

Package ‘TestsSymmetry’

March 1, 2024

Type Package

Title Tests for Symmetry when the Center of Symmetry is Unknown

Version 1.0.0

Date 2024-02-22

Maintainer Jiaojiao Zhou <woerge99@gmail.com>

Description Provides functionality of a statistical testing implementation whether a dataset comes from a symmetric distribution when the center of symmetry is unknown, including Wilcoxon test and sign test procedure. In addition, sample size determination for both tests is provided. The Wilcoxon test procedure is described in Vexler et al. (2023) <<https://www.sciencedirect.com/science/article/abs/pii/S0167947323000579>>, and the sign test is outlined in Gastwirth (1971) <<https://www.jstor.org/stable/2284233>>.

License GPL-3

Encoding UTF-8

Imports Rcpp, MASS, e1071, rpart

LinkingTo Rcpp

RoxygenNote 7.2.2.9000

Depends R (>= 3.5.0)

LazyData true

NeedsCompilation yes

Author Jiaojiao Zhou [aut, cre],
Xinyu Gao [aut],
Albert Vexler [aut]

Repository CRAN

Date/Publication 2024-03-01 15:02:40 UTC

R topics documented:

getV	2
get_quant_H0	2

mod.symm.test	3
model	5
n.symm.test	5
plasma.silicon	7
tree_model_sign	7

Index	8
--------------	----------

getV	<i>Getting asymptotic variance of modified wilcox test statistic</i>
------	--

Description

Getting asymptotic variance of modified wilcox test statistic

Usage

```
getV(x_, y_ = NULL)
```

Arguments

x_	A data set
y_	Another data set, default is NULL

Value

Asymptotic variance of modified wilcoxon test statistic

get_quant_H0	<i>Calculate the necessary quantities under null hypothesis</i>
--------------	---

Description

Calculate the necessary quantities under null hypothesis

Usage

```
get_quant_H0(x)
```

Arguments

x	pilot sample
---	--------------

 mod.symm.test

Wilcoxon and Sign tests for symmetry about an unknown center

Description

R built-in function ‘wilcox.test()’ is designed to perform both the one- and two-sample Wilcoxon test when the center of symmetry is specified. The procedure ‘mod.symm.test()’ extends the capabilities of ‘wilcox.test()’ for situations where the center of symmetry is unknown. Such cases can be found in, e.g., regression residuals evaluations, as well as in the book ‘Nonparametric statistical methods using R’ by Klokke and McKean, and in the scholarly work of Gastwirth.

Usage

```
mod.symm.test(
  x,
  y = NULL,
  alternative = c("two.sided", "left.skewed", "right.skewed"),
  method = "wilcox"
)
```

Arguments

x	numeric vector of data values. Non-finite (e.g., infinite or missing) values will be omitted.
y	an optional numeric vector of data values: as with x non-finite values will be omitted.
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "right.skewed", or "left.skewed". You can specify just the initial letter. "right.skewed": test whether positively skewed, "left.skewed" : test whether negatively skewed.
method	a character string specifying which symmetry test to be used, "wilcox" refers to Wilcoxon signed-rank test, and "sign" is sign test.

Details

When "wilcox", the default method, is used, the test statistic ‘W’ has the form of the Wilcoxon test statistic with the unknown center of symmetry replaced by its sample mean estimator. For more details, see Vexler et al. (2023)

When method="sign" is used, the test statistic ‘S’ is the total number of the observations that smaller than sample mean. For more details, see Gastwirth (1971).

Value

A list of class "htest" containing the following components:

- statistic - the value of the test statistic.

- var - the asymptotic variance.
- alternative - a character string describing the alternative hypothesis.
- p.value - the p-value for the test.
- method - the type of test applied.

Author(s)

Jiaojiao Zhou, Xinyu Gao, Albert Vexler

References

Vexler, A., Gao, X., & Zhou, J. (2023). How to implement signed-rank 'wilcox.test()' type procedures when a center of symmetry is unknown. *Computational Statistics & Data Analysis*, 107746.

Gastwirth, J. L. (1971). On the Sign Test for Symmetry. *Journal of the American Statistical Association*, 66(336), 821-823.

Examples

```
# A study measures the plasma silicon levels before and after silicone implants surgery in 30
# women to evaluate the effect of the surgery. Informally speaking, we can be interested
# in that there is an unknown constant shift such that the the plasma silicon level of
# post-surgery can be explained completely based on that of pre-surgery. This can be stated
# as the null hypothesis 'H_0` The difference of plasma silicon level between post-surgery and
# pre-surgery has a symmetric distribution around a shift that is unknown.
data("plasma.silicon")
post <- plasma.silicon$postoperative
pre <- plasma.silicon$preoperative
# post <- c(0.21,0.24,0.1,0.12,0.28,0.25,0.22,0.21,0.22,0.23,0.22,0.24,0.45,0.38,
#          0.23,0.22,0.18,0.15,0.04,0.14,0.24,0.2,0.24,0.18,0.19,0.15,0.26,0.3,0.22,0.24)
# pre <- c(0.15,0.13,0.39,0.2,0.39,0.42,0.24,0.18,0.26,0.12,0.1,0.11,0.19, 0.15,0.27,
#          0.28,0.11,0.11,0.18,0.18,0.24,0.48,0.27,0.22,0.18,0.19,0.32,0.31,0.19,0.21)
mod.symm.test(x=post, y=pre, alternative = "two.sided", method = "wilcox")

# Result:
# Modified Wilcoxon signed-rank test
# data:  post and pre
# W = 238, p-value = 0.767
# alternative hypothesis: two.sided

# Interpretation:
# Test statistic `W` is the number of walsh average higher than sample mean, see more details
# in paper authored by Vexler, etc.
# p-value is 0.767, which implies there is no clue to reject the null hypothesis that
# the distribution of the difference of plasma silicon levels before and after
# silicone implants surgery is symmetric.
```

model	<i>Tree_model</i>
-------	-------------------

Description

Tree model object for wilcox method to find the optimal alp

Usage

```
model
```

Format

An rpart object that contains a tree model which fits alp to npilot, sk, kur, sd and distance_p2_half.

n.symm.test	<i>Sample size determination for nonparametric tests of symmetry when the center is unknown</i>
-------------	---

Description

Determine the sample size required for one-sample Wilcoxon signed-rank test or Sign test with an unknown symmetry center estimated by sample mean given a target power. The function uses learning sample in order to predict (calculate) sample size needed to reach a preassumed power based on the underlying data that is exemplified by the learning sample.

Usage

```
n.symm.test(
  x,
  sig.level = 0.05,
  power = 0.8,
  method = "wilcox",
  alternative = c("two.sided")
)
```

Arguments

x	learning sample data, numeric vector of data values
sig.level	significance level (Type I error probability), the default value is 0.05
power	power of test (1 minus Type II error probability), the default value is 0.8
method	a character string specifying which symmetry test to be used, "wilcox" refers to Wilcoxon signed-rank test, and "sign" sign test
alternative	a character string specifying the alternative hypothesis "two.sided" : test whether skewed

Details

A Normal approximation to the power requires specification of some unknown quantities in the non-parametric context. In this regard, empirical smoothed CDF and Bootstrap methods were leveraged to estimate these quantities using learning sample 'x'.

Remark: If the test provides a power, say, P , based on the learning data and P is higher than the target power, a warning message will be shown. However, a needed sample size N to reach the target power will be conducted.

Value

A list of class "power.htest" containing the following components:

- N - sample size estimated
- sig.level - significance level (Type I error probability)
- power - power of test (1 minus Type II error probability)
- method - the test method applied
- alternative - two-sided test. Can be abbreviated

References

Chakraborti, S., Hong, B., & van de Wiel, M. A. (2006). A Note on Sample Size Determination for a Nonparametric Test of Location. *Technometrics*, 48(1), 88-94.

Vexler, A., Gao, X., & Zhou, J. (2023). How to implement signed-rank 'wilcox.test()' type procedures when a center of symmetry is unknown. *Computational Statistics & Data Analysis*, 107746.

Gastwirth, J. L. (1971). On the Sign Test for Symmetry. *Journal of the American Statistical Association*, 66(336), 821-823.

Examples

```
data("plasma.silicon")
post <- plasma.silicon$postoperative
pre <- plasma.silicon$preoperative
diff <- post - pre
n.symm.test(diff, sig.level = 0.05, power = 0.5, method = "wilcox", alternative = "two.sided" )

# Result:
# Sample size calculation under wilcox procedure

#           N = 83
#   sig.level = 0.05
#         power = 0.5
#         type = wilcox
# alternative = two.sided

# Interpretation:
# Given the pilot sample `diff` and the significance level 0.05. The sample size of
# the data that is expected to provide the target power 0.5 of the Wilcoxon test procedure
```

is computed as 83.

plasma.silicon *Dataset: plasma.silicon*

Description

A study was conducted to measure the plasma silicon level in blood within 30 women who have been taken a silicone implants surgery. Plasma silicon levels (microg per gram dry weight) were taken prior to surgical placement of the implants. A post-surgery washout period was allowed, and the plasma silicon levels were retaken.

Usage

```
plasma.silicon
```

Format

A data frame contains 30 rows and 3 variables:

PatientNo Patient Number.

preoperative pre-operative measurement of plasma silicion level.

postoperative post-operative measurement of plasma silicion level.

Source

Riffenburgh, R. H., & Gillen, D. L. (2020). Statistics in medicine. Academic press. Appendix 2 databases

References

Riffenburgh, R. H., & Gillen, D. L. (2020). Statistics in medicine. Academic press.

tree_model_sign *Tree_model_sign*

Description

Tree model object to find the optimal alp for sign test method.

Usage

```
tree_model_sign
```

Format

An rpart object that contains a tree model which fits alp to npilot, sk, kur, sd.

Index

* **datasets**

plasma.silicon, 7

* **rpart**

model, 5

tree_model_sign, 7

get_quant_H0, 2

getV, 2

mod.symm.test, 3

model, 5

n.symm.test, 5

plasma.silicon, 7

tree_model_sign, 7