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»

Лабораторная работа № __1__

Дисциплина Математическая статистика

Тема Гистограмма и эмпирическая функция распределения

Вариант №5

Студент _Брянская Е.В.________

Группа _ИУ7-62Б_____________

Оценка (баллы) _______________

Преподаватель _Саркисян П.С.

Москва.
2021 г.

Оглавление

Введение

Цель работы: построение гистограммы и эмпирической функции распределения. **Содержание работы:**

1. Для выборки объёма n из генеральной совокупности M реализовать в виде программы на ЭВМ

   а) вычисление максимального значения M_{max} и минимального значения M_{min};

   б) размаха R выборки;

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

   г) группировку значений выборки в $\mathrm{m} = [log_2 n\ ] + 2$ интервала;

   д) построение на одной координатной плоскости гистограммы и графика функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием $\hat{\mu}$ и дисперсией S^2;

   е) 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.

1. Теоретическая часть

1.1. Формулы для вычисления величин

Выборка: $\vec{x} = (x_1, ..., x_n)$

1.1.1 Максимальное значение выборки

$$M_{max} = max\{x_1, ..., x_n\}$$

1.1.2 Минимальное значение выборки

$$M_{min} = min\{x_1, ..., x_n\}$$

1.1.3 Размах выборки

$$R = M_{max} - M_{min}$$

1.1.4 Выборочное среднее

$$\hat{\mu}(\vec{x}_n) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

1.1.5 Несмещённая оценка дисперсии (состоятельная оценка)

$$S^2(\vec{x}_n) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x}_n)^2$$

1.2. Эмпирическая плотность и гистограмма

Интервальным статистическим рядом называют таблицу:

J_1	...	J_i	...	J_m
n_1	...	n_i	...	n_m

здесь n_i - число элементов выборки $\vec{x}$, которые попали в J_i.

Пусть для выборки $\vec{x}$ построен интервальный статистический ряд, **эмпирической плотностью** называется функция:

$$\hat{f}(x) = \begin{cases} \frac{n_i}{n\Delta}, & x \in J_i; \\ 0, & x \notin J. \end{cases}$$

График импирической плотности называется **гистограммой**.

1.3. Эмпирическая функция распределения

Эмпирической функцией распределения, отвечающей выборке $\vec{x}$, называют функцию:

$$\hat{F}(x) = \frac{n(x, \vec{x})}{n},$$

где $n(x, \vec{x})$ - число элементов вектора $\vec{x}$, которые имеют значение меньше, чем x; n — объём выборки.

2. Практическая часть

2.1. Текст программы

```
function lab_01()
    x=
    [-4.58,-5.10,-4.24,-4.82,-6.05,-4.05,-4.48,-4.65,-3.67,-4.01,-3.22,
  -5.79,-4.20,-5.36,-5.16,-4.31,-3.91,-4.04,-5.65,-5.78,-4.03,-5.15,
  -4.89,-4.60,-5.71,-5.67,-4.40,-5.10,-5.29,-3.90,-3.77,-5.47,-5.57,
  -5.91,-3.62,-2.45,-3.13,-5.60,-4.35,-3.36,-3.87,-4.78,-5.72,-4.66,
  -4.34,-4.60,-5.24,-4.43,-5.15,-5.45,-3.22,-4.61,-5.65,-5.63,-4.73,
  -3.82,-4.34,-4.98,-6.43,-4.25,-4.66,-5.49,-4.98,-6.10,-4.44,-5.25,
  -4.89,-2.97,-7.26,-4.14,-4.45,-5.93,-3.49,-3.96,-5.66,-5.04,-5.40,
  -4.89,-4.61,-4.01,-5.44,-2.76,-4.97,-3.87,-4.33,-5.28,-5.63,-3.94,
  -4.56,-4.67,-5.35,-4.89,-3.79,-6.25,-4.38,-4.28,-4.48,-4.16,-5.67,
  -5.34,-4.78,-4.33,-4.89,-3.16,-4.88,-5.40,-4.64,-4.84,-6.72,-5.25,
  -3.27,-6.44,-3.49,-5.92,-5.21,-3.88,-5.08,-3.50,-6.80,-5.87];
    x = sort(x);

    % Задание а1 - максимальноеминимальное/ значение
    Mmax = max(x);
    fprintf('Mmax = %.2f\n', Mmax);

    Mmin = min(x);
    fprintf('Mmin = %.2f\n', Mmin);

    % Задание б1 - размах R выборки
    R = Mmax - Mmin;
    fprintf('R = %.2f\n', R);

    % Задание в1 - вычисление оценок μ^ и S2 для MX и DX
    n = length(x);
    mu = sum(x)/n;
    fprintf('mu = %.2f\n', mu);

    S2 = sum((x - mu).^2) / (n - 1);
    fprintf('S^2 = %.2f\n', S2);

    % Задание г1 - группировку значений выборки в m = [log2 n] + 2 интервала
    m = floor(log2(n)) + 2;
```

```
    fprintf('m = %.2f\n', m);

    intervals = zeros(1, m + 1);
    count = zeros(1, m + 1);
    delta = (Mmax - Mmin) / m;

    for i = 1:m + 1
        intervals(i) = x(1) + delta * (i - 1);
    end

    for i = 1:length(x)
        for j = 1:m
            if x(i) >= intervals(j) && x(i) < intervals(j + 1)
                count(j) = count(j) + 1;
                break;
            end
        end
    end
    count(m) = count(m) + 1;

    for i = 1:m
        if i < m
            fprintf("[ %.2f ; %.2f ) - %d elements\n", intervals(i), ...
                intervals(i + 1), count(i));
        else
            fprintf("[ %.2f ; %.2f ] - %d elements\n", intervals(i), ...
                intervals(i + 1), count(i));
        end
    end

    % Задание д1 - гистограмма + график функции плотности распределения
    % вероятностей н с в с мат. ож. и дисп.
    for i = 1:m+1
        count(i) = count(i) / (length(x)*delta);
    end

    stairs([intervals(1), intervals], [0 count], 'b');
    grid on;
    hold on;

```

```
    x_mod = (Mmin):(R/length(x)):(Mmax);
    f = normpdf(x_mod, mu, sqrt(S2)); % Фя— плотности распря— нсв...
    plot(x_mod, f, 'r—');

    legend('Гистограмма', 'Фя— плотности распря— нсв...', 'Location','
   northwest');
    hold off;

    % Задание e1 — график эмпирической функции распределения и функции
    % распределения нормальной случайной величины с мат ожиданием и дисперсией
    figure;

    [f, xx] = ecdf(x); % эмпирическая функция
    stairs(xx, f), grid;
    hold on;

    F = normcdf(x_mod, mu, sqrt(S2)); % функция распределения н с в
  plot(x_mod, F, 'r—');

    legend('Эмпирическая функция распределения', 'Функция распределения', '
   Location','northwest');
    hold off;
end
```

2.2. Результат работы программы

$$
\begin{aligned}
M_{\max} &= -2.45; \\
M_{\min} &= -7.26; \\
R &= 4.81; \\
\hat{\mu}(\vec{x}_n) &= -4.76; \\
S^2(\vec{x}_n) &= 0.81; \\
m &= 8; \\
[-7.26; -6.66) &- 3\, elements \\
[-6.66; -6.06) &- 4\, elements \\
[-6.06; -5.46) &- 20\, elements \\
[-5.46; -4.86) &- 29\, elements \\
[-4.86; -4.25) &- 30\, elements \\
[-4.25; -3.65) &- 21\, elements \\
[-3.65; -3.05) &- 10\, elements \\
[-3.05; -2.45] &- 3\, elements
\end{aligned}
$$

2.3. Графики

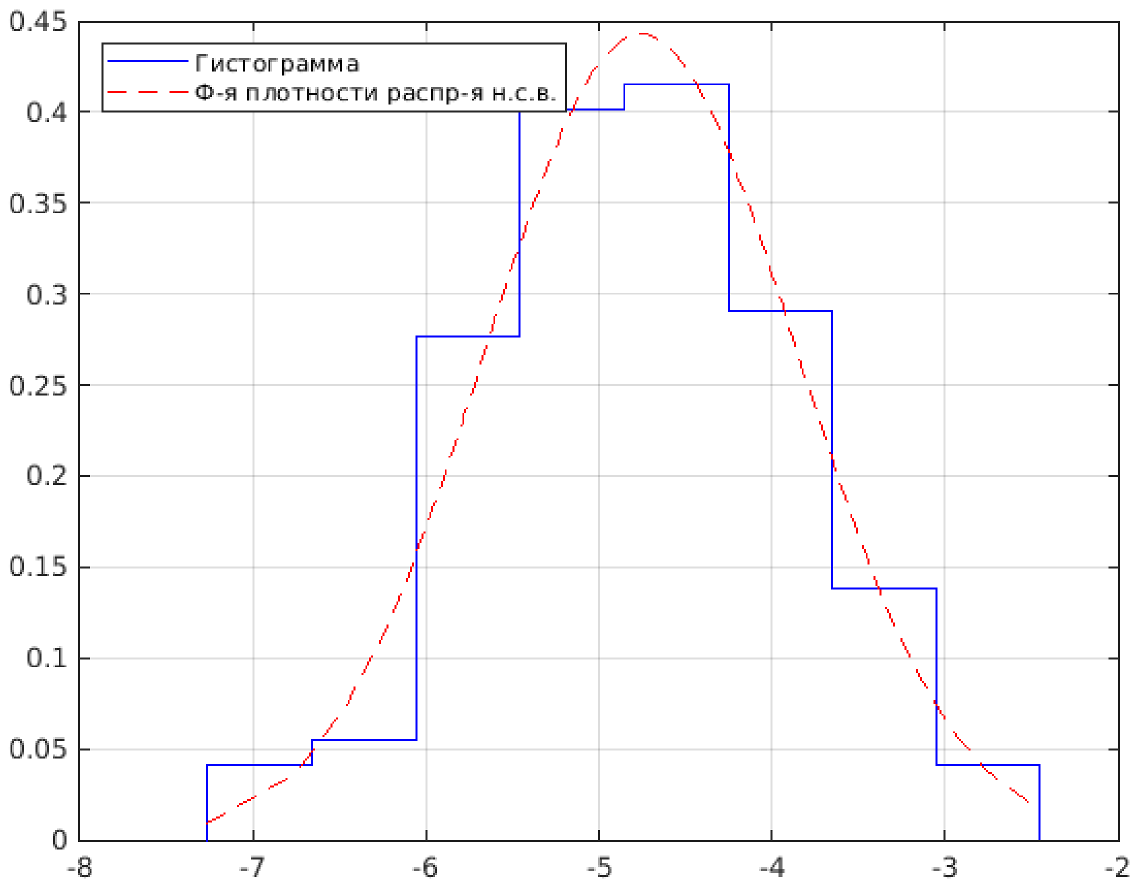

Рис. 2.1 — Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины

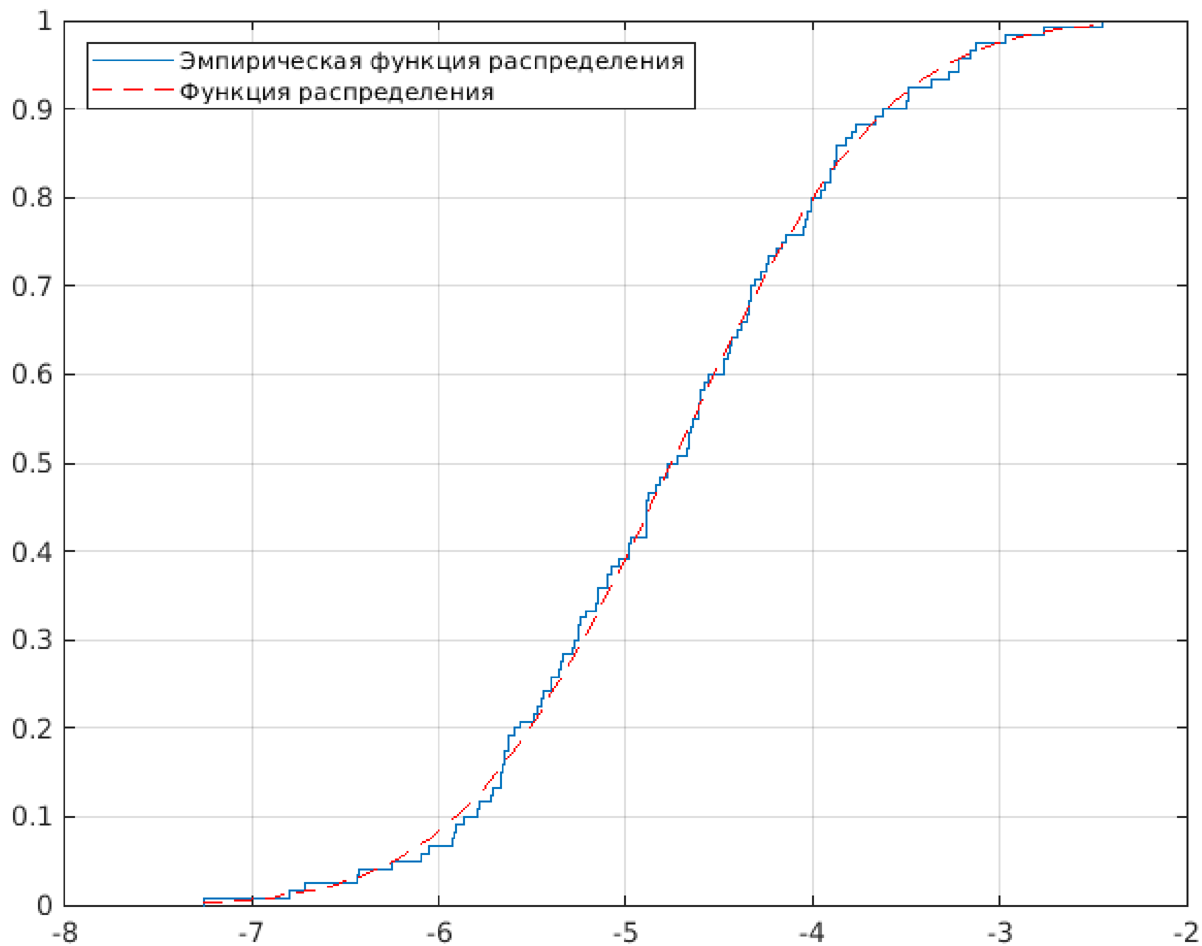

Рис. 2.2 — Эмпирическая функция распределения и функция распределения нормальной случайной величины