

Санкт-Петербургский политехнический университет
Петра Великого

Физико-механический институт

Кафедра «Прикладная математика»

**Отчёт по лабораторной работе №2
по дисциплине «Анализ данных с интервальной
неопределённостью»**

Выполнил студент:
Мишутин Дмитрий Валерьевич
группа: 5040102/20201

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

Санкт-Петербург
2024 г.

Содержание

| | | |
|---|-------------------|---|
| 1 | Постановка задачи | 2 |
| 2 | Теория | 2 |
| 3 | Реализация | 2 |
| 4 | Результаты | 3 |
| 5 | Выводы | 7 |
| 6 | Литература | 7 |
| 7 | Приложения | 8 |

Список иллюстраций

| | | |
|---|--|---|
| 1 | Исходная интервальная выборка $X, (Y_1)$ | 3 |
| 2 | Точечная линейная регрессия выборки $X, (Y_1)$ | 4 |
| 3 | Информационное множество выборки $X, (Y_1)$ | 4 |
| 4 | Коридор совместных зависимостей выборки $X, (Y_1)$ | 5 |
| 5 | Исходная интервальная выборка $X, (Y_2)$ | 5 |
| 6 | Точечная линейная регрессия выборки $X, (Y_2)$ | 6 |
| 7 | Информационное множество выборки $X, (Y_2)$ | 6 |
| 8 | Коридор совместных зависимостей выборки $X, (Y_2)$ | 7 |

1 Постановка задачи

Имеется выборка $(X, (Y))$. X – множество вещественных чисел, Y – множество интервалов. Необходимо восстановить функциональную зависимость.

2 Теория

Для выборки $(X, (Y))$, $X = \{x_i\}_{i=1}^n$, $Y = \{y_i\}_{i=1}^n$ (x_i – точный, y_i – интервальный) линейная регрессионная модель имеет вид:

$$y = \beta_0 + \beta_1 x \quad (1)$$

Для оценки параметров необходимо решить систему вида:

$$\begin{aligned} \underline{y}_i \leq y = \beta_0 + \beta_1 x_i \leq \overline{y}_i \\ i = 1..n \end{aligned} \quad (2)$$

С учетом применения метода вариации неопределенности имеем задачу минимизации:

$$\begin{aligned} \sum_{i=1}^n w_i \rightarrow \min \\ \text{mid}y_i - w_i \text{rad}y_i \leq \beta_0 + \beta_1 x_i \leq \text{mid}y_i + w_i \text{rad}y_i \\ w_i \geq 0, i = 1..n \end{aligned} \quad (3)$$

Информационным множеством называется множество всех значений параметров β_0, β_1 , удовлетворяющих 1. Минимальные и максимальные значения параметров в информационном множестве определяют внешнюю оценку параметров модели.

Коридором совместных зависимостей называется множество всех модельных функций совместных с исходными данными.

3 Реализация

Из языка Python 3.12.2 были использованы следующие модули:

- “numpy” — генерация множества чисел;
- “matplotlib.pyplot” — построение и отображение графиков;
- “scipy” — для выполнения научных и инженерных расчётов.

4 Результаты

Данные были взяты из файлов `-0_25V/-0_25V_13.txt`, `-0_5V/-0_5V_13.txt`, `+0_25V/+0_25V_13.txt` и `+0_5V/+0_5V_13.txt`. С коррекцией при помощи вспомогательных данных из файла `ZeroLine/ZeroLine_13.txt`. Набор значений $X = [-0.5, -0.25, 0.25, 0.5]$. Набор значений Y_1 определяется как интервальная мода данных из соответствующих файлов (исходные данные обинтерваливаются с $eps = 2^{-5}$). Набор значений Y_2 определяется как обинтерваленное среднее из соответствующих файлов ($eps = 2^{-5}$).

Начнем с Y_1 . Итоговая выборка:

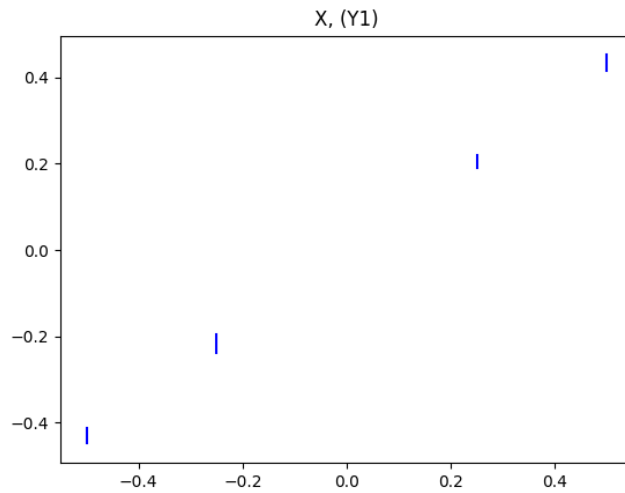


Рис. 1: Исходная интервальная выборка $X, (Y_1)$

Точечная линейная регрессия имеет вид:

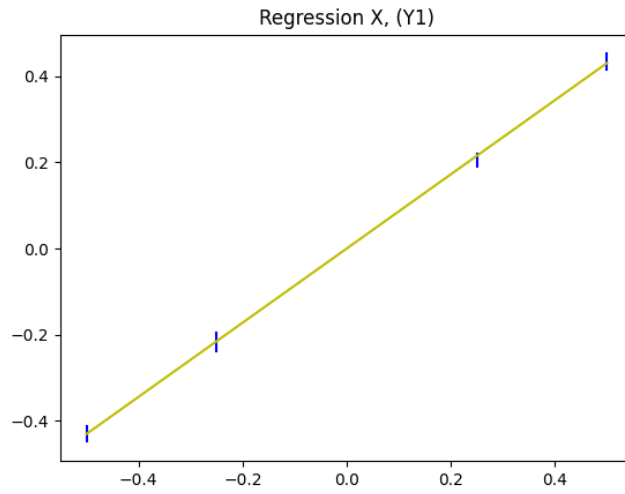


Рис. 2: Точечная линейная регрессия выборки $X, (Y_1)$

Точечные оценки параметров: $\beta_0 = 0.0, \beta_1 = 0.86236$.
 Построим информационное множество:

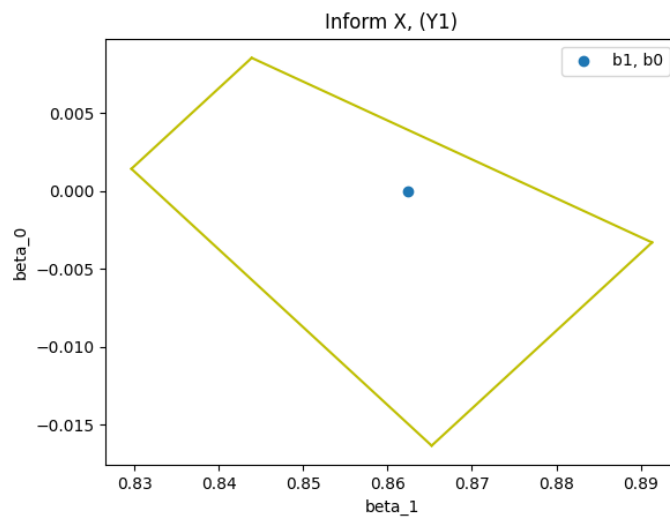


Рис. 3: Информационное множество выборки $X, (Y_1)$

Интервальные оценки параметров: $\beta_0 = [-0.01635, 0.00856], \beta_1 = [0.82965, 0.89135]$.

Коридор совместных зависимостей:

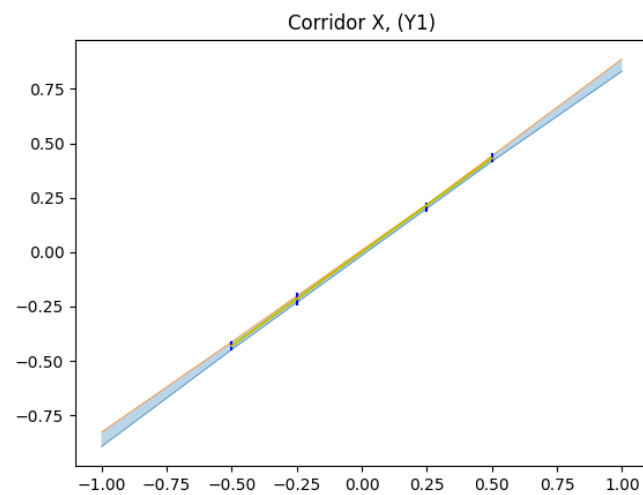


Рис. 4: Коридор совместных зависимостей выборки $X, (Y_1)$

Теперь Y_2 . Итоговая выборка:

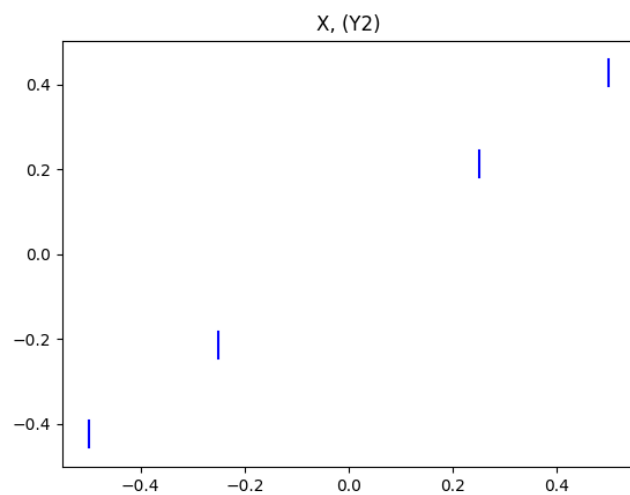


Рис. 5: Исходная интервальная выборка $X, (Y_2)$

Точечная линейная регрессия имеет вид:

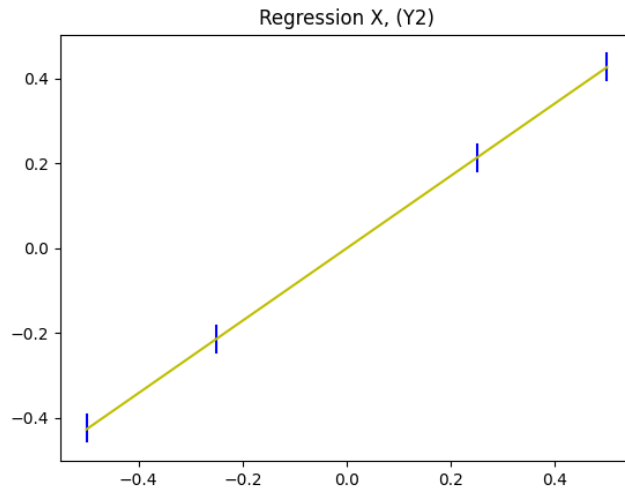


Рис. 6: Точечная линейная регрессия выборки $X, (Y_2)$

Точечные оценки параметров: $\beta_0 = 0.0003, \beta_1 = 0.85377$.
Построим информационное множество:

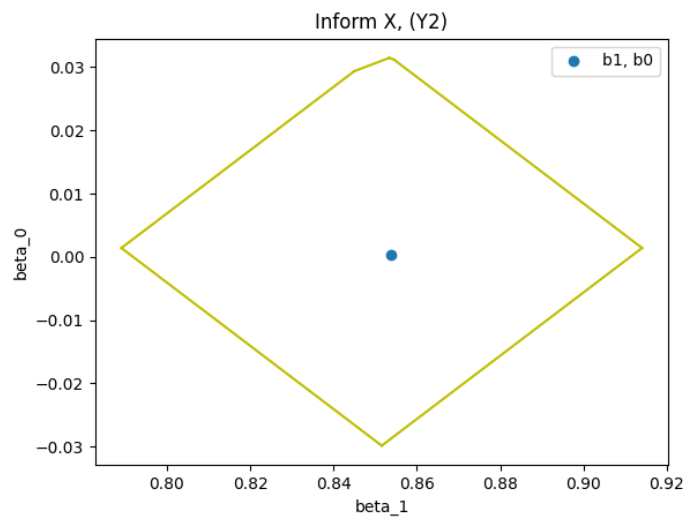


Рис. 7: Информационное множество выборки $X, (Y_2)$

Интервальные оценки параметров: $\beta_0 = [-0.02984, 0.03146], \beta_1 = [0.78907, 0.91407]$.

Коридор совместных зависимостей:

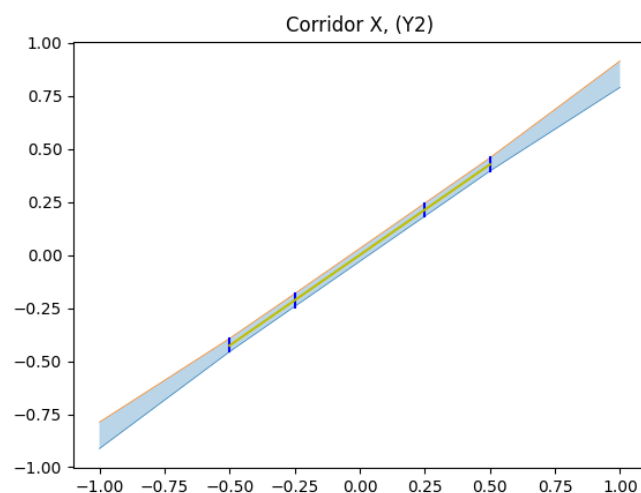


Рис. 8: Коридор совместных зависимостей выборки $X, (Y_2)$

5 Выводы

Из полученных результатов можно заметить, что оценки выборки $X, (Y_1)$ имеют примерно вдвое меньшую неопределенность, чем выборки $X, (Y_2)$. Для обеих выборок точечные оценки параметров модели лежат внутри информационного множества, и как следствие, линия регрессии лежит внутри коридора совместных зависимостей.

6 Литература

- [Баженов А.Н. «Интервальный анализ. Основы теории и учебные примеры: учебное пособие»;](#)
- [Баженов А.Н. «Естественнонаучные и технические применения интервального анализа: учебное пособие»;](#)
- [Баженов А.Н. Репозиторий “Students” на GitHub;](#)

7 Приложения

Исходники лабораторной работы выложены на [GitHub](#).