Package ‘bosfr’

January 29, 2025

Title Computes Exact Bounds of Spearman's Footrule with Missing Data

Version 0.1.0

Author Yijin Zeng [aut, cre, cph]

Maintainer Yijin Zeng <yijinzeng98@gmail.com>

Description Computes exact bounds of Spearman's footrule in the presence of missing data, and performs independence test based on the bounds with controlled Type I error regardless of the values of missing data. Suitable only for distinct, univariate data where no ties is allowed.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.2

Depends R (>= 3.2.1)

Imports gtools, stats

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Repository CRAN

Date/Publication 2025-01-29 19:00:02 UTC

Contents

boundsKendall *Bounds of Kendall's tau in the Presence of Missing Data*

Description

Computes bounds of Kendall's tau in the presence of missing data. Suitable only for univariate distinct data where no ties is allowed.

Usage

```
boundsKendall(X, Y)
```

Arguments

`X, Y`	Numeric vectors of data values with potential missing data. No ties in the data is allowed. Inf and -Inf values will be omitted.

Details

`boundsKendall()` computes bounds of Kendall's tau for partially observed univariate, distinct data. The bounds are computed by first calculating the bounds of Spearman's footrule (*Zeng et al., 2025*), and then applying the combinatorial inequality between Kendall's tau and Spearman's footrule (*Kendall, 1948*). See *Zeng et al., 2025* for more details.

Let $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$ be two vectors of univariate, distinct data. Kendall's tau is defined as the number of discordant pairs between X and Y:

$$\tau(X, Y) = \sum_{i<j} \{I(x_i < x_j)I(y_i > y_j) + I(x_i > x_j)I(y_i < y_j)\}.$$

Scaled Kendall's tau $\tau_{Scale}(X, Y) \in [0, 1]$ is defined as (*Kendall, 1948*):

$$\tau_{Scale}(X, Y) = 1 - 4\tau(X, Y)/(n(n-1)).$$

Value

`bounds`	bounds of Kendall's tau.
`bounds.scaled`	bounds of scaled Kendall's tau.

References

- Zeng Y., Adams N.M., Bodenham D.A. Exact Bounds of Spearman's footrule in the Presence of Missing Data with Applications to Independence Testing. arXiv preprint arXiv:2501.11696. 2025 Jan 20.
- Kendall, M.G. (1948) Rank Correlation Methods. Charles Griffin, London.
- Diaconis, P. and Graham, R.L., 1977. Spearman's footrule as a measure of disarray. Journal of the Royal Statistical Society Series B: Statistical Methodology, 39(2), pp.262-268.

Examples

```
### compute bounds of Kendall's tau between incomplete ranked lists
X <- c(1, 2, NA, 4, 3)
Y <- c(3, NA, 4, 2, 1)
boundsKendall(X, Y)

### compute bounds of Kendall's tau between incomplete vectors of distinct data
X <- c(1.3, 2.6, NA, 4.2, 3.5)
Y <- c(5.5, NA, 6.5, 2.6, 1.1)
boundsKendall(X, Y)
```

---

boundsSFR *Exact bounds of Spearman's footrule in the Presence of Missing Data*

Description

Computes exact bounds of Spearman's footrule in the presence of missing data, and performs independence test based on the bounds with controlled Type I error regardless of the values of missing data. Suitable only for univariate distinct data where no ties is allowed.

Usage

```
boundsSFR(X, Y, pval = TRUE)
```

Arguments

`X`	Numeric vector of data values with potential missing data. No ties in the data is allowed. Inf and -Inf values will be omitted.
`Y`	Numeric vector of data values with potential missing data. No ties in the data is allowed. Inf and -Inf values will be omitted.
`pval`	Boolean for whether to compute the bounds of p-value or not.

Details

`boundsSFR()` computes exact bounds of Spearman's footrule for partially observed univariate, distinct data using the results and algorithms following *Zeng et al., 2025*.

Let $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$ be two vectors of univariate, distinct data, and denote the rank of x_i in X as $R(x_i, X)$, the rank of y_i in Y as $R(y_i, Y)$. Spearman's footrule is defined as the absolute distance between the ranked values of X and Y:

$$D(X,Y) = \sum_{i=1}^{n} |R(x_i, X) - R(y_i, Y)|.$$

Scaled Spearman's footrule is defined as:

$$D_{Scale}(X,Y) = 1 - 3D(X,Y)/(n^2 - 1).$$

When n is odd, $D_{Scale}(X, Y) \in [-0.5, 1]$, but when n is even, $D_{Scale}(X, Y) \in [-0.5\{1+3/(n^2-1)\}, 1]$ (*Kendall, 1948*).

The p-value of the independence test using Spearman's footrule, denoted as p, is computed using the normality approximation result in *Diaconis, P., & Graham, R. L. (1977)*. If `pval = TRUE`, bounds of the p-value, p_l, p_u will be computed in the presence of missing data, such that $p \in [p_l, p_u]$. The independence test method proposed in *Zeng et al., 2025* returns p_u as its p-value. This method controls the Type I error regardless of the values of missing data. See *Zeng et al., 2025* for details.

Value

`bounds`	exact bounds of Spearman's footrule.
`bounds.scaled`	exact bounds of scaled Spearman's footrule.
`pvalue`	the p-value for the test. (Only present if argument `pval = TRUE`.)
`bounds.pvalue`	bounds of the p-value of independence test using Spearman's footrule. (Only present if argument `pval = TRUE`.)

References

- Zeng Y., Adams N.M., Bodenham D.A. Exact Bounds of Spearman's footrule in the Presence of Missing Data with Applications to Independence Testing. arXiv preprint arXiv:2501.11696. 2025 Jan 20.
- Kendall, M.G. (1948) Rank Correlation Methods. Charles Griffin, London.
- Diaconis, P. and Graham, R.L., 1977. Spearman's footrule as a measure of disarray. Journal of the Royal Statistical Society Series B: Statistical Methodology, 39(2), pp.262-268.

Examples

```
### compute exact bounds of Spearman's footrule between incomplete ranked lists
X <- c(1, 2, NA, 4, 3)
Y <- c(3, NA, 4, 2, 1)
boundsSFR(X, Y, pval=FALSE)

### compute exact bounds of Spearman's footrule between incomplete vectors of distinct data,
### and perform independence test
X <- c(1.3, 2.6, NA, 4.2, 3.5)
Y <- c(5.5, NA, 6.5, 2.6, 1.1)
boundsSFR(X, Y, pval=TRUE)
```

Index