Package 'randtests'

April 23, 2024

Type Package

Title Testing Randomness in R

Version 1.0.2

Date 2024-04-22

Maintainer Frederico Caeiro `<fac@fct.unl.pt>`

Description Provides several non parametric randomness tests for numeric sequences.

License GPL (>= 2)

NeedsCompilation no

Author Frederico Caeiro [aut, cre] (<https://orcid.org/0000-0001-8628-7281>),
Ayana Mateus [aut] (<https://orcid.org/0000-0002-5630-3321>)

Repository CRAN

Date/Publication 2024-04-23 15:30:02 UTC

R topics documented:

---

randtests-package   *Testing randomness in R*

Description

The package `randtests` implements several nonparametric randomness tests of hypothesis.

Details

Package:	randomtests
Type:	Package
Version:	1.0.2
Date:	2024-04-22
License:	GPL (>=2)
LazyLoad:	yes

Randomness is a common assumption in many statistical methods. When such assumption is not fulfilled, we may draw wrong conclusions. Although in many datasets a simple graphical analysis is enough to check such assumption, in others a test of hypothesis is required.

Author(s)

Frederico Caeiro and Ayana Mateus

Maintainer: Frederico Caeiro <fac@fct.unl.pt>

References

Mateus A. and Caeiro F. (2014). An R implementation of several Randomness Tests. In T. E. Simos, Z. Kalogiratou and T. Monovasilis (eds.), *AIP Conf. Proc.* **1618**, 531–534.

---

bartels.rank.test   *Bartels Rank Test*

Description

Performs the Bartels rank test of randomness.

Usage

```
bartels.rank.test(x, alternative, pvalue="normal")
```

Arguments

x	a numeric vector containing the observations
alternative	a character string with the alternative hypothesis. Must be one of `"two.sided"` (default), `"left.sided"` or `"right.sided"`. You can specify just the initial letter.
pvalue	a character string specifying the method used to compute the p-value. Must be one of `normal` (default), `beta`, `exact` or `auto`.

Details

Missing values are removed.

This is the rank version of von Neumann's Ratio Test for Randomness (von Neumann, 1941).

The test statistic RVN is

$$RVN = \frac{\sum_{i=1}^{n-1}(R_i - R_{i+1})^2}{\sum_{i=1}^{n}\left(R_i - (n+1)/2\right)^2}$$

where $R_i = rank(X_i), i = 1, \ldots, n$. It is known that $(RVN - 2)/\sigma$ is asymptotically standard normal, where $\sigma^2 = \frac{4(n-2)(5n^2-2n-9)}{5n(n+1)(n-1)^2}$.

The possible `alternative` are `"two.sided"`, `"left.sided"` and `"right.sided"`. By using the alternative "two.sided" the null hypothesis of randomness is tested against nonrandomness. By using the alternative `"left.sided"` the null hypothesis of randomness is tested against a trend. By using the alternative `"right.sided"` the null hypothesis of randomness is tested against a systematic oscillation.

By default (if pvalue is not specified), a normal approximation is used to compute the p-value. With `beta`, the p-value is computed using an approximation given by the Beta distribution. With `exact`, the exact p-value is computed. The option `exact` requires the computation of the exact distribution of the statistic test under the null hypothesis and should only be used for small sample sizes ($n \leq 10$).

Value

A list with class "htest" containing the components:

statistic	the value of the normalized statistic test.
parameter, n	the size of the data, after the remotion of consecutive duplicate values.
p.value	the p-value of the test.
alternative	a character string describing the alternative hypothesis.
method	a character string indicating the test performed.
data.name	a character string giving the name of the data.
rvn	the value of the RVN statistic (not shown on screen).
nm	the value of the NM statistic, the numerator of RVN (not shown on screen).
mu	the mean value of the RVN statistic (not shown on screen).
var	the variance of the RVN statistic (not shown on screen).

Author(s)

Frederico Caeiro

References

Bartels, R. (1982). The Rank Version of von Neumann's Ratio Test for Randomness, *Journal of the American Statistical Association*, **77**(377), 40–46.

Gibbons, J.D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference*, 4th ed. (pp. 97–98). URL: http://books.google.pt/books?id=dPhtioXwI9cC&lpg=PA97&ots=ZGaQCmuEUq

von Neumann, J. (1941). Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *The Annals of Mathematical Statistics* **12**(4), 367–395. doi:10.1214/aoms/1177731677. https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-12/issue-4/Distribution-of-the-Ratio-of-the-Mean-Square-Successive-Difference/10.1214/aoms/1177731677.full

See Also

dbartelsrank, pbartelsrank

Examples

```
##
## Example 5.1 in Gibbons and Chakraborti (2003), p.98.
## Annual data on total number of tourists to the United States for 1970-1982.
##
years <- 1970:1982
tourists <- c(12362, 12739, 13057, 13955, 14123,  15698, 17523, 18610, 19842,
      20310, 22500, 23080, 21916)
plot(years, tourists, pch=20)
bartels.rank.test(tourists, alternative="left.sided", pvalue="beta")
# output
#
#  Bartels Ratio Test
#
#data:  tourists
#statistic = -3.6453, n = 13, p-value = 1.21e-08
#alternative hypothesis: trend


##
## Example in Bartels (1982).
## Changes in stock levels for 1968-1969 to 1977-1978 (in $A million), deflated by the
## Australian gross domestic product (GDP) price index (base 1966-1967).
x <- c(528, 348, 264, -20, -167, 575, 410, -4, 430, -122)
bartels.rank.test(x, pvalue="beta")
```

BartelsRank	*Distribution of the Bartels Rank Test Statistic NM*

Description

Probability function, distribution function for the distribution of the Bartels Rank statistic NM, for a sample of size n.

Usage

```
dbartelsrank(x, n, log = FALSE)
pbartelsrank(q, n, lower.tail = TRUE, log.p = FALSE)
```

Arguments

x, q	a numeric vector of quantiles.
n	number of observations to return.
log, log.p	logical; if TRUE, probabilities p are given as log(p).
lower.tail	logical; if TRUE (default), probabilities are $P[X \leq x]$ otherwise, $P[X > x]$.

Value

dbartelsrank gives the probability function and pbartelsrank gives the distribution function.

Warning

This function can use large amounts of memory and stack (and even crash R if the stack limit is exceeded) if the sample size n is large.

Author(s)

Frederico Caeiro

References

Bartels, R. (1982). The Rank Version of von Neumann's Ratio Test for Randomness, *Journal of the American Statistical Association*, **77**(377), 40–46.

Gibbons, J.D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference*, 4th ed. (pp. 97–98). URL: http://books.google.pt/books?id=dPhtioXwI9cC&lpg=PA97&ots=ZGaQCmuEUq

See Also

bartels.rank.test to calculate the value of the statistic NM from data.

---

cox.stuart.test   *Cox Stuart Trend Test*

Description

Performs the Cox Stuart test of randomness.

Usage

```
cox.stuart.test(x, alternative)
```

Arguments

x   a numeric vector containing the data

alternative   a character string with the alternative hypothesis. Must be one of `"two.sided"`
                    (default), `"left.sided"` or `"right.sided"`. You can specify just the initial let-
                    ter.

Details

Missing values are removed.

Data is grouped in pairs with the ith observation of the first half paired with the ith observation of
the second half of the time-ordered data. If the length of vector X is odd the middle observation is
eliminated. The cox stuart test is then simply a sign test applied to these paired data.

The possible values `"two.sided"`, `"left.sided"` and `"right.sided"` define the alternative hypoth-
esis. By using the alternative `"two.sided"` the null hypothesis of randomness is tested against either
an upward trend or an downward trend. By using the alternative `"left.sided"` the null hypothesis
of randomness is tested against an upward trend. By using the alternative `"right.sided"` the null
hypothesis of randomness is tested against a downward trend.

Value

A list with class "htest" containing the components:

statistic   The number of pairs with a signal "+"

n   The number of pairs, after eliminanting ties.

p.value   the p-value for the test.

alternative   a character string describing the alternative hypothesis.

method   a character string indicating the test performed.

data.name   a character string giving the name of the data.

Author(s)

Ayana Mateus

References

Conover, W.J. (1999). *Practical Nonparametric Statistics*, 3rd edition, John Wiley & Sons (p. 166).

Cox, D. R. and Stuart, A. (1955). Some quick sign test for trend in location and dispersion, *Biometrika*, **42**, 80-95.

Sprent, P. and Smeeton, N.C. (2007). *Applied Nonparametric Statistical Methods*, 4th ed., Chapman and Hall/CRC Texts in Statistical Science (p. 108).

Examples

```
##
## Example 1
## Conover (1999)
## The total annual precipitation recorded each year, for 19 years.
##
precipitation <- c(45.25, 45.83, 41.77, 36.26, 45.37, 52.25, 35.37, 57.16, 35.37, 58.32,
41.05, 33.72, 45.73, 37.90, 41.72, 36.07, 49.83, 36.24, 39.90)
cox.stuart.test(precipitation)

##
## Example 2
## Sweet potato production, harvested in the United States, between 1868 and 1937.
##
data(sweetpotato)
cox.stuart.test(sweetpotato$production)
```

---

difference.sign.test   *Difference Sign Test*

Description

Performs the nonparametric Difference-sign test of randomness.

Usage

```
difference.sign.test(x, alternative)
```

Arguments

x   a numeric vector containing the data

alternative   a character string specifying the alternative hypothesis. Must be one of `"two.sided"`
                (default), `"left.sided"` or `"right.sided"`.

Details

Consecutive equal values are eliminated.

The possible values `"two.sided"`, `"left.sided"` and `"right.sided"` define the alternative hypothesis. By using the alternative `"two.sided"` the null hypothesis of randomness is tested against either an increasing or decreasing trend. By using the alternative `"left.sided"` the null hypothesis of randomness is tested against an decreasing trend. By using the alternative `"right.sided"` the null hypothesis of randomness is tested against an increasing trend

Value

A list with class "htest" containing the components:

`statistic`	the (normalized) value of the statistic test.
`parameter`	the size of the data, after the remotion of consecutive duplicate values.
`p.value`	the p-value of the test.
`alternative`	a character string describing the alternative hypothesis.
`method`	a character string indicating the test performed.
`data.name`	a character string giving the name of the data.
`ds`	the total number of positive diferences (not shown on screen).
`mu`	the mean value of the statistic DS (not shown on screen).
`var`	the variance of the statistic DS (not shown on screen).

Author(s)

Ayana Mateus and Frederico Caeiro

References

Brockwell, P.J. and Davis, R.A. (2002). *Introduction to Time Series and Forecasting*, 2nd edition, Springer (p. 37).

Mateus, A. and Caeiro, F. (2013). Comparing several tests of randomness based on the difference of observations. In T. Simos, G. Psihoyios and Ch. Tsitouras (eds.), *AIP Conf. Proc.* **1558**, 809–812.

Moore, G. H. and Wallis, W. A. (1943). Time Series Significance Tests Based on Signs of Differences, *Journal of the American Statistical Association*, **38**, 153–154.

Examples

```
##
## Example 1
## Annual Canadian Lynx trappings 1821-1934 in Canada.
## Available in datasets package
##
## Not run: plot(lynx)
difference.sign.test(lynx)

##
## Example 2
```

```
## Sweet potato production, harvested in the United States, between 1868 and 1937.
## Available in this package.
##
data(sweetpotato)
difference.sign.test(sweetpotato$production)
```

---

permut   *Generate all permutations of m elements of a vector*

Description

Generate all permutations of x taken m at a time. If argument FUN is not NULL, applies a function given by the argument to each permutation.

Usage

```
permut(x, m=length(x), FUN=NULL,...)
```

Arguments

x	vector source for permutations.
m	number of elements to choose. Default is `m=length(x)`.
FUN	function to be applied to each permutation; default NULL means the identity, i.e., to return the permutation.
...	optionally, further arguments to FUN.

Details

Based on function `permutations` from package `gtools`. This function is required for the computation of the exact p-value of some randomness tests.

Value

A matrix with one permutation, or the value returned by FUN, in each line.

---

rank.test *Mann-Kendall Rank Test*

Description

Performs the Mann-Kendall rank test of randomness.

Usage

```
rank.test(x, alternative)
```

Arguments

x	a numeric vector containing the observations
alternative	a character string specifying the alternative hypothesis. Must be one of "two.sided" (default), "left.sided" or "right.sided".

Details

Missing values are removed.

The possible alternative values are "two.sided", "left.sided" and "right.sided" define the alternative hypothesis. By using the alternative "left.sided" the null of randomness is tested against a downward trend. By using the alternative "right.sided" the null hypothesis of randomness is tested against a upward trend.

Value

A list with class "htest" containing the components:

statistic	the value of the normalized statistic test.
parameter	The size n of the data.
p.value	the p-value of the test.
method	a character string indicating the test performed.
data.name	a character string giving the name of the data.
P	the value of the (non normalized) P statistic.
mu	the mean value of the P statistic.
var	the variance of the P statistic.

Author(s)

Ayana Mateus and Frederico Caeiro

References

Brockwell, P.J. and Davis, R.A. (2002). *Introduction to Time Series and Forecasting*, 2nd edition, Springer (p. 37).

Mann, H.B. (1945). Nonparametric test against trend. *Econometrica*, **13**, 245–259.

Kendall, M. (1990). *Rank correlation methods*, 5th edition. Oxford University Press, USA.

Examples

```
##
## Example 1
## Sweet potato yield per acre, 1868-1937 in the United States.
## Available in this package.
##
data(sweetpotato)
rank.test(sweetpotato$yield)

##
## Example 2
## Old Faithful Geyser Data on Eruption time in mins.
## Available in R package datasets.
##
rank.test(faithful$eruptions)
```

Runs   *Distribution of the Wald Wolfowitz Runs Statistic*

Description

Probability function, distribution function, quantile function and random generation for the distribution of the Runs statistic obtained from samples with n_1 and n_2 elements of each type.

Usage

```
druns(x, n1, n2, log = FALSE)
pruns(q, n1, n2, lower.tail = TRUE, log.p = FALSE)
qruns(p, n1, n2, lower.tail = TRUE, log.p = FALSE)
rruns(n, n1, n2)
```

Arguments

x, q	a numeric vector of quantiles.
p	a numeric vector of probabilities.
n	number of observations to return.
n1, n2	the number of elements of first and second type, respectively.
log, log.p	logical; if TRUE, probabilities p are given as log(p).
lower.tail	logical; if TRUE (default), probabilities are P[X ≤ x] otherwise, P[X > x].

Details

The Runs distribution has probability function

$$P(R = r) = \begin{cases} \dfrac{2\binom{n_1-1}{r/2-1}\binom{n_2-1}{r/2-1}}{\binom{n_1+n_2}{n_1}}, & \text{if } r \text{ is even} \\[3ex] \dfrac{\binom{n_1-1}{(r-1)/2}\binom{n_2-1}{(r-3)/2} + \binom{n_1-1}{(r-3)/2}\binom{n_2-1}{(r-1)/2}}{\binom{n_1+n_2}{n_1}}, & \text{if } r \text{ is odd} \end{cases}$$

for $r = 2, 3, \ldots, 2\min(n_1 + n_2) + c$ with $c = 0$ if $n_1 = n_2$ or $c = 1$ if $n_1 \neq n_2$.

If an element of x is not integer, the result of druns is zero.

The quantile is defined as the smallest value x such that $F(x) \geq p$, where F is the distribution function.

Value

druns gives the probability function, pruns gives the distribution function and qruns gives the quantile function.

References

Swed, F.S. and Eisenhart, C. (1943). Tables for Testing Randomness of Grouping in a Sequence of Alternatives, *Ann. Math Statist.* **14**(1), 66-87.

Examples

```
##
## Example: Distribution Function
## Creates Table I in Swed and Eisenhart (1943), p. 70,
## with n1 = 2 and n1 <= n2 <= 20
##
m <- NULL
for (i in 2:20){
  m <- rbind(m, pruns(2:5,2,i))
}
rownames(m)=2:20
colnames(m)=2:5
#
#                2         3         4 5
# 2  0.333333333 0.6666667 1.0000000 1
# 3  0.200000000 0.5000000 0.9000000 1
# 4  0.133333333 0.4000000 0.8000000 1
# 5  0.095238095 0.3333333 0.7142857 1
# 6  0.071428571 0.2857143 0.6428571 1
# 7  0.055555556 0.2500000 0.5833333 1
# 8  0.044444444 0.2222222 0.5333333 1
# 9  0.036363636 0.2000000 0.4909091 1
# 10 0.030303030 0.1818182 0.4545455 1
# 11 0.025641026 0.1666667 0.4230769 1
# 12 0.021978022 0.1538462 0.3956044 1
# 13 0.019047619 0.1428571 0.3714286 1
# 14 0.016666667 0.1333333 0.3500000 1
```

```
# 15 0.014705882 0.1250000 0.3308824 1
# 16 0.013071895 0.1176471 0.3137255 1
# 17 0.011695906 0.1111111 0.2982456 1
# 18 0.010526316 0.1052632 0.2842105 1
# 19 0.009523810 0.1000000 0.2714286 1
# 20 0.008658009 0.0952381 0.2597403 1
#
```

---

runs.test	*Wald-Wolfowitz Runs Test*

Description

Performs the Wald-Wolfowitz runs test of randomness for continuous data.

Usage

```
runs.test(x, alternative, threshold, pvalue, plot)
```

Arguments

x	a numeric vector containing the observations
alternative	a character string with the alternative hypothesis. Must be one of `"two.sided"` (default), `"left.sided"` or `"right.sided"`. You can specify just the initial letter.
threshold	the cut-point to transform the data into a dichotomous vector
pvalue	a character string specifying the method used to compute the p-value. Must be one of normal (default), or exact.
plot	a logic value to select whether a plot should be created. If 'TRUE', then the graph will be plotted.

Details

Data is transformed into a dichotomous vector according as each values is above or below a given `threshold`. Values equal to the level are removed from the sample.

The default `threshold` value used in applications is the sample median which give us the special case of this test with $n_1 = n_2$, the runs test above and below the median.

The possible `alternative` values are `"two.sided"`, `"left.sided"` and `"right.sided"` define the alternative hypothesis. By using the alternative `"left.sided"` the null of randomness is tested against a trend. By using the alternative `"right.sided"` the null hypothesis of randomness is tested against a first order negative serial correlation.

Value

A list with class "htest" containing the components:

statistic	the value of the normalized statistic test.
parameter	a vector with the sample size, and the values of n_1 and n_2.
p.value	the p-value of the test.
alternative	a character string describing the alternative hypothesis.
method	a character string indicating the test performed.
data.name	a character string giving the name of the data.
runs	the total number of runs (not shown on screen).
mu	the mean value of the statistic test (not shown on screen).
var	the variance of the statistic test (not shown on screen).

Author(s)

Frederico Caeiro

References

Brownlee, K. A. (1965). *Statistical Theory and Methodology in Science and Engineering*, 2nd ed. New York: Wiley.

Gibbons, J.D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference*, 4th ed. (pp. 78–86). URL: http://books.google.pt/books?id=dPhtioXwI9cC&lpg=PA97&ots=ZGaQCmuEUq

Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population, *The Annals of Mathematical Statistics* **11**, 147–162. doi:10.1214/aoms/1177731909. https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-11/issue-2/On-a-Test-Whether-Two-Samples-are-from-the-Same/10.1214/aoms/1177731909.full

Examples

```
##
## Example 1
## Data from example in Brownlee (1965), p. 223.
## Results of 23 determinations, ordered in time, of the density of the earth.
##
earthden <- c(5.36, 5.29, 5.58, 5.65, 5.57, 5.53, 5.62, 5.29, 5.44, 5.34, 5.79,
5.10, 5.27, 5.39, 5.42, 5.47, 5.63, 5.34, 5.46, 5.30, 5.75, 5.68, 5.85)
runs.test(earthden)


##
## Example 2
## Sweet potato yield per acre, harvested in the United States, between 1868 and 1937.
## Data available in this package.
##
data(sweetpotato)
runs.test(sweetpotato$yield)
```

---

sweetpotato	*Sweet potato production*

Description

Sweetpotato Production, Yield per Acre and Acreage harvested in the United States, between 1868 and 1937. This data was already studied in Moore and Wallis (1941).

Usage

```
data(sweetpotato)
```

Format

A list with 70 observations on 4 vectors: `year`, `production`, `yield` and `acreage`.

Source

Agricultural Statistics 1939, p. 243.
URL: <http://archive.org/stream/agriculturalsat00unit#page/243/mode/1up>

Moore, G.H. and Wallis, W.A. (1941). A Significance Test for Time Series and Other Ordered Observations. Technical paper. *NBER*. URL: <http://papers.nber.org/books/wall41-1>

---

turning.point.test	*Turning Point Test*

Description

Performs the nonparametric Turning Point test of randomness.

Usage

```
turning.point.test(x, alternative)
```

Arguments

x	a numeric vector containing the data
alternative	a character string specifying the alternative hypothesis. Must be one of `"two.sided"` (default), `"left.sided"` or `"right.sided"`.

Details

Repeated consecutive observations are removed from data.

The possible values `"two.sided"`, `"left.sided"` and `"right.sided"` define the alternative hypothesis. By using the alternative `"two.sided"` the null hypothesis of randomness is tested against either a positive or negative serial correlation between neighbouring observations.

Value

A list with class "htest" containing the components:

statistic	the (normalized) value of the statistic test.
parameter	the size of the data, after the remotion of consecutive duplicate values.
p.value	the p-value for the test.
alternative	a character string describing the alternative hypothesis.
method	a character string indicating the test performed.
data.name	a character string giving the name of the data.
tp	the value of the TP statistic (not shown on screen).

Author(s)

Ayana Mateus and Frederico Caeiro

References

Brockwell, P.J. and Davis, R.A. (2002). *Introduction to Time Series and Forecasting*, 2nd edition, Springer (p. 36).

Mateus, A. and Caeiro, F. (2013). Comparing several tests of randomness based on the difference of observations. In T. Simos, G. Psihoyios and Ch. Tsitouras (eds.), *AIP Conf. Proc.* **1558**, 809–812.

Moore, G.H. and Wallis, W.A. (1943). Time Series Significance Tests Based on Signs of Differences. *Journal of the American Statistical Association*, **38**, 153–154.

Examples

```
##
## Example 1
##
data(sweetpotato)
turning.point.test(sweetpotato$yield)
```

Index