***Marginal homogeneity***

SPSS macros by Kirill Orlov

kior@akado.ru, ttnphns@gmail.com

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

All rights reserved

*Marginal homogeneity tests.* Statistical tests testing a pair of variables as paired samples for equality of their marginal distributions or locations of those.

# MACRO !KO\_MHTESTS: MARGINAL HOMOGENEITY TESTS

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

!KO\_mhtests vars1= *pre1 pre2 pre3* /\*First list of variables

/vars2= *post1 post2 post3* /\*Second list of variables in correspondence to it

/values= AUTO /\*Values in the variables: list of several values

/\*or AUTO (let the macro find out for itself)

/level= NOM /\*Measurement level of the data: nominal (NOM, default),

/\*ordinal (ORD), discrete interval (INT)

/print= YES /\*Show frequency table in the results output: YES or NO (default)

/save= /\*Save results as a dataset: YES or NO (default).

Minimal specification VARS1, VARS2, VALUES.

The macro performs tests of marginal homogeneity for paired samples. Let there be two variables with the same set of *k* definite, categorical or discrete, values. The variables are understood as paired samples, for example, repeated measurements of each individual’s states “before” and “after” or outcomes in matched individuals of “no treatment” and “treatment”. The issue of testing is whether distributions in population are same between “before” and “after” or between “no treatment” and “treatment”.

If data are nominal, the null hypothesis is “distributions are same” and alternative is “not same”. Stuart–Maxwell and Bhapkar tests check for this. If data are quantitative, the null hypothesis is “distributions coincide in location” and alternative is do not coincide: values in one are higher than in the other”. Agresti’s tests check for this, and there are two versions – one for ordinal data and another for discrete interval data[[1]](#footnote-1).

All these tests are asymptotic and need a “large” sample – not less than about 50 cases in the dataset.

**Algorithm**

Variables and are categorical or quantitative discrete. Values (or codes) in the variables are same, so that and can form the square frequency table (contingency table) sized *k*×*k* (where *k* is the number of distinct values) with the same-labelled rows and columns. In case of ordinal or interval variables the rows/columns of the table are sorted ascendingly of the values.

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

, ,

A. Variables are nominal [2, 5, 6].

Null hypothesis: marginal distributions of and are same in the population.

Let vector **d** be any *k-*1 elements of vector  (*i*= 1, 2, …, *k*).

And let square matrix (sample covariance matrix) be selected with the same numbers *k-*1 rows and columns of matrix with elements

for

(*i*= 1, 2, …, *k*).

Then Stuart–Maxwell test statistic

has chi-square distribution with *k*-1 degrees of freedom, from where we obtain the *p*-value.

Bhapkar’s statistic has the same formula as , and also the chi-square distribution, but elements of are defined differently (now it is estimated covariance matrix):

for .

Both statistics are linked by the relation .

Bhapkar’s statistic is more powerful than Stuart–Maxwell statistic, but as N grows, their difference vanishes. Significant statistic tells that marginal distributions are not equal in the population, without specifying how exactly unequal. If *k*=2, coincides with the McNemar’s test (not Yates corrected).

*Note*. If among *k* categories there are ones where nonzero frequency is only the diagonal one, *nii*, then those categories (rows/columns) are deleted from the table, and the test is computed on the remaining table, however, the number of degrees of freedom remains initial.

B. Variables are ordinal [1, 4].

Null hypothesis: stochastic superiority of one distribution over the other

equals to zero in the population.

The sample value

*,*

where and are ridit scores in the variables and , respectively:

(ridit of the *j*-th level is the half frequency in it plus the summed frequency in more junior levels).

Let . Standard error of in large sample:

*,*

and test statistic

has approximately st. normal distribution, from where we obtain one-sided and two-sided *p*-values. Positive (and *z*) tells about stochastic superiority of over – i.e., that a random observation taken from the population will tend to be higher by variable than by variable . Negative (and *z*) tells about the opposite tendency.

C. Variables are discrete interval [1, 3].

Null hypothesis: means in the population are equal:

.

Means in the sample:

and ,

where is a score (a value itself) in variables and .

Their difference

has large sample standard error:

*,*

and test statistic

has st. normal distribution, from where we obtain one-sided and two-sided *p*-values. Positive (and *z*) tells about higher mean score in than in . Negative (and *z*) tells about the opposite. By its results, the test is quite close to the paired-sample Student’s t-test.

References

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].

***Subcommands***

**VARS1, VARS2**

Two name-by name lists of numeric variables, lists of the same length. The 1st variable of VARS1 will be compared (paired) with the 1st variable of VARS2; the 2nd variable of VARS1 will be compared (paired) with the 2nd variable of VARS2, and so on. Two variables of a pair are the “paired samples”. Variables may repeat in the lists.

The variables must be categorical or quantitative with discrete values. The value set should be common between an i-th variable of VARS1 and an i-th variable of VARS2. Missings are allowed in the data, they are excluded pairwise: a case is excluded from both an i-th variable of VARS1 and from an i-th variable of VARS2, if it is missing in any of the two.

**VALUES**

Specify one of the two:

AUTO - use all valid values which are observed in the variable pair. Value sets must completely coincide between the variables of the pair. Value sets does not have to coincide between the pairs.

*Value list* - enumerate the needed valid values (two minimum) which to analyze in all the variable pairs. These values must exist in all the pairs. Besides these, other values may exist in the variables – they will be ignored.

Attention, with LEVEL= ORD or INT the *value list* should be written ascendingly of the values.

EXAMPLE 1.

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

* Variable pairs (paired samples): *V1* with *VAR1*, *V1* with *VAR2*. Because VALUES=AUTO, all the values in a pair will be analyzed, and since *V1* enters both pairs, it implies all the four variables should completely coincide by thir value sets.

EXAMPLE 2.

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

* Variable pairs (paired samples): *BEFORE\_A* with *AFTER\_A*, *BEFORE\_B* with *AFTER\_B*. Values 1, 2, and 4 in these variables will be analyzed. The values must exist in both pairs.
* LEVEL=INT, so the list VALUES must go necessarily in ascending order.

**LEVEL**

Indicate the measurement level of the input variables (they all must imply the same level). It determines which statistical test will be applied.

NOM - (also default/unspecifying) input variables are nominal. Stuart–Maxwell and Bhapkar tests.

ORD - input variables are ordinal. Agresti marginal homogeneity test for ordinal data (ridits play the role of scores).

INT - input variables are discrete interval. Agresti marginal homogeneity test for interval data (levels’ values play the role of scores).

**PRINT**

Specify PRINT=YES in order the macro to print in the output window the frequency contingency table for each pair of variables. By default, PRINT=NO.

**SAVE**

Specify SAVE=YES in order the macro to save the results obtained, as a new dataset. By default, SAVE=NO.

***Special regimes***

The macro ignores split-file state of the dataset and weighting of the dataset. The macro obeys commands selecting cases (FILTER, USE, SELECT IF, N OF CASES). The macro obeys temporary (standing under TEMPORARY) operations.

1. That latter coincides with the test of marginal homogeneity found in “Nonparametric tests” in SPSS. [↑](#footnote-ref-1)