МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»

ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ

ЛАБОРАТОРНАЯ РАБОТА №1-5

по дисциплине

«Математическая статистика»

Исследование закона распределения случайной величины

Вариант № 41 (блок 4,10)

Выполнили:
Векшин А. И. P3216
Дашкевич Е.В. P3208
Кононова В.В. P3211
Преподаватель:
Танченко Ю.В.

Цель работы

На основании анализа выборки:

1) Построить интервальный ряд исследуемой случайной величины.
2) Построить полигон частот, выборочную функцию распределения и гистограмму.
3) Найти точечные оценки мат. ожидания и дисперсии.
4) Найти доверительные интервалы для мат. ожидания и дисперсии с доверительной вероятностью 0,95.
5) Проверить статистическую гипотезу о виде закона распределения генеральной совокупности.

Исходные данные

-0,644	-0,149	0,365	1,601	1,307	0,041	-2,312	1,023	1,88	-1,422
-0,905	0,577	-0,548	0,732	-0,482	0,413	1,38	-0,489	-0,799	-0,755
-0,716	0,753	0,578	0,555	-1,752	0,597	1,39	-0,402	-0,56	0,157
0,007	-0,167	-1,955	-0,813	-0,926	1,924	-0,453	1,399	1,708	0,378
-2,814	-0,581	0,522	-0,539	0,922	0,714	-0,628	0,28	-0,644	0,178
-0,719	1,202	-1,083	0,606	1,244	-1,547	-0,108	0,856	1,034	-0,127
-0,219	-0,112	0,157	0,074	0,029	-1,071	-0,3	3,343	-0,618	1,019
-0,03	0,673	-0,662	-0,685	-1,675	0,737	1,279	0,894	0,987	0,17
-0,495	-1,322	0,362	0,475	-0,043	-1,698	-0,404	-0,741	-0,237	-0,420
-0,333	-0,216	1,17	0,757	-0,691	-0,591	1,444	1,695	0,307	2,096

Интервальный ряд

Найдем размах выборки

$$x_{max} = 3.343$$

$$x_{min} = -2.814$$

$$w = 6,157$$

Вычислим шаг разбиения по формуле Стерджеса:

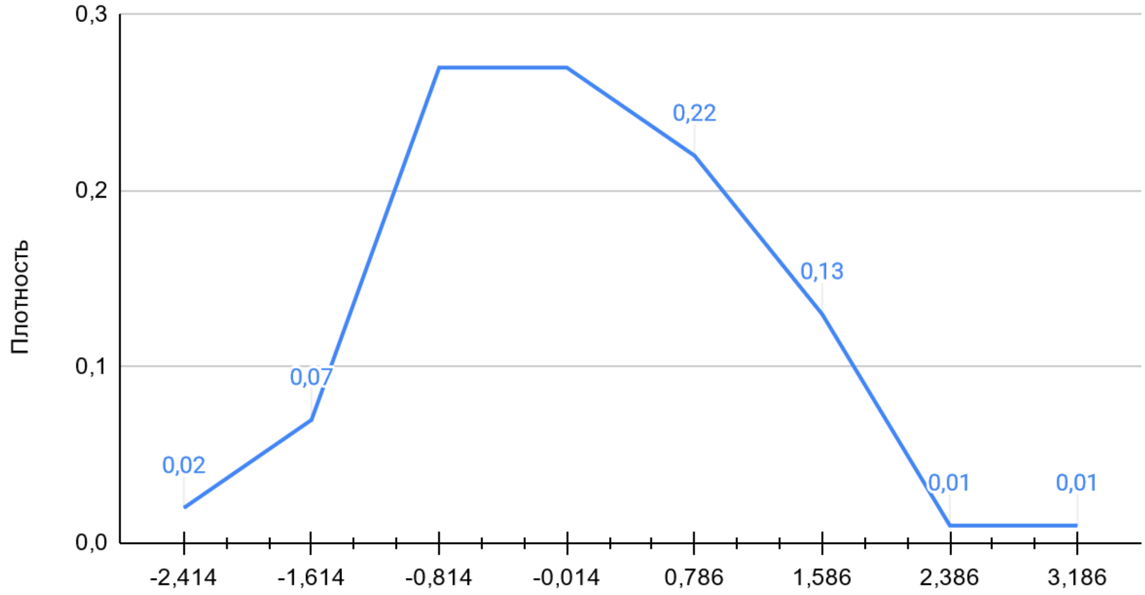
$$h = \frac{w}{1 + \log_2 N} = \frac{6,157}{1 + \log_2 100} \approx 0,8$$

Получим 8 интервалов длинны 0.8.

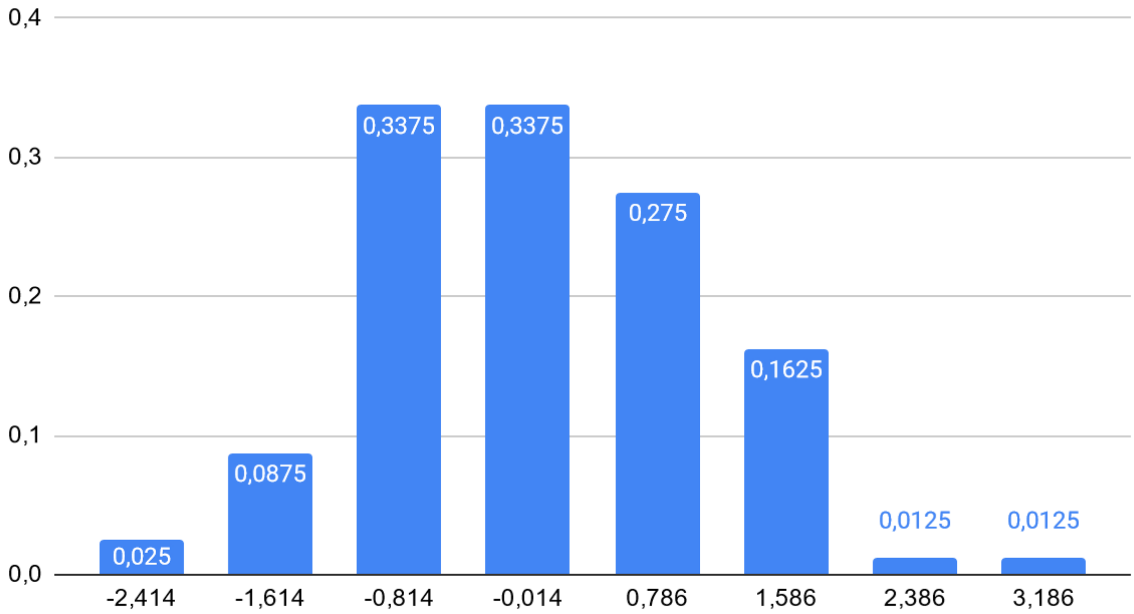
Занесем расчеты для каждого отрезка в таблицу:

Номера интервалов	1	2	3	4	5	6	7	8
Границы интервалов	(-2.8; -2.0)	(-2.0; -1.2)	(-1.2; -0.4)	(-0.4; 0.4)	(0.4; 1.2)	(1.2; 2.0)	(2.0; 2.8)	(2.8; 3.6)
x_i^*	-2,414	-1,614	-0,814	-0,014	0,786	1,586	2,386	3,186
n_i	2	7	27	27	22	13	1	1
p_i^*	0,02	0,07	0,27	0,27	0,22	0,13	0,01	0,01

Полигон частот

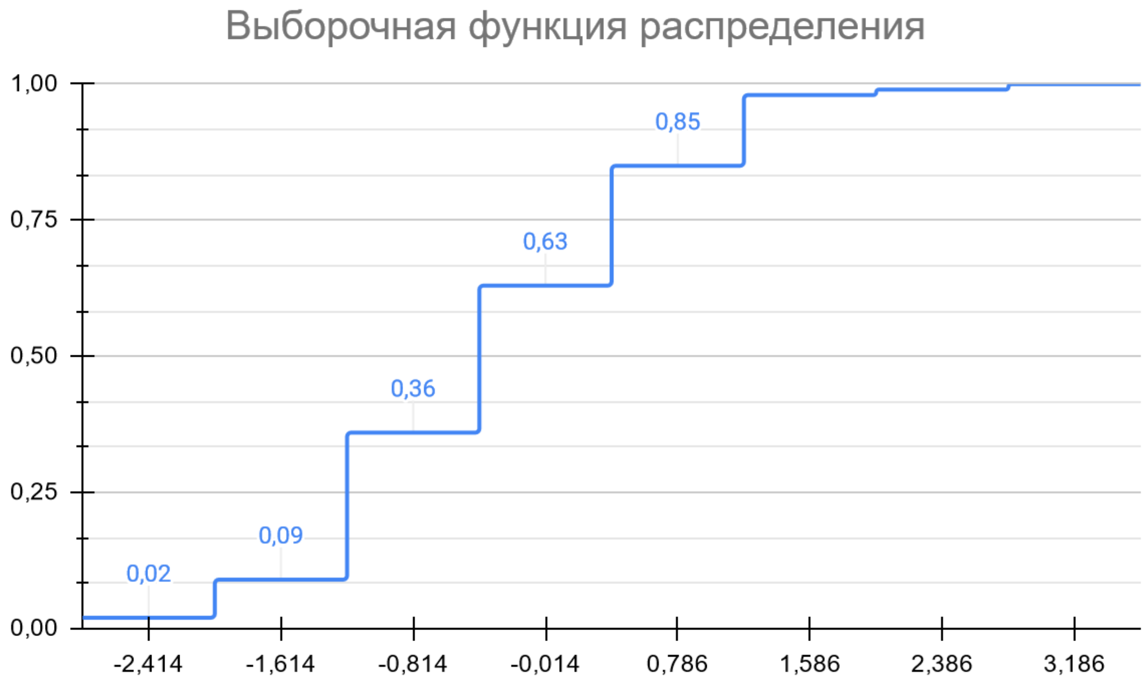


Гистограмма



Построим эмпирическую функцию распределения. Она имеет вид:

$$F^*(x)\begin{cases} 0 & x \le -2.41 \\ 0.02 & -2.41 < x \le -1.61 \\ 0.09 & -1.61 < x \le -0.81 \\ 0.36 & -0.81 < x \le -0.01 \\ 0.63 & -0.01 < x \le 0.78 \\ 0.85 & 0.78 < x \le 1.58 \\ 0.98 & 1.58 < x \le 2.38 \\ 0.99 & 2.38 < x \le 3.18 \\ 1 & x > 3.18 \end{cases}$$



Вычисление точечных оценок мат. ожидания и дисперсии

Найдем точечные оценки математического ожидания и дисперсии. В качестве таких оценок возьмем среднее выборочное значение $\bar{X} = \sum x_i^* p_i^*$ и выборочную дисперсию $S^2 = \sum x_i^{*2} p_i^* - \bar{X}^2$.

№ интервала	1	2	3	4	5	6	7	8	$\sum$
x_i^*	-2,41	-1,61	-0,81	-0,01	0,78	1,58	2,38	3,18	-
P_i^*	0,02	0,07	0,27	0,27	0,22	0,13	0,01	0,01	-
$x_i^* p_i^*$	-0,048	-0,113	-0,220	-0,004	0,173	0,206	0,024	0,032	0,05
$x_i^{*2} p_i^*$	0,117	0,182	0,179	0,000	0,136	0,327	0,057	0,102	1,099

Оценка мат. ожидания: $\bar{X} = \sum x_i^* p_i^* = 0{,}05$

Оценка дисперсии: $S^2 = \sum x_i^{*2} p_i^* - \bar{X}^2 = 1{,}0967$

Доверительные интервалы для мат. ожидания и дисперсии

Доверительный интервал: γ = 0,95

Ввиду большого объема выборки доверительный интервал имеет вид:

$$(\bar{X} - t\frac{S}{\sqrt{n}}, \bar{X} + t\frac{S}{\sqrt{n}})$$

Для рассматриваемого примера будем иметь при $\gamma = 0.95$, n=7 откуда t = 2,36, поэтому в нашем примере имеем:

$$\bar{X} - t\frac{S}{\sqrt{n}} = 0{,}05 - 2{,}36\ \frac{\sqrt{1.096}}{\sqrt{8}} \approx -0{,}823$$

$$\overline{X} + t\frac{S}{\sqrt{n}} = 0{,}05 + 2{,}36\ \frac{\sqrt{1.096}}{\sqrt{8}} \approx 0{,}923$$

Доверительный интервал для мат. ожидания: $-0{,}823 < m < 0.923$

Для дисперсии определим квантили распределения X^2 с 7 степенями свободы:

$$X^2_{1-\frac{\alpha}{2},\ n-1} = X^2_{0.975,\ 7} = 16.01$$

$$X^2_{\frac{\alpha}{2},\ n-1} = X^2_{0.025,\ 7} = 1.7$$

Подставим в формулу

$$\left(\frac{(n-1)s^2}{X^2_{1-\frac{\alpha}{2},\ n-1}}, \quad \frac{(n-1)s^2}{X^2_{\frac{\alpha}{2},\ n-1}}\right)$$

$$\left(\frac{7 * 1{,}0967}{16.01}, \quad \frac{7 * 1{,}0967}{1.7}\right)$$

Доверительный интервал для дисперсии: $0.479 < S^2 < 4.515$

Проверка гипотезы о нормальном распределении генеральной совокупности

На уровне значимости $\alpha = 0.05$ проверим гипотезу H_0 о нормальном распределении генеральной совокупности против конкурирующей гипотезы H_1 о том, что она так не распределена. Используем критерий согласия Пирсона:

$$\chi^2 = \sum \frac{(n_i - n_i')^2}{n_i'}$$

Выборочное среднее: $\bar{X} = \sum \mathrm{x}_\mathrm{i}^* p_i^* = 0{,}05$

Выборочная дисперсия: $S^2 = \sum \mathrm{x}_\mathrm{i}^{*2} p_i^* - \bar{X}^2 = 1{,}0967$

Выборочное отклонение: $\sigma = \sqrt{S^2} = 1.0472$

Так как объём выборки большой, пренебрежём исправлением. Рассчитаем теоретические частоты:

$$n_i' = \frac{h\,n}{\sigma}\ f(z_i) \qquad f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad z_i = \frac{x_i^* - \bar{X}}{\sigma}$$

Номера интервалов	1	2	3	4	5	6	7	8
Границы интервалов	(-2.8; -2.0)	(-2.0; -1.2)	(-1.2; -0.4)	(-0.4; 0.4)	(0.4; 1.2)	(1.2; 2.0)	(2.0; 2.8)	(2.8; 3.6)
x_i^*	-2,414	-1,614	-0,814	-0,014	0,786	1,586	2,386	3,186
n_i	2	7	27	27	22	13	1	1
z_i	-2,353	-1,589	-0,825	-0,061	0,703	1,467	2,231	2,995
$f(z_i)$	0,025	0,113	0,284	0,398	0,312	0,136	0,033	0,005
n_i'	1,914	8,626	21,690	30,428	23,813	10,397	2,532	0,344

Объединим интервалы с малыми x_i^*

n_i	9	27	27	22	15
n_i'	10,540	21,690	30,428	23,813	13,274
$\frac{(n_i - n_i')^2}{n_i'}$	0,2249	1,2998	0,3861	0,1381	0,2245
				$\chi^2_{\text{набл}}$	2,2735

Используем критерий согласия Пирсона

$$\chi^2_{\text{кр}} = \chi^2_{\text{кр}}(0.05; 3) \approx 7.8$$
$$\chi^2_{\text{набл}} < \chi^2_{\text{кр}}$$

Значит на уровне значимости $\alpha = 0.05$ гипотезу H_0 о нормальном распределении генеральной совокупности принимаем.