

Package: rankrate (via r-universe)

October 11, 2024

Title Joint Statistical Models for Preference Learning with Rankings and Ratings

Version 1.2.0

Description Statistical tools for the Mallows-Binomial model, the first joint statistical model for preference learning for rankings and ratings. This project was supported by the National Science Foundation under Grant No. 2019901.

License GPL-3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

LazyData TRUE

Imports stats, gtools, isotone

Suggests knitr, rmarkdown, devtools, ggplot2, pander, reshape2

VignetteBuilder knitr

URL <https://pearce790.github.io/rankrate/>

Repository <https://pearce790.r-universe.dev>

RemoteUrl <https://github.com/pearce790/rankrate>

RemoteRef HEAD

RemoteSha 511e20d3987e28951a4a9d948fa886c4611bad76

Contents

AIBS	2
ASTAR	3
ci_mb	4
dmall	5
dmb	6
fit_mb	7
FV	8

getQ	8
Greedy	9
GreedyLocal	10
kendall	11
psi	11
rsmall	12
rmb	13
ToyData1	14
ToyData2	14
ToyData3	15
to_rankings	15

Index	17
--------------	-----------

AIBS	<i>Real peer review data set from the American Institute of Biological Sciences (AIBS)</i>
------	--

Description

This real data set includes 12 judges (reviewers) and 28 objects (proposals), and demonstrates the ability of the Mallows-Binomial model to combine ratings and rankings for the purpose of demarcating real grant proposals for a funding agency.

Usage

AIBS

Format

A list with three elements: (1) rankings, a 12 x 18 matrix of rankings with one row per judge; (2) ratings, a 12 x 18 matrix of ratings, with one row per judge and one column per object; and (3) M, a number indicating the maximum (worst) integer score.

Source

Originally published in: Gallo, Stephen A.. "Grant Peer Review Scoring Data with Criteria Scores" (2023). https://figshare.com/articles/dataset/Grant_Peer_Review_Scoring_Data_with_Criteria_Scores/12728087/1.

Originally analyzed in: Gallo, Stephen A., et al. "A new approach to peer review assessments: Score, then rank" (2023). *Research Integrity and Peer Review* 8:10 (10). <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-023-00131-7>.

ASTAR	<i>Calculate the exact MLE of a Mallows-Binomial distribution using an A* algorithm</i>
-------	---

Description

This function estimates the exact MLE of a Mallows-Binomial distribution using an A* tree search algorithm proposed in Pearce and Erosheva (2022). Algorithm may be very slow when number of objects exceeds 15, but is often still tractable for larger J when consensus is strong.

Usage

```
ASTAR(rankings, ratings, M, keep_nodes = FALSE)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
ratings	A matrix of ratings, one row per judge and one column per object.
M	Numeric specifying maximum (=worst quality) integer rating.
keep_nodes	Boolean specifying if function should retain the list of open nodes traversed during A* tree search. Defaults to FALSE.

Value

A list with elements π_0 , the estimated consensus ranking MLE, ρ , the estimated object quality parameter MLE, θ , the estimated scale parameter MLE, and `numnodes`, number of nodes traversed during algorithm and a measure of computational complexity. If `keep_nodes == TRUE`, then the list also contains `nodes`, a matrix of open nodes remaining at the end of search. If multiple MLEs are found, π_0 , ρ , and θ are returned a matrix elements, with one row per MLE.

Examples

```
data("ToyData1")
ASTAR(ToyData1$rankings, ToyData1$ratings, ToyData1$M, keep_nodes=TRUE)
```

ci_mb

Bootstrap Confidence Intervals for Mallows-Binomial parameters.

Description

This function calculates confidence intervals for parameters in a Mallows-Binomial model using the nonparametric bootstrap.

Usage

```
ci_mb(
  rankings,
  ratings,
  M,
  interval = 0.9,
  nsamples = 50,
  all = FALSE,
  method = "ASTAR"
)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
ratings	A matrix of ratings, one row per judge and one column per object.
M	Numeric specifying maximum (=worst quality) integer rating.
interval	A numeric entry between 0 and 1 specifying the confidence interval (e.g., .90 indicates a 90% confidence interval). Defaults to 0.90.
nsamples	A numeric entry indicating desired number of bootstrap samples to be used when calculating confidence intervals. Defaults to 50.
all	A boolean indicating if estimated parameters from all bootstrap samples should be returned. Defaults to FALSE.
method	A character string indicating which estimation method to use when estimating parameters. Allowable options are currently "ASTAR", "Greedy", "GreedyLocal", and "FV". Defaults to exact search, "ASTAR".

Value

A list with elements `ci`, a matrix of confidence intervals for Mallows-Binomial parameters, `ci_ranks`, a matrix of confidence intervals for object ranks, `bootstrap_pi0`, a matrix of bootstrap consensus rankings (returned only if `all==TRUE`), and `bootstrap_ptheta`, a matrix of bootstrap estimates of (p, θ) (returned only if `all==TRUE`).

Examples

```
data("ToyData1")
ci_mb(ToyData1$rankings, ToyData1$ratings, ToyData1$M, method="ASTAR", all=TRUE)
```

dmall

Calculate the density of rankings under a Mallows distribution

Description

This function calculates the density of observation(s) under a Mallows distribution.

Usage

```
dmall(rankings, pi0, theta, log = FALSE)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
pi0	A vector specifying the consensus (modal probability) ranking; should be used only for tie-breaking equal values in p.
theta	A numeric entry specifying the Mallows scale parameter.
log	A boolean indicating if the log likelihood should be returned.

Value

A numeric value indicating the (log) likelihood of rankings under a Mallows distribution.

Examples

```
rankings1 <- matrix(c(1,2,3,3,1,2), nrow=2, byrow=TRUE)
rankings2 <- matrix(c(1,2,3,4,2,3,NA,NA), nrow=2, byrow=TRUE)
attr(rankings2, "assignments") <- matrix(c(rep(TRUE, 4), FALSE, TRUE, TRUE, TRUE), nrow=2, byrow=TRUE)
dmall(rankings=c(1,2,3,NA), pi0=c(1,2,3,4), theta=2)
dmall(rankings=rankings1, pi0=c(1,2,3), theta=2)
dmall(rankings=rankings2, pi0=c(1,2,3,4), theta=3, log=TRUE)
```

dmb	<i>Calculate the density of rankings and ratings under a Mallows-Binomial distribution.</i>
-----	---

Description

This function calculates the density of observation(s) under a Mallows-Binomial distribution.

Usage

```
dmb(rankings, ratings, p, theta, M, pi0 = NULL, log = FALSE)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
ratings	A matrix of ratings, one row per judge and one column per object.
p	A vector specifying the underlying object qualities. All values between be between 0 and 1, inclusive.
theta	A numeric entry specifying the Mallows scale parameter.
M	Numeric specifying maximum (=worst quality) integer rating.
pi0	A vector specifying the consensus (modal probability) ranking; should be used only for tie-breaking equal values in p.
log	A boolean indicating if the log likelihood should be returned.

Value

A numeric value indicating the (log) likelihood of rankings and ratings under a Mallows distribution.

Examples

```
data(ToyData1)
dmb(rankings=ToyData1$rankings,ratings=ToyData1$ratings,p=c(.2,.5,.7),theta=1,M=ToyData1$M)
dmb(rankings=ToyData1$rankings,ratings=ToyData1$ratings,p=c(.25,.25,.7),theta=1,M=ToyData1$M,
pi0=c(1,2,3),log=TRUE)
```

fit_mb	<i>Calculate the exact or approximate MLE of a Mallows-Binomial distribution using various methods</i>
--------	--

Description

This function calculates the exact or approximate MLE of a Mallows-Binomial distribution using a user-specified method.

Usage

```
fit_mb(
  rankings,
  ratings,
  M,
  method = c("ASTAR", "Greedy", "GreedyLocal", "FV")
)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
ratings	A matrix of ratings, one row per judge and one column per object.
M	Numeric specifying maximum (=worst quality) integer rating.
method	A character string indicating which estimation method to use when estimating parameters. Allowable options are currently "ASTAR", "Greedy", "GreedyLocal", and "FV". Defaults to exact search, "ASTAR".

Value

A list with elements π_0 , the estimated consensus ranking MLE, p , the estimated object quality parameter MLE, θ , the estimated scale parameter MLE, and numnodes, number of nodes traversed during algorithm and a measure of computational complexity. If multiple MLEs are found, π_0 , p , and θ are returned a matrix elements, with one row per MLE.

Examples

```
data("ToyData1")
fit_mb(ToyData1$rankings, ToyData1$ratings, ToyData1$M, method="ASTAR")
fit_mb(ToyData1$rankings, ToyData1$ratings, ToyData1$M, method="Greedy")
fit_mb(ToyData1$rankings, ToyData1$ratings, ToyData1$M, method="GreedyLocal")
fit_mb(ToyData1$rankings, ToyData1$ratings, ToyData1$M, method="FV")
```

FV	<i>Estimate the MLE of a Mallows-Binomial distribution using the FV method</i>
----	--

Description

This function estimates the MLE of a Mallows-Binomial distribution using the FV method.

Usage

```
FV(rankings, ratings, M)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
ratings	A matrix of ratings, one row per judge and one column per object.
M	Numeric specifying maximum (=worst quality) integer rating.

Value

A list with elements π_0 , the estimated consensus ranking MLE, ρ , the estimated object quality parameter MLE, θ , the estimated scale parameter MLE, and `numnodes`, number of nodes traversed during algorithm and a measure of computational complexity. If multiple MLEs are found, π_0 , ρ , and θ are returned a matrix elements, with one row per MLE.

Examples

```
data("ToyData1")
FV(ToyData1$rankings, ToyData1$ratings, ToyData1$M)
```

getQ	<i>Calculate Q Matrix</i>
------	---------------------------

Description

This function calculates the Q matrix given a collection of (partial) rankings.

Usage

```
getQ(rankings, I, J)
```


Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
I	A numeric entry indicating the total number of judges providing rankings and ratings.
J	A numeric entry or vector of positive integers indicating total number of objects.

Value

A matrix with dimension $J \times J$.

Examples

```
rankings <- matrix(c(1,2,3,4,2,1,NA,NA),byrow=TRUE,nrow=2)
getQ(rankings=rankings,I=2,J=4)
attr(rankings,"assignments") <- matrix(c(rep(TRUE,7),FALSE),byrow=TRUE,nrow=2,ncol=4)
getQ(rankings=rankings,I=2,J=4)
```

Greedy	<i>Estimate the MLE of a Mallows-Binomial distribution using the Greedy method</i>
--------	--

Description

This function estimates the MLE of a Mallows-Binomial distribution using the Greedy method.

Usage

```
Greedy(rankings, ratings, M)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
ratings	A matrix of ratings, one row per judge and one column per object.
M	Numeric specifying maximum (=worst quality) integer rating.

Value

A list with elements π_0 , the estimated consensus ranking MLE, p , the estimated object quality parameter MLE, θ , the estimated scale parameter MLE, and numnodes , number of nodes traversed during algorithm and a measure of computational complexity. If multiple MLEs are found, π_0 , p , and θ are returned a matrix elements, with one row per MLE.

Examples

```
data("ToyData1")
Greedy(ToyData1$rankings, ToyData1$ratings, ToyData1$M)
```

GreedyLocal	<i>Estimate the MLE of a Mallows-Binomial distribution using the "Greedy Local" method</i>
-------------	--

Description

This function estimates the MLE of a Mallows-Binomial distribution using the GreedyLocal method, which is identical to the Greedy method but includes an automatic and targeted post-hoc local search.

Usage

```
GreedyLocal(rankings, ratings, M)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
ratings	A matrix of ratings, one row per judge and one column per object.
M	Numeric specifying maximum (=worst quality) integer rating.

Value

A list with elements π_0 , the estimated consensus ranking MLE, p , the estimated object quality parameter MLE, θ , the estimated scale parameter MLE, and `numnodes`, number of nodes traversed during algorithm and a measure of computational complexity. If multiple MLEs are found, π_0 , p , and θ are returned a matrix elements, with one row per MLE.

Examples

```
data("ToyData1")
GreedyLocal(ToyData1$rankings, ToyData1$ratings, ToyData1$M)
```

kendall	<i>Calculate the Kendall's tau between rankings</i>
---------	---

Description

This function calculates Kendall's tau distance between ranking(s) and a central permutation, π_0

Usage

```
kendall(rankings, pi0)
```

Arguments

rankings	A matrix of rankings, potentially with attribute "assignments" to signify separate reviewer assignments. One ranking per row.
pi0	A vector specifying the consensus (modal probability) ranking.

Value

A vector of the Kendall's tau distance between each ranking in rankings and π_0 .

Examples

```
ranking1 <- c(2,1,3)
ranking2 <- matrix(c(2,1,3,1,2,3),byrow=TRUE,nrow=2)
ranking3 <- matrix(c(1,2,3,4,2,4,NA,NA),byrow=TRUE,nrow=2)
attr(ranking3,"assignments") <- matrix(c(TRUE,TRUE,TRUE,TRUE,
  FALSE,TRUE,FALSE,TRUE),byrow=TRUE,nrow=2)
kendall(ranking1,c(1,2,3))
kendall(ranking2,c(1,2,3))
kendall(ranking3,c(1,2,3,4))
```

psi	<i>Normalizing constant function of a Mallows distribution, psi</i>
-----	---

Description

This function calculates the normalizing constant of a Mallows distribution under the Kendall distance

Usage

```
psi(theta, J, R, log = FALSE)
```

Arguments

theta	A numeric entry specifying the Mallows scale parameter.
J	A numeric entry or vector of positive integers indicating total number of objects each judge has access to. If $\text{length}(J) > 1$, R must be of same length or a single value.
R	A numeric entry or vector of positive integers indicating the length of the ranking provided by each judge. If $\text{length}(R) > 1$, J must be of same length or a single value.
log	A boolean indicating if $\log(\text{psi})$ should be returned.

Value

A numeric value or vector representing normalizing constant of a Mallows distribution.

Examples

```
psi(theta=1,J=10,R=8)
psi(theta=2,J=c(4,4,3),R=c(2,2,1),log=TRUE)
```

 rmall

Random Mallows generation.

Description

This function randomly generates rankings from a Mallows distribution.

Usage

```
rmall(I, pi0, theta, R = NULL)
```

Arguments

I	A numeric entry indicating the number of observations to be drawn, i.e., the number of judges providing rankings and ratings.
pi0	A vector specifying the consensus (modal probability) ranking; should be used only for tie-breaking equal values in p.
theta	A numeric entry specifying the Mallows scale parameter.
R	A numeric entry specifying the length of the rankings to be drawn. When $R \leq \text{length}(p)$, partial rankings are drawn by definition.

Value

A matrix of rankings (orderings) with one row per judge.

Examples

```
rmb(I=5,pi0=1:5,theta=1,R=3)
rmb(I=5,pi0=1:3,theta=.5,R=c(1,1,1,1,3))
rmb(I=5,pi0=1:3,theta=.5)
```

rmb	<i>Random Mallows-Binomial generation</i>
-----	---

Description

This function randomly generates rankings and ratings from a Mallows-Binomial distribution.

Usage

```
rmb(I, p, theta, M, pi0 = NULL, R = NULL)
```

Arguments

I	A numeric entry indicating the number of observations to be drawn, i.e., the number of judges providing rankings and ratings.
p	A vector specifying the underlying object qualities. All values between be between 0 and 1, inclusive.
theta	A numeric entry specifying the Mallows scale parameter.
M	A numeric entry specifying the maximum integer rating.
pi0	A vector specifying the consensus (modal probability) ranking; should be used only for tie-breaking equal values in p.
R	A numeric entry specifying the length of the rankings to be drawn. When $R \leq \text{length}(p)$, partial rankings are drawn by definition.

Value

A list containing elements ratings, a matrix of integer ratings with one row per judge and one column per object, rankings, and matrix of rankings (orderings) with one row per judge, and M, the inputted maximum integer rating.

Examples

```
rmb(I=5,p=c(.1,.3,.4,.7,.9),theta=1,M=10)
rmb(I=10,p=c(.1,.3,.3,.7,.9),pi0=c(1,3,2,4,5),theta=5,M=40,R=3)
```

ToyData1	<i>Toy data set of rankings and ratings demonstrating tie-breaking</i>
----------	--

Description

This toy data set includes 16 judges and 3 objects, and demonstrates the ability of the Mallows-Binomial model to break ties in ratings via rankings.

Usage

ToyData1

Format

list with three elements: (1) rankings, a 16 x 3 matrix of rankings with one row per judge; (2) ratings, a 16 x 3 matrix of ratings, with one row per judge and one column per object; and (3) M, a number indicating the maximum (worst) integer score.

Source

Originally analyzed in: Gallo, Stephen A., et al. "A new approach to peer review assessments: Score, then rank" (2023). Research Integrity and Peer Review 8:10 (10). <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-023-00131-7>.

ToyData2	<i>Toy data set of rankings and ratings demonstrating decision-making with partial rankings</i>
----------	---

Description

This toy data set includes 16 judges and 8 objects, and demonstrates the ability of the Mallows-Binomial model to estimate overall object orderings under partial rankings.

Usage

ToyData2

Format

list with three elements: (1) rankings, a 16 x 8 matrix of rankings with one row per judge; (2) ratings, a 16 x 8 matrix of ratings, with one row per judge and one column per object; and (3) M, a number indicating the maximum (worst) integer score.

Source

Originally analyzed in: Gallo, Stephen A., et al. "A new approach to peer review assessments: Score, then rank" (2023). Research Integrity and Peer Review 8:10 (10). <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-023-00131-7>.

ToyData3	<i>Toy data set of rankings and ratings when judges express internally inconsistent preferences</i>
----------	---

Description

This toy data set includes 16 judges and 3 objects, and demonstrates the ability of the Mallows-Binomial model to estimate overall object orderings even when judges provide sets of rankings and ratings which are internally inconsistent.

Usage

```
ToyData3
```

Format

list with three elements: (1) rankings, a 16 x 3 matrix of rankings with one row per judge; (2) ratings, a 16 x 3 matrix of ratings, with one row per judge and one column per object; and (3) M, a number indicating the maximum (worst) integer score.

Source

Originally analyzed in: Gallo, Stephen A., et al. "A new approach to peer review assessments: Score, then rank" (2023). *Research Integrity and Peer Review* 8:10 (10). <https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-023-00131-7>.

to_rankings	<i>Convert ranks into rankings (orderings)</i>
-------------	--

Description

This function converts a matrix of ranks into a matrix of rankings (i.e., orderings), potentially including reviewer assignments as an attribute of the ranking matrix. Additionally, it can be used to add an assignments matrix to an existing matrix of rankings.

Usage

```
to_rankings(ranks, assignments = NULL, rankings = NULL)
```

Arguments

ranks	A matrix or vector of ranks, such that the (i,j) entry includes the rank given by judge i to proposal j. NA is used to indicate that no rank was assigned to a proposal, which may occur for two reasons: (1) If the assignments matrix is not specified or the (i,j) entry of assignments is TRUE, then an NA indicates that a proposal was considered worse than all ranked proposals. (2) If the (i,j) entry of assignments is FALSE, then NA indicates that a proposal was not considered by the judge and no information can be gleaned from the missing rank.
assignments	A matrix of booleans, such that the (i,j) entry is TRUE if judge i was assigned to review proposal j, and FALSE otherwise. If assignments is NULL, we assume all judges considered all proposals.
rankings	A matrix or vector of rankings. If a matrix, there should be one ranking per row.

Value

A matrix of rankings, with one row per ranking. If assignments argument is specified, then the rankings matrix will have the attribute "assignments".

Examples

```
ranks <- matrix(data=c(4,2,3,1,NA,1,2,3,NA,NA,1,NA),byrow=TRUE,nrow=3)
assignments=matrix(TRUE,byrow=TRUE,nrow=3,ncol=4)
to_rankings(ranks=ranks)
to_rankings(ranks=ranks,assignments=assignments)
to_rankings(assignments=matrix(TRUE,nrow=1,ncol=3),rankings=c(3,2,1))
```


Index

* datasets

- AIBS, [2](#)
- ToyData1, [14](#)
- ToyData2, [14](#)
- ToyData3, [15](#)

- AIBS, [2](#)
- ASTAR, [3](#)

- ci_mb, [4](#)

- dmail, [5](#)
- dmb, [6](#)

- fit_mb, [7](#)
- FV, [8](#)

- getQ, [8](#)
- Greedy, [9](#)
- GreedyLocal, [10](#)

- kendall, [11](#)

- psi, [11](#)

- rmail, [12](#)
- rmb, [13](#)

- to_rankings, [15](#)
- ToyData1, [14](#)
- ToyData2, [14](#)
- ToyData3, [15](#)