



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Московский государственный технический университет имени
Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ «Информатика и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

Отчет по лабораторной работе №1 по курсу "Математическая статистика"

Тема Гистограмма и эмпирическая функция распределения

Студент Пересторонин П.Г.

Группа ИУ7-63Б

Преподаватель Саркисян П.С.

Оглавление

1	Задание	2
2	Теоретические сведения	3
2.1	Формулы для вычисления величин	3
2.1.1	Минимальное и максимальное значения выборки . . .	3
2.1.2	Размах выборки	3
2.1.3	Оценки математического ожидания и дисперсии . . .	3
2.2	Определение эмпирической плотности и гистограммы	3
2.3	Определение эмпирической функции распределения	4
3	Результат работы	5
3.1	Код программы	5
3.2	Результаты расчётов	6

1 Задание

Цель работы: построение гистограммы и эмпирической функции распределения.

1. Для выборки объёма n из генеральной совокупности X реализовать в виде программы на ЭВМ
 - (a) вычисление максимального значения M_{\max} и минимального значения M_{\min} ;
 - (b) размаха R выборки;
 - (c) вычисление оценок $\hat{\mu}$ и S^2 математического ожидания MX и дисперсии DX ;
 - (d) группировку значений выборки в $m = [\log_2 n] + 2$ интервала;
 - (e) построение на одной координатной плоскости гистограммы и графика функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием $\hat{\mu}$ и дисперсией S^2 ;
 - (f) построение на другой координатной плоскости графика эмпирической функции распределения и функции распределения нормальной случайной величины с математическим ожиданием $\hat{\mu}$ и дисперсией S^2 .
2. Провести вычисления и построить графики для выборки из индивидуального варианта.

2 Теоретические сведения

2.1 Формулы для вычисления величин

2.1.1 Минимальное и максимальное значения выборки

$$\begin{aligned}M_{\max} &= X_{(n)} \\ M_{\min} &= X_{(1)}\end{aligned}\tag{2.1}$$

2.1.2 Размах выборки

$$R = M_{\max} - M_{\min}.\tag{2.2}$$

2.1.3 Оценки математического ожидания и дисперсии

$$\begin{aligned}\hat{\mu}(\vec{X}_n) &= \frac{1}{n} \sum_{i=1}^n X_i \\ S^2(\vec{X}_n) &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2\end{aligned}\tag{2.3}$$

2.2 Определение эмпирической плотности и гистограммы

Пусть \vec{x} – выборка из генеральной совокупности X . Если объем n этой выборки велик, то значения x_i группируют в интервальный статистический ряд. Для этого отрезок $J = [x_{(1)}, x_{(n)}]$ делят на m равновеликих частей:

$$J_i = [x_{(1)} + (i - 1) \cdot \Delta, x_{(1)} + i \cdot \Delta), i = \overline{1; m - 1}$$

$$J_m = [x_{(1)} + (m - 1) \cdot \Delta, x_{(n)}]$$

$$\Delta = \frac{|J|}{m} = \frac{x_{(n)} - x_{(1)}}{m}$$

Интервальным статистическим рядом называют таблицу:

J_1	\dots	J_i	\dots	J_m
n_1	\dots	n_i	\dots	n_m

где n_i – количество элементов выборки \vec{x} , которые $\in J_i$.

Обычно выборку разбивают на $m = [\log_2 n] + 2$ интервалов, где n – размер выборки.

Гистограмма – это график эмпирической плотности.

Эмпирической плотностью, отвечающей выборке \vec{x} , называют функцию:

$$\hat{f}(x) = \begin{cases} \frac{n_i}{n\Delta}, x \in J_i, i = \overline{1; m} \\ 0, \text{ иначе} \end{cases} \quad (2.4)$$

где J_i – полуинтервал статистического ряда, n_i – количество элементов выборки, входящих в полуинтервал, n – количество элементов выборки.

2.3 Определение эмпирической функции распределения

Пусть $\vec{x} = (x_1, \dots, x_n)$ – выборка из генеральной совокупности X . Обозначим $n(x, \vec{x})$ – число элементов вектора \vec{x} , которые имеют значения меньше x .

Эмпирической функцией распределения называют функцию $F_n : R \rightarrow R$, определенную как:

$$F_n(x) = \frac{n(x, \vec{x})}{n} \quad (2.5)$$

3 Результат работы

3.1 Код программы

```
1 X = [-2.54, -0.79, -4.27, -3.09, -3.82, -0.61, -0.64, -1.24, -1.73, -2.91, -1.48, -1.28,  
    -0.37, -1.88, -2.19, -1.61, -1.52, -3.17, -1.36, -3.08, -3.11, -3.07, -1.57, -1.51,  
    -2.37, -0.58, -3.05, -2.93, -1.01, -1.40, -2.06, -3.05, -1.84, -1.24, -1.89, -2.06,  
    -1.59, -2.83, -1.07, -2.96, -3.17, -3.08, -0.49, -3.11, -3.14, -2.30, -3.99, -1.56,  
    -1.28, -3.46, -2.63, -0.82, -2.18, -0.89, -3.08, -1.13, -1.62, -1.06, -2.98, -1.55,  
    -1.49, -1.65, -1.45, -2.29, -0.85, -1.44, -2.87, -2.40, -2.13, -3.52, -1.42, -3.64,  
    -3.47, -2.05, -2.39, -2.07, -0.80, -1.52, -3.92, -2.22, -0.78, -2.60, -1.78, -1.61,  
    -1.65, -2.06, -3.33, -3.41, -1.97, -1.74, -2.04, 0.01, -1.37, -3.15, -2.35, -3.66,  
    -1.79, -2.56, -1.87, -1.06, -0.64, -2.49, -1.85, -1.40, -0.86, -0.17, -0.62, -2.85,  
    -2.12, -1.17, -2.48, -1.65, -3.74, -2.87, -3.15, -1.89, -1.34, -4.33, -0.96, -1.79];  
2  
3 %  
4 M_max = max(X);  
5 M_min = min(X);  
6  
7 %  
8 R = M_max - M_min;  
9  
10 %  
11 MX = mean(X);  
12 DX = var(X); % sigma == std == sqrt(var(arg))  
13  
14 %  
15 m = floor(log2(length(X))) + 2;  
16 h = histogram(X, m);  
17 %disp(h);  
18  
19 %  
20 sigma = std(X);  
21 x = (M_min - 1):(sigma / 100):(M_max + 1);  
22 f = normpdf(x, MX, sigma); % normal probability distribution function  
23 figure;  
24 heights = h.Values / (sum(h.Values) * h.BinWidth);  
25 centers = [];  
26 for i = 1:(length(h.BinEdges) - 1)  
27     centers = [centers, (h.BinEdges(i + 1) + h.BinEdges(i)) / 2];  
28 end  
29 disp(centers);  
30 hold on;  
31 bar(centers, heights, 1); % :)  
32 plot(x, f, 'g', 'LineWidth', 2);  
33
```

```

34 % )
35 F = normcdf(x, MX, sigma); % normal cumulative distribution function
36 figure;
37 hold on;
38 ecdf(X); % empiric cumulative distribution function
39 plot(x, F, 'r');

```

3.2 Результаты расчётов

$$M_{\min} = -4,33$$

$$M_{\max} = 0,01$$

$$R = 4,34$$

$$\hat{\mu}(\vec{x}_n) = 2,0585$$

$$S^2(\vec{x}_n) = 0,994$$

$$m = 8$$

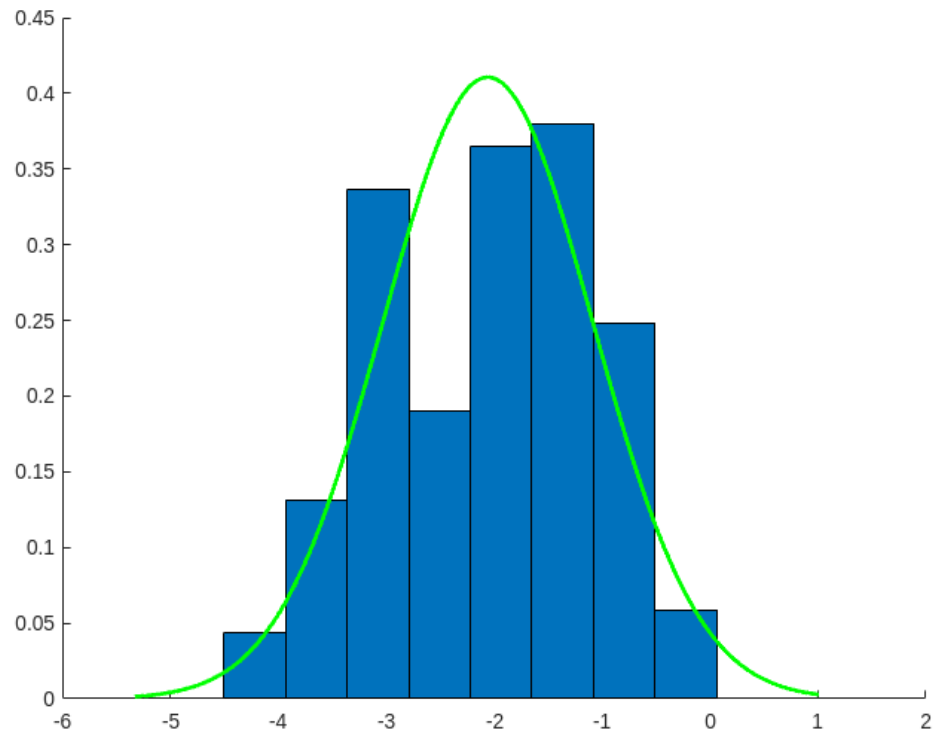


Рис. 3.1: Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с выборочными мат. ожиданием и дисперсией

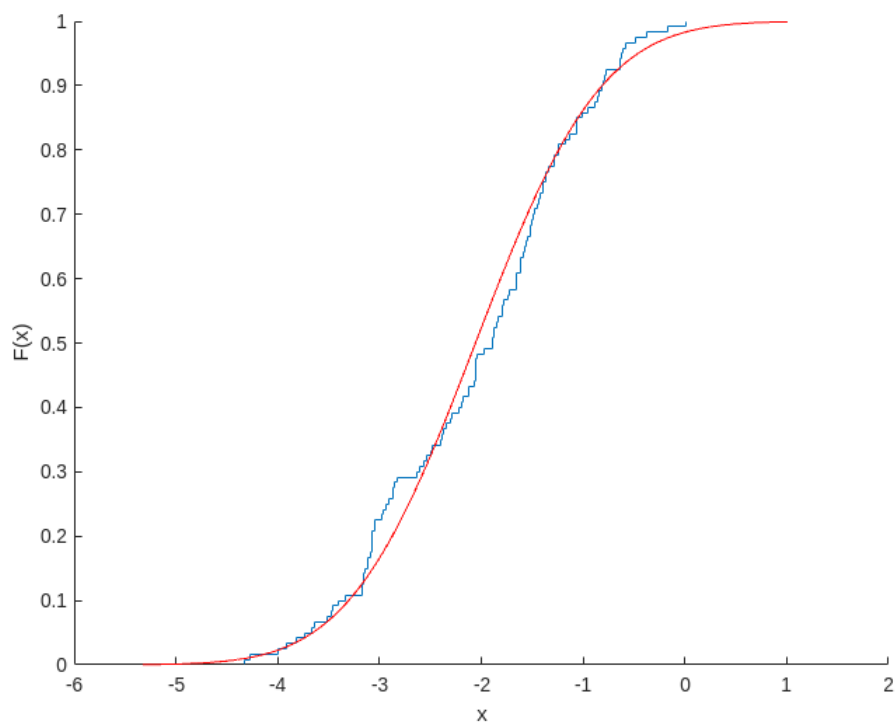


Рис. 3.2: График эмпирической функции распределения и функции распределения нормальной случайной величины с выборочными мат. ожиданием и дисперсией