



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, искусственный интеллект и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

Отчёт по лабораторной работе № 1 по курсу «Математическая статистика»

Тема Гистограмма и эмпирическая функция распределения

Студент Калашков П. А.

Группа ИУ7-66Б

Оценка (баллы)

Преподаватели Андреева Т. В.

1 Задание

1.1 Цель работы

Цель работы: построение гистограммы и эмпирической функции распределения.

1.2 Содержание работы

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 Теоретическая часть

2.1 Формулы для вычисления величин M_{max} , M_{min} , R , $\hat{\mu}$, S^2

Минимальное значение выборки рассчитывается по формуле (2.1); максимальное – (2.2). Размах выборки рассчитывается по формуле (2.3); выборочное среднее – (2.4), исправленная выборочная дисперсия – (2.5).

$$M_{\min} = X_{(1)} \quad (2.1)$$

$$M_{\max} = X_{(n)} \quad (2.2)$$

$$R = M_{\max} - M_{\min}. \quad (2.3)$$

$$\hat{\mu}(\vec{X}_n) = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.4)$$

$$S^2(\vec{X}_n) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \quad (2.5)$$

2.2 Эмпирическая плотность и гистограмма

Пусть \vec{x} – выборка из генеральной совокупности X .

При большом объеме n ($n > 50$) этой выборки значения x_i группируют в интервальный статистический ряд. Для этого отрезок $J = [x_{(1)}, x_{(n)}]$ делят на m равновеликих промежутков по формуле (2.6):

$$J_i = [x_{(1)} + (i-1) \cdot \Delta, x_{(1)} + i \cdot \Delta], i = \overline{1; m-1} \quad (2.6)$$

Последний промежуток определяется по формуле (2.7):

$$J_m = [x_{(1)} + (m-1) \cdot \Delta, x_{(n)}] \quad (2.7)$$

Ширина каждого из таких промежутков определяется по формуле (2.8).

$$\Delta = \frac{|J|}{m} = \frac{x_{(n)} - x_{(1)}}{m} \quad (2.8)$$

Интервальным статистическим рядом называют таблицу 2.1:

Таблица 2.1 – Интервальный статистический ряд

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

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

Гистограмма – это график эмпирической плотности.

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

$$\hat{f}(x) = \begin{cases} \frac{n_i}{n\Delta}, x \in J_i, i = \overline{1; m} \\ 0, x \notin J \end{cases} \quad (2.9)$$

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

2.3 Эмпирическая функция распределения

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

Обозначим $n(t, \vec{x})$ – число элементов вектора \vec{x} , которые имеют значения меньше t .

Эмпирической функцией распределения называют функцию $F_n : \mathbb{R} \rightarrow \mathbb{R}$, определенную как:

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

3 Практическая часть

3.1 Код программы

Листинг 3.1 – Программа для MatLAB

```
1 file = fopen('selection.txt', 'r');
2 X = fscanf(file, '%f,');
3 fclose(file);
4 %% Calculation
5 M_max = max(X);
6 M_min = min(X);
7 R = M_max - M_min;
8 n = length(X);
9 mu = sum(X) / n;
10 Ssquare = sum((X - mu).^2) / (n - 1);
11 m = floor(log2(n)) + 2;
12 step = R / m;
13 sorted_X = sort(X);
14 intervals = cell(1, m);
15 i = 1;
16 for cur = (M_min + step):step:M_max
17     last = cur - step;
18     intervals(i) = {X((last < X) & (X < cur))};
19     i = i + 1;
20 end
21 intervals{m} = [intervals{m}; X(X == M_max)];
22 density = 1:m;
23 for i = 1:m
24     density(i) = length(intervals{i}) / n / step;
25 end
26 %% probability density function
27 histogram(X, m, 'BinEdges', M_min:step:M_max, 'Normalization', 'pdf');
28 hold on;
29 x = (M_min - step):0.1:(M_max+step);
30 f = pdf('Normal', x, mu, Ssquare);
31 plot(x, f, 'LineWidth', 1.5, 'color', 'red')
32 hold off;
33 %% probability function
34 x = (M_min - step):0.1:(M_max+step);
35 [F_empirical, x_empirical] = ecdf(X);
36 F = normcdf((x - mu) / sqrt(Ssquare));
37 hold on;
38 plot(x_empirical, F_empirical, 'LineWidth', 1.5);
39 plot(x, F, 'LineWidth', 1.5, 'color', 'red');
40 hold off;
```

3.2 Результаты расчетов для выборки из индивидуального варианта.

Согласно варианту 3, результаты расчетов для выборки приведены на формулах (3.1), (3.2), (3.3), (3.4), (3.5), (3.6).

$$M_{\min} = -2.77 \quad (3.1)$$

$$M_{\max} = 2.92 \quad (3.2)$$

$$R = 5.69 \quad (3.3)$$

$$\hat{\mu}(\vec{x}_n) = 0.2323 \quad (3.4)$$

$$S^2(\vec{x}_n) = 1.0406 \quad (3.5)$$

$$m = 8 \quad (3.6)$$

На рисунке 3.1 представлены гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с выборочными мат. ожиданием и дисперсией.

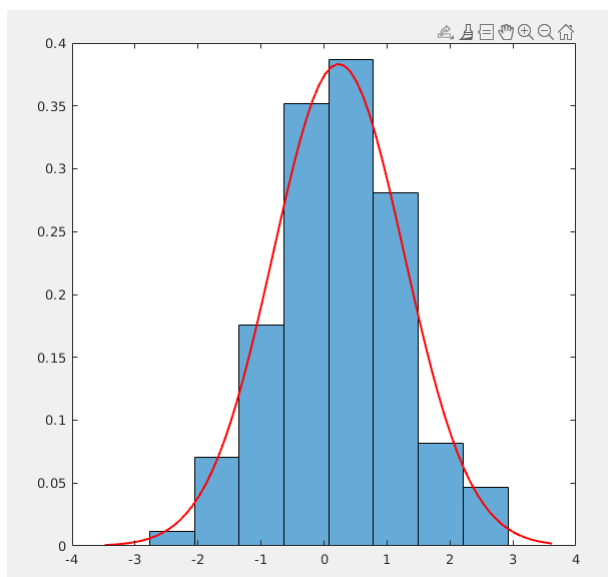


Рисунок 3.1 – Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с выборочными мат. ожиданием и дисперсией.

На рисунке 3.2 представлены график эмпирической функции распределения и функции распределения нормальной случайной величины с выборочными мат. ожиданием и дисперсией.

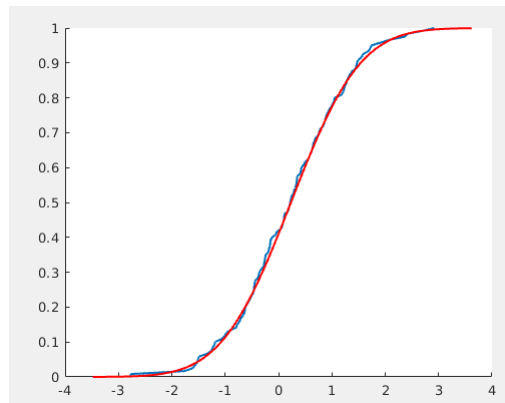


Рисунок 3.2 – График эмпирической функции распределения и функции распределения нормальной случайной величины с выборочными мат. ожиданием и дисперсией.