Package ‘cutoff’

December 20, 2019

Title  Seek the Significant Cutoff Value

Version  1.3

Description  Seek the significant cutoff value for a continuous variable, which will be transformed into a classification, for linear regression, logistic regression, logrank analysis and cox regression. First of all, all combinations will be gotten by combn() function. Then n.per argument, abbreviated of total number percentage, will be used to remove the combination of smaller data group. In logistic, Cox regression and logrank analysis, we will also use p.per argument, patient percentage, to filter the lower proportion of patients in each group. Finally, p value in regression results will be used to get the significant combinations and output relevant parameters. In this package, there is no limit to the number of cutoff points, which can be 1, 2, 3 or more. Still, we provide 2 methods, typical Bonferroni and Duglas G (1994) <doi: 10.1093/jnci/86.11.829>, to adjust the p value, Missing values will be deleted by na.omit() function before analysis.

License  GPL-3

Encoding  UTF-8

LazyData  true

RoxygenNote  6.1.1

Imports  survival, set, do, ROCit

URL  https://github.com/yikeshu0611/cutoff

BugReports  https://github.com/yikeshu0611/cutoff/issues

NeedsCompilation  no

Author  Jing Zhang [aut, cre],
        Zhi Jin [aut]

Maintainer  Jing Zhang <zj391120@163.com>

Repository  CRAN

Date/Publication  2019-12-20 10:10:05 UTC
R topics documented:

- cox .................................................. 2
- cutit .................................................. 3
- judge_123 .......................................... 4
- judge_321 .......................................... 4
- linear .............................................. 5
- logit ............................................... 6
- logrank ............................................ 7
- roc .................................................. 8
- x_ab .................................................. 9

Index 10

---

cox

Significant Cutoff Value for Cox Regression

Description

Significant Cutoff Value for Cox Regression

Usage

```r
cox(data, time, y, x, cut.numb, n.per, y.per, p.cut = 0.05,
    strict = TRUE, include = "low", round = 2, adjust = 1)
```

Arguments

- **data**
- **time**
- **y**
- **x**
- **cut.numb**
- **n.per**
- **y.per**
- **p.cut**
- **strict**
- **include**
- **round**
- **adjust**

- **data**: data
- **time**: name for time variable
- **y**: name for y, must be coded as 1 and 0. The outcome must be 1
- **x**: name for x
- **cut.numb**: number of cutoff points
- **n.per**: the least percentage of the smaller group comprised in all patients
- **y.per**: the least percentage of the smaller outcome patients comprised in each group
- **p.cut**: cutoff of p value, default is 0.05
- **strict**: logical. TRUE means significant differences for each group combination were considered. FALSE means considering for any combination
- **include**: direction of cutoff point. Any left letter of lower or upper
- **round**: digital. Default is 2
- **adjust**: numeric value, adjust method for p value. 1, defaulted, represents Bonferroni. 2 represent formula given by Douglas G in 1994
cutit

Value

A dataframe contains cutoff points value, subject numbers in each group, dumb variable, beta of regression and p value.

Examples

cox(data=mtcars,
   time = 'disp', y='am', x='wt',
   cut.numb=2,
   n.per=0.25,
   y.per=0.10)

cox(data=mtcars,
   time = 'disp', y='am', x='wt',
   cut.numb=2,
   n.per=0.25,
   y.per=0.10,
   p.cut=0.05,
   strict=TRUE,
   include='low',
   round=2)

cutit(mtcars$disp,c(150,190))

cutit(mtcars$disp,c(150,190),labels = TRUE)

---

cutit  
Cut Continuous Vector to Classification

Description

Cut Continuous Vector to Classification

Usage

cutit(x, cut_points, include = "low", labels = FALSE)

Arguments

x numeric vector
cut_points cutting points value
include The direction of cutoff point. Any left letter of lower or upper
labels logical. False is defaulted. TRUE means set range as factor.

Value

numeric vector or factor

Examples

cutit(mtcars$disp,c(150,190))
cutit(mtcars$disp,c(150,190),labels = TRUE)
## judge_123

### Description

Whether the Data Is Arranged from Small to Large

### Usage

```r
judge_123(x)
```

### Arguments

- `x`: numeric vector

### Value

logical

### Examples

```r
c(1,2,3,4,5)
judge_123(c(1,3,2))
```

## judge_321

### Description

Whether the Data Is Arranged from Large to Small

### Usage

```r
judge_321(x)
```

### Arguments

- `x`: numeric vector

### Value

logical

### Examples

```r
c(5,4,3,2,1)
judge_321(c(3,1,2))
```
**Significant Cutoff Value for Linear Regression**

**Description**

Significant Cutoff Value for Linear Regression

**Usage**

```r
linear(data, y, x, cut.numb, n.per, p.cut = 0.05, strict = TRUE,
   include = "low", round = 2, adjust = 1)
```

**Arguments**

- `data`: data
- `y`: name for y
- `x`: name for x
- `cut.numb`: number of cutoff points
- `n.per`: the least percentage of the smaller group comprised in all patients
- `p.cut`: cutoff of p value, default is 0.05
- `strict`: logical. TRUE means significant differences for each group combination were considered. FALSE means considering for any combination
- `include`: direction of cutoff point. Any left letter of lower or upper
- `round`: digital. Default is 2
- `adjust`: numeric value, adjust method for p value. 1, defaulted, represents Bonferroni. 2 represent formula given by Douglas G in 1994

**Value**

a dataframe contains cutoff points value, subject numbers in each group, dummy variable, beta of regression and p value.

**Examples**

```r
linear(data=mtcars,y="qsec",x="disp",
   cut.numb=2,
   n.per=0.25)

linear(data=mtcars,y="qsec",x="disp",
   cut.numb=2,
   n.per=0.25,
   p.cut=0.05,
   strict=TRUE,
   include='low',
   round=2)
```
logit

Significant Cutoff Value for Logistic Regression

Description

Significant Cutoff Value for Logistic Regression

Usage

logit(data, y, x, cut.numb, n.per, y.per, p.cut = 0.05, strict = TRUE,
      include = "low", round = 2, adjust = 1)

Arguments

data data

y name for y, must be coded as 1 and 0. The outcome must be 1

x name for x

cut.numb number of cutoff points

n.per the least percentage of the smaller group comprised in all patients

y.per the least percentage of the smaller outcome patients comprised in each group

p.cut cutoff of p value, default is 0.05

strict logical. TRUE means significant differences for each group combination were
considered. FALSE means considering for any combination

include direction of cutoff point. Any left letter of lower or upper

round digital. Default is 2

adjust numeric value, adjust method for p value. 1, defaulted, represents Bonferroni.
2 represent formula given by Douglas G in 1994

Value

a dataframe contains cutoff points value, subject numbers in each group, dummy variable, or of
regression and p value.
Examples

logit(data=mtcars,
    y='am',
    x='disp',
    cut.numb=1,
    n.per=0.25,
    y.per=0.25)
logit(data=mtcars,
    y='am',
    x='disp',
    cut.numb=1,
    n.per=0.25,
    y.per=0.20,
    p.cut=0.05,
    strict=TRUE,
    include='low',
    round=2)

logrank

Significant Cutoff Value for Logrank Analysis

Description

Significant Cutoff Value for Logrank Analysis

Usage

logrank(data, time, y, x, cut.numb, n.per, y.per, p.cut = 0.05,
        strict = TRUE, include = "low", round = 2, adjust = 1)

Arguments

data  data

time  name for time variable

y  name for y, must be coded as 1 and 0. The outcome must be 1

x  name for x

cut.numb  number of cutoff points

n.per  the least percentage of the smaller group comprised in all patients

y.per  the least percentage of the smaller outcome patients comprised in each group

p.cut  cutoff of p value, default is 0.05

strict  logical. TRUE means significant differences for each group combination were considered. FALSE means considering for any combination

include  direction of cutoff point. Any left letter of lower or upper

round  digital. Default is 2

adjust  numeric value, adjust method for p value. 1, defaulted, represents Bonferroni. 2 represent formula given by Douglas G in 1994
Value

a dataframe contains cutoff points value, subject numbers in each group, dumb variable, beta of regression and p value.

Examples

```r
data=mtcars, 
  time = 'disp', y='am', x='wt',
  cut.numb=2,
  n.per=0.25,
  y.per=0.10)
```

```r
logrank(data=mtcars,
  time = 'disp', y='am', x='wt',
  cut.numb=2,
  n.per=0.25,
  y.per=0.10,
  p.cut=0.05,
  strict=TRUE,
  include='low',
  round=2)
```

---

**roc**  
*To Get the Best Cutoff Value for ROC Curve*

**Description**

Youden index is used for seeking the best cutoff value for ROC Curve.

**Usage**

```r
roc(score, class)
```

**Arguments**

- **score**: continuous value
- **class**: binary value, 0 and 1

**Value**

If the auc of a variate is lower than 0.5, we treat it as negative classification and return information about the negative prediction. Otherwise, The variate will be treated as positive one.

**Examples**

```r
roc(score = mtcars$qsec, class = mtcars$am)
roc(score = mtcars$drat, class = mtcars$am)
```
\textit{x_ab}

\begin{itemize}
  \item \textit{Description} \\
  \hspace{1em} Return \textit{x} Between \textit{a} and \textit{b}
  \item \textit{Usage} \\
  \hspace{1em} \texttt{x_ab(x, a, b, include = "l")}
  \item \textit{Arguments} \\
  \hspace{1em} \textit{x} \hspace{1em} \text{numeric vector} \\
  \hspace{1em} \textit{a} \hspace{1em} \text{one number} \\
  \hspace{1em} \textit{b} \hspace{1em} \text{one number} \\
  \hspace{1em} \textit{include} \hspace{1em} \text{The direction of \textit{a} and \textit{b}. Any left letter of lower or upper} \\
  \item \textit{Value} \\
  \hspace{1em} \text{values of \textit{x} between \textit{a} and \textit{b}}
  \item \textit{Examples} \\
  \hspace{1em} \texttt{x_ab(mtcars$disp,150,190)}
\end{itemize}
Index

cox, 2
cutit, 3

judge_123, 4
judge_321, 4

linear, 5
logit, 6
logrank, 7

roc, 8
x_ab, 9