

Package ‘AnthropMMD’

March 18, 2019

Type Package

Title An R-Shiny Application for Smith's Mean Measure of Divergence (MMD)

Description Offers a graphical user interface for the calculation of the mean measure of divergence, with facilities for trait selection and graphical representations <doi:10.1002/ajpa.23336>.

Version 2.5.3

Date 2019-03-18

Depends R (>= 3.4.0), shiny, smacof, scatterplot3d

License GPL-3 | file LICENSE

Encoding UTF-8

BugReports <https://gitlab.com/f.santos/anthropmmd/issues/new>

NeedsCompilation no

Author Frédéric Santos [aut, cre] (<<https://orcid.org/0000-0003-1445-3871>>)

Maintainer Frédéric Santos <frederic.santos@u-bordeaux.fr>

Repository CRAN

Date/Publication 2019-03-18 11:53:56 UTC

R topics documented:

AnthropMMD-package	2
StartMMD	3
Index	6

AnthropMMD-package *An R-Shiny Application for Smith's Mean Measure of Divergence (MMD)*

Description

Offers a graphical user interface for the calculation of the mean measure of divergence, with facilities for trait selection and graphical representations.

Details

Package: AnthropMMD
Type: Package
Version: 2.5.3
Date: 2019-03-18
License: GPL-3

Author(s)

Frédéric Santos, <frederic.santos@u-bordeaux.fr>

References

- Harris, E. F. and Sjøvold, T. (2004) Calculation of Smith's mean measure of divergence for inter-group comparisons using nonmetric data. *Dental Anthropology*, **17**(3), 83–93.
- Irish, J. (2010) The mean measure of divergence: Its utility in model-free and model-bound analyses relative to the Mahalanobis D2 distance for nonmetric traits. *American Journal of Human Biology*, **22**, 378–395. doi: 10.1002/ajhb.21010
- Nikita, E. (2015) A critical review of the mean measure of divergence and Mahalanobis distances using artificial data and new approaches to the estimation of biodistances employing nonmetric traits. *American Journal of Physical Anthropology*, **157**, 284–294. doi: 10.1002/ajpa.22708
- Santos, F. (2018) AnthropMMD: an R package with a graphical user interface for the mean measure of divergence. *American Journal of Physical Anthropology*, **165**(1), 200–205. doi: 10.1002/ajpa.23336

Examples

```
## Not run: StartMMD()
```

Description

Launches a graphical user interface (GUI) for the calculation of the mean measure of divergence.

Usage

```
StartMMD()
```

Details

The GUI of AnthropMMD is completely autonomous: reading the data file and specifying the parameters of the analysis are done through the interface. Once the dataset is loaded, the output reacts dynamically to any change in the analysis settings.

- AnthropMMD accepts .CSV or .TXT data files, but does not support .ODS or .XLS(X) files. Two types of data input formats can be used:
 - A ‘Raw binary dataset’ (one row for each individual, one column for each variable). The first column must be the group indicator, and the other columns are binary data for the traits studied, where 1 indicates the presence of a trait, and 0 its absence. Row names are optional for this type of file. An example of valid data file can be found here: <http://tinyurl.com/RawBinaryMMD>
 - A ‘Table of n’s and absolute frequencies for each group’, i.e. a dataset of sample sizes and absolute frequencies. This type of dataset has $2 \times K$ rows (K being the number of groups compared) and p columns (p being the number of traits studied). The first K lines must be the group n’s for each trait, and the last K lines are absolute frequencies for each trait (i.e. the number of times the trait is present). Row names are mandatory for this type of file. The first K rows must be labelled with names beginning with ‘N_’, such as: N_GroupA, N_GroupB, ..., N_GroupK. The last K rows should be labelled with names beginning with ‘Freq_’, such as: Freq_GroupA, ..., Freq_GroupK. An example of valid datafile can be found here: <http://tinyurl.com/TableFreqMMD>

For both data types, column names are strongly recommended for better interpretability of the results.

- One can choose between Anscombe or Freeman-Tukey formula for angular transformation (cf. Harris and Sjøvold 2004; Irish 2010).
- ‘Only retain the traits with this minimal number of individuals per group’: the traits with fewer individuals in at least one active group will not be considered in the analysis.
- ‘Exclusion strategy’: a careful selection of traits is crucial when using MMD (cf. Harris and Sjøvold 2004 for a complete explanation), and the user should probably “exclude the traits that are nondiscriminatory across groups” (Irish 2010).
 - ‘Exclude nonpolymorphic traits’ removes all the traits showing no variability at all, i.e. with the same value (‘0’ or ‘1’) for all individuals.

- ‘Exclude quasi-nonpolymorphic traits’ also removes the traits whose variability is only due to a single individual: for example, a trait with only one positive observation in the whole dataset.
- ‘Use Fisher’s exact test’ implements the advice given by Harris and Sjøvold (2004) to select contributory traits, defined as those “showing a statistically significant difference between at least one pair of the groups being evaluated”. Fisher’s exact tests are performed for each pair of groups, and the traits showing no intergroup difference at all are excluded. Note that if you have a large number of groups (say, 10 groups), a trait with strictly equal frequencies for the last 8 groups may be considered as useful according to this criterion if there is a significant difference for the first two groups. This criterion will select all traits that can be useful for a given pair of groups, even if they are nondiscriminatory for all the other ones.
- ‘Exclude traits with overall MD’ lower than a given threshold: it is a simple way of removing the traits with quite similar frequencies across groups (the ‘overall MD’ is defined as the sum of the variable’s measures of divergence over all pairs of groups). This criterion aims to select the traits whose frequency differs substantially across most or all groups.

These four options are designed to avoid negative MMD values.

- Some groups/populations can be manually excluded from the analysis. This may be useful if very few individuals belonging to a given population could be recorded for the variables retained by the criteria described above.
- A MDS plot and a hierarchical clustering, done using MMD dissimilarities as inputs, are displayed in the last two tabs. As MMD can sometimes be negative, those negatives values are replaced by zeros, so that the MMD matrix can be seen as a symmetrical distance matrix. Please note that the classical two-dimensional metric MDS plot cannot be displayed if there is only one positive eigenvalue. Several MDS options are proposed, cf. the help page of the ‘smacofSym’ function from the R package ‘smacof’ for detailed technical information.

Value

The function returns no value by itself, but all results can be individually downloaded through the graphical interface.

- The ‘true’ MMD values (i.e., which can be negative in the case of small samples with similar traits frequencies, cf. Irish 2010) and their standard deviations are presented in the matrix labelled ‘MMD values (upper triangular part) and associated SD values (lower triangular part)’.
- A MMD value can be considered as significant if it is greater than twice its standard deviation. Significance is assessed in another ad-hoc table of results.
- The negative MMD values, if any, are replaced by zeros in the ‘Symmetrical matrix of MMD values’.

Note

The R console is not available when the GUI is active. To exit the GUI, type Echap (on MS Windows systems) or Ctrl+C (on Linux systems) in the R console.

On 14-inch (or smaller) screens, for convenience, it may be necessary to decrease the zoom level of your web browser and/or to turn on fullscreen mode.

Author(s)

Frédéric Santos, <frederic.santos@u-bordeaux.fr>

References

Harris, E. F. and Sjøvold, T. (2004) Calculation of Smith's mean measure of divergence for inter-group comparisons using nonmetric data. *Dental Anthropology*, **17**(3), 83–93.

Irish, J. (2010) The mean measure of divergence: Its utility in model-free and model-bound analyses relative to the Mahalanobis D2 distance for nonmetric traits. *American Journal of Human Biology*, **22**, 378–395. doi: 10.1002/ajhb.21010

Nikita, E. (2015) A critical review of the mean measure of divergence and Mahalanobis distances using artificial data and new approaches to the estimation of biodistances employing nonmetric traits. *American Journal of Physical Anthropology*, **157**, 284–294. doi: 10.1002/ajpa.22708

Santos, F. (2018) AnthroMMD: an R package with a graphical user interface for the mean measure of divergence. *American Journal of Physical Anthropology*, **165**(1), 200–205. doi: 10.1002/ajpa.23336

Examples

```
## Not run: StartMMD()
```

Index

*Topic **MMD, mean measure of
divergence, nonmetric traits**

AnthropMMD-package, [2](#)

AnthropMMD (AnthropMMD-package), [2](#)

AnthropMMD-package, [2](#)

StartMMD, [3](#)