Package 'RiskStratifiedEstimation'

October 4, 2018

Type Package

Title Risk Stratified Effect Estimation

Version 0.1.0

Author Alexandros Rekkas

Maintainer Alexandros Rekkas <a.rekkas@lumc.nl>

Description
The package performs risk stratified analysis of treatment effect based on the common data model.

License Apache License 2.0

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Imports PatientLevelPrediction,
CohortMethod,
Cyclops,
OhdsiRTools,
foreach,
FeatureExtraction,
dplyr,
ff,
caret,
ffbase,
ggplot2,
gridExtra,
tidyr,
doSNOW,
survminer

Depends survival,

Suggests knitr,
rmarkdown

VignetteBuilder knitr

R topics documented:

`absoluteRiskReduction` *Creates a data frame of absolute risk reductions within risk strata*

Description

Creates a data frame of absolute risk reductions within risk strata

Usage

```
absoluteRiskReduction(dataKM, timePoint)
```

Arguments

`dataKM`	A list with the Kaplan-Meier estimates within risk strata
`timePoint`	The time point based on which the absolute risk reductions will be calculated

Value

A data frame with the absolute risk reductions along with confidence intervals

---

`cmToPlpData` *Transforms a cohortMethodData object into a plpData object*

Description

Transforms a cohortMethodData object, as generated from `getDbCohortMethodData`, into a plpData object appropriate for risk stratified analysis

Usage

```
cmToPlpData(cohortMethodData)
```

Arguments

`cohortMethodData`	The cohortMethodData object to be transformed

Value

A plpData object

`comparisonPlot` *Generates a comparison plot within risk quantiles*

Description

Generates a plot with relative and absolute risk reductions. It can handle one or multiple outcomes. In the case of multiple outcomes the input must be a named list.

Usage

```
comparisonPlot(dataARR, dataRRR, cases, treatmentLabel, comparatorLabel,
  ylimARR, ylimRRR, ylimCases = c(0, 1), position_dodge.w = 0.2)
```

Arguments

`dataARR`	A list or dataframe same as the one generated by `absoluteRiskReduction`
`dataRRR`	A list or dataframe same as the one generated by `relativeRiskReduction`
`cases`	The dataframe or list with the outcome rates within risk strata. In the case of multiple outcomes the names of the list will be used as outcome labels
`treatmentLabel`	The name of the treatment drug
`comparatorLabel`	The name of the comparator drug
`ylimARR`	The limits on the y-axis of the sbsolute risk reduction plot
`ylimRRR`	The limits on the y-axis of the hazard ratio plot
`ylimCases`	The limits on the y-axis of the outcome rates bar plot
`position_dodge.w`	The amount of space between bullets in the ploit in the case of multiple outcomes

Value

A plot with hazard ratios and absolute risk reductions across risk strata

---

`createIPW` *Creates Inverse Probability Weights*

Description

Calcuates inverse probability weights based on the propensity score

Usage

```
createIPW(ps, weightsType = "ATE", useStabilizedWeights = TRUE,
  truncationLevels = c(0.01, 0.99))
```

Arguments

ps	A propensity score data frame as created from createPs
weightsType	The type of the weights to be used. Allowed options are 'ATE' for average treatment effect and 'ATT' for average treatment effect on the treated weights
useStabilizedWeights	Should stabilized weights be used?
truncationLevels	The level of truncation expressed in percentiles of the propensity score.

Value

The ps data frame provided as input along with a weights column

---

createStudyPopulation *Create a study population*

Description

Create a study population

Usage

```
createStudyPopulation(cohortMethodData, population = NULL, outcomeId,
  firstExposureOnly = FALSE, restrictToCommonPeriod = FALSE,
  washoutPeriod = 0, removeDuplicateSubjects = FALSE,
  removeSubjectsWithPriorOutcome = TRUE, priorOutcomeLookback = 99999,
  minDaysAtRisk = 1, riskWindowStart = 0, addExposureDaysToStart = FALSE,
  riskWindowEnd = 0, addExposureDaysToEnd = TRUE,
  censorAtNewRiskWindow = FALSE)
```

Arguments

cohortMethodData	An object of type cohortMethodData as generated using getDbCohortMethodData.
population	If specified, this population will be used as the starting point instead of the cohorts in the cohortMethodData object.
outcomeId	The ID of the outcome. If not specified, no outcome-specific transformations will be performed.
firstExposureOnly	Should only the first exposure per subject be included? Note that this is typically done in the createStudyPopulation function,
restrictToCommonPeriod	Restrict the analysis to the period when both treatments are observed?
washoutPeriod	The mininum required continuous observation time prior to index date for a person to be included in the cohort.
removeDuplicateSubjects	Remove subjects that are in both the treated and comparator cohort? See details for allowed values.

`removeSubjectsWithPriorOutcome`
: Remove subjects that have the outcome prior to the risk window start?

`priorOutcomeLookback`
: How many days should we look back when identifying prior outcomes?

`minDaysAtRisk`
: The minimum required number of days at risk.

`riskWindowStart`
: The start of the risk window (in days) relative to the index date (+ days of exposure if the `addExposureDaysToStart` parameter is specified).

`addExposureDaysToStart`
: Add the length of exposure the start of the risk window?

`riskWindowEnd`
: The end of the risk window (in days) relative to the index data (+ days of exposure if the `addExposureDaysToEnd` parameter is specified).

`addExposureDaysToEnd`
: Add the length of exposure the risk window?

`censorAtNewRiskWindow`
: If a subject is in multiple cohorts, should time-at-risk be censored when the new time-at-risk starts to prevent overlap?

Details

Create a study population by enforcing certain inclusion and exclusion criteria, defining a risk window, and determining which outcomes fall inside the risk window.

The `removeduplicateSubjects` argument can have one of the following values:

"keep all" Do not remove subjects that appear in both target and comparator cohort

"keep first" When a subjects appear in both target and comparator cohort, only keep whichever cohort is first in time.

"remove all" Remove subjects that appear in both target and comparator cohort completely from the analysis."

Value

A data frame specifying the study population. This data frame will have the following columns:

rowId A unique identifier for an exposure

subjectId The person ID of the subject

cohortStartdate The index date

outcomeCount The number of outcomes observed during the risk window

timeAtRisk The number of days in the risk window

survivalTime The number of days until either the outcome or the end of the risk window

plotCovariateBalance *Plots the covariate balance before and after balancing*

Description

Plots covariate before and after weighting using the inverse of the propensity score

Usage

```
plotCovariateBalance(ps, cohortMethodData, calculateWeights = TRUE,
  weightsType = "ATE", useStabilizedWeights = TRUE, truncationLevels,
  showNotBalancedCovariateIds = TRUE)
```

Arguments

ps	A propensity score data frame as created from createPs
cohortMethodData	A cohortMethodData object
calculateWeights	Should the weights be calculated?
weightsType	The type of the weights to be used. Allowed options are 'ATE' for average treatment effect and 'ATT' for average treatment effect on the treated weights
useStabilizedWeights	Should stabilized weights be used?
truncationLevels	The level of truncation expressed in percentiles of the propensity score.
showNotBalancedCovariateIds	Show covariate ids that were not balanced after weighting?

Value

The covariate balance plot

plotWeightedKM *Plots the weighted Kaplan-Meier estimate*

Description

Plots the weighted Kaplan-Meier estimate

Usage

```
plotWeightedKM(dataKM, xlim, ylim, ci = TRUE, title = NULL)
```

Arguments

dataKM	A dataframe containing the weigthed Kaplan-Meier estimates as computed by weightedKM
xlim	Limits on x-axis
ylim	Limits on y-axis
ci	Should confidence intervals be displayed?
title	The title on the graph

Value

A weighted Kaplan-Meier plot

---

relativeRiskReduction *Creates a data frame with hazard ratios within risk strata*

Description

Creates a data frame with hazard ratios within risk strata

Usage

```
relativeRiskReduction(ps, calculateWeights = TRUE, weightsType = "ATE",
  useStabilizedWeights = TRUE, truncationLevels)
```

Arguments

ps	A list with the propensity data frames as generated from createPs
calculateWeights	Whether to calculate the weights using createIPW
weightsType	The type of the weights to be used. Allowed options are 'ATE' for average treatment effect and 'ATT' for average treatment effect on the treated weights
useStabilizedWeights	Should stabilized weights be used?
truncationLevels	The level of truncation expressed in percentiles of the propensity score.

Value

A data frame with hazard ratios along with confidence intervals

`runRiskStratifiedEstimation` *Runs a risk stratified analysis*

Description

Runs a risk stratified analysis

Usage

```
runRiskStratifiedEstimation(cohortMethodData, population, modelSettings, save,
  testSplit = "person", testFraction = 0.3, nfold = 10, riskStrata = 4,
  weightsType = "ATE", useStabilizedWeights = TRUE, truncationLevels,
  timePoint, excludeCovariateIds = NULL, binary = TRUE,
  includeAllOutcomes = TRUE, requireTimeAtRisk = TRUE,
  savePlpPlots = FALSE, psThreads = 1, priorType = "laplace",
  verbosity = "INFO", analysisId = NULL, savePlpResult = TRUE,
  saveMapMatrix = TRUE, savePs = TRUE, saveDataKM = TRUE,
  saveAbsoluteRiskRreduction = TRUE, saveRelativeRiskReduction = TRUE)
```

Arguments

`cohortMethodData`	A cohortMethodData object
`population`	The study population to perform the analysis
`modelSettings`	The model settings for the prediction step
`save`	The save directory
`testSplit`	The type of split for the cross validation. Should be either 'person' or 'time'
`testFraction`	The size of the test set
`nfold`	The number of folds for cross validation
`riskStrata`	The number of risk strata on which to perform the analysis
`weightsType`	The type of weights for the balancing of covariates. Should be either 'ATE' or 'ATT'
`useStabilizedWeights`	Should stabilized weights be used?
`truncationLevels`	The level of truncation expressed in percentiles of the propensity score.
`timePoint`	The time point of interest for the calculation of the absolute risk reduction
`excludeCovariateIds`	Covariate Ids to be excluded from calculation of propensity scores
`binary`	Forces the outcomeCount to be 0 or 1 in the prediction step
`includeAllOutcomes`	(binary) indicating whether to include people with outcomes who are not observed for the whole at risk period
`requireTimeAtRisk`	Should subjects without time at risk be removed at the prediction step?
`savePlpPlots`	(binary) Should plots for the prediction step be generated?

psThreads	The number of cores to use for the estimation of the propensity score. If 1 then serial approach is implemented
priorType	The prior for the propensity score model
verbosity	Sets the level of the verbosity. If the log level is at or higher in priority than the logger threshold, a message will print. The levels are: • DEBUGHighest verbosity showing all debug statements • TRACEShowing information about start and end of steps • INFOShow informative information (Default) • WARNShow warning messages • ERRORShow error messages • FATALBe silent except for fatal errors
analysisId	Identifier of the analysis
savePlpResult	Should the prediction result be saved?
saveMapMatrix	Should the map matrix with the risk sratum allocations be saved?
savePs	Should the propensity scores be saved?
saveDataKM	Should the weighted Kaplan-Meier estimates be saved?
saveAbsoluteRiskRreduction	Should the absolute risk reduction estimates be saved?
saveRelativeRiskReduction	Should the hazard ratios be saved?

Value

ps	The propensity scores within risk strata along with patient weights
mapMatrix	The matrix that maps the patients to risk strata
dataKM	The weighted Kaplan-Meier estimates within risk strata
absoluteRiskReduction	The absolute risk reduction within risk strata
relativeRiskReduction	The relative risk reduction within risk strata
predictionResult	The result of the prediction step

---

weightedKM *Calculates and plots weighted Kaplan-Meier estimates*

Description

Calculates the weighted Kaplan-Meier estimates based on: Xie J, Liu C. Adjusted Kaplan-Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. Statistics in Medicine 2005; 2:3089–3110.

Usage

```
weightedKM(ps, calculateWeights = TRUE, weightsType = "ATE",
  useStabilizedWeights = TRUE, truncationLevels)
```

Arguments

ps	A data frame including the propensity scores as generated from createPs
calculateWeights	Whether to calculate the weights using createIPW
weightsType	The type of the weights to be used. Allowed options are 'ATE' for average treatment effect and 'ATT' for average treatment effect on the treated weights
useStabilizedWeights	Should stabilized weights be used?
truncationLevels	The level of truncation expressed in percentiles of the propensity score.

Value

A data frame with the Kaplan-Meier estimates

Index