Package ‘rankFD’

January 28, 2020

Type Package

Title Rank-Based Tests for General Factorial Designs

Version 0.0.5

Date 2020-01-28

Author Frank Konietschke, Sarah Friedrich, Edgar Brunner, Markus Pauly

Maintainer Frank Konietschke <frank.konietschke@charite.de>

Depends R (>= 3.2.2)

Description The rankFD() function calculates the Wald-type statistic (WTS) and the ANOVA-type statistic (ATS) for nonparametric factorial designs, e.g., for count, ordinal or score data in a crossed design with an arbitrary number of factors.

Imports lattice (>= 0.20-33), MASS (>= 7.3-43), Matrix (>= 1.2-2), coin (>= 1.1-2)

LazyData TRUE

Suggests RGtk2 (>= 2.20.31)

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2020-01-28 14:20:03 UTC

R topics documented:

calculateGUI — *A graphical user interface for the package rankFD*

Description

This function provides a graphical user interface for calculating rank-based statistical tests in general factorial designs.

Usage

```
calculateGUI()
```

Details

The function produces a GUI for the calculation of the test statistics and for plotting. Data can be loaded via the "load data" button. The formula and the significance level alpha (default: 0.05) need to be specified. One can choose between two different null hypotheses (H_0^F and H_0^p, see the details to rankFD) to be tested as well as weighted or unweighted effects as discussed in Brunner et al. (2016) (r_i and p_i in their notation). If the plot option is chosen, an additional window opens containing information on the plots.

References

Brunner, E., Konietschke, F., Pauly, M. and Puri, M.L. (2016). Rank-Based Procedures in Factorial Designs: Hypotheses about Nonparametric Treatment Effects. arXiv:1606.03973

Akritas, M. G., Arnold, S. F., and Brunner, E. (1997). Nonparametric hypotheses and rank statistics for unbalanced factorial designs. Journal of the American Statistical Association 92, 258-265.

Brunner, E., Dette, H., and Munk, A. (1997). Box-Type Approximations in Nonparametric Factorial Designs. Journal of the American Statistical Association 92, 1494-1502.

Coal — *Coal Acidity*

Description

Coal acidity values determined under each of three NaOH concentration levels for two different samples from each type of coal

Usage

```
data(Coal)
```

Format

A data frame with 18 rows and 3 variables:

Acidity resulting acidity values

NaOH the NaOH concentration

Type three different types of coal: "Morwell", "Yallourn" and "Maddingley"

Source

Hollander, M., Wolfe, D. A., Chicken, E. (2014) *Nonparametric Statistical Methods*. Wiley Series in Probability and Statistics.

Sternhell, S. (1958) Chemistry of brown coals VI: Further aspects of the chemistry of hydroxyl groups in Victorian brown coals. Australian Journal of Applied Science **9**, 375–379.

`Muco` *Half-Time of Mucociliary Clearance*

Description

Mucociliary efficiency was assessed from the rate of removal of dust in three different groups of subjects

Usage

```
data(Muco)
```

Format

A data frame with 14 rows and 2 variables:

HalfTime Half-Time of Mucociliary clearance, assessed from the rate of removal of dust

Disease normal subjects, subjects with obstructive airways disease (OAD) and subjects with asbestosis

Source

Hollander, M., Wolfe, D. A., Chicken, E. (2014) Nonparametric Statistical Methods. Wiley Series in Probability and Statistics.

Thomson, M. L. and Short, M. D.(1969) Mucociliary function in health, chronic obstructive airway disease, and asbestosis. Journal of Applied Physiology **26**, 535–539.

`noether` *Sample size calculation for the Wilcoxon-Mann-Whitney test using the Noether formula.*

Description

Sample size calculation for the Wilcoxon-Mann-Whitney test using the Noether formula.

Usage

```
noether(alpha, power, t, p, x1 = c(0), ties = FALSE)
```

Arguments

`alpha`	two sided type I error rate
`power`	power: detect a relative effect p at least with probability power
`t`	proportion of subjects in the first group (between 0 and 1)
`p`	relative effect
`x1`	advance information is only needed in case of ties
`ties`	TRUE if ties are possible (non continuous distribution), otherwise FALSE

Value

Returns a data frame with the sample sizes for each group

References

Noether, G. E. (1987). Sample Size Determination for Some Common Nonparametric Tests. Journal of the American Statistical Association 85, 645.647.

`psr` *A function for computing pseudo-ranks of data*

Description

The `psr()` function calculates the pseuo-ranks of data in general factorial designs.

Usage

```
psr(formula, data, psranks)
```

Arguments

formula	A model formula object. The left hand side contains the response variable and the right hand side contains the factor variables of interest. Please use one-way layouts only for the computation of the pseudo-ranks.
data	A data.frame, list or environment containing the variables in formula. The default option is NULL.
psranks	A header specifying the name of the pseudo ranks in the output data set.

Details

The pseudo-ranks are exported within a new column attached to the given data set.

References

Konietschke, F., Hothorn, L. A., & Brunner, E. (2012). Rank-based multiple test procedures and simultaneous confidence intervals. Electronic Journal of Statistics, 6, 738-759.

Kaufmann, J., Werner, C., and Brunner, E. (2005). Nonparametric methods for analysing the accuracy of diagnostic tests with multiple readers. Statistical Methods in Medical Research 14, 129 - 146

See Also

rankFD

Examples

```
data(Muco)
Muco2 <- psr(HalfTime~Disease,data=Muco, psranks="Mypseudos")
```

---

rank.two.samples *A function for analyzing two-sample problems*

Description

The rank.two.samples() function calculates the weighted or unweighted treatment effect in a two-sample problem. In addition to rankFD, the user can specify the alternative and choose from a variety of different possibilities to calculate confidence intervals, see the details below. Furthermore, a Wilcoxon test is calculate with the possibility to consider shift effects.

Usage

```
rank.two.samples(formula, data, conf.level = 0.95,
  alternative = c("two.sided", "less", "greater"), rounds = 3,
  method = c("logit", "probit", "normal", "t.app", "permu"),
  plot.simci = FALSE, info = TRUE, wilcoxon = c("asymptotic",
  "exact"), shift.int = TRUE, nperm = 10000)
```

Arguments

formula	A model formula object. The left hand side contains the response variable and the right hand side contains the factor variables of interest. An interaction term must be specified.
data	A data.frame, list or environment containing the variables in formula. The default option is NULL.
conf.level	A number specifying the confidence level; the default is 0.95.
alternative	Which alternative is considered? One of "two.sided", "less", "greater".
rounds	Value specifying the number of digits the results are rounded to.
method	specifying the method used for calculation of the confidence intervals. One of "logit", "probit", "normal", "t.app" and "permu".
plot.simci	Logical, indicating whether or not confidence intervals should be plotted
info	Logical. info = FALSE suppresses the output of additional information concerning e.g. the interpretation of the test results.
wilcoxon	asymptotic or exact calculation of Wilcoxon test.
shift.int	Logical, indicating whether or not shift effects should be considered.
nperm	Number of permutations used, default is 10000.

Details

The confidence intervals are given for the treatment effect $p = P(X_1 < Y_1) + \frac{1}{2}P(X_1 = Y_1)$ underlying the Wilcoxon-Mann-Whitney test including tied data. Different methods for calculation can be chosen, see Pauly et al.(2016) for the permutation approach, Brunner and Munzel (2000) for the t-approximation and Kaufmann et al.(2005) for the transformations. For plotting, the parameter plot.simci must be set to TRUE.

References

Brunner, E. and Munzel, U. (2000). The nonparametric Behrens-Fisher problem: Asymptotic theory and a small-sample approximation. Biometrical Journal 1, 17 - 21.

Kaufmann, J., Werner, C., and Brunner, E. (2005). Nonparametric methods for analysing the accuracy of diagnostic tests with multiple readers. Statistical Methods in Medical Research 14, 129 - 146

Pauly, M., Asendorf, T., and Konietschke, F. (2016). Permutation tests and confidence intervals for the area under the ROC-curve. Biometrical Journal, to appear.

See Also

rankFD

Examples

```
data(Muco)
Muco2 <- subset(Muco, Disease != "OAD")
twosample <- rank.two.samples(HalfTime ~ Disease, data = Muco2,
```

```
alternative = "greater", method = "probit", wilcoxon = "exact", plot.simci = FALSE,
shift.int = FALSE)
```

---

rankFD *Rank-based tests for general factorial designs*

Description

The package provides the Wald-type as well as the ANOVA-type statistic for rank-based factorial designs, i.e., even for ordinal data. It is implemented for crossed designs and allows for an arbitrary number of factor combinations as well as different sample sizes.

Usage

```
rankFD(formula, data, alpha = 0.05, CI.method = c("Logit", "Normal"),
  effect = c("unweighted", "weighted"), hypothesis = c("H0F", "H0p"),
  Factor.Information = FALSE)
```

Arguments

`formula`	A model `formula` object. The left hand side contains the response variable and the right hand side contains the factor variables of interest. An interaction term must be specified.
`data`	A data.frame, list or environment containing the variables in `formula`. The default option is `NULL`.
`alpha`	A number specifying the significance level; the default is 0.05.
`CI.method`	Either "Logit" or "Normal", specifying the method used for calculation of the confidence intervals.
`effect`	Should the weighted or unweighted effects be calculated?
`hypothesis`	The null hypothesis to test, either "H0F" or "H0p".
`Factor.Information`	Logical. If TRUE, descriptive statistics for the different factor level combinations is printed.

Details

The rankFD() function calculates the Wald-type statistic (WTS) and the ANOVA-type statistic (ATS) for general factorial designs for testing the null hypotheses $H_0^F : CF = 0$ (cf. Akritas et al. (1997) for the WTS and Brunner et al. (1997) for the ATS) based on weighted effect measures, and $H_0^p : Cp = 0$ for the vector of unweighted treatment effects as described in Brunner et al. (2016). In the latter paper, the CIs for the unweighted effects (p_i in their notation) are described and CIs for the weighted effects (r_i in their notation) are obtained similarly.

Value

An rankFD object containing the following components:

- `Descriptive` Some descriptive statistics of the data for all factor level combinations. Displayed are the number of individuals per factor level combination (size), the relative effect (pd), variance and 100*(1-alpha)% confidence intervals.
- `WTS` The value of the WTS along with degrees of freedom of the central chi-square distribution and p-value
- `ATS` The value of the ATS, degrees of freedom of the central F distribution and the corresponding p-value.

References

Brunner, E., Bathke, A. and Konietschke, F. Rank and Pseudo-Rank Procedures for Independent Observations in Factorial Designs. Springer International Publishing, 2018.

Brunner, E., Konietschke, F., Pauly, M., & Puri, M. L. (2017). Rank-based procedures in factorial designs: Hypotheses about non-parametric treatment effects. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79(5), 1463-1485.

Akritas, M. G., Arnold, S. F., and Brunner, E. (1997). Nonparametric hypotheses and rank statistics for unbalanced factorial designs. Journal of the American Statistical Association 92, 258-265.

Brunner, E., Dette, H., and Munk, A. (1997). Box-Type Approximations in Nonparametric Factorial Designs. Journal of the American Statistical Association 92, 1494-1502.

Examples

```
data(Coal)
model <- rankFD(Acidity ~ NaOH * Type, data = Coal, CI.method = "Normal",
effect = "unweighted", hypothesis = "H0F")

data(Muco)
model.oneway <- rankFD(HalfTime ~ Disease, data = Muco, CI.method = "Logit",
effect = "weighted", hypothesis = "H0p")
plot(model.oneway)
```

WMWSSP *Sample size calculation for the Wilcoxon-Mann-Whitney test.*

Description

Sample size calculation for the Wilcoxon-Mann-Whitney test.

Usage

```
WMWSSP(x1, x2, alpha, power, t)
```

Arguments

x1	advance information for the first group
x2	advance information for the second group
alpha	two sided type I error rate
power	power with the sample sizes of each group
t	proportion of subjects in the first group

Value

Returns a data frame

References

Brunner, E., Bathke A. C. and Konietschke, F. Rank- and Pseudo-Rank Procedures in Factorial Designs - Using R and SAS. Springer Verlag.

Examples

```
x1 = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 2)
x2 = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3)
WMWSSP(x1,x2,0.05,0.8,0.5)
```

Index