Санкт-Петербургский политехнический университет
Петра Великого

Физико-механический институт

Высшая школа прикладной математики и вычислительной
физики

Отчет по лабораторной работе №3
"Интервальный анализ"

Выполнили студент группы 5030102/10201: Теплов Андрей Сергеевич

Преподаватель: Баженов Александр Николаевич

Санкт-Петербург
2024

Содержание

1 Постановка задачи

Даны 2 интервальных выборки

$$\mathbf{X} = \{\mathbf{x}_i\}, \tag{1}$$
$$\mathbf{Y} = \{\mathbf{y}_i\}. \tag{2}$$

Взять $\mathbf{X}, \mathbf{Y}$ из файлов данных, задав $\text{rad}\mathbf{x} = \text{rad}\mathbf{y} = \frac{1}{2^N}$В, $N = 14$.
Файлы данных:

- *-0.205_ lvl_ side_ a_ fast_ data.bin*
- *0.225_ lvl_ side_ a_ fast_ data.bin*

Связь кодов данных и В:

$$V = N/16384 - 0.5$$

Сделать оценки констант a, t в уравнениях:

$$\mathbf{X} + a = \mathbf{Y}, \tag{3}$$
$$t\mathbf{X} = \mathbf{Y}, \tag{4}$$

Метод решения:

$$\hat{a} = \text{argmax} F(a, \mathbf{X}, \mathbf{Y}), \tag{5}$$

где F — функционал.
В качестве функционала взять варианты:

$$\text{Ji}(a, \mathbf{X}, \mathbf{Y}), \tag{6}$$
$$\text{Ji}(a, \text{mode}\mathbf{X}, \text{mode}\mathbf{Y}), \tag{7}$$
$$\text{Ji}(a, \text{med}_K\mathbf{X}, \text{med}_K\mathbf{Y}), \tag{8}$$
$$\text{Ji}(a, \text{med}_P\mathbf{X}, \text{med}_P\mathbf{Y}), \tag{9}$$

где Ji — коэффициент Жаккара, mode — интервальная мода, med_K, med_P — интервальные медианы Крейновича и Пролубникова.

Сделать точечные и интервальные оценки, задавшись уровнем α.

2 Необходимая теория

2.1 Интервальная мода

Пусть имеется интервальная выборка

$$\mathbf{X} = \{\mathbf{x}_i\}.$$

Сформируем массив интервалов $\mathbf{z}$ из концов интервалов $\mathbf{X}$.

Для каждого интервала $\mathbf{z}_i$ подсчитываем число μ_i интервалов из выборки $\mathbf{X}_i$, включающих $\mathbf{z}_i$. Максимальные $\mu_i = \max \mu$ достигаются для индексного множества K. Тогда можно найти интервальную моду как мультиинтервал

$$\text{mode}\mathbf{X} = \bigcup_{k \in K} \mathbf{z}_k. \tag{10}$$

2.2 Интервальная медиана Крейновича

Пусть дана выборка $\mathbf{X} = \{\mathbf{x}_i\}$. Пусть $\underline{c} = \{\underline{\mathbf{x}_i}\}$, $\overline{c} = \{\overline{\mathbf{x}_i}\}$ — конфигурация точек, составленные, соответственно, из левых и правых концов интервалов из $\mathbf{X}$.

Тогда медианой Крейновича $\mathrm{med}_K\mathbf{X}$ интервальной выборки $\mathbf{X}$ — это интервал

$$\mathrm{med}_K = [\mathrm{med}\underline{c}, \mathrm{med}\overline{c}]. \tag{11}$$

2.3 Интервальная медиана Пролубникова

Зададим отношение порядка на алгебре $\mathbb{IR}$. Говорят, что неравенство $\mathbf{a} \leqslant \mathbf{b}$ выполняется

1. в сильном смысле, если $\forall \mathbf{a} \in \mathbb{IR}\ \forall \mathbf{b} \in \mathbb{IR} : \overline{\mathbf{a}} \leqslant \underline{\mathbf{b}}$,
2. в слабом смысле, если $\exists \mathbf{a} \in \mathbb{IR}\ \exists \mathbf{b} \in \mathbb{IR} : \underline{\mathbf{a}} \leqslant \overline{\mathbf{b}}$,
3. в $\forall\exists$-смысле, если $\forall \mathbf{a} \in \mathbb{IR}\ \exists \mathbf{b} \in \mathbb{IR} : \overline{\mathbf{a}} \leqslant \overline{\mathbf{b}}$,
4. в $\exists\forall$-смысле, если $\exists \mathbf{a} \in \mathbb{IR}\ \forall \mathbf{b} \in \mathbb{IR} : \underline{\mathbf{a}} \leqslant \underline{\mathbf{b}}$,
5. в центральном смысле, если $(\overline{\mathbf{a}} + \underline{\mathbf{a}})/2 \leqslant (\overline{\mathbf{b}} + \underline{\mathbf{b}})/2$

Для элементов выборки $\mathbf{X}$ можно определить линейный порядок, используя любое из пяти вышеуказанных отношений порядка на $\mathbb{IR}$. То есть, если $i \neq j$, то либо $x_i \leqslant x_j$, либо $x_i \geqslant x_j$ для любого из этих отношений порядка.

Медиана Пролубникова $\mathrm{med}_P\mathbf{X}$ выборки $\mathbf{X}$ — это интервал $\mathbf{x}_m$, для которого половина интервалов из $\mathbf{X}$ лежит слева, а половина — справа.

В ситуации, когда имеются два элемента подинтервала $\mathbf{x}_m$ и $\mathbf{x}_{m+1}$, расположенных посередине вариационного ряда, $\mathbf{x}_m \neq \mathbf{x}_{m+1}$ медиана может быть определена естественным обобщением взятия полусуммы точечных значений, расположенных посередине ряда из точечных значений, в случае интервальной выборки взятие полусуммы интервалов $\mathbf{x}_m$ и $\mathbf{x}_{m+1}$:

$$\mathrm{med}_P\mathbf{X} = (\mathbf{x}_m + \mathbf{x}_{m+1})/2. \tag{12}$$

2.4 Коэффициент Жаккара

Коэффициент Жаккара для двух интервалов $\mathbf{x} \in \mathbb{IR}$ и $\mathbf{y} \in \mathbb{IR}$:

$$\mathrm{Ji}(\mathbf{x}, \mathbf{y}) = \frac{\mathrm{wid}(x \wedge y)}{\mathrm{wid}(x \vee y)} = \frac{\min\{\overline{\mathbf{x}}, \overline{\mathbf{y}}\} - \max\{\underline{\mathbf{x}}, \underline{\mathbf{y}}\}}{\max\{\overline{\mathbf{x}}, \overline{\mathbf{y}}\} - \min\{\underline{\mathbf{x}}, \underline{\mathbf{y}}\}}. \tag{13}$$

Коэффициент Жаккара для множества интервалов $\mathbf{X} \in \mathbb{IR}^n$:

$$\mathrm{Ji}(\mathbf{X}) = \frac{\min \overline{\mathbf{x}_i} - \max \underline{\mathbf{x}_i}}{\max \overline{\mathbf{x}_i} - \min \underline{\mathbf{x}_i}}. \tag{14}$$

Коэффициент Жаккара для двух множеств интервалов $\mathbf{X} \in \mathbb{IR}^n$ и $\mathbf{Y} \in \mathbb{IR}^n$:

$$\mathrm{Ji}_k(\mathbf{X}, \mathbf{Y}) = \frac{\min\{\overline{\mathbf{x}_k}, \overline{\mathbf{y}_k}\} - \max\{\underline{\mathbf{x}_k}, \underline{\mathbf{y}_k}\}}{\max\{\overline{\mathbf{x}_k}, \overline{\mathbf{y}_k}\} - \min\{\underline{\mathbf{x}_k}, \underline{\mathbf{y}_k}\}},\ k \in 1, 2, \dots, |\mathbf{X}|. \tag{15}$$

3 Реализаця

Лабораторная работа выполнена на языке программирования Python. В ходе работы были также использованы библиотеки `numpy` и `matplotlib`.

Ссылка на GitHub репозиторий: https://github.com/vladimir-skvortsov/spbstu-interval-anylysis

3.1 Поиск параметров, при которых функционал достигал наибольших значений

Для поиска параметров, при которых функционал достигал наибольших значений, был использован алгоритм троичного поиска с заданной точностью $\varepsilon = 10^{-3}$ на участках, где функции вели себя как унимодальные.

4 Результаты

Для функцинала 6:

$$\hat{a} = 0.34278 \pm 0.0005,\ F_1(\hat{a}) = -0.94918,$$
$$\hat{t} = -1.01467 \pm 0.0005,\ F_1(\hat{t}) = -0.92734.$$

Для функцинала 7:

$$\hat{a} = 0.34023 \pm 0.0005,\ F_2(\hat{a}) = -0.25437,$$
$$\hat{t} = -0.99871 \pm 0.0005,\ F_2(\hat{t}) = -0.92750.$$

Для функцинала 8:

$$\hat{a} = 0.34415 \pm 0.0005,\ F_3(\hat{a}) = -0.00184,$$
$$\hat{t} = -1.00607 \pm 0.0005,\ F_3(\hat{t}) = 0.63020.$$

Для функцинала 9:

$$\hat{a} = 0.34366 \pm 0.0005,\ F_4(\hat{a}) = -0.12457,$$
$$\hat{t} = -1.00607 \pm 0.0005,\ F_4(\hat{t}) = 0.63021.$$

5 Выводы

В ходе выполнения лабораторной работы были изучены методы оценки параметров в уравнениях с интервальными данными. Используя различные функционалы, такие как коэффициент Жаккара, были найдены оптимальные значения параметров $\hat{a}$ и $\hat{t}$ для уравнений $\mathbf{X}+a = \mathbf{Y}$ и $t\mathbf{X} = \mathbf{Y}$.

Результаты показали, что:

1. Значения параметров $\hat{a}$ и $\hat{t}$ варьируются в зависимости от выбранного функционала. Это демонстрирует важность выбора подходящего критерия оптимальности для конкретной задачи интервального анализа.

2. Наиболее стабильные результаты были получены для функционала 8, где значение $\hat{t}$ показало положительное значение коэффициента Жаккара, что указывает на высокий уровень совпадения интервалов.
3. Выбор интервальной моды и медиан (Крейновича и Пролубникова) как статистических характеристик позволил получить более точные оценки параметров, что подчеркивает их значимость в анализе интервальных данных.

Таким образом, проведенная работа продемонстрировала применимость и эффективность интервального анализа в задачах оценки параметров, а также подчеркнула важность выбора подходящих методов и инструментов для анализа данных.