

Package ‘ZINQL’

September 2, 2024

Type Package
Title A Zero-Inflated Quantile Approach for Differential Abundance Analysis of Longitudinal Microbiome Data
Version 1.0
Author Shuai Li, Runzhe Li, Wodan Ling and Ni Zhao
Maintainer Shuai Li <sl1201@jh.edu>
Description
This package conducts association testing between individual taxa of microbiome data and clinical variable of interest, while adjusting for other covariates and longitudinal effect.
License GPL (>=3).
Encoding UTF-8
LazyData true
Imports lme4, quantreg, Matrix, MASS, dplyr
RoxygenNote 7.2.3
Suggests knitr,
rmarkdown
VignetteBuilder knitr

R topics documented:

metadata	1
otu.tab	2
ZINQL_fit	2
Index	5

metadata	<i>Sample meta data: metadata</i>
----------	-----------------------------------

Description

Meta data for demonstration purposes.

Usage

metadata

Format

A data frame with 30 rows:

Meta data This is the simulated meta data for demonstration purposes, which includes Abx, age and study.

Source

Simulated data

otu.tab	<i>Sample otu table: otu.tab</i>
---------	----------------------------------

Description

A sample otu table for demonstration purposes.

Usage

otu.tab

Format

A Matrix with 30 rows and 20 columns:

rows Each row is a sample.

columns Each column is the abundance of Taxon.

Source

Simulated data

ZINQL_fit	<i>ZINQL_fit</i>
-----------	------------------

Description

A Zero-Inflated Quantile Approach for Differential Abundance Analysis of Longitudinal Microbiome Data. The test combines both logistic mixed-effects model and longitudinal quantile regression model. The test is applied to individual taxon.

Usage

```
ZINQL_fit(y, formula, formula.logistics = NA,
  meta, C, taus = c(0.1, 0.25, 0.5, 0.75, 0.9),
  method = "both", n.positive.cut = 5,
  seed = 2024)
```

Arguments

y	The taxa read count. It should be numeric.
formula	The full model for ZINQ-L. By default, this is the formula for both logistics mixed-effects regression and longitudinal quantile regression. For example, $y \sim X + Y + (1 ID)$, where y is zero-inflated.
formula.logistics	The model of logistics mixed-effects regression. By default the user does not need to input, and this will be the same as formula.
meta	A data.frame including all covariates of interest. All of them should be numeric. The user should transfer factor variables to binary or numeric values.
C	The name of clinical variable of interest. The final P-value will tell whether the abundance is associated with this variable.
taus	A grid of quantile levels, e.g., 0.5 for the median, 0.75 for the 3rd quartile; default is c(0.1, 0.25, 0.5, 0.75, 0.9).
method	Combination method, "MinP" for MinP combination, "Cauchy" for Truncated Cauchy combination, or 'Both' which will return results for both methods. The default is "Both".
n.positive.cut	Minimum number of positive values for y. If the condition is not satisfied, the test will not be conducted and NA P-values will be returned.
seed	A seed for any steps with randomness. The default is 2024.

Details

- ZINQ-L requires at least n.positive.cut of positive y. If this condition is not satisfied, the test will not be conducted and NA P-values will be returned.
- By default, the user does not need to input formula.logistics. This means that the formula of logistics mixed-effects regression is consistent with the longitudinal quantile regression. If the user wants the formula of them to be different, the user can input a different formula for logistics mixed-effects regression, and the formula will be the model for longitudinal quantile regression.
- ZINQ-L will automatically drop covariates with unique values. ZINQ-L will also drop variables that have unique values on positive part of y for longitudinal quantile regression.

Value

The function returns a list:

- Final_P_value - This is the Final P-value of ZINQ-L.
- model - This is a String variable suggesting the status of the test. If the test is successful conducted, this will be either 'ZINQL_MinP', 'ZINQL_Cauchy' or 'Both'. If the test fails, this will be String 'None'.
- Intermediate_P_value - These are intermediate P-values that are combined to form the final P-value by either Truncated Cauchy or MinP approach.

Author(s)

The package is developed by Shuai Li, Runzhe Li, Wodan Ling and Ni Zhao.

References

references

Examples

```
#generate example data
y = c(rep(0,5),1,2,5,6,7)
meta = data.frame(
  ID=c('1','1','1','2','2','2','3','4','3','4'),
  gender=c(0,0,0,1,1,1,0,1,0,1),
  x=c(0,0,0,1,1,1,1,0,1,0),
  age=c(1,2,3,2,3,4,1,2,2,3)
)

#run the function
result = ZINQL_fit(y=y, formula=y~x+age+gender+(1|ID),
  C='x', taus=c(0.1, 0.25, 0.5, 0.75, 0.9), seed=2024,
  meta=meta, method='Both')
print( result )
```

Index

- * **datasets**
 - metadata, [1](#)
 - otu.tab, [2](#)
- metadata, [1](#)
- otu.tab, [2](#)
- ZINQL_fit, [2](#)