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 ФУНДАМЕНТАЛЬНЫЕ НАУКИ

КАФЕДРА

«ВЫЧИСЛИТЕЛЬНАЯ МАТЕМАТИКА И МАТЕМАТИЧЕСКАЯ ФИЗИКА»

Направление: **Математика и компьютерные науки**

Дисциплина: Теория вероятностей и математическая статистика

Домашнее задание №7

«Критерий согласия для проверки простой непараметрической гипотезы»

*Группа*: ФН11-52Б

Вариант №8

Студент: Зеликова В.И.

Преподаватель: Облакова Т.В.

Оценка:

Москва 2022

**Задача 7.** Критерий согласия для проверки простой непараметрической гипотезы

**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.

**Исходные данные**:

Вариант: 28

Основная гипотеза: (A = 0) Выборка получена из распределения

Метрика для группированной выборки: ()

- количество значений, попавших в -ый интервал группировки

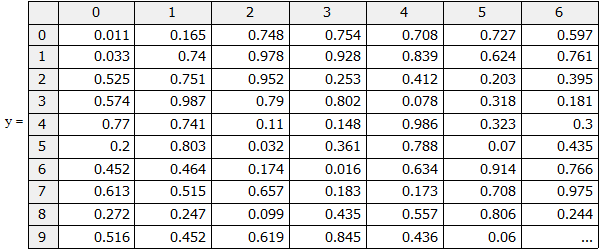
- теоретическая вероятность попадания в -ый интервал группировки

Объем выборки: n = 100

1. **Проведем стохастический эксперимент. Смоделируем выборку, подчиняющуюся основной гипотезе**

Будем проводить его m раз:

Результат моделирования - выборки размера n, распределенные равномерно на отрезке [0,1], запишем в матрицу размера n\*m:

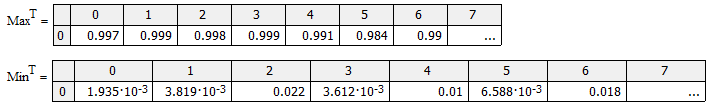


В сгенерированной таблице каждый столбец соответствует одной выборке.

1. **Произведем группировку данных**

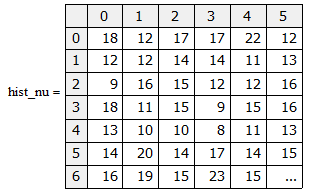
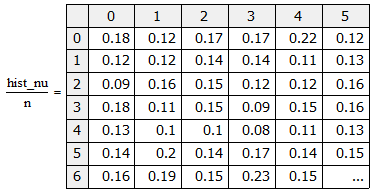
Число интервалов группировки:

Минимальные и максимальные значения:



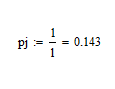
Интервалы: Гистограмма:



Эмпирические частоты для 7 интервалов: Эмпирические вероятности:

Каждый столбец соответствует одной выборке, каждая строка — это определенный интервал размера 1/7

1. **Вычислим значение статистики D для каждой выборки**.

По основной гипотезе выборка подчиняется равномерному распределению, соответственно, теоретические вероятности попадания в каждый из интервалов равны между собой:

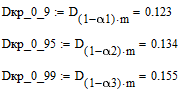
Вычислим значение статистики D для каждой выборки.

1. **Вычислим значения квантилей**

Значения квантилей можно получить, отсортировав полученный массив по возрастанию и выбрав элемент по индексу (1-α)\*m, соответствующему уровню доверия 1-α.



Приближенные значения квантилей:



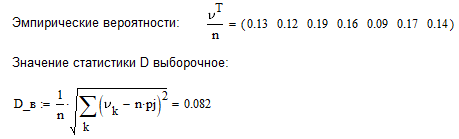
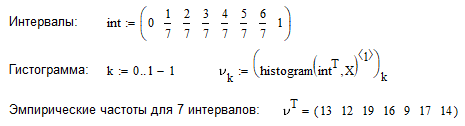
Эти значения практически не меняются при перерасчете выборки, значит m правильно выбрано достаточно большое.

Таким образом был построен критерий соответствия выборки X равномерному распределению – его формулировка (если верна основная гипотеза) имеет вид:

Протестируем его на нескольких примерах.

1. **Рассмотрим случай выборки из равномерного распределения R[0,1]**

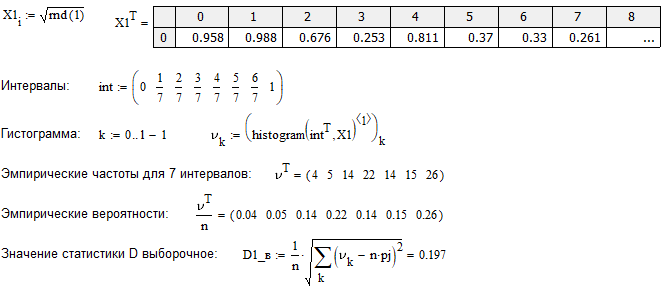




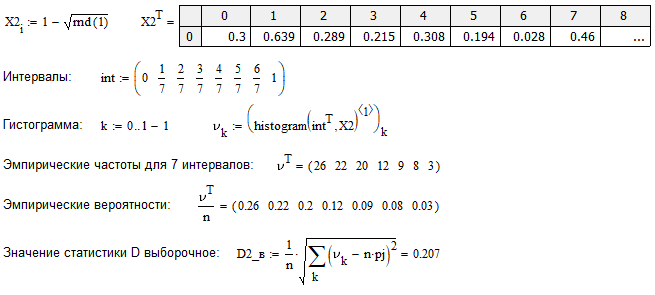
D\_в < Dкр для всех уровней доверия, поэтому гипотеза о том, что данная выборка принадлежит распределению R[0,1] принимается. Критерий сработал успешно.

Рассмотрим теперь работу этого критерия на примере выборок, не подчиняющихся равномерному закону.

1. **Пусть выборка расределена по закону ,**

Полученное значение D1\_в сильно превышает Dкр для всех уровней доверия, поэтому гипотеза о том, что выборка принадлежит распределению   
 отвергается. Критерий сработал успешно.

1. **Пусть выборка расределена по закону ,**

Полученное значение D2\_в сильно превышает Dкр для всех уровней доверия, поэтому гипотеза о том, что выборка принадлежит распределению отвергается. Критерий сработал успешно.

**Вывод:** В ходе решения данной задачи с помощью стохастического эксперимента на основе указанной метрики был построен приближенный критерий для проверки основной гипотезы о том, что выборка получена из закона распределения, совпадающего с распределением R[0,1]. Были получены приближенные значения квантилей уровня 0.9, 0.95, 0.99. Данный критерий был применен к трем различным выборкам. Для первой выборки, согласно результату применения критерия, основная гипотеза принимается, а для второй и третьей выборки, основная гипотеза отвергается. Эти результаты соответствуют нашим ожиданиям. Таким образом, построенный критерий успешно прошел проверку.