Package 'rocbc'

July 23, 2025

Type Package

Title Statistical Inference for Box-Cox Based Receiver Operating
Characteristic Curves

Version 3.1.0

Maintainer Benjamin Brewer <tennisbenj@gmail.com>

Description Generation of Box-Cox based ROC curves and several aspects of inferences and hypothe-
sis testing. Can be used when inferences for one biomarker (Ban-
tis LE, Nakas CT, Reiser B. (2018)<doi:10.1002/bimj.201700107>) are of interest or when com-
parisons of two correlated biomarkers (Ban-
tis LE, Nakas CT, Reiser B. (2021)<doi:10.1002/bimj.202000128>) are of interest. Provides in-
ferences and comparisons around the AUC, the Youden index, the sensitivity at a given speci-
ficity level (and vice versa), the optimal operat-
ing point of the ROC curve (in the Youden sense), and the Youden based cutoff.

License GPL-3

Encoding UTF-8

Imports pracma, clinfun, splancs, mvtnorm, formattable, MRMCaov, pROC

Suggests knitr, markdown, matlab

VignetteBuilder knitr

Author Leonidas Bantis [aut],
Benjamin Brewer [cre, ctb],
Christos Nakas [ctb],
Benjamin Reiser [ctb]

NeedsCompilation no

Repository CRAN

Date/Publication 2024-04-22 15:52:41 UTC

Contents

---

checkboxcox   *Tests whether Box-Cox is appropriate for the given dataset (one-marker version)*

Description

This function tests whether the Box-Cox transformation is able to achieve approximate normality
for your data. That is, it will allow the user to know whether it is appropriate to use all the methods
discussed later on in this package.

Usage

```
checkboxcox(marker, D, plots, printShapiro = FALSE)
```

Arguments

marker	A vector of length n that contains the biomarker scores of all individuals.
D	A vector of length n that contains the true disease status of an individual. It is a binary vector containing 0 for the healthy/control individuals and 1 for the diseased individuals.
plots	Valid inputs are "on" and "off". When set to "on", the user gets the histograms of the biomarker for both the healthy and the diseased group before and after the Box-Cox transformation. In addition, all four corresponding qq-plots are provided.
printShapiro	Boolean. When set to TRUE, the results of the Shapiro-Wilk test will be printed to the console. When set to FALSE, the results are suppressed. Default value is FALSE.

Value

res_shapiro	A results table that contains the results of four Shapiro-Wilk tests for normality testing. Two of these refer to normality testing of the healthy and the diseased groups before the Box-Cox transformation, and the remaining two refer to the Box-Cox transformed biomarkers scores for the healthy and the diseased groups. Thus, this testing process produces four p-values. In addition, if the plots are set to 'on', then the output provides (1) the histograms of the biomarker for both the healthy and the diseased groups before and after the Box-Cox transformation, (2) all four corresponding qq-plots, and (3) a plot with the empirical ROC curve overplotted with the Box-Cox based ROC curve for visual comparison purposes.

transformation.parameter	
	The single transformation parameter, lambda, that is applied for both groups simultaneously.
transx	The Box-Cox transformed scores for the healthy.
transy	The Box-Cox transformed scores for the diseased.
pval_x	The p-value of the Shapiro Wilk test of normality for the healthy group (before the Box-Cox transformation).
pval_y	The p-value of the Shapiro Wilk test of normality for the diseased group (before the Box-Cox transformation).
pval_transx	The p-value of the Shapiro Wilk test of normality for the healthy group (after the Box-Cox transformation).
pval_transy	The p-value of the Shapiro Wilk test of normality for the diseased group (after the Box-Cox transformation).
roc	A function of the estimated Box-Cox ROC curve. You can use this to simply request TPR values for given FPR values.

Author(s)

Leonidas Bantis

References

Bantis LE, Nakas CT, Reiser B. (2021). Statistical inference for the difference between two maximized Youden indices obtained from correlated biomarkers. Biometrical Journal, 63(6):1241-1253. https://doi.org/10.1002/bimj.202000128

Bantis LE, Nakas CT, Reiser B. (2018). Construction of confidence intervals for the maximum of the Youden index and the corresponding cutoff point of a continuous biomarker. Biometrical Journal, 61(1):138-156. https://doi.org/10.1002/bimj.201700107

Bantis LE, Feng Z. (2016). Comparison of two correlated ROC curves at a given specificity or sensitivity level. Statistics in Medicine, 35(24):4352-4367. https://doi.org/10.1002/sim.7008

Bantis LE, Nakas CT, Reiser B. (2014). Construction of confidence regions in the ROC space after the estimation of the optimal Youden index-based cut-off point. Biometrics, 70(1):212-223. https://doi.org/10.1111/biom.12107

Box GEP, Cox DR. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society. 26(2):211-252. https://www.jstor.org/stable/2984418

Examples

```
set.seed(123)
x <- rgamma(100, shape=2, rate = 8) # generates biomarker data from a gamma
                                    # distribution for the healthy group.
y <- rgamma(100, shape=2, rate = 4) # generates biomarker data from a gamma
                                    # distribution for the diseased group.
scores <- c(x,y)
D=c(pracma::zeros(1,100), pracma::ones(1,100))
out=checkboxcox(marker=scores, D, plots="on")
summary(out)
```

---

checkboxcox2	*Tests whether Box-Cox is appropriate for the given dataset (two-marker version)*

Description

This function tests whether the Box-Cox transformation is able to achieve approximate normality for your data. That is, it will allow the user to know whether it is appropriate to use all the methods discussed later on in this package.

Usage

```
checkboxcox2(marker1, marker2, D, plots, printShapiro = FALSE)
```

Arguments

marker1	A vector of length n that contains the biomarker scores of all individuals for the first marker.
marker2	A vector of length n that contains the biomarker scores of all individuals for the second marker.
D	A vector of length n that contains the true disease status of an individual. It is a binary vector containing 0 for the healthy/control individuals and 1 for the diseased individuals.
plots	Valid inputs are "on" and "off". When set to "on", the user gets the histograms of the biomarker for both the healthy and the diseased group before and after the Box-Cox transformation for both marker1 and marker2. In addition, all eight corresponding qq-plots are provided.
printShapiro	Boolean. When set to TRUE, the results of the Shapiro-Wilk test will be printed to the console. When set to FALSE, the results are suppressed. Default value is FALSE.

Value

res_shapiro	A results table that contains the results of eight Shapiro-Wilk tests for normality testing. Four of these refer to normality testing of the healthy and the diseased groups before the Box-Cox transformation, and the remaining four refer to the Box-Cox transformed biomarkers scores for the healthy and the diseased groups. Thus, this testing process produces eight p-values. In addition, if the plots are set to 'on', then the output provides (1) the histograms of the biomarker for both the healthy and the diseased groups before and after the Box-Cox transformation for both marker1 and marker2, (2) all eight corresponding qq-plots, and (3) two plots (one for marker1, one for marker2) with the empirical ROC curve overplotted with the Box-Cox based ROC curve for visual comparison purposes.
transx1	The Box-Cox transformed scores for the first marker and the healthy group.
transy1	The Box-Cox transformed scores for the first marker and the diseased group.

transformation.parameter.1

        The estimated Box-Cox transformation parameter (lambda) for marker 1.

transx2   The Box-Cox transformed scores for the second marker and the healthy group.

transy2   The Box-Cox transformed scores for the second marker and the diseased group.

transformation.parameter.2

        The estimated Box-Cox transformation parameter (lambda) for marker 2.

pval_x_marker1   The p-value of the Shapiro Wilk test of normality for the marker1 healthy group (before the Box-Cox transformation).

pval_y_marker1   The p-value of the Shapiro Wilk test of normality for the marker1 diseased group (before the Box-Cox transformation).

pval_transx_marker1

        The p-value of the Shapiro Wilk test of normality for the marker1 healthy group (after the Box-Cox transformation).

pval_transy_marker1

        The p-value of the Shapiro Wilk test of normality for the marker1 diseased group (after the Box-Cox transformation).

pval_x_marker2   The p-value of the Shapiro Wilk test of normality for the marker2 healthy group (before the Box-Cox transformation).

pval_y_marker2   The p-value of the Shapiro Wilk test of normality for the marker2 diseased group (before the Box-Cox transformation).

pval_transx_marker2

        The p-value of the Shapiro Wilk test of normality for the marker2 healthy group (after the Box-Cox transformation).

pval_transy_marker2

        The p-value of the Shapiro Wilk test of normality for the marker2 diseased group (after the Box-Cox transformation).

roc1   A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.

roc2   A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.

Author(s)

Leonidas Bantis

References

Bantis LE, Nakas CT, Reiser B. (2021). Statistical inference for the difference between two maximized Youden indices obtained from correlated biomarkers. Biometrical Journal, 63(6):1241-1253. https://doi.org/10.1002/bimj.202000128

Bantis LE, Nakas CT, Reiser B. (2018). Construction of confidence intervals for the maximum of the Youden index and the corresponding cutoff point of a continuous biomarker. Biometrical Journal, 61(1):138-156. https://doi.org/10.1002/bimj.201700107

Bantis LE, Feng Z. (2016). Comparison of two correlated ROC curves at a given specificity or sensitivity level. Statistics in Medicine, 35(24):4352-4367. https://doi.org/10.1002/sim.7008

Bantis LE, Nakas CT, Reiser B. (2014). Construction of confidence regions in the ROC space after the estimation of the optimal Youden index-based cut-off point. Biometrics, 70(1):212-223. https://doi.org/10.1111/biom.12107

Box GEP, Cox DR. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society. 26(2):211-252. https://www.jstor.org/stable/2984418

Examples

```
set.seed(123)

nx <- 100
Sx <- matrix(c(1, 0.5, 0.5, 1), nrow = 2, ncol = 2)

mux <- c(X = 10, Y = 12)
X = mvtnorm::rmvnorm(nx, mean = mux, sigma = Sx)

ny <- 100
Sy <- matrix(c(1.1, 0.6, 0.6, 1.1), nrow = 2, ncol = 2)

muy <- c(X = 11, Y = 13.7)
Y = mvtnorm::rmvnorm(ny, mean = muy, sigma = Sy)

dx = pracma::zeros(nx,1)
dy = pracma::ones(ny,1)

markers = rbind(X,Y);
marker1 = markers[,1]
marker2 = markers[,2]
D = c(rbind(dx,dy))

out=checkboxcox2(marker1, marker2, D, plots = "on")
summary(out)
```

---

comparebcAUC	*Performs inference to compare the AUCs of two correlated (or uncorrelated) Box-Cox transformed biomarkers at a given nominal level*

Description

This function provides a comparison of two correlated markers in terms of their AUCs (areas under the Box-Cox based ROC curves). Markers measurements are assumed to be taken on the same individuals for both markers.

Usage

```
comparebcAUC(marker1, marker2, D, alpha, plots)
```

Arguments

marker1	A vector of length n that contains the biomarker scores of all individuals for the first marker.
marker2	A vector of length n that contains the biomarker scores of all individuals for the second marker.
D	A vector of length n that contains the true disease status of an individual. It is a binary vector containing 0 for the healthy/control individuals and 1 for the diseased individuals.
alpha	Nominal level used to calculate the confidence intervals. A common choice is 0.05.
plots	Valid inputs are "on" and "off". When set to "on" it returns the Box-Cox based ROC along with informative information about the two AUCs in the legend of the plot.

Value

resultstable	A summary table of the comparison that contains the AUC of each marker, the p-value of the difference using the probit transformation, the p-value of the difference, and the confidence interval of the difference.
AUCmarker1	The AUC of the first marker.
AUCmarker2	The AUC of the second marker.
pvalue_probit_difference	
	The p-value for the comparison of the AUCs. It employs the probit transformation that has greater power than when not using it (as in pvalue_difference).
pvalue_difference	
	The p-value for the comparisons of the AUC (without the probit transformation). Simulations have shown that this is inferior to the 'pvalue_probit_difference'.
CI_difference	The confidence interval for the difference of the AUCs.
roc1	A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.
roc2	A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.
transx1	The Box-Cox transformed scores for the first marker and the healthy group.
transy1	The Box-Cox transformed scores for the first marker and the diseased group.
transformation.parameter.1	
	The estimated Box-Cox transformation parameter (lambda) for marker 1.
transx2	The Box-Cox transformed scores for the second marker and the healthy group.
transy2	The Box-Cox transformed scores for the second marker and the diseased group.
transformation.parameter.2	
	The estimated Box-Cox transformation parameter (lambda) for marker 2.

Author(s)

Leonidas Bantis

References

Bantis LE, Feng Z. (2016). Comparison of two correlated ROC curves at a given specificity or sensitivity level. Statistics in Medicine, 35(24):4352-4367. https://doi.org/10.1002/sim.7008

Examples

```
#GENERATE SOME BIVARIATE DATA===

set.seed(123)

nx <- 100
Sx <- matrix(c(1,    0.5,
               0.5,  1),
             nrow = 2, ncol = 2)

mux <- c(X = 10, Y = 12)
X=mvtnorm::rmvnorm(nx, mean = mux, sigma = Sx)

ny <- 100
Sy <- matrix(c(1.1,   0.6,
               0.6,   1.1),
             nrow = 2, ncol = 2)

muy <- c(X = 11, Y = 13.7)
Y=mvtnorm::rmvnorm(ny, mean = muy, sigma = Sy)

dx=pracma::zeros(nx,1)
dy=pracma::ones(ny,1)

markers=rbind(X,Y);
marker1=markers[,1]
marker2=markers[,2]
D=c(rbind(dx,dy))

#===COMPARE THE AUCs of Marker 1 vs Marker 2

out=comparebcAUC(marker1, marker2, D, alpha=0.05,  plots="on")
```

---

comparebcJ	*Performs inference to compare the Youden Indices of two correlated (or uncorrelated) Box-Cox transformed biomarkers at a given nominal level*

Description

This function provides a comparison of two correlated markers in terms of their J (Youden indices for Box-Cox based ROC curves). Marker measurements are assumed to be taken on the same individuals for both markers.

Usage

```
comparebcJ(marker1, marker2, D, alpha,  plots)
```

Arguments

marker1   A vector of length n that contains the biomarker scores of all individuals for the first marker.

marker2   A vector of length n that contains the biomarker scores of all individuals for the second marker.

D   A vector of length n that contains the true disease status of an individual, where 0 denotes a healthy/control individual, and 1 denotes a diseased individual.

alpha   Nominal level used to calculate the confidence intervals. A common choice is 0.05.

plots   Valid inputs are "on" and "off". When set to "on" it returns the Box-Cox based ROC along with informative information about the two AUCs in the legend of the plot.

Value

resultstable   A summary table of the comparison that contains the maximized Youden Index of each marker, the p-value of the difference using the probit transformation, the p-value of the difference, and the confidence interval of the difference.

J1   The maximized Youden Index of the first marker.

J2   The maximized Youden Index of the second marker.

pvalue_probit_difference

                  The p-value for the comparison of the Js. It employs the probit transformation that has greater power than when not using it (as in pvalue_difference).

pvalue_difference

                  The p-value for the comparisons of the Js (without the probit transformation). Simulations have shown that this is inferior to the 'pvalue_probit_difference'.

CI_difference   The confidence interval for the difference of the Js.

roc1   A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.

roc2   A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.

transx1   The Box-Cox transformed scores for the first marker and the healthy group.

transy1   The Box-Cox transformed scores for the first marker and the diseased group.

transformation.parameter.1

                  The estimated Box-Cox transformation parameter (lambda) for marker 1.

transx2   The Box-Cox transformed scores for the second marker and the healthy group.

transy2   The Box-Cox transformed scores for the second marker and the diseased group.

transformation.parameter.2

                  The estimated Box-Cox transformation parameter (lambda) for marker 2.

Author(s)

Leonidas Bantis

References

Bantis LE, Nakas CT, Reiser B. (2021). Statistical inference for the difference between two maximized Youden indices obtained from correlated biomarkers. Biometrical Journal, 63(6):1241-1253. https://doi.org/10.1002/bimj.202000128

Examples

```
#GENERATE SOME BIVARIATE DATA===

set.seed(123)

nx <- 100
Sx <- matrix(c(1,    0.5,
               0.5,  1),
             nrow = 2, ncol = 2)

mux <- c(X = 10, Y = 12)
X=mvtnorm::rmvnorm(nx, mean = mux, sigma = Sx)

ny <- 100
Sy <- matrix(c(1.1,   0.6,
               0.6,   1.1),
             nrow = 2, ncol = 2)

muy <- c(X = 11, Y = 13.7)
Y=mvtnorm::rmvnorm(ny, mean = muy, sigma = Sy)

dx=pracma::zeros(nx,1)
dy=pracma::ones(ny,1)

markers=rbind(X,Y);
marker1=markers[,1]
marker2=markers[,2]
D=c(rbind(dx,dy))

#===COMPARE THE Js of Marker 1 vs Marker 2

out=comparebcJ(marker1, marker2, D, alpha=0.05,  plots="on")
```

---

comparebcSens   *Performs inference to compare the sensitivities of two correlated (or uncorrelated) Box-Cox transformed biomarkers at a given specificty and a given nominal level*

Description

This function provides a comparison of two correlated markers in terms of their sensitivities at a given specificity (for Box-Cox based ROC curves). Markers measurements are assumed to be taken on the same individuals for both markers.

Usage

```
comparebcSens(marker1, marker2, D, atSpec, alpha, plots)
```

Arguments

marker1   A vector of length n that contains the biomarker scores of all individuals for the first marker.

marker2   A vector of length n that contains the biomarker scores of all individuals for the second marker.

D   A vector of length n that contains the true disease status of an individual, where 0 denotes a healthy/control individual, and 1 denotes a diseased individual.

alpha   Nominal level used to calculate the confidence intervals. A common choice is 0.05.

atSpec   The value of specificity at which the comparison of sensitivities will take place.

plots   Valid inputs are "on" and "off". When set to "on" it returns the Box-Cox based ROC along with informative information about the two AUCs in the legend of the plot.

Value

resultstable   A summary table of the comparison that contains the sensitivity of each marker at the given specificity, the p-value of the difference using the probit transformation, the p-value of the difference, and the confidence interval of the difference.

Sens1   The sensitivity at the selected specificity for the first marker.

Sens2   The sensitivity at the selected specificity for the second marker.

pvalue_probit_difference

The p-value for the comparison of the sensitivities. It employs the probit transformation that has greater power than when not using it (as in pvalue_difference).

pvalue_difference

The p-value for the comparisons of the sensitivities (without the probit transformation). Simulations have shown that this is inferior to the 'pvalue_probit_difference'.

CI_difference   The confidence interval for the difference of the sensitivities.

roc1   A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.

roc2   A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.

transx1   The Box-Cox transformed scores for the first marker and the healthy group.

transy1   The Box-Cox transformed scores for the first marker and the diseased group.

transformation.parameter.1

                  The estimated Box-Cox transformation parameter (lambda) for marker 1.

transx2   The Box-Cox transformed scores for the second marker and the healthy group.

transy2   The Box-Cox transformed scores for the second marker and the diseased group.

transformation.parameter.2

                  The estimated Box-Cox transformation parameter (lambda) for marker 2.

Author(s)

Leonidas Bantis

References

Bantis LE, Feng Z. (2016). Comparison of two correlated ROC curves at a given specificity or sensitivity level. Statistics in Medicine, 35(24):4352-4367. https://doi.org/10.1002/sim.7008

Examples

```
#GENERATE SOME BIVARIATE DATA===

set.seed(123)

nx <- 100
Sx <- matrix(c(1,   0.5,
               0.5,  1),
             nrow = 2, ncol = 2)

mux <- c(X = 10, Y = 12)
X=mvtnorm::rmvnorm(nx, mean = mux, sigma = Sx)

ny <- 100
Sy <- matrix(c(1.1,   0.6,
               0.6,   1.1),
             nrow = 2, ncol = 2)

muy <- c(X = 11, Y = 13.7)
Y=mvtnorm::rmvnorm(ny, mean = muy, sigma = Sy)

dx=pracma::zeros(nx,1)
dy=pracma::ones(ny,1)

markers=rbind(X,Y);
marker1=markers[,1]
marker2=markers[,2]
D=c(rbind(dx,dy))

#===COMPARE THE Sensitivities of Marker 1 vs Marker 2

out=comparebcSens(marker1=marker1, marker2=marker2, D=D, alpha =0.05, atSpec=0.8, plots="on")
summary(out)
```

comparebcSpec	*Performs inference to compare the specificities of two correlated (or uncorrelated) Box-Cox transformed biomarkers at a given sensitivity and a given nominal level*

Description

This function provides a comparison of two correlated markers in terms of their specificity at a given sensitivity (for Box-Cox based ROC curves). Marker measurements are assumed to be taken on the same individuals for both markers.

Usage

```
comparebcSpec(marker1, marker2, D, atSens, alpha, plots)
```

Arguments

marker1	A vector of length n that contains the biomarker scores of all individuals for the first marker.
marker2	A vector of length n that contains the biomarker scores of all individuals for the second marker.
D	A vector of length n that contains the true disease status of an individual, where 0 denotes a healthy/control individual, and 1 denotes a diseased individual.
alpha	Nominal level used to calculate the confidence intervals. A common choice is 0.05.
atSens	The value of sensitivity at which the comparison of specificities will take place.
plots	Valid inputs are "on" and "off". When set to "on" it returns the Box-Cox based ROC along with informative information about the two AUCs in the legend of the plot.

Value

resultstable	A summary table of the comparison that contains the FPR of each marker at the given sensitivity, the p-value of the difference using the probit transformation, the p-value of the difference, and the confidence interval of the difference.
Spec1	The specificity at the selected sensitivity for the first marker.
Spec2	The specificity at the selected sensitivity for the second marker.
pvalue_probit_difference	
	The p-value for the comparison of the specificities. It employs the probit transformation that has greater power than when not using it (as in pvalue_difference).
pvalue_difference	
	The p-value for the comparisons of the specificities (without the probit transformation). Simulations have shown that this is inferior to the 'pvalue_probit_difference'.
CI_difference	The confidence interval for the difference of the specificities.

roc1	A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.
roc2	A function that refers to the ROC of the first marker. It allows the user to feed in FPR values and the corresponding TPR values.
transx1	The Box-Cox transformed scores for the first marker and the healthy group.
transy1	The Box-Cox transformed scores for the first marker and the diseased group.
transformation.parameter.1	
	The estimated Box-Cox transformation parameter (lambda) for marker 1.
transx2	The Box-Cox transformed scores for the second marker and the healthy group.
transy2	The Box-Cox transformed scores for the second marker and the diseased group.
transformation.parameter.2	
	The estimated Box-Cox transformation parameter (lambda) for marker 2.

Author(s)

Leonidas Bantis

References

Bantis LE, Feng Z. (2016). Comparison of two correlated ROC curves at a given specificity or sensitivity level. Statistics in Medicine, 35(24):4352-4367. https://doi.org/10.1002/sim.7008

Examples

```
#GENERATE SOME BIVARIATE DATA===

set.seed(123)

nx <- 100
Sx <- matrix(c(1,    0.5,
               0.5,  1),
             nrow = 2, ncol = 2)

mux <- c(X = 10, Y = 12)
X=mvtnorm::rmvnorm(nx, mean = mux, sigma = Sx)

ny <- 100
Sy <- matrix(c(1.1,   0.6,
               0.6,   1.1),
             nrow = 2, ncol = 2)

muy <- c(X = 11, Y = 13.7)
Y=mvtnorm::rmvnorm(ny, mean = muy, sigma = Sy)

dx=pracma::zeros(nx,1)
dy=pracma::ones(ny,1)

markers=rbind(X,Y);
marker1=markers[,1]
```

```
marker2=markers[,2]
D=c(rbind(dx,dy))

#===COMPARE THE Specificities of Marker 1 vs Marker 2

out=comparebcSpec(marker1=marker1, marker2=marker2, D=D, alpha =0.05, atSens=0.8, plots="on")
summary(out)
```

---

rocboxcox   *Comprehensive Box-Cox ROC analysis for a single marker*

Description

This function applies the Box-Cox transformation to provide a comprehensive ROC analysis that involves the AUC (and its CI), the maximized Youden index (and its CI), the optimized Youden based cutoff (and its CI), and joint confidence regions for the optimized pair of sensitivity and specificity. Cutoffs are back-transformed so that they lie on the original scale of the data.

Usage

```
rocboxcox(marker, D, alpha, plots, printProgress = FALSE)
```

Arguments

marker	A vector of length n that contains the biomarker scores of all individuals.
D	A vector of length n that contains the true disease status of an individual. It is a binary vector containing 0 for the healthy/control individuals and 1 for the diseased individuals.
alpha	Nominal level used to calculate the confidence intervals. A common choice is 0.05.
plots	Valid inputs are "on" and "off". When set to "on", the function returns a comprehensive figure with the ROC estimate and several point estimates: AUC, Youden index, optimized Youden cutoff, and the Youden based sensitivty and specificity along with the corresponding marginal confidence intervals and the joint confidence region of the estimated sensitivity and specificity.
printProgress	Boolean. When set to TRUE, messages describing the progress of the bootstrapping will be printed to the console window. When set to FALSE, these messages are suppressed. Default value is FALSE.

Value

transx	The Box-Cox transformed scores for the healthy.
transy	The Box-Cox transformed scores for the diseased.
transformation.parameter	
	The estimated Box-Cox transformation parameter (lambda).
AUC	The estimated area under the Box-Cox based ROC curve.

AUCCI	The (1-a)100% CI for the AUC. This CI is based on probit transforming the AUC estimate, finding the CI on the real line, and then back-transforming its endpoints to the ROC space.
pvalueAUC	The corresponding p-value for the AUC estimate.
J	The maximized Youden index.
JCI	The corresponding CI for maximized Youden index. For this CI consider we consider the probit transformation and then back-transform its endpoints to derive a 95% CI for the Youden index itself.
pvalueJ	The corresponding two-tailed p-value for the maximized Youden index.
Sens	The sensitivity that corresponds to the Youden based optimized cutoff.
CImarginalSens	the marginal (1-a)100% CI for the sensitivity that corresponds to the Youden based optimized cutoff. This is derived by first employing the probit transformation, finding a CI on the real line, and then back-transforming its endpoints to the ROC space.
Spec	The Specificity that corresponds to the Youden based optimized cutoff.
CImarginalSpec	The marginal (1-a)100% CI for the specificity that corresponds to the Youden based optimized cutoff. This is derived by first employing the probit transformation, finding a CI on the real line, and then backtransforming its endpoints to the ROC space.
cutoff	The Youden-based optimized cutoff.
CIcutoff	The (1-a)100% CI for the Youden-based optimized cutoff. This is based on the bootstrap. It involves using the Box-Cox transformation for every bootstrap iteration and then using the inverse Box-Cox transformation to obtain the cutoff on its original scale.
areaegg	The area of the (1-a)100% egg-shaped joint confidence region that refers to the optimized pair of sensitivity and specificity. This takes into account the fact that the estimated sensitivity and specificity are correlated as opposed to the corresponding rectangular area that ignores this.
arearect	The area of the (1-a)100% rectangular joint confidence region that refers to the optimized pair of sensitivity and specificity. This ignores the correlation of the optimized sensitivity and specificity and tends to yield a larger area compared to the one of the egg-shaped region.
mxlam	The mean of the marker scores of the healthy group after the Box-Cox transformation.
sxlam	The standard deviation of the marker scores of the healthy group after the Box-Cox transformation.
mylam	The mean of the marker scores of the diseased group after the Box-Cox transformation.
sylam	The standard deviation of the marker scores of the diseased group after the Box-Cox transformation.
results	A table that provides some indicative results: the AUC, the J (maximized Youden index), the estimated cutoff, the sensitivity, and the specificity along with their marginal CIs.
roc	A function of the estimated Box-Cox ROC curve. You can use this to simply request TPR values for given FPR values.

Author(s)

Leonidas Bantis

References

Bantis LE, Nakas CT, Reiser B. (2014). Construction of confidence regions in the ROC space after the estimation of the optimal Youden index-based cut-off point. Biometrics, 70(1):212-223. doi: https://doi.org/10.1111/biom.12107

Box GEP, Cox DR. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society. 26(2):211-252. https://www.jstor.org/stable/2984418

Examples

```
set.seed(123)
x=rgamma(100, shape=2, rate = 8)
y=rgamma(100, shape=2, rate = 4)
scores=c(x,y)
D=c(pracma::zeros(1,100), pracma::ones(1,100))
out=rocboxcox(marker=scores,D, 0.05, plots="on")
```

---

rocboxcoxCI   *Inference around the sensitivity at a given specificity (and vice versa)*
                           for a single Box-Cox transformed biomarker

---

Description

This function applies the Box-Cox transformation and provides the ability to do inferences for the sensitivity at a given specificity (and vice versa).

Usage

```
rocboxcoxCI(marker, D, givenSP, givenSE, alpha, plots)
```

Arguments

marker	A vector of length n that contains the biomarker scores of all individuals.
D	A vector of length n that contains the true disease status of an individual, where 0 denotes a healthy/control individual, and 1 denotes a diseased individual.
givenSP	A vector of specificity values that the user wants to fix/set, at which the sensitivity is to be estimated. In this case, the 'givenSE' argument needs to be set to NA.
givenSE	A vector of sensitivity values that the user want to fix/set, at which the specificity is to be estimated. In this case, the 'givenSP' argument needs to be set to NA.
alpha	Nominal level used to calculate the confidence intervals. A common choice is 0.05.

plots	Valid inputs are "on" and "off". When set to "on", it returns both (1) the Box-Cox based ROC plot along with pointwise 95% confidence intervals for the full spectrum of FPRs and (2) a second plot that visualizes the confidence intervals at the given sensitivities or specificities.

Value

SPandCIs	The specificity values and the CIs around them.
SEandCIs	The sensitivity values and the CIs around them.
SEvalues	The sensitivity values provided by the user at which the specificity was calculated. If the user did not provide any sensitivity values, this argument should be set to NA.
SPvalues	The specificity values provided by the user at which the sensitivity was calculated. If the user did not provide any specificity values, this argument should be set to NA.

Author(s)

Leonidas Bantis

References

Bantis LE, Feng Z. (2016). Comparison of two correlated ROC curves at a given specificity or sensitivity level. Statistics in Medicine, 35(24):4352-4367. https://doi.org/10.1002/sim.7008

Box GEP, Cox DR. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society. 26(2):211-252. https://www.jstor.org/stable/2984418

Examples

```
set.seed(123)
x <- rgamma(100, shape=2, rate = 8)
y <- rgamma(100, shape=2, rate = 4)
scores <- c(x,y)
D=c(pracma::zeros(1,100), pracma::ones(1,100))
givenSP=c(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9)
givenSE=NA
out=rocboxcoxCI(marker=scores, D, givenSP=givenSP, givenSE=NA, alpha=0.05, plots="on")
```

---

threerocs	*Provides visual comparison of three ROC estimation methods (one-marker version)*

Description

This function provides a visual comparison of the Empirical ROC, the Box-Cox ROC, and the Metz binomial semi-parametric estimator of the ROC curve. It also computes the AUC for the curve corresponding to each method.

Usage

```
threerocs(marker, D, plots)
```

Arguments

marker	A vector of length n that contains the biomarker scores of all individuals.
D	A vector of length n that contains the true disease status of an individual. It is a binary vector containing 0 for the healthy/control individuals and 1 for the diseased individuals.
plots	Valid inputs are "on" and "off". When set to "on", the user gets a single plot containing the estimated ROC curves using the Empirical, Box-Cox, and Metz methods for each of the two provided markers.

Value

AUC_Empirical	The AUC of the empirical ROC curve.
AUC_Metz	The AUC of the Metz binormal curve (as calculated by MRMCaov package using the "binormal" option).
AUC_BoxCox	The AUC of the Box-Cox based ROC curve.

Author(s)

Leonidas Bantis

References

Box GEP, Cox DR. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society. 26(2):211-252. https://www.jstor.org/stable/2984418

Smith BJ, Hillis SL, Pesce LL (2023). MCMCaov: Multi-Reader Multi-Case Analysis of Variance. R package version 0.3.0, https://github.com/brian-j-smith/MRMCaov.

Smith BJ, Hillis SL (2020). "Multi-reader multi-case analysis of variance software for diagnostic performance comparison of imaging modalities." In Samuelson F, Taylor-Phillips S (eds.), Proceedings of SPIE 11316, Medical Imaging 2020: Image Perception, Observer Performance, and Technology Assessment, 113160K. doi:10.1117/12.2549075, https://pubmed.ncbi.nlm.nih.gov/32351258.

Examples

```
set.seed(123)
x <- rgamma(100, shape=2, rate = 8) # generates biomarker data from a gamma
                                    # distribution for the healthy group.
y <- rgamma(100, shape=2, rate = 4) # generates biomarker data from a gamma
                                    # distribution for the diseased group.
scores <- c(x,y)
D=c(pracma::zeros(1,100), pracma::ones(1,100))
out=threerocs(marker=scores, D, plots="on")
summary(out)
```

---

threerocs2   *Provides visual comparison of three ROC estimation methods (two-marker version)*

Description

This function provides a visual comparison of the Empirical ROC, the Box-Cox ROC, and the Metz binormal semi-parametric estimator of the ROC curve. It also computes the AUC for the curve corresponding to each method.

Usage

```
threerocs2(marker1, marker2, D, plots)
```

Arguments

marker1   A vector of length n that contains the biomarker scores of all individuals for the first marker.

marker2   A vector of length n that contains the biomarker scores of all individuals for the second marker.

D   A vector of length n that contains the true disease status of an individual. It is a binary vector containing 0 for the healthy/control individuals and 1 for the diseased individuals.

plots   Valid inputs are "on" and "off". When set to "on", the user gets a single plot containing the estimated ROC curves using the Empirical, Box-Cox, and Metz methods for each of the two provided markers.

Value

AUC_Empirical1   The AUC of the empirical ROC curve for the first marker.

AUC_Empirical2   The AUC of the empirical ROC curve for the second marker.

AUC_Metz1   The AUC of the Metz binormal curve (as calculated by MRMCaov package using the "binormal" option) for the first marker.

AUC_Metz2   The AUC of the Metz binormal curve (as calculated by MRMCaov package using the "binormal" option) for the second marker.

AUC_BoxCox1   The AUC of the Box-Cox based ROC curve for the first marker.

AUC_BoxCox2   The AUC of the Box-Cox based ROC curve for the second marker.

Author(s)

Leonidas Bantis

References

Box GEP, Cox DR. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society. 26(2):211-252. https://www.jstor.org/stable/2984418

Smith BJ, Hillis SL, Pesce LL (2023). MCMCaov: Multi-Reader Multi-Case Analysis of Variance. R package version 0.3.0, https://github.com/brian-j-smith/MRMCaov.

Smith BJ, Hillis SL (2020). "Multi-reader multi-case analysis of variance software for diagnostic performance comparison of imaging modalities." In Samuelson F, Taylor-Phillips S (eds.), Proceedings of SPIE 11316, Medical Imaging 2020: Image Perception, Observer Performance, and Technology Assessment, 113160K. doi:10.1117/12.2549075, https://pubmed.ncbi.nlm.nih.gov/32351258.

Examples

```
set.seed(123)

nx <- 100
Sx <- matrix(c(1, 0.5, 0.5, 1), nrow = 2, ncol = 2)

mux <- c(X = 10, Y = 12)
X = mvtnorm::rmvnorm(nx, mean = mux, sigma = Sx)

ny <- 100
Sy <- matrix(c(1.1, 0.6, 0.6, 1.1), nrow = 2, ncol = 2)

muy <- c(X = 11, Y = 13.7)
Y = mvtnorm::rmvnorm(ny, mean = muy, sigma = Sy)

dx = pracma::zeros(nx,1)
dy = pracma::ones(ny,1)

markers = rbind(X,Y);
marker1 = markers[,1]
marker2 = markers[,2]
D = c(rbind(dx,dy))

out=threerocs2(marker1, marker2, D, plots = "on")
summary(out)
```

Index