

Package ‘varoc’

January 16, 2024

Type Package

Title Value Added Receiver Operating Characteristics Curve

Version 0.2.0

Date 2024-01-16

Description A continuous version of the receiver operating characteristics (ROC) curve to visualize and assess the classification and continuity performances of biomarkers, diagnostic tests, or risk prediction models.

License GPL (>= 2)

Depends R (>= 4.2.0), pROC, corrplot, grDevices, graphics, stats, utils

NeedsCompilation no

Author Yunro Chung [aut, cre] (<<https://orcid.org/0000-0001-9125-9277>>)

Maintainer Yunro Chung <yunro.chung@asu.edu>

Repository CRAN

Date/Publication 2024-01-16 15:10:02 UTC

R topics documented:

amd	1
jdp	4
varoc	6

Index	8
--------------	----------

amd	<i>AMD: above mean difference</i>
-----	-----------------------------------

Description

Summary measures to evaluate the continuity performance of biomarkers, diagnostic tests, or risk prediction models.

Usage

```
amd(y, x, fpf=0.3, pval="no", alternative="greater", B=2000, conf.level=0.95)
```

Arguments

y	binary output, where y=1 if disease (or case) and y=0 if non-disease (or control).
x	continuous score, e.g. biomarker, diagnostic test, risk score.
fpf	false positive fraction at which above mean difference is calculated.
pval	"yes" for bootstrap p-value and bootstrap confidence interval.
alternative	alternative hypothesis: "greater" (default), "less", "two.sided".
B	number of bootstrap samples.
conf.level	confidence level of bootstrap confidence interval.

Details

The amd function summarizes a continuity performance of x at each cutoff c as: i) above mean difference (AMD) and ii) integrated AMD (IAMD). For i), $AMD(c)$ is true positive mean(TPM)(c) minus false positive mean(FPM)(c), where $TPM(c)$ is $E(x>c|y=1)$ and $FPM(c)$ is $E(x>c|y=0)$. For ii), IAMD is a global measure of evaluating continuity performance of x over all thresholds.

These measures can be viewed as continuous versions of ROC curve-based measures. Specifically, $TPM(c)$ and $FPM(c)$ are continuous versions of true positive fraction(TPF)(c) (or sensitivity(c)) and false positive fraction(FPF)(c) (or one minus specificity(c)), where $TPF(c)=P(x>c|y=1)$ and $FPF(c)=P(x>c|y=0)$. The useful (or useless) x has $TPF(c)-FPF(c)>0$ and $AMD(c)>0$ (or $TPF(c)-FPF(c)=0$ and $AMD(c)=0$). Similarly, useful (or useless) x has area under the ROC curve(AUC) >0.5 and $IAMD(c)>0$ (or $AUC=0.5$ and $IAMD(c)=0$).

The bootstrap p-value and confidence interval are computed under the null hypothesis: $AMD(c)\leq 0$ or $IAMD(c)\leq 0$, when $pval="yes"$ and $alterantive="greater"$.

The threshold c is determined by setting an acceptable fpf, i.e. $FPF(c)=fpf$. Thus, i) is interpreted as AMD at a FPF of fpf, ii) is interpreted as IAMD (or average AMD) at a FPF range of 0 and 1. The varoc and jdp functions visualize them.

Value

df	data frame with y and x.
fpf	false positive fraction at which above mean difference is calculated.
res	data frame with tpf, fpf, tpm, fpm, amd, lcl, ucl, zAMD, pvalue at each threshold (th), where lcl (lower confidence limit), ucl (upper confidence limit), zAMD (test statistics) and pvalue (one-sided p-value) are for amd.
amd	data frame with tpf, fpf, tpm, fpm, amd, lcl, ucl, zAMD, pvalue at $FPF(th)=fpf$, where lcl, ucl, zAMD and pvalue are for amd. Here, empirical estimator of FPF is used, and amdfpf$ could be different from the fpf value in the argument.
i amd	data frame with auc, itpm, iftm, iamd, lcl, ucl, zIAMD, pvalue, where lcl, ucl, zIAMD and pvalue are for iamd.

Author(s)

Yunro Chung [aut, cre]

References

Danielle Brister and Yunro Chung, Value added receiver operating characteristics curve (in-progress)

Examples

```
set.seed(3)

n1=50
n0=50

#1. marker 1 (useless biomaker)
y1=c(rep(1,n1),rep(0,n0))
x1=abs(c(rnorm(n1,0,1),rnorm(n0,0,1)))

#1.1. amd
fit1=amd(y=y1,x=x1,fpf=0.3)
print(fit1)

#1.2. varoc
varoc(fit1)

#1.3. jdp
jdp(fit1)

#2. marker 2 (useful biomarker)
y2=y1
x2=abs(c(rnorm(n1,1,1),rnorm(n0,0,1)))

#2.1. amd
fit2=amd(y=y2,x=x2,fpf=0.3)

#2.2. varoc for marker 1 vs marker 2
mzr.min=min(c(fit1$res$amd,fit2$res$amd))
mzr.max=max(c(fit1$res$amd,fit2$res$amd))

varoc(fit1,mzr="zAMD",mzr.min=mzr.min,mzr.max=mzr.max)
varoc(fit2,mzr="zAMD",mzr.min=mzr.min,mzr.max=mzr.max)

#2.3. varoc for marker 1 vs marker 2
min=min(c(x1,x2))
max=max(c(x1,x2))
jdp(fit1,min=min,max=max)
jdp(fit2,min=min,max=max)
```

jdp

*JDP: jittered ot plot***Description**

Jittered dot plot to visualize classification and continuity performances of biomarkers.

Usage

```
jdp(fit,
    min=NULL, max=NULL, eps=0.2, seed=1,
    main="JDP", ylab="x", xlab=c("y=0", "y=1"),
    col=c("blue", "red", "gray", "gray"),
    legend="top", lwd=1, lty=3,
    cex.main=1, cex.pt=1.5, cex.lab=1, cex.axis=1, cex.legend=1, digits=2)
```

Arguments

<code>fit</code>	fitted results from the <code>amd()</code> function in the <code>varoc</code> R package.
<code>min</code>	minimum value of y-axis.
<code>max</code>	maximum value of y-axis.
<code>eps</code>	jittered range of x-axis.
<code>seed</code>	seed number for jittering x-axis.
<code>main</code>	title for the plot
<code>ylab</code>	title for the y axis.
<code>xlab</code>	title for the x axis.
<code>col</code>	colors to true positive, false negative, false positive, false negative.
<code>legend</code>	legend location, "bottomright", "bottom", "bottomleft", "left", "topleft", "top", "topright", "right" and "center".
<code>lwd</code>	line width.
<code>lty</code>	line type.
<code>cex.main</code>	main size.
<code>cex.pt</code>	point size.
<code>cex.lab</code>	label size.
<code>cex.axis</code>	axis size.
<code>cex.legend</code>	legend size.
<code>digits</code>	number of decimals.

Details

The `jdp` function plots biomarker (or x) by jittered x-axis ($y=0$ vs $y=1$) and visualizes its classification and continuity metrics. The vertical dotted line is the threshold corresponding to $FPF=fpf$, and the two horizontal lines at $y=0$ and $y=1$ are false positive and true positive means, respectively. See the `amd` function for more details.

Value

No return value, called for side effects.

Author(s)

Yunro Chung [aut, cre]

References

Danielle Brister and Yunro Chung, Value added receiver operating characteristics curve (in-progress)

Examples

```
set.seed(1)

n1=50
n0=50

#1. marker 1 (useless biomaker)
y1=c(rep(1,n1),rep(0,n0))
x1=abs(c(rnorm(n1,0,1),rnorm(n0,0,1)))

#1.1. amd
fit1=amd(y=y1,x=x1,fpf=0.3)
print(fit1)

#1.2. varoc
varoc(fit1)

#1.3. jdp
jdp(fit1)

#2. marker 2 (useful biomarker)
y2=y1
x2=abs(c(rnorm(n1,1,1),rnorm(n0,0,1)))

#2.1. amd
fit2=amd(y=y2,x=x2,fpf=0.3)

#2.2. varoc for marker 1 vs marker 2
mzr.min=min(c(fit1$res$amd,fit2$res$amd))
mzr.max=max(c(fit1$res$amd,fit2$res$amd))

varoc(fit1,mzr="AMD",mzr.min=mzr.min,mzr.max=mzr.max)
varoc(fit2,mzr="AMD",mzr.min=mzr.min,mzr.max=mzr.max)

#2.3. varoc for marker 1 vs marker 2
min=min(c(x1,x2))
max=max(c(x1,x2))
jdp(fit1,min=min,max=max)
jdp(fit2,min=min,max=max)
```

varoc

*VAROC: value added receiver operating characteristics (ROC) curve***Description**

ROC curve to visualize classification and continuity performances of biomarkers, diagnostic tests, or risk prediction models.

Usage

```
varoc(fit,
      mzr,mzr.min=NULL,mzr.max=NULL,
      main="VAROC",ylab="True positive fraction",xlab="False positive fraction",
      col=c("#9932cc","#87ceeb","#ffe135","#f56642"),
      legend="right",lwd=1,
      cex.main=1,cex.axis=1,cex.lab=1,cex.legend=1,digits=2)
```

Arguments

<code>fit</code>	fitted results from the <code>amd()</code> function in the varoc R package.
<code>mzr</code>	<code>mzr="AMD"</code> (or <code>"zAMD"</code>) if VAROC curve adds AMD (or zAMD, i.e. normalized AMD or test statistics). Note that <code>mzr="zAMD"</code> works only when <code>pval="yes"</code> was used for the <code>amd()</code> function.
<code>mzr.min</code>	minimum value of AMD (or ZAMD) that is displayed on the plot.
<code>mzr.max</code>	maximum value of AMD (or ZAMD) that is displayed on the plot.
<code>main</code>	title for the plot
<code>ylab</code>	title for the y axis.
<code>xlab</code>	title for the x axis.
<code>col</code>	color that separates AMD on the plot. Default: <code>c("#9932cc","#87ceeb","#ffe135","#f56642")</code>
<code>legend</code>	legend location, "bottomright", "bottom", "bottomleft", "left", "topleft", "top", "topright", "right" and "center".
<code>lwd</code>	line width
<code>cex.main</code>	main size.
<code>cex.axis</code>	axis size.
<code>cex.lab</code>	label size.
<code>cex.legend</code>	legend size.
<code>digits</code>	number of decimals.

Details

The varoc function plot true positive fraction(c) (or sensitivity(c)) versus false positive fraction(c) (or one minus specificity(c)) at each threshold `c` colored by above mean difference(c). See the `amd` function for more details.

Value

No return value, called for side effects.

Author(s)

Yunro Chung [aut, cre]

References

Danielle Brister and Yunro Chung, Value added receiver operating characteristics curve (in-progress)

Examples

```
set.seed(1)

n1=50
n0=50

#1. marker 1 (useless biomaker)
y1=c(rep(1,n1),rep(0,n0))
x1=abs(c(rnorm(n1,0,1),rnorm(n0,0,1)))

#1.1. amd
fit1=amd(y=y1,x=x1,fpf=0.3)
print(fit1)

#1.2. varoc
varoc(fit1)

#1.3. jdp
jdp(fit1)

#2. marker 2 (useful biomarker)
y2=y1
x2=abs(c(rnorm(n1,1,1),rnorm(n0,0,1)))

#2.1. amd
fit2=amd(y=y2,x=x2,fpf=0.3)

#2.2. varoc for marker 1 vs marker 2
mzr.min=min(c(fit1$res$amd,fit2$res$amd))
mzr.max=max(c(fit1$res$amd,fit2$res$amd))

varoc(fit1,mzr="AMD",mzr.min=mzr.min,mzr.max=mzr.max)
varoc(fit2,mzr="AMD",mzr.min=mzr.min,mzr.max=mzr.max)

#2.3. varoc for marker 1 vs marker 2
min=min(c(x1,x2))
max=max(c(x1,x2))
jdp(fit1,min=min,max=max)
jdp(fit2,min=min,max=max)
```

Index

amd, 1

jdp, 4

varoc, 6