Package 'cmahalanobis'

May 20, 2024

Type Package

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

Version 0.3.0

Author Flavio Gioia

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

Description

It provides functions that calculate Mahalanobis distance, Euclidean distance, Manhattan distance and Chebyshev distance between each pair of species in a list of data frames. These metrics are fundamental in various fields, such as cluster analysis, classification, and other applications of machine learning and data mining, where assessing similarity or dissimilarity between data is crucial. The package is designed to be flexible and easily integrated into data analysis workflows, providing reliable tools for evaluating distances in multidimensional contexts.

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 2024-05-20 13:10:03 UTC

R topics documented:

---

cchebyshev | *Calculate Chebyshev distance*

Description

Calculate Chebyshev distance

Usage

```
cchebyshev(
  dataset,
  plot = FALSE,
  p.value = FALSE,
  num.permutations = 1000,
  plot_title = "Chebyshev Distance Between Groups"
)
```

Arguments

dataset	A list of dataframes
plot	If TRUE, displays a plot of distances
p.value	If TRUE, calculates p-values of distances
num.permutations	
	Number of permutations to calculate p-values
plot_title	The title of plot

Value

A list containing distances and, optionally, p-values

---

ceuclide | *Calculate Euclidean distance*

Description

Calculate Euclidean distance

Usage

```
ceuclide(
  dataset,
  plot = FALSE,
  p.value = FALSE,
  num.permutations = 1000,
  plot_title = "Euclidean Distance Between Groups"
)
```

Arguments

`dataset`	A list of dataframes
`plot`	If TRUE, shows a plot of distances
`p.value`	If TRUE, calculates p-values for distances
`num.permutations`	
	Number of permutations to calculate p-values
`plot_title`	The title of plot

Value

A list containing distances and, optionally, p-values

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

ceuclide(groups, plot = TRUE, p.value = TRUE,
plot_title = "Euclidean Distance Between Groups")
cmanhattan(groups, plot = TRUE, p.value = TRUE,
plot_title = "Manhattan Distance Between Groups")
cchebyshev(groups, plot = TRUE, p.value = TRUE,
plot_title = "Chebyshev Distance Between 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, plot = TRUE, p.value = TRUE,
plot_title = "Mahalanobis Distance Between Groups")
ceuclide(groups, plot = TRUE, p.value = TRUE, num.permutations = 1000,
plot_title = "Euclidean Distance Between Groups")
cmanhattan(groups, plot = TRUE, p.value = TRUE, num.permutations = 1000,
plot_title = "Manhattan Distance Between Groups")
cchebyshev(groups, plot = TRUE, p.value = TRUE, num.permutations = 1000,
plot_title = "Chebyshev Distance Between Groups")
```

---

cmanhattan   *Calculate Manhattan distance*

Description

Calculate Manhattan distance

Usage

```
cmanhattan(
  dataset,
  plot = FALSE,
  p.value = FALSE,
  num.permutations = 1000,
  plot_title = "Manhattan Distance Between Groups"
)
```

Arguments

dataset	A list of dataframes
plot	If TRUE, show a plot of distances
p.value	If TRUE, calculates p-values to distances
num.permutations	
	Number of permutations to calculate p-values
plot_title	The title of plot

Value

A list containing distances and, optionally, p-values

Index