$f_i, f_i = \frac{n_i}{n}$
15-31	9	0.005625	0.09
31-47	10	0.00625	0.19
47-63	11	0.006875	0.3
63-79	13	0.008125	0.43
79-95	16	0.01	0.59
95-111	12	0.0075	0.71
111-127	9	0.005625	0.8
127-143	11	0.006875	0.91
143-159	9	0.005625	1

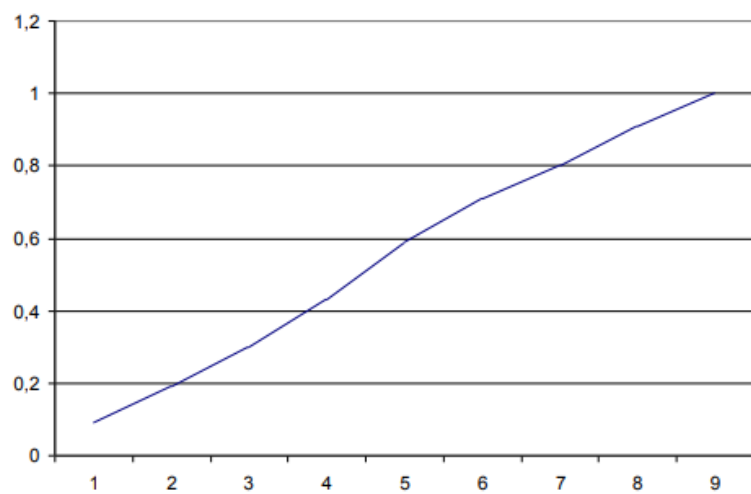
Полигон частот



Гистограмма



Эмпирическая функция распределения



Числовые характеристики выборки

Интервал $[x_i, x_{i+1})$	x'_i	n_i	$n_i x'_i$	$(x'_i)^2$	$n_i (x'_i)^2$
15-31	23	9	207	529	4761
31-47	36	10	360	1296	12960
47-63	55	11	605	3025	33275
63-79	71	13	923	5041	65533
79-95	87	16	1392	7569	121104
95-111	103	12	1236	10609	127308
111-127	119	9	1071	14161	127449
127-143	135	11	1485	18225	200475
143-159	151	9	1359	22801	205209

$$\bar{x} = \sum_{i=1}^9 \frac{x'_i n_i}{n} = 86.68 \quad D = \sum_{i=1}^9 \frac{(x_i - \bar{x})^2 n_i}{n} = 1489.82$$

Приняв в качестве нулевой гипотезу H_0 : генеральная совокупность, из которой извлечена выборка, имеет нормальное распределение, проверить ее, пользуясь критерием Пирсона при уровне значимости $\alpha = 0.025$

Перейдём к новой случайной величине $Z = \frac{X - \bar{x}}{\sqrt{D}}$ и вычислим концы интервалов $z_i = \frac{x_i - \bar{x}}{\sqrt{D}}$

$$n'_i = n \cdot P_i, P_i = \Phi(z_{i+1}) - \Phi(z_i)$$

i	Z_i	$Z_i + 1$	$\Phi(z_i)$	$\Phi(z_{i+1})$	P_i	n'_i	n_i	$\frac{(n_i - n'_i)^2}{n'_i}$
1	$-\infty$	-1.44256	-0.5	-0.4254	0.0745	7.45728	9	0.319149
2	-1.44256	-1.02803	-0.4254	-0.3480	0.0773	7.739544	10	0.660202
3	-1.02803	-0.6135	-0.3480	-0.2302	0.1178	11.78044	11	0.051703
4	-0.6135	-0.19897	-0.2302	-0.0788	0.1513	15.13692	13	0.301675
5	-0.19897	0.215554	-0.0788	0.0853	0.1641	16.41906	16	0.010696
6	0.215554	0.630082	0.0853	0.2356	0.150	15.03471	12	0.612546
7	0.630082	1.04461	0.2356	0.3518	0.116	11.62188	9	0.591491
8	1.04461	1.459137	0.3518	0.4277	0.0758	7.583805	11	1.538856
9	1.459137	∞	0.4277	0.5	0.0722	7.226369	9	0.435318

$$\chi_{\text{набл}} = \sum_{i=1}^9 \frac{(n_i - n'_i)^2}{n'_i} = 4.5216$$

По таблице квантилей распределения χ^2 по заданному уровню значимости α и числу степеней свободы $k = 9 - 3 = 6$, можно найти критическое значение $\chi^2_{\text{крит}} = 14.45$, а так как $\chi^2_{\text{набл}} < \chi^2_{\text{крит}}$, то распределение можно считать нормальным на уровне значимости $\alpha = 0.25$

Доверительные интервалы выборки

Находим доверительные интервалы для математического ожидания и средне-квадратического отклонения с доверительной вероятностью $\gamma = 0.9$

Доверительный интервал находится по формуле:

$$\left(\bar{x} - \frac{t_\gamma S}{\sqrt{n}}; \bar{x} + \frac{t_\gamma S}{\sqrt{n}} \right) \quad S = \sqrt{\frac{n}{n-1} D}$$

$$t_\gamma = 1.66 \quad S = 38.79 \Rightarrow MX \in (80.24; 93.12)$$

Доверительный интервал для среднеквадратического отклонения находится по формуле:

$$\left(\frac{S\sqrt{n-1}}{\chi_2}; \frac{S\sqrt{n-1}}{\chi_1} \right)$$

χ_1 и χ_2 - квантили распределения χ^2 уровня $\frac{1-\gamma}{2} = 0.05$ и $1 - \frac{1-\gamma}{2} = 0.95$ с 99 степенями свободы

$$\chi_1 = 8.77, \chi_2 = 11.1, \text{ тогда } \sigma \in \left(\frac{\sqrt{99} \cdot 38.79}{11.1}; \frac{\sqrt{99} \cdot 38.79}{8.77} \right)$$

$$34.77 < \sigma < 44.01$$

Задание №2

Исходные данные

X	Y								m_x
	110	130	150	170	190	210	230	250	
10	1	3	4	-	-	-	-	-	8
13	-	5	6	5	-	-	-	-	16
16	-	-	4	8	6	-	-	-	18
19	-	-	6	15	9	-	-	-	30
22	-	-	-	-	5	6	7	-	18
25	-	-	-	-	-	1	7	2	10
m_y	1	8	20	28	20	7	14	2	100

Решение

Найдём выборочные средние x и y :

$$x = \frac{1}{n} \sum_{i=1}^6 \sum_{j=1}^8 m_{ij} x_i = \frac{1792}{100} = 17.92 \quad y = \frac{1}{n} \sum_{i=1}^6 \sum_{j=1}^8 m_{ij} y_i = \frac{17900}{100} = 179$$

Найдём выборочные дисперсии:

$$S_x^2 = \frac{1}{n-1} \left(\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} x_i \right)^2 \right) = \frac{1}{99} (33904 - \frac{1}{100} 1792^2) = 18.0945$$

$$S_y^2 = \frac{1}{n-1} \left(\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} y_i^2 - \frac{1}{n} \left(\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} y_i \right)^2 \right) = \frac{1}{99} (3302800 - \frac{1}{100} 17900^2) = 996.97$$

Найдём корреляционный момент:

$$S_{xy} = \frac{1}{n-1} \left(\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} x_i y_j - \frac{1}{n} \left(\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} x_i \right) \left(\sum_{i=1}^6 \sum_{j=1}^8 m_{ij} y_i \right) \right)$$

$$S_{xy} = \frac{1}{99} (331880 - \frac{1}{100} (1792 \cdot 17900)) = 112.242$$

Найдём уравнение эмпирической линии регрессии:

$$S_x = \sqrt{18.0945} = 4.2538 \quad S_y = \sqrt{996.97} = 31.5748 \quad r_{xy} = \frac{S_{xy}}{S_x \cdot S_y} = 0.8357$$

$$y = y_0 + r_{xy} \frac{S_y}{S_x} (x - x_0) = 179 + 0.8357 \cdot \frac{31.5748}{4.2538} \cdot (x - 17.92) = 6.20311x + 67.8403$$

Линия регрессии и случайные точки

