

Санкт-Петербургский политехнический университет Петра Великого
Институт прикладной математики и механики
Высшая школа прикладной математики и вычислительной физики

ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №8

по дисциплине
«Математическая статистика»

Выполнил студент
группы 3630102/80401

Веденичев Дмитрий Александрович

Проверил
Доцент, к.ф.-м.н.

Баженов Александр Николаевич

Санкт-Петербург
2021

СОДЕРЖАНИЕ

СПИСОК ИЛЛЮСТРАЦИЙ	3
1 Постановка задачи	4
2 Теория	4
3 Ход работы	4
4 Программная реализация	5
5 Результаты	5
6 Обсуждение	7
7 Приложение	7
Список литературы	7

СПИСОК ИЛЛЮСТРАЦИЙ

1	Изображение входного сигнала	5
2	Гистограмма сигнала	6
3	Разделение областей для данных сигнала с устранением выбросов	6

1 Постановка задачи

Провести дисперсионный анализ с применением критерия Фишера по данным регистраторов для одного сигнала. Определить области однородности сигнала, переходные области, шум/ фон. Длину сигнала взять равной 1024.

2 Теория

Необходимо вычислить следующие величины:

1. Внутригрупповая дисперсия

$$s_{IntraGroup}^2 = \frac{1}{k} \sum_{i=1}^k s_i^2 = \frac{1}{k} \sum_{i=1}^k \frac{\sum_{j=1}^n (x_{ij} - X_{cp})^2}{k - 1} \quad (1)$$

где X_{cp} – среднее для части выборки; k – количество частей выборки; n – количество элементов в рассматриваемой части выборки. Внутригрупповая дисперсия является дисперсией совокупности и рассматривается как среднее значение выборочных дисперсий.

2. Межгрупповая дисперсия

$$s_{InterGroup}^2 = k \frac{\sum_{i=1}^k (X_{i_{cp}} - X_{cp})^2}{k - 1} \quad (2)$$

где $X_{1cp}, X_{2cp}, \dots, X_{kcp}$ – среднее значение для под-выборок, X_{cp} – среднее значение этих средних значений под-выборок.

3. Значение критерия Фишера

$$F = \frac{s_{InterGroup}^2}{s_{IntraGroup}^2} \quad (3)$$

3 Ход работы

На начальном этапе необходимо извлечь сигнал из исходных данных (wave-ampl.txt). Известно, что сигнал имеет длину 1024, поэтому необходимо выбрать начальный индекс, кратный 1024.

Далее необходимо построить гистограмму, столбцы отвечают за следующие подобласти:

- фон (столбец с наименьшим значением)
- переходы (столбцы с малыми значениями)
- сигнал (второй по величине столбец после фона)

Перед определением областей однородности, необходимо устранить явные выбросы. Для этого был использован медианный фильтр (выброс = среднее арифметическое его соседей). По итогу получим сглаженный сигнал. После устранения выбросов необходимо разделить сигнал на области (сигнал, фон, переходные процессы).

Как только области получены, необходимо определить их тип. Это осуществляется с помощью применения критерия Фишера. Если значение критерия Фишера велико, это будут переходные процессы, если же значение находится вблизи 1, то эти области однородны.

4 Программная реализация

Лабораторная работа выполнена на языке Python версии 3.7 в среде разработки PyCharm. Использовались дополнительные библиотеки:

1. matplotlib
2. numpy

В приложении находится ссылка на GitHub репозиторий с исходным кодом.

5 Результаты

В работе рассматривался сигнал с индексом 15.

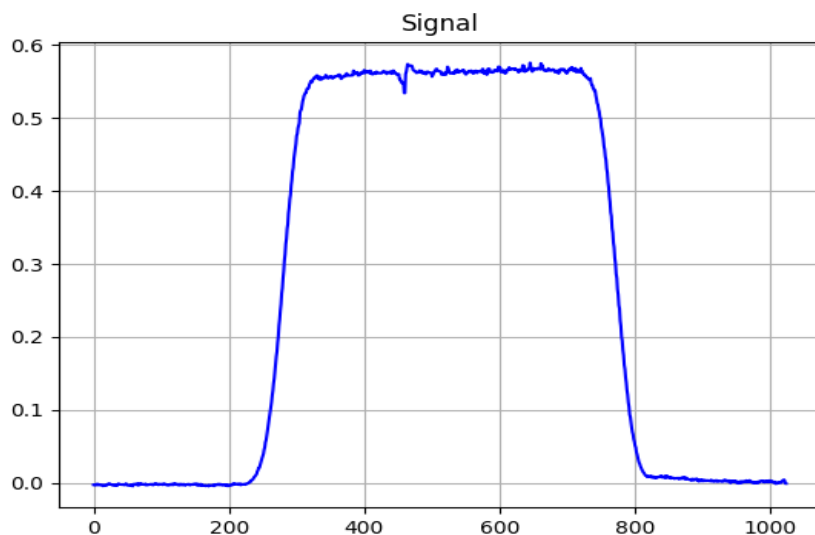


Рис. 1: Изображение входного сигнала

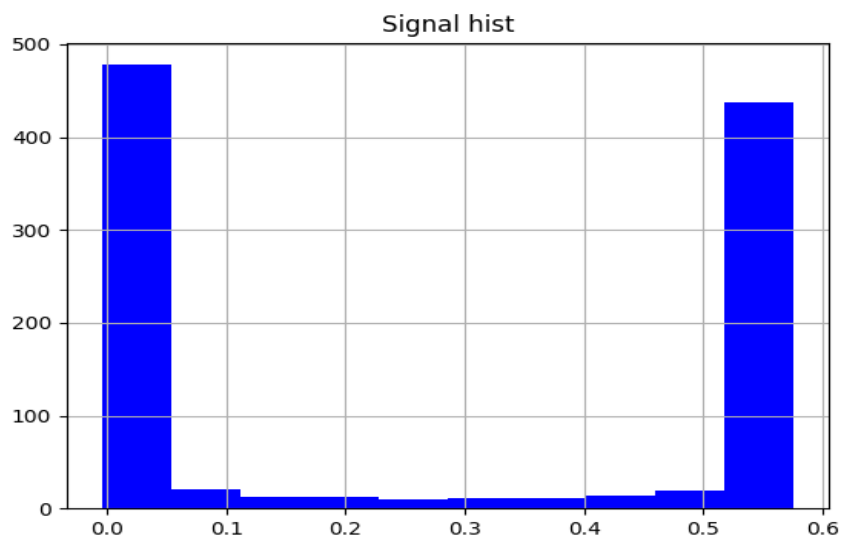


Рис. 2: Гистограмма сигнала

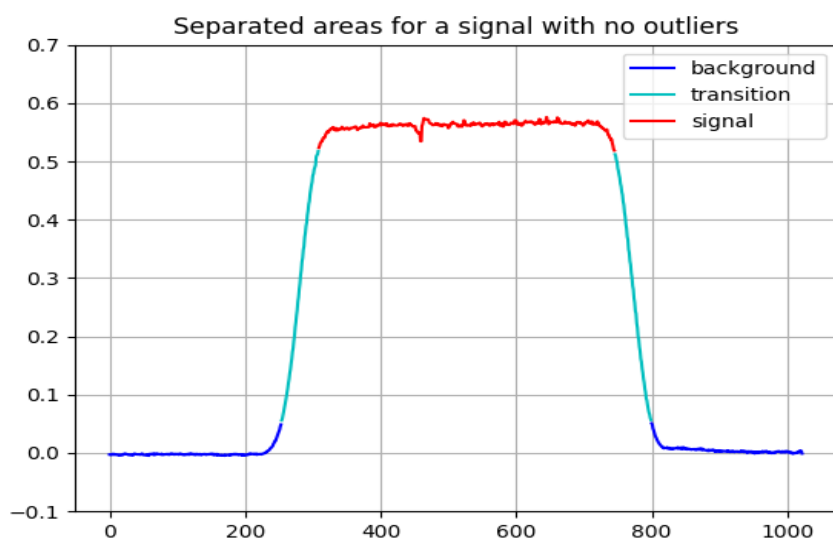


Рис. 3: Разделение областей для данных сигнала с устранением выбросов

Промежуток	Тип	Кол-во разбиений	критерий Фишера
1	background	5	0.23
2	transition	6	63.71
3	signal	19	23.61
4	transition	6	163.41
5	background	4	0.21

Таблица 1: Характеристики выделенных областей

6 Обсуждение

Для входных данных сигнала были получены следующие области однородности: background (слева и справа) и signal, эти области однородны так как значения критерия Фишера находятся вблизи 1.

На transition значения критерия Фишера много больше 1, следовательно, эти области неоднородны.

7 Приложение

Код программы GitHub URL:

https://github.com/PopeyeTheSailorsCat/math_stat_2021/tree/main/lab8

Список литературы

- [1] Гланц, С. Медико-Биологическая статистика. Пер. с англ. — М., Практика, 1998
- [2] Электронный ресурс. Виды дисперсий. <https://math.semestr.ru/group/types-variances.php>