***Marginal homogeneity***

SPSS macros by Kirill Orlov

kior@akado.ru, ttnphns@gmail.com

<https://www.spsstools.net/en/KO-spssmacros>

All rights reserved

*Тесты краевой однородности.* Статистические критерии, тестирующие пару переменных как спаренные выборки на предмет одинаковости их краевых распределений или локаций в них.

# МАКРОС !KO\_MHTESTS: КРИТЕРИИ КРАЕВОЙ ОДНОРОДНОСТИ

Version 1, Sep 2024. Tested on SPSS Statistics 22, 27, 29.

!KO\_mhtests vars1= *pre1 pre2 pre3* /\*Первый список переменных

/vars2= *post1 post2 post3* /\*Ему в соответствие второй список переменных

/values= AUTO /\*Значения в переменных: список из нескольких значений

/\*или AUTO (позволить макросу выяснить самому)

/level= NOM /\*Мерительный уровень данных: номинальный (NOM, тж п/у),

/\*порядковый (ORD), дискретный интервальный (INT)

/print= YES /\*Показывать частотную таблицу в выдаче результатов: YES или NO (тж п/у)

/save= /\*Сохранить результаты как массив данных: YES или NO (тж п/у).

Минимум надо задать VARS1, VARS2, VALUES.

Макрос исполняет тесты краевой однородности для спаренных выборок. Пусть есть две переменные с одинаковым набором из *k* определенных, категориальных или дискретных, значений. Эти переменные понимаются как спаренные выборки, например повторные замеры у каждого индивида его состояний «до» и «после» или исходы у сочетанных индивидов «без воздействия» и «с воздействием». Предмет тестирования – одинаковы ли у «до» vs «после» или у «без воздействия» vs «с воздействием» распределения в популяции.

Если данные номинальные, то нулевая гипотеза: «распределения одинаковы», и альтернативная: «не одинаковы». Это проверяют критерии Stuart–Maxwell и Bhapkar. Если данные количественные, то нулевая гипотеза: «распределения совпадают положением (location)», и альтернативная: «не совпадают: значения в одном выше, чем в другом». Это проверяют критерии Agresti, их два вида – один для порядковых данных и другой для дискретных интервальных данных[[1]](#footnote-1).

Все эти тесты асимптотические и нуждаются в «большой» выборке – не менее примерно 50 наблюдений в массиве данных.

**Алгоритм**

Переменные и категориальные или количественные дискретные. Значения (или коды) в переменных – одинаковы, так что и могут образовать квадратную частотную таблицу (таблицу сопряженности) размером *k*×*k* (где *k* – число различных значений) с одноименными рядами и столбцами. В случае порядковых или интервальных переменных ряды/столбцы таблицы – сортированы по возрастанию значений.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *V2* | | | |  |
| *V1* |  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |  |

, ,

A. Переменные номинальные [2, 5, 6].

Нуль-гипотеза: краевые распределения у и в популяции одинаковы.

Пусть вектор **d** есть любые *k-*1 элементов вектора  (*i*= 1, 2, …, *k*).

И пусть квадратная матрица (выборковая ковариационная матрица) есть выбранные с теми же номерами *k-*1 рядов и столбцов матрицы с элементами

для

(*i*= 1, 2, …, *k*).

Тогда тест-статистика Stuart–Maxwell

имеет хи-квадрат-распределение с *k*-1 степенями свободы, из которого получаем *p*-значение.

Статистика Bhapkar имеет ту же формулу, что , и тоже хи-квадрат-распределение, но элементы определяются иначе (теперь это оцененная ковариационная матрица):

для .

Обе статистики связаны отношением .

Статистика Bhapkar мощнее статистики Stuart–Maxwell, но с ростом N различие между ними сходит на нет. Значимая статистика говорит о том, что в популяции краевые распределения не равны, не уточняя, как именно не равны. Если *k*=2, совпадает с тестом Макнемара (без поправки Йейтса).

*Примечание*. Если среди *k* категорий есть такие, где ненулевая частота – только диагональная, *nii*, то такие категории (ряды/столбцы) удаляются из таблицы, и тест вычисляется на оставшейся таблице, однако число степеней свободы остается исходным.

B. Переменные порядковые [1, 4].

Нуль-гипотеза: стохастическое преобладание одного краевого распределения над другим

равно в популяции нулю.

Выборковое значение

*,*

где и есть ридиты (ridit scores) в переменных и , соответственно:

(ридит *j*-й градации это половина частоты в ней плюс суммарная частота в более младших градациях).

Пусть . Стандартная ошибка в большой выборке:

*,*

и тест-статистика

имеет приблизительно ст. нормальное распределение, из которого получаем односторонее и двусторонее *p*-значения. Положительное (и *z*) говорит о стохастическом преобладании над – т.е. что взятое из популяции случайное наблюдение будет иметь тенденцию быть по переменной выше, чем по переменной . Отрицательное (и *z*) говорит о противоположной тенденции.

C. Переменные дискретные интервальные [1, 3].

Нуль-гипотеза: средние в популяции равны:

.

Средние в выборке:

и ,

где есть балл (само значение) в переменных и .

Их разница

имеет стандартную ошибку в большой выборке:

*,*

и тест-статистика

имеет ст. нормальное распределение, из которого получаем односторонее и двусторонее *p*-значения. Положительное (и *z*) говорит о более высоком среднем балле в , чем в . Отрицательное (и *z*) говорит о противоположном. По своим результатам данный тест весьма близок к парному t-тесту Стьюдента.

Ссылки

1. Agresti, A. Testing marginal homogeneity of ordinal categorical variables // Biometrics, 1983, 39(2), p. 505-510.

2. Agresti, A. Categorical data analysis. 2nd edition. 2002. p. 422 (“Marginal Models for Nominal Classifications”).

3. Agresti, A. Analysis of ordinal categorical data. 2nd edition. 2010. p. 227 (“Comparing Marginal Mean Scores”).

4. Agresti, A. Analysis of ordinal categorical data. 2nd edition. 2010. p. 227 (“Comparing Marginal Mean Ranks or Mean Ridits”).

5. Sun, X. et al. Generalized Mcnemar's test for homogeneity of the marginal distributions // SAS Global Forum 2008, Paper 382-2008.

6. McNemar tests of marginal homogeneity // URL: <https://www.john-uebersax.com/stat/mcnemar.htm#bhapkar> [Reached 03.10.2024].

***Подкоманды***

**VARS1, VARS2**

Два поименных списка числовых переменных одинаковой длины. 1-я переменная VARS1 будет сравнена (спарена) с 1-й переменной VARS2; 2-я переменная VARS1 будет сравнена (спарена) со 2-й переменной VARS2, и так далее. Две переменные пары – это две «спаренные выборки». Переменные могут повторяться в списках.

Переменные должны быть категориальные или количественные с дискретными значениями. Набор значений должен быть общий между i-й переменной VARS1 и i-й переменной VARS2. Пропуски в данных разрешены, они исключаются попарно: в паре i-й переменной VARS1 и i-й переменной VARS2 наблюдение исключается, если оно пропуск в любой из двух.

**VALUES**

Задайте одно из двух:

AUTO - использовать все валидные значения, какие наблюдаются в паре переменных. Наборы значений между переменными пары должны полностью совпадать. Между парами наборы значений не обязаны совпадать.

*Список значений* - перечислите нужные валидные значения (минимум два), которые анализировать во всех парах переменных. Эти значения должны существовать во всех парах. Кроме них в переменных могут существовать и иные значения – они будут проигнорированы.

Внимание, при LEVEL= ORD или INT *список значений* необходимо писать по возрастанию значений.

ПРИМЕР 1.

!KO\_mhtests vars1= v1 v1 /vars2= var1 var2 /values= AUTO /level= NOM.

* Пары переменных (спаренные выборки): *V1* с *VAR1*, *V1* с *VAR2*. Т.к. VALUES=AUTO, все значения в паре будут анализироваться, а поскольку *V1* входит в обе пары, то, значит, все четыре переменных должны полностью совпадать наборами своих значений.

ПРИМЕР 2.

!KO\_mhtests vars1= before\_a before\_b /vars2= after\_a after\_b /values= 1 2 4 /level= INT.

* Пары переменных (спаренные выборки): *BEFORE\_A* с *AFTER\_A*, *BEFORE\_B* с *AFTER\_B*. Значения 1, 2 и 4 в этих переменных будут анализироваться. Эти значения должны присутствовать в обеих парах.
* LEVEL=INT, поэтому список VALUES должен идти непременно по возрастающей.

**LEVEL**

Укажите уровень измерения входящих переменных (они все должны подразумевать единый уровань). Он определяет, какой статистический критерий (тест) будет применен.

NOM - (тж. по умолчанию/незаданию) входящие переменные – номинальные. Тесты Stuart–Maxwell и Bhapkar.

ORD - входящие переменные – порядковые. Тест Agresti маргинальной однородности для порядковых данных (баллами выступают ридиты уровней).

INT - входящие переменные – дискретные интервальные. Тест Agresti маргинальной однородности для дискретных интервальных данных (баллами выступают значения уровней).

**PRINT**

Укажите PRINT=YES, чтобы макрос печатал в окно результатов частотную таблицу сопряженности для каждой пары переменных. П/у PRINT=NO.

**SAVE**

Укажите SAVE=YES, чтобы макрос сохранил полученные результаты как новый массив данных. П/у SAVE=NO.

***Особые режимы***

Макрос игнорирует расщепленное состояние массива данных и взвешивание массива. Макрос слушается команд, выбирающих наблюдения (FILTER, USE, SELECT IF, N OF CASES). Макрос слушается временных (стоящих под TEMPORARY) операций.

1. Этот последний совпадает с тестом краевой однородности, представленным в «Непараметрических критериях» в SPSS. [↑](#footnote-ref-1)