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»

на тему: «Гистограмма и эмпирическая функция распределения»

Вариант № 9

Студент ИУ7-66Б (Группа) ______ (Подпись, дата) Жаворонкова А. А. (И. О. Фамилия)

Преподаватель ______ (Подпись, дата) Саркисян П. С. (И. О. Фамилия)

2024 г.

Формулы

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

где $M_{\max}$ — максимальное значение выборки, $M_{\min}$ — минимальное значение выборки, R — размах выборки, $\hat{\mu}(\vec{X}_n)$ — выборочное среднее, $S^2(\vec{X}_n)$ — исправленная выборочная дисперсия.

Определения

Опр. Эмпирической функцией распределения называется отображение $F_n : R \to R$, заданное формулой:

$$F_n(t) = \frac{l(t, \vec{x})}{n},$$

где n — объем выборки, $l(t, \vec{x})$ — число элементов $\vec{x}$, которые меньше $t (t \in R)$

Опр. Эмпирической плотностью распределения называют функцию:

$$f_n(t) = \begin{cases} \frac{n_i}{n\Delta}, t \in J_i, i \in [1; m] \\ 0, t \notin J \end{cases}$$

где $J = [x_{(1)}, x_{(n)}]$ — отрезок, разбиваемый на m частей, n_i — число элементов реализации, попавших в промежуток J_i.

Опр. Гистограммой называется график функции $f_n(t)$.

Текст программы

```
function lab_1()
X = csvread('X.csv');
% a)
[Mmax, Mmin] = MinMax(X);
% )
```

```
R = range(Mmax, Mmin);
% )
[mu, s2] = estimations(X);
% )
interval_row = grouping(X);
% )
graph1(X, length(interval_row), mu, s2);
% )
graph2(X, mu, s2);
end

function [Mmax, Mmin] = MinMax(X)
fprintf("a)    Mmax    Mmin\n");
Mmax = max(X);
Mmin = min(X);
fprintf("Mmax = %f\n", Mmax);
fprintf("Mmin = %f\n", Mmin);
end

function R = range(Max, Min)
fprintf(")   R\n");
R = Max - Min;
fprintf("R = %f\n", R);
end

function [mu, s2] = estimations(X)
fprintf(")   mu  s2   MX   DX\n");
n = length(X);
mu = sum(X) / n;
s2 = sum((X - mu).^2) / (n - 1);
fprintf("mu = %f\n", mu);
fprintf("s2 = %f\n", s2);
end

function interval_row = grouping(X)
fprintf(")     m = [log2 n] + 2 \n");
n = length(X);
m = floor(log2(n)) + 2;
fprintf("  m = %d\n", m);
X_sorted = sort(X);
delta = (X_sorted(n) - X_sorted(1)) / m;

interval_row = zeros(m, 1);
```

```
for i = 1 : m
    cnt = 0;
    for x = X
        if ((x >= X_sorted(1) + (i - 1) * delta) && (x <
           X_sorted(1) + i * delta)) || ((i == m) && (x >=
           X_sorted(1) + (m - 1) * delta) && (x <= X_sorted(n)))
            cnt = cnt + 1;
        end
    end

    if (i == m)
        fprintf(" %d. [%f; %f], - : %d\n", i, X_sorted(1) + (m -
           1) * delta, X_sorted(n), cnt);
    else
        fprintf(" %d. [%f; %f), - : %d\n", i, X_sorted(1) + (i -
           1) * delta, X_sorted(1) + i * delta, cnt);
    end

    interval_row(i) = cnt;
end
end

function graph1(X, m, mu, s2)
fprintf(")               mu   s2\n");
figure();
h = histogram(X, m, 'Normalization', 'pdf');
delta = (max(X) - min(X)) / m;
h.BinEdges = min(X):delta:max(X);
hold on;

xlist = min(X):.5:max(X);
ylist = pdf('Normal', xlist, mu, s2);
plot(xlist, ylist);
hold off;
end

function graph2(X, mu, s2)
fprintf(")               mu  s2\n");
figure();
ecdf(X);
hold on;

xlist = min(X):.5:max(X);
```

```
ylist = cdf('Normal', xlist, mu, s2);
plot(xlist, ylist);
hold off;
end
```

Результаты расчетов

```
>> lab_1
а) Вычисление максимального значения Mmax и минимального значения Mmin
Mmax = -5.200000
Mmin = -10.110000
б) Вычисление размаха R
R = 4.910000
в) вычисление оценок mu и s2 математического ожидания MX и дисперсии DX
mu = -7.660917
s2 = 0.777892
г) Группировка значений выборки в m = [log2 n] + 2 интервала
Количество интервалов m = 8
 1. [-10.110000; -9.496250), кол-во элементов: 1
 2. [-9.496250; -8.882500), кол-во элементов: 10
 3. [-8.882500; -8.268750), кол-во элементов: 18
 4. [-8.268750; -7.655000), кол-во элементов: 32
 5. [-7.655000; -7.041250), кол-во элементов: 30
 6. [-7.041250; -6.427500), кол-во элементов: 18
 7. [-6.427500; -5.813750), кол-во элементов: 10
 8. [-5.813750; -5.200000], кол-во элементов: 1
```

Рисунок 1 – Результаты расчетов для выборки из индивидуального варианта

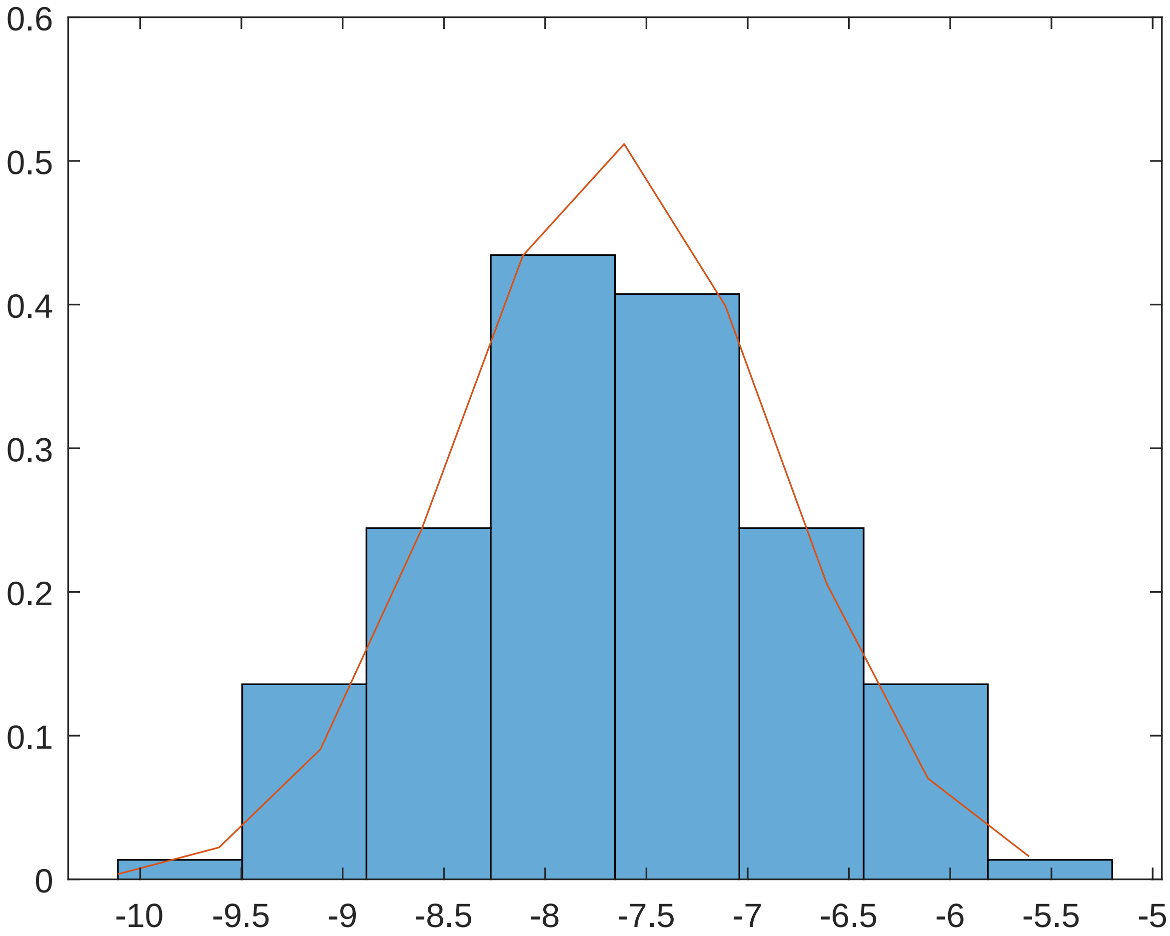

Рисунок 2 – Гистограмма и график функции плотности распределения нормальной случайной величины с выборочными математическим ожиданием и дисперсией

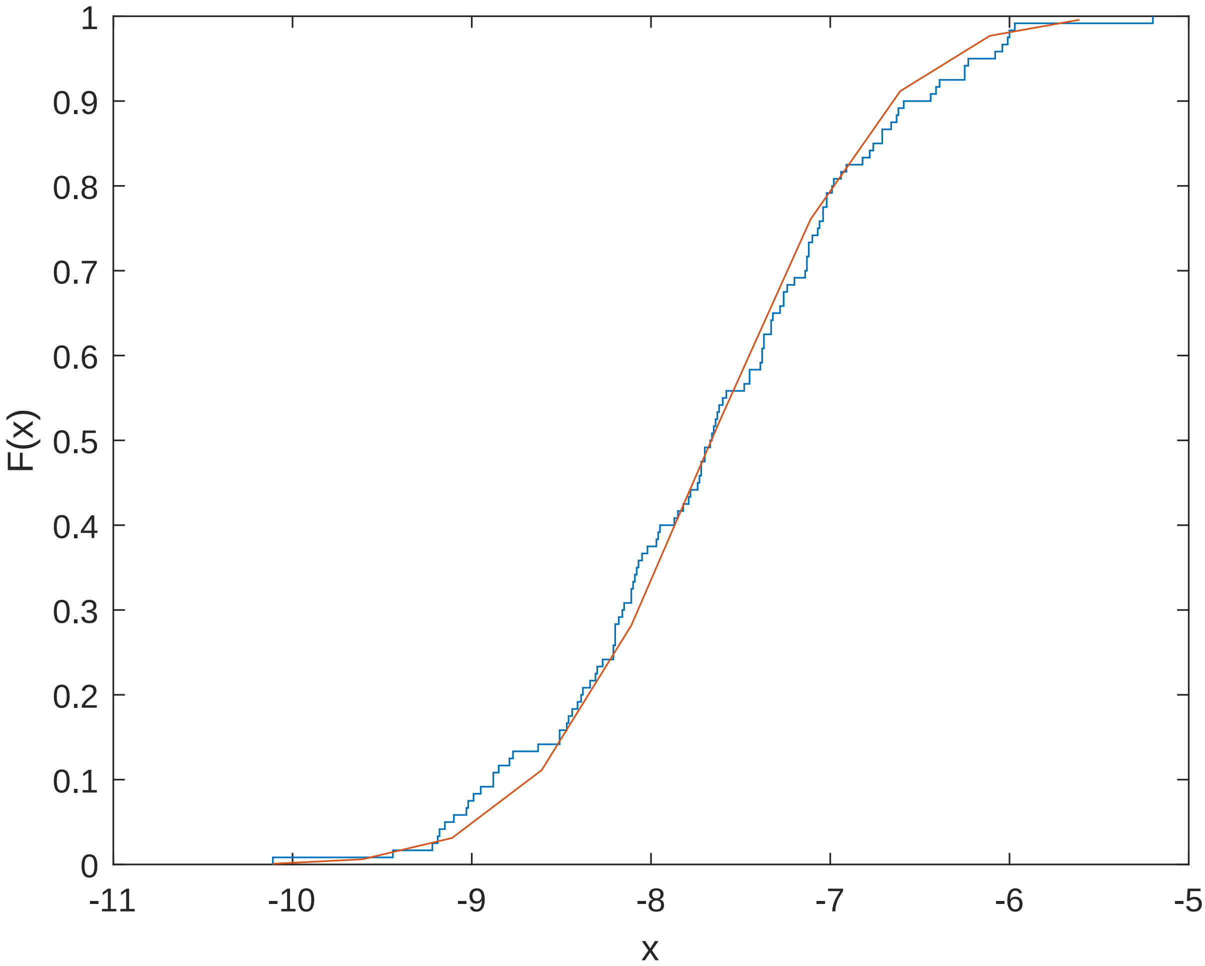

Рисунок 3 – График эмперической функции распределения и функции распределения нормальной случайной величины с выборочными математическим ожиданием и дисперсией