



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 _____ Динь Вьет Ань

Группа _____ ИУ7И-64Б

Преподаватель _____ Андреева Татьяна Владимировна

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

2.1.1 Минимальное и максимальное значения выборки

$$\begin{aligned} M_{\max} &= X_{(n)} \\ M_{\min} &= X_{(1)} \end{aligned} \quad (2.1)$$

2.1.2 Размах выборки

$$R = M_{\max} - M_{\min}. \quad (2.2)$$

2.1.3 Оценки математического ожидания и дисперсии

$$\begin{aligned} \hat{\mu}(\vec{X}_n) &= \frac{1}{n} \sum_{i=1}^n X_i \\ S^2(\vec{X}_n) &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \end{aligned} \quad (2.3)$$

2.2 Эмпирическая плотность и гистограмма

Пусть

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

При большом объеме n этой выборки значения x_i группируют в интервальный статистический ряд.

Отрезок $J = [x_{(1)}, x_{(n)}]$ делят на m равновеликих частей:

$$J_i = [x_{(1)} + (i-1) \cdot \Delta, x_{(1)} + i \cdot \Delta], i = \overline{1; m-1} \quad (2.4)$$

$$J_m = [x_{(1)} + (m-1) \cdot \Delta, x_{(n)}] \quad (2.5)$$

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

Интервальным статистическим рядом называют таблицу:

J_1	\dots	J_i	\dots	J_m
n_1	\dots	n_i	\dots	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:

$$\hat{f}(x) = \begin{cases} \frac{n_i}{n\Delta}, x \in J_i, i = \overline{1; m} \\ 0, \text{ иначе} \end{cases} \quad (2.7)$$

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

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

$$F_n(x) = \frac{n(x, \vec{x})}{n} \quad (2.8)$$

3 | Практическая часть

```
1  % Вариант 17
2  X = [-13.40,-12.63,-13.65,-14.23,-13.39,-12.36,-13.52,-13.44,-13.87,-11.82,
3
4  % a)
5  M_max = max(X);
6  M_min = min(X);
7
8  % dlmwrite('M_max', M_max);
9  % dlmwrite('M_min', M_min);
10
11 % b)
12 R = M_max - M_min;
13
14 % dlmwrite('R', R);
15
16 % c)
17 MX = mean(X);
18 DX = var(X); % var - возвращает дисперсию.
19
20 % dlmwrite('MX', MX);
21 % dlmwrite('DX', DX);
22
23 % d)
24 n = length(X);
25 m = floor(log2(n)) + 2; % округление в сторону – бесконечности.
26 h = histogram(X, m);
27
28 % e)
29 sigma = sqrt(DX);
30 % begin:step:end
31 x = (M_min - 1):(sigma / 100):(M_max + 1);
32 % Функция плотности вероятности стандартного нормального распределения.
33 f = normpdf(x, MX, sigma);
34
35 figure;
36 heights = h.Values / (sum(h.Values) * h.BinWidth);
37 centers = [];
38 for i = 1:(length(h.BinEdges) - 1)
39     centers = [centers, (h.BinEdges(i + 1) + h.BinEdges(i)) / 2];
40 end
41 hold on;
42 bar(centers, heights, 1);
43 plot(x, f, 'g', 'LineWidth', 2);
44
45 % f)
46 F = normcdf(x, MX, sigma);
47 figure;
48 hold on;
49 ecdf(X);
50 plot(x, F, 'r');
```

4 | Экспериментальная часть

4.1 Результаты расчетов

	M_min ×	+
1	-14.6	
2		

	M_max ×	+
1	-10.25	
2		

	R ×	+
1	4.35	
2		

	MX ×	+
1	-12.615	
2		

	DX ×	+
1	0.86533	
2		

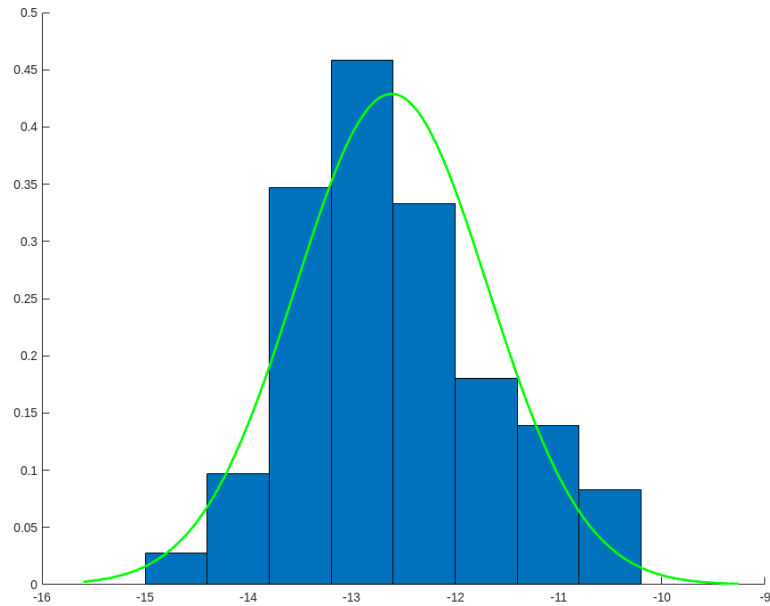
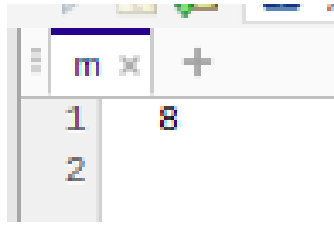


Рис. 4.1: Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с выборочными мат. ожиданием и дисперсией

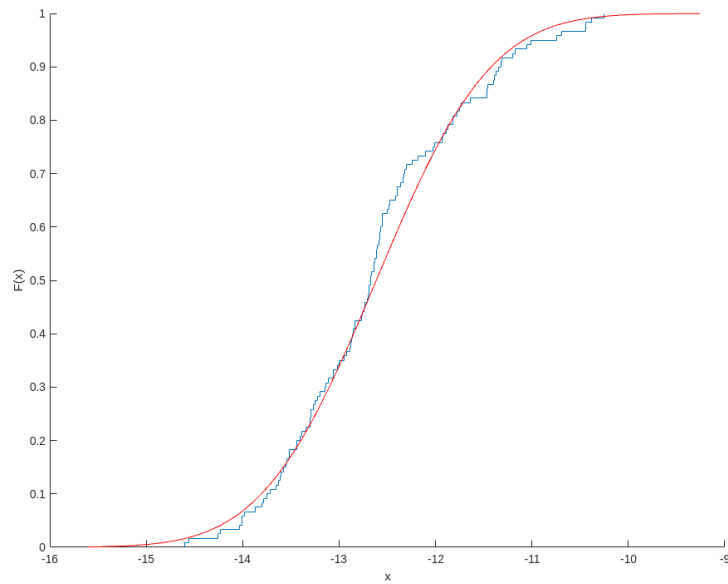


Рис. 4.2: График эмперической функции распределения и функции распределения нормальной случайной величины с выборочными мат. ожиданием и дисперсией