Package 'corTESTsrd'

November 17, 2022

Type Package

Title Significance Testing of Rank Cross-Correlations under SRD

Version 1.0-0

Date 2022-11-17

Depends R (>= 3.5.0)

Imports stats

Description Significance test of Spearman's Rho or Kendall's Tau
between short-range dependent random variables.

NeedsCompilation yes

License LGPL (>= 2.1)

Encoding UTF-8

Author David Lun [aut, cre],
Svenja Fischer [aut],
Alberto Viglione [aut]

Maintainer David Lun `<lun@hydro.tuwien.ac.at>`

RoxygenNote 7.1.2

Repository CRAN

Date/Publication 2022-11-17 16:00:05 UTC

R topics documented:

---

corTESTsrd	*Significance test of rank cross-correlations*

Description

Significance test of Spearman's Rho or Kendall's Tau between time series of short-range dependent random variables. The test is based on the asymptotic normal distributions of the estimators.

Usage

```
corTESTsrd(x, y,
           iid=TRUE, method="spearman",
           alternative="two.sided",
           kernelf=function(z){return(ifelse(abs(z) <= 1, (1 - z^2)^2, 0))},
           bwf=function(n){3*n^(1/4)})
```

Arguments

x	numeric input vector.
y	numeric input vector.
iid	logical, if TRUE, observations are assumed to be iid, if FALSE observations are assumed to be short-range dependent and the long-run variance of the estimator is estimated from the observations.
method	a character string, indicating which correlation coefficient should be used for the test. One of "spearman" or "kendall", cannot be abbreviated.
alternative	a character string indicating the alternative hypothesis. Must be one of "two.sided", "greater" or "less", cannot be abbreviated.
kernelf	a function that is used in the estimation procedure. The default kernel-function is a quartic kernel. Should be a vectorized function.
bwf	a function for choosing the bandwidth, based on the sample size n, that should be used in the estimation procedure. Default is $3n^{1/4}$, $b_n = o(n^{1/2})$ must hold.

Details

Calculates an estimate of the rank correlation coefficient between the inputs x and y, which are assumed to be evenly spaced time series with equal time-increments, and performs a significance test for the rank correlation coefficient with $\mathcal{H}_0 : \rho_S/\tau = 0$ against an alternative specified by the user. The function returns the estimate of the rank correlation coefficient and a p-value. Missing observations (NA) are allowed, but will prompt a warning. Ties are not allowed.

The test statistic and the corresponding p-value are based on the distribution of the respective estimator under the assumption of independence between the inputs x and y, and an additional assumption regarding the dependence structure of the inputs on their own past. The distribution of the test statistic is modelled as a normal distribution.

If the option iid is TRUE, the inputs are assumed to be realizations of independent and identically distributed random variables. In this case the asymptotic variance of the test statistic is given by $\frac{1}{n-1}$ for Spearman's Rho and as $\frac{2(2n+5)}{9n(n-1)}$ for Kendall's Tau, see Gibbons and Chakraborti (2003), equations 3.13 and 2.29 in chapter 11, respectively.

If the option iid is FALSE, the inputs are assumed to be realizations of short-range dependent random variables (see Corollary 1 in Lun et al., 2022). The asymptotic variance of the test statistic is modelled as $\frac{1}{n}(1+2\sum_{h=1}^{\infty}\rho_S^X(h)\rho_S^Y(h))$ for Spearman's Rho and as $\frac{4}{9n}(1+2\sum_{h=1}^{\infty}\rho_S^X(h)\rho_S^Y(h))$ for Kendall's Tau. Here $\rho_S^X(h)$ refers to the Spearman autocorrelation of the first input x for lag h, and the analogue applies to $\rho_S^X(h)$. In this case the asymptotic variance of the test statistic is estimated (see Corollary 2 in Lun et al., 2022). For this estimation procedure a kernel-function together with a bandwidth is used, which can be specified by the user.

Value

Estimate of rank correlation coefficient and p-value of corresponding hypothesis test.

References

J. D. Gibbons, and S. Chakraborti, Nonparametric statistical inference (4th Edition). CRC press, 2003.

D. Lun, S. Fischer, A. Viglione, and G. Blöschl, Significance testing of rank cross-correlations between autocorrelated time series with short-range dependence, Journal of Applied Statistics, 2022, 1-17. doi: 10.1080/02664763.2022.2137115.

Examples

```
#Demonstration
sam_size = 50
nsim = 1000

pval_iid <- rep(NA, nsim)
pval_srd <- rep(NA, nsim)
#iid-simulation: if we have iid observations the modified test
#is able to maintain the desired significance level
for(j in c(1:nsim)) {
  x <- rnorm(n=sam_size)
  y <- rnorm(n=sam_size)
  pval_iid[j] <- corTESTsrd(x, y, iid=TRUE, method="spearman")[2]
  pval_srd[j] <- corTESTsrd(x, y, iid=FALSE, method="spearman")[2]
}
sum(pval_iid <= 0.05)/nsim
sum(pval_srd <= 0.05)/nsim

#ar(1)-simulation: if we have srd-observations the modified test
#counteracts the inflation of type-I-errors
for(j in c(1:nsim)) {
  x <- as.numeric(arima.sim(model=list(ar=c(0.8)), n=sam_size))
  y <- as.numeric(arima.sim(model=list(ar=c(0.8)), n=sam_size))
  pval_iid[j] <- corTESTsrd(x, y, iid=TRUE, method="spearman")[2]
```

```
    pval_srd[j] <- corTESTsrd(x, y, iid=FALSE, method="spearman")[2]
}
sum(pval_iid <= 0.05)/nsim
sum(pval_srd <= 0.05)/nsim

#the test can be made more conservative be choosing a bigger bandwidth,
#but this decreases the power
bwfbig <- function(n) {10*(n^(1/4))}
#ar(1)-simulation: if we have srd-observations the modified test
#counteracts the inflation of type-I-errors
for(j in c(1:nsim)) {
  x <- as.numeric(arima.sim(model=list(ar=c(0.8)), n=sam_size))
  y <- as.numeric(arima.sim(model=list(ar=c(0.8)), n=sam_size))
  pval_iid[j] <- corTESTsrd(x, y, iid=TRUE, method="spearman")[2]
  pval_srd[j] <- corTESTsrd(x, y, iid=FALSE, method="spearman", bwf=bwfbig)[2]
}
sum(pval_iid<=0.05)/nsim
sum(pval_srd<=0.05)/nsim
```

Index