Package ‘tsc’

February 20, 2015

Type Package
Title Likelihood-ratio Tests for Two-Sample Comparisons
Version 1.0-3
Date 2014-11-20
Author Yang Zhao, Albert Vexler, Alan Hutson
Maintainer Yang Zhao <yzhao37@buffalo.edu>
Description Performs the two-sample comparisons using the following exact test procedures: the exact likelihood-ratio test (LRT) for equality of two normal populations proposed in Zhang et al. (2012); the combined test based on the LRT and Shapiro-Wilk test for normality via the Bonferroni correction technique; the newly proposed density-based empirical likelihood (DBEL) ratio test. To calculate p-values of the DBEL procedures, three procedures are used: (a) the traditional Monte Carlo (MC) method implemented in C++, (b) a new interpolation method based on regression techniques to operate with tabulated critical values of the test statistic; (c) a Bayesian type method that uses the tabulated critical values as the prior information and MC generated DBEL-test-statistic’s values as data.
License GPL (>= 2)
LazyData yes
NeedsCompilation yes
Repository CRAN
Date/Publication 2015-01-07 17:51:11

R topics documented:

    table ................................................................. 2
    tsc.test ............................................................. 2

Index 5
**Description**

Stores cutoff information for different target alpha values and various sets of data of varying sample size.

**Format**

data.frame with columns equal to sample size information and rows equal to different target alpha values.

**Details**

This file contains cutoff information for different target alpha values and various sets of data of varying sample size. This table is generated for sample sizes 2-30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 150, 170, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450. The target alphas range from .1 to .9 in increments of .1. The delta is 0.1.

**Note**

This dataset is used within the `tsc.test` function. There is no need for the user to ever call this dataset.

---

**tsc.test**

**Likelihood-ratio Tests for Two-Sample Comparisons**

**Description**

The functions performs the two-sample comparisons using the following exact test procedures: the exact likelihood-ratio test (LRT) for equality of two normal populations proposed in [4]; the combined test based on the LRT and Shapiro-Wilk (S-W) test for normality via the Bonferroni correction technique; the newly proposed density-based empirical likelihood (DBEL) ratio test. To calculate p-values of the DBEL procedures, three procedures are used: (a) the traditional Monte Carlo (MC) method implemented in C++, (b) a new interpolation method based on regression techniques to operate with tabulated critical values of the test statistic; (c) a Bayesian type method that uses the tabulated critical values as the prior information and MC generated DBEL-test-statistic’s values as data.

**Usage**

```
tsc.test(x,y,method="DBEL",t_m=2,mc=3000)
```
**Arguments**

- **x** numeric vector of data for x (missing values are not allowed).
- **y** numeric vector of data for y (missing values are not allowed).
- **method** a character string specifying the method for obtaining the test statistic and corresponding p-value. It must be one of "TAS", "TAS&SW" or "DBEL" (default). "TAS" indicates using the exact LRT; "TAS&SW" indicates using the combined test based on the LRT and the S-W test via the Bonferroni correction technique; "DBEL" indicates using the DBEL ratio test. Similar to the DBEL ratio test procedures can be found in [1], [2], [3].
- **t_m** indicates a method for obtaining the p-value when the "DBEL" method is used. It must have values 1, 2 (default), 3, where t_m=1 corresponds to a traditional MC method; t_m=2 corresponds to the interpolation method based on regression techniques and tabulated critical values, this method is similar to that described in [2]; t_m=3 corresponds to a Bayesian type method that combines the method t_m=1 and t_m=2 in a manner similar to that proposed in [5].
- **mc** number of monte carlo simulations used to obtain p-value when method=“DBEL" and t_m=1 (mc=3000 is default).

**Details**

The function performs the two-sample comparison using exact procedures: for the LRT to test H_0: X~N, Y~N, E(X)=E(Y), Var(X)=Var(Y) vs. H_1: X~N, Y~N, E(X) is not = E(Y), or Var(X) is not = Var(Y); for the LRT combined with the S-W test to test H_0: X~N, Y~N, E(X)=E(Y), Var(X)=Var(Y) vs. H_1: X, or Y does not follow a normal distribution, or E(X) is not = E(Y), or Var(X) is not = Var(Y); for the DBEL ratio test to test H_0: X~N, Y~N, E(X)=E(Y), Var(X)=Var(Y) vs. H_1: X, or Y does not follow a normal distribution, or E(X) is not = E(Y), or Var(X) is not = Var(Y) (Here X~N means X distributed following a normal distribution).

**Value**

Returns a vector of length 2 with a value of the test statistic and the corresponding p-value.

- **test_stat** the value of the test statistic.
- **p_value** the p-value for the test.

**Author(s)**

Yang Zhao, Albert Vexler, Alan Hutson

**References**


Examples

```r
# Ex. 1
x <- rnorm(57, 0, 1)
y <- rnorm(67, 0, 1)
## two-sample comparisons test for whether x and y are from normal distributions,
## and whether the mean of x is equal to the mean of y,
## and whether the variance of x is equal to the variance of y.
## method in [4] is used to obtain the test statistic and corresponding p-value.
test_lrt <- tsc.test(x, y, method = "TAS")
## combined method based on LRT and S-W via the Bonferroni technique
## is used to obtain the p-value.
test_comb <- tsc.test(x, y, method = "TAS & SW")
## DBEL method is used to obtain the test statistics.
## Monte Carlo method is used to obtain the p-value with 1000 Monte Carlo simulations.
test_dbel1 <- tsc.test(x, y, method = "DBEL", t_m = 1, mc = 1000)
## DBEL method is used to obtain the test statistic.
## The interpolation method based on the regression technique and tabulated critical values
## is used to obtain the p-value.
test_dbel2 <- tsc.test(x, y, method = "DBEL", t_m = 2)
## DBEL method is used to obtain the test statistic.
## The Bayesian method is used to obtain the p-value.
test_dbel3 <- tsc.test(x, y, method = "DBEL", t_m = 3)
```

```r
# Ex. 2
A <- rnorm(15, 0, 1)
B <- runif(31, -1, 1)
test_lrt1 <- tsc.test(A, B, method = "TAS")  # p-value is 0.3656844.
test_comb1 <- tsc.test(A, B, method = "TAS & SW")  # p-value is 0.02588757.
test_dbel14 <- tsc.test(A, B, method = "DBEL", t_m = 1, mc = 1000)  # p-value is 0.001.
test_dbel15 <- tsc.test(A, B, method = "DBEL", t_m = 2)  # p-value is 0.001774751.
test_dbel16 <- tsc.test(A, B, method = "DBEL", t_m = 3)  # p-value is 0.000112455.
```

## B is not from the normal distribution, so the null hypothesis should be rejected.
## The LRT method does not reject $H_0$, since this method works just for $X \sim N$ and $Y \sim N$. 


Index

*Topic datasets
  table, 2

*Topic methods
  tsc.test, 2

cutval (table), 2

table, 2

tsc.test, 2