

Федеральное государственное автономное образовательное учреждение
высшего образования
«Национальный исследовательский университет ИТМО»
Факультет программной инженерии и компьютерной техники

Практическая работа №5 по дисциплине
Теория вероятностей и математическая статистика

Вариант №12

Выполнил: Селянта Олег Дмитриевич
Группа: Р3214
Преподаватель: Селина Елена Георгиевна

Санкт-Петербург 2023г.

Задание

Каждый студент получает выборку из 20 чисел. Необходимо определить следующие статистические характеристики: вариационный ряд, экстремальные значения и размах, оценки математического ожидания и среднеквадратического отклонения, эмпирическую функцию распределения и её график, гистограмму и полигон приведенных частот группированной выборки. Для расчета характеристик и построения графиков нужно написать программу на одном из языков программирования. Листинг программы и результаты работы должны быть представлены в отчете по практической работе

Вариант №12

| | | | | | | | | | | | | | | | | | | | |
|------|------|-------|------|------|------|------|------|------|-------|------|------|-------|-------|------|-------|------|------|------|------|
| 0.41 | 1.63 | -1.53 | -0.2 | 0.85 | 0.09 | 1.54 | 0.25 | 1.24 | -0.26 | 1.08 | 0.42 | -0.92 | -0.91 | 1.15 | -0.82 | 0.26 | 0.96 | 1.57 | 0.72 |
|------|------|-------|------|------|------|------|------|------|-------|------|------|-------|-------|------|-------|------|------|------|------|

Выполнение

Вариационный ряд

| | | | | | | | | | | | | | | | | | | | |
|-------|-------|-------|-------|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| -1.53 | -0.92 | -0.91 | -0.82 | -0.26 | -0.2 | 0.09 | 0.25 | 0.26 | 0.41 | 0.42 | 0.72 | 0.85 | 0.96 | 1.08 | 1.15 | 1.24 | 1.54 | 1.57 | 1.63 |
|-------|-------|-------|-------|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|

Статистический ряд

| | | | | | | | | | | | | | | | | | | | |
|-------|-------|-------|-------|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| -1.53 | -0.92 | -0.91 | -0.82 | -0.26 | -0.2 | 0.09 | 0.25 | 0.26 | 0.41 | 0.42 | 0.72 | 0.85 | 0.96 | 1.08 | 1.15 | 1.24 | 1.54 | 1.57 | 1.63 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Экстремальные значения

Минимальное значение: $x_0 = -1.53$; Максимальное значение: $x_{20} = 1.63$

Размах

Размах = наибольшее значение - наименьшее значение = $x_{20} - x_0$; Размах = 3.16

Математическое ожидание

Для данной выборки статистический ряд будет совпадать с вариационным, так как каждое значение встречается только раз

Выборочное среднее (выборочное математическое ожидание) - среднее арифметическое всех значений выборки, считается по формуле:

$$\overline{x_B} = \frac{1}{n} \sum_{i=1}^k x_i \cdot n_i$$

Для исходной выборки

$$\overline{x_B} = 0.3765$$

Выборочная дисперсия

Выборочная дисперсия D_B - среднее арифметическое квадратов отклонений значений выборки от выборочной средней $\overline{x_B}$, считается по формуле:

$$D_B = \frac{1}{n} \sum_{i=1}^k (x_i - \overline{x_B})^2 \cdot n_i$$

Для исходной выборки

$$D_B = 0.3765$$

Среднеквадратическое отклонение

Выборочное среднее квадратическое отклонение выборки определяется формулой

$$\sigma_B = \sqrt{D_B}$$

Для исходной выборки

$$\sigma_B = 0.897927$$

Исправленное выборочное среднее квадратическое отклонение

При решении практических используется величина

$$S^2 = \frac{n}{n-1} \sum_{i=1}^k (x_i - \bar{x}_B)^2 \cdot n_i = \frac{n}{n-1} D_B$$

Которая называется *исправленной выборочной дисперсией*

Величина $S = \sqrt{S^2}$ называется исправленным выборочным средним квадратическим отклонением

Для исходной выборки

$$S = 0.921254$$

Эмпирическая функция распределения

Эмпирической (статистической) функцией распределения называется функция $F_n^x(x)$, определяющая для каждого значения x частность события $\{X < x\}$:

$$F_n^*(x) = p^* \{X < x\}$$

Где $p^x = \frac{n_x}{n}$ - отношение количесвта вариантов $\{X < x\}$ к общему числу вариантов

Для исходной выборки

$$F_n^*(x) = \begin{cases} 0 & \text{при } x < -1.53 \\ 0.05 & \text{при } -1.53 \leq x < -0.92 \\ 0.1 & \text{при } -0.92 \leq x < -0.91 \\ 0.15 & \text{при } -0.91 \leq x < -0.82 \\ 0.2 & \text{при } -0.82 \leq x < -0.26 \\ 0.25 & \text{при } -0.26 \leq x < -0.2 \\ 0.3 & \text{при } -0.2 \leq x < 0.09 \\ 0.35 & \text{при } 0.09 \leq x < 0.25 \\ 0.4 & \text{при } 0.25 \leq x < 0.26 \\ 0.45 & \text{при } 0.26 \leq x < 0.41 \\ 0.5 & \text{при } 0.41 \leq x < 0.42 \\ 0.55 & \text{при } 0.42 \leq x < 0.72 \\ 0.6 & \text{при } 0.72 \leq x < 0.85 \\ 0.65 & \text{при } 0.85 \leq x < 0.96 \\ 0.7 & \text{при } 0.96 \leq x < 1.08 \\ 0.75 & \text{при } 1.08 \leq x < 1.15 \\ 0.8 & \text{при } 1.15 \leq x < 1.24 \\ 0.85 & \text{при } 1.24 \leq x < 1.54 \\ 0.9 & \text{при } 1.54 \leq x < 1.57 \\ 0.95 & \text{при } 1.57 \leq x < 1.63 \\ 1 & \text{при } x \geq 1.63 \end{cases}$$

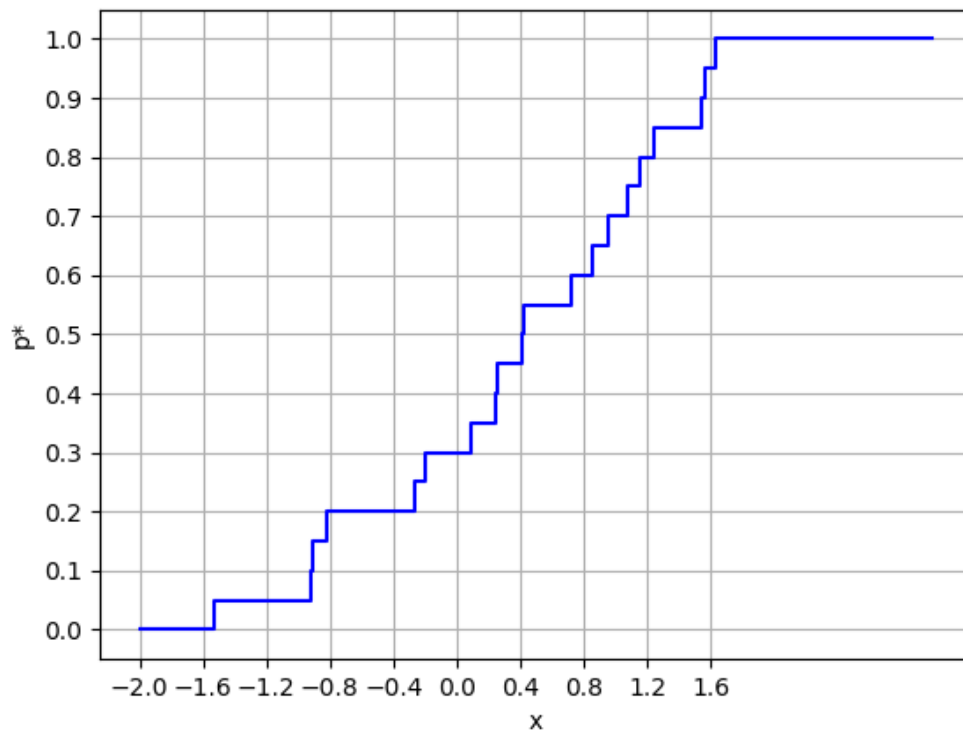


Рис. 1: График эмпирической функции распределения

Интервальный статистический ряд

Так как признак является непрерывным, то имеет смысл составить интервальный статический ряд (для дальнейшего использования в функции распределения). Пользуясь формулой Стерджеса, найдем величину интервала

$$h = \frac{x_{max} - x_{min}}{1 + \log_2 n}$$

Для исходной выборки

$$h = 0.59377$$

Учитывая рекомендацию по выбору начала первого интервала $x = x_{min} - \frac{h}{2}$

| | | | | | |
|-------------------|-------------------|-------------------|------------------|----------------|----------------|
| [-1.8269;-1.2331) | [-1.2331;-0.6393) | [-0.6393;-0.0456) | [-0.0456;0.5482) | [0.5482;1.142) | [1.142;1.7357) |
| 1 | 3 | 2 | 5 | 4 | 5 |

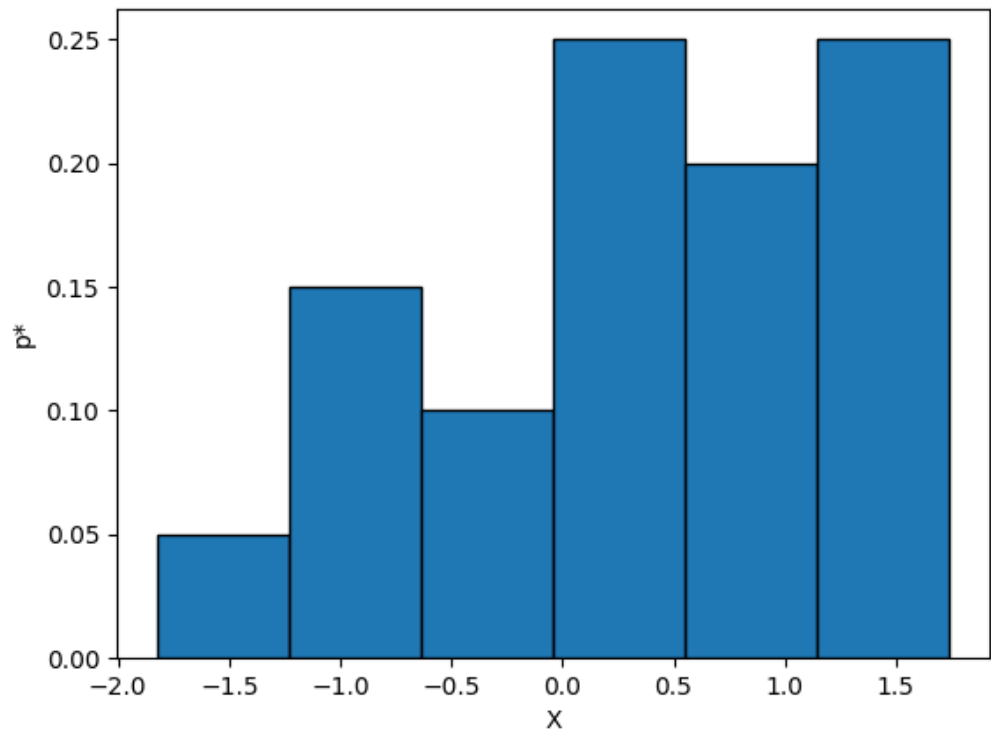


Рис. 2: Гистограмма

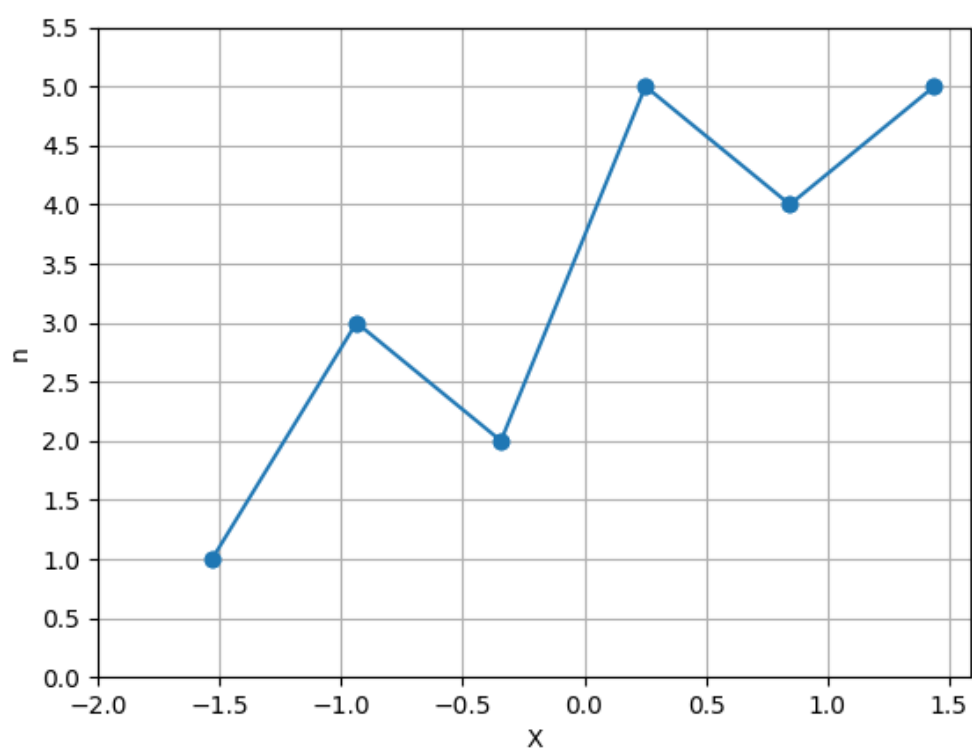


Рис. 3: Полигон приведенных частот группированной выборки