Package 'dbEmpLikeNorm'

February 19, 2015

Version 1.0.0

Date Apr 12, 2013

Title Test for joint assessment of normality

Author Lori A. Shepherd <las65@buffalo.edu>, Wan-Min Tsai

<wan-min.tsai@yale.edu>, Albert Vexler <avexler@buffalo.edu>,
Jeffrey C. Miecznikowski <jcm38@buffalo.edu>

Maintainer Lori A. Shepherd <las65@buffalo.edu>

Depends R (>= 2.10), dbEmpLikeGOF

Description Test for joint assessment of normality

LazyData no

License GPL (>= 2)

URL

NeedsCompilation no

Repository CRAN

Date/Publication 2013-04-25 16:50:23

R topics documented:

dbEmpLikeNorm-package *Empirical Likelihood Joint Assessment of Normality*

Description

Package that has functions to perform a joint assessment of normality across k groups

Author(s)

Lori A. Shepherd, Wan-Min Tsai, Albert Vexler, Jeffrey C. Miecznikowski

References

Vexler A, Gurevich G, Empirical likelihood ratios applied to goodness-of-fit tests based on sample entropy. Computational Statistics and Data Analysis 54(2010) 531-545.

Gurevich G, Vexler A, A two-sample empirical likelihood ratio test based on samples entropy. Statistics and Computing, 2011.

Miecznikowski J,Vexler A, Shepherd L, dbEmpLikeGOF: An R Package for Nonparametric Likelihood Ratio Tests for Goodness-of-Fit and Two Sample Comparisons Based on Sample Entropy.Journal of Statistical Software 2013 (Accepted; to appear)

Tsai WM, Shepherd LA, Miecznikowski J, Hutson A, Vexler A. (2013). An EL based test for normality in multiple groups. Department of Biostatistics. University at Buffalo. Report 1204.

See Also

`dbELnorm`, `returnCutoffValue`

datamat *Pvalue Data Tables*

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 (Type I error) values and various sets of data of varying sample size. Note: *twoGroup* is for data consisting of two group, *threeGroup* is for data consisting of three groups. These tables are generated for sample sizes 10, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, and 300. The target alphas range from .001 to .999 in increments of .001. The default value for delta is 0.5, See [Tsai 2013] for details on setting delta.

Note

This dataset is used within the functions. There is no need for the user to call this dataset.

References

Tsai WM, Shepherd LA, Miecznikowski J, Hutson A, Vexler A. (2013). An EL based test for normality in multiple groups. Department of Biostatistics. University at Buffalo. Report 1204.

`dbELnorm` *Goodness of Fit Test for Normality Among Multiple Groups*

Description

Performs density based empirical likelihood goodness of fit tests for normality among multiple groups

Usage

```
dbELnorm(x,
         delta=0.05,
         num.mc=1000,
         pvl.Table=TRUE,
         vrb=TRUE)
```

Arguments

`x`	list of groups or experiments
`delta`	an option for changing the minimizing range for the EL ratio test statistic
`num.mc`	number of simulations to use when calculating p-value
`pvl.Table`	logical indicating if p-value should be calculated based on estimates from stored data tables or by using Monte Carlo techniques
`vrb`	logical indicating if status messages should be printed

Details

The method employs a density-based empirical likelihood approach to obtain the test statistic and p-values for a goodness-of-fit tests for normality. The null distribution is that the data in *x* is normally distributed with possibly different means and standard deviations.

'x' is a list object where each item in the list is a different set of data.

The 'delta' value must be in the range [0,1]. Essentially this setting controls the range over which a minimum is taken to produce the EL ratio test statistic. The range is from 1 to n^(1-'delta') where 'n' represents the number of observations in 'x'.

The 'pvl.Table' is a binary option where when TRUE, the p-value for the test statistic is determined by imputation from a stored table of test statistics and significance levels for common sample sizes. If 'pvl.Table' is FALSE, then the p-value is determined from Monte-Carlo simulations where the number of resamplings is set by 'num.mc'.

Value

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

teststat	the value of the test statistic
pvalue	the p-value for the test

Author(s)

Lori A. Shepherd, Wan-Min Tsai, Albert Vexler, Jeffrey C. Miecznikowski

References

Tsai WM, Shepherd LA, Miecznikowski J, Hutson A, Vexler A. (2013). An EL based test for normality in multiple groups. Department of Biostatistics. University at Buffalo. Report 1204.

Examples

```
x = rnorm(30, 3,1)
y = rnorm(40, 4,1)

Lst = list(x,y)
dbELnorm(Lst)

y = runif(40)
Lst = list(x,y)
dbELnorm(Lst)
```

---

returnCutoffValue *Estimates The Statistic Cutoff For A Target Alpha*

Description

estimates the test statistic cutoff for significance

Usage

```
returnCutoffValue(numberOfgroups,
                  sample.size,
                  targetalpha=0.05,
                  MC.Method=TRUE,
                  Table.Method=FALSE,
                  Bayes.Method=FALSE,
                  num.mc=1000,
                  delta=0.05,
                  nsims=200,
                  v.threshold=NA)
```

Arguments

numberOfgroups	number of different groups or experiments
sample.size	number of observations
targetalpha	The significance level for the test.
MC.Method	logical indicating if value should be calculated based on Monte Carlo techniques
Table.Method	logical indicating if value should be calculated based on estimates from generated data table
Bayes.Method	logical indicating if value should be calculated using a Bayesian method incorporating elements of MC.Method and Table.Method
num.mc	number of simulations to estimate distribution of statistic in MC.Method
delta	an option for changing the minimizing range for the EL ratio test statistic for the distribution. Utilized in MC.Method
nsims	The number of simulations to generate and investigate in each turn of Bayesian approach
v.threshold	a numeric threshold for the variance. This threshold must be met to accept calculated value of Bayesian approach. If NA, a variance estimate is calculated and used as threshold.

Details

This function is designed to return the cut-off for significance for the statistics obtained from the density-based EL tests. The significance level for the associated cutoffs are specified by the user in 'targetalpha'.

The 'numberOfgroups' is a scalar denoting the number of groups or datasets being tested. The 'sample.size' should be a vector of length equal to the 'numberOfgroups' where sample.size[1] is the number of observations for group 1, sample.size[2] is the number of observations for group 2, etc. If only a single 'sample.size' is specified, it is assumed groups are of equal length.

MC.Method, Table.Method, and Bayes.Method are binary options. When MC.Method is TRUE, the cutoff is determined from a Monte-Carlo simulation where the number of resamplings is controlled by 'num.mc'. When Table.Method is TRUE, the cutoff is determined by imputation from a stored table of test statistics and significance levels for common sample sizes. When Bayes.Method is TRUE, the cutoff is determined through a Bayesian approach where the number of additional observations is controlled by nsims, and the threshold for acceptance is controlled by 'v.threshold'. See [Tsai 2013] for more details on the algorithm

The 'delta' value must be in the range [0,1]. Essentially this setting controls the range over which a minimum is taken to produce the EL ratio test statistic. The range is from 1 to n^(1-'delta') where 'n' represents the number of observations in 'x'.

Value

Returns a statistical cutoff value to assess significance at level 'targetalpha'. If more than one method is selected, a list with value for each method is returned. If only one method is selected, a single numeric value for that method is returned.

Author(s)

Lori A. Shepherd, Wan-Min Tsai, Albert Vexler, Jeffrey C. Miecznikowski

References

Tsai WM, Shepherd LA, Miecznikowski J, Hutson A, Vexler A. (2013). An EL based test for normality in multiple groups. Department of Biostatistics. University at Buffalo. Report 1204.

Examples

```
returnCutoffValue(3, c(10,15,40), MC.Method=TRUE)

returnCutoffValue(3, c(10,15,40), MC.Method=TRUE, Bayes.Method=TRUE, Table.Method=TRUE)
```

Index