

U-критерий Манна – Уитни

Квантованный текст с заданиями в тестовой форме для студентов высшей школы

А. Аймаханова,
Казахский Национальный медицинский университет
им. С.Д. Асфендиярова,
модуль Мед. биофизики и биostatистики

Введение

U-критерий Манна – Уитни — непараметрический критерий различия между 2-мя независимыми выборками. Метод определяет — насколько слабо перекрещиваются (совпадают) значения между двумя выборками. Чем меньше перекрещивающихся значений, тем более вероятно, что различия достоверны.

Впервые непараметрические критерии применили в 30-х годах XX века. Они отличаются простотой проведения, для них не требуется вычислять какие-либо параметры распределения (средние значения, стандартные отклонения и др.).

Границы применимости

Критерий предназначен для оценки различий между двумя выборками по уровню какого-либо количественно измеренного признака, при распределении вариант отличим от нормального. Более того, он позволяет выявлять различия между малыми выборками (когда $n_1, n_2 \geq 3$ или $n_1=2, n_2 \geq 5$).

Основы расчёта непараметрических критериев

В основе расчёта непараметрических критериев лежит упорядочивание (ранжирование) имеющихся значений по отношению друг к другу. Это разграничение значений не предполагает точных количественных соотношений, а следовательно, и ограничений на параметры и вид распределения.

Ранжирование — распределение вариантов внутри вариационного ряда, от меньших величин к большим.

Правила ранжирования

1. Меньшему значению начисляется меньший ранг, как правило, это 1. Наибольшему значению начисляется ранг, соответствующий количеству ранжируемых значений.
2. Если несколько значений равны, им начисляется ранг, представляющий собой среднее значение из тех рангов, которые они получили бы, если бы не были равны.

3. Общая сумма рангов должна совпадать с расчётной, которая определяется по формуле: $\sum (R) = \frac{N(N+1)}{2}$,

где N — общее количество ранжируемых значений. Несовпадение реальной и расчётной сумм рангов будет свидетельствовать об ошибке, допущенной при начислении рангов или их суммировании.

Алгоритм проверки гипотезы

Сформулировать нулевую и альтернативную гипотезы.

1. Задать уровень значимости α .
2. Вычислить значение $U_{эмп}$.
3. Найти критическое значение.
4. Сделать вывод.

Гипотезы

«Нулевая гипотеза» — это предположение о том, что в сравниваемых группах отсутствует различие в распределении частот.

«Альтернативная гипотеза» — это предположение о том, что в сравниваемых группах существует различие в распределении частот.

Уровень значимости

Уровень значимости — это такая вероятность, которую принимают за основу при статистической оценке гипотезы.

Уровень значимости α может быть принят равным 0,05 или 0,01.

Вашему вниманию предлагаются задания, в которых могут быть 1, 2, 3 и большее число правильных ответов. Нажимайте на клавиши с номерами всех правильных ответов:

1. ПОЗВОЛЯЕТ ОПРЕДЕЛИТЬ, НАСКОЛЬКО СЛАБО ПЕРЕКРЕЩИВАЮТСЯ ЗНАЧЕНИЯ МЕЖДУ ДВУМЯ ВЫБОРКАМИ

- 1) критерий Пирсона
- 2) критерий Стьюдента
- 3) критерий Уилкоксона
- 4) критерий Мак-Нимера

Схема подсчёта критерия Манна – Уитни:

1. Создать таблицу. (1 столбец — одна сравниваемая группа, 2 столбец — вторая).

2. Проранжировать значения вариантов в обоих столбцах.

3. Подсчитать сумму рангов для первого и второго столбцов отдельно. Проверить, совпадает ли общая сумма рангов с расчётной.

4. Определить значение $U_{эмп}$ по формуле: $U = (n_1 \cdot n_2) + \frac{n_x(n_x + 1)}{2} - T_x$, где n_1 —

количество вариантов в выборке 1; n_2 — количество вариантов в выборке 2; T_x — большая из ранговых сумм; n_x — количество вариантов в группе с большей суммой рангов.

Критическое значение

Определить критические значения $U_{кр}$ по таблице критических значений распределения Манна — Уитни. Оно должно соответствовать уровню значимости, объёмам выборок n_1 и n_2 соответственно.

Вывод

Сравниваются $U_{эмп}$ и $U_{кр}$. Если $U_{эмп} > U_{кр}$, то нулевая гипотеза принимается, т.е. в сравниваемых группах отсутствует различие в распределении частот.

Если $U_{эмп} < U_{кр}$, то нулевая гипотеза отвергается, т.е. в сравниваемых группах существует различие в распределении частот.

- 5) критерий Манна – Уитни
- 6) нормальное распределение

2. ВПЕРВЫЕ НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ ПРИМЕНИЛИ В

- 1) 20-х
- 2) 30-х
- 3) 40-х
- 4) 50-х

ГОДАХ

- 1) XX
- 2) XXI
- 3) XIX
- 4) XVI
- 5) XVII
- 6) XVIII

СТОЛЕТИЯ.

3. НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ ИСПОЛЬЗУЮТСЯ

- 1) для зависимых выборок
- 2) для независимых выборок
- 3) в случае нормально распределённых выборок
- 4) при неизвестном распределении
- 5) при вычислении коэффициента корреляции
- 6) для определения нормальности выборок
- 7) для определения степени линейности между двумя выборками

4. ОТЛИЧИЕ НЕПАРАМЕТРИЧЕСКИХ МЕТОДОВ от ПАРАМЕТРИЧЕСКИХ

- 1) не требуют предварительного знания вида распределения
- 2) не требуют предварительного расчёта параметров распределения
- 3) позволяют уравнивать совокупности с номинальными и порядковыми признаками
- 4) просты в применении
- 5) сложны в применении
- 6) применяются только для независимых выборок
- 7) применяются только для связанных выборок

5. ГРАНИЦЫ ПРИМЕНИМОСТИ КРИТЕРИЯ МАННА – УИТНИ ДЛЯ ОЦЕНКИ

- 1) средних двух
 - 2) средних трёх
 - 3) различий двух
 - 4) различий более чем двух
 - 5) средних более чем двух
- ВЫБОРОК,**
- 6) подчиняющихся
 - 7) не подчиняющихся
- НОРМАЛЬНОМУ ЗАКОНУ РАСПРЕДЕЛЕНИЯ.**

6. КРИТЕРИЙ ПОЗВОЛЯЕТ ВЫЯВЛЯТЬ РАЗЛИЧИЯ МЕЖДУ

- 1) малыми выборками
- 2) большими выборками

- 3) равными выборками
- 4) независимыми выборками
- 5) зависимыми выборками

КОГДА

- 1) $n_1, n_2 \geq 3$
- 2) $n_1=2, n_2 \geq 5$
- 3) $n_1, n_2 \geq 5$
- 4) $n_1=3, n_2 \geq 5$

7. В ОСНОВЕ РАСЧЁТА НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ ЛЕЖИТ

- 1) оценивание
- 2) упорядочивание
- 3) рандомизация
- 4) репрезентативность
- 5) ограничение
- 6) ранжирование
- 7) точное количественное соотношение

8. РАНЖИРОВАНИЕ – ЭТО _____

Установить правильную последовательность

9. ПРАВИЛА РАНЖИРОВАНИЯ

- меньшему значению начисляется меньший ранг
- общая сумма рангов должна совпадать с расчётной
- если несколько значений равны, им начисляется ранг, представляющий собой среднее значение из тех рангов, которые они получили бы, если бы не были равны
- найти общую сумму рангов
- наибольшему значению начисляется ранг, соответствующий количеству ранжируемых значений
- вычислить расчётную формулу
- определить общее количество ранжируемых значений

10. МЕТОДИКА ПРОВЕРКИ ГИПОТЕЗЫ ПРИ ПОМОЩИ КРИТЕРИЯ МАННА – УИТНИ

- задание уровня значимости α
- вывод
- ранжирование вариантов в обеих выборках
- вычисление статистики критерия
- формулирование нулевой и альтернативной гипотез
- сравнение статистики критерия с критическим значением распределения

11. СХЕМА ПОДСЧЁТА КРИТЕРИЯ МАННА – УИТНИ

- проверить, совпадает ли общая сумма рангов с расчётной формулой
- проранжировать значения вариантов в обоих столбцах
- определить значение $U_{эмп}$
- подсчитать сумму рангов для первого столбца
- создать таблицу (1 столбец – одна сравниваемая группа, 2 столбец – вторая)
- подсчитать сумму рангов для второго столбца

12. {Нулевая гипотеза, уровень значимости} – ЭТО

- 1) в сравниваемых группах отсутствует различие в распределении частот
- 2) основная гипотеза, которая проверяется
- 3) то, что оказывает влияние на конечный результат
- 4) значение наблюдаемой переменной
- 5) значение измеряемого признака
- 6) вероятность, которую принимают за основу при статистической оценке гипотезы
- 7) вероятность допущения ошибки 1-го рода
- 8) вероятность отклонения нулевой гипотезы, когда она ложна

13. УРОВЕНЬ ЗНАЧИМОСТИ α МОЖЕТ БЫТЬ ПРИНЯТ РАВНЫМ

- 1) 0,05
- 2) 0,001
- 3) 0,5
- 4) 0,01
- 5) 0,1
- 6) 0,002
- 7) 0,02

14. НУЛЕВАЯ ГИПОТЕЗА КРИТЕРИЯ МАННА – УИТНИ

- 1) групповые средние равны
- 2) в сравниваемых группах отсутствует различие в распределении частот
- 3) различие средних случайно
- 4) в сравниваемых группах существует различие в распределении частот
- 5) групповые средние не равны
- 6) уровень признака в выборке 2 не ниже уровня признака в выборке 1
- 7) различие средних не случайно
- 8) уровень признака в выборке 2 равен уровню признака в выборке 1

15. АЛЬТЕРНАТИВНАЯ ГИПОТЕЗА КРИТЕРИЯ МАННА – УИТНИ

- 1) групповые средние равны
- 2) в сравниваемых группах отсутствует различие в распределении частот
- 3) различие средних случайно
- 4) в сравниваемых группах существует различие в распределении частот
- 5) групповые средние не равны
- 6) уровень признака в выборке 2 не ниже уровня признака в выборке 1
- 7) различие средних не случайно
- 8) уровень признака в выборке 2 равен уровню признака в выборке 1

16. ФОРМУЛА РАСЧЁТА КРИТЕРИЯ МАННА – УИТНИ

$$1) = n_1 + n_2 + \frac{n_x(n_x + 1)}{2} + T_x$$

$$2) = n_1 n_2 + \frac{n_x(n_x + 2)}{n} - T_x$$

$$3) = n_1 n_2 + \frac{n_x(n_x + 1)}{2} - T_x$$

$$4) = n_1 n_2 + n_x(n_x + 1) - T_x$$

$$5) = n_1 n_2 + n_x(n_x + 1) + T_x$$

17. РАВЕНСТВО

1) $\sum(R) = \frac{(N+1)}{2}$

2) $\sum(R) = \frac{2N(N+1)}{2}$

3) $\sum(R) = \frac{N(N+2)}{2}$

4) $\sum(R) = \frac{N}{2}$

5) $\sum(R) = \frac{N(N+1)}{2}$

6) $\sum(R) = \frac{N+N+1}{2}$

СВИДЕТЕЛЬСТВУЕТ О ВЕРНОСТИ, ДОПУЩЕННОЙ ПРИ НАЧИСЛЕНИИ РАНГОВ ИЛИ ИХ СУММИРОВАНИИ

18. КРИТИЧЕСКОЕ ЗНАЧЕНИЕ КРИТЕРИЯ МАННА – УИТНИ НАХОДИТСЯ С УЧЁТОМ

1) $\alpha = 0,05$

2) $\alpha = 0,01$

3) $\alpha = 0,95$

4) $\alpha = 0,99$

5) n_1

6) n_2

7) n_x

8) N

9) $\sum(R_i)$

19. ЕСЛИ ПРИ ПРИМЕНЕНИИ КРИТЕРИЯ МАННА – УИТНИ ВЫПОЛНЯЕТСЯ НЕРАВЕНСТВО

1) $U_{\text{набл}} > U_{\text{кр}}$

2) $U_{\text{набл}} ? U_{\text{кр}}$

3) $U_{\text{набл}} = U_{\text{кр}}$

4) $U_{\text{набл}} < U_{\text{кр}}$

6) $U_{\text{набл}} \geq U_{\text{кр}}$

ТО НУЛЕВАЯ ГИПОТЕЗА ПРИНИМАЕТСЯ

20. ЕСЛИ ПРИ ПРИМЕНЕНИИ КРИТЕРИЯ МАННА – УИТНИ ВЫПОЛНЯЕТСЯ НЕРАВЕНСТВО

1) $U_{\text{набл}} > U_{\text{кр}}$

2) $U_{\text{набл}} ? U_{\text{кр}}$

3) $U_{\text{набл}} = U_{\text{кр}}$

4) $U_{\text{набл}} < U_{\text{кр}}$

6) $U_{\text{набл}} \geq U_{\text{кр}}$

ТО НУЛЕВАЯ ГИПОТЕЗА ОТВЕРГАЕТСЯ