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»

Вариант № 13

Студент ИУ7-63Б (Группа) ________ (Подпись, дата) Миронов Г. А. (И. О. Фамилия)

Преподаватель ________ (Подпись, дата) Власов П. А. (И. О. Фамилия)

2022 г.

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 Вариационный ряд

Вариационным рядом отвечающим выборке $\overline{x}$ объема n называется вектор случайных величин

$$\vec{x} = (x_{(1)}, \ldots, x_{(n)}) \tag{2.1}$$

полученный из вектора $\overline{x}$ путем упорядочивания его компонентов в порядке неубывания, т.е.

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)} \tag{2.2}$$

Вариационным рядом отвечающим случайной выборке $\overline{X}$ объема n называется вектор случайных величин

$$\vec{X} = (X_{(1)}, \ldots, X(n)) \tag{2.3}$$

Где $X_{(i)}$ - случайная величина, которая для каждой реализации x случайного вектора $\vec{X}$ принимает значегие равное i-му члену вариационного ряда построенного по выборке $\vec{x}$.

2.1.2 Минимальное и максимальное значения выборки

$$\begin{aligned} M_{\max} &= X_{(n)} \\ M_{\min} &= X_{(1)} \end{aligned} \tag{2.4}$$

2.1.3 Размах выборки

$$R = M_{\max} - M_{\min}. \tag{2.5}$$

2.1.4 Оценки математического ожидания и дисперсии

$$\begin{aligned} \hat{\mu}(\vec{X}_n) &= \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X} \\ S^2(\vec{X}_n) &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 \end{aligned} \tag{2.6}$$

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	...	J_i	...	J_m
n_1	...	n_i	...	n_m

где n_i – количество элементов выборки $\vec{x}$, принадлежащих 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:

$$\hat{f}(x) = \begin{cases} \frac{n_i}{n\Delta}, x \in J_i, i = \overline{1; m} \\ 0, \text{иначе} \end{cases} \tag{2.7}$$

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 $\vec{x} = (x_1, ..., x_n)$ – 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 $F_n : \mathbb{R} \to \mathbb{R}$, определенную как:

$$F_n(x) = \frac{n(x, \vec{x})}{n} \tag{2.8}$$

3 Результат работы

3.1 Код программы

Листинг 3.1 – Файл main.m.

```
close all;
clear all;
clc;

pkg load statistics;

X = [-10.82,-9.27,-9.65,-9.36,-9.27,-11.25,-9.89,-9.26,-11.15,
     -8.90,-11.02,-8.28,-9.18,-10.16,-10.59,-10.82,-9.05,-9.47,
     -10.98,-11.50,-8.64,-10.86,-10.76,-11.49,-11.09,-9.33,-9.32,
     -9.66,-8.79,-10.54,-9.12,-10.40,-8.59,-10.22,-9.06,-10.59,
     -10.60,-10.25,-9.35,-11.44,-9.81,-9.32,-9.95,-9.33,-10.64,
     -9.45,-10.99,-10.15,-10.39,-10.36,-10.49,-11.67,-10.00,
     -10.87, -11.11,-9.68,-10.77,-9.13,-8.62,-10.33,-11.36,
     -10.24,-9.41,-11.05,-10.15,-9.35,-11.45,-9.87,-10.41,-10.11,
     -10.84,-11.48,-7.77,-10.79,-9.88,-10.70,-9.07,-9.47,-10.15,
     -9.93,-11.52,-9.04,-10.93,-10.13,-9.56,-11.39,-9.79,-9.19,
     -11.09,-9.86,-10.67,-10.26,-9.07,-10.53,-11.24,-10.16,
     -11.33,-8.76,-8.88,-10.53,-10.12,-8.98,-9.84,-9.90,-10.13,
     -9.32,-9.31,-9.99,-8.55,-11.64,-11.32,-10.51,-11.71,-10.50,
     -10.50,-12.20,-11.68,-10.45,-7.88,-10.84];

% а)
M_max = max(X);
%disp ("The value of Mmax is:"), disp(M_max);
M_min = min(X);
%disp ("The value of Mmin is:"), disp(M_min);

% б)
R = M_max - M_min;
%disp ("The value of R is:"), disp(R);



% в)
MX = mean(X);
DX = var(X); % sigma == std == sqrt(var(arg))

% г)
m = floor(log2(length(X))) + 2;
[nn, xx] = hist(X, m);
nn = nn ./ sum(nn);
```

Листинг 3.2 – Файл `main.m`. Часть 2.

```
% д)
figure;
bar(xx, nn);
hold on;
% normal probability distribution function
sigma = std(X);
x = (M_min - 1):(sigma / 100):(M_max + 1);
plot(x, normpdf(x, MX, sigma), 'g', 'LineWidth', 2);

% е)
figure;
% empirical cumulative distribution function
plot(x, empirical_cdf(x, X), 'r');
hold on;
% normal cumulative distribution function
plot(x, normcdf(x, MX, sigma), 'g', 'LineWidth', 2);
```

3.2 Результаты расчётов

$$M_{\min} = -12,20$$

$$M_{\max} = -7,77$$

$$R = 4,43$$

$$\hat{\mu}(\vec{x}_n) = -10.132$$

$$S^2(\vec{x}_n) = 0.8460$$

$$m = 8$$

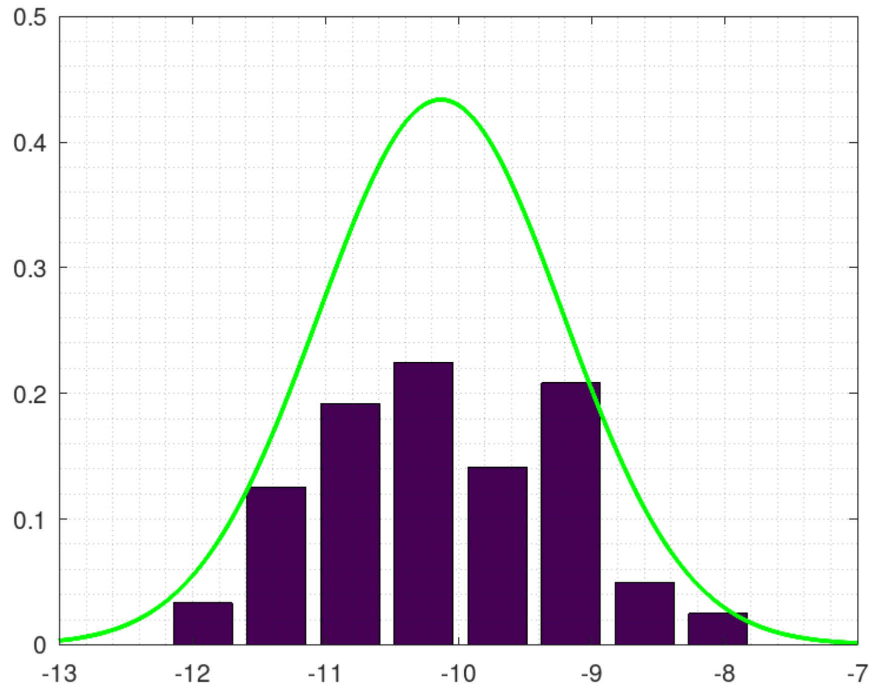

Рисунок 3.1 – Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с выборочными мат. ожиданием и дисперсией

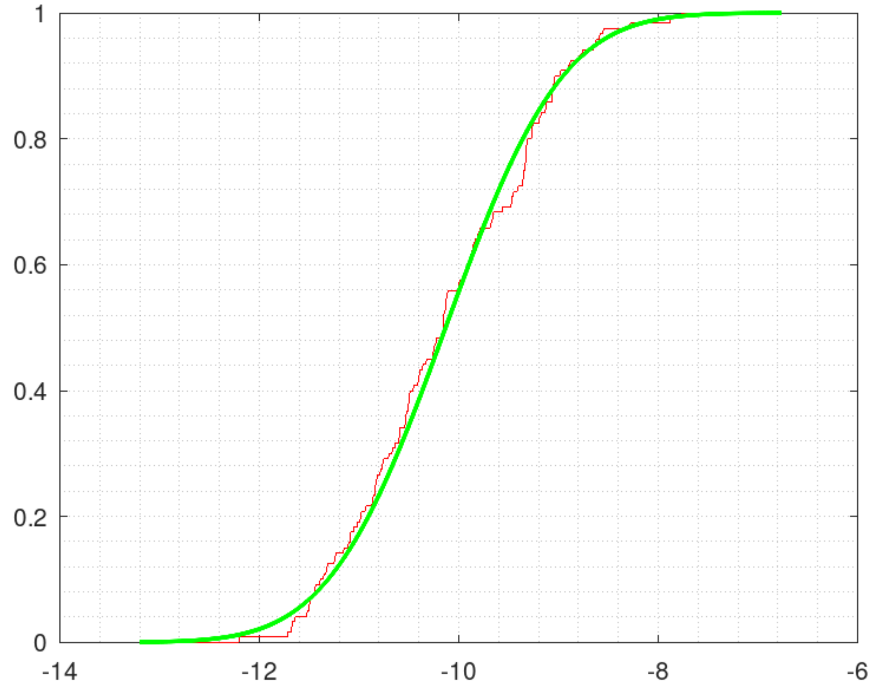

Рисунок 3.2 – График эмперической функции распределения и функции распределения нормальной случайной величины с выборочными мат. ожиданием и дисперсией