Package 'HotellingEllipse'

June 25, 2022

Title Hotelling T-Square and Confidence Ellipse

Version 1.1.0

Description Functions to compute the semi-axes lengths and coordinate points of Hotelling ellipse. Bro and Smilde (2014) <DOI:10.1039/c3ay41907j>. Brereton (2016) <DOI:10.1002/cem.2763>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

LazyDataCompression xz

RoxygenNote 7.2.0

URL https://github.com/ChristianGoueguel/HotellingEllipse

BugReports https://github.com/ChristianGoueguel/HotellingEllipse/issues

Imports dplyr, FactoMineR, ggforce, ggplot2, lifecycle, magrittr, purrr, stats, tibble

Depends R (>= 2.10)

Suggests rmarkdown, knitr, markdown, testthat (>= 3.0.0), spelling, covr

VignetteBuilder knitr

Config/testthat/edition 3

Language en-US

NeedsCompilation no

Author Christian L. Goueguel [aut, cre] (<https://orcid.org/0000-0003-0521-3446>)

Maintainer Christian L. Goueguel <christian.goueguel@gmail.com>

Repository CRAN

Date/Publication 2022-06-24 22:40:18 UTC

R topics documented:

`ellipseCoord` *Coordinate Points Of Hotelling Ellipse*

Description

Get the *x* and *y* coordinates of Hotelling ellipse.

Usage

```
ellipseCoord(data, pcx = 1, pcy = 2, conf.limit = 0.95, pts = 200)
```

Arguments

`data`	Data frame or tibble of PCA, PLS, and ICA scores, or from other feature projection methods.
`pcx`	Integer specifying which component is on the x-axis (by default 1).
`pcy`	Integer specifying which component is on the y-axis (by default 2).
`conf.limit`	Number between 0 and 1 specifying the confidence level (by default 0.95).
`pts`	Integer indicating the number of points for drawing the Hotelling ellipse (by default 200).

Value

Data frame containing the *x* and *y* coordinate points of the Hotelling ellipse.

Author(s)

Christian L. Goueguel, christian.goueguel@gmail.com

Examples

```
## Principal components analysis (PCA)
library(dplyr)
set.seed(123)
pca_mod <- specData %>%
  dplyr::select(where(is.numeric)) %>%
  FactoMineR::PCA(scale.unit = FALSE, graph = FALSE)

## Extract PCA scores
pca_scores <- pca_mod %>%
```

```
    purrr::pluck("ind", "coord") %>%
    tibble::as_tibble()

## Get Hotelling ellipse coordinate points
library(HotellingEllipse)
xy_coord <- ellipseCoord(data = pca_scores, pcx = 1, pcy = 2, conf.limit = 0.95, pts = 200)
```

`ellipseParam` *Lengths Of The Semi-Axes Of Hotelling Ellipse*

Description

Compute the lengths of the semi-axes of Hotelling ellipse.

Usage

```
ellipseParam(data, k = 2, pcx = 1, pcy = 2)
```

Arguments

`data`	Data frame or tibble of PCA, PLS, or ICA scores, or from other feature projection methods.
`k`	Integer specifying the number of components (by default 2).
`pcx`	Integer specifying which component is on the x-axis (by default 1).
`pcy`	Integer specifying which component is on the y-axis (by default 2).

Value

Returns a list that includes:

1. **Tsquare** Data frame containing the T-squared statistic.
2. **Ellipse** Data frame containing the lengths of the semi-minor and semi-major axes.
3. **cutoff.99pct** Number corresponding to the T-square cutoff at 99% confidence level.
4. **cutoff.95pct** Number corresponding to the T-square cutoff at 95% confidence level.

Author(s)

Christian L. Goueguel, christian.goueguel@gmail.com

Examples

```
## Principal components analysis (PCA)
library(dplyr)
set.seed(123)
pca_mod <- specData %>%
  dplyr::select(where(is.numeric)) %>%
  FactoMineR::PCA(scale.unit = FALSE, graph = FALSE)

## Extract PCA scores
pca_scores <- pca_mod %>%
   purrr::pluck("ind", "coord") %>%
   tibble::as_tibble()

## Get Hotelling T2-value and the lengths of the ellipse semi-axes
library(HotellingEllipse)
T2 <- ellipseParam(data = pca_scores, k = 2, pcx = 1, pcy = 2)
```

`specData` *LIBS spectra of 100 soil samples*

Description

A dataset containing the emission spectra of 171 soils measured in laboratory conditions. The samples were cleaned, dried, homogenized, sieved (10 Mesh size) and thereafter pelletized prior to LIBS measurements. LIBS spectra were preprocessed by performing baseline removal.

Usage

```
specData
```

Format

Data frame of 100 rows (soil samples) and 3152 columns (wavelengths).

Source

doi:10.1039/C9JA00090A

Index