

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

Отчет по лабораторной №2 по дисциплине «Стохастические модели и анализ данных»

Выполнил студент:
Коваленко Надежда
группа:
3640102/90201

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

Санкт-Петербург, 2020 г.

Оглавление

Постановка задачи	3
Теория	3
Критерий Фишера	3
Реализация.....	3
Практические результаты	3
Выводы.....	5
Литература	5

Постановка задачи

Есть данные, полученные регистратором. По этим данным нужно провести дисперсионный анализ для значений сигнала с применением критерия Фишера.

Теория

Критерий Фишера

Дисперсию совокупности можно оценить двумя способами [1].

Во-первых, дисперсия, вычисленная для каждой группы — это оценка дисперсии совокупности. Поэтому дисперсию совокупности можно оценить на основании групповых дисперсий. Такая оценка не будет зависеть от различий групповых средних.

Во-вторых, разброс выборочных средних тоже позволяет оценить дисперсию совокупности. Понятно, что такая оценка дисперсии зависит от различий выборочных средних. В качестве оценки дисперсии совокупности возьмем среднее выборочных дисперсий. Эта оценка называется внутригрупповой дисперсией. Обозначим ее $s_{\text{ВНУ}}^2$.

$s_{\text{ВНУ}}^2 = \frac{1}{k} \sum_{i=1}^k s_i^2 = \frac{1}{k} \sum_{i=1}^k \sqrt{\frac{\sum_{j=1}^n (x_{ij} - \bar{X})^2}{k-1}}$, где \bar{X} — среднее для части выборки, k — количество частей, на которое делим сходную выборку, n — количество элементов в подвыборке.

Также нужно вычислить межгрупповую дисперсию. Обозначается она $s_{\text{МЕЖ}}^2$. Вычисление межгрупповой дисперсии происходит в несколько этапов:

1. Вычисление среднего значения для всех выбранных подвыборок $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k)$.
2. Вычисление среднего этих средних: $\bar{\bar{X}} = \frac{1}{k} \sum_{i=1}^k \bar{X}_i$.
3. $s_{\bar{X}} = \sqrt{\frac{\sum_{i=1}^k (\bar{X}_i - \bar{\bar{X}})^2}{k-1}}$
4. $s_{\text{МЕЖ}}^2 = n * s_{\bar{X}}^2$

Тогда $F = \frac{s_{\text{МЕЖ}}^2}{s_{\text{ВНУ}}^2}$.

Реализация

Лабораторная работа выполнена на языке Python (версия 3.8) с использованием библиотек NumPy, matplotlib в среде разработки Visual Studio Code. Детали можно посмотреть на GitHub по ссылке в литературе [2].

Практические результаты

Рассмотрим все сигналы, которые представлены в файле. Их там 800. В каждом сигнале по 1024 значения.

Параметр $k = 32$, значит в каждой подвыборке по 32 элемента.

Вычислим критерий Фишера для всех сигналов и отберем только те, значение которых меньше 23. В итоге получилось следующее:

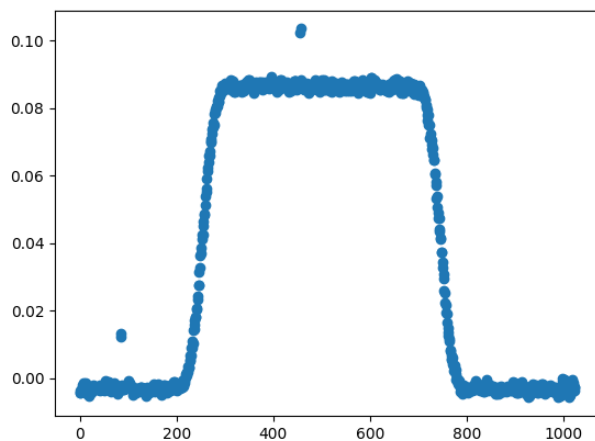


Рисунок 1.432

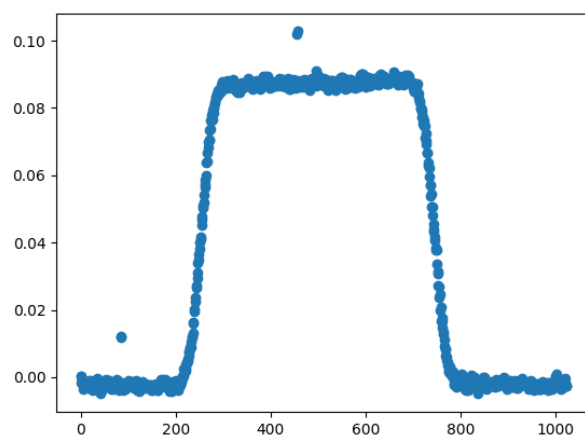


Рисунок 2.433

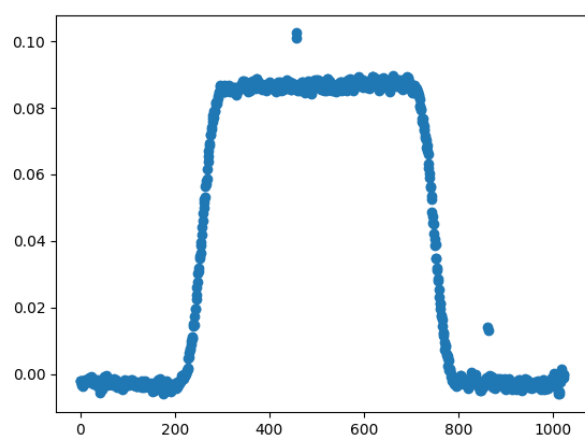


Рисунок 3.449

Номер сигнала	Значение критерия Фишера
432	22.869335
433	22.775491
449	22.502881

Выводы

Значение критерия Фишера для всех сигналов больше 1, поэтому все сигналы не являются константой.

Литература

[1] Гланц, С. Медико-биологическая статистика. Пер. с англ. — М.: Практика, 1998. — 459 с.

[2] Реализация применения критерия хи-квадрат.

URL: <https://github.com/NadezhdaKovalenko/StochasticModels>