Package 'ConcordanceTest'

Title An Alternative to the Kruskal-Wallis Based on the Kendall Tau
Distance

Version 1.0.3

Description The Concordance Test is a non-parametric method for testing whether two o more samples originate from the same distribution. It extends the Kendall Tau correlation coefficient when there are only two groups. For details, see Alcaraz J., Anton-Sanchez L., Monge J.F. (2022) The Concordance Test, an Alternative to Kruskal-Wallis Based on the Kendall-tau Distance: An R Package. The R Journal 14, 26–53 <doi:10.32614/RJ-2022-039>.

Depends R (>= 3.3.2)

Imports Rglpk, stats, graphics

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.0

NeedsCompilation no

Author Javier Alcaraz [aut],
Laura Anton-Sanchez [aut, cre],
Juan Francisco Monge [aut]

Maintainer Laura Anton-Sanchez <l.anton@umh.es>

Repository CRAN

Date/Publication 2024-01-15 14:00:02 UTC

R topics documented:

CT_Coefficient *Concordance Coefficient and Kruskal-Wallis Statistic*

Description

This function computes the Concordance coefficient and the Kruskal-Wallis statistic.

Usage

```
CT_Coefficient(Sample_List, H = 0)
```

Arguments

Sample_List   List of numeric data vectors with the elements of each sample.

H   0 by default. If set to 1, the Kruskal-Wallis statistic is also calculated and re-
               turned.

Value

The function returns a list with the following elements:

1. `Sample_Sizes`: Numeric vector of sample sizes.

2. `order_elements`: Numeric vector containing the elements order.

3. `disorder`: Disorder of the permutation given by `order_elements`.

4. `Concordance_Coefficient`: 1-relative disorder of permutation given by `order_elements`.

5. `H_Statistic`: Kruskal-Wallis statistic (only if H = 1).

Examples

```
## Example
A <- c(12,13,15,20,23,28,30,32,40,48)
B <- c(29,31,49,52,54)
C <- c(24,26,44)
Sample_List <- list(A, B, C)
CT_Coefficient(Sample_List)
CT_Coefficient(Sample_List, H = 1)

## Example with ties
A <- c(12,13,15,20,24,29,30,32,40,49)
B <- c(29,31,49,52,54)
C <- c(24,26,44)
Sample_List <- list(A, B, C)
CT_Coefficient(Sample_List, H = 1)
```

CT_Critical_Values   *Critical Values of the Concordance and Kruskal-Wallis Tests*

Description

This function computes the critical values and the p-values for a desired significance levels of .10, .05 and .01. of the Concordance and Kruskal-Wallis tests. Critical values and p-values can be obtained exactly or by simulation (default option).

Usage

```
CT_Critical_Values(Sample_Sizes, Num_Sim = 10000, H = 0, verbose = TRUE)
```

Arguments

Sample_Sizes	Numeric vector (n1,...,nk) containing the number of repetitions of each element, i.e., the size of each sample in the experiment.
Num_Sim	Number of simulations in order to obtain the probability distribution of the statistics. The default is 10000. If set to 0, the critical values and the p-values are obtained exactly. Otherwise they are obtained by simulation.
H	0 by default. If set to 1, the critical values and the p-values of the Kruskal-Wallis test are also calculated and returned.
verbose	A logical indicating if some "progress report" of the simulations should be given. The default is TRUE.

Value

The function returns a list with the following elements:

1. `C_results`: Concordance coefficient results. Critical values and p-values for a desired significance levels of 0.1, .05 and .01.

2. `H_results`: Kruskal-Wallis results. Critical values and p-values for a desired significance levels of 0.1, .05 and .01 (only if H = 1).

Warning

The computational time in exact calculations increases exponentially with the number of elements and with the number of sets.

Examples

```
Sample_Sizes <- c(3,3,3)
CT_Critical_Values(Sample_Sizes, Num_Sim = 0, H = 1)

CT_Critical_Values(Sample_Sizes, Num_Sim = 1000, H = 1)
```

CT_Density_Plot *Density Plot from the Concordance Coefficient and the Kruskal-Wallis Normalized Statistic*

Description

This function performs the graphical visualization of the density distribution of the Concordance coefficient and the Kruskal-Wallis statistic.

Usage

```
CT_Density_Plot(C_freq = NULL, H_freq = NULL)
```

Arguments

C_freq   Probability distribution of the Concordance coefficient obtained with the function `CT_Distribution`.

H_freq   Probability distribution of the Kruskal-Wallis statistic obtained with the function `CT_Distribution`.

Examples

```
Sample_Sizes <- c(5,5,5)
Distributions <-  CT_Distribution(Sample_Sizes, Num_Sim = 1000, H = 1)
C_freq <- Distributions$C_freq
H_freq <- Distributions$H_freq
CT_Density_Plot(C_freq, H_freq)
```

CT_Distribution *Probability Distribution of the Concordance Coefficient and the Kruskal-Wallis Statistic*

Description

This function computes the probability distribution tables of the Concordance coefficient and Kruskal-Wallis statistic. Probability distribution tables can be obtained exactly or by simulation (default option).

Usage

```
CT_Distribution(Sample_Sizes, Num_Sim = 10000, H = 0, verbose = TRUE)
```

Arguments

Sample_Sizes	Numeric vector (n1,...,nk) containing the number of repetitions of each element, i.e., the size of each sample in the experiment.
Num_Sim	Number of simulations in order to obtain the probability distribution of the statistics. The default is 10000. If set to 0, the probability distribution tables are obtained exactly. Otherwise they are obtained by simulation.
H	0 by default. If set to 1, the probability distribution table of the Kruskal-Wallis statistic is also calculated and returned.
verbose	A logical indicating if some "progress report" of the simulations should be given. The default is TRUE.

Value

The function returns a list with the following elements:

1. `C_freq`: Matrix with the probability distribution of the Concordance coefficient. Each row in the matrix contains the disorder, the value of the coefficient, the frequency and its probability.

2. `H_freq`: Matrix with the probability distribution of the Kruskal-Wallis statistic. Each row in the matrix contains the value of the statistic, the frequency and its probability (only if H = 1).

Warning

The computational time in exact calculations increases exponentially with the number of elements and with the number of sets.

Examples

```
Sample_Sizes <- c(5,4)
CT_Distribution(Sample_Sizes, Num_Sim = 0)
CT_Distribution(Sample_Sizes, Num_Sim = 0, H = 1)

CT_Distribution(Sample_Sizes, Num_Sim = 1000)
CT_Distribution(Sample_Sizes, Num_Sim = 1000, H = 1)
```

CT_Hypothesis_Test	*Hypothesis Test for Testing whether Samples Originate from the Same Distribution*

Description

This function performs the hypothesis test for testing whether samples originate from the same distribution.

Usage

```
CT_Hypothesis_Test(Sample_List, Num_Sim = 10000, H = 0, verbose = TRUE)
```

Arguments

`Sample_List`	List of numeric data vectors with the elements of each sample.
`Num_Sim`	The number of used simulations. The default is 10000.
`H`	0 by default. If set to 1, the Kruskal-Wallis test is also performed and returned.
`verbose`	A logical indicating if some "progress report" of the simulations should be given. The default is TRUE.

Value

The function returns a list with the following elements:

1. `results`: Table with the statistics and the signification levels.

2. `C_p-value`: Concordance test signification level.

3. `H_p-value`: Kruskal-Wallis test signification level (only if H = 1).

References

Myles Hollander and Douglas A. Wolfe (1973), Nonparametric Statistical Methods. New York: John Wiley & Sons. Pages 115-120.

Examples

```
## Hollander & Wolfe (1973), 116.
## Mucociliary efficiency from the rate of removal of dust in normal
##  subjects, subjects with obstructive airway disease, and subjects
##  with asbestosis.
x <- c(2.9, 3.0, 2.5, 2.6, 3.2) # normal subjects
y <- c(3.8, 2.7, 4.0, 2.4)      # with obstructive airway disease
z <- c(2.8, 3.4, 3.7, 2.2, 2.0) # with asbestosis
Sample_List <- list(x, y, z)
CT_Hypothesis_Test(Sample_List, Num_Sim = 1000, H = 1)


## Example
A <- c(12,13,15,20,23,28,30,32,40,48)
B <- c(29,31,49,52,54)
C <- c(24,26,44)
Sample_List <- list(A, B, C)
CT_Hypothesis_Test(Sample_List, Num_Sim = 1000, H = 1)

## Example with ties
A <- c(12,13,15,20,24,29,30,32,40,49)
B <- c(29,31,49,52,54)
C <- c(24,26,44)
Sample_List <- list(A, B, C)
CT_Hypothesis_Test(Sample_List, Num_Sim = 1000, H = 1)
```

CT_Probability_Plot   *Probability Plot for the Concordance Coefficient and the Kruskal-Wallis Statistic*

Description

This function performs the graphical visualization of the probability distribution of the Concordance coefficient and the Kruskal-Wallis statistic.

Usage

```
CT_Probability_Plot(C_freq = NULL, H_freq = NULL)
```

Arguments

C_freq   Probability distribution of the Concordance coefficient obtained with the function [CT_Distribution](#).

H_freq   Probability distribution of the Kruskal-Wallis statistic obtained with the function [CT_Distribution](#).

Examples

```
Sample_Sizes <- c(5,5,5)
Distributions <-  CT_Distribution(Sample_Sizes, Num_Sim = 1000, H = 1)
C_freq <- Distributions$C_freq
H_freq <- Distributions$H_freq
CT_Probability_Plot(C_freq)
CT_Probability_Plot(C_freq, H_freq)
```

LOP   *Linear Ordering Problem (LOP)*

Description

This function computes the solution of the Linear Ordering Problem.

Usage

```
LOP(mat_LOP)
```

Arguments

mat_LOP   Preference matrix defining the Linear Ordering Problem. A numeric square matrix for which we want to obtain the permutation of rows/columns that maximizes the sum of the elements above the main diagonal.

Value

The function returns a list with the following elements:

1. `obj_val`: Optimal value of the solution of the Linear Ordering Problem, i.e., the sum of the elements above the main diagonal under the permutation rows/cols solution.

2. `permutation`: Solution of the Linear Ordering Problem, i.e., the rows/cols permutation.

3. `permutation_matrix`: Optimal permutation matrix of the Linear Ordering Problem.

References

Martí, R. and Reinelt, G. The Linear Ordering Problem: Exact and Heuristic Methods in Combinatorial Optimization. Springer, first edition 2011.

Examples

```
## Square matrix
##
## |  1    2    2  |
## |  2    3    3  |
## |  3    2    2  |
##
## The optimal permutation of rows/cols is (2,3,1),
## and the solution of the Linear Ordering Problem is 8.
## Te permutation matrix of the solution is
## |  0    0    0  |
## |  1    0    1  |
## |  1    0    0  |

mat_LOP <- matrix(c(1,2,3,2,3,2,2,3,2), nrow=3)
LOP(mat_LOP)
```

Permutations_With_Repetition

Enumerate the Permutations of the Elements of a Vector When Some of those Elements are Identical

---

Description

This function enumerates the possible combinations of n elements where the first element is repeated n1 times, the second element is repeated n2 times, the third n3 times, ...

Usage

```
Permutations_With_Repetition(Sample_Sizes)
```

Arguments

Sample_Sizes   Numeric vector (n1,...,nk) that indicates the number of times each element is repeated.

Value

Returns a matrix where each row contains a permutation.

Warning

The number of permutations and the computational time increase exponentially with the number of elements and with the number of sets.

Examples

```
Sample_Sizes <- c(2,2,2)
Permutations_With_Repetition(Sample_Sizes)
```

Index