Package ‘RJafroc’

November 14, 2018

Type Package

Title Modeling, Analysis, Validation and Visualization of Observer Performance Studies in Diagnostic Radiology

Version 1.1.0

Date 2018-11-10

Depends R (>= 3.3.0)

Imports openxlsx, ggplot2, stringr, tools, utils, stats, bbmle, binom, mvtnorm, numDeriv, Rcpp

Suggests testthat, knitr, rmarkdown

LinkingTo Rcpp

Description Tools for quantitative assessment of medical imaging systems, radiologists or computer aided (‘CAD’) algorithms.


‘ROC’ data consists of a single rating per image, where the rating is the perceived confidence level the image is of a diseased patient.

‘FROC’ data consists of a variable number (including zero) of marking pairs per image, where a mark is the location of a clinically reportable suspicious region and the rating is the corresponding confidence level that it is a true lesion. The software supersedes the current Windows version of ‘JAFROC’ software <http://www.devchakraborty.com> which is no longer supported.

‘RJafroc’ is derived from it being an enhanced R version of original Windows ‘JAFROC’. Implemented are a number of figures of merit quantifying performance, functions for visualizing operating characteristics; three ROC ratings data curve-fitting algorithms:

the ‘binormal’ model (‘BM’), the contaminated binormal model (‘CBM’) and the radiological search model (‘RSM’). Also implemented is
maximum likelihood fitting of paired ROC data utilizing the correlated 'CBM' model ('COR-CBM'). Unlike the 'BM', 'CBM', 'CORCBM' and the 'RSM' predict proper ROC curves that do not cross the chance diagonal or display inappropriate hooks, usually near the upper right corner of the plots. 'RSM' fitting yields measures of search and lesion-classification performances, in addition to the usual case-classification performance measured by the area under the 'ROC' curve. Search performance is the ability to find lesions while avoiding finding non-lesions. Lesion-classification performance is the ability to discriminate between found lesions and non-lesions. For fully crossed study designs, termed multiple-reader multiple-case, significance testing of reader-averaged figure-of-merit differences between modalities is implemented via both 'Dorfman', 'Berbaum' and 'Metz' ('DBM') and the 'Obuchowski' and 'Rockette' ('OR') methods, both substantially improved by 'Hillis'. Single treatment analysis allows comparison of performance of a group of radiologists to a specified value, or comparison of 'CAD' performance to a group of radiologists interpreting the same cases. Sample size estimation tools are provided for 'ROC' studies that allow estimation of relevant variances from a pilot study to predict required numbers of readers and cases in a pivotal study. 'FROC' sample size estimation is implemented in Online Appendix Chapter 19 available at <https://github.com/dpc10ster/onlinebookk21778>. Utility and data file manipulation functions allow data to be read in any of the currently used input formats, including Excel, and the results of the analysis can be viewed in text or Excel output files.

VignetteBuilder knitr
License GPL-3
LazyData true
URL http://www.devchakraborty.com
RoxygenNote 6.1.1
Encoding UTF-8
NeedsCompilation yes
Author Dev Chakraborty [cre, aut, cph], Xuetong Zhai [aut], Lucy D'Agostino McGowan [ctb], Alejandro RodriguezRuiz [ctb]
Maintainer Dev Chakraborty <dpc10ster@gmail.com>
Repository CRAN
Date/Publication 2018-11-14 18:30:03 UTC

R topics documented:
RJafroc-package
R topics documented:

binnedData123..........................................................15
binnedData124..........................................................16
binnedData125..........................................................17
ChisqrGoodnessOfFit..................................................18
Compare3ProperRocFits.................................................18
dataset01..............................................................20
dataset02..............................................................21
dataset03..............................................................22
dataset04..............................................................23
dataset05..............................................................24
dataset06..............................................................25
dataset07..............................................................26
dataset08..............................................................27
dataset09..............................................................28
dataset10..............................................................29
dataset11..............................................................30
dataset12..............................................................31
dataset13..............................................................32
dataset14..............................................................33
datasetCadLroc..........................................................34
datasetCrossedModality.................................................35
datasetDegenerate......................................................36
Df2RJafrocDataset.......................................................36
DfBinDataset..........................................................38
DfCreateCorCbmDataset.................................................40
DfExtractCorCbmDataset................................................41
DfExtractDataset......................................................42
DfFroc2Afroc..........................................................43
DfFroc2Roc............................................................44
DfLroc2Roc............................................................46
DfReadCrossedModalities.............................................47
DfReadDataFile.........................................................48
DfReadLrocDataFile....................................................49
DfSaveDataFile........................................................50
FitBinormalRoc........................................................51
FitCbmRoc..............................................................53
FitCorCbm..............................................................55
FitRsmRoc..............................................................56
PlotBinormalFit.........................................................58
PlotCbmFit............................................................59
PlotEmpiricalOperatingCharacteristics..............................60
PlotRsmOperatingCharacteristics.....................................62
SimulateCorCbmDataset.................................................65
SimulateFrocDataset....................................................66
SimulateRocDataset....................................................67
SsPowerGivenJK.........................................................69
SsPowerTable........................................................70
SsSampleSizeKGivenJ..................................................72
RJafroc-package  

**Description**

RJafroc implements software for assessing medical imaging systems, radiologists or computer aided detection algorithms. Models of observer performance are implemented, including the binormal model (BM), the contaminated binormal model (CBM), the correlated contaminated binormal model (CORCBM), and the radiological search model (RSM). The software and applications are described in a book - *Chakraborty DP: Observer Performance Methods for Diagnostic Imaging - Foundations, Modeling, and Applications with R-Based Examples.* Taylor-Francis LLC; 2017 - and its Online Appendices [https://github.com/dpc10ster/onlinebookk21778](https://github.com/dpc10ster/onlinebookk21778).

Observer performance data collection paradigms are the receiver operating characteristic (ROC) and its location specific extensions, primarily free-response ROC (FROC) and the location ROC (LROC). ROC data consists of single ratings per images. A rating is the perceived confidence level that the image is that of a diseased patient. FROC data consists of a variable number (including zero) of mark-rating pairs per image, where a mark is the location of a clinically reportable suspicious region and the rating is the corresponding confidence level that it is a true lesion. LROC data consists of a rating and a forced localization of the most suspicious region on every image. RJafroc supersedes the Windows version of JAFROC software V4.2.1, [http://www.devchakraborty.com](http://www.devchakraborty.com).

Data file related function names are preceded by *Df*, curve fitting functions by *Fit*, included data sets by *dataset*, plotting functions by *Plot*, significance testing functions by *St*, sample size related functions by *Ss*, data simulation functions by *Simulate* and utility functions by *Util*. RJafroc implements a number of figures of merit (FOMs) for quantifying performance, functions for visualizing empirical operating characteristics: e.g., ROC, FROC, alternative FROC (AFROC) and weighted AFROC (wAFROC) curves. Four maximum likelihood curve-fitting algorithms are implemented: the binormal model (BM), the contaminated binormal model (CBM), the correlated contaminated binormal model (CORCBM) and the radiological search model (RSM). Unlike the binormal model, CBM,
CORCBM and RSM predict "proper" ROC curves that do not cross the chance diagonal or display inappropriate "hooks", typically near the upper right corner of the plots. RSM fitting additionally yields measures of search and lesion-classification performances. Search performance is the ability to find lesions while avoiding finding non-lesions. Lesion-classification performance is the ability to correctly classify found lesions from found non-lesions. For fully crossed study designs significance testing of reader-averaged FOM differences between modalities is implemented via both Dorfman-Berbaum-Metz and the Obuchowski-Rockette methods, both substantially improved by Hillis. Also implemented are single treatment analyses, which allow comparison of performance of a group of radiologists to a specified value, or comparison to CAD to a group of radiologists interpreting the same cases. Crossed-modality analysis is implemented wherein there are two crossed treatment factors and the desire is to determined performance in each treatment factor averaged over all levels of the other factor. Sample size estimation tools are provided for ROC studies; these use estimates of the relevant variances from a pilot study to predict required numbers of readers and cases in a pivotal study to achieve a desired power. FROC sample size estimation is implemented in Online Appendix Chapter 19, available at <https://github.com/dpc10ster/onlinebookk21778>. Utility and data file manipulation functions allow data to be read in any of the currently used input formats, including Excel, and the results of the analysis can be viewed in text or Excel output files. The methods are illustrated with several included datasets from the author’s international collaborations. The package is used extensively in the cited book and its online appendices.

Details

Package: RJafrroc
Type: Package
Version: 1.1.0
Date: 2018-06-31
License: GPL-3
URL: http://www.devchakraborty.com

Abbreviations and definitions

- \( a \): The separation or "a" parameter of the binormal model
- AFROC curve: plot of LLF (ordinate) vs. FPF, where FPF is inferred using highest rating of NL marks on non-diseased cases
- AFROC: alternative FROC, see Chakraborty 1989
- AFROC1 curve: plot of LLF (ordinate) vs. FPF1, where FPF1 is inferred using highest rating of NL marks on ALL cases
- \( \alpha \): The significance level \( \alpha \) of the test of the null hypothesis of no treatment effect
- AUC: area under curve; e.g., ROC-AUC = area under ROC curve, an example of a FOM
- \( b \): The width or "b" parameter of the conventional binormal model
- Binormal model: two unequal variance normal distributions, one at zero and one at \( \mu \), for modeling ROC ratings, \( \sigma \) is the std. dev. ratio of diseased to non-diseased distributions
- CAD: computer aided detection algorithm
• CBM: contaminated binormal model (CBM); two equal variance normal distributions for modeling ROC ratings, the diseased distribution is bimodal, with a peak at zero and one at \( \mu \), the integrated fraction at \( \mu \) is \( \alpha \) (not to be confused with \( \alpha \) of NH testing)
• CI: The \((1-\alpha)\) confidence interval for the stated statistic
• Crossed modality: a dataset containing two modality factors, with the levels of the two factors crossed, see paper by Thompson et al
• DBM: Dorfman-Berbaum-Metz, a significance testing method for detecting a treatment effect in MRMC studies
• DBMH: Hillis’ modification of the DBM method
• ddf: Denominator degrees of freedom of appropriate \( F \)-test; the corresponding ndf is \( I - 1 \)
• Empirical AUC: trapezoidal area under curve, same as the Wilcoxon statistic for ROC paradigm
• FN: false negative, a diseased case classified as non-diseased
• FOM: figure of merit, a quantitative measure of performance, performance metric
• FP: false positive, a non-diseased case classified as diseased
• FPF: number of FPs divided by number of non-diseased cases
• FROC curve: plot of LLF (ordinate) vs. NLF
• FROC: free-response ROC (a data collection paradigm where each image yields a random number, 0, 1, 2,..., of mark-rating pairs)
• FRRC: Analysis that treats readers as fixed and cases as random factors
• I: total number of modalities, indexed by \( i \)
• image/case: used interchangeably; a case can consist of several images of the same patient in the same modality
• iMRMC: A text file format used for ROC data by FDA/CDRH researchers
• individual dataset: A single modality single reader dataset.
• Intrinsic: Used in connection with RSM; a parameter that is independent of the RSM \( \mu \) parameter, but whose meaning may not be as transparent as the corresponding physical parameter
• J: total number of readers, indexed by \( j \)
• JAFROC file format: A .xlsx file format, applicable to ROCX, FROC and LROC paradigms
• JAFROC FOM: trapezoidal area under AFROC curve; this term is obsolete; use AFROC-AUC instead
• JAFROC: jackknife AFROC: Windows software for analyzing observer performance data: no longer updated, replaced by current package; the name is a misnomer as the jackknife is used only for significance testing; alternatively, the bootstrap could be used; what distinguishes FROC from ROC analysis is the use of the AFROC-AUC as the FOM. With this change, the DBM or the OR method can be used for significance testing
• JAFROC1 FOM: trapezoidal area under AFROC1 curve; this term is obsolete; use AFROC1-AUC instead
• K: total number of cases, \( K = K1 + K2 \), indexed by \( k \)
• K1: total number of non-diseased cases, indexed by \( k1 \)
• K2: total number of diseased cases, indexed by \( k2 \)
• LL: lesion localization i.e., a mark that correctly locates an existing localized lesion; TP is a special case, when the proximity criterion is lax (i.e., "acceptance radius" is large)
• LLF: number of LLs divided by the total number of lesions
• LROC: location receiver operating characteristic, a data collection paradigm where each image yields a single rating and one location
• lrc/MRMC: A text file format used for ROC data by University of Iowa researchers
• mark: the location of a suspected diseased region
• maxLL: maximum number of lesions per case in dataset
• maxNL: maximum number of NL marks per case in dataset
• MRMC: multiple reader multiple case (each reader interprets each case in each modality, i.e. fully crossed study design)
• ndf: Numerator degrees of freedom of appropriate $F$-test, usually number of treatments minus one
• NH: The null hypothesis that all treatment effects are zero; rejected if the $p$-value is smaller than $\alpha$
• NL: non-lesion localization, of which FP is a special case, i.e., a mark that does not correctly locate any existing localized lesion(s)
• NLF: number of NLs divided by the total number of cases
• Operating characteristic: A plot of normalized correct decisions on diseased cases along ordinate vs. normalized incorrect decisions on non-diseased cases
• Operating point: A point on an operating characteristic, e.g., (FPF, TPF) represents an operating point on an ROC
• OR: Obuchowski-Rockette, a significance testing method for detecting a treatment effect in MRMC studies
• ORH: Hillis’ modification of the OR method
• Physical parameter: Used in connection with RSM; a parameter whose meaning is more transparent than the corresponding intrinsic parameter, but which depends on the RSM $\mu$ parameter
• Proximity criterion / acceptance radius: Used in connection with FROC (or LROC data); the "nearness" criterion is used to determine if a mark is close enough to a lesion to be counted as a LL (or correct localization); otherwise it is counted as a NL (or incorrect localization)
• p-value: the probability, under the null hypothesis, that the observed treatment effects, or larger, could occur by chance
• Proper: a proper fit does not inappropriately fall below the chance diagonal, does not display a "hook" near the upper right corner
• PROPROC: Metz’s binormal model based fitting of proper ROC curves
• RSM, Radiological Search Model: two unit variance normal distributions for modeling NL and LL ratings; four parameters, $\mu$, $\nu'$, $\lambda'$ and $\zeta$
• Rating: Confidence level assigned to a case; higher values indicate greater confidence in presence of disease; $\sim Inf$ is allowed but NA is not allowed
• Reader/observer/radiologist/CAD: used interchangeably
• RJafroc: the current software
• ROC: receiver operating characteristic, a data collection paradigm where each image yields a single rating and location information is ignored
• ROC curve: plot of TPF (ordinate) vs. FPF, as threshold is varied; an example of an operating characteristic
• ROCFIT: Metz software for binormal model based fitting of ROC data
• ROI: region-of-interest (each case is divided into a number of ROIs and the reader assigns an ROC rating to each ROI)
• FRRC: Analysis that treats readers as fixed and cases as random factors
• RRFC: Analysis that treats readers as random and cases as fixed factors
• RRRC: Analysis that treats both readers and cases as random factors
• RSCORE-II: original software for binormal model based fitting of ROC data
• RSM: Radiological search model, also method for fitting a proper ROC curve to ROC data
• RSM-ζ1: Lowest reporting threshold, determines if suspicious region is actually marked
• RSM-λ: Intrinsic parameter of RSM corresponding to λ', independent of µ
• RSM-λ'': Physical Poisson parameter of RSM, average number of latent NLs per case; depends on µ
• RSM-µ: separation of the unit variance distributions of RSM
• RSM-ν': Intrinsic parameter of RSM, corresponding to ν', independent of µ
• RSM-ν'': binomial parameter of RSM, probability that lesion is found
• SE: sensitivity, same as TPF
• Significance testing: determining the p-value of a statistical test
• SP: specificity, same as 1 – FPF
• Threshold: Reporting criteria: if confidence exceeds a threshold value, report case as diseased, otherwise report non-diseased
• TN: true negative, a non-diseased case classified as non-diseased
• TP: true positive, a diseased case classified as diseased
• TPF: number of TPs divided by number of diseased cases
• Treatment/modality: used interchangeably, for example, computed tomography (CT) images vs. magnetic resonance images (MRI)
• wAFROC curve: plot of weighted LLF (ordinate) vs. FPF, where FPF is inferred using highest rating of NL marks on non-diseased cases ONLY
• wAFROC1 curve: plot of weighted LLF (ordinate) vs. FPF1, where FPF1 is inferred using highest rating of NL marks on ALL cases
• wJAFROC FOM: weighted trapezoidal area under AFROC curve: this term is obsolete; use wAFROC-AUC instead; this is the recommended FOM
• wJAFROC1 FOM: weighted trapezoidal area under AFROC1 curve: only use if there are zero non-diseased cases is always number of treatments minus one
Dataset

Dataset, an object, can be created by the user or read from an external text of Excel file. The dataset is a list generally containing 8 elements (9 elements for crossed-modality or LROC datasets): Note: 

- $\text{Inf}$ is used to indicate the ratings of unmarked lesions and/or to indicate unavailable array items. An example of the latter would be if the maximum number of NLs in a dataset was 4, but some images had fewer than 4 NLs, in which case the corresponding "empty" positions would be filled with $\text{Inf}$s. Do not use $\text{NA}$ to denote a rating.

Note: the word "dataset" used in this package always represents an $\mathbb{R}$ object with one of the following structures:

**General data structure, example** dataset02 and dataset05:

- NL: a floating-point array with dimension $c(I, J, K, \ maxNL)$ containing the ratings of NL marks. The first $K_1$ locations of the third index corresponds to NL marks on non-diseased cases and the remaining locations correspond to NL marks on diseased cases. The 4th dimension allows for the possibility of multiple NL marks on a case. For FROC datasets unavailable NL ratings are assigned $\text{Inf}$. For ROC datasets FP ratings are assigned to the first $K_1$ elements of $\text{NL}[,,1:K_1,1]$ and the remaining $K_2$ elements of $\text{NL}[,,(K_1+1):K,1]$ are set to $\text{Inf}$. When converting from FROC to ROC data the software assigns $\text{Inf}$ to cases with no marks.
- LL: a floating-point array with dimension $c(I, J, K_2, \ maxLL)$ that contains the ratings of all LL marks. For ROC datasets TP ratings are assigned to $\text{LL}[,,1:K_2,1]$.
- lesionNum: an integer vector with length $K_2$, whose elements indicate the number of lesions in each diseased case.
- lesionID: an integer array with dimension $[K_2, \ maxLL]$. Its contents label lesions on diseased cases. For example, dataset05$\$lesionID[40,] is $c(1,2,\text{Inf})$, meaning the 40th diseased case in this dataset has two lesions, labeled 1 and 2. The lesionID of an LL in the ‘TP’ or ‘LL’ worksheet must correspond to the lesionID for that case in the ‘Truth’ worksheet. For example, if the lesionID for the 40th diseased case in the ‘TP’ or ‘LL’ worksheet is 2, then the associated rating must correspond to the lesion labeled 2 in the ‘Truth’ worksheet, etc.
- lesionWeight: a floating point array with dimension $c(K_2, \ maxLL)$, representing the relative importance of detecting each lesion. For each case, the weights sum to unity. If zero is assigned to the Weight field in the ‘Truth’ worksheet, the software automatically assigns equal weighting, e.g., dataset05$\$lesionWeight[40,] is $c(0.5,0.5)$, corresponding to equal weights (1/2) to each lesion on an image with two lesions.
- dataType: a string variable: "ROC", "ROI" or "FROC".
- modalityID: a string vector of length I, which labels the modalities in the dataset.
- readerID: a string vector of length J, which labels the readers. For example, $\text{NL}[1,2,,]$ indicates the NL-rating of the reader identified with the second label in readerID using the modality identified with the first label in modalityID.

**LROC data structure, example** datasetCadlroc:

- NL: a floating-point array with dimension $c(I, J, K, 1)$ that contains the ratings of FP marks. For the third index, the first $K_1$ elements contain valid ratings while the rest are filled with $\text{Inf}$s.
- LLCL: a floating-point array with dimension $c(I, J, K_2, 1)$ that contains the ratings of all correct localization (CL) marks. A $\text{Inf}$ indicates a case with no CL mark.
• LL1: a floating-point array with dimension \( c(I, J, K2, 1) \) that contains the ratings of all incorrect localization (IL) marks. A -Inf indicates a case with no IL mark.

• lesionNum: same as general case.

• lesionID: lesionID: an integer vector with length \( K2 \) containing ones.

• lesionWeight: a floating point array with dimension \( c(K2, 1) \) containing ones.

• dataType: a string variable: "LROC".

• modalityID: same as general case.

• readerID: same as general case.

**Crossed modality data structure, example** datasetCrossedModality:

• NL: a floating-point array with dimension \( c(I1, I2, J, K, \text{maxNL}) \) that contains the ratings of NL marks. Note the existence of two modality indices.

• LL: a floating-point array with dimension \( c(I1, I2, J, K2, \text{maxLL}) \) that contains the ratings of all LL marks. Note the existence of two modality indices.

• lesionNum: same as general case.

• lesionID: same as general case.

• lesionWeight: same as general case.

• dataType: a string variable: "ROC" or "FROC".

• modalityID1: same as general case, corresponding to first modality factor.

• modalityID2: same as general case, corresponding to second modality factor.

• readerID: same as general case.

**Data file format**:

The package reads JAFROC, MRMC (ROC data only) and iMRMC (ROC data only) data files. The data can be imported by using the function `DfReadDataFile`.

- **JAFROC data file format** The JAFROC data file is an Excel file containing three worksheets (*.xls and *.xlsx are supported): (1) the ‘Truth’ worksheet, (2) the ‘TP’ or ‘LL’ worksheet and (3) the ‘FP’ or ‘NL’ worksheet. Except for the ‘Truth’ worksheet, where each case must occur at least once, the number of rows in the other worksheets is variable.

  1. ‘Truth’ worksheet consists of
     - ‘CaseID’, an integer field uniquely labeling the cases (images). It must occur at least once for each case, and since a case may have multiple lesions, it can occur multiple times, once for each lesion.
     - ‘LesionID’, an integer field uniquely labeling the lesions in each case. This field is zero for non-diseased cases.
     - ‘Weight’, a floating-point field, which is the relative importance of detecting each lesion. This field is zero for non-diseased cases and for equally weighted lesions; otherwise the weights must sum to unity for each case. Unless a weighted figure of merit is selected, this field is irrelevant.

  2. ‘TP’ worksheet consists of
     - ‘ReaderID’, a string field uniquely labeling the readers (radiologists).
     - ‘ModalityID’, a string field uniquely labeling the modalities.
     - ‘CaseID’, see ‘Truth’ worksheet. A non-diseased case in this field will generate an error.
     - ‘LesionID’, see ‘Truth’ worksheet. An entry in this field that does not appear in the ‘Truth’ worksheet will generate an error. It is the user’s responsibility to ensure that the entries in the ‘Truth’ and ‘TP’ worksheets correspond to the same physical lesions.
- ‘TP_Rating’, a positive floating-point field denoting the rating assigned to a particular lesion-localization mark, with higher numbers represent greater confidence that the location is actually a lesion.

3. ‘FP’ worksheet consists of
- ‘ReaderID’, see ‘TP’ worksheet.
- ‘ModalityID’, see ‘TP’ worksheet.
- ‘CaseID’, see ‘TP’ worksheet.
- ‘FP_Rating’, a positive floating-point field denoting the rating assigned to a particular non-lesion-localization mark, with higher numbers represent greater confidence that the location is actually a lesion.

- **MRMC data file format / LABMRMC format**
  - *Input format for MRMC*. This format is described in the Medical Image Perception Laboratory website, currently http://perception.radiology.uiowa.edu/.
  - *LABMRMC data format*. The data file includes following parts. The file must be saved as plain text file with *.lrc extension. All items in the file are separated by one or more blank spaces.
    1. The first line is a free text description of the file.
    2. The second line is the name or ID of the first reader.
    3. The third line has the names or IDs of all the modalities. Each name or ID must be enclosed by double quotes(" ").
    4. The fourth line must have the letter (l or s) or word (large or small) for each modality. The letter or word indicates that smaller or larger ratings represent stronger confidence of presence of disease.
    5. The following lines contain the ratings in all modalities, separated by spaces or tabs, of the non-diseased cases, one case per line. The cases must appear in the same order for all readers. Missing value is not allowed.
    6. After the last non-diseased case insert a line containing the asterisk (*) symbol.
    7. Repeat steps 5 and 6 for the diseased cases.
    8. Repeat steps 2, 5, 6 and 7 for the remaining readers.
    9. The last line of the data file must be a pound symbol (#).

- **iMRMC data format** This is described in the iMRMC website, currently https://code.google.com/p/imrmc/.

**Df: Datafile Related Functions**

- **Df2RJafrocdataset**: Convert a ratings array to a dataset object.
- **DfBinDataset**: Return a binned dataset.
- **DfCreateCorCbmDataset**: Create paired dataset for testing FitCorCbm.
- **DfExtractDataset**: Extract a subset of modalities and readers from a dataset.
- **DffrocRafroc**: Convert an FROC dataset to an AFROC dataset.
- **DffrocRroc**: Convert an FROC dataset to a highest rating inferred ROC dataset.
- **DflrocRroc**: Convert an LROC dataset to a highest rating inferred ROC dataset.
- **DfReadCrossedModalities**: Read a crossed-modalities data file.
- **DfReadDataFile**: Read a general data file.
- **DfReadLrocDataFile**: Read a LROC data file.
- **DfSaveDataFile**: Save ROC data file in a different format.
- **DfExtractCorCbmDataset**: Extract two arms of a pairing from an MRMC ROC dataset suitable for using FitCorCbm.

### Fitting Functions

- **FitBinormalRoc**: Fit the binormal model to ROC data (R equivalent of ROCFIT or RSCORE).
- **FitCbmRoc**: Fit the contaminated binormal model (CBM) to ROC data.
- **FitRsmRoc**: Fit the radiological search model (RSM) to ROC data.
- **FitCorCbm**: Fit the correlated contaminated binormal model (CORCBM) to paired ROC data.
- **FitRsmRoc**: Fit the radiological search model (RSM) to ROC data.

### Plotting Functions

- **PlotBinormalFit**: Plot binormal-predicted ROC curve with provided BM parameters.
- **PlotEmpiricalOperatingCharacteristics**: Plot empirical operating characteristics for specified dataset.
- **PlotRsmOperatingCharacteristics**: Plot RSM-fitted ROC curves.

### Simulation Functions

- **SimulateFrocDataset**: Simulates an uncorrelated FROC dataset using the RSM.
- **SimulateRocDataset**: Simulates an uncorrelated binormal model ROC dataset.
- **SimulateCorCbmDataset**: Simulates an uncorrelated binormal model ROC dataset.

### Sample size Functions

- **SSPowerGivenJK**: Calculate statistical power given numbers of readers J and cases K.
- **SSPowerTable**: Generate a power table.
- **SSSampleSizeKGivenJ**: Calculate number of cases K, for specified number of readers J, to achieve desired power for an ROC study.

### Significance Testing Functions

- **StSignificanceTesting**: Perform significance testing, DBM or OR.
- **StSignificanceTestingCadVsRadiologists**: Perform significance testing, CAD vs. radiologists.
- **StSignificanceTestingCrossedModalities**: Perform significance testing using crossed modalities analysis.
- **StSignificanceTestingSingleFixedFactor**: Perform significance testing for single fixed factor analysis.
Miscellaneous and Utility Functions

- **Compare3ProperRocFits**: Compare three proper-ROC curve fitting models.
- **UtilAucBinormal**: Binormal model AUC function.
- **UtilAucCbm**: CBM AUC function.
- **UtilAucProproc**: PROPROC AUC function.
- **UtilAucRsm**: RSM ROC/AFROC AUC calculator.
- **UtilFigureOfMerit**: Calculate empirical figures of merit (FOMs) for specified dataset.
- **UtilIntrinsic2PhysicalRsm**: Convert from intrinsic to physical RSM parameters.
- **UtilLesionDistribution**: Calculates the lesion distribution matrix.
- **UtilLesionWeights**: Calculates the lesion weights matrix.
- **UtilMeanSquares**: Calculates the mean squares used in the DBMH and ORH methods.
- **UtilOutputReport**: Generate a formatted report file.
- **UtilPhysical2IntrinsicRsm**: Convert from physical to intrinsic RSM parameters.
- **UtilPseudoValues**: Return jackknife pseudovalues.

Author(s)


References

Basics of ROC


DBM/OR methods and extensions


Obuchowski, NA, & Rockette, HE (1994). HYPOTHESIS TESTING OF DIAGNOSTIC ACCURACY FOR MULTIPLE READERS AND MULTIPLE TESTS: AN ANOVA APPROACH


**FROC paradigm**


Chakraborty DP, Nishikawa RM, Orton CG. Due to potential concerns of bias and conflicts of interest, regulatory bodies should not do evaluation methodology research related to their regulatory missions. Medical Physics. 2017.


Description

A binned dataset suitable for analysis by `FitCorCbm`. It was generated by `DfCreateCorCbmDataset` by setting the `seed` variable to 123. Note the formatting of the data as a single treatment two reader dataset, even though the actual pairing might be different, see `FitCorCbm`. The dataset is intentionally large so as to demonstrate the asymptotic convergence of ML estimates, produced by `FitCorCbm`, to the population values. The data was generated by the following argument values to `DfCreateCorCbmDataset`: `seed = 123`, `K1 = 5000`, `K2 = 5000`, `desiredNumBins = 5`, `muX = 1.5`, `muY = 3`, `alphaX = 0.4`, `alphaY = 0.7`, `rhoNor = 0.3`, `rhoAbn2 = 0.8`.

Usage

`binnedData123`

Format

A list with 8 elements:

- `NL` Ratings array [1, 1:2, 1:10000, 1], of non-lesion localizations, NLs
- `LL` Ratings array [1, 1:2, 1:5000, 1], of lesion localizations, LLs
- `lesionNum` array [1:5000], number of lesions per diseased case, all set to one
- `lesionID` array [1:5000, 1], lesions labels on diseased cases, all set to one
- `lesionWeight` array [1:5000, 1], weights, all set to one
- `dataType` "ROC", the data type
- `modalityID"1", treatment label
- `readerID` [1:2] "1" "2", reader labels

References


Examples

`str(binnedData123)`
binnedData124  

_Binned dataset suitable for checking FitCorCbm; seed = 124_

### Description

A binned dataset suitable for analysis by FitCorCbm. It was generated by DfCreateCorCbmDataset by setting the seed variable to 124. Otherwise similar to binnedData123.

### Usage

binnedData124

### Format

A list with 8 elements:

- NL Ratings array [1, 1:2, 1:10000, 1], of non-lesion localizations, NLs
- LL Ratings array [1, 1:2, 1:5000, 1], of lesion localizations, LLs
- lesionNum array [1:5000], number of lesions per diseased case, all set to one
- lesionID array [1:5000, 1], lesions labels on diseased cases, all set to one
- lesionWeight array [1:5000, 1], weights, all set to one
- dataType "ROC", the data type
- modalityID "1", treatment label
- readerID [1:2] "1" "2", reader labels

### References


### Examples

str(binnedData124)
Description

A binned dataset suitable for analysis by `FitCorCbm`. It was generated by `DfCreateCorCbmDataset` by setting the `seed` variable to 125. Otherwise similar to `binnedData123`.

Usage

`binnedData125`

Format

A list with 8 elements:

- `nL` Ratings array [1, 1:2, 1:10000, 1], of non-lesion localizations, NLs
- `ll` Ratings array [1, 1:2, 1:5000, 1], of lesion localizations, LLs
- `lesionNum` array [1:5000], number of lesions per diseased case, all set to one
- `lesionID` array [1:5000, 1], lesions labels on diseased cases, all set to one
- `lesionWeight` array [1:5000, 1], weights, all set to one
- `dataType` "ROC", the data type
- `modalityID"1", treatment label
- `readerID` [1:2] "1" "2", reader labels

References


Examples

`str(binnedData125)`
ChisqRGoodnessOfFit  Compute the chisquare goodness of fit statistic for ROC fitting model

Description

Compute the chisquare goodness of fit statistic for specified ROC data fitting model

Usage

ChisqRGoodnessOfFit(fpCounts, tpCounts, parameters, model, lesDistr)

Arguments

fpCounts  The FP counts table
.tpCounts  The TP counts table
.parameters  The parameters of the model including cutoffs, see details
.model  The fitting model: "BINORMAL", "CBM" or "RSM"
.lesDistr  The lesion distribution matrix; not needed for "BINORMAL" or "CBM" models

Details

For model = "BINORMAL" the parameters are c(a,b,zetas). For model = "CBM" the parameters are c(mu,alpha,zetas). For model = "RSM" the parameters are c(mu,lambdaP,nuP,zetas).

Value

The return value is a list with the following elements:

chisq  The chi-square statistic
pVal  The p-value of the fit
df  The degrees of freedom

Compare3ProperRocFits  Compare three proper-ROC curve fitting models

Description

Applies the Radiological Search Model (RSM) and the Contaminated Binormal Model (CBM) ROC-curve fitting methods to 14 datasets and compares the fits to Proper ROC (PROPROC) fits obtained using Windows software downloaded from the Univ. of Iowa ROC website ca. June 2017.

Usage

Compare3ProperRocFits(startIndx = 1, endIndx = 14,
showPlot = FALSE, saveProprocLrcFile = FALSE, reAnalyze = FALSE)
Arguments

startIndx  
An integer in the range 1 to 14.

endIndx  
An integer in the range 1 to 14, greater than or equal to startIndx.

showPlot  
If TRUE the three plots are shown along with 95 percent confidence intervals on the lowest and uppermost operating points. The default is FALSE.

saveProprocLrcFile  
If TRUE the binned datasets are saved for subsequent analysis using other ROC software, e.g., Windows DBM-MRMC. The default is FALSE.

reAnalyze  
If TRUE the data is reanalyzed. The default is FALSE in which case the previously saved results are used.

Details

allDatasetsResults is a list-array of length (endIndx - startIndx + 1), where each element of the list-array is a list with 10 elements.

- allDatasetsResults[[1]][[1]] parameters of treatment 1 reader 1 in dataset startIndx
- allDatasetsResults[[1]][[2]] parameters of treatment 1 reader 2 in dataset startIndx
- allDatasetsResults[[1]][[I]] parameters of treatment I reader J in dataset startIndx
- allDatasetsResults[[2]][[1]] parameters of treatment 1 reader 1 in dataset startIndx+1
- allDatasetsResults[[2]][[2]] parameters of treatment 1 reader 2 in dataset startIndx+1
- allDatasetsResults[[2]][[I]] parameters of treatment I reader J in dataset startIndx+1
- allBinnedDatasets[[1]] binned ROC dataset corresponding to dataset startIndx
- allDatasetsResults[[2]][[I]] binned ROC dataset corresponding to dataset startIndx+1

A specific member, e.g., allDatasetsResults[[1]][[1]], has the following structure:

- retRsm The RSM parameters following the output structure of FitRsmRoc
- retCbm The CBM parameters following the output structure of FitCbmRoc
- lesDistr The lesion distribution matrix
- c1 The c-parameter of PROPROC
- da The d_sub_a parameter of PROPROC
- aucProp The PROPROC AUC
- I The number of treatments
- J The number of readers
- K1 The number of non-diseased cases
- K2 The number of diseased cases

The PROPROC parameters were obtained by running Windows software OR DBM-MRMC 2.50 (Sept. 04, 2014, Build 4) with PROPROC and area selected. The RSM and CBM fits are implemented in this package. Chapter 19 of the author's book has further details. If saveProprocLrcFile is TRUE, the .lrc files will be written to the inst-MRMCRuns directory, to appropriate subdirectory, overwriting any existing files.
Value

The returned value is a list of 2: `allDatasetsResults` containing the fitting results and `allBinnedDatasets` containing the binned datasets used in the fitting. See details.

References


Examples

```r
## Not run:
ret <- compare3ProperRocFits(1, 2, reAnalyze = TRUE) # analyze first two datasets
x <- ret$allDatasetsResults
str(x[[1]][[1]]) # parameters for dataset 1 trt 1 and rdr 1
str(x[[1]][[2]]) # parameters for dataset 1 trt 1 and rdr 2
str(x[[1]][[10]]) # parameters for dataset 1 trt 2 and rdr 5
str(x[[1]][[11]]) # error
str(x[[2]][[1]]) # parameters for dataset 2 trt 1 and rdr 1
str(x[[2]][[2]]) # parameters for dataset 2 trt 1 and rdr 2
str(x[[2]][[10]]) # parameters for dataset 2 trt 2 and rdr 5
str(x[[3]][[1]]) # error

## End(Not run)
```

dataset01

*TONY FROC dataset*

Description

This is referred to in the book as the "TONY" dataset. It consists of 185 cases, 89 of which are diseased, interpreted in two treatments ("BT" = breast tomosynthesis and "DM" = digital mammography) by five radiologists using the FROC paradigm. The diseased cases had at most two cancers (lesions) per case while the maximum number of non-lesion localizations (NLs) per case, over the entire dataset, was 3. The example below displays the wAFROC plot for the first treatment and first reader.
Usage

dataset01

Format

A list with 8 elements:

- NL Ratings array [1:2, 1:5, 1:185, 1:3], of non-lesion localizations, NLs
- LL Ratings array [1:2, 1:5, 1:89, 1:2], of lesion localizations, LLs
- lesionNum array [1:89], number of lesions per diseased case
- lesionID array [1:89, 1:2], labels of lesions on diseased cases
- lesionWeight array [1:89, 1:2], weights (or clinical importance) of lesions
- dataType "FROC", the data type
- modalityID [1:2] "BT" "DM", treatment labels
- readerID [1:5] "1" "2" "3" "4" ..., reader labels

References


Examples

str(dataset01)
PlotEmpiricalOperatingCharacteristics(dataset = dataset01, opChType = "wAFROC")$Plot

Dataset

**Van Dyke ROC dataset**

Description

This is referred to in the book as the "VD" dataset. It consists of 114 cases, 45 of which are diseased, interpreted in two treatments ("0" = single spin echo MRI, "1" = cine-MRI) by five radiologists using the ROC paradigm. Each diseased cases had an aortic dissection; the ROC paradigm generates one rating per case. Often referred to in the ROC literature as the Van Dyke dataset, which, along with the Franken dataset, has been widely used to illustrate advances in ROC methodology. The example below displays the ROC plot for the first treatment and first reader.

Usage

dataset02
Format

A list with 8 elements:

- NL Ratings array [1:2, 1:5, 1:114, 1], of false positives, FPs
- LL Ratings array [1:2, 1:5, 1:45, 1], of true positives, TPs
- lesionNum array [1:45], number of lesions per diseased case, all set to 1
- lesionID array [1:45, 1], labels of lesions on diseased cases, all set to 1
- lesionWeight array [1:45, 1], weights (or clinical importance) of lesions, all set to 1
- dataType "ROC", the data type
- modalityID [1:2] "0" "1", treatment labels
- readerID [1:5] "0" "1" "2" ..., reader labels

References


Examples

str(dataset02)
PlotEmpiricalOperatingCharacteristics(dataset = dataset02)$Plot

dataset03

Franken ROC dataset

Description

This is referred to in the book as the "FR" dataset. It consists of 100 cases, 69 of which are diseased, interpreted in two treatments, "0" = conventional film radiographs, "1" = digitized images viewed on monitors, by four radiologists using the ROC paradigm. Often referred to in the ROC literature as the Franken-dataset, which, along the the Van Dyke dataset, has been widely used to illustrate advances in ROC methodology.

Usage

dataset03
Format

A list with 8 elements:

- **NL** Ratings array [1:2, 1:4, 1:100, 1], of false positives, FPs
- **LL** Ratings array [1:2, 1:4, 1:67, 1], of true positives, TPs
- **lesionNum** array [1:67], number of lesions per diseased case, all set to 1
- **lesionID** array [1:67, 1], labels of lesions on diseased cases, all set to 1
- **lesionWeight** array [1:67, 1], weights (or clinical importance) of lesions, all set to 1
- **dataType** "ROC", the data type
- **modalityID** [1:2] "0" "1", treatment labels
- **readerID** [1:4] "0" "1" "2" ..., reader labels

References


Examples

```r
str(dataset04)
PlotEmpiricalOperatingCharacteristics(dataset = dataset04)$Plot
```

Description

This is referred to in the book as the "FED" dataset. It consists of 200 mammograms, 100 of which contained one to 3 simulated microcalcifications, interpreted in five treatments (basically different image processing algorithms), by four radiologists using the FROC paradigm. The maximum number of NLs per case, over the entire dataset was 7 and the dataset contained at least one diseased mammogram with 3 lesions.

Usage

dataset04
Format

A list with 8 elements:

- **NL** Ratings array [1:5, 1:4, 1:200, 1:7], of non-lesion localizations, NLs
- **LL** Ratings array [1:5, 1:4, 1:100, 1:3], of lesion localizations, LLs
- **lesionNum** array [1:100], number of lesions per diseased case
- **lesionID** array [1:100, 1:3], labels of lesions on diseased cases, all set to 1
- **lesionWeight** array [1:100, 1:3] weights (or clinical importance) of lesions, all set to 1
- **dataType** "FROC", the data type
- **modalityID** [1:5] "1" "2" ... treatment labels
- **readerID** [1:4] "1" "3" "4" "5" reader labels

References


Examples

```
str(dataset05)
PlotEmpiricalOperatingCharacteristics(dataset = dataset05, opChType = "wAFROC")$Plot
```

---

**dataset05**

*John Thompson FROC dataset*

Description

This is referred to in the book as the "JT" dataset. It consists of 92 cases, 47 of which are diseased, interpreted in two treatments ("1" = CT images acquired for attenuation correction, "2" = diagnostic CT images), by nine radiographers using the FROC paradigm. Each case was a slice of an anthropomorphic phantom 47 with inserted nodular lesions (max 3 per slice). The maximum number of NLs per case, over the entire dataset was 7.

Usage

```
dataset05
```
dataset06

Format

A list with 8 elements:

- **NL** Ratings array [1:2, 1:9, 1:92, 1:7], of non-lesion localizations, NLs
- **LL** Ratings array [1:2, 1:9, 1:47, 1:3], of lesion localizations, LLs
- **lesionNum** array [1:47], number of lesions per diseased case
- **lesionID** array [1:47, 1:3], labels of lesions on diseased cases, all set to 1
- **lesionWeight** array [1:67, 1] weights (or clinical importance) of lesions, all set to 1
- **dataType** "FROC", the data type
- **modalityID** [1:2] "1" "2", treatment labels
- **readerID** [1:4] "1" "2" "3" "4", reader labels

References


Examples

```
str(dataset06)
PlotEmpiricalOperatingCharacteristics(dataset = dataset06, opChType = "wAFROC")$Plot
```

dataset06

**Magnus FROC dataset**

Description

This is referred to in the book as the "MAG" dataset (after Magnus Bath, who conducted the JAFROC analysis). It consists of 100 cases, 69 of which are diseased, interpreted in two treatments ("1" = conventional chest, "1" = chest tomosynthesis) by four radiologists using the FROC paradigm.

Usage

dataset06
Format

A list with 8 elements:

- **nl** Ratings array [1:2, 1:4, 1:89, 1:17], of non-lesion localizations, NLs
- **ll** Ratings array [1:2, 1:4, 1:42, 1:15], of lesion localizations, LLs
- **lesionNum** array [1:42], number of lesions per diseased case
- **lesionID** array [1:42, 1:15], labels of lesions on diseased cases, all set to 1
- **lesionWeight** array [1:42, 1:15] weights (or clinical importance) of lesions, all set to 1
- **dataType** "FROC", the data type
- **modalityID** [1:2] "1" "2", treatment labels
- **readerID** [1:4] "1" "2" ..., reader labels

References


Examples

```r
str(dataset07)
PlotEmpiricalOperatingCharacteristics(dataset = dataset07, opChType = "wAFROC")$Plot
```

---

dataset07  
*Lucy Warren FROC dataset*

Description

This is referred to in the book as the "OPT" dataset (for OptiMam). It consists of 162 cases, 81 of which are diseased, interpreted in five treatments (see reference, basically different ways of acquiring the images) by seven radiologists using the FROC paradigm.

Usage

dataset07

Format

A list with 8 elements:

- **nl** Ratings array [1:5, 1:7, 1:162, 1:4], of non-lesion localizations, NLs
- **ll** Ratings array [1:5, 1:7, 1:81, 1:3], of lesion localizations, LLs
- **lesionNum** array [1:81], number of lesions per diseased case, all set to 1
- lesionID array [1:81, 1:3], labels of lesions on diseased cases, all set to 1
- lesionWeight array [1:81, 1:3] weights (or clinical importance) of lesions, all set to 1
- dataType "FROC", the data type
- modalityID [1:5] "1" "2" ..., treatment labels
- readerID [1:7] "1" "2" ..., reader labels

References

Warren LM, Mackenzie A, Cooke J, et al. Effect of image quality on calcification detection in

Examples

def str(dataset08)
def PlotEmpiricalOperatingCharacteristics(dataset = dataset07, opChType = "wAFROC")

dataset08

Monica Penedo ROC dataset

Description

This is referred to in the book as the "PEN" dataset. It consists of 112 cases, 64 of which are
diseased, interpreted in five treatments (basically different image compression algorithms) by five
radiologists using the FROC paradigm (the inferred ROC dataset is included; the original FROC
data is lost).

Usage

dataset08

Format

A list with 8 elements:

- NL Ratings array [1:5, 1:5, 1:112, 1], of false positives, FPs
- LL Ratings array [1:5, 1:5, 1:64, 1], of true positives, TPs
- lesionNum array [1:64], number of lesions per diseased case, all set to 1
- lesionID array [1:64, 1], labels of lesions on diseased cases, all set to 1
- lesionWeight array [1:64, 1], weights (or clinical importance) of lesions, all set to 1
- dataType "ROC", the data type
- modalityID [1:5] "0" "1", treatment labels
- readerID [1:5] "0" "1" "2" ..., reader labels
References


Examples

\begin{verbatim}
str(dataset09)
PlotEmpiricalOperatingCharacteristics(dataset = dataset09)$Plot
\end{verbatim}

---

**dataset09**  
*Nico Karssemeijer ROC dataset (CAD vs. radiologists)*

Description

This is referred to in the book as the "NICO" dataset. It consists of 200 mammograms, 80 of which contain one malignant mass, interpreted by a CAD system and nine radiologists using the LROC paradigm. The first reader is CAD. The highest rating method was used to convert this to an ROC dataset. The original LROC data is *datasetCadLroc*. Analyzing this **one-treatment** data requires methods described in the book, specifically, the function *StSignificanceTestingSingleFixedFactor* analyzes such datasets.

Usage

dataset09

Format

A list with 8 elements:

- `nl` Ratings array [1, 1:10, 1:200, 1], of false positives, FPs
- `ll` Ratings array [1, 1:10, 1:80, 1], of true positives, TPs
- `lesionNum` array [1:80], number of lesions per diseased case, all set to 1
- `lesionID` array [1:80, 1], labels of lesions on diseased cases, all set to 1
- `lesionWeight` array [1:80, 1], weights (or clinical importance) of lesions, all set to 1
- `dataType` "ROC", the data type
- `modalityID` [1] "1" treatment label
- `readerID` [1:10] "1" "2" ..., reader labels

References

dataset10

**Examples**

```r
str(dataset09)
PlotEmpiricalOperatingCharacteristics(dataset = dataset09, rdrs = 1:10)$Plot
```

dataset10

**Marc Ruschin ROC dataset**

**Description**

This is referred to in the book as the "RUS" dataset. It consists of 90 cases, 40 of which are diseased, the images were acquired at three dose levels, which can be regarded as treatments. "0" = conventional film radiographs, "1" = digitized images viewed on monitors. Eight radiologists interpreted the cases using the FROC paradigm. These have been reduced to ROC data by using the highest ratings (the original FROC data is lost).

**Usage**

dataset10

**Format**

A list with 8 elements:

- `nl` Ratings array [1:3, 1:8, 1:90, 1], of false positives, FPs
- `ll` Ratings array [1:3, 1:8, 1:40, 1], of true positives, TPs
- `lesionNum` array [1:40], number of lesions per diseased case, all set to 1
- `lesionID` array [1:40, 1], labels of lesions on diseased cases, all set to 1
- `lesionWeight` array [1:40, 1], weights (or clinical importance) of lesions, all set to 1
- `dataType` "ROC", the data type
- `modalityID` [1:3] "1" "2" "3", treatment labels
- `readerID` [1:8] "1" "2" ... reader labels

**References**


**Examples**

```r
str(dataset10)
PlotEmpiricalOperatingCharacteristics(dataset = dataset10)$Plot
```
**Description**

This is referred to in the book as the "DOB1" dataset. Dobbins et al conducted a multi-institutional, MRMC study to compare the performance of digital tomosynthesis (GE's VolumeRad device), dual-energy (DE) imaging, and conventional chest radiography for pulmonary nodule detection and management. All study images were obtained with a flat-panel detector developed by GE. The case set consisted of 158 subjects, of which 43 were non-diseased and the rest had 1 - 20 pulmonary nodules independently verified, using with CT images, by 3 experts who did not participate in the observer study. The study used FROC paradigm data collection. There are 4 treatments labeled 1 - 4 (conventional chest x-ray, CXR, CXR augmented with dual-energy (CXR+DE), VolumeRad digital tomosynthesis images and VolumeRad augmented with DE (VolumeRad+DE)).

**Usage**

dataset11

**Format**

A list with 8 elements:

- **NL** Ratings array [1:4, 1:5, 1:158, 1:4], of non-lesion localizations, NLs
- **LL** Ratings array [1:4, 1:5, 1:115, 1:20], of lesion localizations, LLs
- **lesionNum** array [1:115], number of lesions per diseased case
- **lesionID** array [1:115, 20], labels of lesions on diseased cases, all set to 1
- **lesionWeight** array [1:115, 20] weights (or clinical importance) of lesions, all set to 1
- **dataType** "FROC", the data type
- **modalityID** [1:4] "1" "2" ...., treatment labels
- **readerID** [1:5] "1" "2" ...., reader labels

**References**


**Examples**

`str(dataset11)`
**dataset12**

**Dobbins 2 ROC dataset**

**Description**

This is referred to in the code as the "DOB2" dataset. It contains actionability ratings, i.e., do you recommend further follow up on the patient, one a 1 (definitely not) to 5 (definitely yes), effectively an ROC dataset using a 5-point rating scale.

**Usage**

dataset12

**Format**

A list with 8 elements:

- `nl` Ratings array [1:4, 1:5, 1:152, 1], of false positives, FPs
- `ll` Ratings array [1:4, 1:5, 1:88, 1], of true positives, TPs
- `lesionNum` array [1:88], number of lesions per diseased case, all set to 1
- `lesionID` array [1:88, 1], labels of lesions on diseased cases, all set to 1
- `lesionWeight` array [1:88, 1], weights (or clinical importance) of lesions, all set to 1
- `dataType` "ROC", the data type
- `modalityID` [1:2] "0" "1", treatment labels
- `readerID` [1:4] "0" "1" "2" ..., reader labels

**References**


**Examples**

```r
str(dataset12)
```
dataset13

Dobbins 3 FROC dataset

Description

This is referred to in the code as the "DOB3" dataset. This is a subset of DOB1 which includes data for lesions not-visible on CXR, but visible to truth panel on all treatments.

Usage

dataset13

Format

A list with 8 elements:

- NL Ratings array [1:4, 1:5, 1:158, 1:4], of non-lesion localizations, NLs
- LL Ratings array [1:4, 1:5, 1:106, 1:15], of lesion localizations, LLs
- lesionNum array [1:106], number of lesions per diseased case, all set to 1
- lesionID array [1:106, 15], labels of lesions on diseased cases, all set to 1
- lesionWeight array [1:106, 15] weights (or clinical importance) of lesions, all set to 1
- dataType "FROC", the data type
- modalityID [1:4] "1" "2" ..., treatment labels
- readerID [1:5] "1" "2" ..., reader labels

References


Examples

str(dataset13)
**dataset14**

**Federica Zanca real (as opposed to inferred) ROC dataset**

**Description**

This is referred to in the book as the "FZR" dataset. It is a real ROC study, conducted on the same images and using the same radiologists, on treatments "4" and "5" of dataset04. This was compared to highest rating inferred ROC data from dataset04 to conclude, erroneously, that the highest rating assumption is invalid. See book Section 13.6.2.

**Usage**

`dataset14`

**Format**

A list with 8 elements:

- `nl` Ratings array [1:2, 1:4, 1:200, 1], of false positives, FPs
- `ll` Ratings array [1:2, 1:4, 1:100, 1], of true positives, TPs
- `lesionNum` array [1:100], number of lesions per diseased case, all set to 1
- `lesionID` array [1:100, 1], labels of lesions on diseased cases, all set to 1
- `lesionWeight` array [1:100, 1], weights (or clinical importance) of lesions, all set to 1
- `dataType` "ROC", the data type
- `modalityID` [1:2] "1" "2", treatment labels
- `readerID` [1:4] "1" "2" ...., reader labels

**References**


**Examples**

`str(dataset14)`
datasetCadLroc  

_Nico Karssemeijer LROC dataset (CAD vs. radiologists)_

**Description**

This is the actual LROC data corresponding to `datasetPY`, which was the inferred ROC data. Note that the `ll` field is split into two, `llcl`, representing true positives where the lesions were correctly localized, and `llil`, representing true positives where the lesions were incorrectly localized. The first reader is CAD and the remaining readers are radiologists. The function `StSignificanceTestingSingleFixedFactor` analyzes such datasets.

**Usage**

`datasetCadLroc`

**Format**

A list with 9 elements:

- `nl` Ratings array `[1, 1:10, 1:200, 1]`, of false positives, FPs
- `llcl` Ratings array `[1, 1:10, 1:80, 1]`, of true positives with correct localization, TPCls
- `llil` Ratings array `[1, 1:10, 1:80, 1]`, of true positives with incorrect localization, TPIls
- `lesionNum` array `[1:80]`, number of lesions per diseased case, all set to 1
- `lesionID` array `[1:80, 1]`, labels of lesions on diseased cases, all set to 1
- `lesionWeight` array `[1:80, 1]`, weights (or clinical importance) of lesions, all set to 1
- `dataType` "LROC", the data type
- `modalityID` [1:2] "0" "1", treatment labels
- `readerID` [1:10] "1" "2" ...., reader labels

**References**


**Examples**

`str(datasetCadLroc)`
**datasetCrossedModality**

*John Thompson crossed treatment FROC dataset*

**Description**

This is a crossed treatment dataset, see book Section 18.5. There are two treatment factors. The first treatment factor `modalityID1` can be "F" or "I", which represent two CT reconstruction algorithms. The second treatment factor `modalityID2` can be "20" "40" "60" "80", which represent the mAs values of the image acquisition. The factors are fully crossed. The function `StSignificanceTestingCrossedModalities` analyzes such datasets.

**Usage**

```r
datasetCrossedModality
```

**Format**

A list with 9 elements:

- `NL` Ratings array `[1:2, 1:4, 1:11, 1:68, 1:5]`, of non-lesion localizations, NLs
- `LL` Ratings array `[1:2, 1:4, 1:11, 1:34, 1:3]`, of lesion localizations, LLs
- `lesionNum` array `[1:34]`, number of lesions per diseased case, all set to 1
- `lesionID` array `[1:34, 3]`, labels of lesions on diseased cases, all set to 1
- `lesionWeight` array `[1:34, 3]` weights (or clinical importance) of lesions, all set to 1
- `dataType` "FROC", the data type
- `modalityID1` `[1:2] "F" "I", treatment labels
- `modalityID2` `[1:4] "20" "40" "60" "80", treatment labels
- `readerID` `[1:11] "1" "10" "11" ..., reader labels

**References**


**Examples**

```r
str(datasetCrossedModality)
```
**datasetDegenerate**  
*Simulated degenerate ROC dataset; for testing purposes*

**Description**

A simulated degenerated dataset. A degenerate dataset is defined as one with no interior operating points on the ROC plot. Such data tend to be observed with expert level radiologists. This dataset is used to illustrate the robustness of two fitting models, namely CBM and RSM. The widely used binormal model and PROPROC fail on such datasets.

**Usage**

`datasetDegenerate`

**Format**

A list with 8 elements:

- **NL** Ratings array [1, 1, 1:15, 1], of false positives, FPs
- **LL** Ratings array [1, 1, 1:10, 1], of true positives, TPs
- **lesionNum** array [1:10], number of lesions per diseased case, all set to 1
- **lesionID** array [1:10, 1], labels of lesions on diseased cases, all set to 1
- **lesionWeight** array [1:10, 1], weights (or clinical importance) of lesions, all set to 1
- **dataType** "ROC", the data type
- **modalityID** "1", treatment label
- **readerID** "1", reader label

**Examples**

```r
str(datasetDegenerate)
```

---

**Df2RJafrocDataset**  
*Convert ratings arrays to an RJafroc dataset*

**Description**

Converts ratings arrays, ROC or FROC, not LROC, to an RJafroc dataset, thereby allowing the user to leverage the file I/O, plotting and analyses capabilities of RJafroc.

**Usage**

`Df2RJafrocDataset (NL, LL, ...)`
**Arguments**

- **NL**: Non-lesion localizations array (or FP array for ROC data).
- **LL**: Lesion localizations array (or TP array for ROC data).
- **...**: Other elements of RJafroc dataset that may, depending on the context, need to be specified. lesionNum must be specified if an FROC dataset is to be returned. It is a K2-length array specifying the numbers of lesions in each diseased case in the dataset.

**Details**

The function "senses" the data type (ROC or FROC) from the the absence or presence of lesionNum. ROC data can be NL[1:K1] and LL[1:K2] or NL[1:I,1:J,1:K1] and LL[1:I,1:J,1:K2]. FROC data can be NL[1:I1,K1,1:maxNL] and LL[1:K2,1:maxLL] or NL[1:I,1:J,1:K1,1:maxNL] and LL[1:I,1:J,1:K2,1:maxLL]. Here maxNL/maxLL = maximum numbers of NLs/LLs, per case, over entire dataset. Equal weights are assigned to every lesion (FROC data). Consecutive characters/integers starting from "I" are assigned to lesionID, modalityID and readerID.

**Value**

A dataset with the structure described in RJafroc-package.

**Examples**

```r
set.seed(1)
NL <- rnorm(5)
LL <- rnorm(7)*1.5 + 2
dataset <- Df2RJafrocDataset(NL, LL) # an ROC dataset

I <- 2; J <- 3; set.seed(1)
K1 <- 25; K2 <- 35
z1 <- array(dim = c(I, J, K1))
z2 <- array(dim = c(I, J, K2))
mu <- 2; sigma <- 1.5
for (i in 1:I) {
  for (j in 1:J) {
    z1[i,j,] <- rnorm(K1)
    z2[i,j,] <- rnorm(K2) * sigma + mu
  }
}
dataset <- Df2RJafrocDataset(z1, z2) ## note lesionNum consists of 1s; i.e., an ROC dataset

set.seed(1)
mu <- 1; lambda <- 1; nu <- 1; zeta1 <- 0
K1 <- 5; K2 <- 7
Lmax <- 2; Lk2 <- floor(runif(K2, 1, Lmax + 1))
frocdataraw <- simulateFrocdataset(mu, lambda, nu, zeta1, I = 1, J = 1, K1, K2, lesionNum = Lk2)
NL <- drop(frocdataraw$NL)
LL <- drop(frocdataraw$LL)
dataset <- Df2RJafrocDataset(NL, LL, lesionNum = Lk2)
```
```
## note lesionNum is not all 1s, signalling an FROC dataset

## Simulate FROC dataset, convert to dataset object, display ROC, FROC and AFROC curves
I <- 2; J <- 3; set.seed(1)
K1 <- 25; K2 <- 35
mu <- 1; nuP <- 0.8; lambdaP <- 1; zeta1 <- 0
lambda <- UtilPhysical2IntrinsicRSM(mu, lambdaP, nuP)$lambda
nu <- UtilPhysical2IntrinsicRSM(mu, lambdaP, nuP)$nu
Lmax <- 2; lmax <- floor(runif(K2, 1, Lmax + 1))
z1 <- array(-Inf, dim = c(I, J, K1 + K2, 40))
z2 <- array(-Inf, dim = c(I, J, K2, 40))
dimNL <- array(dim=c(I, J, 2))
## the last value (2) accommodates case and location indices
dimLL <- array(dim=c(I, J, 2))
for (i in 1:I) {
  for (j in 1:J) {
    frcDataRaw <- SimulateFrocDataset(mu, lambda, nu, zeta1, I = 1, J = 1, K1, K2, lesionNum = Lk2)
    dimNL[i, j,] <- dim(drop(frcDataRaw$NL))
    dimLL[i, j,] <- dim(drop(frcDataRaw$LL))
    z1[i, j, 1:dimNL[i, j, 2]] <- drop(frcDataRaw$NL) # drop the excess location indices
    z2[i, j, 1:dimLL[i, j, 2]] <- drop(frcDataRaw$LL)
  }
}
z1 <- z1[, , 1:max(dimNL[, , 2])]
z2 <- z2[, , 1:max(dimLL[, , 2])]
dataset <- Df2RJafrocdataset(z1, z2, lesionNum = Lk2)
retPlot <- PlotEmpiricalOperatingCharacteristics(dataset, trts = seq(1, I), rdrs = seq(1, J), opChType = "ROC")
print(retPlot$Plot)
retPlot <- PlotEmpiricalOperatingCharacteristics(dataset, trts = seq(1, I), rdrs = seq(1, J), opChType = "FROC")
print(retPlot$Plot)
retPlot <- PlotEmpiricalOperatingCharacteristics(dataset, trts = seq(1, I), rdrs = seq(1, J), opChType = "AFROC")
print(retPlot$Plot)
```

---

**DfBinDataset**

*Returns a binned dataset*

**Description**

Bins continuous (i.e. floating point) or quasi-continuous (e.g. integers 0-100) ratings in a dataset and returns the corresponding binned dataset in which the ratings are integers 1, 2......, with higher values representing greater confidence in presence of disease
Usage

DfBinDataset(dataset, desiredNumBins = 7, opChType)

Arguments

dataset The dataset to be binned, with structure as in RJafroc-package.
desiredNumBins The desired number of bins. The default is 7.
opChType The operating characteristic relevant to the binning operation: "ROC", "FROC", "AFROC", or "wAFROC".

Details

For small datasets the number of bins may be smaller than desiredNumBins. The algorithm needs to know the type of operating characteristic relevant to the binning operation. For ROC the bins are FP and TP counts, for FROC the bins are NL and LL counts, for AFROC the bins are FP and LL counts, and for wAFROC the bins are FP and wLL counts. Binning is generally employed prior to fitting a statistical model, e.g., maximum likelihood, to the data. This version chooses ctffs so as to maximize empirical AUC (this yields a unique choice of ctffs which gives the reader the maximum deserved credit).

Value

The binned dataset

References

Miller GA (1956) The Magical Number Seven, Plus or Minus Two: Some limits on our capacity for processing information, The Psychological Review 63, 81-97


Examples

binned <- DfBinDataset(dataset05, opChType = "ROC")
PlotEmpiricalOperatingCharacteristics(dataset05, trts= c(1,2), rdrs = seq(1,9), opChType = "ROC")$Plot
PlotEmpiricalOperatingCharacteristics(binned, trts= c(1,2), rdrs = seq(1,9), opChType = "ROC")$Plot

binned <- DfBinDataset(dataset05, opChType = "AFROC")
PlotEmpiricalOperatingCharacteristics(dataset05, trts= c(1,2), rdrs = seq(1,9), opChType = "AFROC")$Plot
PlotEmpiricalOperatingCharacteristics(binned, trts= c(1,2), rdrs = seq(1,9), opChType = "AFROC")$Plot

## Not run:
library(ggplot2)
DfCreateCorCbmDataset

Create paired dataset for testing FitCorCbm

Description

The paired dataset is generated using bivariate sampling; details are in referenced publication

Usage

DfCreateCorCbmDataset(
  seed = 123, K1 = 50,
  K2 = 50, desiredNumBins = 5,
  muX = 1.5, muY = 3,
  alphaX = 0.4, alphaY = 0.7,
  rhoNor = 0.3, rhoAbn2 = 0.8)

Arguments

seed The seed variable, default is 123; set to NULL for truly random seed
K1 The number of non-diseased cases, default is 50
K2 The number of diseased cases, default is 50
desiredNumBins The desired number of bins; default is 5
muX The CBM $\mu$ parameter in condition X
muY The CBM $\mu$ parameter in condition Y
alphaX The CBM $\alpha$ parameter in condition X
DfExtractCorCbmDataset

alphay  The CBM ‘alpha’ parameter in condition Y
rhoNor  The correlation of non-diseased case z-samples
rhoAbn2 The correlation of diseased case z-samples, when disease is visible in both conditions

Details

The ROC data is binned to 5 bins in each condition.

Value

The return value is the desired dataset, suitable for testing FitCorCbm

References


Examples

## seed <- 1
## this gives unequal numbers of bins in X and Y conditions for 50/50 dataset
dataset <- DfCreateCorCbmDataset()

## Not run:
## this takes long time!! used to show asymptotic convergence of ML estimates
dataset <- DfCreateCorCbmDataset(K1 = 5000, K2 = 5000)

## End(Not run)

DfExtractCorCbmDataset

Extract two arms of a pairing from an MRMC ROC dataset

Description

Extract a paired dataset from a larger dataset. The pairing could be two readers in the same treatment, or different readers in different treatments, or the same reader in different treatments. If necessary The data is binned to 5 bins in each condition.

Usage

DfExtractCorCbmDataset(dataset, trts, rdrs)
Arguments

- **dataset**: The original dataset from which the pairing is to be extracted.
- **trts**: A vector, maximum length 2, contains the indices of the treatment or treatments to be extracted.
- **rdrs**: A vector, maximum length 2, contains the indices of the reader or readers to be extracted.

Details

The desired pairing is contained in the vectors `trts` and `rdrs`. If either has length one, the other must have length two and the pairing is implicit. If both are length two, then the pairing is that implied by the first treatment and the second reader, which is one arm, and the other arm is that implied by the second treatment paired with the first reader. Using this method any allowed pairing can be extracted and analyzed by `FitCorCbm`. The utility of this software is in designing a ratings simulator that is statistically matched to a real dataset.

Value

A new dataset in which the number of treatments is one and the number of readers is two.

Examples

```r
# Extract the paired data corresponding to the second and third readers in the first treatment
# from the include ROC dataset
dataset11_23 <- DFExtractCorCbmDataset(dataset05, trts = 1, rdrs = c(2,3))

# Extract the paired data corresponding to the third reader in the first and second treatments
dataset12_33 <- DFExtractCorCbmDataset(dataset05, trts = c(1,2), rdrs = 3)

# Extract the data corresponding to the first reader in the first treatment paired with the data
# from the third reader in the second treatment
# (the indices are at different positions in the respective arrays)
dataset12_13 <- DFExtractCorCbmDataset(dataset05, trts = c(1,2), rdrs = c(1,3))
```

---

**DFExtractDataset**

*Extract a subset of treatments and readers from a dataset*

Description

Extract a dataset consisting of a subset of treatments/readers from a larger dataset.

Usage

`DFExtractDataset(dataset, trts, rdrs)`
**DfFroc2Afroc**

**Arguments**

- `dataset` The original dataset from which the subset is to be extracted; can be ROC, FROC or LROC
- `trts` A vector contains the indices of the treatments to be extracted. **If this parameter is not supplied, all treatments are extracted.**
- `rdrs` A vector contains the indices of the readers to be extracted. **If this parameter is not supplied, all readers are extracted.**

**Details**

**Note** that `trts` and `rdrs` are the vectors of **indices** not **IDs**. For example, if the ID of the first reader is "0", the corresponding value in `trts` should be 1 not 0.

**Value**

A new dataset containing only the specified treatments and readers that were extracted from the original dataset.

**Examples**

```r
## Extract the data corresponding to the second reader in the
## first treatment from an included ROC dataset
dataset1_2 <- DfExtractDataset(dataset05, trts = 1, rdrs = 2)

## Extract the data of the first and third reader in all
## treatment from the included ROC dataset
datasetA_123 <- DfExtractDataset(dataset05, rdrs = c(1, 3))
```

**DfFroc2Afroc** *Convert an FROC dataset to an AFROC dataset*

**Description**

Converts an FROC dataset to a AFROC dataset, where only the highest rated mark on each non-diseased case is counted and all lesion localizations are counted.

**Usage**

`DfFroc2Afroc(dataset)`

**Arguments**

- `dataset` The dataset to be converted, RJafroc-package.
Details

The first list member of the AFROC dataset is $nl$, whose third dimension has length $Hk1$ $K$ $kR1$, the total number of cases. The ratings of cases $K1$ $+$ $1$ through $K1$ $+$ $K2$ are $-\text{Inf}$. **In an AFROC dataset FPs are only possible on non-diseased cases.** The second member of the list is $ll$. Its third dimension has length $K2$, the total number of diseased cases. This is because LLs are only possible on diseased cases. The structure is shown below:

- **NL** Ratings array $[1:I, 1:J, 1:(K1+K2), 1:maxNL]$, of non-lesion localizations, NLs
- **LL** Ratings array $[1:I, 1:J, 1:K2, 1:maxLL]$, of lesion localizations, LLs
- **lesionNum** array $[1:K2]$, number of lesions per diseased case
- **lesionID** array $[1:K2, 1:maxLL]$, labels of lesions on diseased cases
- **lesionWeight** array $[1:K2, 1:maxLL]$, weights (or clinical importances) of lesions
- **dataType** "FROC", the data type
- **modalityID** $[1:I]$ inherited modality labels
- **readerID** $[1:J]$ inherited reader labels

Value

An AFROC dataset

Examples

```r
afrocdataset <- Dffroc2Rafroc(dataset)
p <- PlotEmpiricalOperatingCharacteristics(afrocdataset, trts = 1, rdrs = 1, opChType = "wAFROC")
print(p$pPlot)
str(afrocdataset)
```

---

**Dffroc2Roc**

Convert an FROC dataset to an ROC dataset

Description

Convert an FROC dataset to a highest rating inferred ROC dataset

Usage

`Dffroc2Roc(dataset)`

Arguments

- **dataset** The FROC dataset to be converted, RJafroc-package.
Details

The first member of the ROC dataset is `nl`, whose 3rd dimension has length \((K_1 + K_2)\), the total number of cases. Ratings of cases \((K_1 + 1)\) through \((K_1 + K_2)\) are \(-\text{Inf}\). **This is because in an ROC dataset FPs are only possible on non-diseased cases.** The second member of the list is `ll`. Its 3rd dimension has length \(K_2\), the number of diseased cases. **This is because TPs are only possible on diseased cases.** For each case the inferred ROC rating is the highest of all FROC ratings on that case. If a case has no marks, a finite ROC rating, guaranteed to be smaller than the rating on any marked case, is assigned to it. The structure is shown below:

- `nl` Ratings array \([1:I, 1:J, 1:(K_1+K_2), 1]\), of false positives, FPs
- `ll` Ratings array \([1:I, 1:J, 1:K_2, 1]\), of true positives, TPs
- `lesionNum` array \([1:K_2]\), number of lesions per diseased case
- `lesionID` array \([1:K_2, 1]\), labels of lesions on diseased cases
- `lesionWeight` array \([1:K_2, 1]\), weights (or clinical importances) of lesions
- `dataType` "ROC", the data type
- `modalityID` \([1:I]\) inherited modality labels
- `readerID` \([1:J]\) inherited reader labels

Value

An ROC dataset

Examples

```r
rocDataSet <- DfFroc2Roc(dataset05)
p <- PlotEmpiricalOperatingCharacteristics(rocDataSet, trts = 1, rdrs = 1)
print(p$pPlot)
str(rocDataSet)
```

```r
## in the following example, because of the smaller number of cases,
## it is easy to see the process at work:
set.seed(1);K1 <- 3;K2 <- 5
mu <- 1;nuP <- 0.5;lambdaP <- 2;zeta1 <- 0
lambda <- UtilPhysical2IntrinsicRSM(mu, lambdaP, nuP)*lambda
nu <- UtilPhysical2IntrinsicRSM(mu, lambdaP, nuP)*nu
Lmax <- 2;Lk2 <- floor(runif(K2, 1, Lmax + 1))
frocdataraw <- simulateFrocDataset(mu, lambda, nu, zeta1, I = 1, J = 1,
K1, K2, lesionNum = Lk2)
hrData <- DfFroc2Roc(frocdataraw)
print("frocdataraw$NL[1,1,] = ");print(frocdataraw$NL[1,1,])
print("hrData$NL[1,1,1:K1,] = ");print(hrData$NL[1,1,1:K1,])
print("frocdataraw$LL[1,1,] = ");print(frocdataraw$LL[1,1,])
print("hrData$LL[1,1,] = ");print(hrData$LL[1,1,])
```

```r
## following is the output
## [1] "frocdataraw$NL[1,1,] = 
## [1,] 2.4046534 0.7635935 -Inf -Inf
## [2,] -Inf -Inf -Inf -Inf
## [3,] 0.2522234 -Inf -Inf -Inf
```
DfLroc2Rroc

Convert an LROC dataset to a ROC dataset

Description

Converts an LROC dataset to a ROC dataset

Usage

DfLroc2Rroc (dataset)

Arguments

dataset The LROC dataset to be converted.

Details

The conversion is effected by taking the maximum rating on each diseased case, which could be a TPCI or a TPII, whichever has the higher rating.

Value

An ROC dataset
**Examples**

```r
dataset <- DfReadLrocDataFile()
str(dataset)
rocDataSet <- Dflroc2Roc(dataset)
str(rocDataSet)
```

---

**DfReadCrossedModalities**  
*Read a crossed-treatment data file*

---

**Description**

Read an crossed-treatment data file, in which the two treatment factors are crossed

**Usage**

```
DfReadCrossedModalities (fileName, renumber = FALSE)
```

**Arguments**

- `fileName`: A string specifying the name of the file that contains the dataset, which must be an extended-JAFROC format data file containing an additional treatment factor.
- `renumber`: If `true`, consecutive integers (starting from 1) will be used as the treatment and reader IDs. Otherwise, treatment and reader IDs in the original data file will be used. The default is `FALSE`.

**Details**

The data format is similar to the JAFROC format (see `RJafroc-package`). The notable difference is that there are two treatment factors. A sample crossed treatment file "includedCrossedModalities-Data.xlsx" is in the `inst/extdata` subdirectory of `RJafroc`.

**Value**

A dataset with the specified structure, similar to a standard `RJafroc` (see `RJafroc-package`). Because of the extra treatment factor, NL and LL are each five dimensional arrays. There are also two treatment IDS: `modalityID1` and `modalityID2`.

**References**


Examples

```r
## Not run:
crossedFileName <- system.file("extdata",
  "includedCrossedModalitiesData.xlsx", package = "RJafroc", mustWork = TRUE)
crossedData <- DfReadCrossedModalities(crossedFileName)
str(crossedData)

## End(Not run)
```

---

**DfReadDataFile**  
*Read a data file*

**Description**

Read a disk file and create a dataset object from it.

**Usage**

```r
DfReadDataFile(fileName, format = "JAFROC",
                delimiter = ",", renumber = FALSE)
```

**Arguments**

- **fileName**: A string specifying the name of the file. The file-extension must match the format specified below.
- **format**: A string specifying the format of the data in the file. It can be "JAFROC" (the default), "MRMC" or "iMRMC". For "MRMC" the format is determined by the data file extension as specified in [http://perception.radiology.uiowa.edu/](http://perception.radiology.uiowa.edu/), i.e., .csv or .txt or .lrc. For file extension .imrmc the format is described in [https://code.google.com/p/imrmc/](https://code.google.com/p/imrmc/).
- **delimiter**: The string delimiter to be used for the "MRMC" format ("," is the default), see [http://perception.radiology.uiowa.edu/](http://perception.radiology.uiowa.edu/). This parameter is not used when reading "JAFROC" or "iMRMC" data files.
- **renumber**: A logical variable: if TRUE, consecutive integers (starting from 1) will be used as the treatment and reader IDs. Otherwise, treatment and reader IDs in the original data file will be used.

**Value**

A dataset with the structure specified in RJafroc-package.
DfReadLrocDataFile

Read a LROC data file

Description

Read the Hupse-Karssemeijer LROC data file, a study comparing standalone performance of breast CAD vs. radiologists; the study actually included radiologists and residents; the following usage includes only the radiologists.

Usage

DfReadLrocDataFile (RADIOLOGISTS = TRUE)

Arguments

RADIOLOGISTS Logical; if TRUE, the default, only radiologists are analyzed otherwise all readers are analyzed

Details

The data format is similar to the JAFROC format (see RJafroc-package) with the crucial difference that there are two types of LL (TP) events: those representing correct localizations and those representing incorrect localizations. Also, every diseased case has one lesion and NLs are not possible on diseased cases. J is one plus the number of readers. The first treatment is CAD, followed by the readers.

The return value is a list with the following elements:
DfSaveDataFile

- NL [1, 1:J, 1:K1, 1] array containing the FP ratings
- LCL1 [1, 1:J, 1:K2, 1] array containing the TP correct localization ratings
- LLI1 [1, 1:J, 1:K2, 1] array containing the TP incorrect localization ratings
- lesionNum array [1:K2], as in standard JAFROC/ROC format dataset, ones
- lesionID array [1:K2], as in standard JAFROC/ROC format dataset, ones
- lesionWeight array [1:K2], weights (or clinical importances) of lesions
- dataType "LROC", the data type
- modalityID [1:I], treatment labels
- readerID [1:J], reader labels

Value

The LROC dataset.

References


Examples

radData <- DfReadLrocDataFile()
str(radData)
allData <- DfReadLrocDataFile(FALSE)
str(allData)

DfSaveDataFile

Save ROC data file in a different format

Description

Save ROC data file in a different format so it can be analyzed with alternate software

Usage

DfSaveDataFile(dataset, fileName, format = "JAFROC",
    dataDescription = paste0(deparse(substitute(dataset)), " Data File"))
fitbinormalroc

Fit the binormal model to selected treatment and reader in an ROC dataset

Description

Fit the binormal model-predicted ROC curve for an individual dataset. This is the R equivalent of ROCFIT or RSCORE

Usage

FitBinormalRoc(dataset, trt = 1, rdr = 1)

Arguments

dataset  The ROC dataset
trt      The desired treatment, default is 1
rdr      The desired reader, default is 1
Details

In the binormal model ratings (more accurately the latent decision variables) from diseased cases are sampled from $N(\alpha, 1)$ while ratings for non-diseased cases are sampled from $N(0, b^2)$. To avoid clutter error bars are only shown for the lowest and uppermost operating points. An FROC dataset is internally converted to a highest rating inferred ROC dataset. To many bins containing zero counts will cause the algorithm to fail; so be sure to bin the data appropriately to fewer bins, where each bin has at least one count.

Value

The returned value is a list with the following elements:

- **a**: The mean of the diseased distribution; the non-diseased distribution is assumed to have zero mean
- **b**: The standard deviation of the non-diseased distribution. The diseased distribution is assumed to have unit standard deviation
- **zetases**: The binormal model cutoffs, zetas or thresholds
- **AUC**: The binormal model fitted ROC-AUC
- **StdAUC**: The standard deviation of AUC
- **NLLIni**: The initial value of negative LL
- **NLLFin**: The final value of negative LL
- **ChisqrFitStats**: The chisquare goodness of fit results
- **covMat**: The covariance matrix of the parameters
- **fittedPlot**: A `ggplot2` object containing the fitted operating characteristic along with the empirical operating points. Use `print()` to display the object

References


Examples

```r
# Test with an included ROC dataset
retFit <- FitBinormalRoc(dataset02); print(retFit$fittedPlot)

# Test with an included FROC dataset; it needs to be binned
# as there are more than 5 discrete ratings levels
binned <- DfBinDataset(dataset05, desiredNumBins = 5, opChType = "ROC")
retFit <- FitBinormalRoc(binned); print(retFit$fittedPlot)

# Test with single interior point data
```
fitcbmroc

Fit the contaminated binormal model (CBM) to selected treatment and reader in an ROC dataset

Description
Fit the CBM-predicted ROC curve for specified treatment and reader.

Usage
fitcbmroc(dataset, trt = 1, rdr = 1)

Arguments
- dataset: The dataset containing the data
- trt: The desired treatment, default is 1
- rdr: The desired reader, default is 1

Details
In CBM ratings from diseased cases are sampled from a mixture distribution: (1) with integrated area $\alpha$ distributed $N(\mu_1)$ and (2) from a distribution with integrated area $1-\alpha$ distributed $N(0, 1)$. Ratings for non-diseased cases are sampled from $N(0, 1)$. The chisqrFitStats consists of a list containing the chi-square value, the p-value and the degrees of freedom.
Value

The return value is a list with the following elements:

- **mu**: The mean of the visible diseased distribution (the non-diseased) has zero mean
- **alpha**: The proportion of diseased cases where the disease is visible
- **zetas**: The cutoffs, zetas or thresholds
- **AUC**: The AUC of the fitted ROC curve
- **stdAUC**: The standard deviation of AUC
- **NLLIni**: The initial value of negative LL
- **NLLFin**: The final value of negative LL
- **ChisqrFitStats**: The chisquare goodness of fit results
- **covMat**: The covariance matrix of the parameters
- **fittedPlot**: A **ggplot2** object containing the fitted operating characteristic along with the empirical operating points. Use `print()` to display the object.

Note

This algorithm is more robust than the binormal model.

References


Examples

```r
## Test with included ROC data
retFit <- FitCbmRoc(dataset02);print(retFit$fittedPlot)

## Test with included degenerate ROC data (yes! CBM can fit such data)
retFit <- FitCbmRoc(datasetDegenerate);print(retFit$fittedPlot)

## Test with single interior point data
fp <- c(rep(1,7), rep(2, 3))
tp <- c(rep(1,5), rep(2, 5))
dataset <- DfZRJafrocDataset(fp, tp)
retFit <- FitCbmRoc(dataset);print(retFit$fittedPlot)

## Test with two interior data points
fp <- c(rep(1,7), rep(2, 5), rep(3, 3))
tp <- c(rep(1,3), rep(2, 5), rep(3, 7))
dataset <- DfZRJafrocDataset(fp, tp)
retFit <- FitCbmRoc(dataset);print(retFit$fittedPlot)

## Test with included ROC data (some bins have zero counts)
retFit <- FitCbmRoc(dataset02, 2, 1);print(retFit$fittedPlot)

## Test with TONY data for which chisqr can be calculated
```
Fit the Correlated Contaminated Binormal Model (CORCBM) to a paired ROC dataset. The ROC dataset has to be formatted as a single treatment, two-reader dataset, even though the actual pairing may be different, see details.

Usage

FitCorCbm(dataset)

Arguments

dataset A paired ROC dataset

Details

The conditions (X, Y) can be two readers interpreting images in the same treatment, the same reader interpreting images in different treatments, or different readers interpreting images in 2 different treatments. Function DfExtractCorCbmDataset can be used to construct a dataset suitable for FitCorCbm. With reference to the returned values, and assuming R bins in condition X and L bins in condition Y, FPCounts is the R x L matrix containing the counts for non-diseased cases; TPCounts is the R x L matrix containing the counts for diseased cases; muX,muY,alphaX,alphaY,rhonor,rhoa2 are the CORCBM parameters; aucX,aucY are the AUCs in the two conditions; stdAucX,stdAucY are the corresponding standard errors; stderr contains the standard errors of the parameters of the model; areaStat, areaPval,covMat are the area-statistic, the p-value and the covariance matrix of the parameters. If a parameter approaches a limit, e.g., rhonor = 0.9999, it is held constant at near the limiting value and the covariance matrix has one less dimension (along each edge) for each parameter that is held constant. The indices of the parameters held fixed are in fitCorCbmRet$fixParam.

Value

The return value is a list containing three objects:

- **fitCorCbmRet** list(FPCounts,TPCounts,muX,muY,alphaX,alphaY,rhonor,rhoa2,zetaX,zetaY,covMat,fixParam)
- **stats** list(aucX,aucY,stdAucX,stdAucY,stderr,areaStat,areaPval)
- **fittedPlot** The fitted plot with operating points, error bars, for both conditions
References


Examples

## Not run:
dataset <- DfExtractCorCbmDataset(dataset05, trts = 1, rdrs = c(4,7))
ret <- FitCorCbm(dataset)
print(ret$fitCorCbmRet)
print(ret$stats)
print(ret$fittedPlot)

ret <- FitCorCbm(binnedData123)
print(ret$fitCorCbmRet)
print(ret$stats)
print(ret$fittedPlot)

## Also try two other datasets ending with 124 and 125

## End(Not run)

---

FitRsmRoc

Fit the radiological search model (RSM) to ROC data

Description

Fit an RSM-predicted ROC curve to a binned ROC dataset

Usage

FitRsmRoc(binnedRocData, lesDistr, trt = 1, rdr = 1)

Arguments

- **binnedRocData**: The binned ROC dataset containing the data
- **lesDistr**: The lesion distribution matrix
- **trt**: The desired treatment, default is 1
- **rdr**: The desired reader, default is 1
Details

If dataset is FROC, first convert it to ROC, using DfFroc2Roc. MLE ROC algorithms require binned datasets. Use DfBinDataset to perform the binning prior to calling this function. In the RSM: (1) The (random) number of latent NLs per case is Poisson distributed with mean parameter lambdaP, and the corresponding ratings are sampled from \( N(0, 1) \). The (2) The (random) number of latent LLs per diseased case is binomial distributed with success probability nuP and trial size equal to the number of lesions in the case, and the corresponding ratings are sampled from \( N(\mu, 1) \). (3) A latent NL or LL is actually marked if its rating exceeds the lowest threshold zeta1. To avoid clutter error bars are only shown for the lowest and uppermost operating points. Because of the extra parameter, and the requirement to have five counts, the chi-square statistic often cannot be calculated.

Value

The return value is a list with the following elements:

- \( \mu \) The mean of the diseased distribution relative to the non-diseased one
- \( \lambda_{\text{dp}} \) The Poisson parameter describing the distribution of latent NLs per case
- \( \nu_{\text{p}} \) The binomial success probability describing the distribution of latent LLs per diseased case
- \( \zetas \) The RSM cutoffs, zetas or thresholds
- AUC The RSM fitted ROC-AUC
- StdAUC The standard deviation of AUC
- NLLIni The initial value of negative LL
- NLLFin The final value of negative LL
- ChisqrFitStats The chisquare goodness of fit results
- covMat The covariance matrix of the parameters
- fittedPlot A ggplot2 object containing the fitted operating characteristic along with the empirical operating points. Use print to display the object

References


Examples

```r
## Test with included ROC data (some bins have zero counts)
lesDistr <- UtilLesionDistribution(dataset02)
retFit <- FitRsmRoc(dataset02, lesDistr)
print(retFit$fittedPlot)

## Test with included degenerate ROC data
lesDistr <- UtilLesionDistribution(datasetDegenerate)
retFit <- FitRsmRoc(datasetDegenerate, lesDistr); print(retFit$fittedPlot)

## Test with single interior point data
fp <- c(rep(1,7), rep(2, 3))
tp <- c(rep(1,5), rep(2, 5))
binnedRocData <- Df2RJafrocDataset(fp, tp)
lesDistr <- UtilLesionDistribution(binnedRocData)
retFit <- FitRsmRoc(binnedRocData, lesDistr); print(retFit$fittedPlot)

## Test with two interior data points
fp <- c(rep(1,7), rep(2, 5), rep(3, 3))
tp <- c(rep(1,3), rep(2, 5), rep(3, 7))
binnedRocData <- Df2RJafrocDataset(fp, tp)
lesDistr <- UtilLesionDistribution(binnedRocData)
retFit <- FitRsmRoc(binnedRocData, lesDistr); print(retFit$fittedPlot)

## Test with three interior data points
fp <- c(rep(1,12), rep(2, 5), rep(3, 3), rep(4, 5)) #25
tp <- c(rep(1,3), rep(2, 5), rep(3, 7), rep(4, 10)) #25
binnedRocData <- Df2RJafrocDataset(fp, tp)
lesDistr <- UtilLesionDistribution(binnedRocData)
retFit <- FitRsmRoc(binnedRocData, lesDistr); print(retFit$fittedPlot)

## test for TONY data, i = 2 and j = 3; only case permitting chisquare calculation
lesDistr <- UtilLesionDistribution(dataset01)
rocData <- DfRroc2Roc(dataset01)
retFit <- FitRsmRoc(rocData, lesDistr, trt = 2, rdr = 3)
print(retFit$fittedPlot)
retFit$ChisqrFitStats
```

PlotBinormalFit

## Plot binormal fit

### Description

Plot the binormal-predicted ROC curve with provided parameters

### Usage

`PlotBinormalFit(a, b)`
Arguments

\( a \) vector: the mean(s) of the diseased distribution(s).

\( b \) vector: the standard deviations(s) of the diseased distribution(s).

Details

\( a \) and \( b \) must have the same length. The predicted ROC curve for each \( a \) and \( b \) pair will be plotted.

Value

A \texttt{ggplot2} object of the plotted ROC curve(s) are returned. Use \texttt{print} function to display the saved object.

Examples

```r
binormalPlot <- PlotBinormalFit(c(1, 2), c(0.5, 0.5))
print(binormalPlot)
```

---

PlotCbmFit

\textit{Plot CBM fitted curve}

Description

Plot the CBM-predicted ROC curve with provided CBM parameters

Usage

\texttt{PlotCbmFit(mu, alpha)}

Arguments

\( \mu \) vector: the mean(s) of the \( z \)-samples of the diseased distribution(s) where the disease is visible

\( \alpha \) vector: the proportion(s) of the diseased distribution(s) where the disease is visible

Details

\( \mu \) and \( \alpha \) must have equal length. The predicted ROC curve for each \( \mu \) and \( \alpha \) pair will be plotted.

Value

A \texttt{ggplot2} object of the plotted ROC curve(s)
References


Examples

cbmPlot <- PlotCbmFit(c(1, 2), c(0.5, 0.5))
print(cbmPlot)

Description

Plot empirical operating characteristics (operating points connected by straight lines) for specified treatments and readers, or if desired, plots only (no operating points) averaged over specified treatments and / or readers

Usage

PlotEmpiricalOperatingCharacteristics(dataset, trts = 1, rdrs = 1, opChType = "ROC")

Arguments

dataset Dataset to be used for plotting
trts List or vector: integer indices of treatments to be plotted
rdrs List or vector: integer indices of readers to be plotted
opChType Type of operating characteristic to be plotted: "ROC" (the default), "FROC", "AFROC", "wAFROC", "AFROC1", or "wAFROC1"

Details

The trts and rdrs are vectors or lists of integer indices, not the corresponding string IDs. For example, if the string ID of the first reader is "0", the value in rdrs should be 1 not 0. The legend shows the string IDs.

If both of trts and rdrs are vectors, all combinations of treatments and readers are plotted. See Example 1.

If both trts and rdrs are lists, they must have the same length. Only the combination of treatment and reader at the same position in their respective lists are plotted. If some elements of the treatments and / or readers lists are vectors, the average operating characteristic over the implied treatments and / or readers are plotted. See Example 2.
**Value**

A `ggplot2` object containing the operating characteristic plot(s) and a data frame containing the points defining the operating characteristics are returned. For example, the returned objects for "ROC" operating characteristics are as follows:

**Plot**

`ggplot2` object. For continuous or averaged data, operating characteristics curves are plotted without showing operating points. For binned individual data, both operating points and connecting lines are shown. To avoid clutter, if there are more than 20 operating points, they are not shown.

**Points**

Data frame with four columns: abscissa, ordinate, class (which codes treatment and reader) and type, which can be "individual", "continuous" or "average"; "individual" refers to a one treatment and one reader.

**Examples**

```r
# Example 1
# Plot individual empirical ROC plots for all combinations of treatments
# 1 and 2 and readers 1, 2 and 3. Six operating characteristics are plotted.

deret <- PlotEmpiricalOperatingCharacteristics(dataset =
dataset02, trts = c(1:2), rdrs = c(1:3))
print(ret$Plot)

# Example 2
# Empirical ROC, FROC, AFROC and wAFROC plots. Each plot consists of
# three parts (see Example 3 for correspondences between indices and string identifiers
# for treatments and readers):
# (1) plot for the 1st treatment (string ID "1") and the 2nd reader (string ID "3")
# (2) plot for the 2nd treatment (string ID "2") AVERAGED over the 2nd and 3rd readers
# (string IDs "3" and "4"), and
# (3) plot AVERAGED over the first two treatments (string IDs "1" and "2") AND over
# the 1st, 2nd and 3rd readers (string IDs "1", "3" and "4")

plotT <- list(1, 2, c(1:2))
plotR <- list(2, c(2:3), c(1:3))
deret <- PlotEmpiricalOperatingCharacteristics(dataset = dataset04, trts = plotT, rdrs = plotR)
print(ret$Plot)

ret <- PlotEmpiricalOperatingCharacteristics(dataset = dataset04, trts = plotT, rdrs = plotR,
opChType = "FROC")
print(ret$Plot)

ret <- PlotEmpiricalOperatingCharacteristics(dataset = dataset04, trts = plotT, rdrs = plotR,
opChType = "AFROC")
print(ret$Plot)

ret <- PlotEmpiricalOperatingCharacteristics(dataset = dataset04, trts = plotT, rdrs = plotR,
opChType = "wAFROC")
print(ret$Plot)

#Example 3
```
## Correspondences between indices and string identifiers for treatments and readers in this dataset. Apparently reader "2" did not complete the study.

```r
str(dataset04)
```

### List of 8

- `$ NL` num [1:5, 1:4, 1:200, 1:7] -Inf -Inf -Inf -Inf -Inf ...
- `$ LL` num [1:5, 1:4, 1:100, 1:3] 5 4 3 5 4 2 4 5 ...
- `$ lesionNum` int [1:100] 1 1 1 1 1 1 1 1 1 1 ...
- `$ lesionID` num [1:100, 1:3] 1 1 1 1 1 1 1 1 1 1 ...
- `$ lesionWeight` num [1:100, 1:3] 1 1 1 1 1 1 1 1 1 1 ...
- `$ dataType` chr [1] "FROC"
- `$ modalityID` chr [1] "1" "2" "3" "4" "5"
- `$ readerID` chr [1] "1" "3" "4" "5"

---

### PlotRsmOperatingCharacteristics

*RSM predicted operating characteristics, ROC pdfs and different FOMs possible with FROC data*

---

**Description**

Visualize predicted ROCs, AFROCs, wAFROCs, FROCs and pdfs (probability density functions of highest ratings, for non-diseased and diseased cases), for up to 2 sets of search model parameters. This function is useful as an instructional tool towards understanding the RSM.

**Usage**

```r
PlotRsmOperatingCharacteristics (mu, lambda, nu, lesDistr,
lesionWeights, type = "ALL", legendPosition = c(1,0),
legendDirection = "horizontal", legendJustification = c(0,1),
nlfRange = NULL, llfRange = NULL, nlfAlpha = NULL,myNegInf = -3)
```

**Arguments**

- **mu** Array, max length 2. The mean(s) of the Gaussian distribution(s) for the ratings of latent LLs (continuous ratings of lesions that are found by the observer's search mechanism)
- **lambda** Array, max length 2. The Poisson distribution *intrinsic* parameter(s), which model the random numbers of latent NLs (suspicious regions that do not correspond to actual lesions) per case, for up to two treatments. The corresponding *physical* parameters are lambda/mu. Two conversion functions are provided: `UtilIntrinsic2PhysicalRSM` and `UtilPhysical2IntrinsicRSM`.
- **nu** Array, max length 2. The binomial distribution success probability *intrinsic* parameters, which model the random numbers of latent LLs (suspicious regions that correspond to actual lesions) per diseased case for up to two treatments; the corresponding *physical* parameter is \( 1 - \exp(nu*mu) \), the success probability of the binomial distribution(s).
PlotRsmOperatingCharacteristics

**lesDistr** Array, [1:maxLL,1:2]. The probability mass function of the lesion distribution for diseased cases. The first column contains the actual numbers of lesions per case. The second column contains the fraction of diseased cases with the number of lesions specified in the first column. The second column must sum to unity.

**lesionWeights** Array, [1:maxLL,1:maxLL]. The weights (or clinical importances) of the lesions. The 1st row contains the weight of the lesion on cases with one lesion only, necessarily 1; the remaining elements of the row are \(-\text{Inf}\). The 2nd row contains the weights of the 2 lesions on cases with 2 lesions only, the remaining elements of the row, if any, are \(-\text{Inf}\). Excluding the \(-\text{Inf}\), each row must sum to 1. The default is equal weighting, e.g., weights are 1/3, 1/3, 1/3 on row 3. This parameter is not to be confused with the lesionWeights field in an FROC dataset with enumerates the weights of lesions on individual cases.

**type** The type of operating characteristic desired: can be "ROC", "AFROC", "wAFROC", "FROC" or "pdfs" or "ALL". The default is "ALL".

**legendPosition** The positioning of the legend: "right", "left", "top" or "bottom". Use "none" to suppress the legend.

**legendDirection** Allows control on the direction of the legend; "horizontal", the default, or "vertical"

**legendJustification** Where to position the legend, default is bottom right corner c(0,1)

**nlfRange** This applies to FROC plot only. The x-axis range, e.g., c(0,2), for FROC plot. Default is "NULL", which means the maximum NLF range, as determined by the data.

**llfRange** This applies to FROC plot only. The y-axis range, e.g., c(0,1), for FROC plot. Default is "NULL", which means the maximum LLF range, as determined by the data.

**nlfAlpha** Upper limit of the integrated area under the FROC plot. Default is "NULL", which means the maximum NLF range is used (i.e., lambda/mu). Attempt to integrate outside the maximum NLF will generate an error.

**myNegInf** How close one approaches the end-point; the default is -3. This is used in the code to demonstrate continuity of the slope of the ROC at the end point; Online Appendix 17.H.3

**Details**

RSM is the Radiological Search Model described in the book.

**Value**

A list of 6 elements containing six ggplot2 objects (ROCPPlot, AFROCPPlot wAFROCPPlot, FROCPPlot and PDFPPlot) and two area measures (each of which can have up to two elements), the area under the search model predicted ROC curves in up to two treatments, the area under the search model predicted AFROC curves in up to two treatments, the area under the search model predicted wAFROC curves in up to two treatments, the area under the search model predicted FROC curves in up to two treatments.
• ROCPlot The predicted ROC plots
• AFROCPlot The predicted AFROC plots
• wAFROCPlot The predicted wAFROC plots
• FROCPlot The predicted FROC plots
• PDFPlot The predicted pdf plots
• aucROC The predicted ROC AUCs
• aucAFROC The predicted AFROC AUCs
• aucwAFROC The predicted wAFROC AUCs
• aucFROC The predicted FROC AUCs

Note
For lesDistr, the sum over the second column must equal one. If all cases contain same number of lesions, simply supply this number instead of the matrix. If the argument is missing, the default value of one lesion per diseased case applies.

In lesionWeights, the sum over each row (excluding -Inf) must be one. The value -Inf should be assigned if the corresponding lesion does not exist. Equal lesion weighting is applied if this argument is missing.

For example, if the maximum number of distinct lesion configurations per case is 3 (e.g., 1, 2 and 4, implying there are no cases with 3 lesions), the first column of lesDistr will be c(1,2,4). The second column might be c(0.8, 0.15, 0.05), which sums to one, meaning 80% of cases have only one lesion, 15% have two lesions and 5% have three lesions. The lesionWeights matrix will be [1:3,1:3], where each row will sum to one (excluding negative infinities).

References

Examples
## Following example is for mu = 2, lambda = 1, nu = 0.6, in one treatment and
## mu = 3, lambda = 1.5, nu = 0.8, in the other treatment. 20% of the diseased
## cases have a single lesion, 40% have two lesions, 10% have 3 lesions,
## and 30% have 4 lesions.
lesDistr <- rbind(c(1, 0.2), c(2, 0.4), c(3, 0.1), c(4, 0.3))

## On cases with one lesion the weights are 1, on cases with 2 lesions the weights
Simulate paired binned data for testing FitCorCbm

Description
Simulates single treatment 2-reader binned ROC dataset, simulated according to the CORCBM model, for the purpose of testing the fitting program FitCorCbm

Usage
```r
SimulateCorCbmdataset(
  seed = 123, K1 = 50, K2 = 50, desiredNumBins = 5,
  muX = 1.5, muY = 3,
  alphaX = 0.4, alphaY = 0.7,
  rhoNor = 0.3, rhoAbn2 = 0.8)
```

Arguments
- `seed` The seed variable, default is 123; set to NULL for truly random seed
- `K1` The number of non-diseased cases, default is 50
- `K2` The number of diseased cases, default is 50
- `desiredNumBins` The desired number of bins; default is 5
- `muX` The CBM mu parameter in condition X
- `muY` The CBM mu parameter in condition Y
- `alphaX` The CBM alpha parameter in condition X
- `alphaY` The CBM alpha parameter in condition Y
- `rhoNor` The correlation of non-diseased case z-samples
- `rhoAbn2` The correlation of diseased case z-samples, when disease is visible in both conditions
Details

X and Y refer to the two arms of the pairing. \( \mu_X \) and \( \alpha_X \) refer to the univariate CBM parameters in condition X. \( \rho_{\text{Nor}} \) is the correlation of ratings of non-diseased cases and \( \rho_{\text{Abn2}} \) is the correlation of ratings of diseased cases when disease is visible in both conditions. The ROC data is binned to 5 bins in each condition. See referenced publication.

Value

The return value is the desired dataset, suitable for testing FitCorCbM.

References


Examples

dataset <- SimulateCorCbMDataset()

## Not run:

dataset <- SimulateCorCbMDataset(K1 = 5000, K2 = 5000)

## End(Not run)

SimulateFrocDataset Simulates an MRMC uncorrelated FROC dataset using the RSM

Description

Simulates an uncorrelated MRMC FROC dataset for specified numbers of readers and treatments

Usage

SimulateFrocDataset(mu, lambda, nu, zeta1, I, J, K1, K2, lesionNum)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mu</td>
<td>The intrinsic mu parameter of the RSM</td>
</tr>
<tr>
<td>lambda</td>
<td>The intrinsic lambda parameter of the RSM (not the physical parameter)</td>
</tr>
<tr>
<td>nu</td>
<td>The intrinsic nu parameter of the RSM (not the physical parameter)</td>
</tr>
<tr>
<td>zeta1</td>
<td>The lowest reporting threshold</td>
</tr>
</tbody>
</table>
**SimulateRocDataset**

The number of treatments

J

The number of readers

K1

The number of non-diseased cases

K2

The number of diseased cases

lesionNum

A K2 length array containing the numbers of lesions per diseased case

**Details**

See book chapters on the Radiological Search Model (RSM) for details. In this code correlations between ratings on the same case are assumed to be zero.

**Value**

The return value is an FROC dataset.

**References**


**Examples**

```r
set.seed(1)
K1 <- 5; K2 <- 7;
maxLL <- 2; lesionNum <- floor(runif(K2, 1, maxLL + 1))
mu <- 1; lambda <- 1; nu <- 1; zeta1 <- -1
I <- 2; J <- 5

frocdataraw <- SimulateFrocdataset(
  mu = mu, lambda = lambda, nu = nu, zeta1 = zeta1,
  I = I, J = J, K1 = K1, K2 = K2, lesionNum = lesionNum )

## plot the data
ret <- PlotEmpiricalOperatingCharacteristics(frocdataraw, trts=1,
  rdrs = 1, opChType = "FROC")
print(ret$Plot)
```

**Description**

Simulates a binormal model ROC dataset for a single treatment and reader.
Usage

SimulateRocDataset(K1, K2, a, b, seed = NULL)

Arguments

K1 The number of non-diseased cases
K2 The number of diseased cases
a The \(a\) parameter of the binormal model
b The \(b\) parameter of the binormal model
seed The initial seed, default is NULL

Details

See book Chapter 6 for details

Value

An ROC dataset

References


Examples

```r
K1 <- 5; K2 <- 7;
a <- 1.5; b <- 0.5

rocDataRaw <- SimulateRocDataset(K1 = K1, K2 = K2,
a = a, b = b)

## plot the data
ret <- PlotEmpiricalOperatingCharacteristics(rocDataRaw, trts = 1,
  rdrs = 1, opChType = "ROC")
print(ret$Plot)
```
SsPowerGivenJK

Statistical power for specified numbers of readers and cases in an ROC study

Description

Calculate the statistical power for specified numbers of readers J, cases K, analysis method and DBM or OR variances components

Usage

SsPowerGivenJK(J, K, effectSize, method, option = "ALL", alpha = 0.05, ...)

Arguments

J  The number of readers in the pivotal study
K  The number of cases in the pivotal study
effectSize  The effect size to be used in the calculation, the sign is unimportant, see Ch 11 in book for guidance.
method  "DBMH" or "ORH"
option  "RRRC", "FRRC", "RRFC" or "ALL"; the default is "ALL"
alpha  The significance level, default is 0.05.
...  Other necessary parameters, OR or DBM variance components, see details

Details

Regarding other parameters (...) needed are either the set of of DBM variance components, i.e, (varYTR, varYTC, and varYEpS), or the set of OR covariance matrix elements, the treatment-reader variance and number of cases in pilot study i.e, (cov1, cov2, cov3, varEps, varTR and KStar).

If both of are given, DBM variance components are used and the OR values are ignored.

Either numeric values, for example, of varYTR, varYTC, varYEpS can be supplied, provided they are in that order, or the function call must explicitly state, for example, cov1 = value1, cov2 = value2, cov3 = value3, varTR = value4, varEps = value5, KStar = value6, i.e., in any order.

Value

The expected statistical power.
References


Examples

```r
## An example of sample size calculation with DBM variance components
retDbm <- StSignificanceTesting(data = dataset02,
    FOM = "Wilcoxon", method = "DBMH")
effectSize <- retDbm$ciDiffTrtRRRC$Estimate
varCompDBM <- retDbm$varComp
varYTR <- varCompDBM$varComp[3]
varYTC <- varCompDBM$varComp[4]
varYEps <- varCompDBM$varComp[6]

## should give close to 80% power for RRRC
SsPowerGivenJK(6, 251, effectSize, "DBMH", varYTR = varYTR, varYTC = varYTC, varYEps = varYEps)

## An example of sample size calculation with OR variance components.
retOR <- StSignificanceTesting(data = dataset02,
    FOM = "Wilcoxon", covEstMethod = "Jackknife", method = "ORH")
effectSize <- retOR$ciDiffTrtRRRC$Estimate
varCompOR <- retOR$varComp
varTR <- varCompOR$varCov[2]
cov1 <- varCompOR$varCov[3]
cov2 <- varCompOR$varCov[4]
cov3 <- varCompOR$varCov[5]
varEps <- varCompOR$varCov[6]

KStar <- length(dataset02$NL[1,1,1])

## same sample size as above, different method, should again give close to 80% power for RRRC
SsPowerGivenJK(6, 251, effectSize, "ORH", cov1 = cov1, cov2 = cov2, cov3 = cov3, varEps = varEps, varTR = varTR, KStar = KStar)
```

SsPowerTable

**Generate a power table**

Description

Generate combinations of numbers of readers J and numbers of cases K for desired power and specified generalizations (i.e., RRRC or FRRC or RRFC)
Usage

SsPowerTable(effectSize, alpha = 0.05, desiredPower = 0.8,
method = "DBMH", option = "ALL", ...)

Arguments

effectSize  The postulated effect size
alpha  The The size of the test, default is 0.05
desiredPower  The desired statistical power, default is 0.8
method  Analysis method, "DBMH" or "ORH", the default is "DBMH"
option  Desired generalization; the default is "RRRC", for random-reader random-cases
...  Other necessary parameters, OR or DBM variance components, see details

Details

Regarding other parameters (...), see details in SsPowerGivenJK

Value

A data frame containing following three columns.

numReaders  The number of readers in the pivotal study.
numCases  The number of cases in the pivotal study.
power  The calculated statistical power corresponding to the indicated numbers of readers and cases.

Note

The procedure is valid for ROC studies only; for FROC studies see Online Appendix Chapter 19.

Examples

```r
## Example of sample size calculation with DBM method
retDbm <- StSignificanceTesting (dataset02, FOM = "Wilcoxon", method = "DBMH")
effectSize <- retDbm$ciDiffTrtRRRC$Estimate
varYTR <- retDbm$varComp$varComp[3]
varYTC <- retDbm$varComp$varComp[4]
varYEps <- retDbm$varComp$varComp[6]
powTab <- SsPowerTable(
  effectSize = effectSize,
  method = "DBMH",
  varYTR = varYTR,
  varYTC = varYTC,
  varYEps = varYEps)
print(powTab)

## Example of sample size calculation with OR method
retOR <- StSignificanceTesting (dataset02, FOM = "Wilcoxon", method = "ORH")
```
SSSampleSizeKGivenJ

Number of cases, for specified number of readers, to achieve desired power

Description

Number of cases to achieve the desired power, for specified number of readers J, and specified DBM or OR variance components.

Usage

SSSampleSizeKGivenJ (J, alpha = 0.05, effectSize = 0.05, desiredPower = 0.8, option = "ALL", method = "DBMH", ...)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>The number of readers in the pivotal study</td>
</tr>
<tr>
<td>alpha</td>
<td>The significance level of the study, default value is 0.05.</td>
</tr>
<tr>
<td>effectSize</td>
<td>The effect size to be used in the study, default value is 0.05.</td>
</tr>
<tr>
<td>desiredPower</td>
<td>The desired statistical power, default value is 0.8.</td>
</tr>
<tr>
<td>option</td>
<td>Desired generalization, &quot;RRRC&quot;, &quot;FRRC&quot;, &quot;RRFC&quot; or &quot;ALL&quot; (the default).</td>
</tr>
<tr>
<td>method</td>
<td>&quot;DBMH&quot; (default) or &quot;ORH&quot;.</td>
</tr>
<tr>
<td>...</td>
<td>Other necessary parameters, OR or DBM variance components, see details</td>
</tr>
</tbody>
</table>

Details

Regarding other parameters (...), see details in SsPowerGivenJK. An additional parameter KStar, the number of cases in the pilot study, is required when using OR variability parameters.
Value

A list of two elements:

K                    The minimum number of cases K in the pivotal study to just achieve the desired statistical power.
power                The predicted statistical power.

Note

The procedure is valid for ROC studies only; for FROC studies see Online Appendix Chapter 19.

References


Examples

```r
## Following is an example of sample size calculation using the DBM variance components and jackknifing (the default) to estimate the variance components
retDbm <- StSignificanceTesting(data = dataset02, FOM = "Wilcoxon", method = "DBMM")
effectSize <- retDbm$c1DiffTrtRRC$Estimate
varCompDBM <- retDbm$varComp
varYTR <- varCompDCBM$varComp[3]
varYTC <- varCompDCBM$varComp[4]
varYEps <- varCompDCBM$varComp[6]
SsSampleSizeKGivenJ(J = 6, varYTR = varYTR, varYTC = varYTC, varYEps = varYEps, effectSize = effectSize)

## Following is an example of sample size calculation using the OR variance components
retOr <- StSignificanceTesting(data = dataset02, FOM = "Wilcoxon", covEstMethod = "Jackknife", method = "ORH")
effectSize <- retOR$c1DiffTrtRRC$Estimate
varCompOR <- retOR$varComp
varTR <- varCompOR$varCov[2]
var1 <- varCompOR$varCov[3]
var2 <- varCompOR$varCov[4]
var3 <- varCompOR$varCov[5]
varEps <- varCompOR$varCov[6]
KStar <- 114
SsSampleSizeKGivenJ(J = 6, cov1 = var1, cov2 = var2, cov3 = var3, varTR = varTR, varEps = varEps, KStar = KStar, effectSize = effectSize, method = "ORH")
```

## Not run:
## Following is an example of power calculations using the DBM variance components, and scanning the number of readers

retDbm <- StSignificanceTesting(data = dataset02, FOM = "Wilcoxon", method = "DBMH")
effectSize <- retDbm$ciDiffTrrRC$Estimate
varyTR <- retDbm$varComp$varComp[3]
varyTC <- retDbm$varComp$varComp[4]
varyEps <- retDbm$varComp$varComp[6]
effectSize <- retDbm$ciDiffTrrRC$Estimate
for (J in 6:10) {
  ret <- SsSampleSizeKGivenJ(J = J, varyTR = varyTR, varyTC = varyTC, varyEps = varyEps, effectSize = effectSize)
  message("# of readers = ", J, " estimated # of cases = ", ret$K, ", predicted power = ",
          signif(ret$powerRRRC,3), "\n")
}

## Following is an example of power calculations using the ORH variance components,
## using bootstrap to estimate variance components
retOR <- StSignificanceTesting(data = dataset02, FOM = "Wilcoxon", covEstMethod = "Bootstrap", method = "ORH")
effectSize <- retOR$ciDiffTrrRC$Estimate
varCompOR <- retOR$varComp
varTR <- varCompOR$varCov[2]
cov1 <- varCompOR$varCov[3]
cov2 <- varCompOR$varCov[4]
cov3 <- varCompOR$varCov[5]
varyEps <- varCompOR$varCov[6]
KStar <- length(dataset02$NL[1,1,1])
SsSampleSizeKGivenJ(J = 6, cov1 = cov1, cov2 = cov2, cov3 = cov3, varTR = varTR, varyEps = varyEps,
                        KStar = KStar, effectSize = effectSize, method = "ORH")

## End(Not run)

### Description
Performs Dorfman-Berbaum-Metz (DBM) or Obuchowski-Rockette (OR) significance testing with
Hillis’ improvements, for specified dataset; significance testing refers to analysis designed to assign
a P-value for rejecting a null hypothesis (NH); the most common NH is that the reader-averaged
figure of merit (FOM) difference between treatments is zero. The results of the analysis are better
visualized in the text or, preferably, Excel-formatted, files produced by UtilOutputReport.

### Usage
StSignificanceTesting (dataset, FOM = "wJAFROC", alpha = 0.05,
                       method = "DBMH", covEstMethod = "Jackknife", nBoots = 200, option = "ALL",
                       VarCompFlag = FALSE, FPFValue = 0.2)
### Arguments

- **dataset**
  The dataset to be analyzed, see RJafroc-package

- **FOM**
  The figure of merit, default "wJAFROC", see UtilFigureOfMerit

- **alpha**
  The significance level of the test of the null hypothesis that all treatment effects are zero; the default alpha is 0.05

- **method**
  The significance testing method to be used. There are two options: "DBMH" (the default) or "ORH", representing the Dorfman-Berbaum-Metz and the Obuchowski-Rockette significance testing methods, respectively.

- **covEstMethod**
  The method used to estimate the covariance matrix in ORH analysis; it can be "Jackknife", "Bootstrap" or "DeLong", the last assumes FOM = "Wilcoxon", otherwise an error results. This parameter is not relevant if the analysis method is "DBMH"

- **nBoots**
  The number of bootstraps (default is 200), relevant only if the "Bootstrap" method is used to estimate the covariance matrix in the ORH method

- **option**
  Determines which factors are regarded as random vs. fixed: "RRRC" = random-reader random-case, "FRRC" = fixed-reader random case, "RRFC" = random-reader fixed case, "ALL" outputs the results of "RRRC", "FRRC" and "RRFC" analyses

- **VarCompFlag**
  If TRUE, only the appropriate (DBM or OR) variance components (six in all) are returned, default is FALSE

- **FPFValue**
  Only needed for LROC data; where to evaluate a partial curve based figure of merit. The default is 0.2.

### Value

For method = "DBMH" the returned value is a list with 22 members:

- **fomArray**
  The figure of merit array for each treatment-reader combination

- **anovaY**
  The ANOVA table of the pseudovalues over all treatments

- **anovaYi**
  The ANOVA table of the pseudovalues for each treatment

- **varComp**
  The variance components of the pseudovalue model underlying the analysis, 6 values, in the following order: c("Var(R)", "Var(C)", "Var(T*R)", "Var(T*C)", "Var(R*C)", "Var(Error)")

- **fRRRC**
  For random-reader random-case (RRRC) analysis, the F-statistic for rejecting the null hypothesis of no treatment effect

- **ddfRRRC**
  For RRRC analysis, the denominator degrees of freedom of the F statistic

- **pRRRC**
  For RRRC analysis, the p-value of the significance test of the NH

- **ciDiffTrtRRRC**
  For RRRC analysis, the confidence intervals and related test statistics for the FOM differences between pairs of treatments

- **ciAvgRdrEachTrtRRRC**
  For RRRC analysis, the confidence intervals and related test statistics for rdr. avg. FOM in each treatment

- **fFRRC**
  For fixed-reader random-case (FRRC) analysis, the F-statistic for rejecting the NH
StSignificanceTesting

/ddfFRRRC/ For FRRC analysis, the denominator degrees of freedom of the F-statistic

/pFRRC/ For FRRC analysis, the p-value of the significance test of the NH

/ciDiffTrtFRRC/ For FRRC analysis, the confidence intervals and related test statistics for the FOM differences between pairs of treatments

/ciAvgRdrEachTrtFRRC/ For FRRC analysis, the confidence intervals and related tests for rdr. avg. FOM in each treatment

/ssAnovaEachRdr/ The sum of squares table of the ANOVA of the pseudovalues for each reader (based on data for the specified reader)

/msAnovaEachRdr/ The mean squares table of the ANOVA of the pseudovalues for each reader (based on data for the specified reader)

/ciDiffTrtEachRdr/ The confidence intervals and related tests of the FOM differences between pairs of treatments for each reader

/fRRFC/ For random-reader fixed-case (RRFC) analysis, the F statistic

/ddfFRRFC/ For RRFC analysis, the denominator degrees of freedom of the F statistic

/pRRFC/ For RRFC analysis, the p-value for rejecting the NH

/ciDiffTrtRRFC/ For RRFC analysis, the confidence intervals and related test statistics for the FOM differences between pairs of treatments

/ciAvgRdrEachTrtRRFC/ For RRFC analysis, the confidence intervals and related tests for reader averaged FOM in each treatment

For method = "ORH" the return value is a list with with 21 members:

/fomArray/ Figures of merit array. See the return of UtilFigureOfMerit

/msT/ Mean square of the figure of merit corresponding to the treatment effect

/msTR/ Mean square of the figure of merit corresponding to the treatment-reader effect

/varComp/ The variance components of the pseudovalue model underlying the analysis, 6 values, in the following order: c("Var(R)", "Var(T*R)", "COV1", "COV2", "COV3","Var(Error)")

/fRRRC/ Same as DBMH method

/ddfFRRRC/ Same as DBMH method

/pRRRC/ Same as DBMH method

/ciDiffTrtRRRC/ Same as DBMH method

/ciAvgRdrEachTrtRRRC/ Same as DBMH method

/fFRRRC/ Same as DBMH method

/ddfFRRRC/ Same as DBMH method

/pFRRRC/ Same as DBMH method

/ciDiffTrtFRRRC/ Same as DBMH method

/ciAvgRdrEachTrtFRRRC/ Same as DBMH method
StSignificanceTesting

ciDiffTrtEachRdr
   Same as DBMH method
varCovEachRdr   Obuchowski-Rockette Variance and Cov1 estimates for each reader
fRRFC          Same as DBMH method
ddfRRFC        Same as DBMH method
pRRFC          Same as DBMH method
ciDiffTrtRRFC  Same as DBMH method
ciAvgRdrEachTrtRRFC  Same as DBMH method

References


Examples

retDbmRoc <- StSignificanceTesting(dataset02,
   FOM = "Wilcoxon", method = "DBMH")

## Not run:
retDbmJAFROC <- StSignificanceTesting(dataset05) # default is weighted JAFROC

retDbmHrAuc <- StSignificanceTesting(dataset05,
   FOM = "HrAuc", method = "DBMH")
print(retDbmHrAuc)

retDbmSongA1 <- StSignificanceTesting(dataset05,
   FOM = "SongA1", method = "DBMH")
print(retDbmSongA1)

retDbmSongA2 <- StSignificanceTesting(dataset05,
   FOM = "SongA2", method = "DBMH")
print(retDbmSongA2)

retDbmWjafroc1 <- StSignificanceTesting(dataset05,
   FOM = "wJAFROC1", method = "DBMH")
print(retDbmWjafroc1)

retDbmJafroc1 <- StSignificanceTesting(dataset05,
Description

Significance testing, comparing CAD vs. a group of radiologists interpreting the same cases, an example of single treatment analysis

Usage

\[
\text{StSignificanceTestingCadVsRadiologists (dataset, FOM = "Wilcoxon", option = "RRRC", method = "singleModality", FPFValue = 0.2)}
\]

Arguments

dataset  \textbf{The dataset must be ROC or LROC.}
FOM  
The desired FOM, default is "Wilcoxon" for ROC data, or ROC data inferred from LROC data; for LROC data the choices are "PCL" and "ALROC".
option  
The desired generalization, the default is "RRRC"; another possibility is "RRFC". 
method  "singleModality", the default, or "dualModality", see details.
FPFValue  
Only needed for LROC data; where to evaluate a partial curve based figure of merit, see details. The default is 0.2.

Details

PCL is the probability of a correct localization. The LROC is the plot of PCL (ordinate) vs. FPF. For LROC data "PCL" means interpolated PCL value at specified "FPFValue". "ALROC" is the trapezoidal area under the LROC from FPF = 0 to FPF = FPFValue. If method = "singleModality" the first reader is assumed to be CAD. If method = "dualModality" the first treatment is
assumed to be CAD. The NH is that the FOM of CAD equals the average of the readers. The method = "singleModality" analysis uses an adaptation of the single-treatment multiple-reader Obuchowski Rockette (OR) model described in a paper by Hillis (2007), section 5.3. The adaptation is characterized by 3 parameters $\text{Var}_R$, $\text{Var}$ and $\text{Cov}_2$, which are returned by the function. The method = "dualModality" analysis replicates CAD data as many times as necessary so as to form one "treatment" of an MRMC pairing, the other "treatment" being the radiologists. Standard RRRC DBMH/ORH analysis is applied. The method, described in Kooi et al gives exactly the same final results (F-statistic, ddf and p-value) as "singleModality" but the intermediate quantities are questionable. The method is characterized by 6 OR parameters $\text{Var}_R$, $\text{VarTR}$, $\text{Var}$, $\text{Cov}_1$, $\text{Cov}_2$ and $\text{Cov}_3$, which are returned by the function.

**Value**

If method = "singleModality" the return value is a list with the following elements:

- `fomCAD` The observed FOM for CAD
- `fomRAD` The observed FOM array for the readers
- `avgRadFom` The average FOM of the readers
- `avgDiffFom` The mean of the difference FOM, RAD - CAD
- `ciAvgDiffFom` The 95-percent CI of the average difference, RAD - CAD
- `varR` The variance of the radiologists
- `varError` The variance of the error term in the single-treatment multiple-reader OR model
- `covR` The covariance of the error term
- `tstat` The observed value of the t-statistic; it's square is equivalent to an F-statistic
- `df` The degrees of freedom of the t-statistic
- `pval` The p-value for rejecting the NH
- `Plots` Empirical operating characteristic plots corresponding to specified FOM

If method = "dualModality" the return value is a list with the following elements:

- `fomCAD` The observed FOM for CAD
- `fomRAD` The observed FOM array for the readers
- `avgRadFom` The average FOM of the readers
- `avgDiffFom` The mean of the difference FOM, RAD - CAD
- `ciDiffFom` A data frame containing the statistics associated with the average difference, RAD - CAD
- `ciAvgRdrEachTrt` A data frame containing the statistics associated with the average FOM in each treatment
- `varR` The variance of the pure reader term in the OR model
- `varTR` The variance of the treatment-reader term error term in the OR model
- `cov1` The covariance 1 of the error term - same reader, different treatments
- `cov2` The covariance 2 of the error term - different readers, same treatment
The covariance of the error term - different readers, different treatments

The variance of the pure error term in the OR model

The observed value of the F-statistic

The numerator degrees of freedom of the F-statistic

The denominator degrees of freedom of the F-statistic

The p-value for rejecting the NH

Empirical operating characteristic plots corresponding to specified FOM, i.e., if FOM = "Wilcoxon" an ROC plot is produced where reader 1 is CAD. If an LROC FOM is selected, an LROC plot is displayed.

Note

The extension of the code to FROC will be addressed in a future update.

References


Examples

```r
ret1M <- StSignificanceTestingCadVsRadiologists (dataset09, FOM = "Wilcoxon", method = "singleModality")

# Not run:
ret2M <- StSignificanceTestingCadVsRadiologists (dataset09, FOM = "Wilcoxon", method = "dualModality")

retLroc1M <- StSignificanceTestingCadVsRadiologists (datasetCadLroc, FOM = "PCL", option = "RRRC", method = "singleModality", FPFValue = 0.05)

retLroc2M <- StSignificanceTestingCadVsRadiologists (datasetCadLroc, FOM = "PCL", option = "RRRC", method = "dualModality", FPFValue = 0.05)

# test with fewer readers
dataset09a <- DFExtractDataset(dataset09, rdrs = seq(1:7))
ret1M7 <- StSignificanceTestingCadVsRadiologists (dataset09a, FOM = "Wilcoxon", method = "singleModality")
ret2M7 <- StSignificanceTestingCadVsRadiologists (dataset09a,
```
StSignificanceTestingCrossedModalities

Perform significance testing using crossed treatments analysis

Description

Performs ORH analysis for specified crossed treatments dataset averaged over specified treatment factor

Usage

StSignificanceTestingCrossedModalities(crossedData, avgIndx, FOM = "wAFROC", alpha = 0.05, option = "ALL")

Arguments

crossedData  The crossed treatments dataset
avgIndx     The index of the treatment to be averaged over
FOM        See StSignificanceTesting.
alpha      See StSignificanceTesting.
option     See StSignificanceTesting.

Value

The return list contains the same items with StSignificanceTesting.

Examples

FOM = "Wilcoxon", method = "dualModality")
datasetCadLroc7 <- DFExtractDataset(datasetCadLroc, rdrs = seq(1:7))
ret1MLroc7 <- StSignificanceTestingCadVsRadiologists (datasetCadLroc7, FOM = "FCL", option = "RRRC", method = "singleModality", FPFValue = 0.05)
ret2MLroc7 <- StSignificanceTestingCadVsRadiologists (datasetCadLroc7, FOM = "FCL", option = "RRRC", method = "dualModality", FPFValue = 0.05)

## End(Not run)
StSignificanceTestingSingleFixedFactor

Perform significance testing for single fixed factor analysis

Description
Significance testing for datasets with single reader in multiple (at least two) treatments, or single treatment with multiple (at least two) readers, where reader or treatment, respectively, is regarded as a fixed factor and a common case-set, regarded as random, is assumed.

Usage
```
StSignificanceTestingSingleFixedFactor(dataset, FOM = "wAFROC", alpha = 0.05)
```

Arguments
- `dataset` A single-treatment or single-reader dataset.
- `FOM` The figure of merit, default "wAFROC", see UtilFigureOfMerit.
- `alpha` The significance level (alpha, default 0.05) of the test of the null hypothesis that FOMs of all levels of the fixed factor are identical.

Details
This function performs implements Hillis et al. 2005, Eqn. 23. Following an overall F-test, reader-pairings are compared using paired t-tests. **In order for a specific pairing to be declared significant, the F-test must also be significant.**

Value
The return value is a list containing:
- `f` The observed F-statistic for testing the null hypothesis of no treatment effect.
- `ddf` The denominator degrees of freedom of the F statistic. The numerator degrees of freedom is always the number of levels of the fixed factor minus one.
- `pValue` The p-value for rejecting the NH.
- `fomStats` Statistics for FOM for each level of the fixed factor.
- `diffFomStats` Statistics for FOM-differences for all distinct pairings of the levels of the fixed factor.
References


Examples

```r
## Create a single treatment ROC dataset with one treatment and four readers
singlefactorData <- DfExtractDataset(dataset02, 1, 1:4)

## Performs single treatment fixed reader analysis
StSignificanceTestingSingleFixedFactor(singlefactorData, FOM = "Wilcoxon")
```

UtilAucBinormal  

Binormal model AUC function

Description

Returns the Binormal model ROC-AUC corresponding to specified parameters. See also UtilAuc-sRSM, UtilAucPROPROC and UtilAucCBM

Usage

UtilAucBinormal (a, b)

Arguments

- **a**  
  The a parameter of the binormal model (separation of non-diseased and diseased pdfs)

- **b**  
  The b parameter of the binormal model (std. dev. of non-diseased diseased pdf; diseased pdf has unit std. dev)

Value

Binormal model-predicted ROC-AUC

References

UtilAucCBM

Examples

```r
a <- 2; b <- 0.7
UtilAucBinormal(a,b)
```

Description

Returns the CBM ROC-AUC. See also UtilAucsRSM, UtilAucPROPROC and UtilAucBinormal.

Usage

```r
UtilAucCBM (mu, alpha)
```

Arguments

- `mu`: The mu parameter of CBM (separation of non-diseased and diseased pdfs).
- `alpha`: The alpha parameter of CBM, i.e., the fraction of diseased cases on which the disease is visible.

Value

CBM-predicted ROC-AUC for the specified parameters.

References


Examples

```r
mu <- 2; alpha <- 0.8
UtilAucCBM(mu, alpha)
```
**UtilAucPROPROC**  
*PROPROC AUC function*

### Description

Returns the PROPROC ROC-AUC corresponding to specified parameters. See also UtilAucSRSM, UtilAucBinormal and UtilAucCBM.

### Usage

```r
UtilAucPROPROC(c1, da)
```

### Arguments

- **c1**: The c-parameter of the PROPROC model, since `c` is a reserved function in R.
- **da**: The da-parameter of the PROPROC model.

### Value

PROPROC model-predicted ROC-AUC for the specified parameters

### References


### Examples

```r
c1 <- .2; da <- 1.5
UtilAucPROPROC(c1, da)
```

---

**UtilAucSRSM**  
*RSM ROC/AFROC AUC calculator*

### Description

Returns the ROC and AFROC AUCs corresponding to specified RSM parameters. See also UtilAucPROPROC, UtilAucBinormal and UtilAucCBM

### Usage

```r
UtilAucSRSM(mu, lambdaP, nuP, lesDistr)
```
Arguments

mu
The mean(s) of the Gaussian distribution(s) for the ratings of latent LLs (continuous ratings of lesions that are found by the search mechanism).

lambdap
The Poisson distribution parameter(s), which describes the random number of latent NLs (suspicious regions that do not correspond to actual lesions) per case; these are the physical parameters.

nuP
The physical nuP parameters, each of which is the success probability of the binomial distribution(s) describing the random number of latent LLs (suspicious regions that correspond to actual lesions) per diseased case.

lesDistr
See PlotRsmOperatingCharacteristics.

Details

The RSM parameters (mu, lambdap and nuP) can be vectors, provided they are of the same length; the first parameter of each array is used, followed by the second, etc; a common lesion distribution is assumed.

Value

A list containing the ROC and AFROC AUCs corresponding to the specified parameters.

References


Examples

mu <- 1; lambdap <- 1; nuP <- 1
lesDistr <- rbnd(c(1, 0.9), c(2, 0.1))
## i.e., 90% of dis. cases have one lesion, and 10% have two lesions
UtilAucsRSM(mu, lambdap, nuP, lesDistr)$aucROC
UtilAucsRSM(mu, lambdap, nuP, lesDistr)$aucAFROC

mu <- c(1,2); lambdap <- c(1,0.5); nuP <- c(1, 0.8)
lesDistr <- rbnd(c(1, 0.9), c(2, 0.1))
## i.e., 90% of dis. cases have one lesion, and 10% have two lesions
UtilAucsRSM(mu, lambdap, nuP, lesDistr)$aucROC
UtilAucsRSM(mu, lambdap, nuP, lesDistr)$aucAFROC
UtilFigureOfMerit  

Calculate empirical figures of merit (FOMs) for specified dataset

Description

Calculate the specified empirical figure of merit for each treatment-reader combination in the ROC, FROC or LROC dataset.

Usage

UtilFigureOfMerit(dataset, FOM = "wAFROC", FPFValue = 0.2)

Arguments

dataset  
The dataset to be analyzed, see RJafroc-package.

FOM  
The figure of merit to be used in the calculation. The default is "wAFROC".

FPFValue  
Only needed for LROC data; where to evaluate a partial curve based figure of merit. The default is 0.2.

Details

The allowed FOMs depend on the type of dataset (i.e., dataType field of dataset object). For ROC datasets: the "wilcoxon" is allowed. For FROC datasets: The following FOMs are allowed: "AFROC1", "AFROC", "wAFROC1", "wAFROC" (the default), "HrAuc", "SongA1", "SongA2", "HrSe", "HrSp", "MaxLLF", "MaxNLF", "MaxNLFallCases", and "ExpTrnsfmSp". The "MaxLLF", "MaxNLF" and "MaxNLFallCases" FOMs correspond to ordinate, and abscissa, respectively, of the highest point on the FROC operating characteristic obtained by counting all the marks). The "ExpTrnsfmSp" FOM is described in the paper by Popescsu. Given the large number of FOMs possible with FROC data, it is appropriate to make a recommendation: it is recommended that one use the wAFROC FOM. For LROC datasets: The following FOMs are allowed: "wilcoxon" for ROC data inferred from LROC data, which ignores localization information; or "PCL" or "ALROC", in which case one needs to specify an additional argument, FPFValue: the desired FPF at which to evaluate PCL or ALROC; the default is 0.2.

Value

An c(I, J) array, where the row names are the IDs of the treatments and column names are the IDs of the readers.

References


Chakraborty DP, Berbaum KS (2004) Observer studies involving detection and localization: modeling, analysis, and validation, Medical Physics, 31(8), 1–18.


Examples

```r
# ROC data
UtilFigureOfMerit(dataset = dataset02, FOM = "Wilcoxon")
# FROC dataset, converted to ROC, Wilcoxon FOM
UtilFigureOfMerit(DFroc2Roc(dataset01), FOM = "Wilcoxon")
# FROC dataset, default wAFROC FOM
UtilFigureOfMerit(dataset = dataset01)
# LROC data
UtilFigureOfMerit(dataset = datasetCadLroc, FOM = "ALROC", FPFValue = 0.2)
```

---

**UtilIntrinsic2PhysicalRSM**

*Convert from intrinsic to physical RSM parameters*

---

**Description**

Convert intrinsic RSM parameters `lambda` and `nu` correspond to the physical RSM parameters `lambda'` and `nu'`. The physical parameters are more meaningful but they depend on `mu`. The intrinsic parameters are independent of `mu`. See book for details.

**Usage**

```r
UtilIntrinsic2PhysicalRSM(mu, lambda, nu)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mu</td>
<td>The mean of the Gaussian distribution for the ratings of latent LLs, i.e. continuous ratings of lesions that were found by the search mechanism ~ N(μ,1). The corresponding distribution for the ratings of latent NLs is N(0,1).</td>
</tr>
<tr>
<td>lambda</td>
<td>The Poisson intrinsic parameter, related to λ', the latter is the mean of the Poisson distribution of numbers of latent NLs (suspicious regions that do not correspond to actual lesions) per case.</td>
</tr>
<tr>
<td>nu</td>
<td>The intrinsic ν parameter; the corresponding physical parameter is the success probability of the binomial distribution of random numbers of latent LLs (suspicious regions that correspond to actual lesions) per diseased case, i.e., the chance that a lesion is &quot;found&quot;.</td>
</tr>
</tbody>
</table>
Details

RSM is the Radiological Search Model described in the book. A latent mark becomes an actual
mark if the corresponding rating exceeds the lowest reporting threshold $\zeta$. See also UtilPhysical-
ical2IntrinsicRSM.

Value

A list containing $\lambda'$ and $\nu'$

References

Chakraborty DP (2006) A search model and figure of merit for observer data acquired according to
the free-response paradigm, Phys Med Biol 51, 3449–3462.

Chakraborty DP (2006) ROC Curves predicted by a model of visual search, Phys Med Biol 51,
3463–3482.

Chakraborty DP (2017) Observer Performance Methods for Diagnostic Imaging - Foundations,
crcpress.com/Observer-Performance-Methods-for-Diagnostic-Imaging-Foundations-Modeling/
Chakraborty/p/book/9781482214840

Examples

mu <- 2; lambda <- 20; nu <- 1.1512925
lambdaP <- UtilIntrinsic2PhysicalRSM(mu, lambda, nu)$lambdaP
nuP <- UtilIntrinsic2PhysicalRSM(mu, lambda, nu)$nuP
## note that the physical values are only constrained to be positive, but the physical variable nuP
## must obey $0 \leq nuP \leq 1$

UtilLesionDistribution

Lesion distribution matrix

Description

Extracts the lesion distribution matrix for a dataset.

Usage

UtilLesionDistribution(dataset)

Arguments

dataset The supplied dataset
Details

The lesion distribution matrix has Lmax rows and two columns. The first column contains the integers 1, 2, ..., Lmax and the second column contains the fraction of diseased cases with the number of lesions per case specified in the first column.

Value

The lesion distribution matrix

Examples

UtilLesionDistribution(dataset@1) # FROC data
UtilLesionDistribution(dataset@2) # ROC data
UtilLesionDistribution(datasetCadLroc) # LROC data

UtilLesionWeights

Lesion weights matrix

Description

Computes the lesion weights matrix, assuming equal weights.

Usage

UtilLesionWeights(lesDistr)

Arguments

lesDistr The supplied lesion distribution matrix

Value

The lesion weights matrix, see PlotRsmOperatingCharacteristics

Examples

UtilLesionWeights(UtilLesionDistribution(dataset@1)) # FROC data
UtilLesionWeights(UtilLesionDistribution(dataset@2)) # ROC data
UtilLesionWeights(UtilLesionDistribution(datasetCadLroc)) # LROC data
UtilMeanSquares

Calculate mean squares

Description

Calculates the mean squares used in the DBMH and ORH methods

Usage

UtilMeanSquares(dataset, FOM = "wJAFROC", method = "DBMH")

Arguments

dataset The dataset to be analyzed, see RJafroc-package.
FOM The figure of merit to be used in the calculation. The default is "wJAFROC". See UtilFigureOfMerit.
method The method, in which the mean squares are calculated. The two valid options are "DBMH" (default) and "ORH".

Details

For DBMH method, msT, msTR, msTC, msTRC will not be available if the dataset contains only one treatment. Similarly, msR, msTR, msRC, msTRC will not be returned for single reader dataset. For ORH method, msT, msR, msTR will be returned for multiple reader multiple treatment dataset. msT is not available for single treatment dataset, and msR is not available for single reader dataset.

Value

A list containing all possible mean squares

Examples

UtilMeanSquares(dataset02, FOM = "Wilcoxon")

UtilMeanSquares(dataset05, method = "ORH")
UtilOutputReport  Generate a formatted report file

Description
Generate a formatted report of the analysis and save to a text file

Usage
UtilOutputReport (dataset, DataFileName, DataFileFormat, delimiter = ",", datadescription = "MyData", ReportFileName, ReportFileFormat = "txt", stMethod = "DBMH", FOM = "wJAFROC", alpha = 0.05, covEstMethod = "Jackknife", nBoots = 200, renumber = FALSE, overwrite = TRUE)

Arguments

- **dataset**: The dataset object to be analyzed (not the file name), RJafroc-package.
- **DataFileName**: A string variable containing the name of the data file to be analyzed, see DfReadDataFile and "Details".
- **DataFileFormat**: The format of the data specified in DataFileName: see DfReadDataFile and "Details". Must be specified if DataFileName is specified.
- **delimiter**: See DfReadDataFile.
- **dataDescription**: A description of the data, default is "MyData"
- **ReportFileName**: The file name of the output report file. If this parameter is missing, the function will use DataFileName or dataDescription followed by the underscore separated concatenation of stMethod and FOM as the output report file.
- **ReportFileFormat**: The format of the output report. The two available formats are "txt" (the default) and "xlsx", corresponding to a formatted text file and an Excel file, respectively.
- **stMethod**: The significance testing method, "ORH" or "DBMH".
- **FOM**: See StSignificanceTesting.
- **alpha**: See StSignificanceTesting.
- **covEstMethod**: See StSignificanceTesting.
- **nBoots**: See StSignificanceTesting.
- **renumber**: A logical variable: if TRUE, consecutive integers (starting from 1) will be used as the treatment and reader IDs in the output report. Otherwise, treatment and reader IDs in the original data file will be used. This option may be needed for aesthetics.
- **overwrite**: A logical variable: if FALSE, a warning will be issued if the report file already exists and the program will wait until the user inputs "y" or "n" to determine whether to overwrite the existing file. If TRUE, an existing file will be silently overwritten.
Details

See examples

Value

A formatted report of the data analysis, patterned roughly on that of OR-DBM MRMC V2.5.

Examples

UtilOutputReport(dataset = dataset03, stMethod = "DBMM", FOM = "Wilcoxon",
dataDescription = "MyROCData0", overwrite = TRUE)

## Not run:
## Generate reports for a dataset object
UtilOutputReport(dataset = dataset02, stMethod = "DBMM", FOM = "Wilcoxon",
dataDescription = "MyROCData1", overwrite = TRUE)

UtilOutputReport(dataset = dataset02, stMethod = "DBMM", FOM = "Wilcoxon",
dataDescription = "MyROCData2", ReportFileFormat = "xlsx", overwrite = TRUE)

UtilOutputReport(dataset = dataset02, stMethod = "ORH", FOM = "Wilcoxon",
dataDescription = "MyROCData3", overwrite = TRUE)

UtilOutputReport(dataset = dataset02, stMethod = "ORH", FOM = "Wilcoxon",
dataDescription = "MyROCData4", ReportFileFormat = "xlsx", overwrite = TRUE)

## Generate report for a data file
fn <- system.file("extdata", "includedRocData.xlsx",
package = "RJafroc", mustWork = TRUE)
UtilOutputReport(DataFileName = fn, DataFileFormat = "JAFROC", stMethod = "DBMM", FOM = "Wilcoxon",
overwrite = TRUE, ReportFileFormat = "xlsx")

## Output report for an existing dataset
## UtilOutputReport(dataset = dataset05, stMethod = "DBMM", FOM = "Wilcoxon") # ERROR! as FOM is incompatible with FROC data

UtilOutputReport(dataset = dataset05, stMethod = "ORH") # OK as default FOM is "wJAFROC"

UtilOutputReport(dataset = dataset05, stMethod = "DBMM", FOM = "HrAuc")

UtilOutputReport(dataset = dataset05, stMethod = "DBMM", FOM = "HrAuc", ReportFileFormat = "xlsx")

## End(Not run)
Description

Convert physical RSM parameters $\lambda'$ and $\nu'$ to the intrinsic RSM parameters $\lambda$ and $\nu$. The physical parameters are more meaningful but they depend on $\mu$. The intrinsic parameters are independent of $\mu$. See book for details.

Usage

UtilPhysical2IntrinsicRSM(mu, lambdaP, nuP)

Arguments

mu
The mean of the Gaussian distribution for the ratings of latent LLs, i.e. continuous ratings of lesions that were found by the search mechanism $\sim N(\mu,1)$. The corresponding distribution for the ratings of latent NLs is $N(0,1)$

lambdaP
The Poisson physical parameter, which describes the distribution of random numbers of latent NLs (suspicious regions that do not correspond to actual lesions) per case; the mean of these random numbers asymptotically approaches lambdaP

nuP
The physical $\nu$ parameter; it is the success probability of the binomial distribution describing the random number of latent LLs (suspicious regions that correspond to actual lesions) per diseased case

Details

@usage UtilIntrinsic2PhysicalRSM (mu, lambda, nu)

RSM is the Radiological Search Model described in the book. A latent mark becomes an actual mark if the corresponding rating exceeds the lowest reporting threshold zeta1. See also UtilIntrinsic2PhysicalRSM.

Value

A list containing $\lambda$ and $\nu$, the physical parameters

References


Examples

\[
\mu \leftarrow 2; \lambda \leftarrow 10; \nu \leftarrow 0.9
\]
\[
\lambda \leftarrow \text{UtilPhysical2IntrinsicRSM}(\mu, \lambda \mu, \nu \mu) \lambda
\]
\[
\nu \leftarrow \text{UtilPhysical2IntrinsicRSM}(\mu, \lambda \mu, \nu \mu) \nu
\]
## note that the physical values are only constrained to be positive, e.g., \( \nu \) is not constrained
## to be between 0 and 1.

---

**UtilPseudoValues** Calculate pseudovalues

**Description**

Calculates centered pseudovalues using the jackknife

**Usage**

```
UtilPseudoValues(dataset, FOM = "wJAFROC")
```

**Arguments**

- `dataset`: The dataset to be analyzed, see RJafroc-package.
- `FOM`: The figure of merit to be used in the calculation. The default is "wJAFROC". See UtilFigureOfMerit.

**Value**

An \( c(I, J, K) \) array containing the pseudovalues of the datasets.

**Examples**

```
UtilPseudoValues(dataset02, FOM = "Wilcoxon")[1,1:10]
```

```
UtilPseudoValues(dataset05)[1,1:10] # default FOM is wAFROC for this FROC dataset
```
Index

*Topic datasets
  binnedData123, 15
  binnedData124, 16
  binnedData125, 17
  dataset01, 20
  dataset02, 21
  dataset03, 22
  dataset04, 23
  dataset05, 24
  dataset06, 25
  dataset07, 26
  dataset08, 27
  dataset09, 28
  dataset10, 29
  dataset11, 30
  dataset12, 31
  dataset13, 32
  dataset14, 33
  dataset15, 34
  datasetCrossedModality, 35
  datasetDegenerate, 36
  Df2RJafrocDataset, 11, 36
  DfBinDataset, 11, 38, 57
  DfCreateCorCbmdataset, 11, 15-17, 40
  DfExtractCorCbmdataset, 12, 41, 55
  DfExtractDataset, 11, 42
  DfFroc2Afroc, 11, 43
  DfFroc2Roc, 11, 44, 57
  DfLroc2Roc, 11, 46
  DfReadCrossedModalities, 11, 47
  DfReadDatasetFile, 10, 11, 48, 92
  DfReadLrocDatasetFile, 12, 49
  DfSaveDatasetFile, 12, 50
  FitBinormalRoc, 12, 51
  FitCbmRoc, 12, 19, 53
  FitCorCbm, 11, 12, 15-17, 40-42, 55, 65
  FitRsmRoc, 12, 19, 56
  PlotBinormalFit, 12, 58
  PlotCbmFit, 59
  PlotEmpiricalOperatingCharacteristics, 12, 60
  PlotRsmOperatingCharacteristics, 12, 62, 86, 90
  RJafroc-package, 4, 37, 39, 43, 44, 47-49, 51, 75, 87, 91, 92, 95
  SimulateCorCbmdataset, 12, 65
  SimulateFrocDataset, 12, 66
  SimulateRocDataset, 12, 67
  SsPowerGivenJK, 12, 69, 71, 72
  SsPowerTable, 12, 70
  SsSampleSizeKGivenJ, 12, 72
INDEX

StSignificanceTesting, 12, 74, 81, 92
StSignificanceTestingCadVsRadiologists, 12, 78
StSignificanceTestingCrossedModalities, 12, 35, 81
StSignificanceTestingSingleFixedFactor, 12, 28, 34, 82

UtilAucBinormal, 13, 83, 84, 85
UtilAucCBM, 13, 83, 84, 85
UtilAucPROPROC, 13, 83–85, 85
UtilAucSRS, 13, 83–85, 85
UtilFigureOfMerit, 13, 75, 76, 82, 87, 91, 95
UtilIntrinsic2PhysicalRSM, 13, 62, 88, 94
UtilLesionDistribution, 13, 89
UtilLesionWeights, 13, 90
UtilMeanSquares, 13, 91
UtilOutputReport, 13, 74, 92
UtilPhysical2IntrinsicRSM, 13, 62, 89, 93
UtilPseudoValues, 13, 95