Package 'cmahalanobis'

May 17, 2024

Type Package

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

Version 0.2.0

Author Flavio Gioia

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

Description It provides a function that calculates the Mahalanobis distance be-
tween each pair of species in a list of data frames. Each data frame contains the observa-
tions of a species with some factors. Mahalanobis distance is a measure of dissimilarity be-
tween two vectors of multivariate random variables, based on the covariance matrix. This dis-
tance is useful for statistical matching or fusion of data, that is the integra-
tion of two data sources that refer to the same target population and that share some variables.

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

Imports stats, ggplot2, reshape2

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

NeedsCompilation no

VignetteBuilder knitr

Config/testthat/edition 3

Repository CRAN

Date/Publication 2024-05-17 17:00:02 UTC

R topics documented:

---

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,
  plot = TRUE,
  p.value = FALSE,
  plot_title = "Mahalanobis Distance Between Groups"
)
```

Arguments

dataset	A list of data frames, where each data frame contains the observations of a species.
plot	Logical, if TRUE, a plot of the Mahalanobis distances is displayed.
p.value	Logical, if TRUE, a matrix of p-values for the distances is returned.
plot_title	The title to be used for the plot if plot is TRUE.

Value

A list containing a matrix with the Mahalanobis distances between each pair of groups, and optionally a matrix of p-values and the plot.

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