

Package: locfdr (via r-universe)

October 16, 2026

Version 1.1-9

Title Computes Local False Discovery Rates

License GPL-2

Encoding UTF-8

Imports graphics, splines, stats

Suggests knitr, rmarkdown, tinytest

VignetteBuilder knitr

LazyData true

Description Computation of local false discovery rates.

NeedsCompilation no

RoxygenNote 7.3.3

Repository <https://bnaras.r-universe.dev>

Date/Publication 2026-04-09 18:58:09 UTC

RemoteUrl <https://github.com/bnaras/locfdr>

RemoteRef HEAD

RemoteSha 4a017db142f3da61f98a9265de1bf0a8a74ffd24

Contents

hivdata	2
lfdrsim	2
loccov	3
loccov2	4
locfdr	4
locmle	7

Index	8
--------------	----------

hivdata	<i>HIV Gene Expression Data</i>
---------	---------------------------------

Description

The data comprises 7680 z-values, each relating to a two-sample t-test. The test compares gene expression values for 4 HIV patients with values for 4 normal subjects; the t-score 'T[i]' for gene i has been transformed to a normal scale, 'z[i] = qnorm(pt(T[i], df=6))', so that the 'z[i]'s theoretically would have a standard $N(0, 1)$ distribution under the null hypothesis.

Usage

```
hivdata
```

Format

A numeric vector containing 7680 z-values.

References

van't Wout, et al., Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4+-T-Cell lines, *Journal of Virology* 77, 1392-1402.

Examples

```
data(hivdata)
hist(hivdata, breaks = 100)
```

lfdrsim	<i>Simulated Data for locfdr</i>
---------	----------------------------------

Description

A simulated dataset that involves 2000 "genes", each of which has yielded a test statistic "zex", with $zex[i] \sim N(\mu_i, 1)$ (independently for $i = 1, 2, \dots, 2000$). Of the 2000 μ_i values, 1800 (90%) are zero (null cases) and the remaining 200 (10%) are drawn from a $N(3, 1)$ distribution (non-null cases).

Usage

```
lfdrsim
```

Format

A data frame with 2000 rows and 2 columns:

mu the true score for each gene

zex the observed z-value

Examples

```
data(lfdrsim)
hist(lfdrsim$zex, breaks = 80)
```

loccov

*Covariance Calculations for Local FDR***Description**

Computes influence functions and covariance matrices for local false discovery rate estimates under the MLE null.

Usage

```
loccov(N, N0, p0, d, s, x, X, f, JV, Y, i0, H, h, sw)
```

Arguments

N	Total number of cases.
N0	Number of cases in the estimation interval.
p0	Estimated proportion of null cases.
d	Estimated null mean (delta).
s	Estimated null standard deviation (sigma).
x	Vector of bin midpoints.
X	Design matrix for the density fit.
f	Fitted density values at bin midpoints.
JV	Jacobian-variance product from MLE.
Y	Two-vector of sufficient statistics from the MLE interval.
i0	Indices of bins in the central matching region.
H	Vector of truncated normal moments.
h	Derivative terms from truncated normal.
sw	Switch: 2 returns parameter derivatives with respect to bin counts; 3 returns influence function of log(fdr); otherwise returns covariance matrix of log(fdr).

Value

Depends on 'sw': derivative matrix (sw=2), influence function matrix (sw=3), or covariance matrix (otherwise).

loccov2

Covariance Calculations for Central Matching Null

Description

Computes covariance and influence functions for local false discovery rate estimates under the central matching or theoretical null.

Usage

```
loccov2(X, X0, i0, f, ests, N)
```

Arguments

X	Design matrix for the mixture density fit.
X0	Design matrix for the null density fit.
i0	Indices of bins in the central matching region.
f	Fitted mixture density values at bin midpoints.
ests	Three-vector of null parameter estimates 'c(delta, sigma, p0)'.
N	Total number of cases.

Value

A list with components:

lIfdr Influence function matrix for log(fdr).

pds. Parameter derivative matrix (3 x K).

stdev Standard deviations of parameter estimates.

Cov Covariance matrix of log(fdr) estimates.

locfdr

Local False Discovery Rate Calculation

Description

Compute local false discovery rates, following the definitions and description in references listed below.

Usage

```
locfdr(
  zz,
  bre = 120,
  df = 7,
  pct = 0,
  pct0 = 1/4,
  nulltype = 1,
  type = 0,
  plot = 1,
  mult,
  mlests,
  main = " ",
  sw = 0
)
```

Arguments

zz	A vector of summary statistics, one for each case under simultaneous consideration. The calculations assume a large number of cases, say ‘length(zz)’ exceeding 200. Results may be improved by transforming zz so that its elements are theoretically distributed as $N(0, 1)$ under the null hypothesis. See the locfdr vignette for tips on creating zz.
bre	Number of breaks in the discretization of the z-score axis, or a vector of break-points fully describing the discretization. If ‘length(zz)’ is small, such as when the number of cases is less than about 1000, set bre to a number lower than the default of 120.
df	Degrees of freedom for fitting the estimated density $f(z)$.
pct	Excluded tail proportions of zz’s when fitting $f(z)$. ‘pct=0’ includes full range of zz’s. pct can also be a 2-vector, describing the fitting range.
pct0	Proportion of the zz distribution used in fitting the null density $f_0(z)$ by central matching. If a 2-vector, e.g. ‘pct0=c(0.25,0.60)’, the range ‘[pct0[1], pct0[2]]’ is used. If a scalar, ‘[pct0, 1-pct0]’ is used.
nulltype	Type of null hypothesis assumed in estimating $f_0(z)$, for use in the fdr calculations. 0 is the theoretical null $N(0, 1)$, 1 is maximum likelihood estimation, 2 is central matching estimation, 3 is a split normal version of 2.
type	Type of fitting used for f ; 0 is a natural spline, 1 is a polynomial, in either case with degrees of freedom df (so total degrees of freedom including the intercept is ‘df+1’).
plot	Plots desired. 0 gives no plots. 1 gives single plot showing the histogram of zz and fitted densities f and $p_0 f_0$. 2 also gives plot of fdr, and the right and left tail area Fdr curves. 3 gives instead the f_1 cdf of the estimated fdr curve; plot=4 gives all three plots.
mult	Optional scalar multiple (or vector of multiples) of the sample size for calculation of the corresponding hypothetical EfdR value(s).
mlests	Optional vector of initial values for (δ_0, σ_0) in the maximum likelihood iteration.

main Main heading for the histogram plot when 'plot>0'.

sw Determines the type of output desired. 2 gives a list consisting of the last 5 values listed under Value below. 3 gives the square matrix of dimension bre-1 representing the influence function of log(fdr). Any other value of sw returns a list consisting of the first 7 (8 if mult is supplied) values listed below.

Value

Depending on the value of 'sw':

For 'sw' other than 2 or 3, a list with components:

fdr the estimated local false discovery rate for each case, using the selected type and nulltype.

fp0 a matrix of estimated parameters delta, sigma, and p0, along with their standard errors, for all three null types.

Efdr the expected false discovery rate for the non-null cases.

cdf1 a 99x2 matrix giving the estimated cdf of fdr under the non-null distribution f_1 .

mat a matrix of estimates at the bin midpoints.

z.2 the interval where $\text{fdr} < 0.2$.

call the function call.

mult if the argument mult was supplied, vector of Efdr ratios for sample size multiples.

For 'sw=2', a list with components 'pds', 'x', 'f', 'pds.', 'stdev'.

For 'sw=3', the influence function matrix of log(fdr).

Author(s)

Bradley Efron, Brit B. Turnbull, and Balasubramanian Narasimhan

References

Efron, B. (2004) "Large-scale simultaneous hypothesis testing: the choice of a null hypothesis", *Jour Amer Stat Assoc*, **99**, pp. 96–104.

Efron, B. (2006) "Size, Power, and False Discovery Rates".

Efron, B. (2007) "Correlation and Large-Scale Simultaneous Significance Testing", *Jour Amer Stat Assoc*, **102**, pp. 93–103.

Examples

```
## HIV data example
data(hivdata)
w <- locfdr(hivdata)
```

locmle

*Maximum Likelihood Estimation for Null Distribution Parameters***Description**

Uses z-values in a specified interval to find maximum likelihood estimates for the null distribution parameters p_0 , δ_0 , and σ_0 .

Usage

```
locmle(z, xlim, Jmle = 35, d = 0, s = 1, ep = 1/1e+05, sw = 0, Cov.in)
```

Arguments

<code>z</code>	Vector of summary statistics (z-values).
<code>xlim</code>	Two-vector 'c(center, half-width)' defining the interval '[center - half-width, center + half-width]' used for estimation. If missing, computed automatically from 'z'.
<code>Jmle</code>	Number of iterations for the MLE algorithm.
<code>d</code>	Initial value for delta (null mean).
<code>s</code>	Initial value for sigma (null standard deviation).
<code>ep</code>	Convergence tolerance.
<code>sw</code>	If 1, returns correlation matrix along with MLE.
<code>Cov.in</code>	Optional list with components 'x', 'X', 'f', 'sw' for covariance and influence function calculations.

Value

If 'sw' is not 1 and 'Cov.in' is not supplied, a named vector of length 6: 'del0', 'sig0', 'p0', 'sd.del0', 'sd.sig0', 'sd.p0'.

If 'sw=1', a list with 'mle' (the estimates) and 'Cor' (correlation matrix).

If 'Cov.in' is supplied, a list with 'mle' and either 'Cov.lfdr', 'pds.', or 'lfdr' depending on 'Cov.in\$sw'.

Index

* **datasets**

 hivdata, [2](#)

 lfdrsim, [2](#)

hivdata, [2](#)

lfdrsim, [2](#)

loccov, [3](#)

loccov2, [4](#)

locfdr, [4](#)

locmle, [7](#)