

Санкт-Петербургский политехнический университет  
Петра Великого

Физико-механический институт

Кафедра «Прикладная математика»

**Отчёт по лабораторной работе №3  
по дисциплине «Анализ данных с интервальной  
неопределённостью»**

Выполнил студент:  
Куксенко Кирилл Сергеевич  
группа: 5040102/20201

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

Санкт-Петербург  
2023 г.

## Содержание

<b>1</b>	<b>Постановка задачи</b>	<b>2</b>
<b>2</b>	<b>Теория</b>	<b>2</b>
2.1	Классификация измерений . . . . .	2
2.2	Взаимные отношения интервалов наблюдения и прогноз- ного интервала модели . . . . .	2
<b>3</b>	<b>Реализация</b>	<b>3</b>
<b>4</b>	<b>Результаты</b>	<b>3</b>
<b>5</b>	<b>Обсуждение</b>	<b>9</b>

## Список иллюстраций

1	Первая выборка, $Y_1$ . . . . .	3
2	Точечная линейная регрессия и коридор совместных зна- чений для $Y_1$ . . . . .	4
3	Точечная линейная регрессия и коридор совместных зна- чений для $\mathcal{E}_1$ . . . . .	4
4	Диаграмма статусов измерений выборки $\mathcal{E}_1$ . . . . .	5
5	Диаграмма статусов измерений выборки $\mathcal{E}_1$ (Приближеие) .	6
6	Вторая выборка, $Y_2$ . . . . .	6
7	Точечная линейная регрессия и коридор совместных зна- чений для $Y_2$ . . . . .	7
8	Точечная линейная регрессия и коридор совместных зна- чений для $\mathcal{E}_2$ . . . . .	8
9	Диаграмма статусов измерений выборки $\mathcal{E}_2$ . . . . .	8
10	Диаграмма статусов измерений выборки $\mathcal{E}_2$ (Приближение)	9

# 1 Постановка задачи

Провести анализ остатков интервальных измерений.

## 2 Теория

### 2.1 Классификация измерений

Измерения можно классифицировать следующим образом. Измерения, добавление которых к выборке не приводит к модификации модели, называются *внутренними*. Те, которые изменяют модель, называются *внешними*. Измерения, которые определяют какую-либо границу информационного множества, называются *граничными*. *Выбросами* называются те измерения, которые делают информационное множество пустым. Граничные измерения - подмножество внутренних, выбросы - внешних.

Для удобства анализа взаимоотношения информационных множеств работу с ними заменяют на анализ взаимоотношения интересующего интервального измерения и интервального прогнозируемого значения модели (коридора совместных значений).

### 2.2 Взаимные отношения интервалов наблюдения и прогнозного интервала модели

Существует несколько характеристик, определяющих это взаимоотношение.

*Размахом* называется следующее отношение 1.

$$l(x, \mathbf{y}) = \frac{\Upsilon(x)}{rad(\mathbf{y})} \quad (1)$$

*Относительным остатком* называется отношение 2.

$$r(x, \mathbf{y}) = \frac{mid(\mathbf{y}) - mid(\Upsilon(x))}{rad(\mathbf{y})} \quad (2)$$

здесь  $x$  - точечное значение,  $\mathbf{y}$  - интервальное значение интересующей величины (отклик  $x$ ),  $\Upsilon(x)$  - интервальная оценка интересующей величины (значение коридора совместных значений).

Для внутренних наблюдений выполняется неравенство 3.

$$|r(x, \mathbf{y})| \leq 1 - l(x, (y)) \quad (3)$$

В случае равенства 3 измерение будет граничным.

Выбросы определяются неравенством 4

$$|r(x, \mathbf{y})| > 1 + l(x, \mathbf{y}) \quad (4)$$

### 3 Реализация

Весь код написан на языке Python (версии 3.7.3). [Ссылка на GitHub с исходным кодом](#).

### 4 Результаты

Данные  $S_X$  были взяты из файлов *data/dataset2/XV\_spN.txt*, где  $X \in P = \{-0.45, -0.35, -0.25, -0.15, -0.05, 0.05, 0.15, 0.25, 0.35, 0.45\}$ . Набор  $\delta_i$  получен из файла *data/dataset2/0.0V\_sp443.txt*.

Рассмотрим первую выборку  $Y_1$ .  $Y_1$  получена следующим образом.  $\mathbf{y}_i = [\min S_i, \max S_i]$ ,  $i \in P$ ,  $\mathbf{y}_i \in Y_1$ .

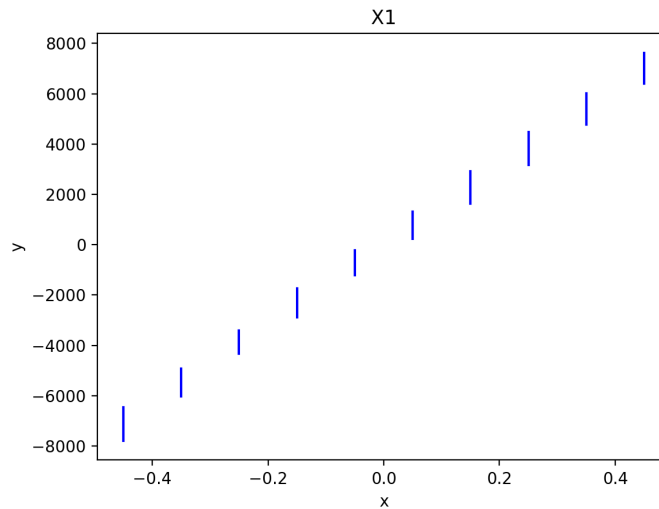


Рис. 1: Первая выборка,  $Y_1$

Построим точечную линейную регрессию и коридор совместных значений.

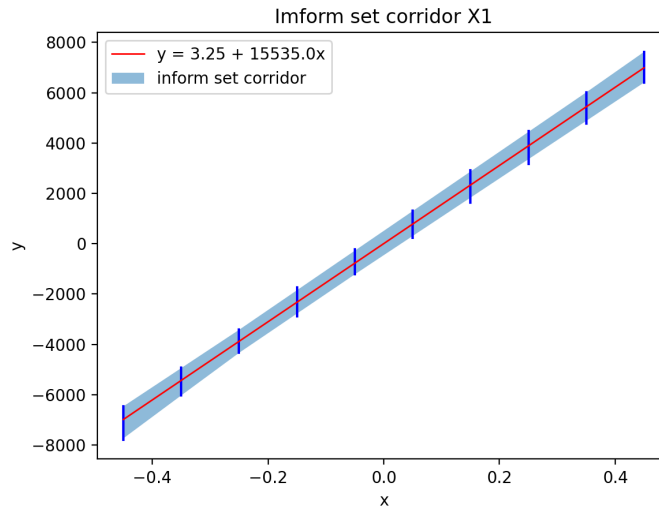


Рис. 2: Точечная линейная регрессия и коридор совместных значений для  $Y_1$

Построим выборку остатков  $\mathcal{E}_1$ ,  $\varepsilon_i = \mathbf{y}_i - (\beta_0 + \beta_1 x_i)$ ,  $\varepsilon_i \in \mathcal{E}_1$ .  
 Выборка  $\mathcal{E}_1$  и Коридор совместных значений для  $\mathcal{E}_1$  имеют вид.

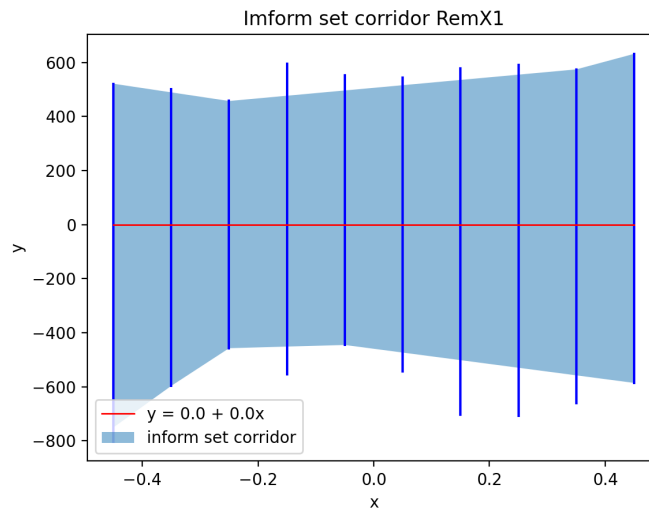


Рис. 3: Точечная линейная регрессия и коридор совместных значений для  $\mathcal{E}_1$

Теперь построим диаграмму статусов для выборки  $\mathcal{E}_1$ .

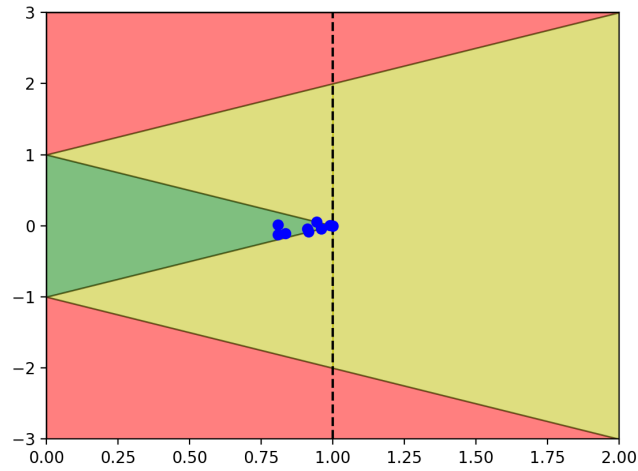


Рис. 4: Диаграмма статусов измерений выборки  $\mathcal{E}_1$

Для данной выборки  $\mathcal{E}_1$  и простейшей линейной модели граничными являются измерения, соответствующие следующим значениям переменной  $x$ :  $[-0.45, -0.35, -0.25, -0.05, 0.35]$ . Измерение, соответствующее переменной  $x = 0.45$ , возможно, является внешним или также граничным, а все остальные измерения внутренние (рис. 5).

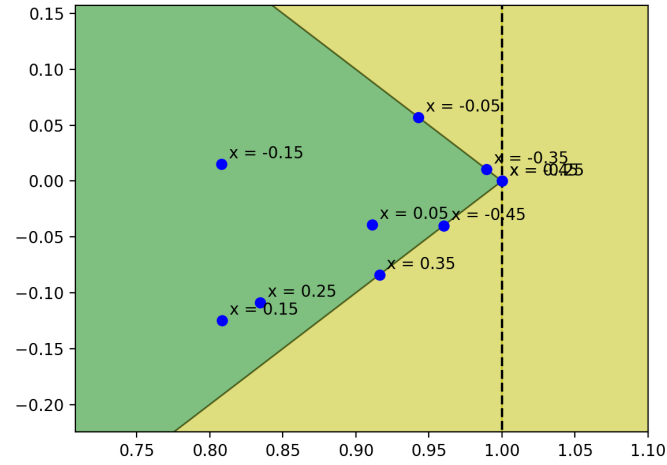


Рис. 5: Диаграмма статусов измерений выборки  $\mathcal{E}_1$  (Приближение)

Для наглядности проведём аналогичные измерения для другой выборки  $Y_2$ .  $Y_2$  получена следующим образом.  $\mathbf{y}_i = [\text{median}(S_i) - \varepsilon, \text{median}(S_i) + \varepsilon]$ ,  $\varepsilon = 25.0$ ,  $i \in P$ ,  $\mathbf{y}_i \in Y_2$ .

$Y_2$  имеет вид.

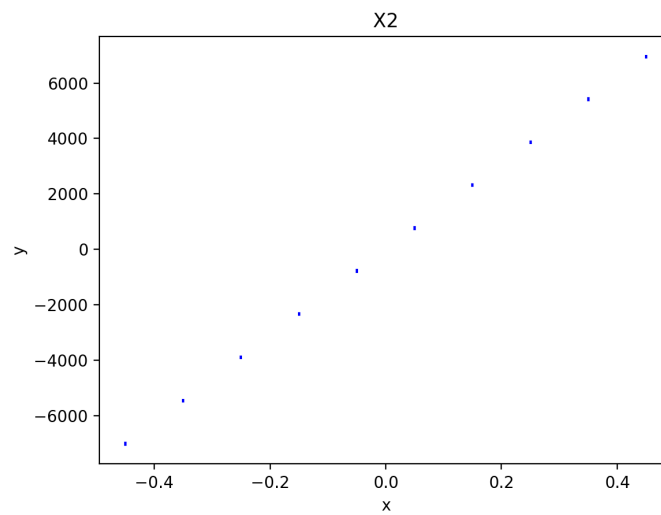


Рис. 6: Вторая выборка,  $Y_2$

Построим точечную линейную регрессию и коридор совместных значений для  $Y_2$ .

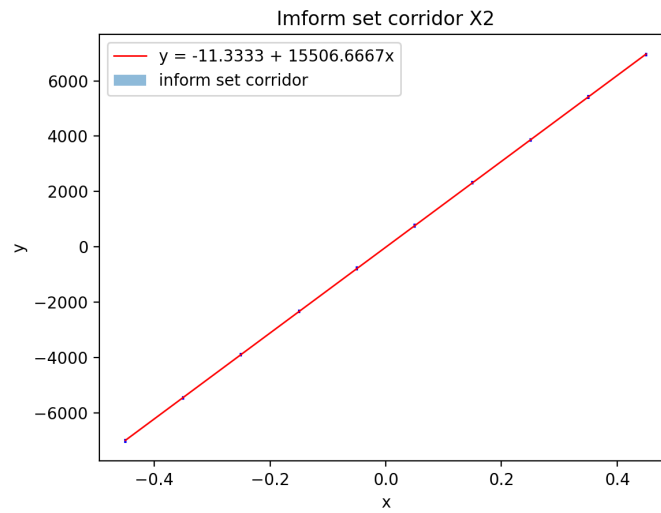


Рис. 7: Точечная линейная регрессия и коридор совместных значений для  $Y_2$

Выборка остатков  $\mathcal{E}_2$  и коридор совместных значений для  $\mathcal{E}_2$  имеют вид.



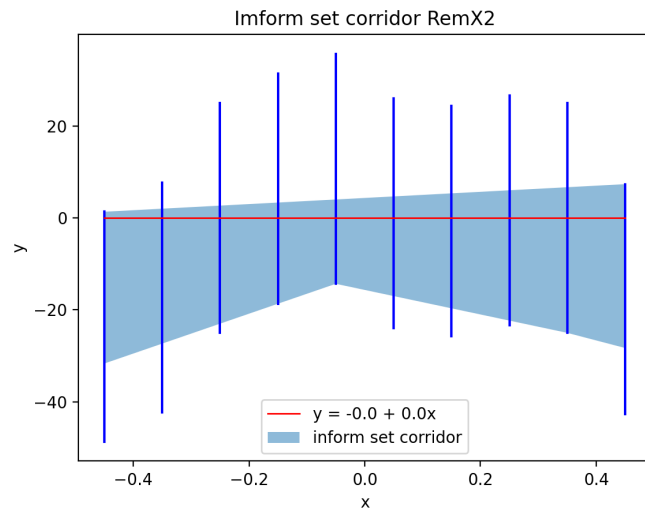


Рис. 8: Точечная линейная регрессия и коридор совместных значений для  $\mathcal{E}_2$

Построим диаграмму статусов для  $\mathcal{E}_2$ .

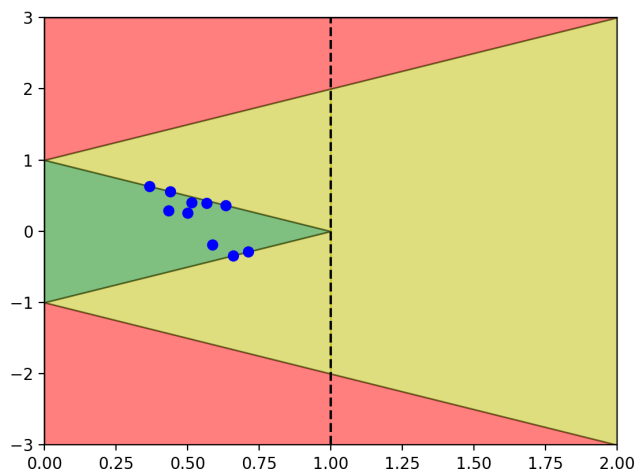


Рис. 9: Диаграмма статусов измерений выборки  $\mathcal{E}_2$

Для выборки  $\mathcal{E}_2$  граничными являются измерения, соответствующие

значениям переменной  $x \in [-0.45, -0.15, -0.05, 0.35, 0.45]$ . Остальные являются внутренними.

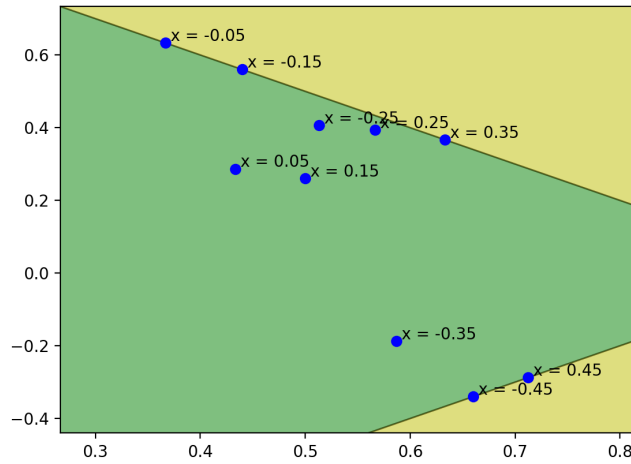


Рис. 10: Диаграмма статусов измерений выборки  $\mathcal{E}_2$  (Приближение)

## 5 Обсуждение

Из полученных результатов можно заметить следующее. Для первой выборки на диаграмме статусов измерений статусы находятся вблизи точки  $(1, 0)$  (рис. 4), в то время как для второй выборки статусы расположились дальше от точки  $(1, 0)$  и имеют большие значения для плеча 1. Это вполне сочетается с тем, как выглядят коридоры совместных значений для каждой выборки (рис. 3, 8). Также стоит отметить, что ни для одной выборки не было обнаружено выбросов или явных внешних измерений.