



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

Вариант 22

Дисциплина	Математическая статистика.
Тема	
Студент	Тимонин А. С.
Группа	ИУ7-62Б
Оценка (баллы)	
Преподаватель	Власов П.А.

Москва, 2020 г.

Теоретическая часть

Формулы для вычисления

Выборка из генеральной совокупности $\vec{x} = (x_1, \dots, x_n)$

Формула для вычисления максимального значения M_{\max} :

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

Формула для вычисления минимального значения M_{\min} :

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

Размах выборки R вычисляется по формуле:

$$R = M_{\max} - M_{\min}$$

Вычисление оценки математического ожидания $M\bar{X}$:

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

Вычисление оценки дисперсии $D\bar{X}$:

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

Определение эмпирической плотности и гистограммы

Интервальный статистический ряд

Пусть \vec{x} - выборка из генеральной совокупности X . Если объем n этой выборки велик ($n \geq 50$), то значения x_i группируют не только в статистический ряд, но и в так называемый интервальный статистический ряд. Для этого отрезок $J = [x_{(1)}, x_{(n)}]$ делят на p равновеликих частей:

$$J_i = [a_i, a_{i+1}), i = \overline{0; p-1}$$

$$J_p = [a_{p-1}, a_p]$$

$$\text{где } a_i = x_{(1)} + i\Delta, i = \overline{0; p}, \Delta = \frac{|J|}{p} = \frac{x_{(n)} - x_{(1)}}{p}$$

Опр Интервальным статистическим рядом называют таблицу

J_1	\dots	J_i	\dots	J_p
n_1	\dots	n_i	\dots	n_p

Здесь n_i - количество элементов выборки \vec{x} , которые $\in J_i$

Замечание

1. Очевидно, что $\sum_{i=1}^p n_i = n$

2. Для выборки p - числа интервалов можно пользоваться формулой $p = [\log_n n] + 1$

где $[a]$ - целая часть числа a

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

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

Опр **Гистограммой** называют график эмпирической плотности

Определение эмперической функции распределения

Опр **Эмпирической функцией распределения** называют функцию

$$\mathcal{F}_n : \mathbb{R} \rightarrow \mathbb{R}$$

определенную условием

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

Практическая часть

Листинги программы

Листинг 1: Содержимое генеральной совокупности X

```
1 7.76, 5.96, 4.58, 6.13, 5.05,  
2 6.40, 7.46, 5.55, 5.01, 3.79,  
3 7.65, 8.87, 5.94, 7.25, 6.76,  
4 6.92, 6.68, 4.89, 7.47, 6.53,  
5 6.76, 6.96, 6.58, 7.92, 8.47,  
6 6.27, 8.05, 5.24, 5.60, 6.69,  
7 7.55, 6.02, 7.34, 6.81, 7.22,  
8 6.39, 6.40, 8.28, 5.39, 5.68,  
9 6.71, 7.89, 5.69, 5.18, 7.84,  
10 7.18, 7.54, 6.04, 4.58, 6.82,  
11 4.45, 6.75, 5.28, 7.42, 6.88,  
12 7.10, 5.24, 9.12, 7.37, 5.50,  
13 5.52, 6.34, 5.31, 7.71, 6.88,  
14 6.45, 7.51, 6.21, 7.44, 6.15,  
15 6.25, 5.59, 6.68, 6.52, 4.03,  
16 5.35, 6.53, 3.68, 5.91, 6.68,  
17 6.18, 7.80, 7.17, 7.31, 4.48,  
18 5.69, 7.11, 6.87, 6.14, 4.73,  
19 6.60, 5.61, 7.32, 6.75, 6.28,  
20 6.41, 7.31, 6.68, 7.26, 7.94,  
21 7.67, 4.72, 6.01, 5.79, 7.38,  
22 5.98, 5.36, 6.43, 7.25, 5.54,  
23 6.66, 6.47, 6.84, 6.13, 6.21,  
24 5.52, 6.33, 7.55, 6.24, 7.84
```

Листинг 2: Точка входа программы

```
1 function lab01(  
2     X = csvread('/Users/antontimonin/Desktop/MatStat/lab_01/Data12.csv');  
3     X = sort(X);  
4  
5     params(X);  
6     intervals(X);  
7     graphs(X);  
8 end
```

Листинг 3: Функции для вычисления параметров

```
1 function params(X)  
2     X = sort(X);  
3     Xmin = X(1); fprintf('Mmin = %s\n', num2str(Xmin));  
4     Mmax = X(end); fprintf('Mmax = %s\n', num2str(Mmax));  
5     R = Mmax - Xmin; fprintf('R = %s\n', num2str(R));  
6     mu = expect(X); fprintf('mu = %s\n', num2str(mu));  
7     sigSqr = populVar(X); fprintf('sigma^2 = %s\n', num2str(sigSqr));  
8     sqr = unbisVariance(X); fprintf('S^2: %s\n', num2str(sqr));  
9     m = numSubintervals(length(X)); fprintf('m = %s\n ', num2str(m));  
10 end  
11  
12 function m = expect(X)
```

```

13  n = length(X);
14  sum = 0;
15
16  for i = 1:n
17      sum = sum + X(i);
18  end
19
20  m = sum / n;
21 end
22
23 function sigSqr = populVar(X)
24     n = length(X);
25     sum = 0;
26
27     for i = 1:n
28         sum = sum + (X(i))^2;
29     end
30
31     mu = expect(X);
32     sigSqr = sum / n - mu^2;
33 end
34
35 function sqr = unbisVariance(X)
36     sigSqr = populVar(X);
37     n = length(X);
38
39     sqr = n / (n - 1) * sigSqr;
40 end
41
42 function m = numSubintervals(size)
43     m = floor(log2(size) + 2);
44 end

```

Листинг 4: Функция вычисляющая интервалы

```

1  function intervals(X)
2      m = numSubintervals(length(X));
3
4      count = zeros(1, m+2);
5      delta = (X(end) - X(1)) / m;
6
7      J = X(1):delta:X(end);
8      fprintf('%d\n', X(end));
9      J(length(J)+1) = 20;
10
11     j = 1;
12     n = length(X);
13
14     for i = 1:n
15         if (j ~ = m)
16             if ((not (X(i) >= J(j) && X(i) < J(j+1))))
17                 j = j + 1;
18                 fprintf(' [%.2f;%.2f)\n', J(j-1), J(j));
19             end
20         end

```

```

21     count(j) = count(j) + 1;
22 end
23
24 fprintf(' [%2.2f;%2.2f] -> %d\n', J(m), J(m + 1), count);
25
26 Xbuf = count(1:m+2);
27 for i = 1:m+2
28     Xbuf(i) = count(i) / (n*delta);
29 end
30
31 stairs(J, Xbuf), grid;
32 end

```

Листинг 5: Отрисовка графов

```

1 function graphs(X)
2     hold on;
3     f(X, expect(X), ...
4         unbisVariance(X), ...
5         numSubintervals(length(X)));
6
7     figure;
8     empiricF(X);
9     hold on;
10    F(X, expect(X), ...
11        unbisVariance(X), ...
12        numSubintervals(length(X)));
13 end
14
15 function f(X, MX, DX, m)
16     R = X(end) - X(1);
17     delta = R/m;
18     Sigma = sqrt(DX);
19     Xn = (MX - R): delta/50 :(MX + R);
20     Xn(length(Xn)+1) = 20;
21     Y = normpdf(Xn, MX, Sigma);
22     plot(Xn, Y);
23 end
24
25 function F(X, MX, DX, m)
26     R = X(end) - X(1);
27     delta = R/m;
28     Xn = (MX - R): delta :(MX + R);
29     Y = 1/2 * (1 + erf((Xn - MX) / sqrt(2*DX)));
30     p2 = plot(Xn, Y, 'Color', [.1 .7 .7], 'LineWidth', 1);
31     hold off;
32 end
33
34 function empiricF(X)
35     [yy, xx] = ecdf(X);
36     yy(length(yy)+1) = 1;
37     xx(length(xx)+1) = 20;
38     stairs(xx, yy);
39 end

```

Результат выполнения программы

$$M_{\max} = 9.12$$

$$M_{\min} = 3.68$$

$$R = 5.44$$

$$\mu = 6.4596$$

$$S^2 = 1.1013$$

[3.68; 4.36)	3
[4.36; 5.04)	8
[5.04; 5.72)	20
[5.72; 6.40)	22
[6.40; 7.08)	30
[7.08; 7.76)	25
[7.76; 8.44)	9
[8.44; 9.12]	3

Таблица 1: Результаты расчетов для выборки

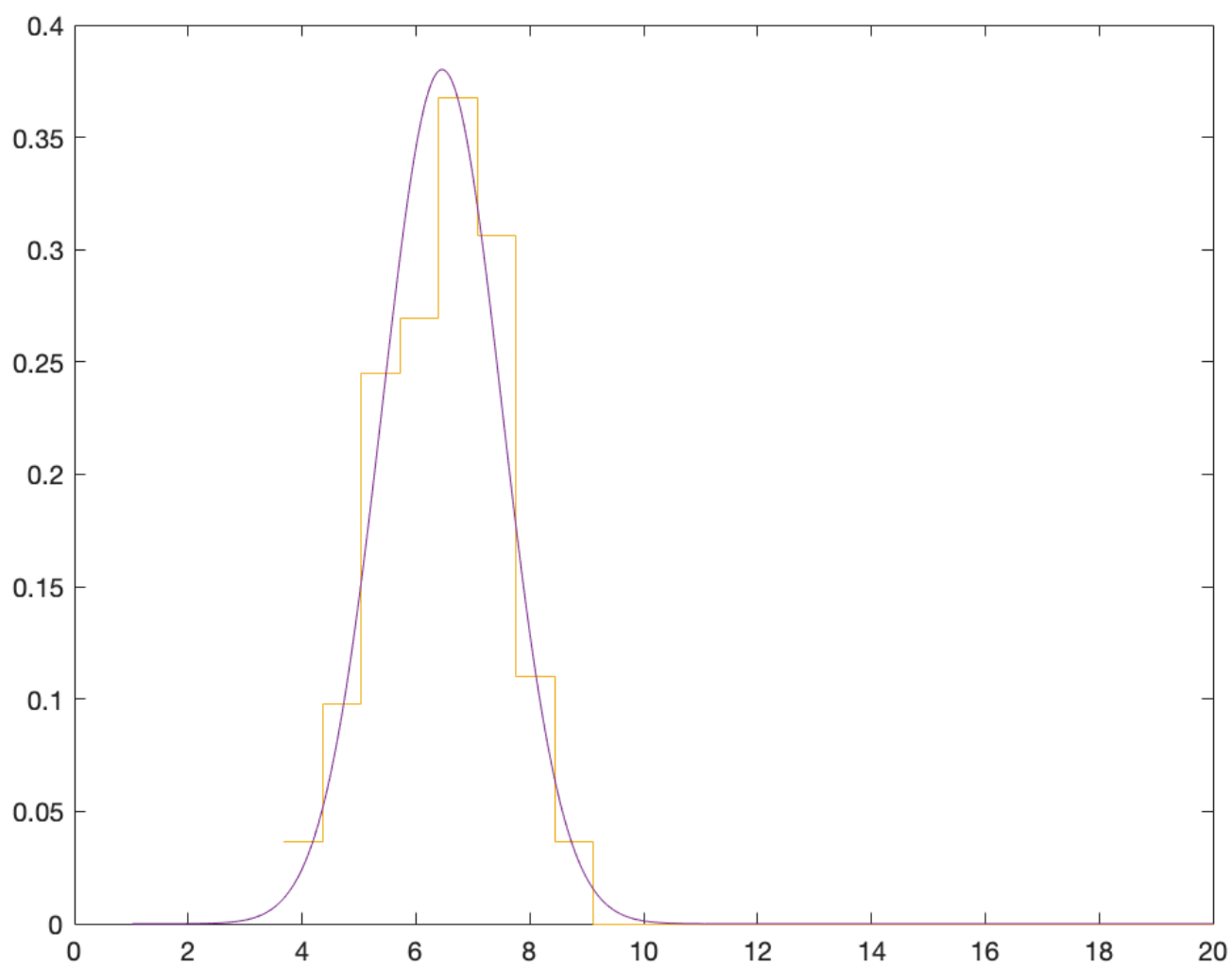


Рис. 1: Гистограмма

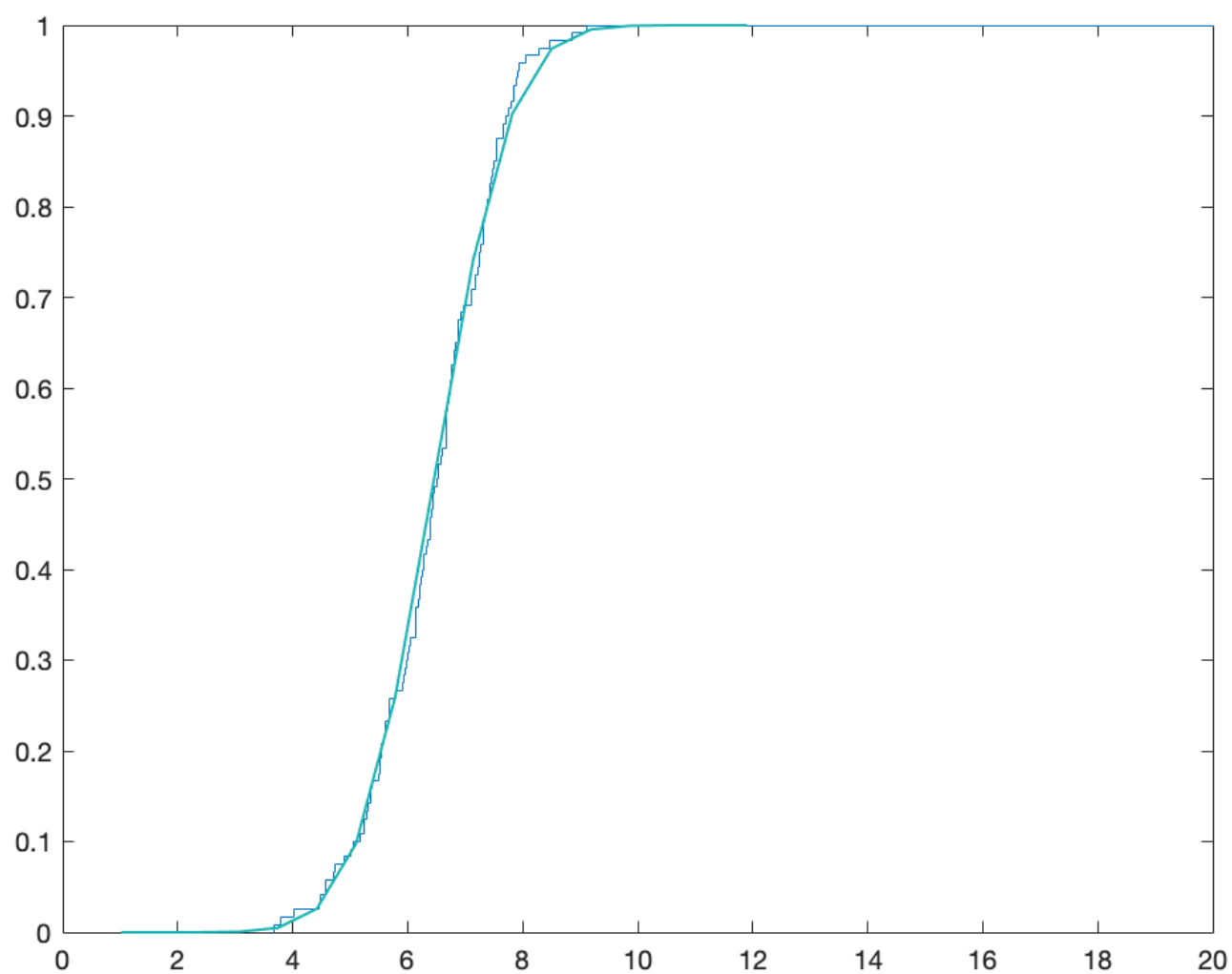


Рис. 2: Эмпирическая функция распределения