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 курсу: «Математическая статистика»

Тема: «Гистограмма и эмпирическая функция распределения»

Студент:	Пронин А. С.
Группа:	ИУ7-62Б
Преподаватель:	Власов П. А.
Оценка:	____________

Москва

2022

Введение

Цель работы: построение гистограммы и эмпирической функции рапределения. **Содержание работы:**

1. Для выборки объема n из генеральной совокупности X реализовать в виде программы на ЭВМ

   а) вычисление максимального значения M_{max} и минимального значения M_{min};

   б) размаха R выборки;

   в) вычисление оценок $\overset{\wedge}{\mu}$ и S^2 математического ожидания MX и дисперсии DX;

   г) группировку значений выборки в $m = [log_2 n] + 2$ интервала;

   д) построение на одной координатной плоскости гистограммы и графика функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием $\overset{\wedge}{\mu}$ и дисперсией S^2;

   е) построение на другой координатной плоскости графика эмпирической функции распределения и функции распределения нормальной случайной величины математическим ожиданием $\overset{\wedge}{\mu}$ и дисперсией S^2.

2. Провести вычисления и построить графики для выборки из индивидуального варианта.

Отчёт

Формулы для вычисления величин

$M_{max} = max(\overrightarrow{x_n})$,
$M_{min} = min(\overrightarrow{x_n})$,
$R = M_{max} - M_{min}$
$\overset{\wedge}{\mu}(\overrightarrow{x_n}) = \overline{x_n} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$
$S^2(\overrightarrow{x_n}) = \frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \overline{x_n})^2 = \frac{n}{n-1}\overset{\wedge}{\sigma}^2(\overrightarrow{x})$, где $\overset{\wedge}{\sigma}^2(\overrightarrow{x}) = \frac{1}{n}\sum\limits_{i=1}^{n}(x_i - \overline{x_n})^2$

Определения эмпирической плотности и гистограммы

Определение Эмпирической плотонстью, отвечающей интервальному ряду $(J_i, n_i), \overline{1, m}$ называется функция

$$\hat{f}(x) = \begin{cases} \frac{n_i}{n\Delta}, & \text{если } x \in J_i; \\ 0, & \text{иначе,} \end{cases}$$

где $\Delta = \frac{x_{(n)} - x_{(1)}}{m}$,
$J_i = [x_{(1)}] + (i-1)\Delta, x_{(1)} + i\Delta], i = \overline{1, m}$,
$J_m = [x_{(1)}] + (m-1)\Delta, x_{(n)}]$,
а $m = [log_2 n] + 2$ по заданию

Определение График эмпирической плотности называется гистограммой.

Определения эмпирической плотности и гистограммы

Пусть $\overrightarrow{x} = (x_1, ..., x_n)$ - выборка из генеральной совокупности X.

Обозначим $n(t, \overrightarrow{x})$ - число компонент вектора $\overrightarrow{x}$, которые меньше чем t

Определение Эмпирической функцие распределения, построенной по выборке $\overrightarrow{x}$, называют функцию

$F_n(t) : \mathbb{R} \to \mathbb{R}$,

определенную правилом $F_n(t) = \frac{n(t, \overrightarrow{x})}{n}$

Текст программы

```
close all
clear
X=[-13.40 -12.63 -13.65 -14.23 -13.39 -12.36 -13.52 -13.44
   -13.87 -11.82 -12.01 -11.40 -13.02 -12.61 -13.06 -13.75
   -13.55 -14.01 -11.75 -12.95 -12.59 -13.60 -12.76 -11.05
   -13.15 -13.61 -11.73 -13.00 -12.66 -12.67 -12.60 -12.47
   -13.52 -12.61 -11.93 -13.11 -13.22 -11.87 -13.44 -12.70
   -11.78 -12.30 -12.89 -13.29 -12.48 -10.44 -12.55 -12.64
   -12.03 -14.60 -14.56 -13.30 -11.32 -12.24 -11.17 -12.50
   -13.25 -12.55 -12.85 -12.67 -12.41 -12.58 -12.10 -13.54
   -12.69 -12.87 -12.71 -12.77 -13.30 -12.74 -12.73 -12.64
   -12.18 -11.20 -12.40 -13.78 -13.71 -10.74 -11.89 -13.20
   -11.31 -14.26 -10.38 -12.88 -11.39 -11.35 -12.55 -12.84
   -10.25 -12.40 -14.01 -11.47 -13.14 -12.69 -11.92 -12.86
   -13.06 -12.57 -13.63 -12.34 -12.84 -14.03 -13.34 -11.64
   -13.58 -10.44 -11.37 -11.01 -13.80 -13.27 -12.32 -10.69
   -12.92 -13.29 -12.58 -13.98 -11.46 -11.82 -12.33 -11.47]
%Task1
Mmin=min(X)
Mmax=max(X)
%Task2
R=Mmax-Mmin
%Task3
n = length(X)
mu = sum(X)/n
```

```
%sigmasqr = sum(power(X-mu, 2))/n
Ssqr = 1/(n-1)*sum(power(X-mu, 2))
%Task4
m = fix(log(n)/log(2)+2)
delta = (Mmax-Mmin)/m
J=[Mmin:delta:Mmax-delta; Mmin+delta:delta:Mmax]
IntStatR=[1:m]*0
for i=1:m
    for x=X
        if (x>=J(1,i) && x<J(2,i))
            IntStatR(i)=IntStatR(i)+1
        end
    end
end
IntStatR(m)=IntStatR(m)+1
%Task5
figure('Position', [380 400 560 420])
Ni = [0 IntStatR/(n*delta) 0]
Ji = [Mmin Mmin:delta:Mmax]
stairs(Ji, Ni)
hold on
Xn = Mmin:R/100:Mmax
Y = normpdf(Xn, mu, sqrt(Ssqr))
plot(Xn, Y)
grid
legend('histogramm','density function')
hold off
%Task6
figure('Position', [980 400 560 420])
[y, x] = ecdf(sort(X))
stairs(x, y)
hold on
Xn = Mmin:R/100:Mmax
Y = normcdf(Xn, mu, sqrt(Ssqr))
plot(Xn, Y)
grid
legend('histogramm','distribution function')
legend('Location','northwest')
hold off
```

Результаты расчетов

$M_{max} = -10.2500$

$M_{mun} = -14.6000$

$R = 4.3500$

$\overset{\wedge}{\mu} = -12.6148$

$S^2 = 0.8653$

$m = 8$

$(J_i, n_i) =$

J_i	[-14.6000;-14.0563)	[-14.0563;-13.5125)	[-13.5125;-12.9688)
n_i	4	18	20

J_i	[-12.9688;-12.4250)	[-12.4250;-11.8812)	[-11.8812;-11.3375)
n_i	36	16	14

J_i	[-11.3375;-10.7937)	[-10.7937;-10.2500]
n_i	6	6