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

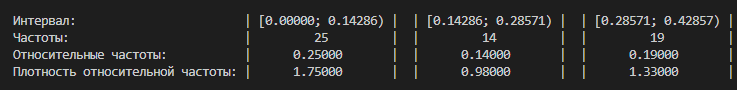
Функция распределения:

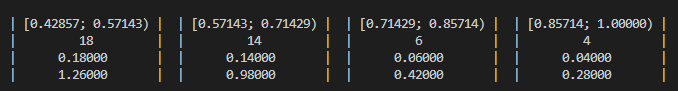
Проведём первый эксперимент и произведём группировку получившейся выборки, где количество интервалов группировки определяется по формуле Стёрджеса

Первые 10 элементов выборки:

[0.60517721 0.45670289 0.36446797 0.12496954 0.16485094 0.4655559

0.14830709 0.57785907 0.44138689 0.65023162] 





Теоретические вероятности: [0.26530612244897955, 0.22448979591836737, 0.18367346938775508, 0.1428571428571429, 0.10204081632653061, 0.061224489795918324, 0.020408163265306145]

n\*p\_j: [26.530612244897956, 22.44897959183674, 18.36734693877551, 14.28571428571429, 10.204081632653061, 6.122448979591832, 2.0408163265306145]

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 :

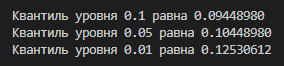
[0.0071428571428571, 0.007959183673469497, 0.008775510204081564, 0.009591836734693856, 0.009591836734693856, 0.010408163265306145, 0.010408163265306145, 0.010408163265306145, 0.011224489795918436, 0.011224489795918436]

Последние 10 элементов массива :

[0.17469387755102045, 0.17530612244897956, 0.17530612244897956, 0.17551020408163262, 0.18469387755102046, 0.18469387755102046, 0.18469387755102046, 0.18530612244897957, 0.20469387755102045, 0.20530612244897956]

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.

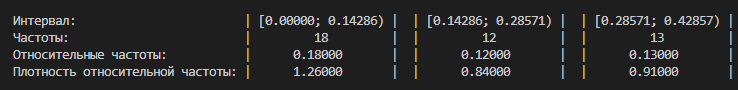
Протестируем критерий на значениях A = 0 и A = 1:

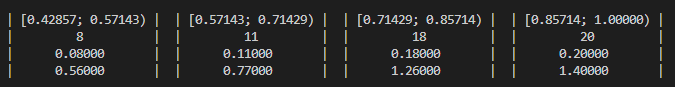
A = 0:

Первые 10 элементов выборки:

[0.89235679 0.80559679 0.7208767 0.81984404 0.62703732 0.383491

0.77002471 0.46479148 0.83465801 0.86093889] 





Величина в данном случае получается равна

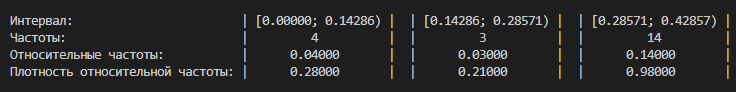
Так как для , можно говорить, что выборка не получена из распределения ,

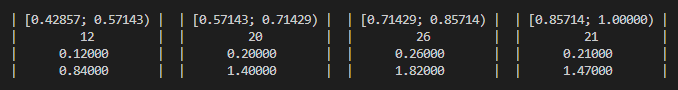
A = 1:

Первые 10 элементов выборки:

[0.71050631 0.71018169 0.39983749 0.46194587 0.73863561 0.82788274

0.81178959 0.56521997 0.84677451 0.73921791] 





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.

