Министерство образования и науки РФ

Федеральное государственное бюджетное образовательное

Учреждение высшего образования

«Тверской государственный университет»

Факультет прикладной математики и кибернетики

Направление «Прикладная математика и информатика»

**Курсовая работа по дисциплине:  
«Теория вероятностей и математическая статистика»**

**Автор:**

Осинский Роман Русланович

III курс, 34 группа

**Проверил:**

к.ф.-м.н, доцент

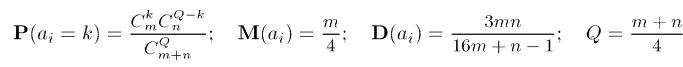
Сидорова Оксана Игоревна

Тверь, 2019

Теоретическая часть

Квартильный критерий Барнетта-Эйсена

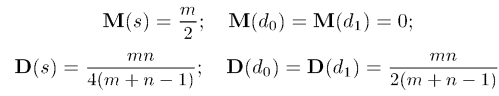
Предположим, что имеются две выборки – X и Y, объема m и n соответственно и m + n делится на 4. Объединим эти выборки в одну. Обозначим через ai количество членов выборки Х, попавших в i-ю квартиль объединенной выборки (i = 1, 2, 3, 4). Если обе выборки принадлежат одному распределению, то ai имеет гипергеометрическое распределение и:



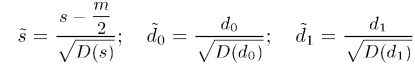
Введем обозначения:

s1 = a1 + a4; d0 = a4 – a1; d1 = a3 – a2

Тогда:



Обозначим:



В качестве критерия согласия возьмем статистику D =

При n, m -> inf D имеет распределение хи-квадрат с тремя степенями свободы.

Если D > (3), то гипотеза H0 о согласии эмпирических распределений отвергается.

Пример

Имеются две выборки, объем каждой равен 16:

X = 1, 3, 5, 7, 9, 11, 13, 18, 21, 23, 25, 30, 32, 34, 35, 37

Y = -8, -7, 4, 5, 6, 7, 8, 12, 14, 15, 18, 19, 20, 21, 23, 25

Проверим гипотезу о согласии эмпирических распределений критерием Бернетта-Эйсена. Объединим выборку:

|  |  |  |  |
| --- | --- | --- | --- |
| -8, -7, 1, 3, 4, 5, 5, 6 | 7, 7, 8, 9, 11, 12, 13, 14 | 15, 18, 18, 19, 20, 21, 21, 23 | 23, 25, 25, 30, 32, 34, 35, 37 |

Разделим ее на четыре равный части.

В первой квартили находится а1 = 3 элементов Х (1, 3, 5). Во второй а2 = 4 элементов Х (7, 9, 11, 13). В третьей а3 = 3 элемента выборки Х (18, 21, 23). В четвертой а4 = 6 элементов выборки Х (25, 30, 32, 34, 35, 37).

Получаем s = 3 + 6 = 9, d0 = 6 – 3 = 3, d1 = 3 – 4 = -1

M(s) = 16/2 = 8, M(d0) = M(d1) = 0

D(s) = 2.0645, D(d0) = D(d1) = 4.129

Š = (9 – 9)/1.4368 = 0.6959

Ď0 = 3/2.032 = 1.4763

Ď1 = -1/2.032 = -0.4921

Тогда D = 2.9

Найдем критическую точку при α = 0.9:

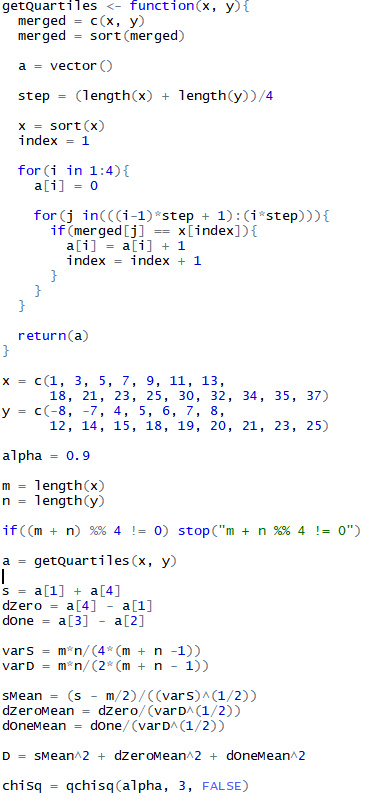
(3) = 6.251

Так как D = 2.9 < (3) = 6.251, то оснований отвергать Н0 нет.

Замечание: критерий сильно от связанных групп, которые попадают в разные квартили. Если бы в этом примере мы бы отнесли значение 23 не к третьей квартили, а к четвертой, то получили бы другую статистику D = 6.7813 и отвергли бы гипотезу H0.

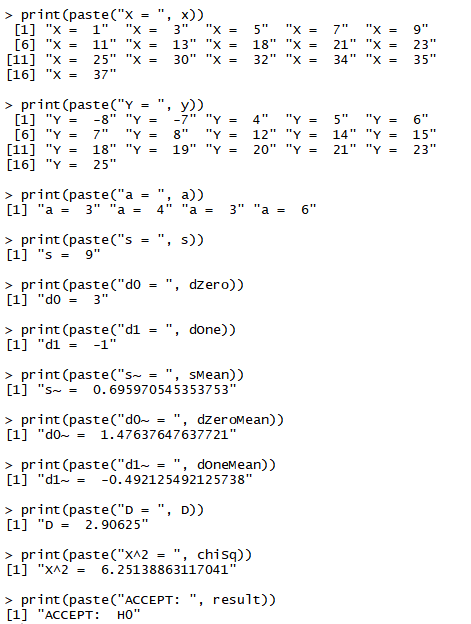
Практическая часть

Напишем программу на языке программирования R, принимающую на вход две выборки и сравнивающую их эмпирические функции распределения с помощью критерия Барнетта-Эйсена. Листинг:

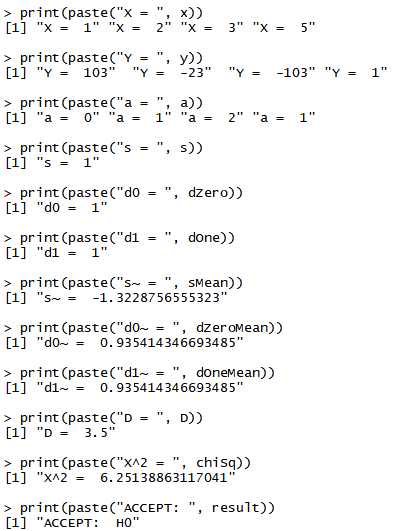


Проверим результаты выполнения программы на примерах:

1. α = 0.9



1. α = 0.9



1. α = 0.95

