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 №9

Выполнила: Минкова Алина Андреевна
Группа: P3214
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 №14

-1.35	0.38	0.35	0.8	-1.49	-0.36	-0.8	0.55	-0.46	-0.6	-0.42	1.21	1.56	0.14	0.35	-1.57	-0.15	0.2	-1.63	0.05

Выполнение

Вариационный ряд

-1.63	-1.57	-1.49	-1.35	-0.8	-0.6	-0.46	-0.42	-0.36	-0.15	0.05	0.14	0.2	0.35	0.35	0.38	0.55	0.8	1.21	1.56

Статистический ряд

-1.63	-1.57	-1.49	-1.35	-0.8	-0.6	-0.46	-0.42	-0.36	-0.15	0.05	0.14	0.2	0.35	0.38	0.55	0.8	1.21	1.56
1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1

Экстремальные значения

Минимальное значение: x_0 = -1.63; Максимальное значение: x_{20} = 1.56

Размах

Размах = наибольшее значение - наименьшее значение = x_{20} - x_0; Размах = 3.19

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

$$\overline{x_B} = -0.16199999999999998$$

Выбороная дисперсия

Выбороная 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

$$D_B = 0.7735660000000002$$

Среднеквадратическое отклонение

Выборочное среднее квадратическое отклонение выборки определяется формулой

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

Для исходной выборки

$$\sigma_B = 0.879526$$

Исправленное выборочное среднее квадратическое отклонение

При решении практических используется величина

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

Которая называется *исправленной выборочной дисперсией*

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

Для исходной выборки

$$S = 0.902375$$

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

$$F_n^*(x) = \begin{cases} 0 & \text{при } x < -1.63 \\ 0.05 & \text{при } -1.63 \le x < -1.57 \\ 0.1 & \text{при } -1.57 \le x < -1.49 \\ 0.15 & \text{при } -1.49 \le x < -1.35 \\ 0.2 & \text{при } -1.35 \le x < -0.8 \\ 0.25 & \text{при } -0.8 \le x < -0.6 \\ 0.3 & \text{при } -0.6 \le x < -0.46 \\ 0.35 & \text{при } -0.46 \le x < -0.42 \\ 0.4 & \text{при } -0.42 \le x < -0.36 \\ 0.45 & \text{при } -0.36 \le x < -0.15 \\ 0.5 & \text{при } -0.15 \le x < 0.05 \\ 0.55 & \text{при } 0.05 \le x < 0.14 \\ 0.6 & \text{при } 0.14 \le x < 0.2 \\ 0.65 & \text{при } 0.2 \le x < 0.35 \\ 0.7 & \text{при } 0.35 \le x < 0.35 \\ 0.75 & \text{при } 0.35 \le x < 0.38 \\ 0.8 & \text{при } 0.38 \le x < 0.55 \\ 0.85 & \text{при } 0.55 \le x < 0.8 \\ 0.9 & \text{при } 0.8 \le x < 1.21 \\ 0.95 & \text{при } 1.21 \le x < 1.56 \\ 1 & \text{при } x \ge 1.56 \end{cases}$$

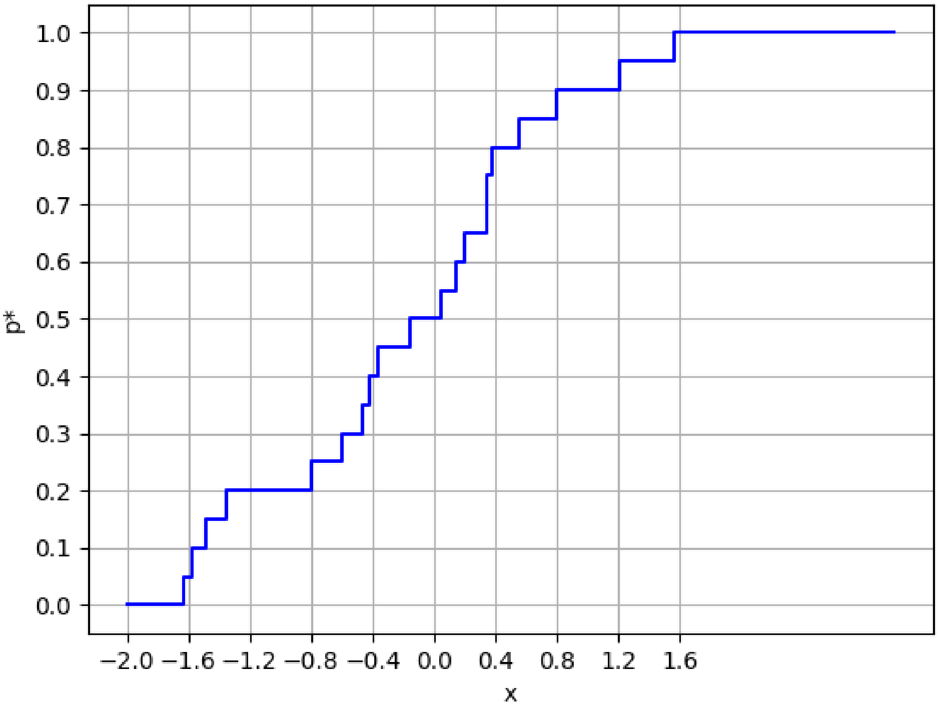

Рис. 1: График эмпирической функции распределения

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

$$h = 0.599407$$

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

[-1.9297;-1.3303)	[-1.3303;-0.7309)	[-0.7309;-0.1315)	[-0.1315;0.4679)	[0.4679;1.0673)	[1.0673;1.6667)
4	1	5	6	2	2

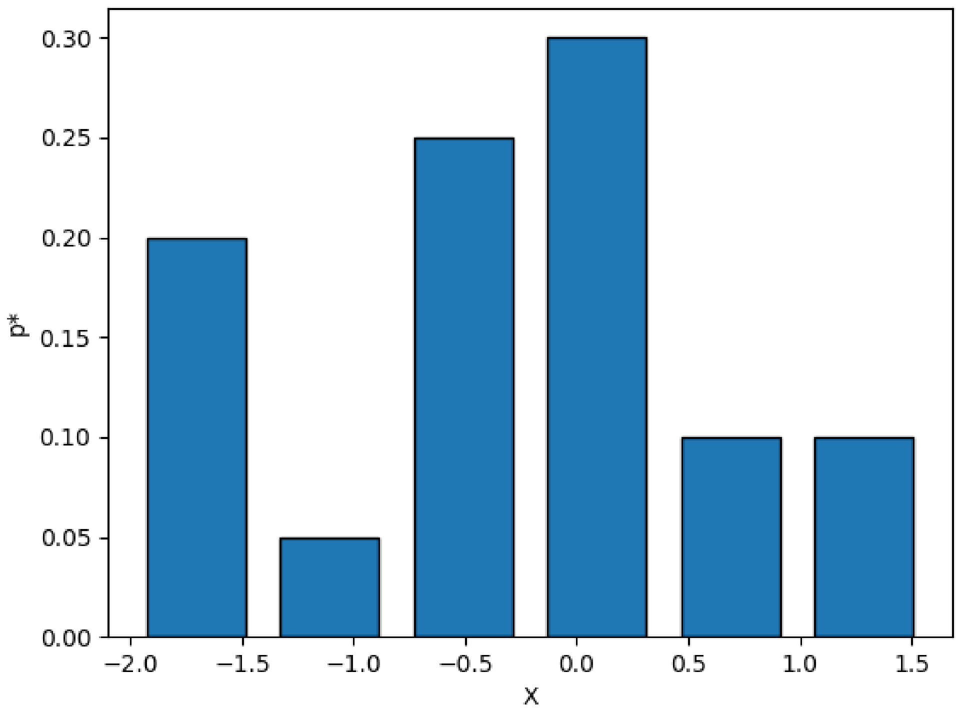

Рис. 2: Гистограмма

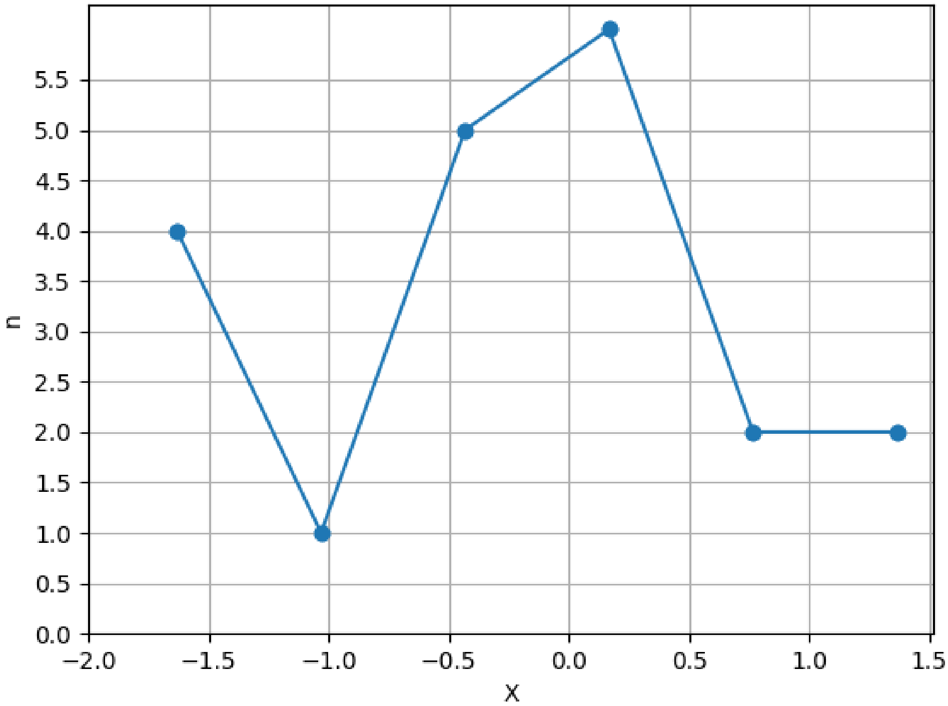

Рис. 3: Полигон приведенных частот группированной выборки