Package ‘clubpro’

June 7, 2024

Type Package

Title Classification Using Binary Procrustes Rotation

Version 0.6.2

Description Implements a classification method described by Grice (2011, ISBN:978-0-12-385194-9) using binary procrustes rotation; a simplified version of procrustes rotation.

License GPL (>= 3)

URL https://timbeechey.github.io/clubpro/

BugReports https://github.com/timbeechey/clubpro/issues

Encoding UTF-8

LazyData true

RoxygenNote 7.3.1

LinkingTo Rcpp, RcppArmadillo, RcppProgress

Imports Rcpp, lattice, stats, graphics

Suggests knitr, rmarkdown, tinytest

SystemRequirements C++17

Depends R (>= 2.10)

VignetteBuilder knitr

NeedsCompilation yes

Author Timothy Beechey [aut, cre] (<https://orcid.org/0000-0001-8858-946X>)

Maintainer Timothy Beechey <tim.beechey@proton.me>

Repository CRAN

Date/Publication 2024-06-07 15:50:02 UTC
Contents

accuracy ................................................................. 2
as.data.frame.clubprocsi ........................................... 3
as.data.frame.clubprorand ......................................... 4
caffeine ............................................................ 4
club ................................................................. 5
compare ........................................................... 6
csi ................................................................. 6
cval ............................................................... 7
individual_results .................................................. 8
jellyfish ............................................................ 8
median_csi .......................................................... 9
n_ambiguous ......................................................... 9
n_correct .......................................................... 10
n_incorrect ......................................................... 11
pcc ................................................................. 11
pcc_replicates ....................................................... 12
plot.clubproaccuracy ............................................... 13
plot.clubprocomparison .......................................... 13
plot.clubprocsi ..................................................... 14
plot.clubprofit ...................................................... 15
plot.clubpropredictions .......................................... 16
plot.clubprorand ................................................... 16
plot.clubprothreshold ............................................ 17
predict.clubprofit .................................................. 18
summary.clubprocomparison .................................. 18
summary.clubprofit .............................................. 19
threshold .......................................................... 20

Index ................................................................. 21

accuracy ................................................................. 2

Classification accuracy for each observation.

Description

Classification accuracy for each observation.

Usage

accuracy(m)

Arguments

m an object of class "clubprofit" produced by club()
Details

Returns a character vector containing a string corresponding to each observation indicating whether classification of that observation was "correct", "incorrect", or "ambiguous".

Value

a table

Examples

mod <- club(rate ~ dose, data = caffeine)
accuracy(mod)

as.data.frame.clubprocsi

Convert the output of csi() to a data.frame.

Description

Convert the output of csi() to a data.frame.

Usage

## S3 method for class 'clubprocsi'
as.data.frame(x, row.names = NULL, optional = FALSE, ...)

Arguments

x an object of class "clubprocsi"
row.names ignored
optional ignored
... ignored

Details

This function is useful to format pcc replicates data for plotting.

Examples

mod <- club(rate ~ dose, data = caffeine)
z <- csi(mod)
as.data.frame(z)
as.data.frame.clubprorand

Convert the output of pcc_replicates() to a data.frame.

Description

Convert the output of pcc_replicates() to a data.frame.

Usage

## S3 method for class 'clubprorand'
as.data.frame(x, row.names = NULL, optional = FALSE, ...)

Arguments

x an object of class "clubprorand"
row.names ignored
optional ignored
... ignored

Details

This function is useful to format pcc replicates data for plotting.

Examples

mod <- club(rate ~ dose, data = caffeine)
z <- pcc_replicates(mod)
as.data.frame(z)

caffeine Caffeine data

Description

Effect of three different doses of caffeine on finger tapping rate.

Usage

caffeine

Format

A data frame with 30 rows and 2 columns:

dose dose of caffeine in mg
rate finger taps per minute
Source

Description
club() is used to classify observations using binary procrustes rotation.

Usage
club(
  f,            # a formula.
  data,         # a data.frame.
  imprecision,  # a number indicating the margin of imprecision allowed in classification.
  nreps,        # the number of replicates to use in the randomisation test.
  normalise_cols, # a boolean indicating whether to normalise matrix columns.
  reorder_obs,  # a string indicating the method for reordering observations to calculate c-values.
  display_progress  # a boolean indicating whether a progress bar should be displayed.
)

Arguments
f            # a formula.
data         # a data.frame.
imprecision  # a number indicating the margin of imprecision allowed in classification.
nreps        # the number of replicates to use in the randomisation test.
normalise_cols # a boolean indicating whether to normalise matrix columns.
reorder_obs  # a string indicating the method for reordering observations to calculate c-values.
display_progress  # a boolean indicating whether a progress bar should be displayed.

Value
an object of class "clubprofit" is a list containing the following components:

- prediction  a character vector of predicted classifications.
- accuracy    a character vector indicating whether each classification is "correct", "incorrect", or "ambiguous".
- pcc          the percentage of correct classifications.
- cval         the chance of randomly reordered data producing a PCC >= the observed PCC.
- pcc_replicates a vector of PCCs generated from randomly reordered data used to calculate cval.
- call         the matched call.
Examples

```r
mod <- club(rate ~ dose, data = caffeine)
```

---

**compare**  
*Compare models.*

**Description**  
Compare models.

**Usage**  
```r
compare(m1, m2)
```

**Arguments**

- `m1`: an object of class "clubprofit" produced by `club()`
- `m2`: an object of class "clubprofit" produced by `club()`

**Details**  
Compare the PCC of two clubprofit models and compute the chance-value of the difference.

**Value**  
an object of type "clubprocomparison"

**Examples**

```r
m1 <- club(width ~ location, jellyfish)
m2 <- club(length ~ location, jellyfish)
compare(m1, m2)
```

---

**csi**  
*Classification strength indices.*

**Description**  
Classification strength indices.

**Usage**  
```r
csi(m)
```

**Arguments**

- `m`: an object of class "clubprofit" produced by `club()`
cval

Details

Returns a vector containing the classification strength index for each observation.

Value

a numeric vector.

Examples

```r
mod <- club(rate ~ dose, data = caffeine)
csi(mod)
```

---

cval

|Chance value.|

Description

Chance value.

Usage

cval(m)

Arguments

m an object of class "clubprofit" produced by club()

Details

Compute the chance that randomly reordered data results in a percentage of correctly classified observations at least as high as the observed data.

Value

a numeric value.

Examples

```r
mod <- club(rate ~ dose, data = caffeine)
cval(mod)
```
### individual_results

*Individual level classification results.*

#### Description

Individual level classification results.

#### Usage

`individual_results(m, digits)`

#### Arguments

- `m` an object of class "clubprofit" produced by `club()`
- `digits` an integer

#### Details

Returns a data.frame containing predicted classifications and classification accuracy for each individual observation.

#### Value

a data.frame containing columns of predictions and prediction accuracy

#### Examples

```r
mod <- club(rate ~ dose, data = caffeine)
individual_results(mod)
```

### jellyfish

*Jellyfish dimension data*

#### Description

Sizes of jellyfish from two locations in the Hawkesbury River, New South Wales, Australia.

#### Usage

`jellyfish`

#### Format

A data frame with 46 rows and 3 columns:

- **location** location where jellyfish was caught
- **width** jellyfish width in mm
- **length** jellyfish length in mm
**median_csi**

**Source**

**Description**
Median classification strength index.

**Usage**

\[
\text{median}_\text{csi}(m)
\]

**Arguments**

- \( m \) an object of class "clubprofit" produced by \( \text{club()} \)

**Details**
Returns the median classification strength index.

**Value**
a numeric vector.

**Examples**

```r
mod <- \text{club}(\text{rate} - \text{dose}, \text{data} = \text{caffeine})
\text{median}_\text{csi}(mod)
```

**n_ambiguous**

**Description**
Number of ambiguous classifications.

**Usage**

\[
\text{n}_\text{ambiguous}(m)
\]

**Arguments**

- \( m \) an object of class "clubprofit" produced by \( \text{club()} \)
n_correct

Details
Returns the number of observations which were classified ambiguously by the model.

Value
an integer.

Examples
mod <- club(rate ~ dose, data = caffeine)
n_ambiguous(mod)

| n_correct | Number of correct classifications.

Description
Number of correct classifications.

Usage
n_correct(m)

Arguments
m an object of class "clubprofit" produced by club()

Details
Returns the number of observations which were classified correctly by the model.

Value
an integer.

Examples
mod <- club(rate ~ dose, data = caffeine)
n_correct(mod)
**n_incorrect**  
*Number of incorrect classifications.*

**Description**  
Number of incorrect classifications.

**Usage**  
```
n_incorrect(m)
```

**Arguments**  
- `m` an object of class "clubprofit" produced by `club()`

**Details**  
Returns the number of observations which were classified incorrectly by the model.

**Value**  
an integer.

**Examples**  
```r
mod <- club(rate ~ dose, data = caffeine)
n_incorrect(mod)
```

---

**pcc**  
*Percentage of correct classifications.*

**Description**  
Percentage of correct classifications.

**Usage**  
```
pcc(m)
```

**Arguments**  
- `m` an object of class "clubprofit" produced by `club()`

**Details**  
Returns the percentage of correctly classified observations.
Value

a numeric value.

Examples

```r
mod <- club(rate ~ dose, data = caffeine)
pcc(mod)
```

---

### pcc_replicates

**PCC replicates.**

**Description**

PCC replicates.

**Usage**

```r
pcc_replicates(m)
```

**Arguments**

- `m` an object of class "clubprofit" produced by `club()`

**Details**

Returns an object containing a vector of PCC replicates used to calculate the chance-value.

**Value**

an object of class clubprorand.

**Examples**

```r
mod <- club(rate ~ dose, data = caffeine)
head(pcc_replicates(mod))
```
plot.clubproaccuracy  Plot accuracy.

Description
Plot accuracy.

Usage
## S3 method for class 'clubproaccuracy'
plot(x, ...)

Arguments
x  an object of class "clubproaccuracy"
...  ignored

Details
Produces a mosaic plot of prediction accuracy by category

Value
called for side-effects only

Examples
mod <- club(rate ~ dose, data = caffeine)
z <- accuracy(mod)
plot(z)

plot.clubprocomparison  Plot model comparison.

Description
Plot model comparison.

Usage
## S3 method for class 'clubprocomparison'
plot(x, ...)

plot.clubproaccuracy  Plot accuracy.

plot.clubprocomparison  Plot model comparison.
Arguments

x an object of class "clubprocomparison".
...
ignored

Details

Plot a distribution of PCCs computed from randomly reordered data used to calculate the chance-value for a model comparison.

Value

no return value, called for side effects only.

Examples

m1 <- club(width ~ location, jellyfish)
m2 <- club(length ~ location, jellyfish)
z <- compare(m1, m2)
plot(z)
plot.clubprofit

Description

Plot classification accuracy.

Usage

## S3 method for class 'clubprofit'
plot(x, ...)

Arguments

x an object of class "clubprofit" produced by club()

... ignored

Details

Produces bar plot showing counts of individuals against observed values within each target grouping. Fill colours indicate whether each individual was classified correctly, incorrectly or ambiguously.

Value

called for side-effects only

Examples

mod <- club(rate ~ dose, data = caffeine)
z <- csi(mod)
plot(z)
Description

Plot predictions.

Usage

## S3 method for class 'clubpropredictions'
plot(x, ...)

Arguments

x       an object of class "clubpropredictions"
...
ignored

Details

Produces a mosaic plot of observed versus predicted categories.

Value

called for side-effects only

Examples

mod <- club(rate ~ dose, data = caffeine)
z <- predict(mod)
plot(z)

Description

Plot PCC replicates.

Usage

## S3 method for class 'clubprorand'
plot(x, ...)

plot.clubprorand
Arguments

x an object of class "clubprofit" produced by club()

Details

Plot the distribution of PCCs computed from randomly reordered data used to calculate the chance-value.

Value

no return value, called for side effects only.

Examples

mod <- club(rate ~ dose, data = caffeine)
plot(pcc_replicates(mod))

plot.clubprothreshold

Plot PCC as a function of binary category boundary location.

Description

Plot PCC as a function of binary category boundary location.

Usage

## S3 method for class 'clubprothreshold'
plot(x, ...)

Arguments

x an object of class "clubprothreshold"

Details

Produces an xyplot showing the PCC returned for each possible category boundary.

Value

called for side-effects only

Examples

mod <- club(rate ~ dose, data = caffeine)
z <- threshold(mod)
plot(z)
predict.clubprofit  Predicted category for each observation.

Description
Predicted category for each observation.

Usage

## S3 method for class 'clubprofit'
predict(object, ...)

Arguments

object  an object of class "clubprofit" produced by club()
...

Details
Returns a character vector containing the name of the predicted category for each observed value.

Value

a table

Examples

mod <- club(rate ~ dose, data = caffeine)
predict(mod)

---------

summary.clubprocomparison
Generate a summary of a comparison of clubprofit models.

Description
Generate a summary of a comparison of clubprofit models.

Usage

## S3 method for class 'clubprocomparison'
summary(object, ...)

Arguments

object  an object of class "clubprocomparison".
...

ignored
Value

No return value, called for side effects.

Examples

```r
m1 <- club(width ~ location, jellyfish)
m2 <- club(length ~ location, jellyfish)
z <- compare(m1, m2)
summary(z)
```

Description

Generate a summary of results from a fitted clubpro model.

Usage

```r
## S3 method for class 'clubprofit'
summary(object, ...)
```

Arguments

- `object` an object of class "clubprofit".
- `...` ignored

Value

No return value, called for side effects.

Examples

```r
mod <- club(rate ~ dose, data = caffeine)
summary(mod)
```
threshold

Classification strength indices.

Description
Classification strength indices.

Usage
threshold(m)

Arguments
m an object of class "clubprofit" produced by club()

Details
Returns a vector containing the classification strength index for each observation.

Value
an object of class clubprofit

Examples
mod <- club(width ~ location, data = jellyfish)
threshold(mod)
Index

* datasets
  caffeine, 4
  jellyfish, 8
accuracy, 2
as.data.frame.clubprocsi, 3
as.data.frame.clubprorand, 4
caffeine, 4
club, 5
compare, 6
csi, 6
cval, 7
individual_results, 8
jellyfish, 8
median_csi, 9
n_ambiguous, 9
n_correct, 10
n_incorrect, 11
pcc, 11
pcc_replicates, 12
plot.clubproaccuracy, 13
plot.clubprocomparison, 13
plot.clubprocsi, 14
plot.clubprofit, 15
plot.clubpropredictions, 16
plot.clubprorand, 16
plot.clubprothreshold, 17
predict.clubprofit, 18
summary.clubprocomparison, 18
summary.clubprofit, 19
threshold, 20