Package ‘socceR’

July 3, 2019

Type Package
Title Evaluating Sport Tournament Predictions
Version 0.1.1
Date 2019-07-01
Maintainer Claus Thorn Ekstrøm <ekstrom@sund.ku.dk>
Description Functions for evaluating tournament predictions, simulating results from individual soccer matches and tournaments. See <http://sandsynligvis.dk/2018/08/03/world-cup-prediction-winners/> for more information.
License GPL (>= 2)
Depends R (>= 3.1.0)
Imports Rcpp (>= 1.0.0)
LinkingTo Rcpp
LazyData true
RoxygenNote 6.1.1
Encoding UTF-8
URL https://github.com/ekstroem/socceR
BugReports https://github.com/ekstroem/socceR/issues
NeedsCompilation yes
Author Claus Thorn Ekstrøm [aut, cre]
Repository CRAN
Date/Publication 2019-07-03 11:50:03 UTC

R topics documented:
collapse_prediction .................................................... 2
fifa2018 ................................................................. 3
fifa2018result .......................................................... 3
logloss ................................................................. 4
optimize_weights ...................................................... 5
socceR ................................................................. 6
ttrs ................................................................. 6
collapse_prediction

Description

Creates a matrix to collapse the rows of a tournament prediction matrix

Usage

collapse_prediction(ranks = c(1L, 2L, 3L, 4L, 8L, 16L, 32L))

Arguments

ranks An integer vector of R ordered elements giving the cut off of the ranks to create

Details

Returns a vector of numeric values. Elements in the input factor that cannot be converted to numeric will produce NA.

Value

Returns a numeric matrix with R rows and T columns that can be multiplied on a square prediction matrix to obtain the collapsed predictions

Author(s)

Claus Ekstrom <ekstrom@sund.ku.dk>

Examples

m2 <- matrix(c(.5, .5, 0, 0, .5, .5, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1), 4)
# Collapse into ranks 1, 2, and 3+4
collapse <- collapse_prediction(c(1L, 2L, 4L))

collapsed_prediction <- collapse %*% m2
collapsed_prediction
**fifa2018**  

*FIFA 2018 prediction matrices*

**Description**

A list containing five predictions for the FIFA 2018 World Cup.

**Usage**

fifa2018

**Format**

A list with 5 predictions (each a 7 by 32 matrix) containing the predictions probabilities of 1st, 2nd, 3rd, 4th, 5th-8th, 9th-12th, and 17th-32nd place.

- **flat**  A prediction with equal probability of winning for all teams
- **ekstrom1**  Ekstrom’s prediction (based on the Skellam distribution)
- **ekstrom2**  Ekstrom’s prediction (based on the ELO rankings)
- **GLSE1**  Prediction of Groll et all
- **GLSE2**  Updated prediction of Groll et all

**Source**

http://sandsynligvis.dk/2018/08/03/world-cup-prediction-winners/

---

**fifa2018result**  

*FIFA 2018 end results*

**Description**

A named vector sorted in the ranking of the teams in the FIFA 2018 World Cup. The value correspond to the corresponding columns in the prediction matrices of fifa2018

**Usage**

fifa2018result

**Format**

A vector of the final rankings

**Source**

http://sandsynligvis.dk/2018/08/03/world-cup-prediction-winners/
logloss

Computes the log loss score for a tournament prediction

Description

Compute the (weighted) rank probability score for a tournament.

Usage

logloss(m, outcome, rankweights = 1)

Arguments

m
An R*T prediction matrix where the R rows represent the ordered ranks and each column is a team. Each column should sum to 1, and each row should sum to the number of teams that can attain a given rank.

outcome
A vector of length T containing the integers 1 to R giving the ranks that were obtained by each of the T teams

rankweights
A vector of length R of rank weights or a single weight which will be reused for all ranks (defaults to 1)

Value

The rank probability score. Zero means a perfect score.

Author(s)

Claus Ekstrom <ekstrom@sund.ku.dk>

Examples

m1 <- matrix(c(1, 0, 0, 0, 1, 0, 0, 0, 0, .5, .5, 0, 0, .5, .5, 4)
m1 # Prediction where certain on the top ranks
logloss(m1, c(1, 2, 3, 4))
**optimize_weights**

*Optimize weights from list of prediction matrices*

**Description**

Computes the optimal weights to obtain the minimal loss function from a list of prediction matrices.

**Usage**

```r
optimize_weights(predictionlist, outcome, FUN = trps)
```

**Arguments**

- `predictionlist`: A list of R x T prediction matrices where each column sum to 1 and each row sums to
- `outcome`: An integer vector listing the
- `FUN`: The function used for optimizing the predictions. The default is top use rps for the rank probability score. Another option is logloss for log loss.

**Value**

Returns a numeric vector containing an optimal vector of weights that sum to 1 and that minimizes the loss function.

**Author(s)**

Claus Ekstrom <ekstrom@sund.ku.dk>

**Examples**

```r
m1 <- matrix(c(1, 0, 0, 0, 1, 0, 0, 0, .5, .5, 0, .5, .5, .5), 4)
m1 # Prediction where certain on the top ranks
m2 <- matrix(c(.5, .5, 0, .5, .5, 0, 0, 0, 1, 0, 0, 0, 1), 4)
m2 # Prediction where the groups are okay
m3 <- matrix(c(.5, .5, 0, .5, .5, 0, 0, 0, .5, .5, 0, .5, .5), 4)
m3 # Prediction where no clue about anything
m4 <- matrix(rep(1/4, 16), 4)

optimize_weights(list(m1, m2, m3, m4), 1:4)
```
socceR

Evaluating sport tournament predictions

Description
Functions for evaluating sport tournament predictions, the tournament rank probability score, and working with models for prediction sport matches.

Author(s)
Claus Ekstrom <ekstrom@sund.ku.dk>

trps

Computes the rank probability score for a tournament

Description
Compute the (weighted) rank probability score for a tournament.

Usage
trps(m, outcome, rankweights = 1L)

Arguments
- m: An R*T prediction matrix where the R rows represent the ordered ranks and each column is a team. Each column should sum to 1, and each row should sum to the number of teams that can attain a given rank.
- outcome: A vector of length T containing the integers 1 to R giving the ranks that were obtained by each of the T teams
- rankweights: A vector of length R of rank weights or a single weight which will be reused for all ranks (defaults to 1)

Value
The rank probability score. Zero means a perfect score.

Author(s)
Claus Ekstrom <ekstrom@sund.ku.dk>

Examples
m1 <- matrix(c(1L, PL, PL, PL, PL, 1L, PL, PL, PL, PL, NUL, NUL, PL, PL, NUL, NU), 4)
m1 # Prediction where certain on the top ranks
trps(m1, c(1, 2, 3, 4))
Index

*Topic **datasets**
  fifa2018, 3
  fifa2018result, 3
*Topic **manip**
  collapse_prediction, 2
  optimize_weights, 5

  collapse_prediction, 2
  fifa2018, 3
  fifa2018result, 3
  logloss, 4
  optimize_weights, 5
  socceR, 6
  socceR-package (socceR), 6
  trps, 6