

Package ‘cmahalanobis’

February 20, 2024

Type Package

Title Calculate the Mahalanobis Distance for a Given List of Data
Frames with Factors

Version 0.1.0

Description It provides a function that calculates the Mahalanobis distance between each pair of species in a list of data frames. Each data frame contains the observations of a species with some factors. Mahalanobis distance is a measure of dissimilarity between two vectors of multivariate random variables, based on the covariance matrix. This distance is useful for statistical matching or fusion of data, that is the integration of two data sources that refer to the same target population and that share some variables.
- ``Fisher, R.A. (1922) On the mathematical foundations of theoretical statistics. <[doi:10.1098/rsta.1922.0009](https://doi.org/10.1098/rsta.1922.0009)>".
- ``Mahalanobis, P.C. (1936) On the generalized distance in statistics. <[doi:10.1007/s13171-019-00164-5](https://doi.org/10.1007/s13171-019-00164-5)>".

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

Imports stats

Suggests testthat

NeedsCompilation no

Author Flavio Gioia [aut, cre] (<<https://orcid.org/0009-0000-0326-3840>>)

Maintainer Flavio Gioia <flaviogioia.fg@gmail.com>

Repository CRAN

Date/Publication 2024-02-20 21:10:09 UTC

R topics documented:

cmahalanobis	2
Index	4

`cmahalanobis`*Calculate the Mahalanobis distance for each species*

Description

. This function takes a list of data frames as input, where each data frame contains the observations of a species, and returns a matrix with the Mahalanobis distances between each pair of species.

Usage

```
cmahalanobis(dataset)
```

Arguments

`dataset` A list of data frames, where each data frame contains the observations of a species.

Value

A matrix with the Mahalanobis distances between each pair of species.

Examples

```
# Example with the iris dataset
library(stats)
# Split the data into 3 parts for each species
setosa <- subset(iris, Species == "setosa")
setosa <- setosa[,-5]
versicolor <- subset(iris, Species == "versicolor")
versicolor <- versicolor[,-5]
virginica <- subset(iris, Species == "virginica")
virginica <- virginica[,-5]

# Create a list with the three groups of flowers
groups <- list(setosa, versicolor, virginica)

# Calculate the Mahalanobis distance with the cmahalanobis function
cmahalanobis(groups)

# Example with the mtcars dataset
library(stats)
# Split the data into 2 parts for each type of transmission
auto <- subset(mtcars, am == 0)
auto <- auto[,-9]
manual <- subset(mtcars, am == 1)
manual <- manual[,-9]

# Create a list with the two groups of cars
groups <- list(auto, manual)
```

```
# Calculate the Mahalanobis distance with the cmahalanobis function  
cmahalanobis(groups)
```

Index

cmahalanobis, [2](#)