**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»

«Гистограмма и эмпирическая функции распределения»

Студент Маслова Марина Дмитриевна

Группа ИУ7-63Б

Оценка (баллы) ____________________

Преподаватель Власов Павел Александрович

2022 г.

# 1 Задание

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

**Содержание работы**

1. Для выборки объема 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.

**Содержание отчета**

1. формулы для вычисления величин $M_{max}$, $M_{min}$, $R$, $\hat{\mu}$, $S^2$;
2. определение эмпирической плотности и гистограммы;
3. определение эмпирической функции распределения;
4. текст программы;
5. результаты расчетов для выборки1 из индивидуального варианта.

# 2 Теоретическая часть

## 2.1 Формулы для вычисления величин

Пусть $\vec{x} = (x_1, ..., x_n)$ — выборка объема $n$ из генеральной совокупности $X$.

Максимальное и минимальное значения выборки:

$$M_{max} = x_{(n)} = \max\{x_1, ..., x_n\}, \tag{2.1}$$

$$M_{min} = x_{(1)} = \min\{x_1, ..., x_n\}, \tag{2.2}$$

где $x_{(1)}, x_{(n)}$ — крайние члены вариационного ряда, отвечающего выборке $\vec{x}$.

Размах выборки:

$$R = M_{max} - M_{min}. \tag{2.3}$$

Оценка математического ожидания — выборочное среднее:

$$\hat{\mu}(\vec{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{2.4}$$

Оценка дисперсии — исправленная выборочная дисперсия:

$$S^2(\vec{x}) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2, \tag{2.5}$$

где $\overline{x} = \hat{\mu}$.

## 2.2 Определение эмпирической плотности и гистограммы

Если объем выборки достаточно велик ($n > 50$), то элементы выборки группируют в так называемый интервальный статистический ряд. Для этого отрезок $J = [x_{(1)}, x_{(n)}]$ разбивают на $m$ равновеликих промежутков. Ширина каждого из них:

$$\Delta = \frac{|J|}{m} = \frac{x_{(n)} - x_{(1)}}{m}. \tag{2.6}$$

Далее полагают:

$$J_i = [x_{(1)} + (i-1)\Delta; x_{(1)} + i\Delta], \;\; i = \overline{1, m-1},$$
$$J_m = [x_{(1)} + (m-1)\Delta; x_{(n)}]. \tag{2.7}$$

**Определение.** Интервальным статистическим рядом, отвечающим выборке $\vec{x}$, называется таблица вида:

| | | | | |
|---|---|---|---|---|
| $J_1$ | ... | $J_i$ | ... | $J_m$ |
| $n_1$ | ... | $n_i$ | ... | $n_m$ |

Здесь $n_i$ — число элементов выборки $\vec{x}$, попавших в промежуток $J_i$, $i = \overline{1, m}$.

Количество интервалов определяется формулой:

$$m = [\log_2 n] + 2. \tag{2.8}$$

Пусть для данной выборки $\vec{x}$ построен интервальный статистический ряд $(J_i, n_i), i = \overline{1, m}$.

**Определение.** Эмпирической плотностью распределения (соответствующей выборке $\vec{x}$) называется функция:

$$f(x) = \begin{cases} \frac{n_i}{n\Delta}, & x \in J_i, \; i = \overline{1, m}, \\ 0, & \text{иначе.} \end{cases} \tag{2.9}$$

**Определение.** График эмпирической функции плотности называется гистограммой.

## 2.3 Определение эмпирической функции распределения

Пусть $\vec{x} = (x_1, ..., x_n)$ — выборка из генеральной совокупности $X$.

Обозначим $n(t, \vec{x})$ — число компонент вектора $\vec{x}$, которые меньше, чем $t$.

**Определение.** Эмпирической функцией распределения, построенной по выборке $\vec{x}$, называют функцию $F_n : \mathbb{R} \to \mathbb{R}$, определенную правилом:

$$F_n(t) = \frac{n(t, \vec{x})}{n}. \tag{2.10}$$

# 3 Практическая часть