Package ‘MVTests’

November 3, 2023

Type Package
Date 2023-11-03
Title Multivariate Hypothesis Tests
Version 2.2.2
Author Hasan BULUT [aut,cre]
Maintainer Hasan BULUT <hasan.bulut@omu.edu.tr>
Description Multivariate hypothesis tests and the confidence intervals. It can be used to test the hypotheses about mean vector or vectors (one-sample, two independent samples, paired samples), covariance matrix (one or more matrices), and the correlation matrix. Moreover, it can be used for robust Hotelling $T^2$ test at one sample case in high dimensional data. For this package, we have benefited from the studies Rencher (2003), Nel and Merwe (1986) <DOI:10.1080/03610928608829342>, Tatlidil (1996), Tsagris (2014), Vil lasenor Alva and Estrada (2009) <DOI:10.1080/03610920802474465>.
License GPL-2
Depends R (>= 4.0)
Imports stats, rrcov, mvtnorm, robustbase
Encoding UTF-8
LazyData true
RoxygenNote 7.2.1
NeedsCompilation no
Repository CRAN
Date/Publication 2023-11-03 16:40:05 UTC

R topics documented:

Bcov ................................................................. 2
BoxM ............................................................. 3
Bspe ............................................................. 4
Coated .......................................................... 5
iris ............................................................... 5
Bcov function tests whether the covariance matrix is equal to a given matrix or not.

**Usage**

Bcov(data, Sigma)

**Arguments**

- **data**: a data frame.
- **Sigma**: The covariance matrix in NULL hypothesis.

**Details**

This function computes Bartlett’s test statistic for the covariance matrix of one sample.

**Value**

A list with 3 elements:

- **ChiSquare**: The value of Test Statistic
- **df**: The Chi-Square statistic’s degree of freedom
- **p.value**: p value

**Author(s)**

Hasan BULUT <hasan.bulut@omu.edu.tr>

**References**

**Examples**

```r
data(iris)
S<-matrix(c(5.71,-0.8,-0.6,-0.5,-0.8,4.09,-0.74,-0.54,-0.6,
           -0.74,7.38,-0.18,-0.5,-0.54,-0.18,8.33),ncol=4,nrow=4)
result <- Bcov(data=iris[,1:4],Sigma=S)
summary(result)
```

---

**BoxM**

**Box's M Test**

**Description**

BoxM function tests whether the covariance matrices of independent samples are equal or not.

**Usage**

```r
BoxM(data, group)
```

**Arguments**

- `data`: a data frame.
- `group`: grouping vector.

**Details**

This function computes Box-M test statistic for the covariance matrices of independent samples. The hypotheses are defined as H0: The Covariance matrices are homogeneous and H1: The Covariance matrices are not homogeneous.

**Value**

A list with 3 elements:

- `ChiSquare`: The value of Test Statistic.
- `df`: The Chi-Square statistic’s degree of freedom.
- `p.value`: p value.

**Author(s)**

Hasan BULUT <hasan.bulut@omu.edu.tr>

**References**

Examples

data(iris)
results <- Bsper(data=iris[,1:4],group=iris[,5])
summary(results)

Bsper

Bartlett’s Sphericity Test

Description

Bsper function tests whether a correlation matrix is equal to the identity matrix or not.

Usage

Bsper(data)

Arguments

data a data frame.

Details

This function computes Bartlett’s test statistic for Sphericity Test. The hypotheses are H0: R is equal to I and H1: R is not equal to I.

Value

a list with 4 elements:

ChiSquare The value of Test Statistic
df The Chi-Square statistic’s degree of freedom
p.value p value
R Correlation matrix

Author(s)

Hasan BULUT <hasan.bulut@omu.edu.tr>

References


Examples

data(iris)
results <- Bsper(data=iris[,1:4])
summary(results)
**Coated**

**Description**

The data set is given in Table 5.3 in Rencher (2003). The data set consists of 2 variables (Depth and Number), 2 treatments and 15 observations. The first column of the data is Location numbers.

**Usage**

Coated

**Format**

A data frame with 15 rows and 5 columns. The columns are as follows:

- **Location** The location numbers of observations.
- **Coating1.Depth1** The Depth values in the first treatment
- **Coating1.Number1** The Number values in the first treatment
- **Coating2.Depth2** The Depth values in the second treatment
- **Coating2.Number2** The Number values in the second treatment

**Source**

The data set is used in the book entitled Methods of Multivariate Analysis (Rencher, 2003).

**References**


---

**Iris**

**Iris Data**

**Description**

The Iris dataset is consists of 4 variables, 3 groups and 150 observations. The last column of the data is Iris species.

**Usage**

iris
Format
A data frame with 150 rows and 5 columns. The columns are as follows:

- **Sepal.Length** The Sepal length values of iris flowers
- **Sepal.Width** The Sepal width values of iris flowers
- **Petal.Length** The Petal length values of iris flowers
- **Petal.Width** The Petal width values of iris flowers
- **Species** The species of iris flowers

Source
https://archive.ics.uci.edu/ml/datasets/Iris

Mpaired

Multivariate Paired Test

Description
Mpaired function computes the value of test statistic based on Hotelling T Square approach in multivariate paired data sets.

Usage
Mpaired(T1, T2)

Arguments
- **T1** The first treatment data.
- **T2** The second treatment data.

Details
This function computes one sample Hotelling T^2 statistics for paired data sets.

Value
a list with 7 elements:
- **HT2** The value of Hotelling T^2 Test Statistic
- **F** The value of F Statistic
- **df** The F statistic’s degree of freedom
- **p.value** p value
- **Descriptive1** The descriptive statistics of the first treatment
- **Descriptive2** The descriptive statistics of the second treatment
- **Descriptive.Difference** The descriptive statistics of the differences
Author(s)
Hasan BULUT <hasan.bulut@omu.edu.tr>

References

Examples

data(Coated)
X<-Coated[,2:3]; Y<-Coated[,4:5]
result <- Mpaired(T1=X,T2=Y)
summary(result)

OneSampleHT2

One Sample Hotelling T^2 Test

Description
OneSampleHT2 computes one sample Hotelling T^2 statistics and gives confidence intervals

Usage
OneSampleHT2(data, mu0, alpha = 0.05)

Arguments
data
  a data frame.
mu0
  mean vector that is used to test whether population mean parameter is equal to it.
alpha
  Significance Level that will be used for confidence intervals. default alpha=0.05.

Details
This function computes one sample Hotelling T^2 statistics that is used to test whether population mean vector is equal to a vector given by a user. When H0 is rejected, this function computes confidence intervals for all variables.

Value
a list with 7 elements:
HT2
  The value of Hotelling T^2 Test Statistic
F
  The value of F Statistic
df
  The F statistic’s degree of freedom
p.value
  p value
The lower and upper limits of confidence intervals obtained for all variables

alpha  The alpha value using in confidence intervals

Descriptive  Descriptive Statistics

Author(s)

Hasan BULUT <hasan.bulut@omu.edu.tr>

References


Examples

data(iris)

mean0<-c(6,3,1,0.25)
result <- OneSampleHT2(data=iris[1:50,-5],mu0=mean0,alpha=0.05)
summary(result)

RHT2  Robust Hotelling T^2 Test for One Sample in High Dimensional Data

Description

Robust Hotelling T^2 Test for One Sample in high Dimensional Data

Usage

RHT2(data, mu0, alpha = 0.75, d, q)

Arguments

data  the data. It must be matrix or data.frame.
mu0  the mean vector which will be used to test the null hypothesis.
alpha  numeric parameter controlling the size of the subsets over which the determinant is minimized. Allowed values are between 0.5 and 1 and the default is 0.75.
d  the constant in Equation (11) in the study by Bulut (2021).
q  the second degree of freedom value of the approximate F distribution in Equation (11) in the study by Bulut (2021).

Details

RHT2 function performs a robust Hotelling T^2 test in high dimensional test based on the minimum regularized covariance determinant estimators. This function needs the q and d values. These values can be obtained simRHT2 function. For more detailed information, you can see the study by Bulut (2021).
RobCat

Value

- a list with 3 elements:
  - T2: The Robust Hotelling $T^2$ value in high dimensional data
  - Fval: The F value based on T2
  - pval: The p value based on the approximate F distribution

Author(s)

Hasan BULUT <hasan.bulut@omu.edu.tr>

References


Examples

```r
library(rrcov)
data(octane)
mu.clean<-colMeans(octane[-c(25,26,36,37,38,39),])

RHT2(data=octane,mu0=mu.clean,alpha=0.84,d=1396.59,q=1132.99)
```

RobCat

### Robust CAT Algorithm

Description

RobCat computes p value based on robust CAT algorithm to compare two means vectors under multivariate Behrens-Fisher problem.

Usage

```r
RobCat(X, Y, M = 1000, alpha = 0.75)
```

Arguments

- **X**: a matrix or data frame for first group.
- **Y**: a matrix or data frame for second group.
- **M**: iteration number and the default is 1000.
- **alpha**: numeric parameter controlling the size of the subsets over which the determinant is minimized; roughly alpha*n, observations are used for computing the determinant. Allowed values are between 0.5 and 1 and the default is 0.75.
Details

This function computes the p-value based on robust CAT algorithm to compare two means vectors under multivariate Behrens-Fisher problem. When p-value < 0.05, it means the difference of two mean vectors is significant statistically.

Value

- a list with 2 elements:
  - `Cstat`: Calculated value of test statistic
  - `pval`: The p-value

Author(s)

Hasan BULUT <hasan.bulut@omu.edu.tr>

Examples

```r
data(iris)
RobCat(X=iris[1:20,-5],Y=iris[81:100,-5])
```

---

**RperT2**  
*Robust Permutation Hotelling $T^2$ Test in High Dimensional Data*

Description

Robust Permutation Hotelling $T^2$ Test for Two Independent Samples in high Dimensional Data

Usage

```
RperT2(X1, X2, alpha = 0.75, N = 100)
```

Arguments

- **X1**: the data matrix for the first group. It must be matrix or data.frame.
- **X2**: the data matrix for the first group. It must be matrix or data.frame.
- **alpha**: numeric parameter controlling the size of the subsets over which the determinant is minimized. Allowed values are between 0.5 and 1 and the default is 0.75.
- **N**: the permutation number

Details

`RperT2` function performs a robust permutation Hotelling $T^2$ test for two independent samples in high dimensional test based on the minimum regularized covariance determinant estimators.
Value

a list with 2 elements:

T2 The calculated value of Robust Hotelling T^2 statistic based on MRCD estimations
p.value p value obtained from test process

Author(s)

Hasan BULUT <hasan.bulut@omu.edu.tr>

References

Bulut, H (2023). A robust Hotelling test statistic for two samples case in high dimensional data. (Unpublished)

Examples

library(rrcov)
x<-mvtnorm::rmvnorm(n=10,sigma=diag(20),mean=rep(0,20))
y<-mvtnorm::rmvnorm(n=10,sigma=diag(20),mean=rep(1,20))
RperT2(X1=x,X2=y)$p.value

simRHT2

 Monte Carlo Simulation to obtain d and q constants for RHT2 function

Description

Monte Carlo Simulation to obtain d and q constants for RHT2 function

Usage

simRHT2(n, p, nrep = 500)

Arguments

n the sample size
p the number of variables
nrep the number of iteration. The default value is 500.

Details

simRHT2 function computes d and q constants to construct an approximate F distribution of robust Hotelling T^2 statistic in high dimensional data. These constants are used in RHT2 function. For more detailed information, you can see the study by Bulut (2021).
Value
  a list with 2 elements:
  q The q value
  d The d value

Author(s)
  Hasan BULUT <hasan.bulut@omu.edu.tr>

References

summary.MVTests
  Summarizing Results in MVTests Package

Description
  summary.MVTests function summarizes of results of functions in this package.

Usage
  ## S3 method for class 'MVTests'
  summary(object, ...)

Arguments
  object an object of class MVTests.
  ... additional parameters.

Details
  This function prints a summary of the results of multivariate hypothesis tests in the MVTests package.

Value
  the input object is returned silently.

Author(s)
  Hasan BULUT <hasan.bulut@omu.edu.tr>
Examples

# One Sample Hotelling T Square Test
data(iris)
X<-iris[1:50,1:4]
mean0<-c(6,3,1,0.25)
result.onesample <- OneSampleHT2(data=X,mu0=mean0, alpha=0.05)
summary(result.onesample)

# Two Independent Sample Hotelling T Square Test
data(iris)
G<-c(rep(1,50),rep(2,50))
result.twosamples <- TwoSamplesHT2(data=iris[1:100,1:4], group=G, alpha=0.05)
summary(result.twosamples)

# Box's M Test
data(iris)
result.BoxM <- BoxM(data=iris[,1:4], group=iris[,5])
summary(result.BoxM)

# Bartlett's Test of Sphericity
data(iris)
result.Bsper <- Bsper(data=iris[,1:4])
summary(result.Bsper)

# Bartlett's Test for One Sample Covariance Matrix
data(iris)
S<-matrix(c(5.71,-0.8,-0.6,-0.5,-0.8,4.09,-0.74,-0.54,-0.6,-0.74,
7.38,-0.18,-0.5,-0.54,-0.18,8.33),ncol=4,nrow=4)
result.bcov<- Bcov(data=iris[,1:4], Sigma=S)
summary(result.bcov)

---

TR2

Robust Hotelling T^2 Test Statistic

Description

Robust Hotelling T^2 Test Statistic for Two Independent Samples in high Dimensional Data

Usage

TR2(x1, x2, alpha = 0.75)

Arguments

x1 the data matrix for the first group. It must be matrix or data.frame.
x2 the data matrix for the first group. It must be matrix or data.frame.
alpha numeric parameter controlling the size of the subsets over which the determinant is minimized. Allowed values are between 0.5 and 1 and the default is 0.75.
TwoSamplesHT2

Details

TR2 function calculates the robust Hotelling $T^2$ test statistic for two independent samples in high dimensional data based on the minimum regularized covariance determinant estimators.

Value

a list with 2 elements:

TR2 The calculated value of Robust Hotelling $T^2$ statistic based on MRCD estimations

Author(s)

Hasan BULUT <hasan.bulut@omu.edu.tr>

References

Bulut, H (2023). A robust Hotelling test statistic for two samples case in high dimensional data. (Unpublished)

Examples

library(rrcov)
x<-mvtnorm::rmvnorm(n=10,sigma=diag(20),mean=rep(0,20))
y<-mvtnorm::rmvnorm(n=10,sigma=diag(20),mean=rep(1,20))
TR2(x1=x,x2=y)

TwoSamplesHT2 Two Independent Samples Hotelling $T^2$ Test

Description

TwoSamplesHT2 function computes Hotelling $T^2$ statistic for two independent samples and gives confidence intervals.

Usage

TwoSamplesHT2(data, group, alpha = 0.05, Homogenity = TRUE)

Arguments

data a data frame.
group a group vector consisting of 1 and 2 values.
alpha Significance Level that will be used for confidence intervals. default=0.05
Homogenity a logical argument. If sample covariance matrices are homogeneity, then Homogenity=TRUE. Otherwise Homogenity=FALSE The homogeneity of covariance matrices can be investigated with BoxM function.
Details

This function computes two independent samples Hotelling T^2 statistics that is used to test whether two population mean vectors are equal to each other. When \( H_0 \) is rejected, this function computes confidence intervals for all variables to determine variable(s) affecting on rejection decision. Moreover, when covariance matrices are not homogeneity, the approach proposed by D. G. Nel and V. D. Merwe (1986) is used.

Value

a list with 8 elements:

- \( HT2 \) The value of Hotelling T^2 Test Statistic
- \( F \) The value of F Statistic
- \( df \) The F statistic’s degree of freedom
- \( p.value \) p value
- \( CI \) The lower and upper limits of confidence intervals obtained for all variables
- \( alpha \) The alpha value using in confidence intervals
- \( Descriptive1 \) Descriptive Statistics for the first group
- \( Descriptive2 \) Descriptive Statistics for the second group

Author(s)

Hasan BULUT <hasan.bulut@omu.edu.tr>

References

D.G. Nel & C.A. Van Der Merwe (1986) A solution to the multivariate behrens fisher problem, Communications in Statistics: Theory and Methods, 15:12, 3719-3735

Examples

data(iris)
G<-c(rep(1,50),rep(2,50))
# When covariances matrices are homogeneity
results1 <- TwoSamplesHT2(data=iris[1:100,1:4],group=G,alpha=0.05)
summary(results1)
# When covariances matrices are not homogeneity
results2 <- TwoSamplesHT2(data=iris[1:100,1:4],group=G,Homogenity=FALSE)
summary(results2)
Index

* datasets
  Coated, 5
  iris, 5
Bcov, 2
BoxM, 3
Bsper, 4
Coated, 5
iris, 5
Mpaired, 6
OneSampleHT2, 7
RHT2, 8
RobCat, 9
RperT2, 10
simRHT2, 11
summary.MVTests, 12
TR2, 13
TwoSamplesHT2, 14