Package 'ZIPG'

May 7, 2026

Type Package

Title Zero-Inflated Poisson-Gamma Regression

Version 1.1

Author Roulan Jiang [aut, cre],
Tianying Wang [aut]

Maintainer Roulan Jiang <roulan2000@gmail.com>

Description We provide a flexible Zero-inflated Poisson-Gamma Model (ZIPG) by connecting both the mean abundance and the variability to different covariates, and build valid statistical inference procedures for both parameter estimation and hypothesis testing. These functions can be used to analyze microbiome count data with zero-inflation and overdispersion. The model is discussed in Jiang et al (2023) <doi:10.1080/01621459.2022.2151447>.

License GPL (>= 2)

URL https://github.com/roulan2000/ZIPG

BugReports https://github.com/roulan2000/ZIPG/issues

Encoding UTF-8

LazyData true

Depends R (>= 2.10)

RoxygenNote 7.2.1

Imports pscl, MASS

NeedsCompilation no

Repository CRAN

Date/Publication 2024-02-05 22:40:18 UTC

Contents

Dietary *A diet-microbiome data*

Description

A diet-microbiome data with shotgun metagenomic sequencing results of fecal samples and daily dietary records of 34 subjects in 17 consecutive days. See original data from article "Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans" (Johnson et al. 2019).

Usage

```
Dietary
```

Format

A list with OTU table and other covariates

OTU OTU table

COV Major covariates extracted from original data

M Sequencing depth

Source

https://github.com/knights-lab/dietstudy_analyses

ZIPG_CI *Get confidence interval from ZIPG model*

Description

Get confidence interval from ZIPG model

Usage

```
ZIPG_CI(ZIPG_res, type = "Wald", CI_type = "normal", alpha = 0.05)
```

Arguments

ZIPG_res	Result from ZIPG_main()
type	Type of hypothesis testing method, 'Wald' or 'bWald'.
CI_type	Type of confidence interval, 'Wald','bWald' or 'pbWald'.
alpha	We construct (1- alpha)% confidence interval by alpha/2 and (1-alpha/2).

Value

Table of confidence interval

Examples

```
data(Dietary)
dat = Dietary
ZIPG_res <- ZIPG_main(data = dat$COV,
X = ~ALC01+nutrPC1+nutrPC2, X_star = ~ ALC01,
W = dat$OTU[,100], M = dat$M )
ZIPG_CI(ZIPG_res)
```

ZIPG_main *Fit zero-inflated poisson-gamma model via EM Algorithm*

Description

Fit zero-inflated poisson-gamma model via EM Algorithm

Usage

```
ZIPG_main(
  data,
  W,
  M,
  X,
  X_star,
  return_model = TRUE,
  pbWald_list = NULL,
  bWald_list = NULL
)
```

Arguments

data	Data.frame for covariates of interest
W	Count data
M	Sequencing depth, ZIPG use log(M) as offset by default
X	Formula of covariates of differential abundance
X_star	Formula of covariates of differential varibility
return_model	whether return full complete imfomation for fitted model
pbWald_list	A list of arguments for parameteric bootstrap Wald test, B for bootstrap sample size, X0 and X_star0 for formula of covariates included in H0
bWald_list	A list of arguments for non-parameteric bootstrap Wald test, B for bootstrap sample size,

Value

A list of ZIPG fitted model. Use ZIPG_summary() for a quick look at the results.

Examples

```
data(Dietary)
dat = Dietary
ZIPG_res <- ZIPG_main(data = dat$COV,
X = ~ALC01+nutrPC1+nutrPC2, X_star = ~ ALC01,
W = dat$OTU[,100], M = dat$M )
ZIPG_summary(ZIPG_res)
```

ZIPG_simulate *Simulate W from ZIPG model*

Description

Simulate W from ZIPG model

Usage

```
ZIPG_simulate(
  M,
  X,
  X_star,
  A = 1,
  d,
  d_star,
  parms,
  N,
  zi = TRUE,
  returnU = FALSE
)
```

Arguments

M	Sequencing depth
X	Covariates matrix with intercept, n * (d+1)
X_star	Covariates matrix with intercept, n * (d_star+1)
A	no use, reserved for multi-taxa
d	number of covariates in X
d_star	number of covariates in X_star
parms	model paraneters, input c(beta,beta*,gamma)
N	repetition times
zi	whether generate zero-inflated distribution
returnU	whether return fluctuation factor U

Value

A list of W generated from ZIPG model with input parameter

Examples

```
data(Dietary)
dat = Dietary
sim_M = sample(dat$M,100,replace = TRUE)
sim_pre = rep(sample(rep(c(0,1),each = 10)),each = 5)
sim_PC1_mean = rep(rnorm(20,mean = 0,sd = 1),each = 5)
sim_PC1_error = rnorm(100,0,0.1)
sim_PC1 = sim_PC1_mean + sim_PC1_error
X = as.matrix(cbind(1,data.frame(X1 = sim_pre,X2 = sim_PC1)))
parms = c(-4.23,1,0.45,0.6,1,0,0) #p = 0.5
W_sim <- ZIPG_simulate(M = sim_M,X=X,X_star=X,d=2,d_star=2,parms = parms,N=100)
hist(W_sim$W_list[[1]])
ZIPG_res <- ZIPG_main(data = data.frame(X1 = sim_pre,X2 = sim_PC1),
X = ~X1+X2, X_star = ~ X1,W = W_sim$W_list[[2]], M = sim_M )
ZIPG_summary(ZIPG_res)
```

ZIPG_summary *Summary for ZIPG_main() result.*

Description

Summary for ZIPG_main() result.

Usage

```
ZIPG_summary(ZIPG_res, type = "Wald")
```

Arguments

ZIPG_res	Result from ZIPG_main()
type	Type of hypothesis testing method, 'Wald','bWald' or 'pbWald'.

Value

pvalue

Examples

```
data(Dietary)
dat = Dietary
ZIPG_res <- ZIPG_main(data = dat$COV,
X = ~ALC01+nutrPC1+nutrPC2, X_star = ~ ALC01,
W = dat$OTU[,100], M = dat$M )
ZIPG_summary(ZIPG_res)
```

Index