Package 'SLTCA'

October 12, 2022

Type Package

Title Scalable and Robust Latent Trajectory Class Analysis

Description Conduct latent trajectory class analysis with longitudinal data. Our method supports longitudinal continuous, binary and count data. For more methodological details, please refer to Hart, K.R., Fei, T. and Hanfelt, J.J. (2020), Scalable and robust latent trajectory class analysis using artificial likelihood. Biometrics <doi:10.1111/biom.13366>.

Depends R (>= 3.3.0)

Imports stats, geepack, VGAM, Matrix, mvtnorm

Version 0.1.0

Maintainer Teng Fei <tfei@emory.edu>

BugReports https://github.com/tengfei-emory/SLTCA/issues

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

NeedsCompilation no

Author Kari Hart [aut],
Teng Fei [cre, aut] (<https://orcid.org/0000-0001-7888-1715>),
John Hanfelt [aut] (<https://orcid.org/0000-0002-8488-6771>)

Repository CRAN

Date/Publication 2020-09-23 15:30:07 UTC

R topics documented:

simulation *Simulate a dataset which can be analyzed by SLTCA*

Description

Simulate a dataset with longitudinal observations.

Usage

```
simulation(n)
```

Arguments

n	Sample size.

Value

Returns a data frame with 6 longitudinal features y.1 - y.6, including count (y.1 and y.2), binary (y.3 and y.4) and continuous (y.5 and y.6) type. Variable baselinecov is the baseline risk factor of latent classes. Variable latent is the true latent class labels.

Author(s)

Teng Fei. Email: tfei@emory.edu

References

Hart, K.R., Fei, T. and Hanfelt, J.J. (2020), Scalable and robust latent trajectory class analysis using artificial likelihood. Biometrics. Accepted Author Manuscript <doi:10.1111/biom.13366>.

Examples

```
dat <- simulation(500)
```

SLTCA *Scalable and Robust Latent Trajectory Class Analysis Using Artificial Likelihood*

Description

Conduct latent trajectory class analysis with longitudinal observations.

Usage

```
SLTCA(
  k = 20,
  dat,
  num_class,
  id,
  time,
  num_obs,
  features,
  Y_dist,
  covx,
  ipw,
  stop,
  tol = 0.005,
  max = 50,
  varest = TRUE,
  balanced = TRUE,
  MSC = "EQIC",
  verbose = TRUE
)
```

Arguments

k	Number of random initialization to start the algorithm.
dat	Input data matrix.
num_class	Number of latent classes in the fitted model.
id	Column name in the data matrix ‘dat‘ for the patient id.
time	Column name in the data matrix ‘dat‘ for the time of longitudinal observations.
num_obs	Column name in the data matrix ‘dat‘ for the number of longitudinal observations (number of visits).
features	A vector of column names in the data matrix ‘dat‘ for the longitudinal observations.
Y_dist	A vector indicating the type of longitudinal observations. An element of Y_dist can be ’normal’,’bin’, and ’poi’ for continuous, binary and count data.
covx	A vector of column names in the data matrix ‘dat‘ for baseline latent class risk factors.
ipw	Column name in the data matrix ‘dat‘ for the inverse probability weights for missingness. ipw=1 if not specified.
stop	Stopping criterion for the algorithm. stop can be either ’tau’ based on posterior probabilities or ’par’ based on point estimation.
tol	A constant such that the algorithm stops if the stopping criterion is below this constant.
max	Maximum number of iterations if the algorithm does not converge.
varest	True or False: whether conduct variance estimation or not.

balanced	True or False: whether the longitudinal observations are equally spaced.
MSC	Model selection criteria: 'AQIC','BQIC' or 'EQIC'.
verbose	Output progress of fitting the model.

Value

A list with point estimates (alpha, beta0, beta1, phi, gamma), variance estimates (ASE), posterior membership probabilities (tau), QICs (qic) of the latent trajectory class model, and stopping criteria (diff) at the last iteration. Point estimates and variance estimates are provided in matrix format, where columns represent latent classes and rows represent covariates or longitudinal features.

Author(s)

Teng Fei. Email: <tfei@emory.edu>

References

Hart, K.R., Fei, T. and Hanfelt, J.J. (2020), Scalable and robust latent trajectory class analysis using artificial likelihood. Biometrics. Accepted Author Manuscript <doi:10.1111/biom.13366>.

Examples

```
# In this illustrative example the sample size is set as n=50,
# variance estimation is skipped by setting varest=FALSE, and
# the maximum number of iterations is set as max=1 in order to pass CRAN test.
# Please use n=500, varest=TRUE and max=50 for more reliable results.

dat <- simulation(n=50)
res <- SLTCA(k=1,dat,num_class=2,"id","time","num_obs",paste("y.",1:6,sep=''),
             Y_dist=c('poi','poi','bin','bin','normal','normal'),
             "baselinecov",1,stop="tau",tol=0.005,max=1,
             varest=FALSE,balanced=TRUE,MSC='EQIC',verbose=FALSE)
```

Index