Государственное образовательное учреждение высшего профессионального образования
«Московский Государственный Технический Университет имени Н. Э. Баумана»

Отчет по лабораторной работе №1 «Гистограмма и эмпирическая функция распределения» по курсу «Математическая статистика»

Вариант 1

Студент: Анисимов Н.С. ИУ7-62
Преподаватель: Велищанский М.А.

4 мая 2018 г.

Оглавление

1 Формулы и определения **2**

1.1 Формулы 2

1.2 Определение эмпирической плотности и гистограммы 2

1.3 Определение эмпирической функции распределения 3

2 Листинг **4**

3 Результаты **5**

1 Формулы и определения

1.1 Формулы

Минимальное значение выборки

$$M_{\min} = \min\{x_1, \ldots, x_n\} \tag{1.1}$$

- $(x_1, \ldots, x_n)$ — реализация случайной выборки.

Максимальное значение выборки

$$M_{\max} = \max\{x_1, \ldots, x_n\} \tag{1.2}$$

- $(x_1, \ldots, x_n)$ — реализация случайной выборки.

Размах выборки

$$R = M_{\max} - M_{\min} \tag{1.3}$$

- $M_{\max}$ — максимальное значение выборки;
- $M_{\min}$ — минимальное значение выборки.

Оценка математического ожидания

$$\hat{\mu}(\vec{x}_n) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1.4}$$

Смещённая оценка дисперсии

$$\hat{\sigma}^2(\vec{x}_n) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 \tag{1.5}$$

Несмещённая оценка дисперсии

$$S^2(\vec{x}_n) = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 \tag{1.6}$$

1.2 Определение эмпирической плотности и гистограммы

Эмпирической плотностью распределения, соответствующей выборке $\vec{x}$, называется функция:

$$f_n(x) = \begin{cases} \frac{n_i}{n\Delta}, & x \in J_i \\ 0, & x \notin \bigcup\limits_{i=1}^{m} J_i \end{cases} \tag{1.7}$$

- n_i — число элементов выборки $\vec{x}$, которые попали в интервал J_i;
- n — объем выборки;

График функции $f_n(x)$ называется гистограммой.

1.3 Определение эмпирической функции распределения

Пусть $\vec{X} = (X_1, \ldots, X_n)$ – случайная выборка, $\vec{x} = (x_1, \ldots, x_n)$ – реализация этой случайной выборки. Обозначим $n(x, \vec{x})$ – количество элементов вектора $\vec{x}$, которые меньше x.

Эмпирической функцией распределения, построенной по выборке $\vec{x}$, называют отображение $F_n \colon \Re \to \Re$, определенное правилом:

$$F_n(x) = \frac{n(x, \vec{x})}{n} \tag{1.8}$$

2 Листинг

```
data = csvread('data');

N = length(data);
M_min = min(data);
M_max = max(data);
R = range(data);
MX = mean(data);
DX = var(data,1);
S = var(data);

J_m = floor(log2(N)) + 2;
[gr, x1] = hist(data, J_m);
delta = R / J_m;
y1 = gr ./ (delta*N);

x2 = M_min-R : R/50 : M_max+R
y2 = normpdf(x2, MX, sqrt(S))

y3 = normcdf(x2, MX, sqrt(S));

figure
bar(x1, y1, 1);
hold on;
plot(x2, y2);
hold off;

figure
cdfplot(data);
hold on;
plot(x2, y3, 'r');
hold off;

printf("min: %f\nmax: %f\nrange: %f\nmx: %f\ndx: %f\ns: %f\n",
        M_min, M_max, R, MX, DX, S)

mxi = sum(x1 .* gr) / N;
dxi = sum((x1 - mxi).^2 .* gr) / N;
si = sum((x1 - mxi).^2 .* gr) / (N - 1);

printf("mxi: %f\ndxi: %f\nsi: %f\n",
        mxi,dxi,si)
```

3 Результаты

$$
\begin{aligned}
M_{\min} =& -4.110000\\
M_{\max} =& 1.400000\\
R =& 5.510000\\
\hat{\sigma}^2 =& -1.604583\\
S^2 =& 1.034091
\end{aligned}
$$

Интервальный статистический ряд

$[-4.110000;\ -3.421250)$	4
$[-3.421250;\ -2.732500)$	11
$[-2.732500;\ -2.043750)$	26
$[-2.043750;\ -1.355000)$	33
$[-1.355000;\ -0.666250)$	26
$[-0.666250;\ 0.022500)$	15
$[0.022500;\ 0.711250)$	2
$[0.711250;\ 1.400000]$	3

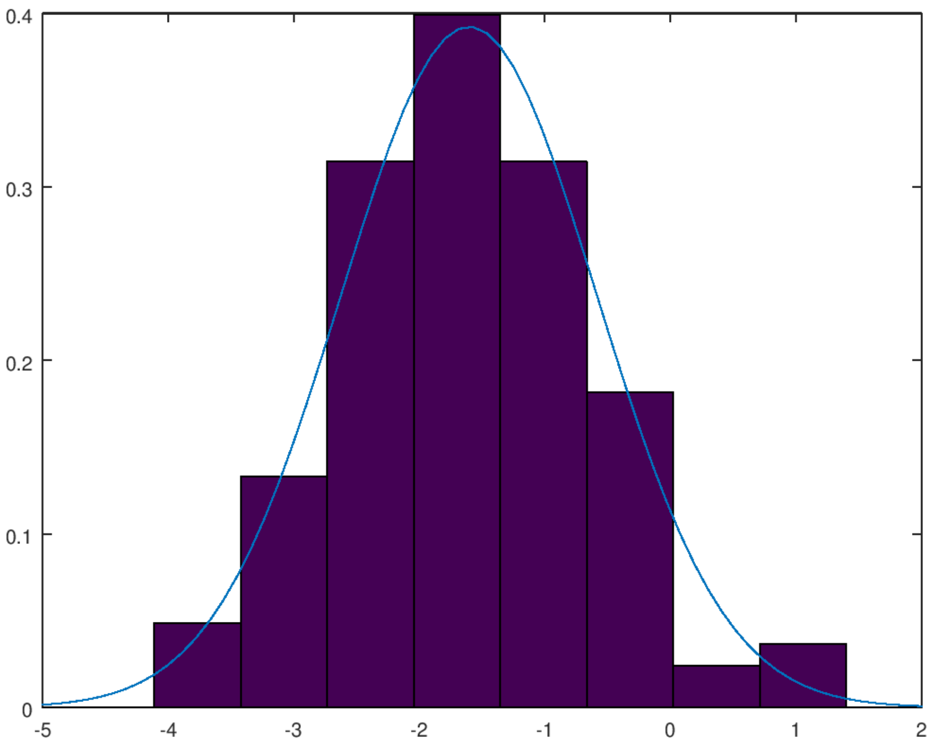

Рис. 3.1: Гистограмма и график функции плотности распределения вероятностей нормальной случайной величины

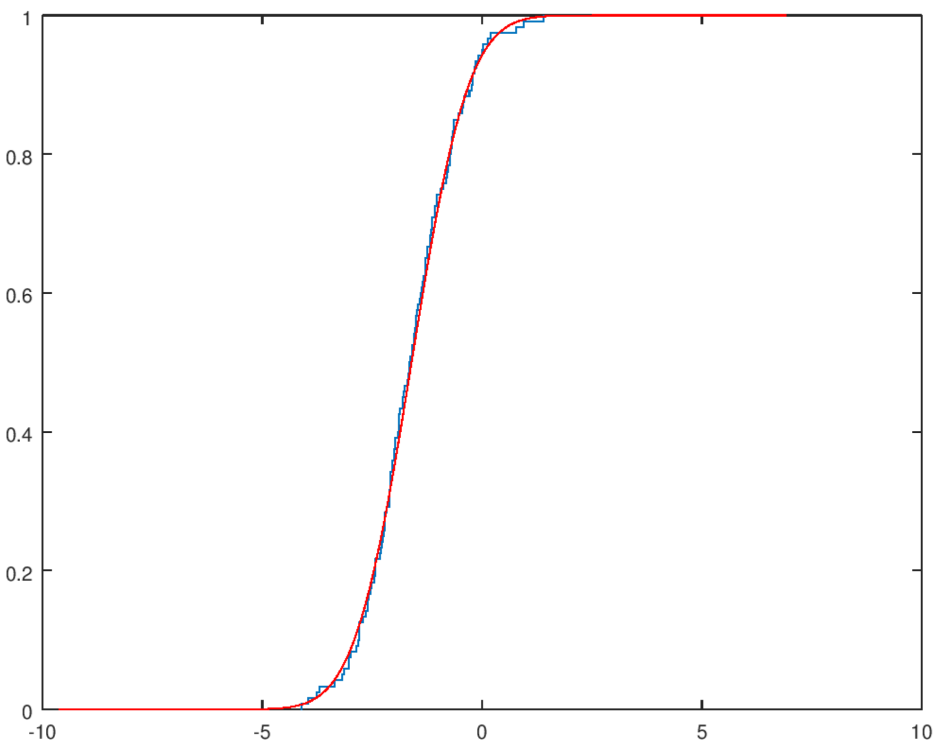

Рис. 3.2: Графики эмпирической функции распределения и функции распределения нормальной случайной величины