

Федеральное государственное автономное
образовательное учреждение высшего образования
«Научно-образовательная корпорация ИТМО»

**ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И КОМПЬЮТЕРНОЙ
ТЕХНИКИ**

**Индивидуальное домашнее задание №1
«Обработка одномерной выборки. Исследование распределения
случайной величины.»**

Вариант № 2

Работу выполнили:
студент группы Р3209
Зайцева И. С.
студент группы Р3217
Русакова Е. Д.

Преподаватель:
Милованович Е. В.

г. Санкт-Петербург
2024 г.

Цель работы:

На основании анализа выборки:

1. Построить интервальный ряд исследуемого признака
2. Полигон частот, выборочную функцию распределения и гистограмму
3. Найти точечные оценки математического ожидания и дисперсии
4. Построить доверительные интервалы для математического ожидания и дисперсии с доверительной вероятностью 0,95
5. Проверить статистическую гипотезу о виде закона распределения

Исходные данные:

0,866	-0,005	0,403	1,908	0,448	0,169	-0,731	-1,189	0,905	0,283
2,431	1,409	0,191	-0,165	0,889	0,804	-2,131	-0,754	1,458	1,65
0,026	0,885	0,011	-0,99	-0,104	0,174	-0,052	-0,182	1,813	0,346
0,11	1,757	-0,693	-0,732	1,073	-1,724	-1,81	0,947	-1,118	0,666
0,97	1,14	-1,105	0,894	1,547	-0,484	-0,086	-0,066	0,15	-0,264
0,866	-0,005	0,403	1,908	0,448	0,169	-0,731	-1,189	0,905	0,283
2,431	1,409	0,191	-0,165	0,889	0,804	-2,131	-0,754	1,458	1,65
0,11	1,757	-0,693	-0,732	1,073	-1,724	-1,81	0,947	-1,118	0,666
0,026	0,885	0,011	-0,99	-0,104	0,174	-0,052	-0,182	1,813	0,346
0,97	1,14	-1,105	0,894	1,547	-0,484	-0,086	-0,066	0,15	-0,264

Вариационный ряд:

-2,131	-1,105	-0,693	-0,104	0,011	0,174	0,448	0,889	1,073	1,65
-2,131	-1,105	-0,693	-0,104	0,011	0,174	0,448	0,889	1,073	1,65
-1,81	-0,99	-0,484	-0,086	0,026	0,191	0,666	0,894	1,14	1,757
-1,81	-0,99	-0,484	-0,086	0,026	0,191	0,666	0,894	1,14	1,757
-1,724	-0,754	-0,264	-0,066	0,11	0,283	0,804	0,905	1,409	1,813
-1,724	-0,754	-0,264	-0,066	0,11	0,283	0,804	0,905	1,409	1,813
-1,189	-0,732	-0,182	-0,052	0,15	0,346	0,866	0,947	1,458	1,908
-1,189	-0,732	-0,182	-0,052	0,15	0,346	0,866	0,947	1,458	1,908
-1,118	-0,731	-0,165	-0,005	0,169	0,403	0,885	0,97	1,547	2,431
-1,118	-0,731	-0,165	-0,005	0,169	0,403	0,885	0,97	1,547	2,431

Количество элементов: $n = 100$

Количество интервалов: $k = \sqrt{n} = \sqrt{100} = 10$

$$x_{min} = -2,5$$

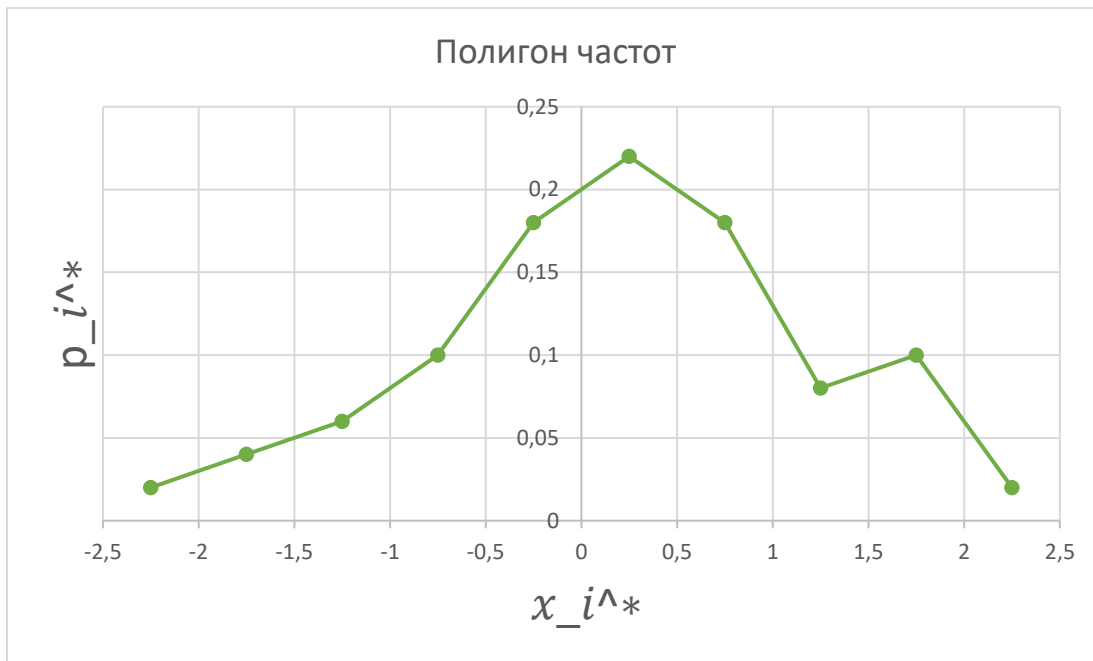
$$x_{max} = 2,5$$

Длина интервала: $\Delta x_i = \frac{x_{max} - x_{min}}{k} = 0,5$

номер интервала	1	2	3	4	5	6	7	8	9	10
границы интервала	[-2,5 ; -2)	[-2 ; -1,5)	[-1,5 ; -1)	[-1 ; -0,5)	[-0,5 ; 0)	[0 ; 0,5)	[0,5 ; 1)	[1 ; 1,5)	[1,5 ; 2)	[2 ; 2,5]
x_i^* представитель	-2,25	-1,75	-1,25	-0,75	-0,25	0,25	0,75	1,25	1,75	2,25
n_i	2	4	6	10	18	22	18	8	10	2
p_i^* частота	0,02	0,04	0,06	0,1	0,18	0,22	0,18	0,08	0,1	0,02
h_i	0,04	0,08	0,12	0,2	0,36	0,44	0,36	0,16	0,2	0,04

Полигон частот:

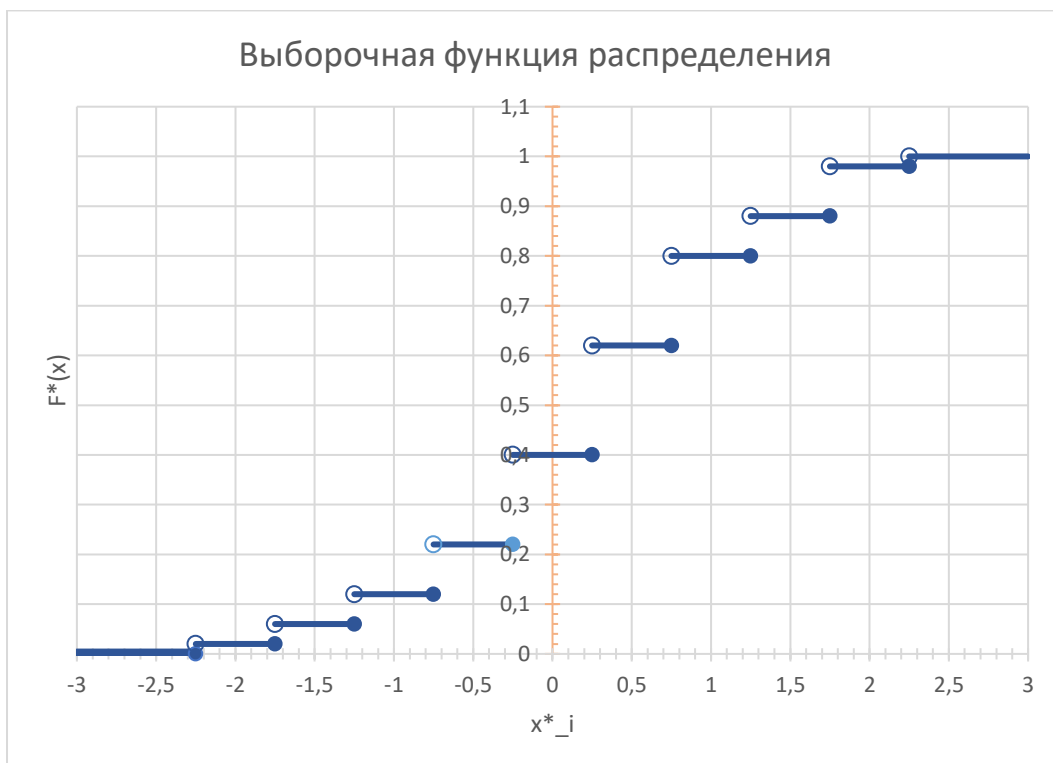
x_i^*	-2,25	-1,75	-1,25	-0,75	-0,25	0,25	0,75	1,25	1,75	2,25
p_i^*	0,02	0,04	0,06	0,1	0,18	0,22	0,18	0,08	0,1	0,02



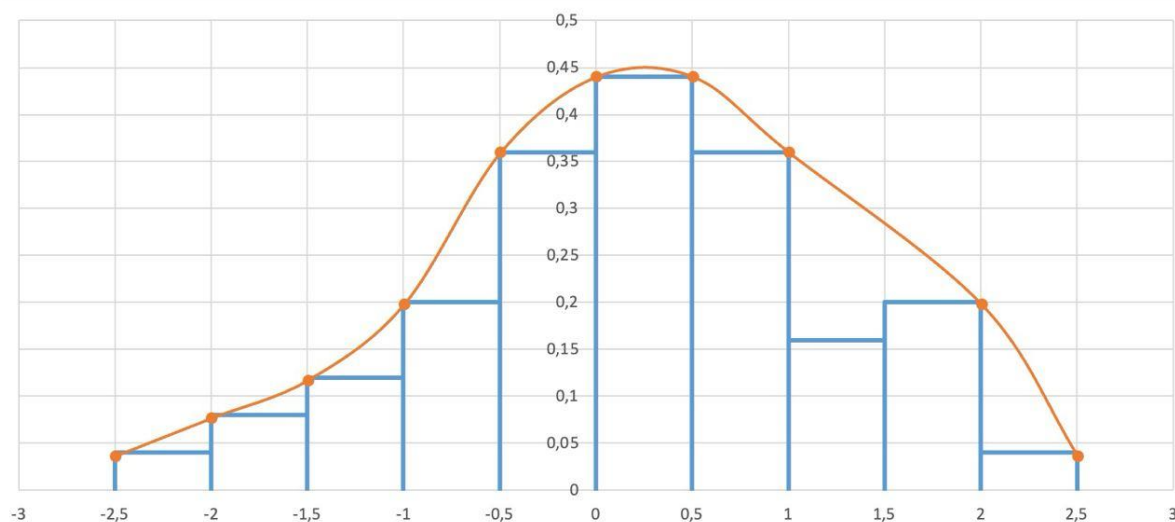
Выборочная функция распределения:

$$F^*(x) = p^*(x < x_i^*)$$

$$F^*(x) = \begin{cases} 0, & x \leq -2.25 \\ 0.02, & -2.25 < x \leq -1.75 \\ 0.06, & -1.75 < x \leq -1.25 \\ 0.12, & -1.25 < x \leq -0.75 \\ 0.22, & -0.75 < x \leq -0.25 \\ 0.4, & -0.25 < x \leq 0.25 \\ 0.62, & 0.25 < x \leq 0.75 \\ 0.8, & 0.75 < x \leq 1.25 \\ 0.88, & 1.25 < x \leq 1.75 \\ 0.98, & 1.75 < x \leq 2.25 \\ 1, & 2.25 < x \end{cases}$$



Гистограмма:



Точечная оценка математического ожидания:

$$\tilde{m} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{100} * 23,876 = 0,23876$$

Точечная оценка дисперсии:

$$\tilde{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \tilde{m})^2} = \sqrt{\frac{1}{100-1} \sum_{i=1}^n (x_i - 23,876)^2} = \sqrt{\frac{1}{99} * 100,954} = 1,0098$$

Доверительный интервал для математического ожидания с доверительным интервалом $\beta = 0,95$ при неизвестной дисперсии:

$$\tilde{m} - t_{n-1,\beta} * \frac{\tilde{\sigma}}{\sqrt{n}} < m < \tilde{m} + t_{n-1,\beta} * \frac{\tilde{\sigma}}{\sqrt{n}}$$

где $t_{n-1,\beta}$ – коэффициент Стьюдента с $n-1$ степень свободы и доверительным интервалом $\beta = 0,95$

$$0,23876 - 1,98 * \frac{1,0098}{\sqrt{100}} < m < 0,23876 + 1,98 * \frac{1,0098}{\sqrt{100}}$$
$$(0,0388; 0,4387)$$

Доверительный интервал для дисперсии с доверительным интервалом $\beta = 0,95$:

$$(n-1) * \frac{\tilde{\sigma}^2}{h_1} < \sigma^2 < (n-1) * \frac{\tilde{\sigma}^2}{h_2}$$

Где h_1 и h_2 – коэффициенты распределения хи-квадрат

$$h_1 = \chi^2_{n-1, \frac{1+\beta}{2}} = \chi^2_{99, 0.975}$$

$$h_2 = \chi^2_{n-1, \frac{1-\beta}{2}} = \chi^2_{99, 0.025}$$

$$99 * \frac{1,0197}{128,422} < \sigma^2 < 99 * \frac{1,0197}{73,361}$$

$$(0,7861; 1,3761)$$

Проверка статической гипотезы о виде распределения:

1. H_0 – случайная величина X имеет нормальное распределение
 H_1 – случайная величина X имеет распределение, отличное от нормального
2. Уровень значимости $\alpha = 0,05$
3. Критерий Пирсона

$$U = \sum_{i=1}^n \frac{(n_i - np_i)^2}{np_i} = 5,865556$$

номер интервала ла i	1	2	3	4	5	6	7	8	9	10
n_i	2	4	6	10	18	22	18	8	10	2
x_i	-2,5	-2	-1,5	-1	-0,5	0	0,5	1	1,5	2
x_{i+1}	-2	-1,5	-1	-0,5	0	0,5	1	1,5	2	2,50
p_i	0,0102	0,0291	0,06795	0,12515	0,1713	0,19355	0,17275	0,1206	0,0651	0,0283
np_i	1,02	2,91	6,795	12,515	17,13	19,355	17,275	12,06	6,51	2,83
U_i	0,94156	0,40828	0,09301	0,50541	0,04418	0,36145	0,03042	1,36679	1,87098	0,24342
	9	2	3	2	6	8	7	9	3	8

$$p_i = \Phi\left(\frac{x_{i+1} + \tilde{m}}{\tilde{\sigma}}\right) - \Phi\left(\frac{x_i + \tilde{m}}{\tilde{\sigma}}\right)$$

$\Phi(x)$ - функция Лапласа

$$U_{кр} = \chi^2_{7, 0.05} = 14,067$$

$$U < U_{кр} \Rightarrow U \in O(0; 14,067)$$

Случайная величина X имеет нормальное распределение

Результаты:

Точечная оценка математического ожидания: $\tilde{m} = 0,23876$

Точечная оценка дисперсии: $\tilde{\sigma} = 1,0098$

Доверительный интервал для математического ожидания: $m \in (0,0388; 0,4387)$

Доверительный интервал для дисперсии: $\sigma \in (0,7861; 1,3761)$

Вывод:

Проанализировав исходную выборку, мы нашли значение точечных оценок математического ожидания и дисперсии и их доверительные интервалы.

Вычислив значение статистического критерия по исходным данным, мы выявили, что он вошел в допустимую область O , следовательно, нашу гипотезу о виде закона распределения (нормального распределения) следует принять.