Package ‘PsiHat’

October 5, 2015

Type Package

Title Several Local False Discovery Rate Estimators

Version 1.0

Date 2015-09-30

Author Alaa Ali, Kyle Leckett, Marta Padilla, David R. Bickel (contributions from Bradley Efron, Brit B. Turnbull, Balasubramanian Narasimhan from package locfdr).

Maintainer Marta Padilla <padilla.mpf@gmail.com>

Description Suite of R functions for the estimation of local false discovery rate (LFDR) using several methods.

License GPL-3

Depends methods, qvalue, splines


NeedsCompilation no

Repository CRAN

Date/Publication 2015-10-05 12:58:12

R topics documented:

   PsiHat-package ......................................................... 2
   lfdrc.bbe ............................................................ 3
   lfdrc.elfdr .......................................................... 4
   lfdrc.empiricalNull ............................................... 6
   p0est ................................................................. 8
   rvalue ............................................................... 9

Index 11
Several Local False Discovery Rate Estimators

Description

Suite of R functions for the estimation of local false discovery rate (LFDR) using different methods.

Details

Package: PsiHat
Type: Package
Version: 1.0
Date: 2015-09-30
License: GPL-3
Depends: R (>= 2.14.0), methods, qvalue, splines

Author(s)

Code: David R. Bickel and M. Padilla (modifications).
Authors of included code from other R packages: Bradley Efron, Brit B. Turnbull, Balasubramanian Narasimhan (package: locfdr).
Maintainer: M. Padilla <padilla.mpf@gmail.com>

References

Bickel, D. R. (2013). Simple estimators of false discovery rates given as few as one or two p-values without strong parametric assumptions. Statistical Applications in Genetics and Molecular Biology, 12(4), pp. 529-43.

See Also

R packages: qvalue and locfdr (v1.1.7).

Examples

# a vector of p-values with missing elements:
y <- runif(55); y[c(13, 2, 5)] <- NA
z1 <- lfdr.bbe1(pvalue = y, monotonic = TRUE)
z2 <- lfdr.bbe1(pvalue = y, robust = TRUE)
z3 <- lfdr.bbe(pvalue = y)
z4 <- rvalue(pvalue = y)
z8 <- lfdr.assumedNull(pvalue=y,plot=0,df=3)

# a statistic vector with missing elements:
y <- runif(55, -1, 5); y[c(13, 2, 5)] <- NA

z5 <- lfdr.empiricalNull(stat = y, plot = 0, df = 3)
z6 <- lfdr.assumedNull(stat = y, plot = 0, df = 5)
z7 <- lfdr.elfdr(stat = y, plot = 0, df = 6)

---

**lfdr.bbe**  
*Binomial-based LFDR estimator.*

**Description**

A binomial-based estimator of the local false discovery rate (BBE) with p0 estimation (use `lfdr.bbe`). When p0 = 1 the estimator is called BBE1 (use `lfdr.bbe1`).

**Usage**

```r
lfdr.bbe(pvalue, p0 = NULL, robust = FALSE, monotonic = FALSE, ...)
lfdr.bbe1(pvalue, robust = FALSE, monotonic = FALSE, ...)
```

**Arguments**

- `pvalue`: Input numeric vector of p-values.
- `p0`: Proportion of non-affected features, it is NULL or a value within [0,1]. For `lfdr.bbe1` p0=1, while for `lfdr.bbe` it can be set or it can be estimated internally by `p0est`, setting p0=NULL (the default value).
- `robust`: Logical. If robust=TRUE, the estimate is more robust for small p-values. If robust=TRUE, the estimate is more robust for small p-values (see package qvalue).
- `monotonic`: Logical. If monotonic=TRUE, LFDR estimates are monotonically increasing with p-values.
- `...`: Further arguments passed to function `p0est`.

**Value**

A list with:

- `LFDR.hat`: Vector of estimates of the LFDR,
- `p0.hat`: Estimated p0 (for BBE), 1 (for BBE1)
- `pvalue`: Vector of p-values.
- `info`: Method name and information about computation failure.
Note

If computation fails for all features, \( p_\emptyset \) is set to NA and LFDR is set to a vector of NA with length equal to the number of features. If it fails for a given feature, only the resulting LFDR for that feature is set to NA. Error messages are not suppressed.

Author(s)

Code: David R. Bickel and Marta Padilla (modifications).

References

Bickel, D. R. (2013). Simple estimators of false discovery rates given as few as one or two p-values without strong parametric assumptions. Statistical Applications in Genetics and Molecular Biology, Statistical Applications in Genetics and Molecular Biology, 12(4), pp. 529-43.

See Also

qvalue package and p0est and nqvalue.

Examples

```r
# a p-values vector with missing elements
pp <- runif(15); pp[c(13, 2, 3)] <- NA

#BBE1:
z1 <- lfdr.bbe1(pvalue = pp, robust = TRUE)
z2 <- lfdr.bbe1(pvalue = pp, robust = FALSE, monotonic = TRUE)

#estimating or setting p0:
z3 <- lfdr.bbe(pvalue = pp, p0 = NULL)
z4 <- lfdr.bbe(pvalue = pp, p0.method = "bootstrap")
z5 <- lfdr.bbe(pvalue = pp, p0 = 0.8, robust = TRUE, monotonic = TRUE)
```

lfdr.elfdr

Expected LFDR (ELFDR).

Description

Expected LFDR (ELFDR) of HBE with empirical estimation of the null hypothesis distribution (HBEE) from function lfdr.empiricalNull.
lfdr.elfdr

Usage

lfdr.elfdr(stat = NULL, pvalue = NULL, nulltype = 1, bre = 120,
          df = 7, plot = 0, ...)

Arguments

stat        Input numeric vector of statistics. At least one of the inputs (pvalue or stat) must be non null.
pvalue      Optional input numeric vector of p-values. At least one of the inputs (pvalue or stat) must be non null. If the non empty input is a vector of p-values, the function qnorm is applied to it (see manual of package locfdr).
nulltype    Parameter for selection of the type of null hypothesis distribution in Efrons method. The value nulltype = 1 is the default in lfdr.elfdr (See function locfdr in locfdr package).
bre         Number of breaks in the discretization of the z-score axis in Efrons method (see package locfdr). Useful when the number of features is small.
df          Degrees of freedom for fitting the estimated density $f(z)$ in Efrons method (see package locfdr). Useful when the number of features is small.
plot        Optional. If plot=0, no plots are made (see package locfdr).
...          Further arguments to pass to function locfdr to compute HBEE (see locfdr R package).

Value

A list with:

lfdr.hat    Vector of estimates of the HBE.
p0.hat      Estimated proportion of unaffected features p0 (null hypothesis).
stat        Input vector of statistics.
info        Method name and information about computation failure.

Note

If computation fails for all features, p0.hat is set to NA and lfdr.hat is set to a vector of NA with length equal to the number of features. If it fails for a given feature, only the resulting LFDR for that feature is set to NA. Error messages are not suppressed.

Author(s)

Code: David R. Bickel and Marta Padilla (modifications).

References

See Also
Function `lfdr.empiricalNull`.

Examples

```r
# a statistic vector with missing elements
ss <- runif(55, -1, 5) ; ss[c(13, 25)] <- NA
# a p-values vector with missing elements
pp <- runif(55) ; pp[c(13, 25)] <- NA

z1 <- lfdr.elfdr(stat = ss, df = 3)
z2 <- lfdr.elfdr(pvalue = pp, df = 3)
```

### Description

Wrapper of Efrons LFDR (here called Histogram-based estimator (HBE)) with empirical estimation of the null hypothesis distribution (HBE): `lfdr.empiricalNull` (nulltype=1 by default) or with assumed theoretical null hypothesis distribution as N(0,1) (HBEA): `lfdr.assumedNull`.

### Usage

```r
lfdr.assumedNull(stat = NULL, pvalue = NULL, bre = 120, df = 7, plot = 0, ...)
lfdr.empiricalNull(stat = NULL, pvalue = NULL, nulltype = 1, bre = 120,
                   df = 7, plot = 0, ...)
```

### Arguments

- **stat**: Input numeric vector of statistics. At least one of the inputs (pvalue or stat) must be non null.
- **pvalue**: Optional input numeric vector of p-values. At least one of the inputs (pvalue or stat) must be non null. If the non empty input is a vector of p-values, the function `qnorm` is applied to it (see manual of package `locfdr`).
- **nulltype**: Parameter for selection of the type of null hypothesis distribution in Efrons method. The value nulltype = 1 is the default in `lfdr.elfdr` (See function `locfdr` in `locfdr` package).
- **bre**: Number of breaks in the discretization of the z-score axis in Efrons method (see package `locfdr`). Useful when the number of features is small.
- **df**: Degrees of freedom for fitting the estimated density $f(z)$ in Efrons method (see package `locfdr`). Useful when the number of features is small.
If the empirical null distribution is desired, the function `lfdr.empiricalNull()` can be used. It takes the following arguments:

- `stat`: Input vector of statistics.
- `pvalue`: Vector of p-values.
- `plot`: Optional. If plot = 0, no plots are made (see package `locfdr`). Further arguments to pass to function `locfdr` to compute HBEE (see `locfdr` R package).

**Value**

A list with:

- `lfdrNhat`: Vector of estimates of the HBE.
- `p0hhat`: Estimated proportion of affected features p0 (null hypothesis).
- `stat`: Input vector of statistics.
- `info`: Method name and information about computation failure.

**Note**

Functions `lfdr.assumedNull` and `lfdr.empirical` are based on function `locfdr` from package `locfdr` (v1.1.7). See references.

If computation fails for all features, `p0hhat` is set to NA and `lfdrNhat` is set to a vector of NA with length equal to the number of features. If it fails for a given feature, only the resulting LFDR for that feature is set to NA. Error messages are not suppressed.

**Author(s)**

Code: David R. Bickel and Marta Padilla (modifications).
Authors of R package `locfdr` on which these functions are based: Bradley Efron, Brit B. Turnbull, Balasubramanian Narasimhan.

**References**


**See Also**

`lfdr.elfdr`, `locfdr` function and package, and R package `splines`.

**Examples**

```r
# a statistic vector with missing elements
ss <- runif(55,-1,5) ; ss[c(13,2,5)]<-NA
# a p-values vector with missing elements
pp <- runif(55) ; pp[c(13,2,5)]<-NA
z1 <- lfdr.assumedNull(stat = ss, df = 3)
z2 <- lfdr.empiricalNull(pvalue = pp, df=3)
```
Estimates the proportion of unaffected features (p0).

Description

Estimates the proportion of null hypothesis (unaffected features) (p0). It is a wrapper of function qvalue from R package qvalue.

Usage

p0est(pvalue, lambda = seq(0, 0.9, 0.05), pi0.method = "smoother", smooth.df = 3, smooth.log.pi0 = FALSE, fdr.level = NULL)

Arguments

pvalue Vector of p-values.
lambda Tuning parameter in [0,1) to estimate p0.
pi0.method Method to estimate the proportion of true null hypothesis (p0): "smoother" or "bootstrap".
smooth.df If pi0.method= "smoother", number of degrees-of-freedom.
smooth.log.pi0 If pi0.method= "smoother" and smooth.log.pi0=TRUE, p0 is estimated by applying a smoother to log p0 estimates against the tuning parameter lambda.
fdr.level A level at which to control the FDR. Must be in (0,1].

Value

A number in [0,1]

Author(s)

Code: Marta Padilla (wrapper of function qvalue from R package qvalue)

See Also

qvalue package, lfdr.bbe and nqvalue.

Examples

# a p-value with missing elements:
y <- runif(15); y[c(13, 2,5)] <- NA
z1 <- p0est(pvalue = y, lambda = 0.5)
z2 <- p0est(pvalue = y, pi0.method = "bootstrap")
**Description**

*rvalue* gives adjusted p-values based on qvalues, whereas *nqvalue* is a wrapper of function *qvalue* from R package *qvalue* with simplified output.

**Usage**

*rvalue*(pvalue, robust = FALSE, ...)

*nqvalue*(pvalue, robust = FALSE, ...)

**Arguments**

- **pvalue**: Input numeric vector of p-values.
- **robust**: Logical. If robust=TRUE, the estimate is more robust for small p-values (see package *qvalue*).
- **...**: Further arguments passed to function *qvalue* from R package *qvalue*.

**Value**

A list with:

- **LFDR.hat**: Vector of estimates of the LFDR.
- **p0.hat**: Estimated p0.
- **pvalue**: Vector of p-values.
- **info**: Method name and information about computation failure.

**Note**

If computation fails for all features, p0.hat is set to NA and LFDR.hat is set to a vector of NA with length equal to the number of features. If it fails for a given feature, only the resulting LFDR.hat for that feature is set to NA. Error messages are not suppressed.

**Author(s)**

Code: Marta Padilla  

**References**

See Also

`p0est` and `qvalue` package.

Examples

```r
# a p-values vector with missing elements
pp <- runif(15) ; pp[c(13,2,5)]<-NA
z1 <- rvalue(pvalue = pp, robust = TRUE)
z2 <- rvalue(pvalue = pp, pi0.method = "bootstrap")

z3 <- nqvalue(pvalue = pp, robust = TRUE)
z4 <- nqvalue(pvalue = pp, pi0.method = "bootstrap")
```
Index

*Topic **BBE1**
  lfd.bbe, 3
*Topic **BBE**
  lfd.bbe, 3
  PsiHat-package, 2
*Topic **Binomial-Based Estimator**
  lfd.bbe, 3
*Topic **ELFDR**
  lfd.elfdr, 4
*Topic **Expected Local False Discovery Rate**
  lfd.elfdr, 4
*Topic **HBEA**
  lfd.empiricalNull, 6
  PsiHat-package, 2
*Topic **HBEE**
  lfd.empiricalNull, 6
*Topic **HBE**
  lfd.empiricalNull, 6
  PsiHat-package, 2
*Topic **Histogram-Based Estimator**
  lfd.empiricalNull, 6
*Topic **LFDR-BBE1**
  lfd.bbe, 3
*Topic **LFDR-BBE**
  lfd.bbe, 3
*Topic **LFDR**
  PsiHat-package, 2
*Topic **Local False Discovery Rate**
  PsiHat-package, 2
*Topic **list(lfd.bbe)**
  lfd.bbe, 3
*Topic **list(lfd.bbe1)**
  lfd.bbe, 3
*Topic **p0**
  p0est, 8
*Topic **proportion of null hypothesis**
  p0est, 8
*Topic **qvalue**

rvalue, 9
*Topic **rvalue**
  rvalue, 9
lfd.assumedNull(lfd.empiricalNull), 6
lfd.bbe, 3, 8
lfd.bbe1(lfd.bbe), 3
lfd.elfdr, 4, 7
lfd.empiricalNull, 6, 6
nqvalue, 4, 8
nqvalue(rvalue), 9
p0est, 4, 8, 10
PsiHat(PsiHat-package), 2
PsiHat-package, 2
qvalue, 2, 4, 10
rvalue, 9