# Package: MRMCaov (via r-universe)

September 9, 2024

**Type** Package

**Title** Multi-Reader Multi-Case Analysis of Variance

**Version** 0.3.0

**Date** 2023-01-09

**Author** Brian J Smith [aut, cre], Stephen L Hillis [aut], Lorenzo L Pesce [ctb]

**Maintainer** Brian J Smith `<brian-j-smith@uiowa.edu>`

**Description** Estimation and comparison of the performances of diagnostic tests in multi-reader multi-case studies where true case statuses (or ground truths) are known and one or more readers provide test ratings for multiple cases. Reader performance metrics are provided for area under and expected utility of ROC curves, likelihood ratio of positive or negative tests, and sensitivity and specificity. ROC curves can be estimated empirically or with binormal or binormal likelihood-ratio models. Statistical comparisons of diagnostic tests are based on the ANOVA model of Obuchowski-Rockette and the unified framework of Hillis (2005) <doi:10.1002/sim.2024>. The ANOVA can be conducted with data from a full factorial, nested, or partially paired study design; with random or fixed readers or cases; and covariances estimated with the DeLong method, jackknifing, or an unbiased method. Smith and Hillis (2020) <doi:10.1117/12.2549075>.

**Depends** R (>= 3.5.0)

**Imports** ggplot2, methods, mvtnorm, progress, tibble, trust

**Suggests** knitr

**LazyData** true

**License** GPL-3

**URL** `https://github.com/brian-j-smith/MRMCaov`

**BugReports** `https://github.com/brian-j-smith/MRMCaov/issues`

**RoxygenNote** 7.2.3

**VignetteBuilder** knitr

**Encoding** UTF-8

**Repository** https://brian-j-smith.r-universe.dev

**RemoteUrl** https://github.com/brian-j-smith/mrmcaov

**RemoteRef** HEAD

**RemoteSha** d79e49bcef42ef2a17195aae84ca84ae569f8702

# Contents

---

## `MRMCaov-package` *MRMCaov: Multi-Reader Multi-Case Analysis of Variance*

### Description

Estimation and comparison of the performances of diagnostic tests in multi-reader multi-case studies where true case statuses (or ground truths) are known and one or more readers provide test ratings for multiple cases. Reader performance metrics are provided for area under and expected utility of ROC curves, likelihood ratio of positive or negative tests, and sensitivity and specificity. ROC curves can be estimated empirically or with binormal or binormal likelihood-ratio models. Statistical comparisons of diagnostic tests are based on the ANOVA model of Obuchowski-Rockette and the unified framework of Hillis (2005) doi:10.1002/sim.2024. The ANOVA can be conducted with data from a full factorial, nested, or partially paired study design; with random or fixed readers or cases; and covariances estimated with the DeLong method, jackknifing, or an unbiased method. Smith and Hillis (2020) doi:10.1117/12.2549075.

## Details

The functions below are available in **MRMCaov** for estimation and comparison of test performance metrics in studies involving multiple cases and one or more readers. Examples of their use can be found in the online guide at `https://brian-j-smith.github.io/MRMCaov/`.

Statistical Inference:

| | |
|---|---|
| `mrmc` | Multi-reader multi-case ANOVA |
| `srmc` | Single-reader multi-case ANOVA |
| `stmc` | Single-test (single-reader) multi-case Estimation |

Tabular and Graphical Summaries:

| | |
|---|---|
| `parameters` | ROC curve parameters |
| `plot` | ROC curve plots |
| `roc_curves` | ROC curves |
| `summary` | Statistical analysis summaries |

Performance Metrics (Binary Rating):

| | |
|---|---|
| `binary_sens` | Sensitivity |
| `binary_spec` | Specificity |

Performance Metrics (Ordinal or Numeric Rating):

| | |
|---|---|
| `binormal_auc` | Binormal ROC AUC |
| `binormal_sens` | ... sensitivity |
| `binormal_spec` | ... specificity |
| `binormalLR_auc` | Binormal likelihood ratio ROC AUC |
| `binormalLR_sens` | ... sensitivity |
| `binormalLR_spec` | ... specificity |
| `empirical_auc` | Empirical ROC AUC |
| `empirical_sens` | ... sensitivity |
| `empirical_spec` | ... specificity |
| `trapezoidal_auc` | Empirical ROC AUC |
| `trapezoidal_sens` | ... sensitivity |
| `trapezoidal_spec` | ... sensitivity |

Performance Metric Covariance Estimation Methods:

`DeLong`
`jackknife`
`unbiased`

ROC Curves:

| | |
|---|---|
| `roc_curves` | Estimate one or more curves |
| `parameters` | Extract curve parameters |
| `points` | Extract curve points |
| `mean` | Compute the mean of multiple curves |
| `plot` | Plot curves |

Conversion of MRMC Model Parameters:

| | |
|---|---|
| `OR_to_RMH` | Obuchowski-Rockette to Roe, Metz & Hillis parameters |
| `RMH_to_OR` | Roe, Metz & Hillis to Obuchowski-Rockette parameters |

## Note

This research was supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) of the National Institutes of Health under Award Number R01EB025174

## Author(s)

**Maintainer**: Brian J Smith <`brian-j-smith@uiowa.edu`>

Authors:

- Stephen L Hillis <`steve-hillis@uiowa.edu`>

Other contributors:

- Lorenzo L Pesce <`lorenzo.pesce@northwestern.edu`> [contributor]

## See Also

Useful links:

- `https://github.com/brian-j-smith/MRMCaov`
- Report bugs at `https://github.com/brian-j-smith/MRMCaov/issues`

---

## `cov_methods` *Covariance Methods*

---

### Description

Reader performance metric covariance estimation methods to be used with package-supplied multicase statistical analysis functions.

### Usage

```
DeLong()

jackknife()

unbiased(abar = FALSE)
```

### Arguments

| | |
|---|---|
| `abar` | logical indicating whether to compute mean covariance components. |

### Value

Returns a function of class `cov_method` specifying a covariance method for `mrmc`, `srmc`, or `stmc`.

### References

DeLong ER, DeLong DM, and Clarke-Pearson DL (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics, 44: 837–45.

Efron B (1982). The Jackknife, the Bootstrap and Other Resampling Plans. Philadelphia: SIAM.

Gallas BD, Pennello GA, and Meyers KJ (2007). Multireader multicase variance analysis for binary data. JJournal of the Optical Society of America A, 24: B70–80.

### See Also

`mrmc`, `srmc`, `stmc`

---

`Franken` *Multi-reader multi-case dataset*

---

## Description

Multi-reader multi-case dataset

## Usage

```
Franken
```

## Format

A data frame with 800 rows and 5 variables:

**reader** reader identifier

**treatment** treatment identifier

**case** case identifier

**truth** true case status (1 = abnormal, 0 = normal)

**rating** ordinal reader ratings of abnormal case status (1 = definitely normal, 5 = definitely abnormal)

## References

Franken EA Jr, Berbaum KS, Marley SM, Smith WL, Sato Y, Kao SC, Milam SG (1992). Evaluation of a digital workstation for interpreting neonatal examinations: a receiver operating characteristic study. Investigational Radiology, 27(9): 732-737.

---

`Kundel` *Multi-reader multi-case dataset*

---

## Description

Data from a multi-reader multi-case study conducted by Kundel et al. (1997) to compare the clinical diagnostic accuracy of hard-copy readings of screen-film bedside chest radiographs and both hard- and soft-copy readings of computed radiographs from separate groups of patients in a medical intensive care unit.

## Usage

```
Kundel
```

## Format

A data frame with 760 rows and 5 variables:

**reader** reader identifier

**treatment** numeric identifier for treatment modality

**treatment_name** name of treatment modality

**case** patient identifier

**truth** true case status (1 = diseased, 0 = non-diseased)

**rating** ordinal reader ratings of case status

## References

Kundel HL, Gefter W, Aronchick J, Miller W Jr, Hatabu H, Whitfill CH, Miller W Sr (1997). Accuracy of bedside chest hard-copy screen-film versus hard- and soft-copy computed radiographs in a medical intensive care unit: receiver operating characteristic analysis. Radiology, 205(3): 859-63.

---

# `metrics` *Performance Metrics*

## Description

Estimated performance metrics from ROC curves.

## Usage

```
binary_sens(truth, rating)

binary_spec(truth, rating)

binormal_auc(
  truth,
  rating,
  partial = FALSE,
  min = 0,
  max = 1,
  normalize = FALSE
)

binormal_eu(truth, rating, slope = 1)

binormal_sens(truth, rating, spec)

binormal_spec(truth, rating, sens)
```

```
binormalLR_auc(
  truth,
  rating,
  partial = FALSE,
  min = 0,
  max = 1,
  normalize = FALSE
)

binormalLR_eu(truth, rating, slope = 1)

binormalLR_sens(truth, rating, spec)

binormalLR_spec(truth, rating, sens)

empirical_auc(
  truth,
  rating,
  partial = FALSE,
  min = 0,
  max = 1,
  normalize = FALSE
)

empirical_eu(truth, rating, slope = 1)

empirical_sens(truth, rating, spec)

empirical_spec(truth, rating, sens)

trapezoidal_auc(
  truth,
  rating,
  partial = FALSE,
  min = 0,
  max = 1,
  normalize = FALSE
)

trapezoidal_sens(truth, rating, spec)

trapezoidal_spec(truth, rating, sens)
```

### Arguments

| | |
|---|---|
| `truth` | vector of true binary statuses. |
| `rating` | vector of 0-1 binary ratings for the binary metrics and ranges of numeric ratings for the others. |

| | |
|---|---|
| `partial` | character string `"sensitivity"` or `"specificity"` for calculation of partial AUC, or `FALSE` for full AUC. Partial matching of the character strings is allowed. `"specificity"` results in area under the ROC curve between the given `min` and `max` specificity values, whereas `"sensitivity"` results in area to the right of the curve between the given sensitivity values. |
| `min`, `max` | minimum and maximum sensitivity or specificity values over which to calculate partial AUC. |
| `normalize` | logical indicating whether partial AUC is divided by the interval width (`max - min`) over which it is calculated. |
| `slope` | slope of the iso-utility line at which to compute expected utility of the ROC curve. |
| `sens`, `spec` | numeric sensitivity/specificity at which to calculate specificity/sensitivity. |

## Details

Performance metrics measure the degree to which higher case ratings are associated with positive case statuses, where positive status is taken to be the highest level of `truth`. Available metrics include area under the ROC curve (auc), expected utility of the ROC curve (eu) at a given iso-utility line (Abbey, 2013), sensitivity (sens) at a given specificity, and specificity (spec) at a given sensitivity.

## Value

Returns a numeric value.

## References

Abbey CK, Samuelson FW and Gallas BD (2013). Statistical power considerations for a utility endpoint in observer performance studies. Academic Radiology, 20: 798-806.

## See Also

`mrmc`, `srmc`, `stmc`

---

# `mrmc` *Multi-Reader Multi-Case ROC Analysis*

## Description

Estimation and comparison of ROC performance metrics for multi-reader multi-case studies.

## Usage

```
mrmc(response, test, reader, case, data, cov = jackknife, design = NULL)
```

### Arguments

| | |
|---|---|
| `response` | response metric expressed in terms of a package-supplied performance `metric`. |
| `test` | variable of test identifiers. |
| `reader` | variable of reader identifiers. |
| `case` | variable of case identifiers. |
| `data` | data frame containing the `response`, `test`, `reader`, and `case` variables. |
| `cov` | function, function call, or character string naming the `method` to use in calculating performance metric covariances. |
| `design` | one of the following study designs: 1 = factorial, 2 = cases nested within readers, 3 = cases nested within tests, or `NULL` to automatically set the design based on variable codings in data. |

### Details

Readers and cases are treated as random factors by default. Either one may be designated as fixed in calls to `mrmc` with the syntax `fixed(<variable name>)`, where `<variable name>` is the name of the reader or case variable.

### Value

Returns an `mrmc` class object with the following elements.

`design` experimental study design: 1 = factorial, 2 = cases nested within readers, 3 = cases nested within tests.

`vars` character names of the analysis factors and reader performance metric.

`fixed` logicals indicating whether the reader and case factors are treated as fixed in the analysis.

`aov` results from an ordinary analysis of variance.

`data` data frame of computed reader performance metrics for the analysis of variance.

`num_obs` number of case observations for each of the computed metrics.

`cov` reader performance covariance matrix.

`mrmc_data` data frame of case-specific reader ratings.

`levels` character levels of the true case statuses.

### References

Dorfman DD, Berbaum KS, and Metz CE (1992). Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. Investigative Radiology, 27: 723–731.

Obuchowski NA and Rockette HE (1995). Hypothesis testing of diagnostic accuracy for multiple readers and multiple tests: an ANOVA approach with dependent observations. Communications in Statistics–Simulation and Computation 24: 285–308.

Hillis SL, Obuchowski NA, Schartz KM, and Berbaum KS (2005). A comparison of the Dorfman-Berbaum-Metz and Obuchowski-Rockette methods for receiver operating characteristic (ROC) data. Statisticsin Medicine, 24: 1579–1607.

Hillis SL (2007). A comparison of denominator degrees of freedom methods for multiple observer ROC analysis. Statistics in Medicine, 26: 596–619.

Hillis SL, Berbaum KS, and Metz CE (2008). Recent developments in the Dorfman-Berbaum-Metz procedure for multireader ROC study analysis. Academic Radiology, 15: 647–661.

### See Also

`metrics`, `cov_methods`, `parameters`, `plot`, `roc_curves`, `summary`

### Examples

```
## Random readers and cases
(est <- mrmc(empirical_auc(truth, rating), treatment, reader, case,
             data = VanDyke))
plot(est)
summary(est)

## Fixed readers and random cases
est <- mrmc(empirical_auc(truth, rating), treatment, fixed(reader), case,
            data = VanDyke)
summary(est)
```

---

## `OR_to_RMH` *Convert Obuchowski-Rockette Parameters to Roe & Metz Parameters*

### Description

Determines Roe & Metz (RM) simulation model parameters for simulating multireader multicase likelihood-of-disease rating data based on real-data or conjectured Obuchowski-Rockette (OR) parameter estimates that describe the distribution of the empirical AUC reader performance measure. The algorithm assumes the constrained unequal-variance RM model (Hillis, 2012) which generalizes the original RM model (Roe and Metz, 1997) by allowing the diseased and nondiseased decision-variable distributions to have unequal variances for each reader, with the variance components involving diseased cases constrained to differ by a factor of 1/b^2 from corresponding variance components involving nondiseased cases. This algorithm is described in Hillis (2020). *Throughout we refer to the Hillis (2012) RM model as the **RMH model**.*

### Usage

```
OR_to_RMH(...)

## Default S3 method:
OR_to_RMH(
  ...,
  AUC1,
  AUC2,
```

```
  var_R,
  var_TR,
  corr1,
  corr2,
  corr3,
  var_error = NULL,
  n0,
  n1,
  b_method = c("unspecified", "mean_to_sigma", "specified"),
  mean_sig_input = NULL,
  b_input = NULL,
  b_le_1 = TRUE
)

## S3 method for class 'data.frame'
OR_to_RMH(params, ...)
```

## Arguments

| | |
|---|---|
| `...` | arguments passed to the default method. |
| `AUC1`, `AUC2` | test 1 and 2 expected empirical AUCs. |
| `var_R`, `var_TR` | OR reader and test-by-reader variance components. |
| `corr1`, `corr2`, `corr3` | OR error correlations. |
| `var_error` | OR error variance. |
| `n0`, `n1` | number of nondiseased and diseased cases. |
| `b_method` | method of estimating RMH b parameter. |
| `mean_sig_input` | mean-to-sigma ratio, required only if `b_method` = `"mean_to_sigma"`. |
| `b_input` | binormal *b* value, required only if `b_method` = `"specified"`. |
| `b_le_1` | logical indicating whether the algorithm searches first for b <= 1 and then, if no solution, for b >= 1; if FALSE, the algorithm searches only for for b >= 1. Required only if `b_method` = `"unspecified"`. |
| `params` | data frame of above OR parameter values in the columns. |

## Details

Hillis (2012) modified the original RM model (Roe and Metz, 1997) by allowing variance components involving case to depend on truth (diseased/nondiseased), with variance components involving diseased cases set equal to those involving nondiseased cases multiplied by the factor 1/b^2, b>0. *We refer to the Hillis (2012) model as the **RMH** model.* Hillis (2018) derived analytical formulas that express OR parameters describing the distribution of empirical AUC outcomes computed from RMH simulated data as functions of the RMH model parameters. The reverse mapping from the RMH parameters to the OR parameters is implemented in R by the RMH_to_OR function. The OR_to_RMH function uses an iterative search procedure.

`b_method` indicates the method for estimating the RMH *b* parameter. Note that *b* is the conventional binormal-curve slope, i.e., the slope of each reader's true ROC curve plotted in probit space.

- `b_method = "unspecified"` should be used when the goal is to determine RM parameters that result in simulated data for which the empirical AUC distribution is described by the inputted values for the OR parameter vector

  $$\beta_O R = (AUC1, AUC2, var_R, var_T R, var_e rror, corr1, corr2, corr3).$$

- `b_method = "mean_to_sigma"` should be used when the goal is to determine RM parameters that result in simulated data for which the empirical AUC distribution is described by the inputted values for the OR parameter vector

  $$\beta 1_O R = (AUC1, AUC2, var_R, var_T R, corr1, corr2, corr3),$$

  and such that the median mean-to-sigma ratio across readers is equal to `mean_sig_input` for the test having the lowest AUC. Note that $\beta 1_O R$ differs from $\beta_O R$ in that it does not contain the OR error variance.

- `b_method = "specified"` should be used when the goal is to determine RM parameters that result in simulated data for which the empirical AUC distribution is described by the inputted values for the OR parameter vector $\beta 1_O R$ (see above) with *b* equal to `mean_sig_input`. (E.g., set `b_input = 1` for symmetric ROC curve.)

For `b_method = "mean_to_sigma"` or `"specified"`, the simulated empirical AUC estimate distribution is specified by the parameter values in `params`, except for `var_error`. Thus for these two options, `var_error` can be equal to `NA` or excluded from `params`.

Parameter `mean_sig_input` is the inputted mean-to-sigma ratio needed for `b_method = "mean_to_sig"`. See Hillis & Berbaum (2011) for more information.

Parameter `b_input` is the inputted binormal *b* value needed for `b_method = "specified"`.

There may not be a solution for a set of OR parameters values. When this occurs, the function will either produce an approximate solution or indicate what OR input needs to be changed.

A related function is the RMH_to_OR function, which determines OR parameters that describe the distribution of empirical AUC estimates computed from inputted RM model parameter values, based on the analytical mapping provided by Hillis (2018).

## Value

The RMH model parameters are returned in a data frame with the following elements.

**delta1** mean separation of nondiseased and diseased decision-variable distributions for test 1 across reader population.

**delta2** mean separation of nondiseased and diseased decision-variable distributions for test 2 across reader population.

**var_R** RMH reader variance compnent.

**var_TR** RMH text-by-reader variance component.

**var_C** RMH case variance component.

**var_TC** RMH test-by-case variance.

**var_RC** RMH reader-by-case variance.

**var_error** RMH error variance.

**b** variance components involving diseased cases are constrained to differ by a factor of 1/b^2 from corresponding variance components involving nondiseased cases.

Related quantities that are also returned in the data frame:

**b_method** method used to estimate b.

**n0** number of nondiseased cases per simulated sample.

**n1** number of diseased cases per simulated sample.

**mean_to_sig1** expected mean-to-sigma ratio across readers for test 1.

**mean_to_sig2** expected mean-to-sigma ratio across readers for test 2.

**Pr1_improper** probability that the test 1 ROC curve for a random reader will be noticeably improper (i.e, |mean-to-sigma ratio| < 2).

**Pr2_improper** probability that the test 2 ROC curve for a random reader will be noticeably improper (i.e, |mean-to-sigma ratio| < 2).

## Author(s)

Stephen L. Hillis, Departments of Radiology and Biostatistics, University of Iowa, `<steve-hillis@uiowa.edu>`

Brian J. Smith, Department of Biostatistics, University of Iowa, `<brian-j-smith@uiowa.edu>`

## References

Hillis SL (2012). Simulation of unequal-variance binormal multireader ROC decision data: an extension of the Roe and Metz simulation model. *Academic Radiology*, 19(12): 1518-1528. doi: 10.1016/j.acra.2012.09.011

Hillis SL (2018). Relationship between Roe and Metz simulation model for multireader diagnostic data and Obuchowski-Rockette model parameters. *Statistics in Medicine*, 37(13): 2067-2093. doi: 10.1002/sim.7616

Hillis SL (2020). Determining Roe and Metz model parameters for simulating multireader multicase confidence-of-disease rating data based on read-data or conjectured Obuchowski-Rockette parameter estimates. Vol. 11316, SPIE Medical Imaging: SPIE. doi.org/10.1117/12.2550541

Hillis SL and Kevin SB (2011). Using the mean-to-sigma ratio as a measure of the improperness of binormal ROC curves. *Academic Radiology*, 18(2): 143-154. doi: 10.1016/j.acra.2010.09.002

Roe CA and Metz CE (1997). Dorfman-Berbaum-Metz method for statistical analysis of multireader, multimodality receiver operating characteristic data: validation with computer simulation. *Academic Radiology*, 4(4): 298-303. doi: 10.1016/S1076-6332(97)80032-3

## See Also

`RMH_to_OR`

### Examples

```
## Example 1: Computing RM parameters from OR parameters directly
##-----------------------------------------------------------------
## Example 1a: Using b_method ="unspecified" (the default)
RM <- OR_to_RMH(n0 = 69, n1 = 45, AUC1 = 0.897, AUC2 = 0.941,
                corr1 = 0.433, corr2 = 0.430, corr3 = 0.299,
                var_R = 0.00154, var_TR = 0.000208, var_error = 0.000788)
RM
##  We recommend also computing the OR parameter values ("true values")
# that describe the distribution of simulated data based on above RM parameters,
# using the RMH_to_OR function. Ideally the true values will be the same as the
# inputted OR values used for deriving the RM parameter values. We recommend
# always performing this check.  This check is carried out below, as shown below.
true_values = RMH_to_OR(RM)
true_values
#   From the output we see, for this example, that the true OR values are identical to the
# inputted OR values


# Example 1b: Using b_method = "specified" with b_input = 1
#   Note that the error variance does not need to be specified since this b_method
# does not utilize it.
RM <- OR_to_RMH(n0 = 69, n1 = 45, AUC1 = 0.897, AUC2 = 0.941,
                corr1 = 0.433, corr2 = 0.430, corr3 = 0.299,
                var_R = 0.00154, var_TR = 0.000208,
                b_method = "specified", b_input = 1)
RM
true_values <- RMH_to_OR(RM)
true_values
#  From the output we see, for this example, that the true values are identical
# (within rounding error) to the inputted OR values (but note that var_error was
# not inputted)

## Example 1c: Using b_method = "mean_to_sigma" with mean_to_sig_input = 4.5
#   Note the error variance does not need to be  specified since this b_method
# does not utilize it.
RM <- OR_to_RMH(n0 = 69, n1 = 45, AUC1 = 0.897, AUC2 = 0.941,
                corr1 = 0.433, corr2 = 0.430, corr3 = 0.299,
                var_R = 0.00154, var_TR = 0.000208,
                b_method = "mean_to_sigma", mean_sig_input = 4.5)
RM
true_values <- RMH_to_OR(RM)
true_values
#   From the output we see for this example that the true OR values are identical
# (within rounding error) to the inputted OR values (but note that var_error was
# not inputted)

##------------------------------------------------------------------------

## Example 2: Computing RM parameters from a data frame of OR parameters
## ------------------------------------------------------------------------
## Example 2a: One study
```

```
vandyke_OR <- data.frame(n0 = 69, n1 = 45, AUC1 = 0.897, AUC2 = 0.941,
                         corr1 = 0.433, corr2 = 0.430, corr3 = 0.299,
                         var_R = 0.00154, var_TR = 0.000208, var_error = 0.000788)
vandyke_RM <- OR_to_RMH(vandyke_OR)
vandyke_RM
true_values <- RMH_to_OR(vandyke_RM)
true_values

## Example 2b: Three studies
three_studies_OR <- data.frame(
  rbind(
    vandyke = c(69, 45, 0.89793704, 0.94083736, 0.432, 0.429, 0.298, 0.00154,
                0.0002, 0.00080229),
    franken = c(33, 67, .8477498869, 0.8368950701, 0.521430051, 0.319691199,
                0.3386375697, 0.0000433385, 0.0, 0.0014967254),
    kundel = c(66, 29, 0.8038793103, 0.8413662487, 0.507695244, 0.3843523762,
               0.4035662578, 0.0007340122, 0, 0.002148844)
  )
)
colnames(three_studies_OR) <- c("n0", "n1", "AUC1", "AUC2", "corr1", "corr2",
                                "corr3", "var_R", "var_TR", "var_error")
three_studies_OR
three_studies_RM <- OR_to_RMH(three_studies_OR)
three_studies_RM
true_values <- RMH_to_OR(three_studies_RM)
true_values
##   Note above that the true values for corr2 and corr3 for the Franken study
# differ slightly from the inputted values; this is because corr2 < corr3 for the
# inputted OR values, which is not possible for simulated RM model data.

##Example 2c: Examples 1a, 1b and 1c run using one data frame
vandyke_OR <- data.frame(n0 = 69, n1 = 45, AUC1 = 0.897, AUC2 = 0.941,
                         corr1 = 0.433, corr2 = 0.430, corr3 = 0.299,
                         var_R = 0.00154, var_TR = 0.000208, var_error = 0.000788)
vandyke_OR_x3 <- vandyke_OR[c(1,1,1),] #has 3 rows, each equal to vandyke_OR
b_method = c("unspecified","mean_to_sigma","specified")
mean_sig_input = c(NA,4.5,NA)
b_input = c(NA,NA,1)
vandyke_OR_3ex <- cbind(vandyke_OR_x3,b_method,mean_sig_input,b_input)
vandyke_OR_3ex
vandyke_OR_3ex_RM <- OR_to_RMH(vandyke_OR_3ex)
vandyke_OR_3ex_RM
true_values <- RMH_to_OR(vandyke_OR_3ex_RM)
true_values


## Example 3: Printing the alternative x1 -- x7 parameter values
## -------------------------------------------------------
## The OR_to_RMH function first finds the solutions using the alternative RM
## parameterization consisting of b and the alternative parameters
## x1, x2, x3, x4, x5, x6, and x7, and then solves for the conventional RM
## parameters in terms of these alternative parameters.  (See Hillis (2020) for details.)
## Although the user generally has no need to know these parameter values, they
```

```
## can be printed out using the all = TRUE print option, as shown below
## using Example 1a:

RM <- OR_to_RMH(n0 = 69, n1 = 45, AUC1 = 0.897, AUC2 = 0.941,
               corr1 = 0.433, corr2 = 0.430, corr3 = 0.299,
               var_R = 0.00154, var_TR = 0.000208, var_error = 0.000788)
print(RM,all = TRUE)
```

---

## `plot` *ROC Plots*

### Description

Plots of ROC curves.

### Usage

```
## S3 method for class 'roc_curve'
plot(x, n = 100, emp_points = FALSE, ...)

## S3 method for class 'roc_curves'
plot(x, n = 100, emp_points = FALSE, ...)

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

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

## S3 method for class 'roc_points'
plot(x, coord_fixed = TRUE, ...)

## S3 method for class 'mrmc'
plot(x, n = 100, ...)

## S3 method for class 'stmc'
plot(x, n = 100, ...)
```

### Arguments

| | |
|---|---|
| `x` | object to plot. |
| `n` | number of equally spaced false-positive rate points at which to calculate true-positive rates and interpolate through for display of the curve. |
| `emp_points` | logical indicating whether to overlay empirical ROC points on parametric curves. |
| `...` | arguments passed to other methods. |
| `coord_fixed` | logical indicating whether to fix the scales of x and y axes. |

### Value

Returns a ggplot object.

### See Also

`roc_curves`

### Examples

```
curves <- with(VanDyke,
  roc_curves(truth, rating, groups = list(Test = treatment, Reader = reader))
)
plot(curves)
```

---

## `print.roc_curve` *Print ROC Objects*

### Description

Print ROC objects from the **MRMCaov** package.

### Usage

```
## S3 method for class 'roc_curve'
print(x, n = 11, ...)

## S3 method for class 'roc_curves'
print(x, n_curves = 5, n = 11, ...)
```

### Arguments

| | |
|---|---|
| `x` | object to print. |
| `n` | number of ROC curve points to print. |
| `...` | arguments passed to other methods. |
| `n_curves` | number of ROC curves to print. |

### Value

Returns the printed object `x`.

### Examples

```
curves <- with(VanDyke,
  roc_curves(truth, rating, groups = list(Test = treatment, Reader = reader))
)
print(curves)
```

---

`RMH_to_OR` *Convert Roe & Metz Parameters to Obuchowski-Rockette Parameters*

---

## Description

Determines Obuchowski-Rockette (OR) model parameter values that describe the distribution of empirical AUC reader performance outcomes computed from multireader multicase likelihood-of-disease rating data simulated from the Roe & Metz (RM) simulation model, based on the analytical mapping provided by Hillis (2018). The function assumes the RM model proposed by (Hillis, 2012), which generalizes the orginal RM model (Roe and Metz, 1997) by allowing the latent confidence-of-disease rating distributions to have unequal diseased-case and nondiseased-case variances, with the variance components involving diseased cases constrained to differ by a factor of 1/b^2, b>0, from corresponding variance components involving nondiseased cases. *Throughout we refer to the Hillis (2012) RM model as the* ***RMH model****.*

## Usage

```
RMH_to_OR(...)

## Default S3 method:
RMH_to_OR(
  ...,
  n0,
  n1,
  b,
  delta1,
  delta2,
  var_R,
  var_TR,
  var_C,
  var_TC,
  var_RC,
  var_error
)

## S3 method for class 'data.frame'
RMH_to_OR(params, ...)
```

## Arguments

| | |
|---|---|
| `...` | arguments passed to the default method. |
| `n0`, `n1` | numbers of nondiseased and diseased cases. |
| `b` | b>0, with 1/b^2 = ratio of each diseased-case variance component to the corresponding diseased-case variance component. It follows that b is also the conventional binormal-curve slope, i.e., the slope of each reader's true ROC curve plotted in probit space. |

| | |
|---|---|
| `delta1`, `delta2` | test 1 and test 2 separations of the RMH-model nondiseased and diseased latent likelihood-of-disease rating distribution means. |
| `var_R`, `var_TR` | RMH-model reader and test-by-reader variance components. |
| `var_C`, `var_TC`, `var_RC`, `var_error` | RMH-model case, test-by-case, reader-by-case and error variance components for nondiseased cases. |
| `params` | data frame of above RM parameter values in the columns. |

## Details

Hillis (2012) modified the original RM model (Roe and Metz, 1997) by allowing variance components involving case to depend on truth (diseased/nondiseased), with variance components involving diseased cases set equal to those involving nondiseased cases multiplied by the factor 1/b^2, b>0. Assuming this model, *which we refer to as the **RMH model***, Hillis (2018) derived analytical formulas expressing OR parameters that describe the distribution of empirical AUC outcomes computed from RMH model simulated data as functions of the RMH parameters. This mapping from the RMH parameters to the OR parameters is implemented in R by the RMH_to_OR function.

A related function is the OR_to_RMH function, which determines RM parameter values corresponding to real-data or conjectured Obuchowski-Rockette (OR) parameter estimates.

## Value

The OR model parameters are returned in a data frame with the following elements.

**...** arguments passed to the default method.

**AUC1, AUC2** test 1 and 2 expected empirical AUCs.

**var_R, var_TR** OR reader and test-by-reader variance components.

**corr1, corr2, corr3** OR error correlations.

**var_error** OR error variance.

**n0, n1** number of nondiseased and diseased cases.

Related quantities describing the true reader ROC curves that are also returned in the data frame:

**b** b > 0, with 1/b^2 = (RM diseased variance component) / (corresponding RM nondiseased variance component).

**mean_to_sig1** expected mean-to-sigma ratio across readers for test 1.

**mean_to_sig2** expected mean-to-sigma ratio across readers for test 2.

**Pr1_improper** probability that the test 1 ROC curve for a random reader will be noticeably improper (i.e, |mean-to-sigma ratio| < 2).

**Pr2_improper** probability that the test 2 ROC curve for a random reader will be noticeably improper (i.e, |mean-to-sigma ratio| < 2).

## Author(s)

Stephen L. Hillis, Departments of Radiology and Biostatistics, University of Iowa, `<steve-hillis@uiowa.edu>`

Brian J. Smith, Department of Biostatistics, University of Iowa, `<brian-j-smith@uiowa.edu>`

## References

Hillis SL (2012). Simulation of unequal-variance binormal multireader ROC decision data: an extension of the Roe and Metz simulation model. *Academic Radiology*, 19(12): 1518-1528. doi: 10.1016/j.acra.2012.09.011

Hillis SL (2018). Relationship between Roe and Metz simulation model for multireader diagnostic data and Obuchowski-Rockette model parameters. *Statistics in Medicine*, 37(13): 2067-2093. doi: 10.1002/sim.7616

Roe CA and Metz CE (1997). Dorfman-Berbaum-Metz method for statistical analysis of multireader, multimodality receiver operating characteristic data: validation with computer simulation. *Academic Radiology*, 4(4): 298-303. doi: 10.1016/S1076-6332(97)80032-3

## See Also

`OR_to_RMH`

## Examples

```
##  Example 1: Computing OR parameters from RMH parameters directly
# RMH parameters from first line (A_z = 0.702) of Table 1 in Roe & Metz (1997)
# with 50 diseased and 50 nondiseased cases.
OR <- RMH_to_OR(n0 = 50, n1 = 50, delta1 = 0.75, delta2 = 0.75,
                var_R = 0.0055, var_TR = 0.0055, var_C = 0.3, var_TC = 0.3,
                var_RC = 0.2, var_error = 0.2, b = 1)
OR

##  Example 2: Computing OR parameters from a data frame of RMH parameters
##  -------------------------------------------------------------------
## Example 2a:  RMH parameters from first line (A_z = 0.702) of Table 1 in
# Roe & Metz (1997) with 50 diseased and 50 nondiseased cases
RM_parms_line1 <- data.frame(n0 = 50, n1 = 50, delta1 = 0.75, delta2 = 0.75,
                             var_R = 0.0055, var_TR = 0.0055, var_C = 0.3, var_TC = 0.3,
                             var_RC = 0.2, var_error = 0.2, b = 1)
OR <- RMH_to_OR(RM_parms_line1)
OR
## Note below that applying the OR_to_RMH function to the above OR parameters
# results in the original RMH parameters within rounding error:
check <- OR_to_RMH(OR)
check

## Example 2b: RMH parameters from last 3 lines of Table 1 in Roe & Metz (1997)
# using 10 diseased and 25 nondiseased cases
RM_3_models <- data.frame(
  rbind(
    line6 = c(25, 10, 0.75, 0.75, 0.011, 0.011, 0.1, 0.1, 0.2, 0.6, 1),
    line7 = c(25, 10, 1.50, 1.50, 0.03, 0.03, 0.1, 0.1, 0.2, 0.6, 1),
    line8 = c(25, 10, 2.5, 2.5, 0.056, 0.056, 0.1, 0.1, 0.2, 0.6, 1)
  )
)
colnames(RM_3_models) <- c("n0", "n1", "delta1", "delta2", "var_R", "var_TR",
                           "var_C", "var_TC", "var_RC", "var_error", "b")
RM_3_models
```

```
OR_3_models <- RMH_to_OR(RM_3_models)
OR_3_models

## Example 2c: RMH parameters from last 3 lines of Table 1 in Hillis (2012)
# using 10 diseased and 25 nondiseased cases
RM_3_models_Hillis <- data.frame(
  rbind(
    line6 = c(25, 25, 0.821, 0.821, 0.0132, 0.0132, 0.1, 0.1, 0.2, 0.6, 0.84566),
    line7 = c(25, 25, 1.831, 1.831, 0.0447, 0.0447, 0.1, 0.1, 0.2, 0.6, 0.71082),
    line8 = c(25, 25, 3.661, 3.611, 0.1201, 0.1201, 0.1, 0.1, 0.2, 0.6, 0.55140)
  )
)
colnames(RM_3_models_Hillis) <- c("n0", "n1", "delta1", "delta2", "var_R", "var_TR",
                                  "var_C", "var_TC", "var_RC", "var_error", "b")
RM_3_models_Hillis
OR_3_models_Hillis <- RMH_to_OR(RM_3_models_Hillis)
OR_3_models_Hillis
```

---

## `roc_curves` *ROC Performance Curves*

---

### Description

Calculation of ROC curve true positive rate (TPR) and false positive rate (FPR) pairs for values of a numeric rating of a true binary response.

### Usage

```
roc_curves(...)

## Default S3 method:
roc_curves(truth, rating, groups = list(), method = "empirical", ...)

## S3 method for class 'mrmc'
roc_curves(x, ...)

## S3 method for class 'stmc'
roc_curves(x, ...)

parameters(x, ...)

## S3 method for class 'roc_curve'
parameters(x, ...)

## S3 method for class 'roc_curves'
parameters(x, ...)
```

```
## S3 method for class 'mrmc'
parameters(x, ...)

## S3 method for class 'stmc'
parameters(x, ...)

## S3 method for class 'roc_curve'
points(
  x,
  metric = c("specificity", "sensitivity"),
  values = seq(0, 1, length = 101),
  ...
)

## S3 method for class 'roc_curves'
points(
  x,
  metric = c("specificity", "sensitivity"),
  values = seq(0, 1, length = 101),
  ...
)

## S3 method for class 'empirical_curve'
points(
  x,
  metric = c("specificity", "sensitivity"),
  values = NULL,
  which = c("curve", "curves", "observed"),
  ties = max,
  ...
)

## S3 method for class 'empirical_curves'
points(
  x,
  metric = c("specificity", "sensitivity"),
  values = NULL,
  which = c("curve", "curves", "observed"),
  ties = max,
  ...
)

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

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

```
## S3 method for class 'binormal_curves'
mean(x, method = c("points", "parameters"), ...)
```

### Arguments

| | |
|---|---|
| `...` | arguments passed from the `mean()` method to `points()`. |
| `truth` | vector of true binary statuses. |
| `rating` | vector of numeric ratings. |
| `groups` | list or data frame of grouping variables of the same lengths as `truth` and `rating`. |
| `method` | character string indicating the curve type as `"binormal"`, `"binormalLR"`, `"empirical"`, or `"trapezoidal"` or the averaging of binormal curves over `"points"` or `"parameters"`. |
| `x` | object returned by `mrmc` or `roc_curves` for which to compute points on or to average over the curves. |
| `metric` | reader performance metric to which the `values` correspond. |
| `values` | numeric vector of values at which to compute ROC curve points, or `NULL` for default empirical values as determined by `which`. |
| `which` | character string indicating whether to use curve-specific observed values and 0 and 1 (`"curve"`), the combination of these values over all curves (`"curves"`), or only the observed curve-specific values (`"observed"`). |
| `ties` | function determining empirical roc points returned in cases of ties. |

### Value

Function `roc_curves` returns an `roc_curve` class object of a single estimated ROC curve or an `roc_curves` class object of grouped ROC curves. Function `parameters` extracts the parameters that define the curves, `points` returns a data frame of points on individual curves, and `means` returns points on averaged curves (Chen and Samuelson, 2014).

### References

Chen W and Samuelson FW (2014). The average receiver operating characteristic curve in multi-reader multicase imaging studies. The British Journal of Radiology, 87(1040): 20140016.

### See Also

`plot`

### Examples

```
curves <- with(VanDyke,
  roc_curves(truth, rating, groups = list(Test = treatment, Reader = reader))
)
points(curves)
mean(curves)
```

`srmc` *Single-Reader Multi-Case ROC Analysis*

### Description

Estimation and comparison of ROC performance metrics for single-reader multi-case studies.

### Usage

```
srmc(response, test, case, data, cov = jackknife)
```

### Arguments

| | |
|---|---|
| `response` | response metric expressed in terms of a package-supplied performance `metric`. |
| `test` | variable of test identifiers. |
| `case` | variable of case identifiers. |
| `data` | data frame containing the `response`, `test`, and `case` variables. |
| `cov` | function, function call, or character string naming the `method` to use in calculating performance metric covariances. |

### Value

Returns an `mrmc` class object.

### See Also

`metrics`, `cov_methods`, `parameters`, `plot`, `roc_curves`, `summary`

### Examples

```
VanDyke1 <- subset(VanDyke, reader == "1")
est <- srmc(empirical_auc(truth, rating), treatment, case, data = VanDyke1)
plot(est)
summary(est)
```

---

## `stmc` *Single-Test (Single-Reader) Multi-Case ROC Analysis*

---

### Description

Estimation of ROC performance metrics for a single test of multiple cases.

### Usage

```
stmc(response, case, data, cov = jackknife)
```

### Arguments

| | |
|---|---|
| `response` | response metric expressed in terms of a package-supplied performance `metric`. |
| `case` | optional variable of case identifiers. |
| `data` | data frame containing the `response`, `test`, `reader`, and `case` variables. |
| `cov` | function, function call, or character string naming the `method` to use in calculating performance metric covariances. |

### Value

Returns a `stmc` class object with the following elements.

`metric` character name of the performance metric.

`est` estimated mean.

`se` estimated standard error.

`stmc_data` data frame of case-specific ratings.

### See Also

`metrics`, `cov_methods`, `parameters`, `plot`, `roc_curves`, `summary`

### Examples

```
VanDyke11 <- subset(VanDyke, treatment == "1" & reader == "1")
est <- stmc(empirical_auc(truth, rating), data = VanDyke11)
plot(est)
summary(est)
```

---

## `summary` *Summary Estimates and Statistical Tests*

---

### Description

Summary estimates and statistical tests from single and multi-reader multi-case analyses.

### Usage

```
## S3 method for class 'mrmc'
summary(object, conf.level = 0.95, ...)

## S3 method for class 'stmc'
summary(object, conf.level = 0.95, ...)
```

### Arguments

| | |
|---|---|
| `object` | object to summarize. |
| `conf.level` | confidence level for confidence intervals. |
| `...` | additional arguments affecting the summary. |

### Value

The respective method functions return `summary.mrmc` and `summary.stmc` class objects.

Object `summary.mrmc` is a list of the following elements.

**data_name** character name of the original data frame supplied for the analysis.

**cov_method** character name of the covariance method.

**design** experimental study design: 1 = factorial, 2 = cases nested within readers, 3 = cases nested within tests.

**vars** character names of the analysis factors and reader performance metric.

**conf.level** numeric confidence interval level.

**vcov_comps** data frame of estimated variances, covariances, and correlations.

**test_equality** data frame of the ANOVA global test of equality.

**test_diffs** data frame of pairwise test differences.

**test_means** data frame of estimated test means.

**reader_test_diffs** data frame of (fixed) reader-specific pairwise test differences.

**reader_means** data frame of (fixed) reader-specific test means.

Object `summary.stmc` is a numeric vector of the estimated reader performance mean, standard error, lower confidence limit, and upper confidence limit.

### See Also

`mrmc`, `srmc`, `stmc`

---

`VanDyke` *Multi-reader multi-case dataset*

---

## Description

Multi-reader multi-case dataset

## Usage

```
VanDyke
```

## Format

A data frame with 1140 rows and 7 variables:

**reader** reader identifier

**treatment** treatment identifier

**case** case identifier (factorial design)

**case2** case identifier (cases nested within readers)

**case3** case identifier (cases nested within treatments)

**truth** true case status (1 = positive, 0 = negative)

**rating** ordinal reader ratings of case status

## References

Van Dyke CW, White RD, Obuchowski NA, Geisinger MA, Lorig RJ, Meziane MA (1993). Cine MRI in the diagnosis of thoracic aortic dissection. 79th Radiological Society of North America Meetings, Chicago, IL.

# Index