Санкт-Петербургский Политехнический Университет Петра Великого

Физико-механический институт

Кафедра прикладной математики и вычислительной физики

Отчёт по лабораторной работе №1 по дисциплине

«Многомерный статистический анализ»

Выполнил студент гр. 5030102/90401: Веселый Д. В.

Преподаватель: Павлова Л. В.

Санкт-Петербург

2023

# Постановка задачи: не вижу задачу!

Далее – перечень, что делали. А задача?

Коротко задача – обосновать и исследовать модель распределения случайно величины

Задача: по имеющейся выборке посчитать статистические характеристики, сформировать гипотезы о виде распределения, и проверить их на основе критерия хи-квадрат Фишера, обосновать и исследовать модель распределения случайно величины.

Используя данные из файла “Number\_3.txt”:

1) Найти выборочные характеристики исследуемой с.в. :  выборочное среднее, выборочная дисперсия, выборочные коэффициенты асимметрии и эксцесса.

2) Построить э.ф.р. и нормированную гистограмму (гистограмма - красивая! без провалов и "неровностей").

3) По э.ф.р построить (в одних и тех же коор­динатных осях) доверительные полосы для теор. функции распределения  (т.ф.р.) с доверительными вероятностями  0.90 и 0.95.

4) После анализа выборочных характеристик и вида гистограммы выдвинуть гипотезу (или гипотезы) о виде распределения исследуемой с.в.

5) Проверить гипотезу (гипотезы) о виде распределения на основе критерия **хи-квадрат Фишера.**

6) После того, как принято решение о виде распределения, найти МП- оценки параметров распределения с. в.

7)  С этими оценками построить гипотетические теоретические кривые: ф.р.и плотность вероятности. **Наложить эти кривые на э.ф.р. и нормированную гистограмму, соответственно.**

8) Привести анализ полученных результатов.

# 1) Выборочные характеристики

Выборочное среднее: 3.867406990333334

Выборочная дисперсия: 6.814024540508929

Выборочный коэффициент ассиметрии : 1.4101517630007825

Коэффицеинт эксцесса: 2.504227644811291

Изображение выглядит как текст

Автоматически созданное описание

2) Эмпирическая функция распределения и гистограмма:

Гистограмма строится следующим образом:

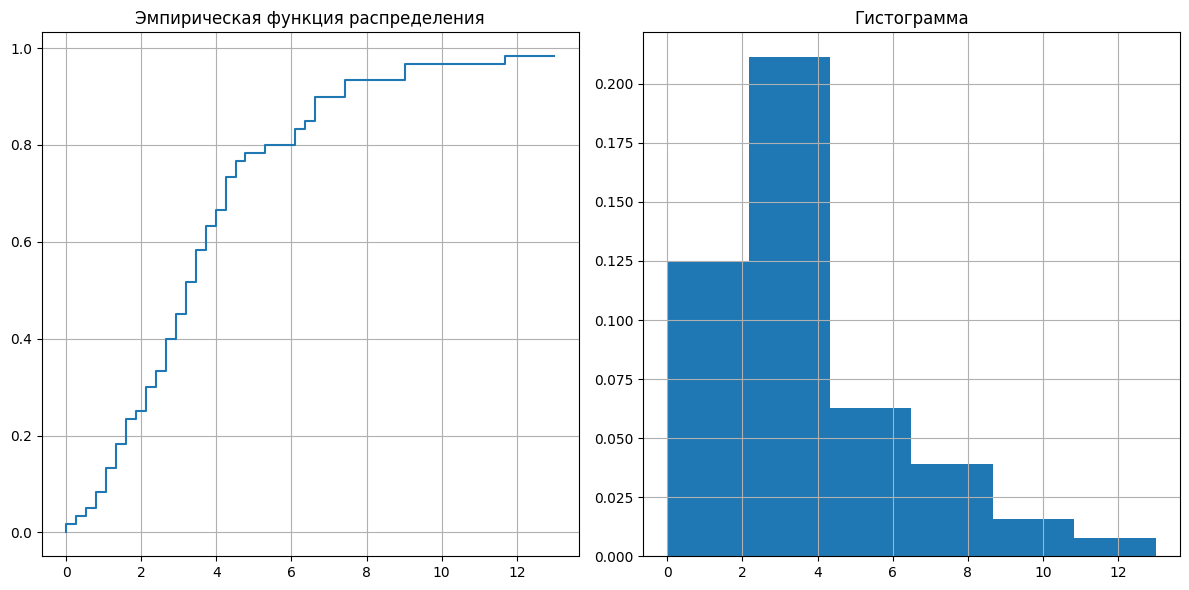
* Высота каждого столбца считается по формуле ниже, при этом сумма площадей всех интервалов равна 1:

где – правая и левая границы интервала, – высота столбца, – количество элементов выборки, попавших в интервал , – размер выборки.

Эмпирическая функция распределения в каждой точке рассчитывается как относительная частота элементов, расположенных на числовой прямой левее аргумента:

Изображение выглядит как текст

Автоматически созданное описание

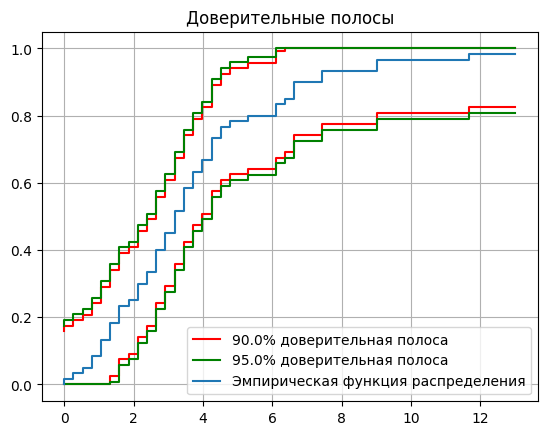


# 3) Доверительные полосы

где – доверительная вероятность, – размер выборки.

Изображение выглядит как текст

Автоматически созданное описание



# 4) Гипотеза о виде распределения

Анализируя данные, полученные выше, можно придти к гипотезе о том, что это может быть гамма-распределение (ассиметрия > 0, эксцесс показывает умеренную остроту пика, а визуально гистограмма и э.ф.р напоминают гамма распределение)

# 5) Проверка гипотезы о виде распределения

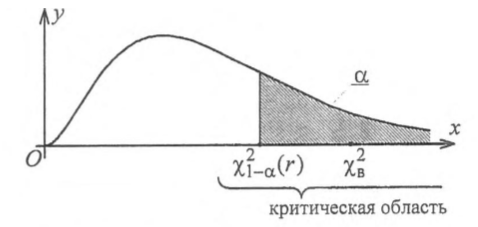
Гипотеза о виде распределения проверяется на основе критерия хи-квадрат Фишера

*- неизвестно!! Как нащли?? Это – ключевой момент работы. Находилось с помощью минимизации функции, состоящее как раз из критерия хи-квадрат, те я минимизировал значение критерия Фишера с предполагаемым распределением, аргументами минизируемой функции и были параметры . На следующем шаге(пункте) были получены оценки для параметров методом максимального правдоподобия*

*.* Все это и нужно было расписать как следует. Ок, принято.

Значения в предельном случае распределены по закону хи-квадрат с степенями свободы, где – число интервалов, r – число параметров:

Далее строится критическая область:



Если вычисленное значение меньше , то гипотеза о виде распределени принимается, в ином случае попадает в критическую область – принимаем альтернативу.

Изображение выглядит как текст

Автоматически созданное описание

Вычисленное значение 0.2803092563497795

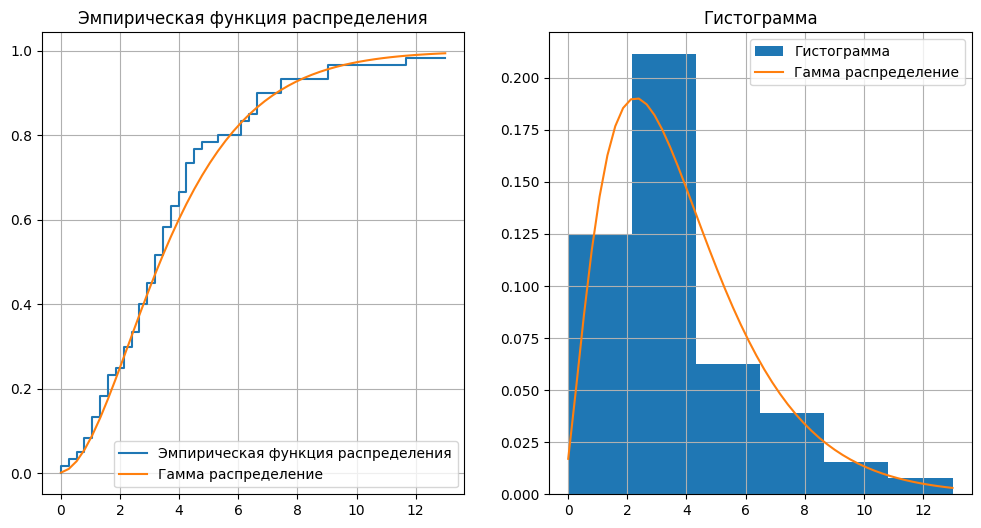
Квантиль 9.487729036781154

Видим, что полученное значение меньше, значит принимаем гипотезу

# 6) Оценка параметров и результаты

В результате исследования были получены выборочные статистики, построена эмпирическая функция распределения и гистограмма. Построены доверительные полосы э.ф.р. Была предложена гипотеза о принадлежности распределения случайной величины гамма-распределению. Далее гипотеза была принята на основании критерия хи-квадрат Фишера.

Используя метод максимального правдоподобия, были получены ( как получены? Метод?) оценки параметров :



Помимо этого, среднее у гамма-распределения = 4.06, а дисперсия = = 6.41. При этом выборочные характеристики были:

Выборочное среднее: 3.867406990333334

Выборочная дисперсия: 6.814024540508929

Как видим , значения очень близки друг к другу