# Package 'PAMA'

**Title** Rank Aggregation with Partition Mallows Model

**Version** 1.2.0

**Description**

Rank aggregation aims to achieve a better ranking list given multiple observations. 'PAMA' implements Partition-Mallows model for rank aggregation where the rankers' quality are different. Both Bayesian inference and Maximum likelihood estimation (MLE) are provided. It can handle partial list as well. When covariates information is available, this package can make inference by incorporating the covariate information. More information can be found in the paper ``Integrated Partition-Mallows Model and Its Inference for Rank Aggregation''.

**Depends** R (>= 3.1.0)

**License** GPL (>= 2)

**Encoding** UTF-8

**LazyData** false

**RoxygenNote** 7.1.1

**Suggests** knitr, rmarkdown

**LinkingTo** Rcpp

**Imports** ExtMallows, mc2d, PerMallows, rankdist, Rcpp, stats

**NeedsCompilation** yes

**Author** Wanchuang Zhu [cre, aut]

**Maintainer** Wanchuang Zhu <andy.chou.sub@gmail.com>

**Repository** CRAN

**Date/Publication** 2021-05-06 13:30:02 UTC

## R topics documented:

---

NBANFL *Dataset of NBA and NFL.*

---

### Description

Dataset of NBA and NFL.

### Usage

```
NBANFL()
```

### Value

List. It contains four data sets used in literature.

1. NBA: full ranking lists of NBA power ranking.

2. NBAPL: The partial lists of NBA power ranking

3. NFLdata: all the ranking lists of NFL data

4. NFLcov: covariates of NFL players

### References

1. Ke Deng, Simeng Han, Kate J. Li & Jun S. Liu (2014) Bayesian Aggregation of Order-Based Rank Data, Journal of the American Statistical Association, 109:507, 1023-1039

2. Li, X., Yi, D., and Liu, J. S. (2021). Bayesian analysis of rank data with covariates andheterogeneous rankers.Statistical Science, (In press)

### Examples

```
NBANFL()
```

---

PAMA.B *This function implements Bayesian inference of PAMA model.*

---

### Description

This function implements Bayesian inference of PAMA model.

### Usage

```
PAMA.B(datfile, nRe, iter = 1000, init = "EMM")
```

## Arguments

| | |
|---|---|
| `datfile` | A matrix or dataframe. This is the data where our algorithm will work on. Each colomn denotes a ranker's ranking. The data should be in entity-based format. |
| `nRe` | A number. Number of relevant entities |
| `iter` | A number. Numner of iterations of MCMC |
| `init` | A string. This indicates which method is used to initiate the starting point of the aggregated ranking list. "mean" uses the sample mean. "EMM" uses the method from R package 'ExtMallows'. |

## Value

List. It contains Bayesian posterior samples of all the parameters and log-likelihood.

1. I.mat: posterior samples of I
2. phi.mat: posterior samples of phi
3. smlgamma.mat: posterior samples of gamma
4. l.mat: posterior samples of log-likelihood

## Author(s)

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng

## References

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng (2021) Partition-Mallows Model and Its Inference for Rank Aggregation. Journal of the American Statistical Association

## Examples

```
dat=t(PerMallows::rmm(10,1:20,0.5))
PAMA.B(dat,10,iter=10)
PAMA.B(dat,10,iter=1000)
```

---

| PAMA.Cov | *This function implements Bayesian inference of PAMA model with covariates.* |
|---|---|

---

## Description

This function implements Bayesian inference of PAMA model with covariates.

## Usage

```
PAMA.Cov(datfile, Covdatfile, nRe, iter = 1000, init = "EMM")
```

## Arguments

| | |
|---|---|
| `datfile` | A matrix or dataframe. This is the data where our algorithm will work on. Each colomn denotes a ranker's ranking. The data should be in entity-based format. |
| `Covdatfile` | A matrix or dataframe. Each column denotes a covariate. |
| `nRe` | A number. Number of relevant entities |
| `iter` | A number. Numner of iterations of MCMC. Defaulted as 1000. |
| `init` | A string. This indicates which method is used to initiate the starting point of the aggregated ranking list. "mean" uses the sample mean. "EMM" uses the method from R package 'ExtMallows'. |

## Details

The covariates are incoporated in the PAMA framework as indicators of group membership. That is, covariates are associated to group members via a logistic regression.

## Value

List. It contains Bayesian posterior samples of all the parameters and log-likelihood.

1. I.mat: posterior samples of I

2. phi.mat: posterior samples of phi

3. smlgamma.mat: posterior samples of gamma

4. l.mat: posterior samples of log-likelihood.

5. theta.mat: posterior samples of coefficients of covariates.

## Author(s)

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng

## References

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng (2021) Partition-Mallows Model and Its Inference for Rank Aggregation. Journal of the American Statistical Association

## Examples

```
a=NBANFL()
PAMA.Cov(t(a$NFLdata),a$NFLcov,nRe=10,iter=10)
## Not run: PAMA.Cov(t(a$NFLdata),a$NFLcov,nRe=10,iter=1000)
```

| PAMA.F | *This function implements Maximum Likelihood estimation of PAMA model.* |
|---|---|

## Description

This function implements Maximum Likelihood estimation of PAMA model.

## Usage

```
PAMA.F(datfile, nRe, threshold, iter = 1000, init = "EMM")
```

## Arguments

| | |
|---|---|
| datfile | A matrix or dataframe. This is the data where our algorithm will work on. Each row denotes a ranker's ranking. The data should be in entity-based format. |
| nRe | A number. Number of relevant entities. |
| threshold | A number (positive). The stopping threshold in determining convergence of MLE. If the two consecutive iterations of log-likelihood is smaller than the threshold, then the convergence is satisfied. |
| iter | A number. Numner of iterations of MCMC. |
| init | A string. This indicates which method is used to initiate the starting point of the aggregated ranking list. "mean" uses the sample mean. "EMM" uses the method from R package 'ExtMallows'. |

## Value

List. It contains MLE of all the parameters and log-likelihood. We use an iterative procedure to find the MLEs, so there are several values for each parameter until convergence.

1. I.mat: samples of I
2. phi.mat: samples of phi
3. smlgamma.mat: samples of gamma
4. l.mat: samples of log-likelihood

## Author(s)

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng

## References

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng (2021) Partition-Mallows Model and Its Inference for Rank Aggregation. Journal of the American Statistical Association

## Examples

```
a=NBANFL()
PAMA.F(a$NBA,nRe=10,threshold=0.1,iter=10)
```

---

| PAMA.PL | *This function implements Bayesian inference of PAMA model with partial lists.* |

---

### Description

This function implements Bayesian inference of PAMA model with partial lists.

### Usage

```
PAMA.PL(datfile, PLdatfile, nRe, iter = 1000, init = "EMM")
```

### Arguments

| | |
|---|---|
| datfile | A matrix or dataframe. This is the data where our algorithm will work on. Each colomn denotes a ranker's ranking. The data should be in entity-based format. |
| PLdatfile | A matrix or dataframe. It contains all the partial lists. Each colomn denotes a partial list. |
| nRe | A number. Number of relevant entities. |
| iter | A number. Numner of iterations of MCMC. Defaulted as 1000. |
| init | A string. This indicates which method is used to initiate the starting point of the aggregated ranking list. "mean" uses the sample mean. "EMM" uses the method from R package 'ExtMallows'. |

### Details

The partial lists are handle by Data Augmentation strategy.

### Value

List. It contains Bayesian posterior samples of all the parameters and log-likelihood.

1. I.mat: posterior samples of I
2. phi.mat: posterior samples of phi
3. smlgamma.mat: posterior samples of gamma
4. l.mat: posterior samples of log-likelihood.

### Author(s)

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng

### References

Wanchuang Zhu, Yingkai Jiang, Jun S. Liu, Ke Deng (2021) Partition-Mallows Model and Its Inference for Rank Aggregation. Journal of the American Statistical Association

## Examples

```
a=NBANFL()
PAMA.PL(a$NBA,a$NBAPL,nRe=10,iter=1)
PAMA.PL(a$NBA,a$NBAPL,nRe=10,iter=100)
```

---

| PAMAlike | *this function returns the log-likelihood of an observed ranking list given the parameters.* |
|---|---|

---

## Description

this function returns the log-likelihood of an observed ranking list given the parameters.

## Usage

```
PAMAlike(bsrkr, I, phi, smlgamma)
```

## Arguments

| | |
|---|---|
| bsrkr | It is a observed ranking list. |
| I | It is the true classification of entities. 0 denotes the corresponding entity is a background entity. The positive integer denotes the relative rankings of a relevant entity. |
| phi | It is a positive number. It is the common disperse parameter in Mallows model for all the rankers |
| smlgamma | A positive number. It is the quality parameter of the ranker. It is used to distinguish relative and background entities |

## Value

The lon-likelihood of barkr given I, phi and smlgamma

## Examples

```
dat=t(PerMallows::rmm(10,1:20,0.5))
I=c(1:10,rep(0,10))
like=PAMAlike(bsrkr=dat[,1],I=I,phi=0.2,smlgamma=1)
```

# Index