Package ‘ZIBseq’

October 12, 2022

Type Package

Title Differential Abundance Analysis for Metagenomic Data via Zero-Inflated Beta Regression

Version 1.2

Date 2017-3-12

Author XIAOLING PENG, GANG LI, ZHENQIU LIU, HONGLIANG CHEN

Maintainer Hongliang Chen<294236451@qq.com>

Description Detects abundance differences across clinical conditions. Besides, it takes the sparse nature of metagenomic data into account and handles compositional data efficiently.

License GPL (>= 2)

Copyright Hongliang Chen

LazyLoad yes

Depends R (>= 3.3.1), gamlss, nlme

Imports stats, gamlss.dist

Repository CRAN

Date/Publication 2017-06-14 13:06:05 UTC

NeedsCompilation no

R topics documented:

ZIBseq-package .................................................. 2
calc_qvalues ..................................................... 3
testdata ........................................................... 4
ZIBseq ............................................................ 5

Index 6
Description

Detects abundance differences across clinical conditions. Besides, it takes the sparse nature of metagenomic data into account and handles compositional data efficiently.

Index of help topics:

- **ZIBseq**
  - Conducts the zero-inflated beta regression based on the general count 'data' and categorical vector 'outcome'.
- **ZIBseq-package**
  - Identify differentially abundant features
- **calc_qvalues**
  - a function used to calculate q values
- **testdata**
  - Real metagenomic data

~~ An overview of how to use the package, including the most important functions ~~

Author(s)

XIAOLING PENG, GANG LI, ZHENQIU LIU, HONGLIANG CHEN

Maintainer: Hongliang Chen <294236451@qq.com>

References


See Also

~~ Optional links to other man pages, e.g. ~~ ~~ ZIBseq ~~

Examples

```r
## Not run:
data(testdata)
x=testdata[,9:248]
p=dim(x)[2]
for (i in 1:p){x[,i]=as.numeric(as.character(x[,i]))}
gr=testdata[,2]
gr=as.numeric(gr)
gr[which(gr<4)]=0
gr[which(gr==4)]=1
result=ZIBseq(data=x,outcome=gr)
## End(Not run)
```
Description
Estimates their q-values based on a list of p-values resulting from the simultaneous testing of many hypothesis.

Usage
calc_qvalues(pvalues)

Arguments
pvalues input the p value

Details
To control the false discovery rate(FDR), q-value has been widely accepted as an alternative approach for multiple hypothesis testing correction in recent years.

Value
qvalues

Author(s)
chen hongliang

References

Examples
## Should be DIRECTLY executable !! ----
##-- ==> Define data, use random, 
##--or do help(data=index) for the standard data sets.

## The function is currently defined as
function (pvalues)
{
  nrows = length(pvalues)
  lambdas <- seq(0, 0.95, 0.01)
  pi0_hat <- array(0, dim = c(length(lambdas)))
  for (l in 1:length(lambdas)) {
    count = 0
    for (i in 1:nrows) {
      if (pvalues[i] > lambdas[l]) {
```r
count = count + 1
pi0_hat[l] = count/(nrows * (1 - lambdas[l]))

f <- unclass(smooth.spline(lambdas, pi0_hat, df = 3))
f_spline <- f$y
pi0 = f_spline[length(lambdas)]
ordered_ps <- order(pvalues)
pvalues <- pvalues
qvalues <- array(0, dim = c(nrows))
ordered_qs <- array(0, dim = c(nrows))
ordered_qs[nrows] <- min(pvalues[ordered_ps[nrows]] * pi0, 1)
for (i in (nrows - 1):1) {
  p = pvalues[ordered_ps[i]]
  new = p * nrows * pi0/i
  ordered_qs[i] <- min(new, ordered_qs[i + 1], 1)
}
for (i in 1:nrows) {
  qvalues[ordered_ps[i]] = ordered_qs[i]
}
return(qvalues)
```

### Description

The metagenomic dataset was downloaded from dbGaP under study ID phs000258. The data and analytical results were first reported by Zupancic et al. (2012). There were a total of 310 Amish adult samples with 112 males and 198 females. And there were a total of 240 taxa at the genus level.

### Usage

```r
data(testdata)
```

### Format

`testdata` is a data frame with 310 cases(rows) and 248 variables(columns). Among 248 variables, 240 of them are taxa at the genus level and 8 of them are clinical phenotypes.
ZIBseq

Conducts the zero-inflated beta regression based on the general count data and categorical vector outcome.

Description

zero-inflated beta regression

Usage

ZIBseq(data, outcome, transform = F, alpha = 0.05)

Arguments

data a matrix records the count data
outcome a categorical vector of a specific kind of clinical condition
transform square-root transform of the compositional matrix
alpha customized threshold while calculating q values

Details

The function takes the sparse nature of metagenomics data into account and handle the compositional data efficiently.

Value

sigFeature output the significant feature
useFeature features being concerned
qvalue qvalue
pvalue pvalue

Author(s)

Hongliang Chen

References


See Also
calc_qvalues
Index

* datasets
  testdata, 4
* package
  ZIBseq-package, 2

calc_qvalues, 3, 5

testdata, 4

ZIBseq, 2, 5
ZIBseq-package, 2