



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 дисциплине "Математическая статистика"

Тема Гистограмма и эмпирическая функция распределения

Студент Малышев И. А.

Группа ИУ7-61Б

Оценка (баллы) _____

Преподаватель: Власов П. А.

Москва — 2022 г.

Задание

Цель работы

Построение гистограммы и эмпирической функции распределения.

Постановка задачи

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.

Теоретические сведения

Формулы для вычисления величин

Минимальное и максимальное значения выборки

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

Размах выборки

$$R = M_{\max} - M_{\min}. \quad (2)$$

Оценки выборочного математического ожидания и исправленной выборочной дисперсии

$$\begin{aligned} \hat{\mu}(\vec{X}_n) &= \frac{1}{n} \sum_{i=1}^n X_i \\ S^2(\vec{X}_n) &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \end{aligned} \quad (3)$$

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:

где n_i – количество элементов выборки \vec{x} , которые $\in J_i$.

Обычно выборку разбивают на $m = \lceil \log_2 n \rceil + 2$ интервалов, где n – размер выборки.

J_1	\dots	J_i	\dots	J_m
n_1	\dots	n_i	\dots	n_m

Гистограмма – это график эмпирической плотности.

Эмпирической плотностью, отвечающей выборке \vec{x} , называют функцию:

$$\hat{f}(x) = \begin{cases} \frac{n_i}{n\Delta}, x \in J_i, i = \overline{1; m} \\ 0, \text{ иначе} \end{cases} \quad (4)$$

где J_i – полуинтервал статистического ряда, n_i – количество элементов выборки, входящих в полуинтервал, n – количество элементов выборки.

Определение эмпирической функции распределения

Пусть $\vec{x} = (x_1, \dots, x_n)$ – 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 $F_n : \mathbb{R} \rightarrow \mathbb{R}$, определенную как:

$$F_n(x) = \frac{n(x, \vec{x})}{n} \quad (5)$$

Результаты работы программы

Текст программы

Результаты расчётов

- $M_{\max} = 12.4100$
- $M_{\min} = 6.8100$
- $R = 5.6000$
- $\hat{\mu}(\vec{x}_n) = 9.4872$
- $S^2(\vec{x}_n) = 1.2173$
- $m = 8$