МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДАНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

«Московский государственный технический

университет имени Н.Э. Баумана»

(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ \_\_\_ФН\_\_\_

КАФЕДРА  
«ВЫЧИСЛИТЕЛЬНАЯ МАТЕМАТИКА И МАТЕМАТИЧЕСКАЯ ФИЗИКА»

Направление: Математика и компьютерные науки

Дисциплина: Теория вероятности и математическая статистика

Домашняя работа №7

Группа: \_ФН11-52Б\_

Вариант №16

Студент: Хаписов М.Х.

Преподаватель: Облакова Т.В.

Москва 2022

**Задача 7**

**Критерий согласия для проверки простой непараметрической гипотезы**

**Исходные данные:**

**Основная гипотеза:**

1(A=0) Выборка получена из распределения

2.(A=1) Выборка получена из закона распределения, совпадающего с распределением ,

3. (A=2) Выборка получена из закона распределения, совпадающего с распределением ,

**Варианты значений**

**Варианты метрик для группированной выборки**

1. ()
2. ()
3. ()
4. ()
5. ()

-количество значений, попавших в -ый интервал группировки

-теоретическая вероятность попадания в -ый интервал группировки

**Задание.**

Постройте с помощью стохастического эксперимента на основе указанной метрики приближенный критерий для проверки основной гипотезы. Найдите критические значения для трех уровней значимости .

Протестируйте критерий на трех-четырех примерах и сформулируйте выводы.

**Вариант**

|  |  |  |  |
| --- | --- | --- | --- |
| 36 |  |  |  |

Основная гипотеза:

Выборка получена из закона распределения, совпадающего с распределением ,

Используемая метрика:

Размер выборки

Эксперимент повторяется раз.

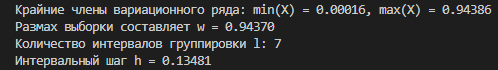
Плотность

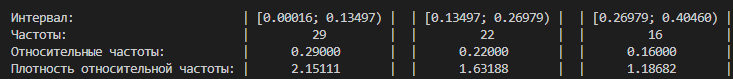
Функция распределения:

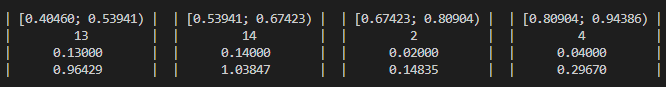
Проведём первый эксперимент и произведём группировку получившейся выборки, где количество интервалов группировки определяется по формуле Стёрджеса

Первые 10 элементов выборки:

[0.80204289 0.27132761 0.54201963 0.05825051 0.13176683 0.04122369

0.29815764 0.06618956 0.11685672 0.18990775] 





Величина в данном случае получается равна

Аналогично получим значение ещё раз, после чего составим массив, состоящий из значений параметра и отсортируем его.

Первые 10 элементов массива :

[0.008039003560420355, 0.008152589843139957, 0.00933476980192733, 0.009397743145683047, 0.00939885901711211, 0.009505671561296261, 0.00952375336102559, 0.009740987869091758, 0.009858080353757988, 0.010004169390070886]

Последние 10 элементов массива :

[0.17898776471469316, 0.17959828942215297, 0.18018467063413787, 0.18334544163141114, 0.18353978529242188, 0.1898147055794825, 0.19116154116396078, 0.19170474444201674, 0.2042743345094617, 0.20854853536034884]

Значения квантилей можно получить, отсортировав полученный массив по возрастанию и выбрав элемент по индексу , соответствующему уровню доверия





Теперь можно говорить, что если значение параметра D больше, чем эти квантили, то выборка не распределена по закону , с соответствующей этим квантилям точностью. В случае A = 2 видим, что , в следствие чего принимаем основную гипотезу.

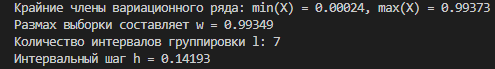
Протестируем критерий на значениях A = 0 и A = 1:

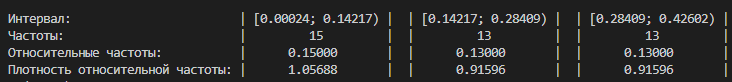
A = 0:

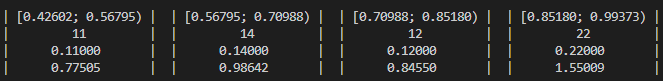
Первые 10 элементов выборки:

[0.7788702 0.40187949 0.94941642 0.04589628 0.45923576 0.38925202

0.6832384 0.55924094 0.82320583 0.54268833]







Величина в данном случае получается равна

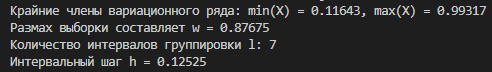
Так как для , можно говорить, что выборка не получена из распределения ,

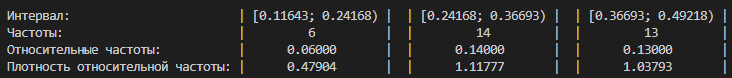
A = 2:

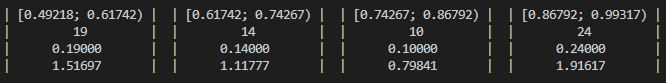
Первые 10 элементов выборки:

[0.89239922 0.31624654 0.79148515 0.55964894 0.88464897 0.47411339

0.23727819 0.5219018 0.18865649 0.87922426]







Величина в данном случае получается равна

Так как для , можно говорить, что выборка не получена из распределения ,

Вывод:

С помощью стохастического эксперимента на основе указанной метрики был построен приближенный критерий для проверки гипотезы о принадлежности выборки равномерному распределению на отрезке [0,1]. Были получены приближенные значения квантилей уровня 0.9, 0.95, 0.99, а также было проведено тестирование полученного критерия на выборках, подчиняющихся распределениям, приведенным в здании. Данный критерий показал себя достаточно точным – заключение о том, что выборка принадлежит распределению , , было вынесено только в случае выборки A, подчинявшейся данному закону. В остальных случаях было принято решение в пользу альтернативы.