Package ‘ircor’

August 21, 2017

Title Correlation Coefficients for Information Retrieval

Version 1.0


Depends R (>= 3.2.0)

License MIT + file LICENSE

BugReports https://github.com/julian-urbano/ircor/issues

URL https://github.com/julian-urbano/ircor/

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Suggests testthat

NeedsCompilation no

Author Julián Urbano [aut, cre],
    Mónica Marrero [aut]

Maintainer Julián Urbano <urbano.julian@gmail.com>

Repository CRAN

Date/Publication 2017-08-21 09:07:29 UTC

R topics documented:

  tau ................................................................. 2
  tauAP .............................................................. 3

Index 5
Description

tau is the rank correlation coefficient by Kendall, where neither vector can contain tied items. tau_a and tau_b are the versions developed to cope with ties under the scenarios of accuracy and agreement, respectively. See the references for details.

Usage

tau(x, y)
tau_a(x, y)
tau_b(x, y)

Arguments

x  a numeric vector. In tau_a this is the vector of true scores.
y  a numeric vector of the same length as x. In tau_a this is the vector of estimated scores.

Value

The correlation coefficient.

References


See Also

tauAP for AP correlation coefficients.

Examples

# No ties
x <- c(0.67, 0.45, 0.29, 0.12, 0.57, 0.24, 0.94, 0.75, 0.08, 0.54)
y <- c(0.48, 0.68, 0.32, 0.09, 0.06, 0.61, 0.87, 0.22, 0.44, 0.84)
tau(x, y)
tau_a(x, y)  # same as tau
tau_b(x, y)  # same as tau

# Ties in y
y <- round(y, 1)
tau_a(x, y)
tau_b(x, y)
# Ties in \( x \) too
\[
x \leftarrow \text{round}(x, 1)
\]
\[
tau_b(x, y)
\]

---

**tauAP**

**AP Rank Correlation Coefficients**

**Description**

\( \text{tauAP} \) is the AP rank correlation coefficient by Yilmaz et al., where neither vector can contain tied items. \( \text{tauAP\_a} \) and \( \text{tauAP\_b} \) are the versions developed by Urbano and Marrero to cope with ties under the scenarios of accuracy and agreement, respectively. See the references for details.

**Usage**

\[
\text{tauAP}(x, y, \text{decreasing} = \text{TRUE})
\]

\[
\text{tauAP\_a}(x, y, \text{decreasing} = \text{TRUE})
\]

\[
\text{tauAP\_b}(x, y, \text{decreasing} = \text{TRUE})
\]

**Arguments**

- \( x \): a numeric vector. In \( \text{tauAP\_a} \) this is the vector of true scores.
- \( y \): a numeric vector of the same length as \( x \). In \( \text{tauAP\_a} \) this is the vector of estimated scores.
- \( \text{decreasing} \): logical. Should the sort order be increasing or decreasing (default)?

**Details**

Note that the sorting order is decreasing by default, as should be for instance if the scores represent the effectiveness of systems. When the sorting order is ascending, as is for instance when the vectors represent ranks, the parameter \text{decreasing} must be set to \text{FALSE}.

**Value**

The correlation coefficient.

**References**


J. Urbano and M. Marrero (2017). The Treatment of Ties in AP Correlation. ACM ICTIR.

**See Also**

- \text{tau} for Kendall correlation coefficients.
Examples

# No ties
x <- c(0.67, 0.45, 0.29, 0.12, 0.57, 0.24, 0.94, 0.75, 0.08, 0.54)
y <- c(0.48, 0.68, 0.32, 0.09, 0.06, 0.61, 0.87, 0.22, 0.44, 0.84)
tauAP(x, y)
tauAP_a(x, y) # same as tauAP

# Ties in y
y <- round(y, 1)
tauAP_a(x, y)
tauAP_b(x, y)

# Ties in x too
x <- round(x, 1)
tauAP_b(x, y)

# Set decreasing to FALSE when x and y already represent ranks
x <- rank(-x)
y <- rank(-y)
tauAP_b(x, y, FALSE) # same as above
Index

tau, 2, 3
tau_a (tau), 2
tau_b (tau), 2
tauAP, 2, 3
tauAP_a (tauAP), 3
tauAP_b (tauAP), 3