Package: ROCpsych (via r-universe)

October 15, 2026

Type Package

Title Compute and Compare Diagnostic Test Statistics Across Groups

Version 1.4

Date 2025-05-16

Author Shenghai Dai [aut, cre], Olasunkanmi J. Kehinde [aut], Maureen Schmitter-Edgecombe [aut], Brian F. French [aut]

Maintainer Shenghai Dai <s.dai@wsu.edu>

Description Functions for (1) computing diagnostic test statistics (sensitivity, specificity, etc.) from confusion matrices with adjustment for various base rates or known prevalence based on McCaffrey et al (2003) <doi:10.1007/978-1-4615-0079-7_1>, (2) computing optimal cut-off scores with different criteria including maximizing sensitivity, maximizing specificity, and maximizing the Youden Index from Youden (1950) <doi:10.1002/1097-0142(1950)3:1%3C32::AID-CNCR2820030106%3E3.0.CO;2-3>, and (3) displaying and comparing classification statistics and area under the receiver operating characteristic (ROC) curves or area under the curves (AUC) across consecutive categories for ordinal variables.

Depends R (>= 3.5.0), reportROC, pROC, stats

Encoding UTF-8

NeedsCompilation no

LazyData true

License GPL (>= 2)

URL https://daiscode.github.io/ROCpsych/

RoxygenNote 7.3.2

Repository https://daiscode.r-universe.dev

Date/Publication 2026-09-15 08:35:28 UTC

RemoteUrl https://github.com/daiscode/rocpsych

RemoteRef HEAD

RemoteSha 9ee2024c7db0772a0e6a1d4956d6ae9460faf743

Contents

cutscores 2
group.auc.test 3
group.to.vars 5
PV.BR 6
ROC.data.ex 7
ROC.stats 7

Index **10**

`cutscores` *Function to compute optimal cut-off scores*

Description

This function computes the optimal cut-off scores based on sensitivity, specificity, and the Youden Index (Youden, 1950) <doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3>.

Usage

```
cutscores(outcome, predictor)
```

Arguments

`outcome`	The outcome variable indicating the status in the form of a data frame or matrix. This variable is typically coded as 0 (positive) and 1 (negative).
`predictor`	A numerical vector of scores used to predict the status of the outcome. This variable should be of the same length as the outcome variable (i.e., two variables are from the same data set and also of the same number of data rows).

Value

A list of two objects: (1) summary statistics of selected cut scores, and (2) detailed information of each used cut score and corresponding classification statistics.

`Summary`	Summary statistics of selected cut scores. Specifically, * Cut.off, the select cut-off scores according to different criteria * SEN, Sensitivity, also true positive rate, the y-axis of the ROC. * SPE, Specificity, also true negative rate. * 1-SPE, the x-axis of the ROC. * Youden.Index. * TP, true positives. * FP, false positives. * FN, false negatives. * TN, true negatives.
`Details`	Detailed information of each used cut score and corresponding classification statistics.

References

Youden, W.J. (1950). "Index for rating diagnostic tests." Cancer,3, 32-35. doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.

Examples

```
#read the example data
data(ROC.data.ex)
#run the function
result<-cutscores(ROC.data.ex$outcome, ROC.data.ex$predictor)
#obtain results
result$Summary
result$Details
```

`group.auc.test` *Function to compare AUC across all consecutive categories of an ordinal scale*

Description

This function computes commonly used classification statistics of a confusion matrix and compares the area under the curve (AUC) across all consecutive categories of an ordinal variable. The function of roc.test () from the pROC package (https://cran.r-project.org/package=pROC) is used for AUC comparison.

Usage

```
group.auc.test(outcome,predictor,
                      groups, cut.off='max.Youden',BR=1)
```

Arguments

`outcome`	The outcome variable indicating the status in the form of a data frame or matrix. This variable is typically coded as 0 (positive) and 1 (negative).
`predictor`	A numerical vector of scores used to predict the status of the outcome. This variable should be of the same length as the outcome variable (i.e., two variables are from the same data set and also of the same number of data rows).
`groups`	A data frame that contains all created indicator variables using the function group.to.vars () in this package.
`cut.off`	Specification of the criterion used to select the optimal cut score. Three options available: (1) 'max.Youden' returns the cut score that maximizes the Youden Index (the default); (2) 'max.sen' returns the cut score that maximizes the sensitivity; and (3) 'max.spe' returns the cut score that maximizes the specificity.
`BR`	Base rates or known prevalence. Multiple values can be specified simultaneously. By default BR=1.

Value

A list of two objects: (1) descriptive and classification statistics, and (2) results of the AUC comparison for each pair of the consecutive categories.

`Summary.Stats` Summary and classification statistics for all participants and all the consecutive groups. The first row is the results of the entire sample and has a row name of "All", followed by results for each pair of the groups specified by group.to.vars (). For example, if the first indicator of age is age.40, then the second row of results will have the row name of "age.40" and includes results for participants with age at or below 40, the third row will have the row name of "age.40.1" and includes results for those with age beyond 40.
The results include the following statistics:
* N, the sample size for each category.
* TP, true positives.
* FP, false positives.
* FN, false negatives.
* TN, true negatives.
* Cut.off, the optimal cut score.
* AUC, Area under the ROC curve.
* AUC.SE, Standard error of AUC.
* AUC.low & AUC.up, '95 * Sensitivity, also true positive rate, y-axis of the ROC.
* Specificity, also true negative rate.
* Youden.Index.
* PPV or positive predictive value for each specified base rate.
* NPV or negative predictive value for each specified base rate.
* PPV for the sample.
* NPV for the sample.
* FNR, false negative rate, or miss rate.
* FPR, false positive rate, or fall-out rate.
* FOR, false omission rate.
* FDR, false discovery rate.
* Prevalence.
* Accuracy.
* PLR, positive likelihood ratio.
* NLR, negative likelihood ratio.
* DOR, Diagnostic odds ratio.

`AUC.test` Results of the AUC comparison for each pair of the consecutive categories.

Examples

```
#read the example data
data(ROC.data.ex)
#create new binary variables for the ordinal variable
data.new.age<-group.to.vars(ROC.data.ex,
                            ROC.data.ex$age,
                           root.name='age')
#run the function
result.age<-group.auc.test(ROC.data.ex$outcome,ROC.data.ex$predictor,
```

```
                          groups=data.new.age[,5:ncol(data.new.age)],
                          cut.off='max.Youden', BR=1)
#obtain results
result.age$Summary.Stats
result.age$AUC.test
```

---

group.to.vars *Function to create new variables from the ordinal variable for further analysis*

Description

This function collapses group memberships or categories of the ordinal variable into binary variables (or indicators) for each category and appends the new variables to the end of the original data. For each new variable, 0 represents participants at or below the selected category and 1 represents participants above the selected category. For example, age.40 = 0 means participants with age at or below 40, whereas age.40 = 1 indicates participants with age beyond 40.

Usage

```
group.to.vars(data, group, root.name=NULL)
```

Arguments

data	A data frame or matrix that contains the ordinal variable.
group	The ordinal variable in the 'data' object.
root.name	Indicate whether a root name is used to name the new variables. If not specified (by default, root.name=NULL), the variable name will be used as the root.

Value

A data frame with the original data and newly created variables.

Examples

```
#read the example data
data(ROC.data.ex)
#create new binary variables for the ordinal variable
data.new.age<-group.to.vars(ROC.data.ex,
                            ROC.data.ex$age,
                           root.name='age')
```

PV.BR *Function to compute PPV and NPV with specified base rates*

Description

This function computes positive predictive values (PPV) and negative predictive values (NPV) with provided base rates (or known prevalence).

Usage

```
PV.BR(outcome, predictor,cut.off='max.Youden', BR=1)
```

Arguments

outcome	The outcome variable indicating the status in the form of a data frame or matrix. This variable is typically coded as 0 (positive) and 1 (negative).
predictor	A numerical vector of scores used to predict the status of the outcome. This variable should be of the same length as the outcome variable (i.e., two variables are from the same data set and also of the same number of data rows).
cut.off	Specification of the criterion used to select the optimal cut score. Three options available: (1) 'max.Youden' returns the cut score that maximizes the Youden Index (the default); (2) 'max.sen' returns the cut score that maximizes the sensitivity; and (3) 'max.spe' returns the cut score that maximizes the specificity.
BR	Base rates or known prevalence. Multiple values can be specified simultaneously. By default BR=1.

Value

An object that contains results of classification statistics.

Result	* Cut.off, the optimal cut score. * Sensitivity, also true positive rate, the y-axis of the ROC. * Specificity, also true negative rate. * Youden.Index. * PPV or positive predictive value for each specified base rate. * NPV or negative predictive value for each specified base rate. * PPV for the sample. * NPV for the sample.

References

McCaffrey R.J., Palav A.A., O'Bryant S.E., Labarge A.S. (2003). "A Brief Overview of Base Rates. In: McCaffrey R.J., Palav A.A., O'Bryant S.E., Labarge A.S. (eds) Practitioner's Guide to Symptom Base Rates in Clinical Neuropsychology. Critical Issues in Neuropsychology. ." Springer, Boston, MA. doi:10.1007/978-1-4615-0079-7_1.

Examples

```
#read the example data
data(ROC.data.ex)
#run the function
PV.BR(ROC.data.ex$outcome, ROC.data.ex$predictor,
      cut.off='max.Youden', BR=1)
```

`ROC.data.ex` *Example data*

Description

This hypothetical dataset contains records of the outcome, the predictor, gender, and age from 241 participants.

Usage

```
data("ROC.data.ex")
```

Format

A data frame with 241 observations on the following 4 variables.

`outcome` a numeric vector

`predictor` a numeric vector

`gender` a numeric vector

`age` a numeric vector

Examples

```
data(ROC.data.ex)
## maybe str(ROC.data.ex) ; plot(ROC.data.ex) ...
```

`ROC.stats` *Function to compute statistics from a confusion matrix*

Description

This function computes all diagnostic statistics from a confusion matrix.

Usage

```
ROC.stats(outcome, predictor,cut.off='max.Youden',BR=1)
```

Arguments

outcome	The outcome variable indicating the status in the form of a data frame or matrix. This variable is typically coded as 0 (positive) and 1 (negative).
predictor	A numerical vector of scores used to predict the status of the outcome. This variable should be of the same length as the outcome variable (i.e., two variables are from the same data set and also of the same number of data rows).
cut.off	Specification of the criterion used to select the optimal cut score. Three options available: (1) 'max.Youden' returns the cut score that maximizes the Youden Index (the default); (2) 'max.sen' returns the cut score that maximizes the sensitivity; and (3) 'max.spe' returns the cut score that maximizes the specificity.
BR	Base rates or known prevalence. Multiple values can be specified simultaneously. By default BR=1.

Value

An object that contains the results.

ROC.stats

Summary and classification statistics for all participants and all the consecutive groups. Specifically.
* N, sample size for each category.
* TP, true positives.
* FP, false positives.
* FN, false negatives.
* TN, true negatives.
* Cut.off, the optimal cut score.
* AUC, Area under the ROC curve.
* AUC.SE, Standard error of AUC.
* AUC.low & AUC.up, '95 * Sensitivity, also true positive rate, the y-axis of the ROC.
* Specificity, also true negative rate.
* Youden.Index.
* PPV or positive predictive value for each specified base rate.
* NPV or negative predictive value for each specified base rate.
* PPV for the sample.
* NPV for the sample.
* FNR, false negative rate, or miss rate.
* FPR, false positive rate, or fall-out rate.
* FOR, false omission rate.
* FDR, false discovery rate.
* Prevalence.
* Accuracy.
* PLR, positive likelihood ratio.
* NLR, negative likelihood ratio.
* DOR, Diagnostic odds ratio.

Examples

```
#read the example data
```

```
data(ROC.data.ex)
#run the function
ROC.stats(ROC.data.ex$outcome, ROC.data.ex$predictor,
          cut.off='max.Youden',BR=1)
```

Index

∗ **datasets**
ROC.data.ex, 7

cutscores, 2

group.auc.test, 3
group.to.vars, 5

PV.BR, 6

ROC.data.ex, 7
ROC.stats, 7