Package 'autohd'

May 10, 2021

Title High Dimensional Bayesian Survival Mediation Analysis

Version 0.1.0

Date 2021-04-30

Depends R (>= 2.10)

Imports broom,hdbm,Hmisc,ICBayes,icenReg,missForest,SurvRegCensCov,survival,schoolmath, tibble,Rdpack,rjags,usethis,coxme,mlr3

LazyData Yes

LazyDataCompression xz

ByteCompile Yes

Description Perform mediation analysis for time to event high-dimensional data.
Mediation Analysis proposed by Miocevic et al.(2017) <doi:10.1080/10705511.2017.1342541>
as a statistical tool in the Bayesian framework.
Time to event data analysis methods like Cox proportional hazard model,
accelerated failure time model to work with high dimensional data
with Bayesian approaches are provided. Missing data imputation techniques tool
to work with high dimensional data coupled for mediation analysis by
presented by the active mediator variables.

License GPL-3

Encoding UTF-8

NeedsCompilation yes

Maintainer Atanu Bhattacharjee <atanustat@gmail.com>

RoxygenNote 7.1.1.9000

Author Atanu Bhattacharjee [aut, cre, ctb],
Akash Pawar [aut, ctb]

Repository CRAN

Date/Publication 2021-05-10 08:02:24 UTC

R topics documented:

`coxmulti` *Multivariate cox proportional hazard data analysis*

Description

Given the dimensions of the variables and survival informations. The function performs multivariate Cox PH by taking 5 variables at a time.

Usage

```
coxmulti(
  C1 = NULL,
  C2 = NULL,
  C3 = NULL,
  C4 = NULL,
  C5 = NULL,
  survdur,
  event,
  data
)
```

Arguments

`C1`	Covar1
`C2`	Covar2
`C3`	Covar3

C4	Covar4
C5	Covar5
survdur	"Column/Variable name" consisting duration of survival.
event	"Column/Variable name" consisting survival event.
data	High dimensional data containing survival observations and covariates.

Value

Data set containing the survival estimates and Pvalue.

Examples

```
##
coxmulti(C1="PGC",C2="C7",C3="HPN",C4="DDC",C5=NULL,survdur="os",event="death",data=hnscc2)
```

---

hdaftma *High dimensional multivariate accelerated failure time model with bayesian mediation analysis*

Description

Given the dimension of variables and survival information the function filters significant variables by fitting AFT model. Further, it performs mediation analysis among the signifiant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
hdaftma(m, n, survdur, event, ths, sig, b, d, data)
```

Arguments

m	Starting column number from where high dimensional variates to be selected.
n	Ending column number till where high dimensional variates to be selected.
survdur	"Column/Variable name" consisting duration of survival.
event	"Column/Variable name" consisting survival event.
ths	A numeric between 0 to 100.
sig	Level of significance.
b	Number of MCMC iterations to burn.
d	Number of draws for the iterations.
data	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
hdaftma(m=8,n=80,survdur="os",event="death",sig=0.05,ths=0.02,b=10,d=10,data=hnscc2)
##
```

`hdcoxma` *High dimensional multivariate cox proportional model with bayesian mediation analysis.*

Description

Given the dimension of variables and survival information the function filters significant variables by fitting multivariate Cox PH with 5 variables at a time. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
hdcoxma(m, n, survdur, event, ths, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`survdur`	"Column/Variable name" consisting duration of survival.
`event`	"Column/Variable name" consisting survival event.
`ths`	A numeric between 0 to 100.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws for the iterations.
`data`	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
hdcoxma(m=8,n=106,survdur="os",event="death",ths=0.02,b=10,d=10,data=hnscc2)
```

`hdraftma` *High dimensional competing risk analysis using multivariate accelerated failure time model with mediation analysis.*

Description

Given the dimension of variables and survival information including the competing risks the function filters significant variables, allowing the user to fit multivariate AFT model. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
hdraftma(m, n, survdur, event, sig, ths, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`survdur`	"Column/Variable name" consisting duration of survival.
`event`	"Column/Variable name" consisting survival event.
`sig`	Level of significance pre-determined by the user.
`ths`	A numeric between 0 to 100.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws for the iterations.
`data`	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
hdraftma(m=8,n=100,survdur="os",event="death2",sig=0.1,ths=0.02,b=10,d=10,data=hnscc2)
##
```

hdsurvma *High dimensional survival analysis using SurvMCmulti with mediation analysis*

Description

Given the dimension of variables and survival information the function filters significant variables, allowing the user to perform survival analysis with high number of iterations. Further, it performs mediation analysis among the signifiant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
hdsurvma(m, n, Surv, Event, ths, chn, i, adp, b, d, data)
```

Arguments

m	Starting column number from where high dimensional variates to be selected.
n	Ending column number till where high dimensional variates to be selected.
Surv	"Column/Variable name" consisting duration of survival.
Event	"Column/Variable name" consisting survival event.
ths	A numeric between 0 to 100.
chn	Number of MCMC chains to perform survival analysis.
i	Number of MCMC iterations to perform survival analysis.
adp	Number of MCMC adaptations to perform survival analysis.
b	Number of MCMC iterations to burn.
d	Number of draws.
data	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
## Not run:
data2 <- hnscc2[1:50,]
hdsurvma(m=8,n=15,Surv="os",Event="death",ths=0.02,chn=4,i=10,adp=100,b=10,d=10,data=data2)

## End(Not run)
```

hidimcox *High dimensional univariate cox proportional hazard analysis.*

Description

Given the dimension of variables and survival information the function performs univariate Cox PH.

Usage

```
hidimcox(m, n, survdur, event, sig, data)
```

Arguments

m	Starting column number form where study variables of high dimensional data will get selected.
n	Ending column number till where study variables of high dimensional data will get selected.
survdur	Column name of survival duration event, a string value. i.e. "os"
event	Column name of survival event, a string value. i.e "death"
sig	Level of significance pre-determined by the user
data	High dimensional data containing the survival, progression and genomic observations.

Value

Data set containing the list of selected variables with HR, LCL,UCL and Pvalues through survival analysis.

Examples

```
##
data(hnscc2)
hidimcox(m=8,n=50,survdur="os",event="death",sig=0.05,data=hnscc2)
##
```

hnscc *Head and neck cancer data*

Description

High dimensional head and neck cancer gene expression data

Usage

```
data(hnscc)
```

Format

A dataframe with 565 rows and 104 variables

ID ID of subjects

leftcensoring Initial censoring time

death Survival event

death2 Survival event with competing risk

os Duration of overall survival

PFS Duration of progression free survival

Prog Progression event

GJB1,...,HMGCS2 High dimensional covariates

Examples

```
data(hnscc)
```

hnscc2 *Head and neck cancer data*

Description

High dimensional head and neck cancer gene expression data

Usage

```
data(hnscc)
```

Format

A dataframe with 565 rows and 104 variables

ID ID of subjects

leftcensoring Initial censoring time

death Survival event

death2 Survival event with competing risk

os Duration of overall survival

PFS Duration of progression free survival

Prog Progression event

GJB1,...,HMGCS2 High dimensional covariates

Examples

```
data(hnscc2)
```

`imphdaft` *High dimensional missing data imputation and performing mediation analysis with univariate accelerated failure time model using Weibull distribution.*

Description

Given the dimension of variables and survival information the function performs imputations using missForest function and filters significant variables, allowing the user to fit AFT model. Further, it performs mediation analysis among the signifiant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
imphdaft(m, n, survdur, event, time, sig, ths, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`survdur`	"Column/Variable name" consisting duration of survival.
`event`	"Column/Variable name" consisting survival event.
`time`	"Column/Variable name" consisting time of repeated observations.
`sig`	Level of significance pre-determined by the user
`ths`	A numeric between 0 to 100.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws for the iterations.
`data`	High dimensional data containing survival observations with multiple covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
## Not run:
imphdaft(m=6,n=25,survdur="OS",event="event",time="Visit",sig=0.5,ths=0.02,b=10,d=10,data=srdata)

## End(Not run)
##
```

`imphdcox` *High dimensional missing data imputation and performing mediation analysis with multivariate cox proportional hazard model. It works in a multivariate setup.*

Description

Given the dimension of variables and survival information the function performs imputations using missForest function and filters significant variables, allowing the user to fit multivariate CoxPH model with 5 variables. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
imphdcox(m, n, survdur, event, time, sig, ths, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`survdur`	"Column/Variable name" consisting duration of survival.
`event`	"Column/Variable name" consisting survival event.
`time`	"Column/Variable name" consisting time of repeated observations.
`sig`	Level of significance pre-determined by the user.
`ths`	A numeric between 0 to 100.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws for the iterations.
`data`	High dimensional data containing survival observations with multiple covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
## Not run:
imphdcox(m=11,n=25,survdur="OS",event="event",time="Visit",sig=0.5,ths=20,b=10,d=10,data=srdata)
##

## End(Not run)
```

`imphdsurv` *High dimensional missing data imputation and survival analysis using survMCmulti with mediation analysis*

Description

Given the dimension of variables and survival information the function performs imputations using missForest function and filters significant variables, allowing the user to do survival analysis with higher number of iterations. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
imphdsurv(m, n, Surv, Event, time, ths, chn, i, adp, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`Surv`	"Column/Variable name" consisting duration of survival.
`Event`	"Column/Variable name" consisting survival event.
`time`	"Column/Variable name" consisting time of repeated observations.
`ths`	A numeric between 0 to 100.
`chn`	Number of MCMC chains to perform survival analysis.
`i`	Number of MCMC iterations to perform survival analysis.
`adp`	Number of MCMC adaptations to perform survival analysis.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws.
`data`	High dimensional data containing survival observations with multiple covariates.

Details

High dimensional missing data imputation and performing mediation analysis using survMCmulti. It works in a multivariate setup.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
## Not run:
imphdsurv(m=11,n=25,Surv="OS",Event="event",time="Visit",ths=0.02,chn=6,i=10,
          adp=100,b=10,d=10,data=srdata)
##

## End(Not run)
```

`impuni` *High dimensional missing data imputation and performing the mediation analysis with Bayesian univariate cox proportional modeling.*

Description

Given the dimension of variables and survival information the function performs imputations using missForest function and filters significant variables, allowing the user to do univariate survival analysis with higher number of iterations. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
impuni(m, n, Survdur, event, time, lcr = NULL, t, i, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`Survdur`	"Column/Variable name" consisting duration of survival.
`event`	"Column/Variable name" consisting survival event.
`time`	"Column/Variable name" consisting time of repeated observations.
`lcr`	"Leftcensoring information"
`t`	A numeric threshold value between 0 to 1.
`i`	Number of MCMC iteration to perform in obtaining posterior estimates of HR by CoxPH.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws for the iterations.
`data`	High dimensional data containing survival observations with multiple covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
## Not run:
##
impuni(m=8,n=25,Survdur="os",event="death",lcr=,t=0.02,i=6,b=10,d=10,data=hnscc)
##

## End(Not run)
```

`impuniaft` *High dimensional missing data imputation and performing the mediation analysis with univariate accelerated failure time model with Weibull distribution.*

Description

Given the dimension of variables and survival information the function performs imputations using missForest function and filters significant variables, allowing the user to fit univariate AFT. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
impuniaft(m, n, survdur, event, t, sig, ths, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`survdur`	"Column/Variable name" consisting duration of survival.
`event`	"Column/Variable name" consisting survival event.
`t`	"Column/Variable name" consisting time of repeated observations.
`sig`	Level of significance pre-determined by the user.
`ths`	A numeric between 0 to 1.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws for the iterations.
`data`	High dimensional data containing survival observations with multiple covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
## Not run:
impuniaft(m=11,n=25,survdur="OS",event="event",t="Visit",sig=0.5,ths=0.02,b=10,d=10,data=srdata)
##

## End(Not run)
```

`impunicox` *High dimensional missing data imputation and performing the mediation analysis with univariate cox proportional modeling.*

Description

Given the dimension of variables and survival information the function performs imputations using missForest function and filters significant variables, allowing the user to fit univariate CoxPH model. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
impunicox(m, n, survdur, event, t, sig, ths, b, d, data)
```

Arguments

`m`	Starting column number from where high dimensional variates to be selected.
`n`	Ending column number till where high dimensional variates to be selected.
`survdur`	"Column/Variable name" consisting duration of survival.
`event`	"Column/Variable name" consisting survival event.
`t`	"Column/Variable name" consisting time of repeated observations.
`sig`	Level of significance pre-determined by the user.
`ths`	A numeric between 0 to 1.
`b`	Number of MCMC iterations to burn.
`d`	Number of draws for the iterations.
`data`	High dimensional data containing survival observations with multiple covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
## Not run:
impunicox(m=11,n=25,survdur="OS",event="event",t="Visit",sig=.2,ths=0.02,b=10,d=10,data=srdata)
##

## End(Not run)
```

`srdata` *protein data*

Description

High dimensional pprotein gene expression data

Usage

```
data(srdata)
```

Format

A dataframe with 288 rows and 252 columns

ID ID of subjects

Visit Number of times observations recorded

event death as event 1 if died or 0 if alive

OS Duration of overall survival

leftcensored Left censoring information

lc Left censoring information

C6kine,.....,GFRalpha4 High dimensional covariates

Examples

```
data(srdata)
```

unicrma *High dimensional competing risk analysis by univariate accelerated failure time model with mediation analysis with Weibull distribution*

Description

Given the dimension of variables and survival information including the competing risks the function filters significant variables,allowing the user to fit univariate AFT model. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
unicrma(m, n, survdur, event, sig, t, b, d, data)
```

Arguments

m	Starting column number from where high dimensional variates to be selected.
n	Ending column number till where high dimensional variates to be selected.
survdur	"Column/Variable name" consisting duration of survival.
event	"Column/Variable name" consisting survival event.
sig	Level of significance pre-determined by the user.
t	A numeric between 0 to 100.
b	Number of MCMC iterations to burn.
d	Number of draws for the iterations.
data	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
data(hnscc2)
unicrma(m=8,n=100,survdur="os",event="death2",sig=0.05,t=20,b=10,d=10,data=hnscc2)
##
```

unihdaftma *High dimensional competing risk analysis by univariate accelerated failure time model with mediation analysis*

Description

Given the dimension of variables and survival information risks the function filters significant variables, allowing the user to fit univariate AFT model. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
unihdaftma(m, n, survdur, event, ths, b, d, data)
```

Arguments

m	Starting column number from where high dimensional variates to be selected.
n	Ending column number till where high dimensional variates to be selected.
survdur	"Column/Variable name" consisting duration of survival.
event	"Column/Variable name" consisting survival event.
ths	A numeric between 0 to 100.
b	Number of MCMC iterations to burn.
d	Number of draws.
data	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
data(hnscc)
unihdaftma(m=8,n=80,survdur="os",event="death",ths=0.5,b=1000,d=10,data=hnscc2)
##
```

unihdcoxma *High dimensional univariate cox proportional hazard model.*

Description

Given the dimension of variables and survival information risks the function filters significant variables, allowing the user to fit univariate COx PH model. Further, it performs mediation analysis among the significant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
unihdcoxma(m, n, survdur, event, sig, ths, b, d, data)
```

Arguments

m	Starting column number from where high dimensional variates to be selected.
n	Ending column number till where high dimensional variates to be selected.
survdur	"Column/Variable name" consisting duration of survival.
event	"Column/Variable name" consisting survival event.
sig	Level of significance pre-determined by the user.
ths	A numeric between 0 to 100.
b	Number of MCMC iterations to burn.
d	Number of draws for the iterations.
data	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
data(hnscc)
unihdcoxma(m=8,n=105,survdur="os",event="death",sig=0.5,ths=0.02,b=1000,d=10,data=hnscc2)
##
```

unihdma *High dimensional survival analysis using Bayesian univariate cox prportional hazard with mediation analysis*

Description

Given the dimension of variables and survival information the function filters significant variables allowing the user to perform survival anlaysis with high number of iterations. Further, it performs mediation analysis among the signifiant variables and provides handful variables with their alpha.a values which are mediator model exposure coefficients and beta.a coefficients.

Usage

```
unihdma(m, n, survdur, event, LC = NULL, t, i, b, d, data)
```

Arguments

m	Starting column number from where high dimensional variates to be selected.
n	Ending column number till where high dimensional variates to be selected.
survdur	"Column/Variable name" consisting duration of survival.
event	"Column/Variable name" consisting survival event.
LC	"Initial time of getting in to the study.
t	A numeric between 0 to 100.
i	Number of MCMC iteration to perform in obtaining posterior extimates of HR by CoxPH.
b	Number of MCMC iterations to burn.
d	Number of draws for the iterations.
data	High dimensional data containing survival observations and high dimensional covariates.

Value

Data frame containing the beta and alpha values of active variables among the significant variables.

Examples

```
##
unihdma(m=8,n=15,survdur="os",event="death",LC="leftcensoring",t=0.02,i=6,b=10,d=10,data=hnscc2)
##
```

Index