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 №14

Выполнил: Состанов Тимур Айратович
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

-0.53	-0.87	-0.93	-0.41	0.48	0.81	-1.55	-1.42	-1.34	-0.61	-0.04	-0.33	-0.84	-1.33	0.57	0.62	0.76	-0.48	0.3	-0.35

Выполнение

Вариационный ряд

-1.55	-1.42	-1.34	-1.33	-0.93	-0.87	-0.84	-0.61	-0.53	-0.48	-0.41	-0.35	-0.33	-0.04	0.3	0.48	0.57	0.62	0.76	0.81

Статистический ряд

-1.55	-1.42	-1.34	-1.33	-0.93	-0.87	-0.84	-0.61	-0.53	-0.48	-0.41	-0.35	-0.33	-0.04	0.3	0.48	0.57	0.62	0.76	0.81
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 = -1.55; Максимальное значение: x_{20} = 0.81

Размах

Размах = наибольшее значение - наименьшее значение = x_{20} - x_0; Размах = 2.36

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

$$\overline{x_B} = -0.3745$$

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

$$D_B = -0.3745$$

Среднеквадратическое отклонение

Выборочное среднее квадратическое отклонение выборки определяется формулой

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

Для исходной выборки

$$\sigma_B = 0.743428$$

Исправленное выборочное среднее квадратическое отклонение

При решении практических используется величина

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

Которая называется *исправленной выборочной дисперсией*

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

Для исходной выборки

$$S = 0.762741$$

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

$$F_n^*(x) = \begin{cases} 0 & \text{при } x < -1.55 \\ 0.05 & \text{при } -1.55 \leq x < -1.42 \\ 0.1 & \text{при } -1.42 \leq x < -1.34 \\ 0.15 & \text{при } -1.34 \leq x < -1.33 \\ 0.2 & \text{при } -1.33 \leq x < -0.93 \\ 0.25 & \text{при } -0.93 \leq x < -0.87 \\ 0.3 & \text{при } -0.87 \leq x < -0.84 \\ 0.35 & \text{при } -0.84 \leq x < -0.61 \\ 0.4 & \text{при } -0.61 \leq x < -0.53 \\ 0.45 & \text{при } -0.53 \leq x < -0.48 \\ 0.5 & \text{при } -0.48 \leq x < -0.41 \\ 0.55 & \text{при } -0.41 \leq x < -0.35 \\ 0.6 & \text{при } -0.35 \leq x < -0.33 \\ 0.65 & \text{при } -0.33 \leq x < -0.04 \\ 0.7 & \text{при } -0.04 \leq x < 0.3 \\ 0.75 & \text{при } 0.3 \leq x < 0.48 \\ 0.8 & \text{при } 0.48 \leq x < 0.57 \\ 0.85 & \text{при } 0.57 \leq x < 0.62 \\ 0.9 & \text{при } 0.62 \leq x < 0.76 \\ 0.95 & \text{при } 0.76 \leq x < 0.81 \\ 1 & \text{при } x \geq 0.81 \end{cases}$$

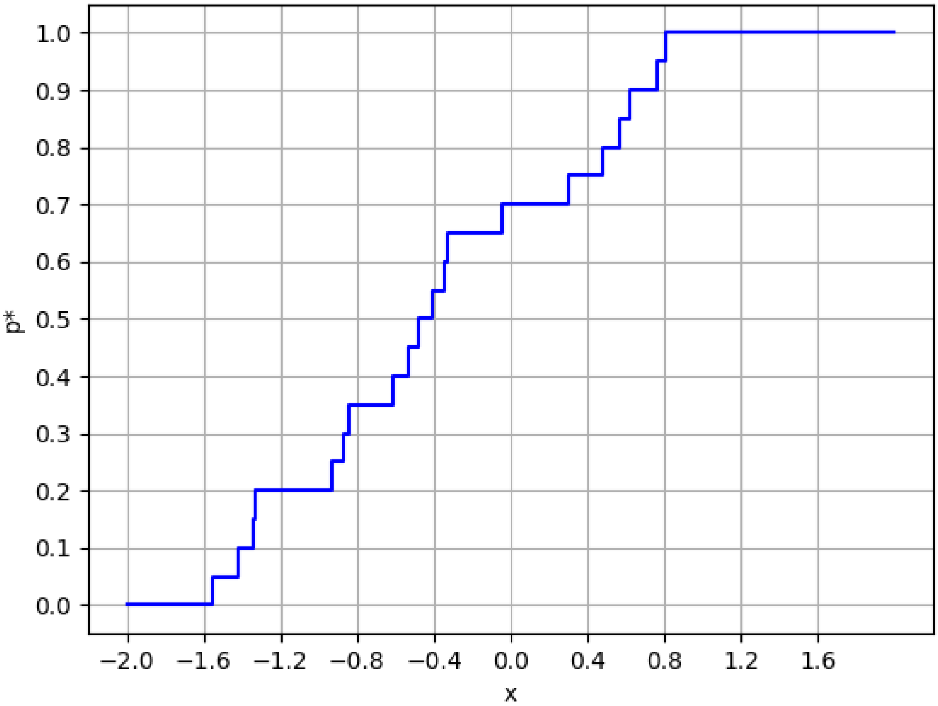

Рис. 1: График эмпирической функции распределения

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

$$h = 0.443448$$

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

[-1.7717;-1.3283)	[-1.3283;-0.8848)	[-0.8848;-0.4414)	[-0.4414;0.0021)	[0.0021;0.4455)	[0.4455;0.889)
4	1	5	4	1	5

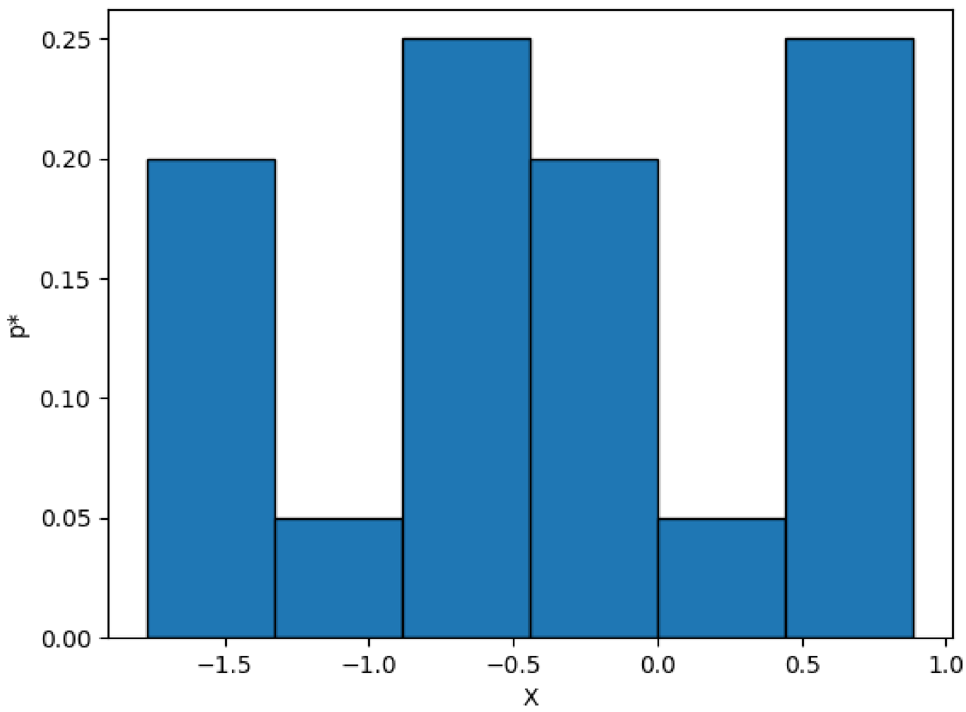

Рис. 2: Гистограмма

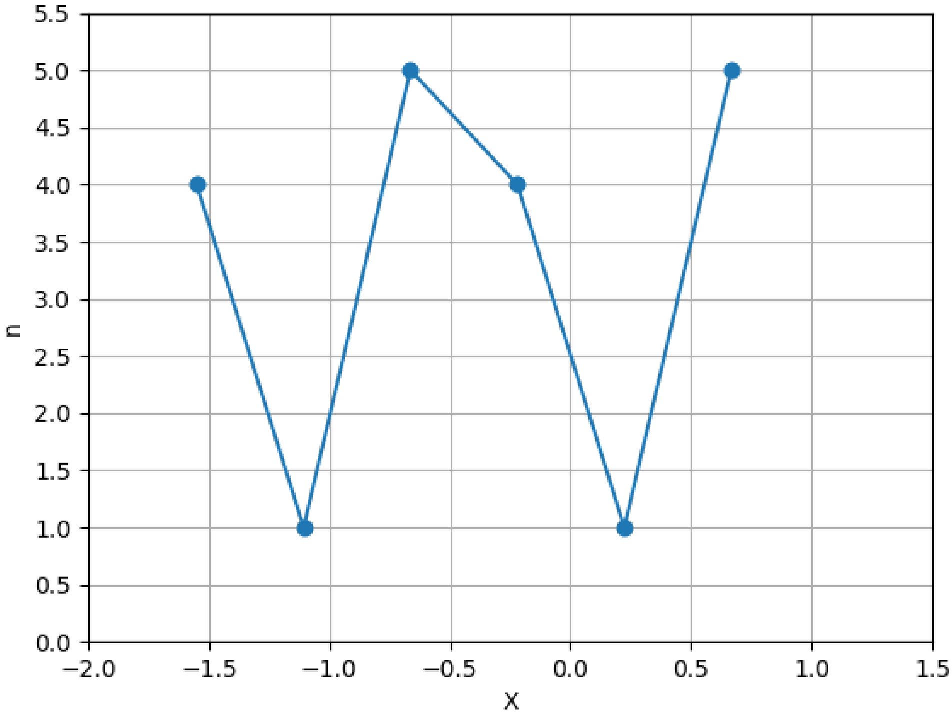

Рис. 3: Полигон приведенных частот группированной выборки