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 Виноградов А. О.

Группа ИУ7-66Б

Оценка (баллы) ______

Преподаватели Андреева Т. В.

Москва — 2023 г.

1 Постановка задачи

Цель работы: построение гистограммы и эмпирической функции распределения.

1.1 Содержание работы

1) Для выборки объема n из генеральной совокупности X реализовать в виде программы на ЭВМ

    (a) вычисление максимального значения Mmax и минимального значения Mmin;

    (b) размаха R выборки;

    (c) вычисление оценок $\hat{\mu}$ и S^2 математического ожидания MX и дисперсии DX;

    (d) группировку значений выборки в m = [log2 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 Теоретическая часть

2.1 Формулы для вычисления величин

Реализация случайной выборки:

$$\overrightarrow{x} = (x_1, x_2, \ldots, x_n). \tag{2.1}$$

Минимальное значение выборки:

$$M_{min} = min(x_1, x_2, \ldots, x_n). \tag{2.2}$$

Максимальное значение выборки:

$$M_{max} = max(x_1, x_2, \ldots, x_n). \tag{2.3}$$

Размах выборки:

$$R = M_{max} - M_{min}. \tag{2.4}$$

Оценка математического ожидания выборки:

$$\hat{\mu}(\overrightarrow{x}) = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{2.5}$$

Несмещенная(исправленная) оценка дисперсии выборки:

$$S^2(\overrightarrow{x}) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2. \tag{2.6}$$

2.2 Эмпирическая плотность и гистограмма

Опредление 1. *Эмпирической функцией плотности, отвечающей выборке $\overline{x}$, называют функцию*

$$f_n(x) = \begin{cases} \frac{m_i}{n\delta}, \textit{если } x \in J_i, \\ 0, \textit{иначе}, \end{cases} \tag{2.7}$$

где

- m — количество полуинтервалов интервала J = $[M_{min}, M_{max}]$, $\overline{x}$, которые приняли значение меньше x,
- m_i — количество элементов выборки $\overline{x}$, принадлежащих полуинтервалу J_i, $i = \overline{1, m}$,
- n — количество элементов выборки $\overline{x}$,
- $\delta = \frac{M_{min} - M_{max}}{m} = \frac{|J|}{m}$.

Опредление 2. *График функции $f_n(x)$ называют гистограммой.*

2.3 Эмпирическая функция распределения

Опредление 3. *Эмпирической функцией распределения, отвечающей выборке $\overline{x}$, называют функцию*

$$F_n(x, \overline{x}) = \frac{n(x, \overline{x})}{n}, x \in \mathbb{R}, \tag{2.8}$$

где

- $n(x, \overline{x})$ — количество элементов выборки $\overline{x}$, которые приняли значение меньше x,
- n — количество элементов выборки.

3 Практическая часть

3.1 Текст программы

Листинг 3.1 – Текст программы

```
f = fopen('file.txt','r');
% f = fopen('test.txt', 'r');
X = fscanf(f, '%f,');
Z = sort(X);
fclose(f);
n = length(X);


M_min = min(X);
M_max = max(X);
fprintf('Mmin = %f\n', M_min);
fprintf('Mmax = %f\n', M_max);


R = M_max - M_min;
fprintf('Range = %f\n', R);


MX = mean(X);
DX = sum((X - MX).^2) /(n - 1);
fprintf('Mu = %f\n', MX);
fprintf('S2 = %f\n', DX);


m = floor(log2(n))+ 2;
h = R / m;
intervals = cell(1, m);
i = 1;
fprintf('m = %d, delta = %f\n', m, h);
bracket = ') ';
```

```
format short;
for cur = (M_min):h:(M_max-h)
    next = cur + h;
    intervals(i) = {X((cur <= X) & (X < next))};
    fprintf('[%5.2f, %5.2f%s', cur, next,bracket);
    i = i + 1;
    if i == m
        bracket = ']';
    end
end
intervals{m} = [intervals{m}; X(X == M_max)];
intervals(1);
cellsz = cellfun(@size,intervals,'uni',false);
cellsz{5}(1);

fprintf('\n');
s = size(cellsz);
for i = 1:s(2)
    fprintf('%8d        ', cellsz{i}(1));
end
fprintf('\n');

figure('Name', 'Graph 1');
histogram(X, m, 'BinEdges', M_min:h:M_max, 'Normalization', 'pdf');
hold on;
x = (M_min - h):0.1:(M_max+h);
% f = exp(-(x-MX).^2./(2*DX))./(sqrt(DX * 2 * pi));
Xn = (M_min - h): 0.1 :(M_max + h);
y = pdf('normal', Xn, MX, DX);
plot(Xn,y), grid;


figure('Name', 'Graph 2');
[yy, xx] = ecdf(X);
stairs(xx, yy,'LineWidth',1.5);
hold on;
Y = normcdf((Xn - MX) / DX);
plot(Xn, Y, 'r'), grid;
```

3.2 Результаты расчетов для выборки из индивидуального варианта

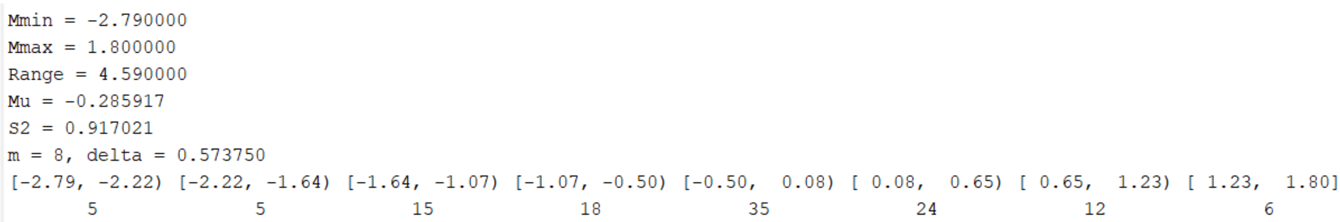
Задание выполнялось по варианту №2.

Таблица 3.1 – Значения параметров для выборки из индивидуального варианта

M_{min}	-2.79
M_{max}	1.8
R	4.59
$\hat{\mu}(\overrightarrow{x})$	-0.285917
$S^2(\overrightarrow{x})$	0.917021
m	8
δ	0.57375

Таблица 3.2 – Интервальная группировка значений выборки при m = 8

[-2.79, -2.22)	[-2.22, -1.64)	[-1.64, -1.07)	[-1.07, -0.50)
5	5	15	18
[-0.50, 0.08)	[0.08, 0.65)	[0.65, 1.23)	[1.23, 1.80]
35	24	12	6



```
Mmin = -2.790000
Mmax = 1.800000
Range = 4.590000
Mu = -0.285917
S2 = 0.917021
m = 8, delta = 0.573750
[-2.79, -2.22) [-2.22, -1.64) [-1.64, -1.07) [-1.07, -0.50) [-0.50,  0.08) [ 0.08,  0.65) [ 0.65,  1.23) [ 1.23,  1.80]
      5              5             15             18             35             24             12              6
```

Рисунок 3.1 – Результат работы программы

Результаты построения гистограммы и графика функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием $\hat{\mu}$ и дисперсией S^2, а также построения графика эмпирической функции распределения и функции распределения нормальной случайной величины с математическим ожиданием $\hat{\mu}$ и дисперсией S^2 приведены на рисунках 3.2, 3.3.

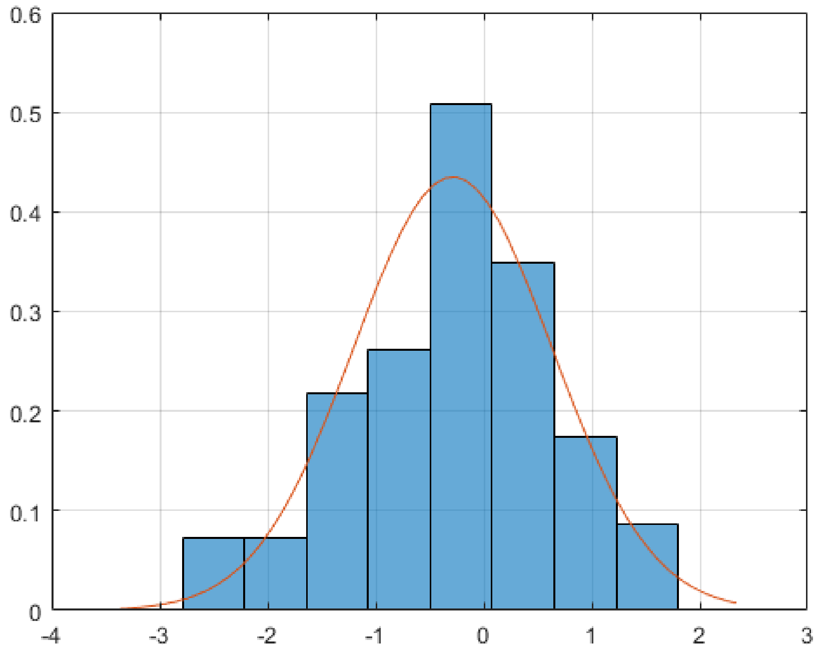


Рисунок 3.2 – Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием $\hat{\mu}$ и дисперсией S^2

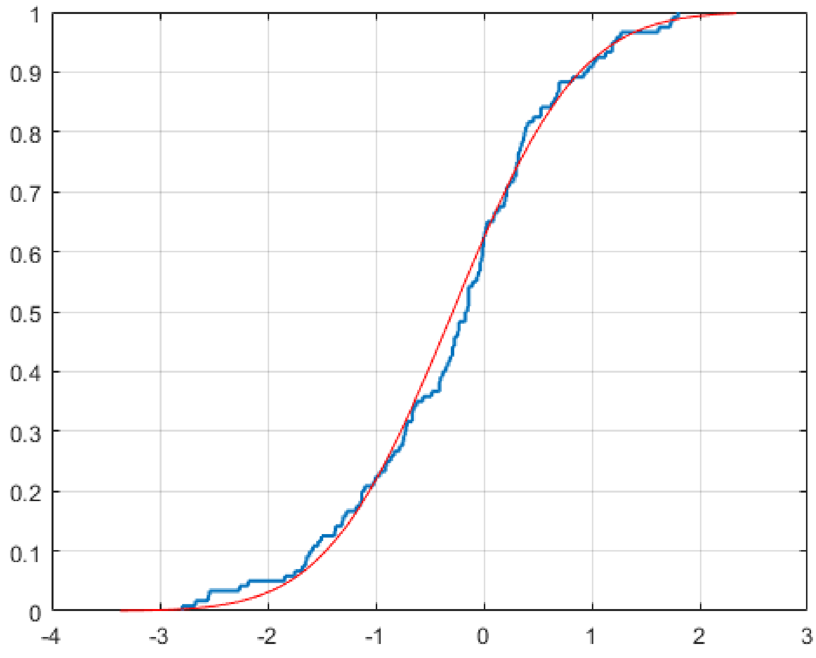


Рисунок 3.3 – График эмпирической функции распределения и функции распределения нормальной случайной величины с математическим ожиданием $\hat{\mu}$ и дисперсией S^2