

Санкт-Петербургский Политехнический университет
Петра Великого

Физико-механический институт

Высшая школа прикладной математики и вычислительной физики

Отчёт
по лабораторной работе №1
по дисциплине
"Анализ данных с интервальной
неопределенностью"

Выполнил студент:

Иванов Андрей Игоревич,
группа 5040102\20201

Проверил:

к.ф.-м.н., доцент
Баженов Александр Николаевич

Санкт-Петербург, 2023

Содержание

1	Постановка задачи	3
2	Теория	4
2.1	Интервальная выборка	4
2.2	Индекс Жаккара	4
2.3	Оптимальный корректирующий множитель	4
3	Реализация	6
4	Результаты	7
4.1	Выборка X_+	7
4.2	Выборка X_-	8
4.3	Поиск корректирующего множителя R_{opt}	10
4.4	Выборка $X_+ \cup R_{opt} \cdot X_-$	12

Список иллюстраций

1	График интервальной выборки	7
2	Гистограмма отступов	7
3	График частоты моды	8
4	График интервальной выборки	9
5	Гистограмма отступов	9
6	График частоты моды	10
7	График зависимости $JK(X)$ от R_{opt}	11
8	График интервальной выборки	12
9	График частоты моды	12

1 Постановка задачи

Даны две вещественные выборки X_+, X_- . Необходимо:

- Сформировать интервальные выборки $\mathbf{X}_+, \mathbf{X}_-$;
- Реализовать алгоритм вычисления индекса Жаккара;
- Найти оптимальный корректирующий множитель R_{opt} такой, что выборка $\mathbf{X}_+ \cup R_{opt} \cdot \mathbf{X}_-$ была наиболее совместной в смысле индекса Жаккара;
- Реализовать алгоритм поиска моды выборки;

2 Теория

2.1 Интервальная выборка

Формирование выборки происходит следующим образом:

Пусть X - выборка вещественных чисел. Тогда соответствующая интервальная выборка \mathbf{X} определяется как:

$$\mathbf{X} = \{(x_i - \delta; x_i + \delta) | x_i \in X, i \in \overline{1, |X|}, \delta = \frac{1}{2^{14}}\} \quad (1)$$

2.2 Индекс Жаккара

Степень совместности двух интервалов x и y может быть определена как коэффициент Жаккара:

$$JK(x, y) = \frac{wid(x \wedge y)}{wid(x \vee y)} \quad (2)$$

где \wedge и \vee — операции взятия минимума и максимума по включению в полной интервальной арифметике Каухера. Для данной меры совместности справедливо:

$$-1 \leq JK(x, y) \leq 1 \quad (3)$$

Введенная числовая характеристика $JK(x, y)$ может быть естественным образом обобщена на случай интервальной выборки $X = \{x_k\}_{k=1}^n, k = \overline{1, n}$ для определения ее меры совместности

$$JK(X) = \frac{wid(\bigwedge_k x_k)}{wid(\bigvee_k x_k)} \quad (4)$$

Выражение индекса Жаккара можно переписать используя операции взятия минимума и максимума в соответствии с определением ширины интервала в классической интервальной арифметике:

$$JK(X) = \frac{\min(\overline{x_k}) - \max(\underline{x_k})}{\max(\overline{x_k}) - \min(\underline{x_k})} \quad (5)$$

2.3 Оптимальный корректирующий множитель

Для нахождения R_{opt} необходимо получить нижнюю и верхнюю границы для последующего уточнения:

$$\underline{R} = \frac{\min(\underline{x_i})}{\max(\overline{x_i})} \quad (6)$$

$$\overline{R} = \frac{\max(\underline{x_i})}{\min(\overline{x_i})} \quad (7)$$

Уточнение R_{opt} производится либо итерационно с заданным шагом, либо методом половинного деления. Ввиду лучшей скорости для уточнения был использован метод половинного деления, для наглядного отображения зависимости величины R_{opt} от индекса Жаккара использовался итерационный способ с шагом $\frac{wid(\bigcup X)}{100}$.

3 Реализация

Лабораторная работа выполнена на языке Python 3.10 с помощью загружаемых пакетов NumPy и Matplotlib. Исходный код лабораторной работы находится на GitHub репозитории.

4 Результаты

4.1 Выборка X_+

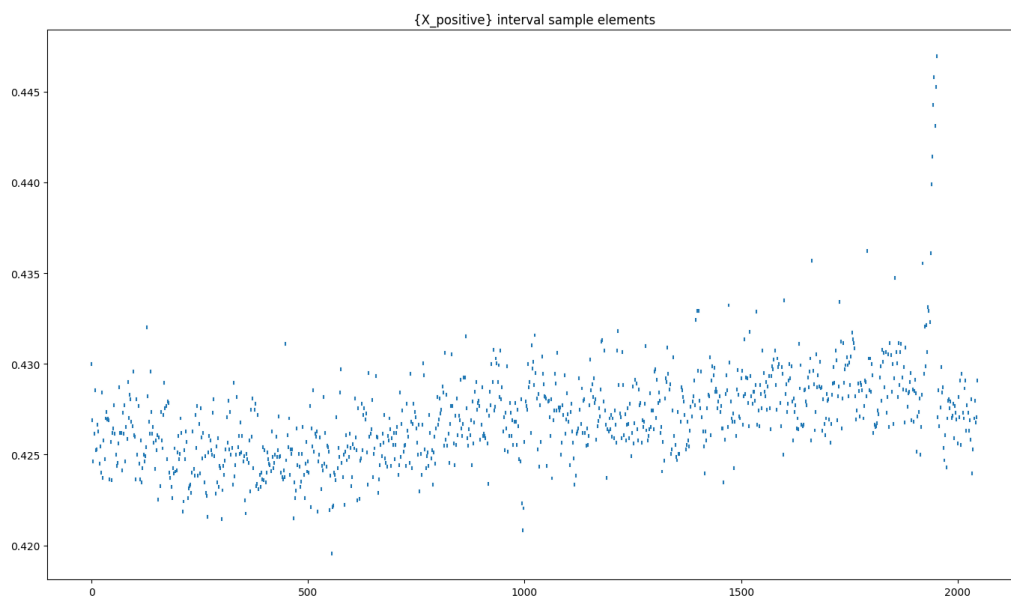


Рис. 1: График интервальной выборки

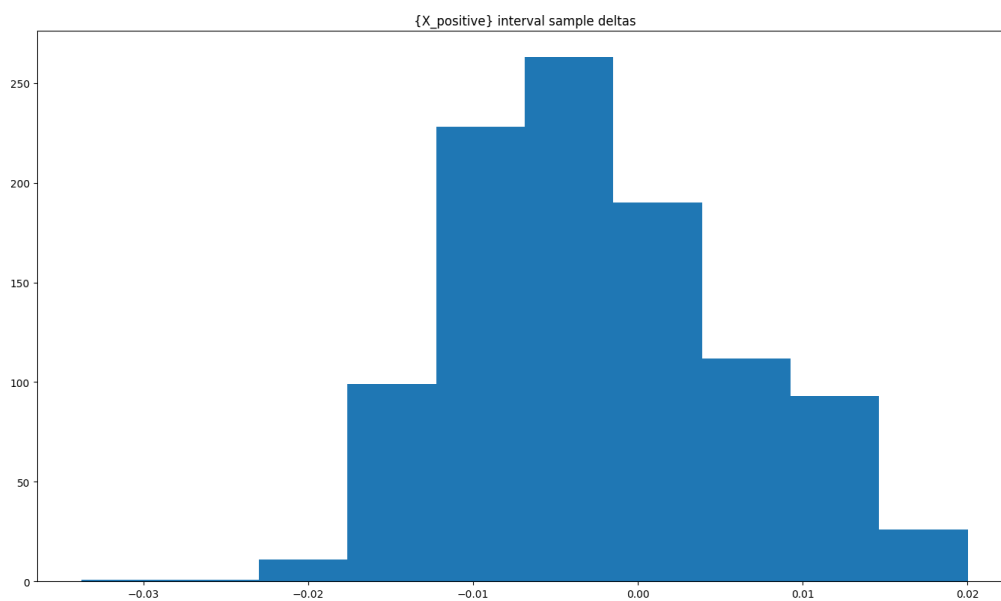


Рис. 2: Гистограмма отступов

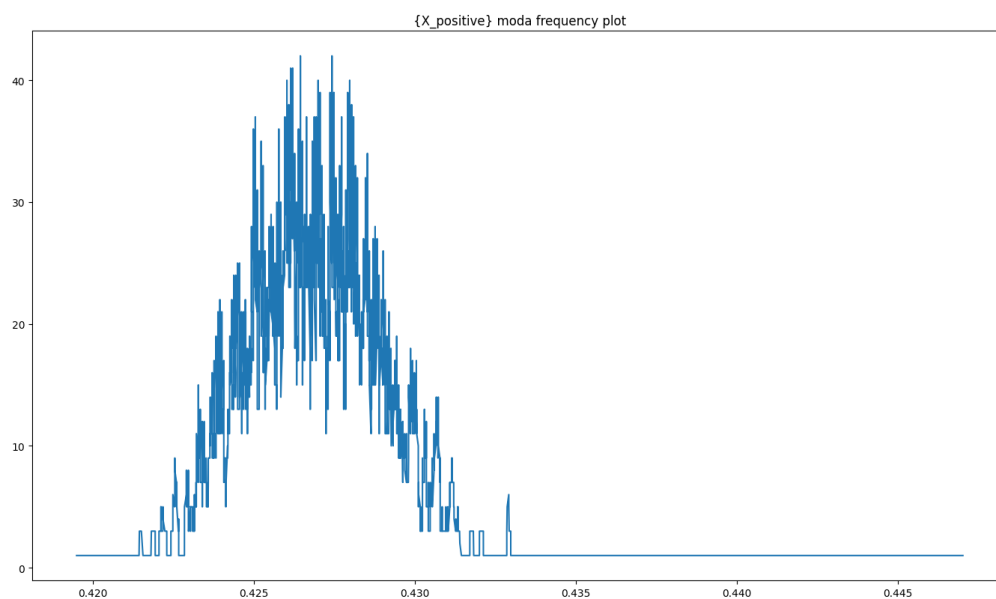


Рис. 3: График частоты моды

Данные для формирования интервальной выборки X_+ были считаны из файлов

- +0_5V_4_.txt
- ZeroLine_4.txt

Мода выборки: $[0.42645, 0.42743]$

На показаниях с порядковым номером в районе 1950 наблюдаются выбросы.

4.2 Выборка X_-

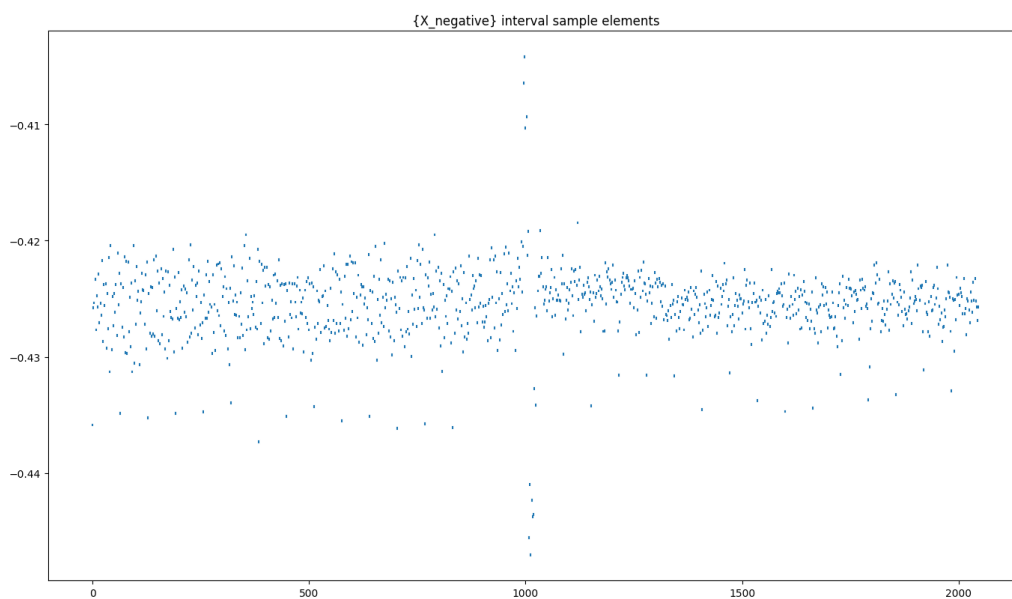


Рис. 4: График интервальной выборки

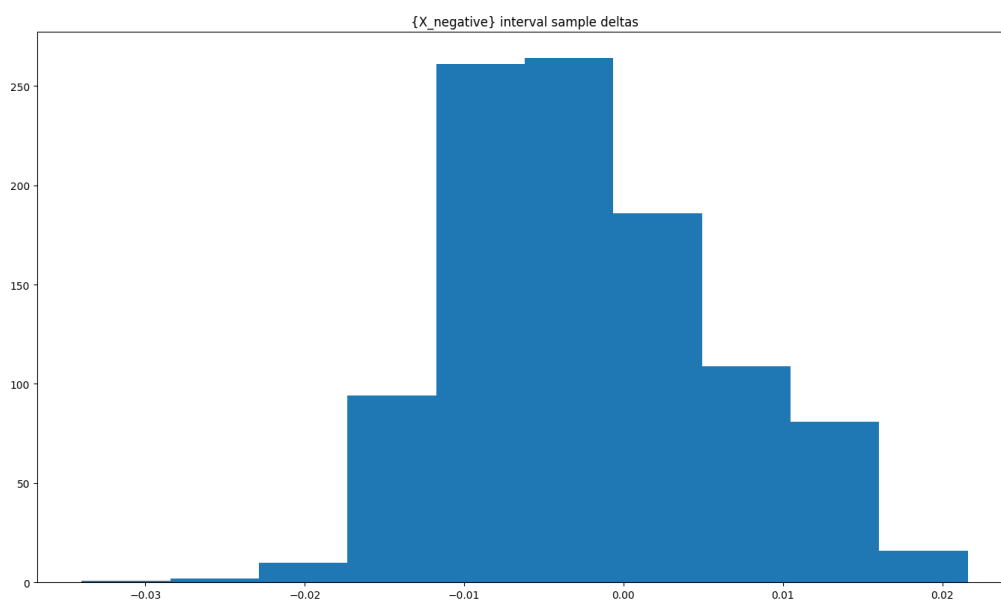


Рис. 5: Гистограмма отступов

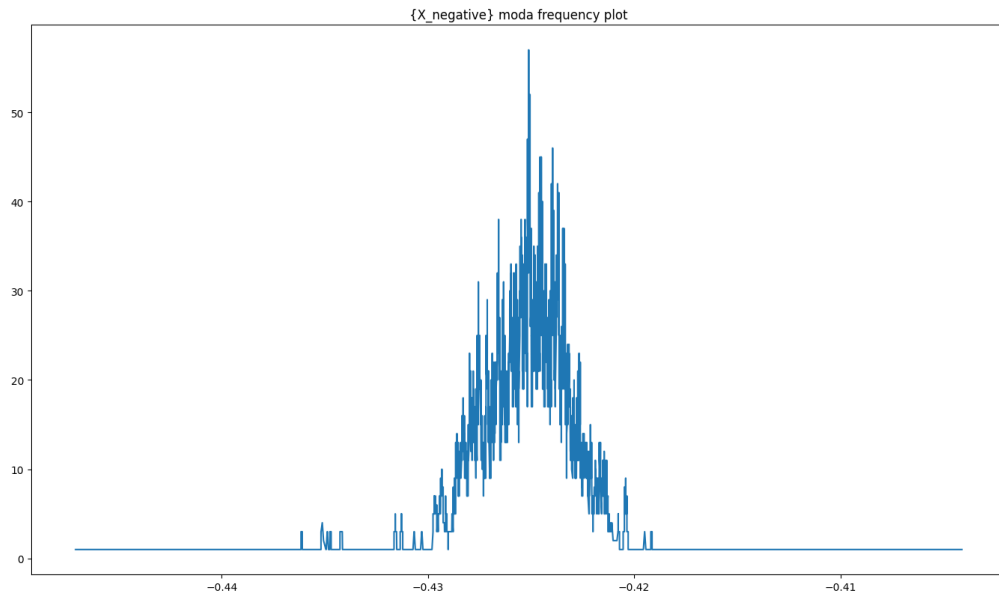


Рис. 6: График частоты моды

Данные для формирования интервальной выборки X_+ были считаны из файлов

- -0_5V_8_.txt
- ZeroLine_8.txt

Мода выборки: $[-0.42511, -0.42511]$

На показаниях с порядковым номером в районе 1000 наблюдаются выбросы.

4.3 Поиск корректирующего множителя R_{opt}

Было получено значение корректирующего множителя

$$R_{opt} = -1.0000042926186499 \quad (8)$$

с соответствующим индексом Жаккара:

$$JK(X_+ \cup R_{opt} \cdot X_-) = -0.9971593104169201 \quad (9)$$

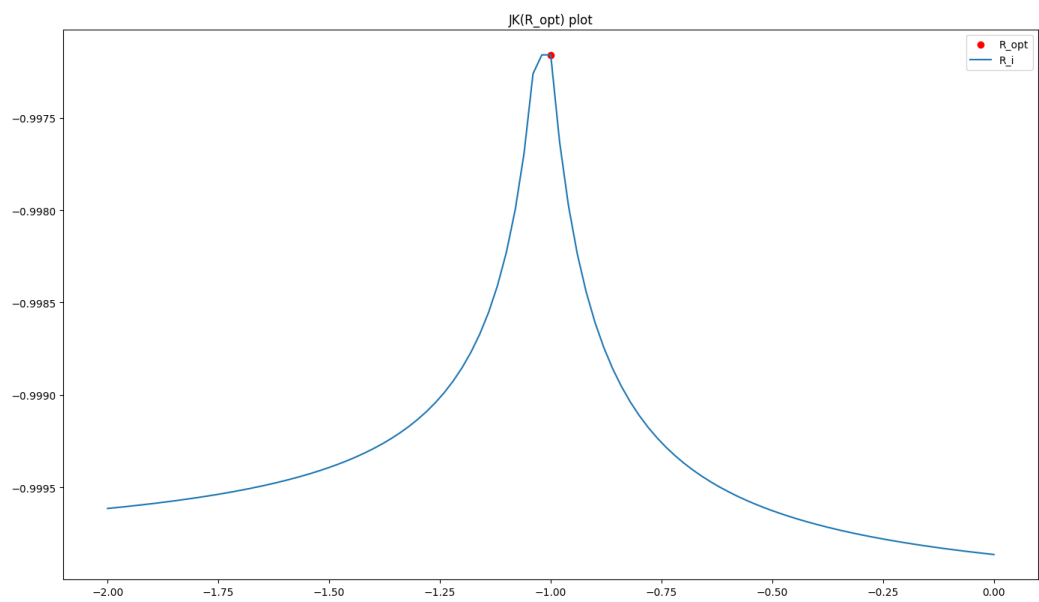


Рис. 7: График зависимости $JK(X)$ от R_{opt}

4.4 Выборка $X_+ \cup R_{opt} \cdot X_-$

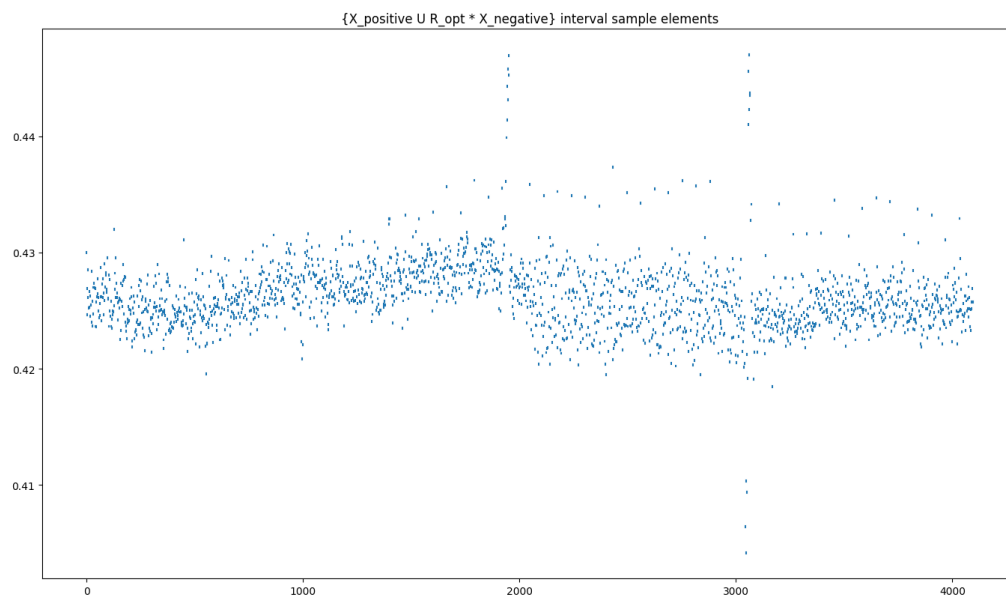


Рис. 8: График интервальной выборки

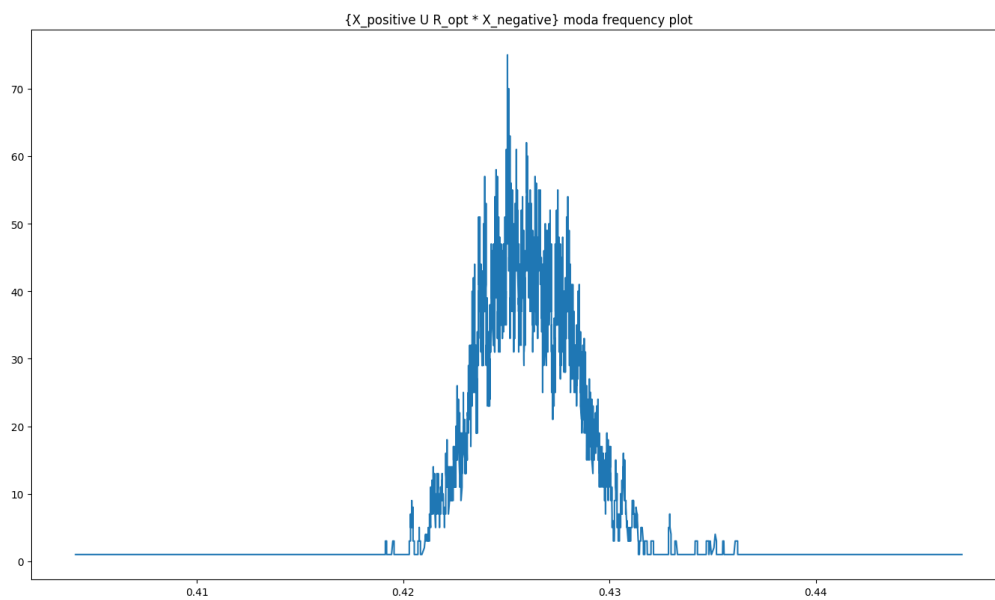


Рис. 9: График частоты моды