Package ‘MVN’

March 18, 2015

Type Package
Title Multivariate Normality Tests
Version 3.9
Date 2015-03-18
Author Selcuk Korkmaz [aut, cre], Dincer Goksuluk [aut], Gokmen Zararsiz [aut]
Maintainer Selcuk Korkmaz <selcuk.korkmaz@hacettepe.edu.tr>
Depends R (>= 2.15.0)
VignetteBuilder knitr
Suggests knitr
Imports methods, nortest, moments, MASS, robustbase, mvoutlier, plyr, psych

Description Performs multivariate normality tests and graphical approaches and implements multivariate outlier detection and univariate normality of marginal distributions through plots and tests.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN
Date/Publication 2015-03-18 18:01:27

R topics documented:

MVN-package ................................................................. 2
hz-class ................................................................. 2
hzTest ................................................................. 3
mardia-class .......................................................... 4
mardiaTest ........................................................... 5
mvnPlot ................................................................. 6
Description

Performs multivariate normality tests and graphical approaches and implements multivariate outlier detection and univariate normality of marginal distributions through plots and tests.

Details

Package: MVN
Type: Package
License: GPL (>= 2)

Author(s)

Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz
Maintainer: Selcuk Korkmaz <selcuk.korkmaz@hacettepe.edu.tr>

Class "hz"

Description

An S4 class for Henze-Zirkler’s Multivariate Normality Test

Slots

HZ: stores the value of Henze-Zirkler statistic
p.value: stores the p-value for the HZ test
dname: stores the data set name
dataframe: stores the data set
hzTest

**Author(s)**
Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

---

**hzTest**  
*Henze-Zirkler’s Multivariate Normality Test*

**Description**
This function performs Henze-Zirkler’s Multivariate Normality Test.

**Usage**
```haskell
hzTest(data, cov = TRUE, qqplot = FALSE)
```

**Arguments**
- `data`: a numeric matrix or data frame
- `cov`: if `TRUE` covariance matrix is normalized by \( n \), if `FALSE` it is normalized by \( n^{-1} \)
- `qqplot`: if `TRUE` it creates a chi-square Q-Q plot

**Details**
The Henze-Zirkler test is based on a non-negative functional distance that measures the distance between two distribution functions. If the data is multivariate normal, the test statistic \( HZ \) is approximately lognormally distributed. It proceeds to calculate the mean, variance and smoothness parameter. Then, mean and variance are lognormalized and the p-value is estimated.

**Value**
- `HZ`: the value of Henze-Zirkler statistic at significance level \( 0.05 \)
- `p-value`: a p-value for the \( HZ \) test

**Author(s)**
Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

**References**


See Also

roystonTest mardiaTest mvnPlot mvOutlier uniPlot uniNorm

Examples

setosa = iris[1:50, 1:4]  # Iris data only for setosa and four variables
result = hzTest(setosa, qqplot = TRUE)
result

mardia-class Class "mardia"

Description

An S4 class for Mardia’s Multivariate Normality Test

Slots

g1p: stores the Mardia’s multivariate skewness statistic
chi.skew: stores the chi-square value of the skewness statistic
p.value.skew: stores the p-value of skewness statistic
g2p: stores the Mardia’s multivariate kurtosis statistic
z.kurtosis: stores the z value of kurtosis statistic
p.value.kurt: stores the p-value of kurtosis statistic
p.value.skew: stores the p-value of skewness statistic
chi.small.skew: stores the chi-square value of the small sample skewness statistic
dname: stores the data set name
dataframe: stores the data set

Author(s)

Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz
**mardiaTest**  

**Mardia’s Multivariate Normality Test**

**Description**

This function performs Mardia’s Multivariate Normality Test.

**Usage**

```r
mardiaTest(data, cov = TRUE, qqplot = FALSE)
```

**Arguments**

- `data` : a numeric matrix or data frame
- `cov` : if TRUE covariance matrix is normalized by n, if FALSE it is normalized by n-1
- `qqplot` : if TRUE it creates a chi-square Q-Q plot

**Details**

This function calculate the Mardia’s multivariate skewness and kurtosis coefficients as well as their corresponding statistical significance. It can also calculate corrected version of skewness coefficient for small sample size (n< 20).

For multivariate normality, both p-values of skewness and kurtosis statistics should be greater than 0.05.

If sample size less than 20 then `p.value.skew.small` should be used as significance value of skewness instead of `p.value.skew`.

**Value**

- `g1p` : Mardia's multivariate skewness statistic
- `chi.skew` : Chi-square value of the skewness statistic
- `p.value.skew` : p-value of the skewness statistic
- `g2p` : Mardia's multivariate kurtosis statistic
- `z.kurtosis` : z value of the kurtosis statistic
- `p.value.kurt` : p-value of kurtosis statistic
- `chi.skew.small` : Chi-square value of the small sample skewness statistic
- `p.value.skew.small` : p-value of small sample skew statistic

**Author(s)**

Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz
References


See Also

roystonTest hzTest mvdPlot mvoutlier uniPlot uniNorm

Examples

setosa = iris[1:50, 1:4] # Iris data only for setosa and four variables
result = mardiaTest(setosa, qqplot = TRUE)
result

---

mvnPlot

Perspective and Contour Plots

Description

This function creates perspective and contour plots for a bivariate data set.

Usage

mvnPlot(object, type = c("persp", "contour"), default = TRUE, ...)

Arguments

object an object of mardia, hz or royston class

type if type is selected as persp it creates a perspective plot, if type is selected as contour it creates a contour plot.

default when default is TRUE it creates plots in default settings

Details

After set the default=FALSE option, users can define their own plot settings with changing theta, phi, border and shade.
mvOutlier

Note
Please be careful that this function creates perspective and contour plots when there are only two
variables.

Author(s)
Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

See Also
roystonTest hzTest mardiaTest mvOutlier uniPlot uniNorm

Examples

```r
setosa = iris[1:50, 1:2] # Iris data only for setosa and two variables
result = hzTest(setosa)
### Perspective Plot ###
mvnPlot(result, type = "persp", default = TRUE)
### Contour Plot ###
mvnPlot(result, type = "contour", default = TRUE)
```

mvOutlier

Multivariate Outlier Detection

Description
This function detects multivariate outliers based on Mahalanobis distance and adjusted Mahalanobis
distance.

Usage

```r
mvOutlier(data, qqplot = TRUE, alpha = 0.5, method = c("quan", "adj.quan"))
```

Arguments

- `data`: a numeric matrix or data frame
- `qqplot`: if TRUE it creates a chi-square Q-Q plot
- `alpha`: a numeric parameter controlling the size of the subsets over which the determin-
ant is minimized. Allowed values for the alpha are between 0.5 and 1 and the
default is 0.5.
- `method`: quan for Mahalanobis distance and adj.quan for adjusted Mahalanobis dis-
tance.

Value

- `outlier`: an outlier set
- `newData`: new data set without possible outliers
Author(s)
Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

See Also
mardiaTest roystonTest hzTest mvnPlot uniPlot uniNorm

Examples
setosa = iris[1:50, 1:3]  # Iris data only for setosa and three variables
result = mvOutlier(setosa, qqplot = TRUE, method = "quan")
result

royston-class  
Class "royston"

Description
An S4 class for Royston’s Multivariate Normality Test

Slots
- h: stores the the value of Royston’s H statistic
- p.value: stores the p-value for the Royston test
- dname: stores the data set name
- dataframe: stores the data set

Author(s)
Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

roystonTest  
Royston’s Multivariate Normality Test

Description
This function performs Royston’s Multivariate Normality Test.

Usage
roystonTest(data, qqplot = FALSE)

Arguments
- data: a numeric matrix or data frame
- qqplot: if TRUE it creates a chi-square Q-Q plot
Details

A function to generate the Shapiro-Wilk’s W statistic needed to feed the Royston’s H test for multivariate normality. However, if kurtosis of the data greater than 3 then Shapiro-Francia test is used for leptokurtic samples else Shapiro-Wilk test is used for platykurtic samples.

Value

H  the value of Royston’s H statistic at significance level 0.05

p-value  an approximate p-value for the test with respect to equivalent degrees of freedom (edf)

Author(s)

Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

References


See Also

hzTest mardiaTest mvnPlot mvOutlier uniPlot uniNorm

Examples

setosa = iris[1:50, 1:4] # Iris data only for setosa and four variables
result = roystonTest(setosa, qqplot = TRUE)
result
uniNorm  

Univariate Normality Tests

Description

This function performs univariate normality tests, including Shapiro-Wilk, Cramer-von Mises, Lilliefors (Kolmogorov-Smirnov), Shapiro-Francia and Anderson-Darling.

Usage

uniNorm(data, type = c("sw", "cvm", "lillie", "sf", "ad"), desc = TRUE)

Arguments

data: a vector, data frame or matrix

type: select one of the univariate normality tests: sw: Shapiro-Wilk, cvm: Cramer-von Mises, lillie: Lilliefors (Kolmogorov-Smirnov), sf: Shapiro-Francia, ad: Anderson-Darling

desc: if TRUE, it displays descriptive statistics including mean, standard deviation, median, minimum, maximum, 25th and 75th percentiles, skewness and kurtosis.

Details

sw is default.

Author(s)

Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

See Also

mardiaTest roystonTest hzTest mvnPlot mvOutlier uniPlot

Examples

setosa = iris[1:50, 1:4] # Iris data only for setosa and four variables
uniNorm(setosa, type = "sw", desc = TRUE)
uniPlot  

Uniivariate Plots

Description
This function creates univariate plots, including Q-Q plot, histogram, box-plot and scatterplot matrices.

Usage
uniPlot(data, type = c("qqplot", "histogram", "box", "scatter"), mfrow = NULL, ...)

Arguments
- data: a vector, data frame or matrix
- type: select one of the univariate plots: qqplot: Q-Q plot, histogram: histogram with a normal curve, box: box-plot and scatter: scatterplot matrix
- mfrow: multi-paneled plotting window
- ...: optional arguments

Details
Box-Plots are based on standardized values, variables are centered and scaled before plotting. qqplot is default.

Author(s)
Selcuk Korkmaz, Dincer Goksuluk, Gokmen Zararsiz

See Also
mardiaTest roystonTest hzTest mvnPlot mvOutlier uniNorm

Examples
setosa = iris[1:50, 1:4] # Iris data only for setosa and four variables
uniPlot(setosa, type = "qqplot")
Index

hz (hz-class), 2
hz-class, 2
hzTest, 3, 6–11

mardia (mardia-class), 4
mardia-class, 4
mardiaTest, 4, 5, 7–11
MVN (MVN-package), 2
MVN-package, 2
mvnPlot, 4, 6, 6, 8–11
mvOutlier, 4, 6, 7, 7, 9–11

royston (royston-class), 8
royston-class, 8
roystonTest, 4, 6–8, 8, 10, 11

uniNorm, 4, 6–9, 10, 11
uniPlot, 4, 6–10, 11