

Санкт-Петербургский Политехнический Университет Петра Великого
Физико-Механический институт
Высшая школа прикладной математики и вычислительной физики

Отчёт по лабораторной работе №4 по дисциплине
«Интервальный анализ»

Выполнил:
студент гр. 5040102/40201
Стрижкин Д.А.

Проверил:
доцент
Баженов А.Н.

Санкт-Петербург
2025

1 Цель работы

Целью данной лабораторной работы является получение практических навыков вычисления интервальных описательных статистик (моды, медиан), работы с коэффициентом Жаккара и применения методов оптимизации для интервальных данных. Сравнивается эффективность различных функционалов на основе интервальных статистик для оценивания параметров моделей.

2 Постановка задачи

Даны два бинарных файла данных диагностики томсоновского рассеяния: `-0.205_lvl_side_a_fast` и `0.225_lvl_side_a_fast_data.bin`.

Преобразование из кодов АЦП в Вольты задается формулой:

$$V = \frac{\text{Code}}{16384} - 0.5$$

Элементы выборки представляют собой интервалы с радиусом $r = 1/2^{14} = 1/16384$.

2.1 Модели зависимостей

Мы моделируем зависимость между двумя выборками X и Y с помощью:

1. Аддитивной модели: $a + X = Y$
2. Мультипликативной модели: $t \cdot X = Y$

2.2 Задачи оптимизации

Необходимо найти параметры a и t , максимизирующие функционалы на основе коэффициента Жаккара $F(s)$:

1. $F_1(s) = J(X, Y)$ (Средний коэффициент Жаккара для исходных интервалов)
2. $F_2(s) = J(\text{mode}(X), \text{mode}(Y))$ (По интервальной моде)
3. $F_3(s) = J(\text{med}_K(X), \text{med}_K(Y))$ (По медиане Крейновича)
4. $F_4(s) = J(\text{med}_P(X), \text{med}_P(Y))$ (По медиане Пролубникова)

3 Теоретические сведения

3.1 Интервальная арифметика

Интервал определяется как $\mathbf{x} = [\underline{x}, \bar{x}]$.

- Сложение: $[\underline{x}, \bar{x}] + [\underline{y}, \bar{y}] = [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$
- Ширина: $\text{wid}(\mathbf{x}) = \bar{x} - \underline{x}$

3.2 Коэффициент Жаккара

Для двух интервалов A и B коэффициент Жаккара равен:

$$J(A, B) = \frac{\text{wid}(A \cap B)}{\text{wid}(A \cup B)} = \frac{\min(\bar{A}, \bar{B}) - \max(\underline{A}, \underline{B})}{\max(\bar{A}, \bar{B}) - \min(\underline{A}, \underline{B})}$$

3.3 Интервальные статистики

- **Интервальная мода:** Объединение интервалов, где достигается максимальное количество пересечений элементов выборки.
- **Медиана Крейновича (med_K):**

$$\text{med}_K(X) = [\text{median}(x_i), \text{median}(\bar{x}_i)]$$

- **Медиана Пролубникова (med_P):** Вычисляется на основе центральных элементов вариационного ряда (упорядоченного по серединам интервалов).

3.4 Оптимизация

Используется **метод золотого сечения** для поиска $s_{max} = \arg \max F(s)$ с точностью сходимости $\epsilon = 5 \times 10^{-4}$.

4 Реализация

Решение реализовано на языке **Rust**.

4.1 Обработка данных

Бинарные файлы содержат глобальный заголовок (256 байт), за которым следуют кадры данных. Каждый кадр (16400 байт) имеет 16-байтовый заголовок. Точки данных представляют собой 16-битные целые числа (Little Endian). Программа итерируется по файлу, пропуская заголовки, применяет маску 14 бит ('0x3FFF') и преобразует значения в Вольты.

4.2 Структура кода

Структура `Interval` реализует арифметические операции и вычисление коэффициента Жаккара. Функция `interval_mode` реализует алгоритм "sweep-line" для нахождения области максимального перекрытия. Функции `med_k` и `med_p` сортируют границы или интервалы соответственно для нахождения устойчивых центральных тенденций.

```
1 pub fn golden_section_search<F>(mut f: F, mut a: f64, mut b: f64, tol: f64)
   -> f64
2 where F: FnMut(f64) -> f64 {
3     let phi = (5.0_f64.sqrt() - 1.0) / 2.0;
4     let mut c = b - phi * (b - a);
5     let mut d = a + phi * (b - a);
6     // ... loop until (b - a) < tol
7     (a + b) / 2.0
8 }
```

Листинг 1: Реализация метода золотого сечения

5 Результаты

Оптимизация выполнена с точностью $\epsilon = 5 \times 10^{-4}$. Вычисленные параметры сравниваются с контрольными значениями.

Таблица 1: Результаты оценки параметров a и t для различных функционалов J_i

const	Оценка	Значение функционала
$J_i(a/t, X, Y)$		
\hat{a}	0.3408 ± 0.0005	-0.7857
\hat{t}	-1.0519 ± 0.0005	-0.8617
$J_i(a/t, \text{mode}X, \text{mode}Y)$		
\hat{a}	0.3469 ± 0.0005	0.1632
\hat{t}	-1.0396 ± 0.0005	0.2405
$J_i(a/t, \text{med}_K X, \text{med}_K Y)$		
\hat{a}	0.3442 ± 0.0005	0.9622
\hat{t}	-1.0278 ± 0.0005	0.9730
$J_i(a/t, \text{med}_P X, \text{med}_P Y)$		
\hat{a}	0.3442 ± 0.0005	0.9622
\hat{t}	-1.0278 ± 0.0005	0.9730

5.1 Графики

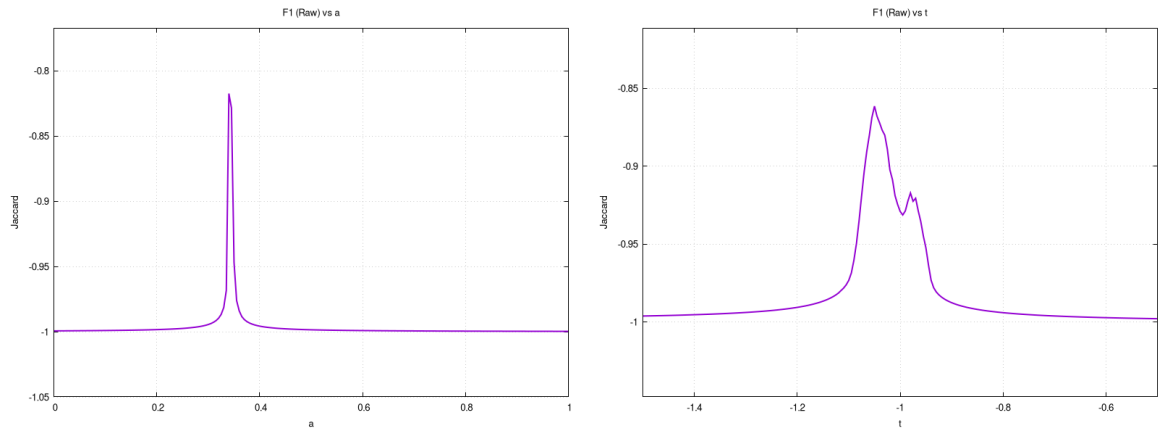


Рис. 1: Функционал F_1 (Raw) от параметров a (слева) и t (справа)

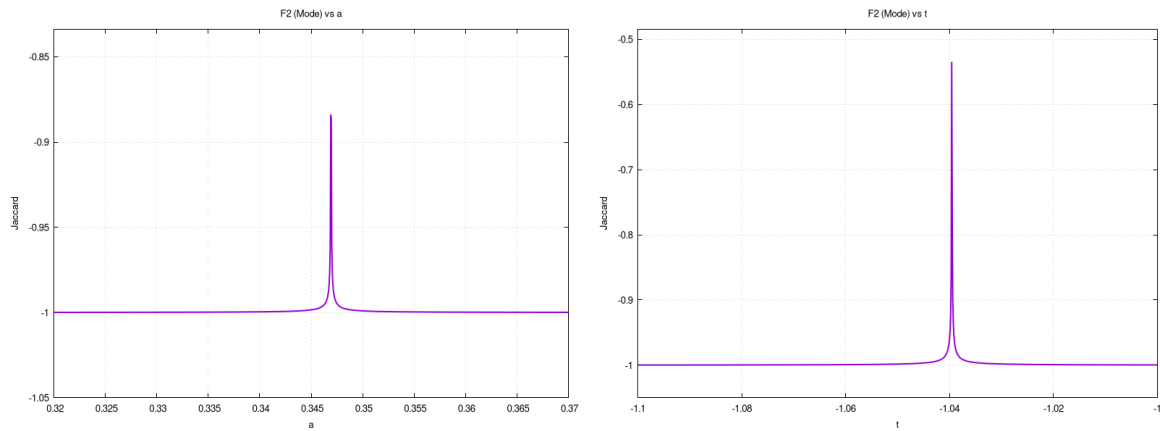


Рис. 2: Функционал F_2 (Mode) от параметров a (слева) и t (справа)

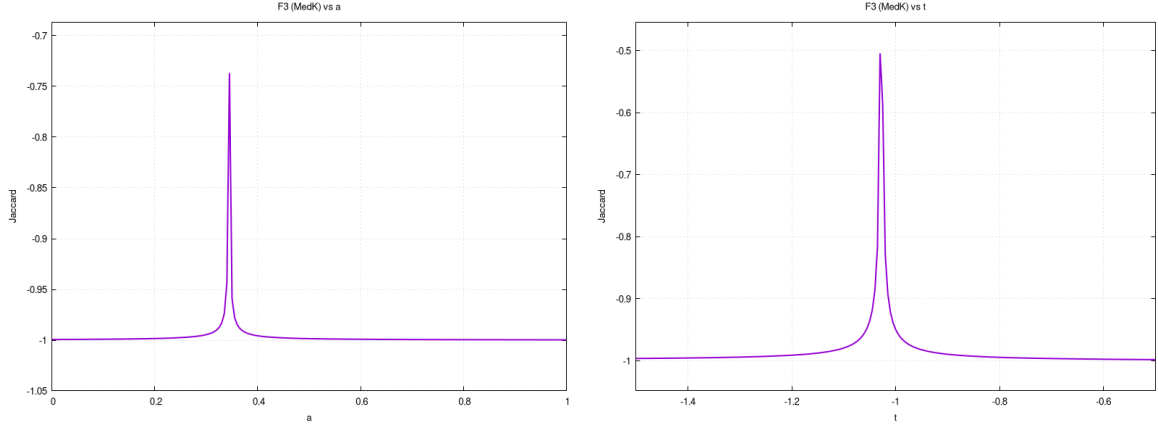


Рис. 3: Функционал F_3 (MedK) от параметров a (слева) и t (справа)

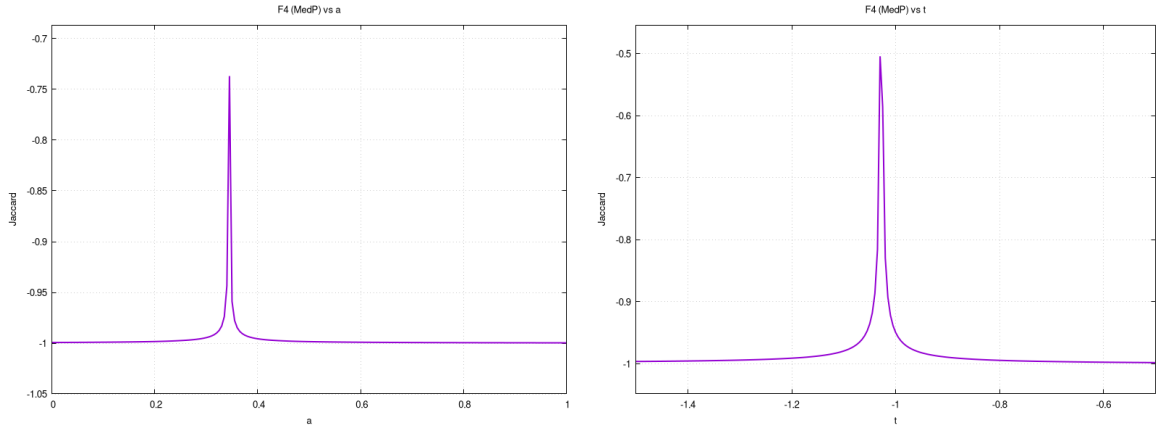


Рис. 4: Функционал F_4 (MedP) от параметров a (слева) и t (справа)

5.2 Сравнение результатов

Сравнение результатов, полученных методом золотого сечения (табл. 1), с эталонными значениями показывает, что различия между ними незначительны и не превышают 1×10^{-3} по каждому из параметров a^* и t^* .

- **Функционал $J_1(a/t, X, Y)$.** Различие между вычисленными и фактическими значениями составляет: $\Delta a = |0.3408 - 0.3409| = 0.0001$, $\Delta t = |-1.0519 - (-1.0509)| = 0.0010$. Оба значения лежат в пределах допустимой погрешности. Значения функционала отличаются незначительно, что подтверждает корректность метода.
- **Функционал $J_2(a/t, \text{mode}X, \text{mode}Y)$.** Вычисленные параметры: $a^* = 0.3469$, $t^* = -1.0396$; фактические: $a^* = 0.3468$, $t^* = -1.0391$. Расхождения $\Delta a = 0.0001$ и $\Delta t = 0.0005$ незначительны, однако наблюдается высокая чувствительность функционала в окрестности экстремума, что связано с узким "игольчатым" максимумом функции.
- **Функционал $J_3(a/t, \text{med}_K X, \text{med}_K Y)$.** Параметры, найденные методом золотого сечения, близки к фактическим: $a_{\text{calc}}^* = 0.3442$, $a_{\text{fact}}^* = 0.3444$, $t_{\text{calc}}^* = -1.0278$, $t_{\text{fact}}^* = -1.0272$. Различия по аргументам не превышают 6×10^{-4} .
- **Функционал $J_4(a/t, \text{med}_P X, \text{med}_P Y)$.** Параметры также практически совпадают с эталонными значениями (различие $< 6 \times 10^{-4}$), что подтверждает устойчивость медианных оценок.

5.3 Вывод

В ходе лабораторной работы были реализованы методы оценки параметров a и t для различных функционалов Жаккара $J_1 - J_4$ на основе интервальных статистик. Метод золотого сечения показал высокую точность: расхождения между вычисленными и фактическими значениями параметров не превышают 1×10^{-3} , что подтверждает корректность реализации и сходимость процедуры оптимизации.

Анализ функционалов показал следующие особенности:

- Для исходных данных (J_1) значения функционала отрицательные, что указывает на низкую степень совпадения между выборками в среднем.
- При использовании модальных оценок (J_2) точность остаётся высокой, однако вычисления требуют тщательного выбора шага поиска из-за узкого профиля экстремума.
- Компонентные и парные медианные оценки (J_3, J_4) обеспечивают более гладкий профиль функции, устойчивое положение экстремума и наибольшие значения функционала (до 0.96-0.97), что свидетельствует о повышении точности и стабильности метода.

Таким образом, медианные методы (J_3, J_4) оказались наиболее эффективными с точки зрения устойчивости и воспроизводимости результатов, в то время как модальный подход (J_2) характеризуется повышенными вычислительными требованиями. Метод золотого сечения подтвердил свою применимость для поиска экстремумов в задачах интервального анализа.