Package 'MMMS'

February 19, 2015

Type Package

Title Multi-Marker Molecular Signature for Treatment-specific Subgroup Identification

Version 0.1

Date 2014-03-18

Author Lin Li, Tobias Guennel,Scott Marshall, Leo Wang-Kit Cheung

Maintainer Lin Li <lli@biostatsolutions.com>

Description The package implements a multi-marker molecular signature (MMMS) approach for treatment-specific subgroup identification.

Reference Lin Li, Tobias Guennel,Scott Marshall, Leo Wang-Kit Cheung (2014) A multi-marker molecular signature approach for treatment-specific subgroup identification with survival outcomes. The Pharmacogenomics Journal, http://dx.doi.org/10.1038/tpj.2014.9

License GPL-3

Depends survival, glmnet

NeedsCompilation no

Repository CRAN

Date/Publication 2014-03-28 17:27:21

R topics documented:

`MMMS-package` *Multi-Marker Molecular Signature for Treatment-specific Subgroup Identification*

Description

The package implements a multi-marker molecular signature (MMMS) approach for treatment-specific subgroup identification. Time-to-event outcomes are currently supported, based on the approach described in Li et al. (2014). Other types of outcomes (e.g. binary and continuous outcomes) may be supported in future versions.

Details

Package:	MMMS
Type:	Package
Version:	0.1
Date:	2014-03-18
License:	GPL-3

The package contains one main function: `MMMS()`, which estimates composite scores for a multi-marker molecular signature (MMMS), identifies a subgroup based on the scores, and assesses the significance of the treatment-specific subgroup effect on the outcome of interest. `MMMS()` calls several functions that can also be called separately: `get.score()`, `get.score.main()`, `get.subgroup()`, etc.

Author(s)

Author: Lin Li, Tobias Guennel,Scott Marshall, Leo Wang-Kit Cheung

Contributors: Brigid M. Wilson, Dilan C. Paranagama

Maintainer: Lin Li <lli@biostatsolutions.com>

References

Lin Li, Tobias Guennel, Scott Marshall, Leo Wang-Kit Cheung (2014) A multi-marker molecular signature approach for treatment-specific subgroup identification with survival outcomes. *The Pharmacogenomics Journal*. http://dx.doi.org/10.1038/tpj.2014.9

`get.score` *Calculation of composite scores based on an interaction model*

Description

This function is to calculate composite scores of a multi-marker molecular signature based on an interaction model.

Usage

```
get.score(time, event, treat, bio, covar = NULL, nfolds = 5, alpha = 0.5,
          pos.direction = FALSE)
```

Arguments

time	A numeric vector containing the follow up time for right censored data.
event	A numeric vector containing the status indicator, normally 0=alive, 1=dead.
treat	A numeric vector containing the treatment indicator: 1=treatment of interest, 0=alternative treatment (e.g. placebo or standard of care).
bio	A numeric data frame or matrix containing biomarker values.
covar	A numeric matrix containing clinical covariates. Default is NULL for not including any covariates.
nfolds	The number of folds for cross validation in choosing tuning parameters. The function cv.glmnet() in the "glmnet" package is called, which requires cross validation to choose the tuning parameter "lambda". Default is 5.
alpha	A scalar for the elasticnet mixing parameter as in the "glmnet" package (0=ridge, 1=lasso). A fixed value is supposed to be used, without searching for the optimal alpha value. Default is 0.5.
pos.direction	A logical value indicating whether a subgroup with hazard ratio > 1 is desirable. Default is FALSE, i.e. a hazard ratio < 1 is desirable.

Details

This function is a function called by MMMS() to calculate MMMS composite scores. An interaction model is considered by assuming that a treatment-specific subgroup exists. The composite scores based on interaction terms and main-effect terms are both calculated via elastic net as implemented by the "glmnet" package. The composite scores based on interaction terms are used for identifying treatment-specific subgroups, while those based on main-effect terms are used for adjusting for biomarker main effect.

Value

A list with the following elements:

score	The composite scores based interaction terms for the treatment arm of interest (treat==1).
score.all	The composite scores based on interaction terms for all patients.
score.main	The composite scores based on main-effect terms.
coefs	Elnet coefficient estimates for interaction terms.
coefs.main	Elnet coefficient estimates for main-effect terms.
fit	The glmnet fitted object for obtaining the MMMS composite scores.
lam.best	The optimal lambda value chosen for obtaining the MMMS composite scores.
treat	The treatment variable in the input data.
alpha	The alpha value used for obtaining the MMMS composite scores.

Author(s)

Author: Lin Li, Tobias Guennel,Scott Marshall, Leo Wang-Kit Cheung

Contributors: Brigid M. Wilson, Dilan C. Paranagama

Maintainer: Lin Li <lli@biostatsolutions.com>

References

Lin Li, Tobias Guennel, Scott Marshall, Leo Wang-Kit Cheung (2014) A multi-marker molecular signature approach for treatment-specific subgroup identification with survival outcomes. *The Pharmacogenomics Journal*. http://dx.doi.org/10.1038/tpj.2014.9

See Also

MMMS

Examples

```
# load the dataset
data(simdat)
attach(simdat)

# get composite scores using an interaction model
score = get.score(time,event,treat,bio,covar,nfolds=5,alpha=0.5,
                  pos.direction=FALSE)
```

`get.score.main` *Calculation of composite scores based on a main-effect model*

Description

This function is to fit a main-effect model assuming no treatment-specific subgroups exist (under the null).

Usage

```
get.score.main(time, event, treat, bio, covar = NULL, nfolds = 5,
               alpha = 0.5)
```

Arguments

`time`	A numeric vector containing the follow up time for right censored data.
`event`	A numeric vector containing the status indicator, normally 0=alive, 1=dead.
`treat`	A numeric vector containing the treatment indicator: 1=treatment of interest, 0=alternative treatment (e.g. placebo or standard of care).
`bio`	A numeric data frame or matrix containing biomarker values.

covar	A numeric matrix containing clinical covariates. Default is NULL for not including any covariates.
nfolds	The number of folds for cross validation in choosing tuning parameters. The function cv.glmnet() in the "glmnet" package is called, which requires cross validation to choose the tuning parameter "lambda". Default is 5.
alpha	A scalar for the elasticnet mixing parameter as in the "glmnet" package (0=ridge, 1=lasso). A fixed value is supposed to be used, without searching for the optimal alpha value. Default is 0.5.

Details

This function is a function called by MMMS() to obtain bootstrap-based p-values. A main-effect model is considered by assuming that no treatment-specific subgroups exist. This function is used for obtaining (semi)parametric bootstrap samples under the null.

Value

A list with the following elements:

fit	The glmnet fitted object assuming no subgroups exist.
lam.best	The optimal lambda value chosen when assuming no subgroups exist.
fit.selected	An object returned by coxph() using selected biomarkers when assuming no subgroups exist.
sfit	An object returned by survfit() for bootstrap sampling.

Author(s)

Author: Lin Li, Tobias Guennel,Scott Marshall, Leo Wang-Kit Cheung

Contributors: Brigid M. Wilson, Dilan C. Paranagama

Maintainer: Lin Li <lli@biostatsolutions.com>

References

Lin Li, Tobias Guennel, Scott Marshall, Leo Wang-Kit Cheung (2014) A multi-marker molecular signature approach for treatment-specific subgroup identification with survival outcomes. *The Pharmacogenomics Journal*. http://dx.doi.org/10.1038/tpj.2014.9

See Also

MMMS, get.score

Examples

```
# load the dataset
data(simdat)
attach(simdat)

# get composite scores using a main-effect model
main.only=get.score.main(time,event,treat,bio,covar,nfolds=5,alpha=0.5)
```

`get.subgroup` *Searching for a treatment-specific subgroup*

Description

Searching for a treatment-specific subgroup based on MMMS composite scores.

Usage

```
get.subgroup(score.info, data, lb = 20, ub = 80)
```

Arguments

`score.info`	The object returned by `get.score()`.
`data`	A list containing input data used for calculating composite scores. The list should have at the least the following elements: `$time`, `$event`, `$treat`, and `$covar`, which are the same as those used for calculating the composite scores.
`lb`	A scalar indicating the lower bound of the search range for desired subgroup sizes in percentage (e.g. 20 means 20%).
`ub`	A scalar indicating the lower bound of the search range for desired subgroup sizes in percentage (e.g. 80 means 80%).

Details

This function is a function called by `MMMS()` to search for an optimal treatment-specific subgroup. A range of desired subgroup sizes can be specified. When no subgroup can be found within the range, values of `NA` will be returned.

Value

A list with the following elements:

`pct`	All possible subgroup sizes within the desired range.
`cutoff`	All possible cutoffs for the composite scores (based on interaction terms) that give subgroup sizes within the desired range.
`i.best`	The index with respect to `pct` and `cutoff` that corresponds to the optimal subgroup.
`fit.best`	The fitted model based on the subgroup defined by the optimal cutoff.
`pct.best`	The subgroup size of the subgroup defined by the optimal cutoff.
`cutoff.best`	The optimal cutoff.
`chisq.best`	The chi-square statistic for the treatment-by-subgroup interaction for the subgroup defined by the optimal cutoff.

Author(s)

Author: Lin Li, Tobias Guennel,Scott Marshall, Leo Wang-Kit Cheung

Contributors: Brigid M. Wilson, Dilan C. Paranagama

Maintainer: Lin Li <lli@biostatsolutions.com>

References

Lin Li, Tobias Guennel, Scott Marshall, Leo Wang-Kit Cheung (2014) A multi-marker molecular signature approach for treatment-specific subgroup identification with survival outcomes. *The Pharmacogenomics Journal*. http://dx.doi.org/10.1038/tpj.2014.9

See Also

MMMS, get.score

Examples

```
# load the dataset
data(simdat)
attach(simdat)

# get composite scores using an interaction model
score = get.score(time,event,treat,bio,covar,nfolds=5,alpha=0.5,
                  pos.direction=FALSE)

# get subgroup
subgrp = get.subgroup(score,simdat,lb=20,ub=80)
```

MMMS *Subgroup identification using a multi-marker molecular signature approach*

Description

Identification of a treatment-specific subgroup for time-to-event outcomes using a multi-marker molecular signature approach

Usage

```
MMMS(time, event, treat, bio, covar = NULL, pct.lb = 20, pct.ub = 80, n.boot = 1000,
    pos.direction = FALSE, nfolds = 5, alpha = 0.5, verbose = FALSE, seed = NULL)
```

Arguments

time	A numeric vector containing the follow up time for right censored data.
event	A numeric vector containing the status indicator, normally 0=alive, 1=dead.
treat	A numeric vector containing the treatment indicator: 1=treatment of interest, 0=alternative treatment (e.g. placebo or standard of care).
bio	A numeric data frame or matrix containing biomarker values.
covar	A numeric matrix containing clinical covariates. Default is NULL for not including any covariates.
pct.lb	A scalar indicating the lower bound of the search range for desired subgroup sizes in percentage (e.g. 20 means 20%).
pct.ub	A scalar indicating the upper bound of the search range for desired subgroup sizes in percentage (e.g. 80 means 80%).
n.boot	A scalar indicating the number of bootstraps for calculating the bootstrap p-value for the subgroup effect. Default is 1000, which is a recommended value.
pos.direction	A logical value indicating whether a subgroup with hazard ratio > 1 is desirable. Default is FALSE, i.e. a hazard ratio < 1 is desirable.
nfolds	The number of folds for cross validation in choosing tuning parameters. The function cv.glmnet() in the "glmnet" package is called, which requires cross validation to choose the tuning parameter "lambda". Default is 5.
alpha	A scalar for the elasticnet mixing parameter as in the "glmnet" package (0=ridige, 1=lasso). A fixed value is supposed to be used, without searching for the optimal alpha value. Default is 0.5.
verbose	A logical value indicating whether bootstrap progress should be printed. Default is FALSE.
seed	An integer for setting random seed, if provided. Default is NULL for not setting any seed.

Details

MMMS() calls several functions that could also be used separately: get.score(), get.score.main(), get.subgroup(), etc.

As is described in Li et al. (2014), the bootstrap p-value is based on a statistically valid test whose type I error is approximately controlled at the nominal level. However, caution is needed for interpreting the estimates of subgroup size and treatment-by-subgroup interaction effect, as bias has been observed in these estimates. Approaches for correcting bias in the estimates may be implemented in future versions of the "MMMS" package.

Value

A list with the following elements:

score.obj	The object returned by get.score().
score	The composite scores based on interaction terms.
score.main	The composite scores based on main-effect terms.

coefs	Elnet coefficient estimates for interaction terms.
coefs.main	Elnet coefficient estimates for main-effect terms.
fit	The glmnet fitted object for obtaining the MMMS composite scores.
lambda	The optimal lambda value chosen for obtaining the MMMS composite scores.
alpha	The alpha value used for obtaining the MMMS composite scores.
subgrp.obj	The object returned by get.subgroup().
subgrp.size	The size (in percentage) of the optimal subgroup identified.
subgrp.fit	The fitted model object for the optimal subgroup identified.
subgrp.cut	The cutpoint of the composite score score for the optimal subgroup identified.
subgrp.pval	The p-value of the treatment-by-subgroup effect based on n.boot bootstraps. NA is returned if n.boot=0.
n.boot	The number of bootstraps considered for calculating subgrp.pval.

Author(s)

Author: Lin Li, Tobias Guennel,Scott Marshall, Leo Wang-Kit Cheung

Contributors: Brigid M. Wilson, Dilan C. Paranagama

Maintainer: Lin Li <lli@biostatsolutions.com>

References

Lin Li, Tobias Guennel, Scott Marshall, Leo Wang-Kit Cheung (2014) A multi-marker molecular signature approach for treatment-specific subgroup identification with survival outcomes. *The Pharmacogenomics Journal*. http://dx.doi.org/10.1038/tpj.2014.9

See Also

get.score, get.subgroup.

Examples

```
# load the dataset
data(simdat)

# estimate the MMMS (No bootstrap is considered for a quick illustration)
mmms = with(simdat,MMMS(time,event,treat,bio,covar,pct.lb=20,pct.ub=80,n.boot=0,
                        pos.direction=FALSE,nfolds=5,alpha=0.5,verbose=TRUE,seed=12345))
```

simdat *An example dataset*

Description

An example dataset with time-to-event outcomes, biomarker values and covariates.

Usage

```
data(simdat)
```

Format

The data contains a list of length 9:

`time`: Time variable.

`event`: Event variable.

`treat`: Treatment variable.

`cont`: Biomarker data – continuous values.

`burden`: Biomarker data – burden scores of mutations.

`snps`: Biomarker data – genotypes of single nucleotide polymorphisms.

`covar`: covariates.

`sub.true`: true subgroup size.

`bio`: combined biomarker data (`cbind(cont,burden,snps)`).

References

Lin Li, Tobias Guennel, Scott Marshall, Leo Wang-Kit Cheung (2014) A multi-marker molecular signature approach for treatment-specific subgroup identification with survival outcomes. *The Pharmacogenomics Journal*. http://dx.doi.org/10.1038/tpj.2014.9

Examples

```
data(simdat)
```

Index