Package ‘SMAHP’

March 24, 2025

Type Package

Title Survival Mediation Analysis of High-Dimensional Proteogenomic
Data

Version 0.0.5

Description
SMAHP (pronounced as SOO-MAP) is a novel multi-omics framework for causal mediation analysis of high-dimensional proteogenomic data with survival outcomes. The full methodological details can be found in our recent preprint by Ahn S et al. (2025) <doi:10.48550/arXiv.2503.08606>.

License GPL-3

Encoding UTF-8

Imports data.table, fdrtool, ncvreg, penAFT, survival, stats, glmnet

RoxygenNote 7.3.2

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Language en-US

VignetteBuilder knitr

Depends R (>= 3.5.0)

NeedsCompilation no

Author Seungjun Ahn [aut] (<`https://orcid.org/0000-0002-4816-8924`>),
Weijia Fu [cre, aut] (<`https://orcid.org/0009-0002-5955-3534`>),
Zhigang Li [ctb]

Maintainer Weijia Fu <`weijia.fu@mountsinai.org`>

Repository CRAN

Date/Publication 2025-03-24 12:00:12 UTC

Contents

`example_dat` *Example Dataset*

Description

A example simulated dataset used in the SMAP package.

Usage

```
data(example_dat)
```

Format

`example_dat`

Examples

```
data(example_dat)
```

`get_adjusted_p` *A function that returns adjusted p-values*

Description

A function that returns adjusted p-values.

Usage

```
get_adjusted_p(res_SMAHP)
```

Arguments

`res_SMAHP`	Outputs from SMAHP

Value

A matrix of adjusted p-values.

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_SMAHP <- SMAHP(example_dat$X, example_dat$M, example_dat$C, time = surv_dat$time,
status = surv_dat$status)
get_adjusted_p(res_SMAHP)
```

`get_coef` *A function that extracts coefficients.*

Description

A function that extracts coefficients from mediation and outcome model.

Usage

```
get_coef(res_SMAHP)
```

Arguments

`res_SMAHP`	Outputs from SMAHP

Value

A data frame which includes the name of exposure and mediator with coefficients from mediation and outcome model.

`beta1`	The coefficient estimate of exposure in mediation model
`beta2`	The coefficient estimate of mediator in outcome model
`beta3`	The coefficient estimate of exposure in outcome model

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_SMAHP <- SMAHP(example_dat$X, example_dat$M, example_dat$C, time = surv_dat$time,
status = surv_dat$status)
get_coef(res_SMAHP)
```

`get_pairs` *A function that returns mediation-exposure matrix*

Description

A function that returns final mediation-exposure matrix.

Usage

```
get_pairs(res_SMAHP)
```

Arguments

`res_SMAHP`	Outputs from SMAHP

Value

Mediation-exposure matrix.

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_SMAHP <- SMAHP(example_dat$X, example_dat$M, example_dat$C, time = surv_dat$time,
status = surv_dat$status)
get_pairs(res_SMAHP)
```

`get_raw_p` *A function that returns unadjusted p-values*

Description

A function that returns unadjusted p-values.

Usage

```
get_raw_p(res_SMAHP)
```

Arguments

`res_SMAHP`	Outputs from SMAHP

Value

A matrix of unadjusted p-values.

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_SMAHP <- SMAHP(example_dat$X, example_dat$M, example_dat$C, time = surv_dat$time,
status = surv_dat$status)
get_raw_p(res_SMAHP)
```

`get_sig_pathway` *A function that returns significant pathways.*

Description

A function that returns significant pathways.

Usage

```
get_sig_pathway(res_SMAHP)
```

Arguments

`res_SMAHP`	Outputs from SMAHP

Value

A data frame which includes the name of exposure and mediator with adjusted p-value from identified significant pathways.

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_SMAHP <- SMAHP(example_dat$X, example_dat$M, example_dat$C, time = surv_dat$time,
status = surv_dat$status)
get_sig_pathway(res_SMAHP)
```

`SIS_combo` *SIS combo*

Description

A function for SIS + SIS combo.

Usage

```
SIS_combo(X, M, C, time, status, p_adjust_option = "BH", thres = 0.05)
```

Arguments

`X`	An n by p matrix of exposures.
`M`	An n by p matrix of mediators.
`C`	An n by p matrix of covariates.
`time`	A vector of survival time of samples.
`status`	A vector of status indicator: 0=alive, 1=dead.
`p_adjust_option`	The method for multiple correction. Option include q-value, holm, hochberg, hommel, bonferroni, BH, BY, and fdr. Default is BH.
`thres`	Threshold for determining significance.

Value

A list which includes the final p-value matrix (p_final_matrix), adjusted p-value matrix (p_adjusted_matrix) and mediation-exposure matrix (p_med_matrix).

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
SIS_combo(example_dat$X, example_dat$M, example_dat$C, time = surv_dat$time,
status = surv_dat$status)
```

`SMAHP` *Main function*

Description

An approach for survival mediation analysis of high-dimensional proteogenomic data.

Usage

```
SMAHP(
  X,
  M,
  C,
  time,
  status,
  model_option = "MCP",
  SIS_thres = "n/log(n)",
  p_adjust_option = "BH",
  p_thres = 0.05
)
```

Arguments

`X`	An n by p matrix of exposures.
`M`	An n by p matrix of mediators.
`C`	An n by p matrix of covariates. If there are no covariates, set C = NULL.
`time`	A vector of survival time of samples.
`status`	A vector of status indicator: 0=alive, 1=dead.
`model_option`	The penalization method used when selecting X for M. Options include MCP, elastic net and lasso. Default is MCP.
`SIS_thres`	SIS thresholds. Default is "n/log(n)". Other options include "n-1", "n/2log(n)", "2n/log(n)", "3n/log(n)".
`p_adjust_option`	The method for multiple correction. Option include q-value, holm, hochberg, hommel, bonferroni, BH, BY, and fdr. Default is BH.
`p_thres`	Threshold for determining significance.

Value

A list with the following components:

`p_final_matrix`	raw p-value matrix before adjustment
`p_adjusted_matrix`	adjusted p-value matrix
`p_med_matrix`	mediation-exposure matrix
`outcome_model`	coefficient estimates from outcome model
`med_results`	coefficient estimates from mediation model

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
SMAHP(example_dat$X, example_dat$M, example_dat$C, time = surv_dat$time, status = surv_dat$status)
```

`StepOne` *Penalized mediation and outcome models*

Description

An auxiliary function which conducts variable selection of X for Y using penAFT, X for M using the selected penalized model and M for Y using penAFT.

Usage

```
StepOne(X, M, time, status, model_option = "MCP")
```

Arguments

`X`	An n by p matrix of exposures.
`M`	An n by p matrix of mediators.
`time`	A vector of survival time of samples.
`status`	A vector of status indicator: 0=alive, 1=dead.
`model_option`	The penalization method used when selecting X for M. The options include MCP, elastic net and lasso. Default is MCP.

Value

A list with the following components:

`X_sel_Y_s1`	X selected for Y using penAFT
`M_X_s1`	X selected for M using the selected penalized model
`M_sel_Y_s1`	M selected for Y using penAFT

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_step1 <- StepOne(X = example_dat$X, M = example_dat$M, time = surv_dat$time,
status = surv_dat$status, model_option = "MCP")
```

`StepThree` *Post-hoc model*

Description

An auxiliary function which generates p-value based on outcome and mediation model.

Usage

```
StepThree(X, M, C, time, status, X_sel_Y_s1, M_X_sel_s2)
```

Arguments

`X`	An n by p matrix of exposures.
`M`	An n by p matrix of mediators.
`C`	An n by p matrix of covariates. If there are no covariates, set C = NULL.
`time`	A vector of survival time of samples.
`status`	A vector of status indicator: 0=alive, 1=dead.
`X_sel_Y_s1`	Outputs from StepOne: A vector of selected X for Y.
`M_X_sel_s2`	Outputs from StepTwo: A data table with selected M, X pairs and related effect size.

Value

A list with the following components:

`p_beta_m`	p-values generated from outcome model
`p_alpha_x`	p-values generated from mediation modell
`outcome_model`	coefficient estimates from outcome model
`med_results`	coefficient estimates from mediation model

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_step1 <- StepOne(X = example_dat$X, M = example_dat$M, time = surv_dat$time,
status = surv_dat$status, model_option = "MCP")
M_X_sel_s2 <- StepTwo(X = example_dat$X, M = example_dat$M, time = surv_dat$time,
status = surv_dat$status, X_sel_Y_s1 = res_step1$X_sel_Y_s1, M_X_s1 = res_step1$M_X_s1,
M_sel_Y_s1 = res_step1$M_sel_Y_s1)
res_step3 <- StepThree(X = example_dat$X, M = example_dat$M, C = example_dat$C,
time = surv_dat$time, status = surv_dat$status, X_sel_Y_s1 = res_step1$X_sel_Y_s1,
M_X_sel_s2 = M_X_sel_s2)
```

`StepTwo` *Select M-X pairs via SIS*

Description

An auxiliary function which selects M-X paris using SIS.

Usage

```
StepTwo(
  X,
  M,
  time,
  status,
  X_sel_Y_s1,
  M_X_s1,
  M_sel_Y_s1,
  SIS_thres = "n/log(n)"
)
```

Arguments

`X`	An n by p matrix of exposures.
`M`	An n by p matrix of mediators.
`time`	A vector of survival time of samples.
`status`	A vector of status indicator: 0=alive, 1=dead.
`X_sel_Y_s1`	Outputs from StepOne: A vector of selected X for Y.
`M_X_s1`	Outputs from StepOne: A list of selected X for M.
`M_sel_Y_s1`	Outputs from StepOne: A vector of selected M for Y.
`SIS_thres`	SIS thresholds. Default is "n/log(n)". Other options include "n-1", "n/2log(n)", "2n/log(n)", "3n/log(n)".

Value

A data table with selected M, X pairs and related effect size.

Examples

```
data(example_dat)
surv_dat <- example_dat$surv_dat
res_step1 <- StepOne(X = example_dat$X, M = example_dat$M, time = surv_dat$time,
status = surv_dat$status, model_option = "MCP")
M_X_sel_s2 <- StepTwo(X = example_dat$X, M = example_dat$M, time = surv_dat$time,
status = surv_dat$status, X_sel_Y_s1 = res_step1$X_sel_Y_s1, M_X_s1 = res_step1$M_X_s1,
M_sel_Y_s1 = res_step1$M_sel_Y_s1)
```

Index