

**Цель работы:** построение гистограммы и эмпирической функции распределения.

### Содержание работы

1. Для выборки объёма  $n$  из генеральной совокупности  $X$  реализовать в виде программы на ЭВМ:
  - вычисление максимального значения  $M_{\max}$  и минимального значения  $M_{\min}$ ,
  - размаха  $R$  выборки,
  - вычисление оценок  $\hat{\mu}$  и  $S^2$  математического ожидания  $MX$  и дисперсии  $DX$ ,
  - группировку значений выборки в  $m = [\log_2 n] + 2$  интервала,
  - построение на одной координатной плоскости гистограммы и графика функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием  $\hat{\mu}$  и дисперсией  $S^2$ ,
  - 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.

# 1      Вариант 21, выборка

$\vec{x} = ($   
-14.34, -16.97, -14.09, -14.74, -16.69, -13.85, -15.55, -14.62, -13.30, -15.52,  
-14.75, -16.51, -17.15, -16.87, -15.06, -13.60, -14.48, -14.71, -14.17, -13.88,  
-14.55, -15.37, -14.81, -16.05, -17.06, -15.86, -15.12, -15.98, -14.16, -15.81,  
-15.06, -16.19, -16.22, -16.19, -14.87, -15.62, -15.86, -15.25, -16.34, -14.44,  
-14.72, -15.17, -15.24, -14.44, -15.93, -14.87, -16.53, -15.76, -15.12, -12.91,  
-16.06, -16.06, -14.89, -15.57, -13.59, -16.84, -13.88, -14.33, -15.45, -16.58,  
-16.05, -14.34, -13.55, -16.78, -14.15, -14.28, -14.40, -13.98, -16.23, -15.35,  
-14.77, -15.61, -15.59, -15.64, -14.76, -17.18, -15.13, -15.01, -14.21, -13.91,  
-16.55, -15.44, -14.03, -16.44, -15.57, -15.07, -16.28, -16.30, -15.74, -14.03,  
-14.85, -15.73, -15.81, -14.42, -14.14, -15.14, -15.49, -16.42, -14.22, -14.20,  
-17.17, -15.82, -14.96, -14.75, -14.98, -13.64, -14.00, -17.29, -14.51, -16.18,  
-15.70, -15.07, -14.28, -14.55, -13.85, -15.36, -15.74, -14.61, -16.32, -15.34  
)

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

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

1. Максимальное значение выборки  $M_{max} = \max(x_1, \dots, x_n)$
2. Минимальное значение выборки  $M_{min} = \min(x_1, \dots, x_n)$
3. Размах выборки  $R = M_{max} - M_{min}$
4. Выборочное среднее (математическое ожидание)  $\hat{\mu}(\vec{x}) = \frac{1}{n} \sum_{i=1}^n x_i$
5. Состоятельная оценка дисперсии  $S^2(\vec{x}) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$   
где  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

## 3      Определение эмпирической плотности и гистограмм

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

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

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

Пусть  $\vec{x}$  – выборка из генеральной совокупности  $X$ . Если объем  $n$  этой выборки велик, то значения  $x_i$  группируют не только в статистический ряд, но и в интервальный статистический ряд. Для этого отрезок  $J = [x_{(1)}, x_{(n)}]$  (где  $x_{(1)} = \min(x_1, \dots, x_n)$ ,  $x_{(n)} = \max(x_1, \dots, x_n)$ ) делят на  $m$  равновеликих частей:

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

$$J_m = [a_m, a_{m+1}]$$

$$a_i = x_{(1)} + (i-1) \cdot \Delta, i = \overline{1; m+1}$$

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

Интервальным статистическим рядом называют таблицу

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

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

Требуемый интервал  $m = \lceil \log_2 n \rceil + 2$

Гистограмма – это график эмпирической плотности.

## 4 Определение эмпирической функции 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 условием  $F_n(x) = \frac{n(x, \vec{x})}{n}$ .

## 5 Листинг

Листинг 1: Реализация

```

1 function main()
2     X=[-14.34, -16.97, -14.09, -14.74, -16.69, ...
3         -13.85, -15.55, -14.62, -13.30, -15.52, ...
4         -14.75, -16.51, -17.15, -16.87, -15.06, ...
5         -13.60, -14.48, -14.71, -14.17, -13.88, ...
6         -14.55, -15.37, -14.81, -16.05, -17.06, ...
7         -15.86, -15.12, -15.98, -14.16, -15.81, ...
8         -15.06, -16.19, -16.22, -16.19, -14.87, ...
9         -15.62, -15.86, -15.25, -16.34, -14.44, ...
10        -14.72, -15.17, -15.24, -14.44, -15.93, ...
11        -14.87, -16.53, -15.76, -15.12, -12.91, ...
12        -16.06, -16.06, -14.89, -15.57, -13.59, ...
13        -16.84, -13.88, -14.33, -15.45, -16.58, ...

```

```

14         -16.05, -14.34, -13.55, -16.78, -14.15, ...
15         -14.28, -14.40, -13.98, -16.23, -15.35, ...
16         -14.77, -15.61, -15.59, -15.64, -14.76, ...
17         -17.18, -15.13, -15.01, -14.21, -13.91, ...
18         -16.55, -15.44, -14.03, -16.44, -15.57, ...
19         -15.07, -16.28, -16.30, -15.74, -14.03, ...
20         -14.85, -15.73, -15.81, -14.42, -14.14, ...
21         -15.14, -15.49, -16.42, -14.22, -14.20, ...
22         -17.17, -15.82, -14.96, -14.75, -14.98, ...
23         -13.64, -14.00, -17.29, -14.51, -16.18, ...
24         -15.70, -15.07, -14.28, -14.55, -13.85, ...
25         -15.36, -15.74, -14.61, -16.32, -15.34];
26
27     % 1)
28     % а) Максимальное и минимальное значения
29     Mmax = max(X);
30     Mmin = min(X);
31     fprintf("\na) Mmax (максимальное значение) = %f; Mmin (минимальное
        значение) = %f ", Mmax, Mmin);
32
33     % б) Размах
34     R = Mmax - Mmin;
35     fprintf("\nb) R (размах) = %f ", R);
36
37     % в) Оценки
38     mu = sum(X) / length(X);
39     s2 = sum((X - mean(X)).^2) / (length(X) - 1);
40     fprintf("\nv) mu (оценка математического ожидания) = %f; s^2 (оценк
        а дисперсии) = %f ", mu, s2);
41
42     % г) Группировка значений выборки
43     % Нахождение количества интервалов
44     m = floor(log2(length(X))) + 2;
45     fprintf("\ng) Группировка значений выборки в m = [log2 n] + 2 интер
        вала: m = %f\n", m);
46
47     % Разбиение выборки на интервалы от min до max, с помощью
        BinLimits
48     % объединяем только те значения, которые находятся в интервале от
        мин до макс
49     [counts, edges] = histcounts(X, m, 'BinLimits', [min(X), max(X)]);
50
51     for i = 1: length(counts)
52         fprintf("[%f : %f] - %d\n", edges(i), edges(i + 1), counts(i))
            ;
53     end
54
55     % д) Построение гистограммы
56     hist = histogram();
57     hist.BinEdges = edges;
58     hist.BinCounts = counts / length(X) / ((max(X) - min(X)) / m);
59

```

```

60 hold on; % Продолжаем работать с той же системой
61
62 % График функции плотности рапределения вероятностей нормальной сл
    учайно величины
63 delta = R/m;
64 sigma = sqrt(s2);
65 Xn = min(X):delta/20:max(X);
66 Y = normpdf(Xn, mu, sigma);
67 plot(Xn, Y, 'blue');
68
69 % е) График распределения сормальной случайной величины
70 figure;
71 [yy, xx] = ecdf(X);
72 stairs(xx, yy);
73
74 hold on;
75
76 % График эмпирической функции распределения нормальной случайной в
    еличины
77 delta = R/m;
78 Xn = min(X):delta/20:max(X);
79 Y = 1/2 * (1 + erf((Xn - mu) / sqrt(2*s2)));
80 plot(Xn, Y, 'black');
81
82 end

```

## 6 Результаты расчетов для выборки варианта 21

1. Максимальное значение  $M_{max} = -12.91$ ,
2. Минимальное значение выборки  $M_{min} = -17.29$ ,
3. Размах выборки  $R = 4.38$ ,
4. Оценки:  $\hat{\mu} = -15.220917, S^2 = 0.968029$ ,
5. Группировка значений выборки в  $m = [\log_2 n] + 2 = 8$  интервала предоставлена на рисунке 1.

```

а) Mmax (максимальное значение) = -12.910000; Mmin (минимальное значение) = -17.290000
б) R (размах) = 4.380000
в) m1 (оценка математического ожидания) = -15.220917; s^2 (оценка дисперсии) = 0.968029
г) Группировка значений выборки в m = [log2 n] + 2 интервала: m = 8.000000
[-17.290000 : -16.742500] - 9
[-16.742500 : -16.195000] - 13
[-16.195000 : -15.647500] - 19
[-15.647500 : -15.100000] - 22
[-15.100000 : -14.552500] - 21
[-14.552500 : -14.005000] - 23
[-14.005000 : -13.457500] - 11
[-13.457500 : -12.910000] - 2

```

Рис. 1: Поток вывода программы.

6. Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием  $\hat{\mu}$  и дисперсией  $S^2$ .

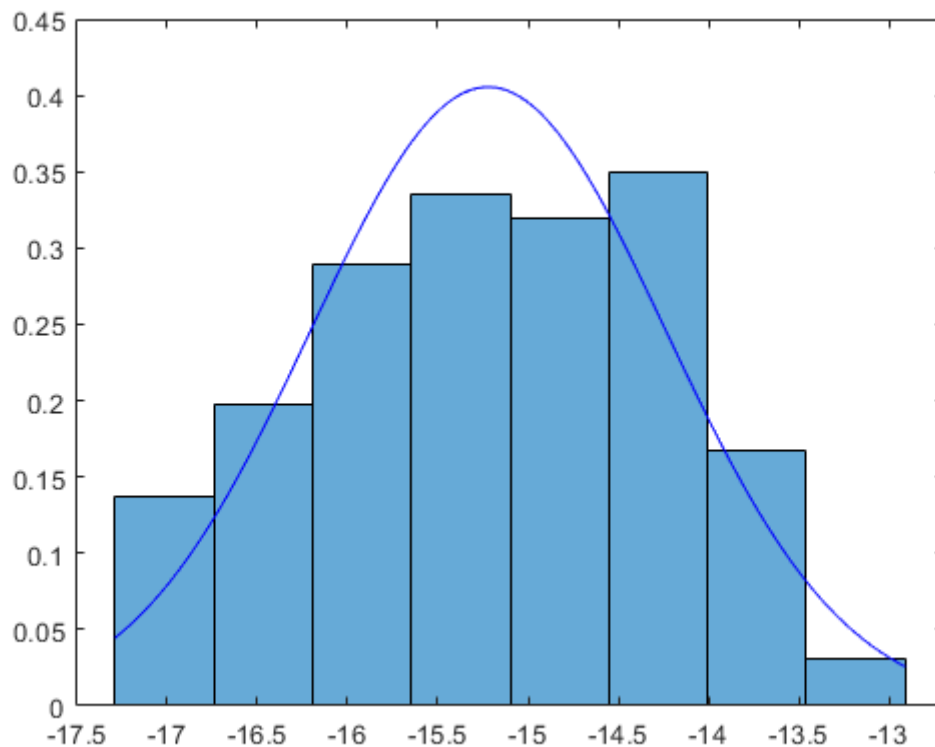


Рис. 2: Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины с математическим ожиданием  $\hat{\mu}$  и дисперсией  $S^2$ .

7. График эмпирической функции распределения и функции распределения нормальной случайной величины с математическим ожиданием  $\hat{\mu}$  и дисперсией  $S^2$ .

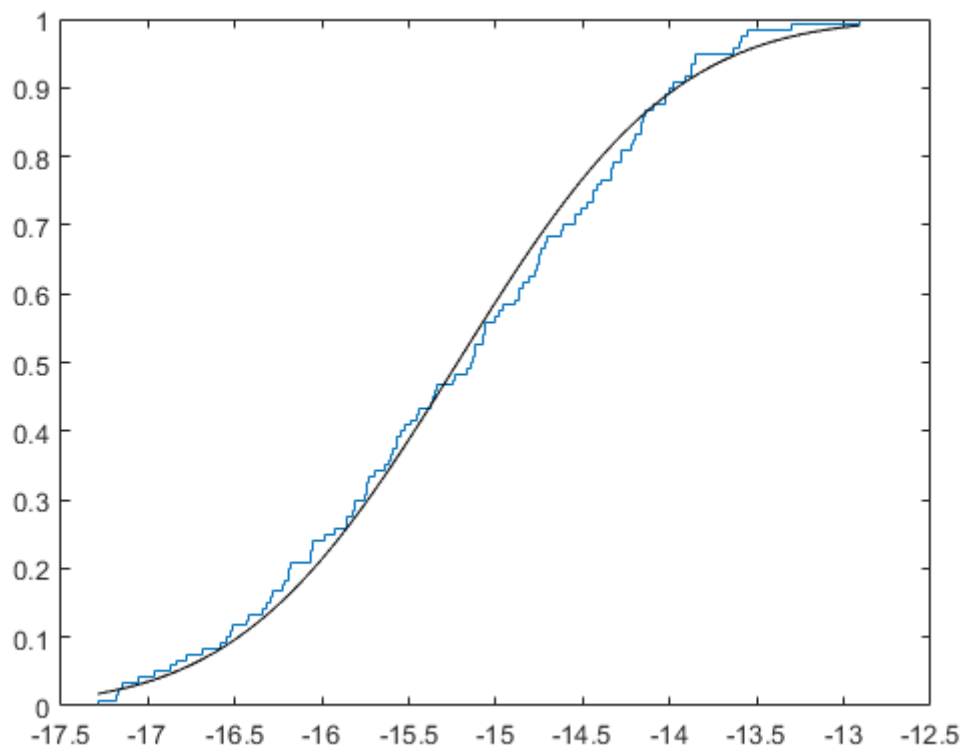


Рис. 3: График эмпирической функции распределения и функции распределения нормальной случайной величины с математическим ожиданием  $\hat{\mu}$  и дисперсией  $S^2$ .