Package 'ROCket'

May 9, 2025

Type Package

Title Simple and Fast ROC Curves

Version 1.0.3

Description A set of functions for receiver operating characteristic (ROC) curve estimation and area under the curve (AUC) calculation.
All functions are designed to work with aggregated data;
nevertheless, they can also handle raw samples.
In 'ROCket', we distinguish two types of ROC curve representations:
1) parametric curves - the true positive rate (TPR) and the false positive rate (FPR) are functions of a parameter (the score),
2) functions - TPR is a function of FPR.
There are several ROC curve estimation methods available. An introduction to the mathematical background of the implemented methods (and much more) can be found in de Zea Bermudez, Gonçalves, Oliveira & Subtil (2014) and Cai & Pepe (2004).

License GPL-3

Encoding UTF-8

URL https://github.com/da-zar/ROCket

BugReports https://github.com/da-zar/ROCket/issues

Imports data.table (>= 1.13.0)

Suggests testthat

RoxygenNote 7.2.0

Collate 'ROCket.R' 'generics.R' 'auc.R' 'mwu.R' 'rkt_ecdf.R' 'rkt_prep.R' 'rkt_roc.R' 'roc_methods.R' 'utils.R' 'zzz.R'

NeedsCompilation no

Author Daniel Lazar [aut, cre]

Maintainer Daniel Lazar <da-zar@gmx.net>

Repository CRAN

Date/Publication 2025-05-09 16:20:05 UTC

Contents

---

auc *Calculate the AUC*

Description

Calculate the AUC

Usage

```
auc(x, ...)

## S3 method for class '`function`'
auc(x, ...)

## S3 method for class 'curve'
auc(x, lower, upper, n = 10000, ...)

## S3 method for class 'rkt_roc'
auc(x, exact = TRUE, ...)
```

Arguments

x	An R object.
...	Further parameters.
lower, upper	The limits of integration.
n	The number of integration points.
exact	Logical. If the exact formula should be used for calculating the AUC instead of numerical approximation.

Value

The area under the curve as a numeric value.

mwu.test *Mann-Whitney U test*

Description

Performs the Mann-Whitney U test with a normal approximation.

Usage

```
mwu.test(prep, alternative = c("two.sided", "less", "greater"), correct = TRUE)
```

Arguments

prep	A rkt_prep object.
alternative	The alternative hypothesis type. One of: "two.sided", "less", "greater".
correct	Logical. Whether to apply continuity correction.

Value

A list of the class "htest".

rkt_ecdf *Empirical estimate of the CDF*

Description

Calculate an empirical cumulative distribution function based on a sample x and optionally a vector w of weights.

Usage

```
rkt_ecdf(x, w)

## S3 method for class 'rkt_ecdf'
print(x, ...)

## S3 method for class 'rkt_ecdf'
mean(x, ...)

## S3 method for class 'rkt_ecdf'
variance(x, ...)

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

Arguments

x	Numeric vector containing the sample. Alternatively, if w is supplied, distinct values within the sample. For S3 methods, a function of class rkt_ecdf.
w	Optional. Numeric vector containing the weights of each value in x.
...	Further parameters.

Details

The weights vector w can contain the counts of each distinct value in x, this is the most natural use case. In general the weights are describing the jumps of the final ecdf. Normalization is handled internally.

If x contains duplicates, corresponding values in w will be summed up. Only positive weights are allowed. Elements in x with non-positive weights will be ignored.

Value

A function of class rkt_ecdf.

Examples

```
require(ROCket)

plot(rkt_ecdf(rnorm(100)))
plot(rkt_ecdf(c(0, 1)))
plot(rkt_ecdf(c(0, 1), c(1, 10)))
```

---

rkt_prep *ROC points*

Description

Calculate the ROC points for all meaningful cutoff values based on predicted scores.

Usage

```
rkt_prep(scores, positives, negatives = totals - positives, totals = 1)

## S3 method for class 'rkt_prep'
print(x, ...)

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

Arguments

scores	Numeric vector containing the predicted scores.
positives	Numeric vector of the same length as scores. The number of positive entities associated with each score. If data is not aggregated, a vector of 0's and 1's.
negatives	Similar to positives. Defaults to totals - positives.
totals	How many times each score was predicted. Defaults to 1 (assuming data is not aggregated). If any value in positives is greater than 1 (aggregated data), totals must be a vector. Not needed if negatives is supplied.
x	An environment of class rkt_prep for S3 methods.
...	Further parameters.

Details

In a situation where many of the predicted scores have the same value it might be easier and faster to use aggregated data.

Value

An environment of class rkt_prep.

Examples

```
require(ROCket)

plot(rkt_prep(1:4, c(0, 1, 0, 1)))
plot(rkt_prep(1:4, c(0, 1000, 0, 1000), totals = 1000))
plot(rkt_prep(1:4, c(100, 200, 300, 400), totals = c(1000, 800, 600, 400)))
```

---

rkt_roc *Empirical estimate of the ROC*

Description

Calculate the empirical estimate of the ROC from raw sample or aggregated data.

Usage

```
rkt_roc(prep, method = 1)

## S3 method for class 'rkt_roc'
print(x, ...)

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

Arguments

prep	A rkt_prep object.
method	A number specifying the type of ROC estimate. Possible values can be viewed with show_methods().
x	An object of class rkt_roc.
...	Further parameters passed to plot and lines

Value

An object of class rkt_roc, i.e. a function or a list of two functions (for method = 1).

Examples

```
require(ROCket)

scores <- c(1, 2, 3, 4)
positives <- c(0, 1, 0, 1)
prep <- rkt_prep(scores, positives)

roc1 <- rkt_roc(prep, method = 1)
roc2 <- rkt_roc(prep, method = 2)
roc3 <- rkt_roc(prep, method = 3)

plot(roc1)
plot(roc2)
plot(roc3)
```

---

show_methods *Available ROC estimation methods*

Description

Show the implemented ROC estimation methods.

Usage

```
show_methods()
```

Value

A data.table containing the number and a short description of each implemented method.

`variance` *Sample Variance*

Description

Sample Variance

Usage

```
variance(x, ...)

## Default S3 method:
variance(x, ...)
```

Arguments

`x`	An R object.
`...`	Further parameters.

Value

The (biased) sample variance as a numeric value.

See Also

`variance.rkt_ecdf`, `var`

Index