

Package ‘countdata’

September 14, 2021

Type Package

Title The Beta-Binomial Test for Count Data

Version 1.2

Date 2021-09-14

Maintainer Thang Pham <t.pham@amsterdamumc.nl>

Description The beta-binomial test is used for significance analysis of independent samples by Pham et al. (2010) <[doi:10.1093/bioinformatics/btp677](https://doi.org/10.1093/bioinformatics/btp677)>. The inverted beta-binomial test is used for paired sample testing, e.g. pre-treatment and post-treatment data, by Pham and Jimenez (2012) <[doi:10.1093/bioinformatics/bts394](https://doi.org/10.1093/bioinformatics/bts394)>.

Depends R (>= 2.10)

License BSD_3_clause + file LICENSE

Encoding UTF-8

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation yes

Author Thang Pham [aut, cre, cph, ctb]
(<<https://orcid.org/0000-0003-0333-2492>>)

Repository CRAN

Date/Publication 2021-09-14 19:30:02 UTC

R topics documented:

bb.test	2
fold.change	3
ibb.test	4
normalize	5
Index	7

bb.test *The beta-binomial test*

Description

Performs the beta-binomial test for count data.

Usage

```
bb.test(x, tx, group, alternative = c("two.sided", "less", "greater"),
        n.threads = -1, verbose = TRUE)
```

Arguments

x	A vector or matrix of counts. When x a matrix, the test is performed row by row.
tx	A vector or matrix of the total sample counts. When tx is a matrix, the number of rows must be equal to the number of rows of x.
group	A vector of group indicators.
alternative	A character string specifying the alternative hypothesis: "two.sided" (default), "greater" or "less".
n.threads	The number of threads to be used. When n.threads is 0, the maximal number of CPU cores is used. When n.threads is -1 (default), one CPU core less than the maximum is used, and so on.
verbose	A logical value. If TRUE (default), status information is printed.

Details

This test is designed for independent samples, two or more groups.

Value

A list with a single component is returned

p.value	The p-value of the test.
---------	--------------------------

Author(s)

Thang V. Pham

References

Pham TV, Piersma SR, Warmoes M, Jimenez CR (2010) On the beta binomial model for analysis of spectral count data in label-free tandem mass spectrometry-based proteomics. *Bioinformatics*, 26(3):363-369.

Examples

```
x <- c(1, 5, 1, 10, 9, 11, 2, 8)
tx <- c(19609, 19053, 19235, 19374, 18868, 19018, 18844, 19271)
group <- c(rep("cancer", 3), rep("normal", 5))
bb.test(x, tx, group)
# p.value = 0.01568598
```

fold.change	<i>Fold change calculation</i>
-------------	--------------------------------

Description

Calculates the fold changes between two numerical matrices row by row.

Usage

```
fold.change(d1, d2, BIG = 1e4)
```

Arguments

d1	The first data matrix.
d2	The second data matrix.
BIG	A number representing a big value of the result, i.e. black-and-white regulation.

Details

The two matrices d1 and d2 must have the same number of rows. A positive value means up-regulation where the average of d2 is higher than that of d1. Conversely, a negative value means down-regulation where the the average of d1 is higher than that of d2. If one group contains all zeros, a positive or negative BIG value is returned.

Value

A vector of fold changes is returned.

Author(s)

Thang V. Pham

References

Pham TV (2021). countdata: The Beta-Binomial Test for Count Data. R package version 1.1. <https://CRAN.R-project.org/package=countdata>

Examples

```
x <- rbind(c(1.5, 1.2, 10.2),
           c(9.1, 2.2, 8.1))

fold.change(x, 2 * x)
# returns a column vector of 2

fold.change(x, 0.5 * x)
# returns a column vector of -2
```

ibb.test

The inverted beta-binomial test

Description

Performs the inverted beta-binomial test for paired count data.

Usage

```
ibb.test(x, tx, group, alternative = c("two.sided", "less", "greater"),
         n.threads = -1, BIG = 1e4, verbose = TRUE)
```

Arguments

x	A vector or matrix of counts. When x is a matrix, the test is performed row by row.
tx	A vector or matrix of the total sample counts. When tx is a matrix, the number of rows must be equal to the number of rows of x.
group	A vector of group indicators. There should be two groups of equal size. The samples are matched by the order of appearance in each group.
alternative	A character string specifying the alternative hypothesis: "two.sided" (default), "greater" or "less".
n.threads	The number of threads to be used. When n.threads is 0, the maximal number of CPU cores is used. When n.threads is -1 (default), one CPU core less than the maximum is used, and so on.
BIG	A number representing a big value of the result, i.e. black-and-white regulation.
verbose	A logical value. If TRUE (default), status information is printed.

Details

This test is designed for paired samples, for example data acquired before and after treatment.

Value

A list of values is returned

p.value	The p-value of the test.
fc	An estimation of the common fold change for all sample pairs. A positive value means up-regulation, i.e. the second group is higher, and a negative value down-regulation. A black-and-white regulation is denoted by the BIG value.

Author(s)

Thang V. Pham

References

Pham TV, Jimenez CR (2012) An accurate paired sample test for count data. *Bioinformatics*, 28(18):i596-i602.

Examples

```
x <- c(33, 32, 86, 51, 52, 149)
tx <- c(7742608, 15581382, 20933491, 7126839, 13842297, 14760103)
group <- c(rep("cancer", 3), rep("normal", 3))

ibb.test(x, tx, group)
# p.value = 0.004103636
# fc = 2.137632
```

normalize

Global normalization of count data

Description

Normalize a numerical matrix by scaling each column so that the scaled column sums are equal.

Usage

```
normalize(d)
```

Arguments

d A numerical matrix.

Details

The average of column sums is computed. A scaling factor is calculated for each column so that the scaled column sum is equal to the computed average value.

Value

A matrix of the same size as `d` is returned. The column sums of this matrix are equal.

Author(s)

Thang V. Pham

References

Pham TV (2021). `countdata`: The Beta-Binomial Test for Count Data. R package version 1.1.
<https://CRAN.R-project.org/package=countdata>

Examples

```
d <- rbind(c(2.5, 11.2, 7.2),  
          c(9.1, 2.2, 7.1))
```

```
colSums(d)  
# 11.6 13.4 14.3
```

```
colSums(normalize(d))  
# 13.1 13.1 13.1
```

Index

- * **beta-binomial**
 - bb.test, 2
 - * **count data**
 - bb.test, 2
 - ibb.test, 4
 - * **data normalization**
 - normalize, 5
 - * **fold change**
 - fold.change, 3
 - * **independent sample test**
 - bb.test, 2
 - * **inverted beta-binomial test**
 - ibb.test, 4
 - * **paired sample test**
 - ibb.test, 4
 - * **significance analysis**
 - bb.test, 2
 - ibb.test, 4
- bb.test, 2
- fold.change, 3
- ibb.test, 4
- normalize, 5