Package ‘dks’

February 26, 2026

Title The double Kolmogorov-Smirnov package for evaluating multiple testing procedures.

Version 1.56.0

Author Jeffrey T. Leek <jleek@jhsph.edu>

Description The dks package consists of a set of diagnostic functions for multiple testing methods. The functions can be used to determine if the p-values produced by a multiple testing procedure are correct. These functions are designed to be applied to simulated data. The functions require the entire set of p-values from multiple simulated studies, so that the joint distribution can be evaluated.

Maintainer Jeffrey T. Leek <jleek@jhsph.edu>

Depends R (>= 2.8)

Imports cubature

License GPL

biocViews MultipleComparison, QualityControl

git_url https://git.bioconductor.org/packages/dks

git_branch RELEASE_3_22

git_last_commit d5a1582

git_last_commit_date 2025-10-29

Repository Bioconductor 3.22

Date/Publication 2026-02-26

Contents

cred.set *Calculate a credible set for the posterior distribution on the Beta hyperparameters.*

Description

This function accepts a distribution calculated with pprob.dist and calculates a credible set of the specified level for the hyperparameters. If the credible set includes the value (1,1) the sample is likely to be uniform.

Usage

```
cred.set(dist,delta=NULL,level=0.95)
```

Arguments

`dist`	The posterior distribution for the hyperparameters computed with pprob.dist.
`delta`	The grid size, must match the grid size from pprob.dist.
`level`	The level of the credible set.

Details

The cred.set function calculates a credible set of the specified level based on the distribution calculated with pprob.dist. The grid size, delta, should match the grid size from the call to pprob.dist. The result is a matrix of the same size as dist which indicates whether each point is in the credible set.

Value

`cred`	The credible set for the hyper-parameters of the beta distribution.
`level`	The user specified level of the set.
`elevel`	The empirical level of the set, the smaller delta is, the closer elevel will be to level.

Author(s)

Jeffrey T. Leek `<jleek@jhsph.edu>`

References

J.T. Leek and J.D. Storey, "The Joint Null Distribution of Multiple Hypothesis Tests."

See Also

`dks`, `dks.pvalue`, `pprob.dist`,`cred.set`

Examples

```
## Load data
data(dksdata)

## Calculate the posterior distribution
dist1 <- pprob.dist(P[,1])

delta = 0.1
## Calculate a 95% credible set
cred1 <- cred.set(dist1,delta=0.1)

## Plot the posterior and the credible set

alpha <- seq(0.1,10,by=delta)
beta <- seq(0.1,10,by=delta)

par(mfrow=c(1,2))
image(log10(alpha),log10(beta),dist1,xaxt="n",yaxt="n",xlab="Alpha",ylab="Beta")
axis(1,at=c(-2,-1,0,1,2),labels=c("10^-2","10^-1","10^0","10^1","10^2"))
axis(2,at=c(-2,-1,0,1,2),labels=c("10^-2","10^-1","10^0","10^1","10^2"))
points(0,0,col="blue",cex=1,pch=19)

image(log10(alpha),log10(beta),cred1$cred,xaxt="n",yaxt="n",xlab="Alpha",ylab="Beta")
axis(1,at=c(-2,-1,0,1,2),labels=c("10^-2","10^-1","10^0","10^1","10^2"))
axis(2,at=c(-2,-1,0,1,2),labels=c("10^-2","10^-1","10^0","10^1","10^2"))
points(0,0,col="blue",cex=1,pch=19)
```

---

dks *Frequentist and Bayesian diagnostic tests for multiple testing p-values.*

Description

This function accepts a matrix of simulated null p-values where each column corresponds to the p-values from a single simulated study. The null p-values should represent a subset of all the simulated p-values corresponding to the tests with no signal.

Usage

```
dks(P,alpha=c(0.1,10),beta=c(0.1,10),plot=TRUE,eps=1e-10)
```

Arguments

`P` An m0 x B matrix of null p-values, each column corresponds to the p-values from a single simulated study.

`alpha` The range of the first parameter for the prior on the beta distribution.

`beta` The range of the second parameter for the prior on the beta distribution.

`plot` Should diagnostic plots be displayed.

`eps` Maximum integration error when computing the posterior distribution.

Details

The dks function performs the Bayesian and Frequentist diagnostic tests outlined in Leek and Storey (2009). The result of the function is a double Kolmogorov-Smirnov p-value as well as posterior probability of uniformity estimates for each of the studies. The p-values should be simulated from a realistic distribution and only the null p-values should be passed to the dks function.

Value

`dkspvalue`	The double Kolmogorov-Smirnov p-value.
`postprob`	A B-vector of the posterior probability that each study's null p-values are uniform.

Author(s)

Jeffrey T. Leek `<jleek@jhsph.edu>`

References

J.T. Leek and J.D. Storey, "The Joint Null Distribution of Multiple Hypothesis Tests."

See Also

`pprob.uniform`, `dks.pvalue`, `pprob.dist`,`cred.set`

Examples

```
## Load data
data(dksdata)

## Perform the diagnostic tests with plots
dks1 <- dks(P)
dks1$dkspvalue
```

---

dks.pvalue *Frequentist diagnostic test for multiple testing p-values.*

Description

This function accepts a matrix of simulated null p-values where each column corresponds to the p-values from a single simulated study. The null p-values should represent a subset of all the simulated p-values corresponding to the tests with no signal.

Usage

```
dks.pvalue(P)
```

Arguments

`P`	An m0 x B matrix of null p-values, each column corresponds to the p-values from a single simulated study.

Details

The dks.pvalue function performs the double Kolmogorov-Smirnov test outlined in Leek and Storey (2009). The p-values should be simulated from a realistic distribution and only the null p-values should be passed to the dks.pvalue function.

Value

`dkspvalue`	The double Kolmogorov-Smirnov p-value.
`kspvalue`	A B-vector of the Kolmogorov-Smirnov p-values one for each test.

Author(s)

Jeffrey T. Leek `<jleek@jhsph.edu>`

References

J.T. Leek and J.D. Storey, "The Joint Null Distribution of Multiple Hypothesis Tests."

See Also

`pprob.uniform`, `dks`, `pprob.dist`,`cred.set`

Examples

```
## Load data
data(dksdata)

## Calculate the double KS p-value
dksp <- dks.pvalue(P)
dksp$dkspvalue

## Histogram of the distribution of KS test p-values
hist(dksp$kspvalue)
```

`dksdata` *Simulated null p-values from the uniform distribution.*

Description

This data set can be used to illustrate the behavior of the functions in the dks package. P is a matrix of null p-values, where each column corresponds to the p-values from a single study.

Usage

```
P
```

P — *Simulated null p-values from the uniform distribution.*

Description

This data set is a simulated 200 x 100 matrix of null p-values where each of the 100 columns corresponds to a distinct study and each column contains 200 simulated p-values.

Usage

```
P
```

Format

matrix

pprob.dist — *The posterior distribution for the hyper-parameters of the Beta distribution.*

Description

This function accepts a vector of simulated null p-values from a single simulated study. The null p-values should represent a subset of all the simulated p-values corresponding to the tests with no signal. The result is an estimated posterior distribution for the parameters of the Beta distribution. A posterior centered at (1,1) suggests a uniform distribution.

Usage

```
  pprob.dist(p,alpha=c(0.1,10),beta=c(0.1,10),delta=0.10,eps=1e-10)
```

Arguments

`p`	An vector of null p-values from a single simulated study.
`alpha`	The range of the first parameter for the prior on the beta distribution.
`beta`	The range of the second parameter for the prior on the beta distribution.
`delta`	The grid size, the posterior is calculated over the range of the parameters at grid points separated by delta.
`eps`	Maximum integration error when computing the posterior distribution.

Details

The pprob.dist function calculates the posterior probability for the parameters of the beta distribution given the sample p. The prior is assumed to be uniform on the range specified by the user. A posterior distribution is returned in the form of a matrix, where element (i,j) is the posterior at (alpha[1] + i*delta, beta[1] + j*delta). The null p-values should be simulated from a realistic distribution and only the null p-values should be passed to the pprob.dist function.

Value

`dist`	The posterior distribution in the form of a matrix.

Author(s)

Jeffrey T. Leek `<jleek@jhsph.edu>`

References

J.T. Leek and J.D. Storey, "The Joint Null Distribution of Multiple Hypothesis Tests."

See Also

`dks`, `dks.pvalue`, `pprob.uniform`,`cred.set`

Examples

```
## Load data
data(dksdata)

## Calculate the posterior distribution
dist1 <- pprob.dist(P[,1])

delta <- 0.1

## Plot the posterior distribution
alpha <- seq(0.1,10,by=delta)
beta <- seq(0.1,10,by=delta)
image(log10(alpha),log10(beta),dist1,xaxt="n",yaxt="n",xlab="Alpha",ylab="Beta")
axis(1,at=c(-2,-1,0,1,2),labels=c("10^-2","10^-1","10^0","10^1","10^2"))
axis(2,at=c(-2,-1,0,1,2),labels=c("10^-2","10^-1","10^0","10^1","10^2"))
points(0,0,col="blue",cex=1,pch=19)
```

`pprob.uniform` *Bayesian diagnostic test for multiple testing p-values.*

Description

This function accepts a vector of simulated null p-values from a single simulated study. The null p-values should representa subset of all the simulated p-values corresponding to the tests with no signal.

Usage

```
pprob.uniform(p,alpha=c(0.1,10),beta=c(0.1,10),eps=1e-10)
```

Arguments

`p`	An vector of null p-values from a single simulated study.
`alpha`	The range of the first parameter for the prior on the beta distribution.
`beta`	The range of the second parameter for the prior on the beta distribution.
`eps`	Maximum integration error when computing the posterior distribution.

Details

The pprob.uniform function calculates the posterior probability that a set of null p-values come from the uniform distribution as described in Leek and Storey (2009). The p-values should be simulated from a realistic distribution and only the null p-values should be passed to the pprob.uniform function.

Value

pp	The posterior probability that p is a sample from the uniform distribution.

Author(s)

Jeffrey T. Leek <jleek@jhsph.edu>

References

J.T. Leek and J.D. Storey, "The Joint Null Distribution of Multiple Hypothesis Tests."

See Also

dks, dks.pvalue, pprob.dist,cred.set

Examples

```
## Load data
data(dksdata)
pp <- pprob.uniform(P[,1])
hist(pp)
```

Index