

Package ‘CvM2SL2Test’

February 14, 2012

Version 2.0-1

Date 2012-01-02

Title Cramer-von Mises Two Sample Tests

Author Yuanhui Xiao <yxiao@gsu.edu>

Maintainer Yuanhui Xiao <yxiao@gsu.edu>

Depends

Suggests

Description The package contains two functions: one computes the Cramer-von Mises two sample test scores, the other computes the exact p-value(s) for given Cramer-von Mises two-sample test score(s) under the assumption that the populations under comparison have the same probability distribution. Like Kolmogorov-Smirnov two-sample test, Cramer-von Mises test is also distribution free. In certain cases, the Cramer-von Mises test is more powerful than the Kolmogorov-Smirnov test, but it is less widely used than the latter.

License GPL-2

Repository CRAN

Date/Publication 2012-01-03 06:32:57

R topics documented:

cvmts.pval	2
cvmts.test	4

Index	5
--------------	----------

`cvmts.pval`*Computing Exact P-value for Cramer-von Mises Two Sample Tests*

Description

The Cramer-von Mises two sample test is to test whether two independent samples were drawn from the same population. The function `cvmts.pval()` evaluates the p-value for given test statistic.

Usage

```
cvmts.pval(cvmstats, m, n)
```

Arguments

<code>cvmstats</code>	an R object holding a list of computed Cramer-von Mises test scores.
<code>m</code>	sample size of the first sample.
<code>n</code>	sample size of the second sample.

Value

The returned value the p-value(s)

Note

The function `cvmts.pval()` first construct the distribution function, which is represented as sequence of tables in my implementation. This step may be slow, depending the sample sizes `n` and `m` as well as the capacity of the computer. Once the distribution function is established, the function use convolution operation to compute the p-value(s) for given Cramer-von Mises test statistics. So, if you have a sequence of pairs of samples of the same sample sizes `n` and `m`, it is best to compute the test scores for all pairs of samples, and then call this function. If you compute the p-values separately, the process may be slow.

Author(s)

Yuanhui Xiao, Department of Mathematics and Statistics, Georgia State University, Atlanta, Georgia, 30302 <yxiao@gsu.edu>

References

Yuanhui Xiao, Alexander Gordon and Andrei Yakovlev (2007), A C++ program for the Cramer-von Mises two sample test, Journal of Statistical Software, 17(8), <http://www.jstatsoft.org>

See Also

`cvmts.pval`

Examples

```

## sample size of the first sample
n <- 10

## create a sample x of size n from the normal distribution with mean 0 and
## standard deviation 1
x <- rnorm(n, 0, 1)

## sample size of the second sample
m <- 10

## create a sample y of size m from the normal distribution with mean 1 and
## standard deviation 1
y <- rnorm(m, 1, 1)

## compute the Cramer-von Mises test statistic
cvm <- cvmts.test(x, y)

## compute the p-value for the test.
pval <- cvmts.pval(cvm, n, m)

## Now suppose x is a list of samples of the same size (n), and y is a list
## of samples of the same size (m), and you want to test whether each pair
## (x[i], y[i]) were drawn from the same population, for i = 1, 2, ...

## a bad way to use cvmts.pval()

## for(i <- 1; i<=length(x); i++){

##     cvm <- cvmts.test(x[i], y[i])
##     pval <- cvmts.pval(cvm, n, m)
## }

## Why? In each call to the function cvmts.pval(), in the first step established is
## the distribution of the Cramer-von Mises test statistics under the
## assumption that the two sample were drawn from the same population. Then
## the distribution is used to calculate the p-value. The first step may
## be expensive if the sample sizes n and m are large.

## I prefer the following way

## initialize
# cvms <- seq(1, length(x))

## compute test scores
# for(i <- 1; i<=length(x); i++){
#   cvms[i] <- cvmts.test(x[i], y[i])
# }

# compute p-values
# pvals <- cvmts.pval(cvms, n, m)

```

`cvmts.test`*Computing Cramer-von Mises Two Sample Test Statistics*

Description

The Cramer-von Mises two sample test is to test whether two independent samples were drawn from the same population.

Usage

```
cvmts.test(x, y)
```

Arguments

`x` an R object for the first sample, which is a vector
`y` an R object for the second sample, which is also a vector

Value

score of the Cramer-von Mises two-sample test for given samples

Author(s)

Yuanhui Xiao, Department of Mathematics and Statistics Georgia State University Atlanta, GA 30302 <yxiao@gsu.edu>

See Also

`cvmts.test`

Examples

```
## create a sample x of size 10 from the normal distribution with mean 0 and
## standard deviation 1
x <- rnorm(10, 0, 1)

## create a sample y of size 10 from the normal distribution with mean 1 and
## standard deviation 1
y <- rnorm(10, 1, 1)

## Call the function to get the Cramer-von Mises test statistic
cvm <- cvmts.test(x, y)
```

Index

*Topic **nonparametric**

cvmts.pval, [2](#)

cvmts.test, [4](#)

cvmts.pval, [2](#)

cvmts.test, [4](#)