



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический
университет имени Н.Э. Баумана»
(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 функций.
Студент	Степанов А. О.
Группа	ИУ7-63Б
Оценка (баллы)	
Преподаватель	Власов П.А.

Москва, 2020 г.

ФОРМУЛЫ ВЕЛИЧИН

Все формулы применяются для выборки из генеральной совокупности $\vec{X} = (x_1, \dots, x_n)$.

Максимальное значение выборки вычисляется по формуле 1.

$$M_{\max} = \max\{x_1, \dots, x_n\} \quad (1)$$

Минимальное значение выборки вычисляется по формуле 2.

$$M_{\min} = \min\{x_1, \dots, x_n\} \quad (2)$$

Размах выборки вычисляется по формуле 3.

$$R = M_{\max} - M_{\min} \quad (3)$$

Оценка математического ожидания вычисляется по формуле 4.

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (4)$$

Оценка дисперсии вычисляется по формуле 5.

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (5)$$

ЭМПИРИЧЕСКАЯ ПЛОТНОСТЬ И ГИСТОГРАММА

Эмпирической плотностью (отвечающей выборке \vec{x}) называют функцию

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

Гистограммой называют график эмпирической плотности.

ЭМПИРИЧЕСКАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

Эмпирической функцией распределения называют функцию

$$F_n : \mathcal{R} \rightarrow \mathcal{R}$$

определенную условием

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

ТЕКСТ ПРОГРАММЫ

Листинг 1: Точка входа в программу

```
1 function lab1()
2     X = [
3         -7.50, -6.61, -7.85, -7.72, -8.96, ...
4         -6.55, -7.82, -6.55, -6.87, -5.95, ...
5         -5.05, -4.56, -6.14, -6.83, -6.33, ...
6         -7.67, -4.65, -6.30, -8.01, -5.88, ...
7         -5.38, -7.06, -6.85, -5.53, -7.83, ...
8         -5.89, -7.57, -6.76, -6.02, -4.62, ...
9         -8.55, -6.37, -7.52, -5.78, -6.12, ...
10        -8.82, -5.14, -7.68, -6.14, -6.48, ...
11        -7.14, -6.25, -7.32, -5.51, -6.97, ...
12        -7.86, -7.04, -6.24, -6.41, -6.00, ...
13        -7.46, -6.00, -6.06, -5.94, -5.39, ...
14        -5.06, -6.91, -8.06, -7.24, -6.42, ...
15        -8.73, -6.20, -7.35, -5.90, -5.02, ...
16        -5.93, -7.56, -7.49, -6.26, -6.06, ...
17        -7.35, -5.10, -6.52, -7.97, -5.71, ...
18        -7.62, -7.33, -5.31, -6.21, -7.28, ...
19        -7.99, -4.65, -7.07, -7.31, -7.72, ...
20        -5.22, -7.00, -7.17, -6.64, -7.00, ...
21        -6.12, -6.57, -6.07, -6.65, -7.60, ...
22        -6.92, -6.78, -6.85, -7.90, -7.40, ...
23        -5.32, -6.58, -6.71, -5.07, -5.80, ...
24        -4.87, -5.90, -7.43, -7.03, -6.67, ...
25        -7.72, -5.83, -7.49, -6.68, -6.71, ...
```

```

26         -7.31, -7.83, -7.92, -5.97, -6.34, ...
27     ];
28
29     Params(X);
30     Intervals(X);
31     MakeGraphs(X);
32 end

```

Листинг 2: Функция для вычисления параметров

```

1  % Подсчет параметров выборки и вывода их на экран
2  % [in] X - генеральная совокупность
3  function Params(X)
4      % Максимальное значение выборки
5      maxX = max(X);
6      fprintf("Mmax= %.3f\n", maxX);
7
8      % Минимальное значение выборки
9      minX = min(X);
10     fprintf("Mmin= %.3f\n", minX);
11
12     % Разброс выборки
13     R = maxX - minX;
14     fprintf("R= %.3f\n", R);
15
16     % Оценка математического ожидания выборки
17     M = mean(X);
18     fprintf("M= %.3f\n", M);
19
20     % Оценка дисперсии выборки
21     D = FindD(X);
22     fprintf("D= %.3f\n", D);
23 end
24
25 % Функция для вычисления оценки дисперсии выборки
26 % [in] X - генеральная совокупность
27 % [return] оценка дисперсии выборки
28 function D = FindD(X)
29     D = sum((X - mean(X)) .^ 2) / (length(X) - 1);
30     return
31 end

```

Листинг 3: Вычисление интервалов

```
1 % Функция для группировки значений в m интервалов
2 % и вывода на экран количества элементов в каждом интервале
3 % (m = [log2(n)] + 2)
4 % [in] X - генеральная совокупность
5 function Intervals(X)
6     m = floor(log2(length(X))) + 2;
7     delta = (max(X) - min(X)) / m;
8     borders = min(X) : delta : max(X);
9
10    fprintf('\n%d intervals:\n', m);
11    for i = 1:(length(borders) - 1)
12        count = 0;
13        for x = X
14            % Последний интервал включает в себя правое значение
15            if (i == length(borders) - 1) && (x >= borders(i)) && ...
16                (x <= borders(i + 1))
17                count = count + 1;
18            % Остальные интервалы включительно только слева
19            elseif (x >= borders(i)) && (x < borders(i + 1))
20                count = count + 1;
21        end
22    end
23
24    printableString = ' [%.3f; %.3f) -> %d\n';
25    if (i == length(borders) - 1)
26        printableString = ' [%.3f; %.3f] -> %d\n';
27    end
28
29    fprintf(printableString, borders(i), borders(i + 1), count);
30 end
31 end
```

Листинг 4: Рисование графиков

```
1 % Функция для рисования графиков
2 % [in] X - генеральная совокупность
3 function MakeGraphs(X)
4     x = sort(X);
5
6     subplot(2, 1, 1);
7     % Гистограмма
8     histogram(X, 'Normalization', 'pdf');
```

```

9      hold on;
10     % График функции плотности распределения нормальной случайной величины
11     f = normpdf(x, mean(x), sqrt(FindD(x)));
12     p1 = plot(x, f);
13     p1.LineWidth = 2;
14     hold off;
15     legend({'Гистограмма', 'Функция_плотности_распределения'}, ...
16           'Location','northwest');
17
18     subplot(2, 1, 2);
19     % Эмперическая функция распределения
20     histogram(X, 'Normalization', 'cdf');
21     hold on;
22     % Функция распределения нормальной случайной величины
23     f = normcdf(x, mean(x), sqrt(FindD(x)));
24     p2 = plot(x, f);
25     p2.LineWidth = 2;
26     hold off;
27     legend({'Эмперическая_функция_распределения', ...
28           'Функция_распределения'}, ...
29           'Location','northwest');
30 end

```

РЕЗУЛЬТАТЫ РАСЧЕТОВ

$$M_{\max} = -4.560$$

$$M_{\min} = -8.960$$

$$R = 4.400$$

$$\hat{\mu} = -6.649$$

$$S^2 = 0.948$$

$[-8.960; -8.410)$	4
$[-8.410; -7.860)$	6
$[-7.860; -7.310)$	25
$[-7.310; -6.760)$	20
$[-6.760; -6.210)$	23
$[-6.210; -5.660)$	24
$[-5.660; -5.110)$	8
$[-5.110; -4.560]$	10

Таблица 1: Группировка значений выборки в 8 интервалов

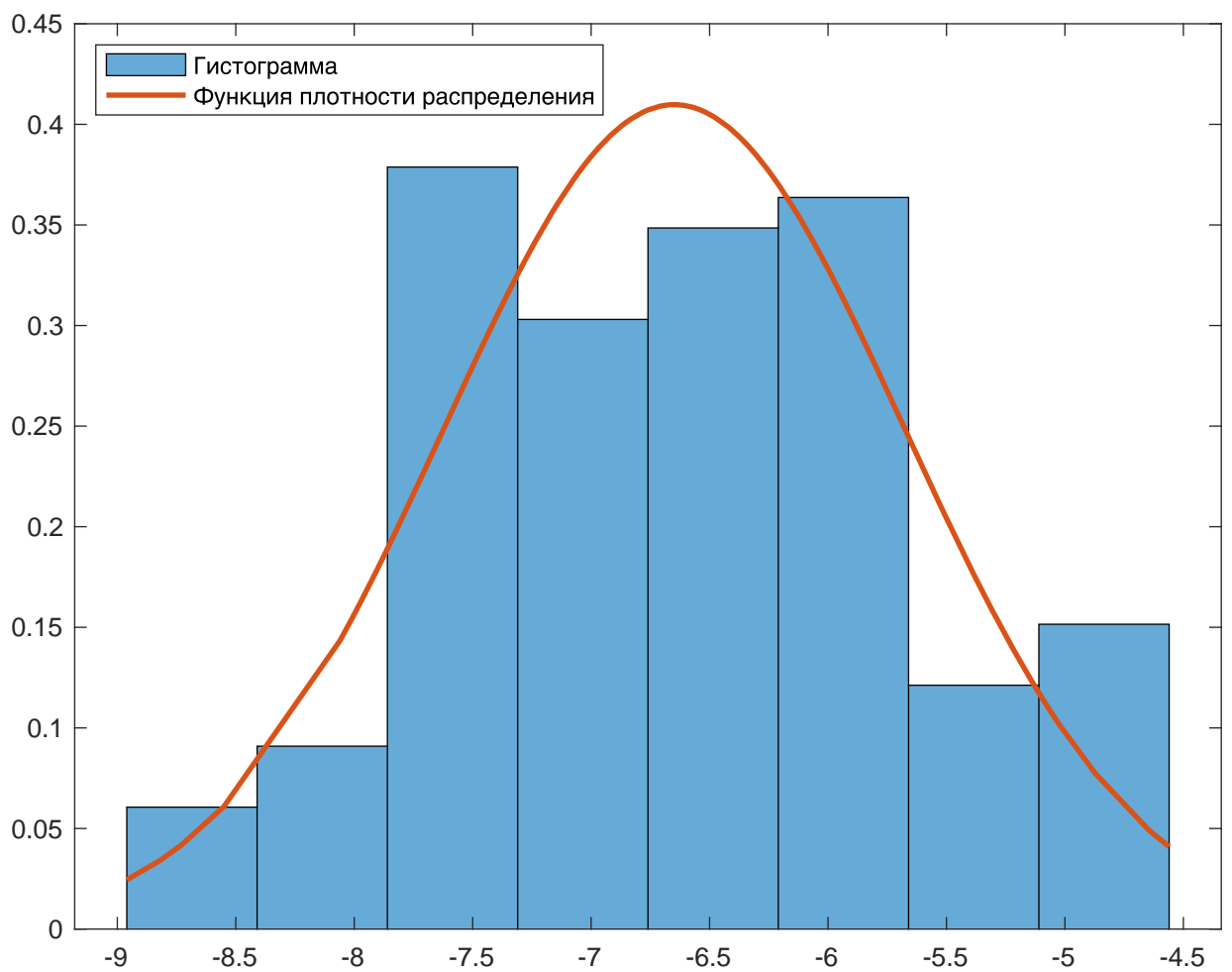


Рис. 1: Гистограмма

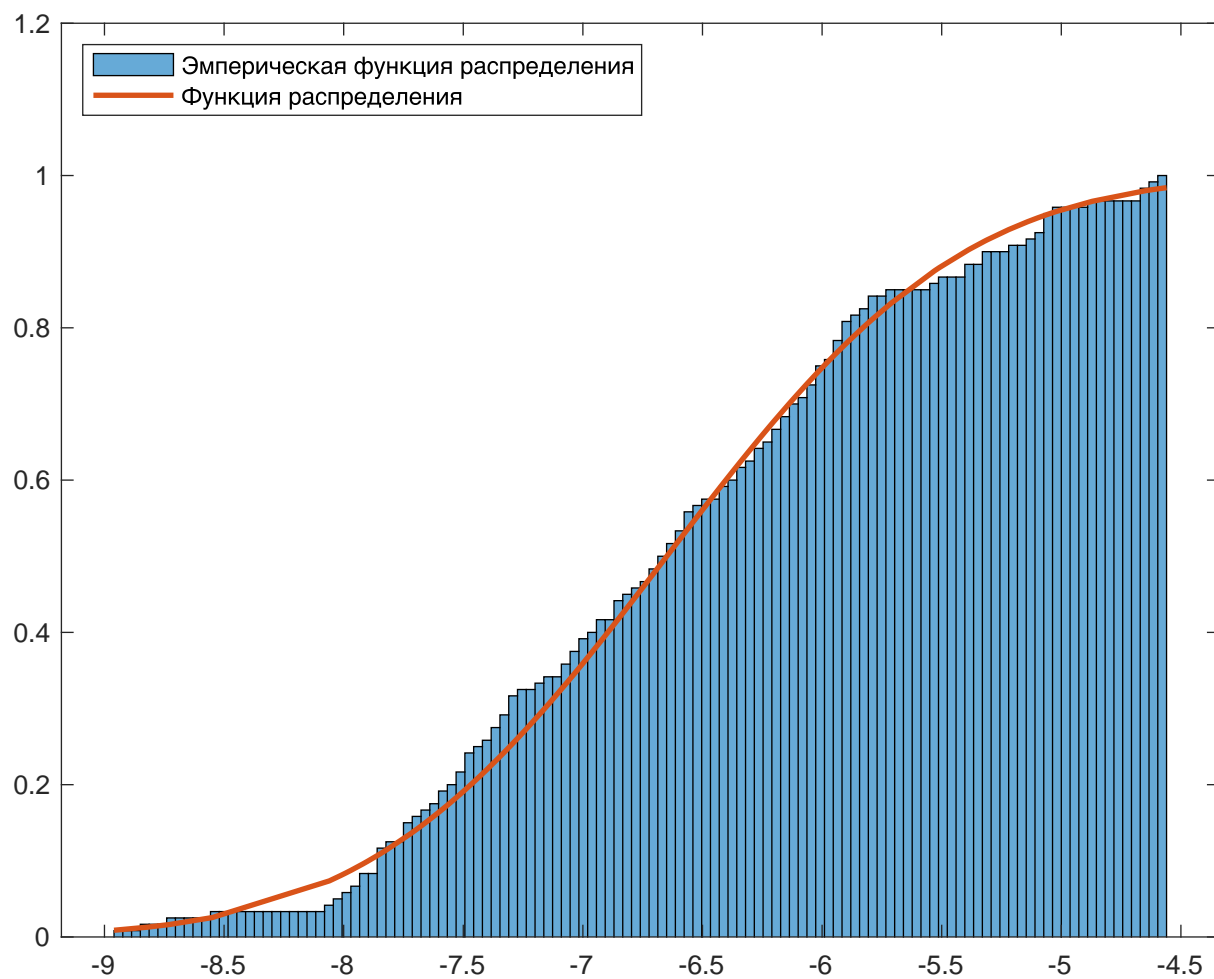


Рис. 2: Эмперическая функция распределения