Package 'SelfControlledCaseSeries'

March 22, 2021

Type Package

```
Title Self-Controlled Case Series
Version 3.0.0
Date 2021-03-22
Maintainer Martijn Schuemie <schuemie@ohdsi.org>
Description SelfControlledCaseSeries is an R package for performing self-
      controlled case series (SCCS) analyses in an observational database in the OMOP
      Common Data Model. It extracts all necessary data from the database and
      transforms it to the format required for SCCS. Age and season can be modeled
      using splines assuming constant hazard within calendar months. Event-dependent
      censoring of the observation period can be corrected for. Many exposures can be
      included at once (MSCCS), with regularization on all coefficients except for the
      exposure of interest.
VignetteBuilder knitr
URL https://github.com/OHDSI/SelfControlledCaseSeries
BugReports https://github.com/OHDSI/SelfControlledCaseSeries/issues
Depends R (>= 3.2.2),
      Cyclops (>= 3.0.0),
      DatabaseConnector (>= 4.0.0),
      Andromeda
Imports SqlRender (>= 1.6.0),
      dplyr (>= 1.0.0),
      Rcpp (>= 0.11.2),
      ParallelLogger (\geq 2.0.0),
      splines,
      ggplot2,
      methods,
      cli,
      pillar
Suggests testthat,
      knitr,
      rmarkdown,
      EmpiricalCalibration
Remotes ohdsi/EmpiricalCalibration
License Apache License 2.0
```

LinkingTo Rcpp
NeedsCompilation yes
RoxygenNote 7.1.1
$\textbf{Roxygen} \ \ list(markdown = TRUE)$
Encoding UTF-8

R topics documented:

. 3/1	2
omputeMdrr	3
reateAgeCovariateSettings	4
reateControlIntervalSettings	4
reateCreateSccsIntervalDataArgs	5
reateCreateScriIntervalDataArgs	6
reateCreateStudyPopulationArgs	7
reateEraCovariateSettings	7
reateExposureOutcome	9
reateFitSccsModelArgs	9
ϵ	10
	11
	12
ϵ	13
	14
Ç	15
	16
	17
yclicSplineDesign	18
tSccsModel	18
etAttritionTable	19
etDbSccsData	19
etModel	22
SccsData	23
SccsIntervalData	23
	24
	24
	25
	25
	26
	26
e i	27
•	28
	28
	29
·	29
	32
	32
·	32 33
	33
	33 34
	34 34
	34 35
ummarizeSccsAnalyses	35

computeMdrr 3

Index 36

computeMdrr

Compute the minimum detectable relative risk

Description

Compute the minimum detectable relative risk

Usage

```
computeMdrr(
  sccsIntervalData,
  exposureCovariateId,
  alpha = 0.05,
  power = 0.8,
  twoSided = TRUE,
  method = "binomial"
)
```

Arguments

sccsIntervalData

An object of type SccsIntervalData as created using the createSccsIntervalData

function.

exposureCovariateId

Covariate Id for the health exposure of interest.

alpha Type I error.

power 1 - beta, where beta is the type II error.

twoSided Consider a two-sided test?

method The type of sample size formula that will be used. Allowable values are "pro-

portion", "binomial", "SRL1", "SRL2", or "ageEffects". Currently "ageEffects"

is not supported.

Details

Compute the minimum detectable relative risk (MDRR) for a given study population, using the observed time at risk and total time in days and number of events. Five sample size formulas are implemented: sampling proportion, binomial proportion, 2 signed root likelihood ratio methods, and likelihood extension for age effects. The expressions by Musonda (2006) are used.

Value

A data frame with the MDRR, number of events, time at risk, and total time.

References

Musonda P, Farrington CP, Whitaker HJ (2006) Samples sizes for self-controlled case series studies, Statistics in Medicine, 15;25(15):2618-31

 ${\tt createAgeCovariateSettings}$

Create age covariate settings

Description

Create age covariate settings

Usage

```
createAgeCovariateSettings(
  ageKnots = 5,
  allowRegularization = FALSE,
  computeConfidenceIntervals = FALSE
)
```

Arguments

ageKnots

If a single number is provided this is assumed to indicate the number of knots to use for the spline, and the knots are automatically spaced according to equal percentiles of the data. If more than one number is provided these are assumed to be the exact location of the knots in age-days

allowRegularization

When fitting the model, should the covariates defined here be allowed to be regularized?

computeConfidenceIntervals

Should confidence intervals be computed for the covariates defined here? Setting this to FALSE might save computing time when fitting the model. Will be turned to FALSE automatically when allowRegularization = TRUE.

Details

Create an object specifying whether and how age should be included in the model. Age can be included by splitting patient time into calendar months. During a month, the relative risk attributed to age is assumed to be constant, and the risk from month to month is modeled using a cubic spline.

Value

An object of type AgeCovariateSettings.

 ${\tt createControlIntervalSettings}$

Create control interval settings

Description

Create control interval settings

Usage

```
createControlIntervalSettings(
  includeEraIds = NULL,
  excludeEraIds = NULL,
  start = 0,
  startAnchor = "era start",
  end = 0,
  endAnchor = "era end",
  firstOccurrenceOnly = FALSE
)
```

Arguments

includeEraIds One or more IDs of variables in the SccsData object that should be used to

construct this covariate. If no IDs are specified, all variables will be used.

excludeEraIds One or more IDs of variables in the [SccsData] object that should not be used to

construct this covariate.

start The start of the control interval (in days) relative to the startAnchor.

startAnchor The anchor point for the start of the control interval. Can be "era start" or

"era end".

end The end of the control interval (in days) relative to the endAnchor.

endAnchor The anchor point for the end of the control interval. Can be "era start" or

"era end".

firstOccurrenceOnly

Should only the first occurrence of the exposure be used?

Details

Create an object specifying how to create a control interval for the self-controlled risk interval (SCRI) design.

Value

An object of type ControlSettings.

createCreateSccsIntervalDataArgs

 ${\it Create~a~parameter~object~for~the~function~create Sccs Interval Data}$

Description

Create a parameter object for the function createSccsIntervalData

Usage

```
createCreateSccsIntervalDataArgs(
  eraCovariateSettings,
  ageCovariateSettings = NULL,
  seasonalityCovariateSettings = NULL,
  minCasesForAgeSeason = 10000,
  eventDependentObservation = FALSE
)
```

Arguments

eraCovariateSettings

Either an object of type EraCovariateSettings as created using the createEraCovariateSettings() function, or a list of such objects.

ageCovariateSettings

An object of type ageCovariateSettings as created using the createAgeCovariateSettings() function.

seasonalityCovariateSettings

An object of type seasonalityCovariateSettings as created using the createSeasonalityCovariateSettings() function.

minCasesForAgeSeason

Minimum number of cases to use to fit age and season splines. If needed (and available), cases that are not exposed will be included.

eventDependentObservation

Should the extension proposed by Farrington et al. be used to adjust for event-dependent observation time?

Details

Create an object defining the parameter values.

create Create ScriInterval Data Args

Create a parameter object for the function createScriIntervalData

Description

Create a parameter object for the function createScriIntervalData

Usage

create Create ScriInterval Data Args (era Covariate Settings, control Interval Settings)

Arguments

eraCovariateSettings

Either an object of type EraCovariateSettings as created using the createEraCovariateSettings() function, or a list of such objects.

controlIntervalSettings

An object of type ControlIntervalSettings as created using the createControlIntervalSettings() function.

Details

Create an object defining the parameter values.

createCreateStudyPopulationArgs

Create a parameter object for the function createStudyPopulation

Description

Create a parameter object for the function createStudyPopulation

Usage

```
createCreateStudyPopulationArgs(
  firstOutcomeOnly = FALSE,
  naivePeriod = 0,
  minAge = NULL,
  maxAge = NULL
)
```

Arguments

firstOutcomeOnly

Whether only the first occurrence of an outcome should be considered.

naivePeriod

The number of days at the start of a patient's observation period that should not be included in the risk calculations. Note that the naive period can be used to determine current covariate status right after the naive period, and whether an

outcome is the first one.

minAge

Minimum age at which patient time will be included in the analysis. Note that information prior to the min age is still used to determine exposure status after the minimum age (e.g. when a prescription was started just prior to reaching the minimum age). Also, outcomes occurring before the minimum age is reached will be considered as prior outcomes when using first outcomes only. Age should be specified in years, but non-integer values are allowed. If not

specified, no age restriction will be applied.

maxAge

Maximum age at which patient time will be included in the analysis. Age should be specified in years, but non-integer values are allowed. If not specified, no age restriction will be applied.

Details

Create an object defining the parameter values.

createEraCovariateSettings

Create era covariate settings

Description

Create era covariate settings

Usage

```
createEraCovariateSettings(
  includeEraIds = NULL,
  excludeEraIds = NULL,
  label = "Covariates",
  stratifyById = TRUE,
  start = 0,
  startAnchor = "era start",
  end = 0,
  endAnchor = "era end",
  firstOccurrenceOnly = FALSE,
  splitPoints = c(),
  allowRegularization = FALSE
)
```

Arguments

includeEraIds One or more IDs of variables in the SccsData object that should be used to

construct this covariate. If no IDs are specified, all variables will be used.

excludeEraIds One or more IDs of variables in the [SccsData] object that should not be used to

construct this covariate.

label A label used to identify the covariates created using these settings.

stratifyById Should a single covariate be created for every ID in the SccsData object, or

should a single covariate be constructed? For example, if the IDs identify exposures to different drugs, should a covariate be constructed for every drug, or a single covariate for exposure to any of these drugs. Note that overlap will be

considered a single exposure.

start The start of the risk window (in days) relative to the startAnchor.

startAnchor The anchor point for the start of the risk window. Can be "era start" or "era

end".

end The end of the risk window (in days) relative to the endAnchor.

endAnchor The anchor point for the end of the risk window. Can be "era start" or "era

end".

firstOccurrenceOnly

Should only the first occurrence of the exposure be used?

splitPoints To split the risk window into several smaller windows, specify the end of each

sub- window relative to the start of the main risk window. If add Exposed-DaysToStart is TRUE, the split points will be considered to be relative to the

end of the main risk window instead.

allowRegularization

When fitting the model, should the covariates defined here be allowed to be regularized?

Details

Create an object specifying how to create a (set of) era-based covariates.

Value

An object of type EraCovariateSettings.

createExposureOutcome Create a exposure-outcome combination.

Description

Create a exposure-outcome combination.

Usage

```
createExposureOutcome(exposureId, outcomeId, ...)
```

Arguments

exposureId A concept ID identifying the target drug in the exposure table. If multiple strate-

gies for picking the exposure will be tested in the analysis, a named list of numbers can be provided instead. In the analysis, the name of the number to be used can be specified using the exposureType parameter in the createSccsAnalysis

function.

outcomeId A concept ID identifying the outcome in the outcome table.

... Custom variables, to be used in the analyses.

Details

Create a set of hypotheses of interest, to be used with the runSccsAnalyses function.

```
createFitSccsModelArgs
```

Create a parameter object for the function fitSccsModel

Description

Create a parameter object for the function fitSccsModel

Usage

```
createFitSccsModelArgs(
  prior = createPrior("laplace", useCrossValidation = TRUE),
  control = createControl(cvType = "auto", selectorType = "byPid", startingVariance =
     0.1, noiseLevel = "quiet")
)
```

Arguments

prior The prior used to fit the model. See Cyclops::createPrior for details.

control The control object used to control the cross-validation used to determine the

hyperparameters of the prior (if applicable). See Cyclops::createControl for de-

tails.

Details

Create an object defining the parameter values.

createGetDbSccsDataArgs

Create a parameter object for the function getDbSccsData

Description

Create a parameter object for the function getDbSccsData

Usage

```
createGetDbSccsDataArgs(
  useCustomCovariates = FALSE,
  useNestingCohort = FALSE,
  nestingCohortId = NULL,
  deleteCovariatesSmallCount = 100,
  studyStartDate = "",
  studyEndDate = "",
  maxCasesPerOutcome = 0,
  exposureIds = "exposureId",
  customCovariateIds = ""
```

Arguments

useCustomCovariates

Create covariates from a custom table?

useNestingCohort

Should the study be nested in a cohort (e.g. people with a specific indication)? If not, the study will be nested in the general population.

nestingCohortId

A cohort definition ID identifying the records in the nestingCohortTable to use as nesting cohort.

deleteCovariatesSmallCount

The minimum count for a covariate to appear in the data to be kept.

studyStartDate A calendar date specifying the minimum date where data is used. Date format is 'yyyymmdd'.

studyEndDate A calendar date specifying the maximum date where data is used. Date format is 'yyyymmdd'.

maxCasesPerOutcome

If there are more than this number of cases for a single outcome cases will be sampled to this size. maxCasesPerOutcome = 0 indicates no maximum size.

exposureIds

A list of identifiers to define the exposures of interest. If exposureTable = DRUG_ERA, exposureIds should be CONCEPT_ID. If exposureTable <> DRUG_ERA, exposureIds is used to select the cohort_concept_id in the cohort-like table. If no exposureIds are provided, all drugs or cohorts in the exposureTable are included as exposures.

customCovariateIds

A list of cohort definition IDS identifying the records in the customCovariateTable to use for building custom covariates.

createSccsAnalysis 11

Details

Create an object defining the parameter values.

createSccsAnalysis

Create a SelfControlledCaseSeries analysis specification

Description

Create a SelfControlledCaseSeries analysis specification

Usage

```
createSccsAnalysis(
  analysisId = 1,
  description = "",
  exposureType = NULL,
  outcomeType = NULL,
  getDbSccsDataArgs,
  createStudyPopulationArgs,
  design = "SCCS",
  createSccsIntervalDataArgs = NULL,
  fitSccsModelArgs
)
```

Arguments

analysisId An integer that will be used later to refer to this specific set of analysis choices.

description A short description of the analysis.

exposureType If more than one exposure is provided for each exposureOutcome, this field

should be used to select the specific exposure to use in this analysis.

 $\hbox{outcomeType} \qquad \hbox{ If more than one outcome is provided for each $\tt exposureOutcome$, this field}$

should be used to select the specific outcome to use in this analysis.

getDbSccsDataArgs

An object representing the arguments to be used when calling the getDbSccs-Data function.

 ${\tt createStudyPopulationArgs}$

An object representing the arguments to be used when calling the getDbSccs-

Data function.

design Either "SCCS" for the general self-controlled case series design, or "SCRI" for

the self-controlled risk interval design.

createSccsIntervalDataArgs

An object representing the arguments to be used when calling the createSccsIntervalData function. Ignored when design = "SCRI".

create ScriInterval Data Args

An object representing the arguments to be used when calling the createScriIntervalData function. Ignored when design = "SCCS".

fitSccsModelArgs

An object representing the arguments to be used when calling the fitSccsModel function.

12 createSccsIntervalData

Details

Create a set of analysis choices, to be used with the runSccsAnalyses function.

createSccsIntervalData

Create SCCS era data

Description

Create SCCS era data

Usage

```
createSccsIntervalData(
   studyPopulation,
   sccsData,
   eraCovariateSettings,
   ageCovariateSettings = NULL,
   seasonalityCovariateSettings = NULL,
   minCasesForAgeSeason = 10000,
   eventDependentObservation = FALSE
)
```

Arguments

studyPopulation

An object created using the createStudyPopulation() function.

sccsData An object of type SccsData as created using the getDbSccsData function.

eraCovariateSettings

Either an object of type EraCovariateSettings as created using the createEraCovariateSettings function, or a list of such objects.

ageCovariateSettings

An object of type ageCovariateSettings as created using the createAgeCovariateSettings() function.

seasonality Covariate Settings

An object of type seasonalityCovariateSettings as created using the createSeasonalityCovar function.

minCasesForAgeSeason

Minimum number of cases to use to fit age and season splines. If needed (and available), cases that are not exposed will be included.

eventDependentObservation

Should the extension proposed by Farrington et al. be used to adjust for event-dependent observation time?

Details

This function creates covariates based on the data in the sccsData argument, according to the provided settings. It chops patient time into periods during which all covariates remain constant. The output details these periods, their durations, and a sparse representation of the covariate values.

Value

An object of type SccsIntervalData.

References

Farrington, C. P., Anaya-Izquierdo, A., Whitaker, H. J., Hocine, M.N., Douglas, I., and Smeeth, L. (2011). Self-Controlled case series analysis with event-dependent observation periods. Journal of the American Statistical Association 106 (494), 417-426

 ${\tt createSccsSimulationSettings}$

Create SCCS simulation settings

Description

Create SCCS simulation settings

Usage

```
createSccsSimulationSettings(
  meanPatientTime = 4 * 365,
  sdPatientTime = 2 * 365,
  minAge = 18 * 365,
  maxAge = 65 * 365,
  minBaselineRate = 0.001,
  maxBaselineRate = 0.01,
  eraIds = c(1, 2),
  patientUsages = c(0.2, 0.1),
  usageRate = c(0.01, 0.01),
  meanPrescriptionDurations = c(14, 30),
  sdPrescriptionDurations = c(7, 14),
  simulationRiskWindows = list(createSimulationRiskWindow(relativeRisks = 1),
    createSimulationRiskWindow(relativeRisks = 1.5)),
  includeAgeEffect = TRUE,
  ageKnots = 5,
  includeSeasonality = TRUE,
  seasonKnots = 5,
  outcomeId = 10
)
```

Arguments

meanPatientTime

Mean number of observation days per patient.

sdPatientTime Standard deviation of the observation days per patient.

minAge The minimum age in days.

maxAge The maximum age in days.

minBaselineRate

The minimum baseline rate (per day).

14 createScriIntervalData

maxBaselineRate

The maximum baseline rate (per day).

eraIds The IDs for the covariates to be generated.

patientUsages The fraction of patients that use the drugs.

usageRate The rate of prescriptions per person that uses the drug.

meanPrescriptionDurations

The mean duration of a prescription, per drug.

sdPrescriptionDurations

The standard deviation of the duration of a prescription, per drug.

simulationRiskWindows

One or a list of objects of type simulationRiskWindow as created using the createSimulationRiskWindow function.

include Age Effect

Include an age effect for the outcome?

ageKnots Number of knots in the age spline.

includeSeasonality

Include seasonality for the outcome?

seasonKnots Number of knots in the seasonality spline.

outcomeId The ID to be used for the outcome.

Details

Create an object of settings for an SCCS simulation.

Value

An object of type sccsSimulationSettings.

createScriIntervalData

Create Self-Controlled Risk Interval (SCRI) era data

Description

Create Self-Controlled Risk Interval (SCRI) era data

Usage

```
createScriIntervalData(
   studyPopulation,
   sccsData,
   eraCovariateSettings,
   controlIntervalSettings)
```

Arguments

studyPopulation

An object created using the createStudyPopulation() function.

sccsData An object of type SccsData as created using the getDbSccsData function.

eraCovariateSettings

Either an object of type EraCovariateSettings as created using the createEraCovariateSettings function, or a list of such objects.

 ${\tt controlIntervalSettings}$

An object of type ControlIntervalSettings as created using the createControlIntervalSettin function.

Details

This function creates interval data according to the elf-Controlled Risk Interval (SCRI) design. Unlike the generic SCCS design, where all patient time is used to establish a background rate, in the SCRI design a specific control interval (relative to the exposure) needs to be defined. The final model will only include time that is either part of the risk interval (defined using the eraCovariateSettings argument, or the control interval (defined using controlIntervalSettings).

Value

An object of type SccsIntervalData.

References

Greene SK, Kulldorff M, Lewis EM, Li R, Yin R, Weintraub ES, Fireman BH, Lieu TA, Nordin JD, Glanz JM, Baxter R, Jacobsen SJ, Broder KR, Lee GM. Near real-time surveillance for influenza vaccine safety: proof-of-concept in the Vaccine Safety Datalink Project. Am J Epidemiol. 2010 Jan 15;171(2):177-88. doi: 10.1093/aje/kwp345.

createSeasonalityCovariateSettings

Create seasonality settings

Description

Create seasonality settings

Usage

```
createSeasonalityCovariateSettings(
  seasonKnots = 5,
  allowRegularization = FALSE,
  computeConfidenceIntervals = FALSE
)
```

Arguments

seasonKnots

If a single number is provided this is assumed to indicate the number of knots to use for the spline, and the knots are automatically equally spaced across the year. If more than one number is provided these are assumed to be the exact location of the knots in days relative to the start of the year.

allowRegularization

When fitting the model, should the covariates defined here be allowed to be regularized?

 ${\tt computeConfidenceIntervals}$

Should confidence intervals be computed for the covariates defined here? Setting this to FALSE might save computing time when fitting the model. Will be turned to FALSE automatically when allowRegularization = TRUE.

Details

Create an object specifying whether and how seasonality should be included in the model. Seasonality can be included by splitting patient time into calendar months. During a month, the relative risk attributed to season is assumed to be constant, and the risk from month to month is modeled using a cyclic cubic spline.

Value

An object of type seasonalitySettings.

createSimulationRiskWindow

Create a risk window definition for simulation

Description

Create a risk window definition for simulation

Usage

```
createSimulationRiskWindow(
  start = 0,
  end = 0,
  endAnchor = "era end",
  splitPoints = c(),
  relativeRisks = c(0)
)
```

Arguments

start Start of the risk window relative to exposure start.

end The end of the risk window (in days) relative to the endAnchor.

endAnchor The anchor point for the end of the risk window. Can be "era start" or "era

end".

splitPoints Subdivision of the risk window in to smaller sub-windows.

relativeRisks Either a single number representing the relative risk in the risk window, or when

splitPoints have been defined a vector of relative risks, one for each sub-window.

createStudyPopulation 17

Value

An object of type simulationRiskWindow.

createStudyPopulation Create a study population

Description

Create a study population

Usage

```
createStudyPopulation(
  sccsData,
  outcomeId = NULL,
  firstOutcomeOnly = FALSE,
  naivePeriod = 0,
  minAge = NULL,
  maxAge = NULL
)
```

Arguments

sccsData An object of type SccsData as created using the getDbSccsData function.

outcomeId The outcome to create the era data for. If not specified it is assumed to be the

one outcome for which the data was loaded from the database.

firstOutcomeOnly

Whether only the first occurrence of an outcome should be considered.

naivePeriod The number of days at the start of a patient's observation period that should not

be included in the risk calculations. Note that the naive period can be used to determine current covariate status right after the naive period, and whether an

outcome is the first one.

minAge Minimum age at which patient time will be included in the analysis. Note that

information prior to the min age is still used to determine exposure status after the minimum age (e.g. when a prescription was started just prior to reaching the minimum age). Also, outcomes occurring before the minimum age is reached will be considered as prior outcomes when using first outcomes only. Age should be specified in years, but non-integer values are allowed. If not

specified, no age restriction will be applied.

maxAge Maximum age at which patient time will be included in the analysis. Age should

be specified in years, but non-integer values are allowed. If not specified, no age

restriction will be applied.

18 fitSccsModel

cyclicSplineDesign Create a design matrix for a cyclic spline

Description

Create a design matrix for a cyclic spline

Usage

```
cyclicSplineDesign(x, knots, ord = 4)
```

Arguments

x Vector of coordinates of the points to be interpolated.

knots Location of the knots.

ord Order of the spline function.

Details

This function is used by other functions in this package.

fitSccsModel

Fit the SCCS model

Description

Fit the SCCS model

Usage

```
fitSccsModel(
  sccsIntervalData,
  prior = createPrior("laplace", useCrossValidation = TRUE),
  control = createControl(cvType = "auto", selectorType = "byPid", startingVariance =
     0.1, noiseLevel = "quiet")
)
```

Arguments

sccsIntervalData

An object of type SccsIntervalData as created using the createSccsIntervalData

function.

prior The prior used to fit the model. See Cyclops::createPrior for details.

control The control object used to control the cross-validation used to determine the

hyperparameters of the prior (if applicable). See Cyclops::createControl for de-

tails.

getAttritionTable 19

Details

Fits the SCCS model as a conditional Poisson regression. When allowed, coefficients for some or all covariates can be regularized.

Value

An object of type SccsModel. Generic functions print, coef, and confint are available.

References

Suchard, M.A., Simpson, S.E., Zorych, I., Ryan, P., and Madigan, D. (2013). Massive parallelization of serial inference algorithms for complex generalized linear models. ACM Transactions on Modeling and Computer Simulation 23, 10

getAttritionTable

Get the attrition table for a population

Description

Get the attrition table for a population

Usage

getAttritionTable(object)

Arguments

object

Either an object of type SccsData, a population object generated by functions like createStudyPopulation, or an object of type outcomeModel.

Value

A tibble specifying the number of people and exposures in the population after specific steps of filtering.

 ${\tt getDbSccsData}$

Load data for SCCS from the database

Description

Load all data needed to perform an SCCS analysis from the database.

20 getDbSccsData

Usage

```
getDbSccsData(
  connectionDetails,
  cdmDatabaseSchema,
 oracleTempSchema = cdmDatabaseSchema,
  outcomeDatabaseSchema = cdmDatabaseSchema,
  outcomeTable = "condition_era",
  outcomeIds,
  exposureDatabaseSchema = cdmDatabaseSchema,
  exposureTable = "drug_era",
  exposureIds = c(),
  useCustomCovariates = FALSE,
  customCovariateDatabaseSchema = cdmDatabaseSchema,
  customCovariateTable = "cohort",
  customCovariateIds = c(),
  useNestingCohort = FALSE,
 nestingCohortDatabaseSchema = cdmDatabaseSchema,
 nestingCohortTable = "cohort",
  nestingCohortId = NULL,
  deleteCovariatesSmallCount = 100,
  studyStartDate = "",
  studyEndDate = ""
  cdmVersion = "5",
 maxCasesPerOutcome = 0
)
```

Arguments

connectionDetails

An R object of type ConnectionDetails created using the function DatabaseConnector::createConnection.

cdmDatabaseSchema

The name of the database schema that contains the OMOP CDM instance. Requires read permissions to this database. On SQL Server, this should specify both the database and the schema, so for example 'cdm_instance.dbo'.

oracleTempSchema

A schema where temp tables can be created in Oracle.

outcomeDatabaseSchema

The name of the database schema that is the location where the data used to define the outcome cohorts is available. If outcomeTable = CONDITION_ERA, outcomeDatabaseSchema is not used. Requires read permissions to this database.

outcomeTable

The tablename that contains the outcome cohorts. If outcomeTable is not CON-DITION_OCCURRENCE or CONDITION_ERA, then expectation is outcomeTable has format of COHORT table: COHORT_DEFINITION_ID, SUBJECT_ID, COHORT_START_DATE, COHORT_END_DATE.

outcomeIds

A list of ids used to define outcomes. If outcome Table = CONDITION_OCCURRENCE, the list is a set of ancestor CONCEPT_IDs, and all occurrences of all descendant concepts will be selected. If outcome Table <> CONDITION_OCCURRENCE, the list contains records found in COHORT_DEFINITION_ID field.

exposureDatabaseSchema

The name of the database schema that is the location where the exposure data used to define the exposure cohorts is available. If exposureTable = DRUG_ERA,

getDbSccsData 21

exposureDatabaseSchema is not used but assumed to be cdmSchema. Requires read permissions to this database.

exposureTable The tablename that contains the exposure cohorts. If exposureTable <> DRUG_ERA,

then expectation is exposure Table has format of COHORT table: cohort_concept_id,

SUBJECT_ID, COHORT_START_DATE, COHORT_END_DATE.

exposureIds A list of identifiers to define the exposures of interest. If exposureTable =

DRUG_ERA, exposureIds should be CONCEPT_ID. If exposureTable <> DRUG_ERA, exposureIds is used to select the cohort_concept_id in the cohort-like table. If no exposureIds are provided, all drugs or cohorts in the exposureTable are included

as exposures.

useCustomCovariates

Create covariates from a custom table?

customCovariateDatabaseSchema

The name of the database schema that is the location where the custom covariate data is available.

customCovariateTable

Name of the table holding the custom covariates. This table should have the same structure as the cohort table.

customCovariateIds

A list of cohort definition IDS identifying the records in the customCovariateTable to use for building custom covariates.

useNestingCohort

Should the study be nested in a cohort (e.g. people with a specific indication)? If not, the study will be nested in the general population.

nestingCohortDatabaseSchema

The name of the database schema that is the location where the nesting cohort is defined.

nestingCohortTable

Name of the table holding the nesting cohort. This table should have the same structure as the cohort table.

nestingCohortId

A cohort definition ID identifying the records in the nestingCohortTable to use as nesting cohort.

deleteCovariatesSmallCount

The minimum count for a covariate to appear in the data to be kept.

studyStartDate A calendar date specifying the minimum date where data is used. Date format is 'yyyymmdd'.

studyEndDate A calendar date specifying the maximum date where data is used. Date format is 'yyyymmdd'.

cdmVersion Define the OMOP CDM version used: currently support "4" and "5".

maxCasesPerOutcome

If there are more than this number of cases for a single outcome cases will be sampled to this size. maxCasesPerOutcome = 0 indicates no maximum size.

Details

This function downloads several types of information:

• Information on the occurrences of the outcome(s) of interest. Note that information for multiple outcomes can be fetched in one go, and later the specific outcome can be specified for which we want to build a model.

22 getModel

- Information on the observation time and age for the people with the outcomes.
- Information on exposures of interest which we want to include in the model.

Five different database schemas can be specified, for five different types of information: The

- cdmDatabaseSchema is used to extract patient age and observation period. The
- outcomeDatabaseSchema is used to extract information about the outcomes, the
- exposureDatabaseSchema is used to retrieve information on exposures, and the
- **customCovariateDatabaseSchema** is optionally used to find additional, user-defined covariates. All four locations could point to the same database schema.
- **nestingCohortDatabaseSchema** is optionally used to define a cohort in which the analysis is nested, for example a cohort of diabetics patients.

All five locations could point to the same database schema.

Value

An SccsData object.

getModel

Output the full model

Description

Output the full model

Usage

```
getModel(sccsModel)
```

Arguments

sccsModel

An object of type sccsModel as created using the fitSccsModel function.

Value

A data frame with the coefficients and confidence intervals (when not-regularized) for all covariates in the model.

isSccsData 23

isSccsData

Check whether an object is a SccsData object

Description

Check whether an object is a SccsData object

Usage

```
isSccsData(x)
```

Arguments

Χ

The object to check.

Value

A logical value.

 $\verb"isSccsIntervalData"$

Check whether an object is a SccsIntervalData object

Description

Check whether an object is a SccsIntervalData object

Usage

```
isSccsIntervalData(x)
```

Arguments

Χ

The object to check.

Value

A logical value.

24 loadSccsAnalysisList

 ${\tt loadExposureOutcomeList}$

Load a list of exposureOutcome from file

Description

Load a list of objects of type exposureOutcome from file. The file is in JSON format.

Usage

loadExposureOutcomeList(file)

Arguments

file

The name of the file

Value

A list of objects of type exposureOutcome.

loadSccsAnalysisList Load a list of sccsAnalysis from file

Description

Load a list of objects of type sccsAnalysis from file. The file is in JSON format.

Usage

loadSccsAnalysisList(file)

Arguments

file

The name of the file

Value

A list of objects of type sccsAnalysis.

loadSccsData 25

loadSccsData

Load the cohort method data from a file

Description

Loads an object of type SccsData from a file in the file system.

Usage

```
loadSccsData(file)
```

Arguments

file

The name of the file containing the data.

Value

An object of class SccsData.

 ${\tt loadSccsIntervalData} \quad \textit{Load the cohort method data from a file}$

Description

Loads an object of type SccsIntervalData from a file in the file system.

Usage

```
loadSccsIntervalData(file)
```

Arguments

file

The name of the file containing the data.

Value

An object of class SccsIntervalData.

26 plotAgeSpans

age effect
uge ejjeci

Description

Plot the age effect

Usage

```
plotAgeEffect(sccsModel, rrLim = c(0.1, 10), fileName = NULL)
```

Arguments

sccsModel An object of type sccsModel as created using the fitSccsModel function.

rrLim The limits on the incidence rate ratio scale in the plot.

fileName Name of the file where the plot should be saved, for example 'plot.png'. See the

function ggsave in the ggplot2 package for supported file formats.

Details

Plot the spline curve of the age effect.

Value

A Ggplot object. Use the ggsave function to save to file.

plotAgeSpans	Plot the age ranges spanned by each observation period.	

Description

Plot the age ranges spanned by each observation period.

Usage

```
plotAgeSpans(studyPopulation, maxPersons = 10000, fileName = NULL)
```

Arguments

studyPopulation

An object created using the createStudyPopulation() function.

maxPersons The maximum number of persons to plot. If there are more than this number of

persons a random sample will be taken to avoid visual clutter.

fileName Name of the file where the plot should be saved, for example 'plot.png'. See the

function ggplot2::ggsave() for supported file formats.

Details

Plots a line per patient from their age at observation start to their age at observation end.

Value

A ggplot object. Use the ggplot2::ggsave() function to save to file in a different format.

plotEventObservationDependence

Plot time from event to observation end for censored and uncensored time.

Description

Plot time from event to observation end for censored and uncensored time.

Usage

plotEventObservationDependence(studyPopulation, fileName = NULL)

Arguments

studyPopulation

An object created using the createStudyPopulation() function.

fileName

Name of the file where the plot should be saved, for example 'plot.png'. See the function ggplot2::ggsave() for supported file formats.

Details

This plot shows whether there is a difference in time between (first) event and the observation period end for periods that are 'censored' and those that are 'uncensored'. By 'censored' we mean periods that end before we would normally expect. Here, we define periods to be uncensored if they end at either the study end date (if specified), database end date (i.e. the date after which no data is captured in the database), or maximum age (if specified). All other periods are assumed to be censored.

As proposed by Farrington et al., by comparing the two plots, we can gain some insight into whether the censoring is dependent on the occurrence of the event.

Value

A ggplot object. Use the ggplot2::ggsave() function to save to file in a different format.

References

Farrington P, Whitaker H, Ghebremichael Weldeselassie Y (2018), Self-controlled case series studies: A modelling guide with R, Taylor & Francis

28 plotExposureCentered

```
plotEventToCalendarTime
```

Plot the count of events over calendar time.

Description

Plot the count of events over calendar time.

Usage

```
plotEventToCalendarTime(studyPopulation, fileName = NULL)
```

Arguments

studyPopulation

An object created using the createStudyPopulation() function.

fileName Name of the file where the plot should be saved, for example 'plot.png'. See the

function ggplot2::ggsave() for supported file formats.

Value

A ggplot object. Use the ggplot2::ggsave() function to save to file in a different format.

plotExposureCentered Plot information centered around the start of exposure

Description

Plot information centered around the start of exposure

Usage

```
plotExposureCentered(
   studyPopulation,
   sccsData,
   exposureEraId = NULL,
   fileName = NULL
)
```

Arguments

studyPopulation

An object created using the createStudyPopulation() function.

sccsData An object of type SccsData as created using the getDbSccsData function.

exposureEraId The exposure to create the era data for. If not specified it is assumed to be the

one exposure for which the data was loaded from the database.

fileName Name of the file where the plot should be saved, for example 'plot.png'. See the

function ggplot2::ggsave() for supported file formats.

plotSeasonality 29

Details

This plot shows the number of events and the number of subjects under observation in week-sized intervals relative to the start of the first exposure.

Value

A ggplot object. Use the ggplot2::ggsave() function to save to file in a different format.

plotSeasonality

Plot the seasonality effect

Description

Plot the seasonality effect

Usage

```
plotSeasonality(sccsModel, rrLim = c(0.1, 10), fileName = NULL)
```

Arguments

sccsModel An object of type sccsModel as created using the fitSccsModel function.

rrLim The limits on the incidence rate ratio scale in the plot.

fileName Name of the file where the plot should be saved, for example 'plot.png'. See the

function ggsave in the ggplot2 package for supported file formats.

Details

Plot the spline curve of the seasonality effect.

Value

A Ggplot object. Use the ggsave function to save to file.

runSccsAnalyses

Run a list of analyses

Description

Run a list of analyses

30 runSccsAnalyses

Usage

```
runSccsAnalyses(
  connectionDetails,
  cdmDatabaseSchema,
  oracleTempSchema = cdmDatabaseSchema,
  exposureDatabaseSchema = cdmDatabaseSchema,
  exposureTable = "drug_era",
  outcomeDatabaseSchema = cdmDatabaseSchema,
  outcomeTable = "condition_era",
  customCovariateDatabaseSchema = cdmDatabaseSchema,
  customCovariateTable = "cohort",
 nestingCohortDatabaseSchema = cdmDatabaseSchema,
 nestingCohortTable = "cohort",
  cdmVersion = 5,
  outputFolder = "./SccsOutput",
  sccsAnalysisList,
  exposureOutcomeList,
  combineDataFetchAcrossOutcomes = TRUE,
  getDbSccsDataThreads = 1,
  createStudyPopulationThreads = 1,
  createSccsIntervalDataThreads = 1,
  fitSccsModelThreads = 1,
  cvThreads = 1
```

Arguments

connectionDetails

 $An\ R\ object\ of\ type\ {\tt ConnectionDetails}\ created\ using\ the\ function\ {\tt DatabaseConnector::creat$

cdmDatabaseSchema

The name of the database schema that contains the OMOP CDM instance. Requires read permissions to this database. On SQL Server, this should specify both the database and the schema, so for example 'cdm_instance.dbo'.

oracleTempSchema

A schema where temp tables can be created in Oracle.

exposureDatabaseSchema

The name of the database schema that is the location where the exposure data used to define the exposure cohorts is available. If exposureTable = "DRUG_ERA", exposureDatabaseSchema is not used but assumed to be cdmDatabaseSchema. Requires read permissions to this database.

exposureTable

The table name that contains the exposure cohorts. If exposureTable <> "DRUG_ERA", then expectation is exposureTable has format of COHORT table: cohort_concept_id, SUBJECT_ID, COHORT_START_DATE, COHORT_END_DATE.

outcomeDatabaseSchema

The name of the database schema that is the location where the data used to define the outcome cohorts is available. If outcomeTable = "CONDITION_ERA", outcomeDatabaseSchema is not used. Requires read permissions to this database.

outcomeTable

The table name that contains the outcome cohorts. If outcomeTable is not CONDITION_OCCURRENCE or CONDITION_ERA, then expectation is outcomeTable has format of COHORT table: COHORT_DEFINITION_ID, SUBJECT_ID, COHORT_START_DATE, COHORT_END_DATE.

runSccsAnalyses 31

customCovariateDatabaseSchema

The name of the database schema that is the location where the custom covariate data is available.

customCovariateTable

Name of the table holding the custom covariates. This table should have the same structure as the cohort table.

nestingCohortDatabaseSchema

The name of the database schema that is the location where the nesting cohort is defined.

nestingCohortTable

Name of the table holding the nesting cohort. This table should have the same structure as the cohort table.

cdmVersion Define the OMOP CDM version used: currently support "4" and "5".

outputFolder Name of the folder where all the outputs will written to.

sccsAnalysisList

A list of objects of sccsAnalysis as created using the createSccsAnalysis() function.

exposureOutcomeList

A list of objects of type exposureOutcome as created using the createExposureOutcome() function.

combineDataFetchAcrossOutcomes

Should fetching data from the database be done one outcome at a time, or for all outcomes in one fetch? Combining fetches will be more efficient if there is large overlap in the subjects that have the different outcomes.

getDbSccsDataThreads

The number of parallel threads to use for building the SccsData objects.

createStudyPopulationThreads

The number of parallel threads to use for building the studyPopulation objects.

createSccsIntervalDataThreads

The number of parallel threads to use for building the SccsIntervalData objects.

fitSccsModelThreads

The number of parallel threads to use for fitting the models.

cvThreads

The number of parallel threads to use for the cross-validation when estimating the hyperparameter for the outcome model. Note that the total number of CV threads at one time could be fitSccsModelThreads * cvThreads.

Details

Run a list of analyses for the drug-comparator-outcomes of interest. This function will run all specified analyses against all hypotheses of interest, meaning that the total number of outcome models is length(cmAnalysisList) * length(drugComparatorOutcomesList) (if all analyses specify an outcome model should be fitted). When you provide several analyses it will determine whether any of the analyses have anything in common, and will take advantage of this fact. For example, if we specify several analyses that only differ in the way the outcome model is fitted, then this function will extract the data and fit the propensity model only once, and re-use this in all the analysis.

32 saveSccsAnalysisList

Value

A tibble describing for each exposure-outcome-analysisId combination where the intermediary and outcome model files can be found, relative to the outputFolder.

saveExposureOutcomeList

Save a list of exposureOutcome to file

Description

Write a list of objects of type exposureOutcome to file. The file is in JSON format.

Usage

```
saveExposureOutcomeList(exposureOutcomeList, file)
```

Arguments

exposureOutcomeList

The exposureOutcome list to be written to file

file

The name of the file where the results will be written

saveSccsAnalysisList Save a list of sccsAnalysis to file

Description

Write a list of objects of type sccsAnalysis to file. The file is in JSON format.

Usage

```
saveSccsAnalysisList(sccsAnalysisList, file)
```

Arguments

sccsAnalysisList

The sccsAnalysis list to be written to file

file The name of the file where the results will be written

saveSccsData 33

saveSccsData

Save the cohort method data to file

Description

Saves an object of type SccsData to a file.

Usage

```
saveSccsData(SccsData, file)
```

Arguments

SccsData An object of type SccsData as generated using getDbSccsData().

file The name of the file where the data will be written. If the file already exists it

will be overwritten.

Value

Returns no output.

 ${\tt save SccsInterval Data} \quad \textit{Save the cohort method data to file}$

Description

Saves an object of type SccsIntervalData to a file.

Usage

```
saveSccsIntervalData(SccsIntervalData, file)
```

Arguments

SccsIntervalData

An object of type SccsIntervalData as generated using createSccsIntervalData().

file

The name of the file where the data will be written. If the file already exists it will be overwritten.

Value

Returns no output.

34 SccsIntervalData-class

SccsData-class

SCCS Data

Description

SccsData is an S4 class that inherits from Andromeda. It contains information on the cases and their covariates.

A SccsData is typically created using getDbSccsData(), can only be saved using saveSccsData(), and loaded using loadSccsData().

Usage

```
## S4 method for signature 'SccsData'
show(object)
## S4 method for signature 'SccsData'
summary(object)
```

Arguments

object

An object of type SccsData.

SccsIntervalData-class

SCCS Interval Data

Description

SccsIntervalData' is an S4 class that inherits from Andromeda. It contains information on the cases and their covariates, divided in non-overlapping time intervals.

A SccsIntervalData is typically created using createSccsIntervalData(), can only be saved using saveSccsIntervalData(), and loaded using loadSccsIntervalData().

Usage

```
## S4 method for signature 'SccsIntervalData'
show(object)
## S4 method for signature 'SccsIntervalData'
summary(object)
```

Arguments

object

An object of type SccsIntervalData.

simulateSccsData 35

simulateSccsData

Simulate SCCS data

Description

Simulate SCCS data

Usage

```
simulateSccsData(nCases, settings)
```

Arguments

nCases

The number of cases to simulate.

settings

An object of type sccsSimulationSettings as created using the createSccsSimulationSettings

Value

An object of type sccsData.

summarizeSccsAnalyses Create a summary report of the analyses

Description

Create a summary report of the analyses

Usage

```
summarizeSccsAnalyses(referenceTable, outputFolder)
```

Arguments

referenceTable A tibble as created by the runSccsAnalyses function.

outputFolder Name of the folder where all the outputs have been written to.

Value

A tibble containing summary statistics for each exposure-outcome-analysis combination.

Index

Andromeda, 34	loadExposureOutcomeList, 24
	loadSccsAnalysisList, 24
computeMdrr, 3	loadSccsData, 25
<pre>createAgeCovariateSettings, 4</pre>	loadSccsData(), 34
<pre>createAgeCovariateSettings(), 12</pre>	loadSccsIntervalData, 25
<pre>createControlIntervalSettings, 4</pre>	loadSccsIntervalData(), 34
<pre>createControlIntervalSettings(), 15</pre>	200000001110011001000(7, 0 7
<pre>createCreateSccsIntervalDataArgs, 5</pre>	plotAgeEffect, 26
<pre>createCreateScriIntervalDataArgs, 6</pre>	plotAgeSpans, 26
<pre>createCreateStudyPopulationArgs, 7</pre>	plotEventObservationDependence, 27
<pre>createEraCovariateSettings, 7</pre>	plotEventToCalendarTime, 28
<pre>createEraCovariateSettings(), 12, 15</pre>	plotExposureCentered, 28
<pre>createExposureOutcome, 9</pre>	plotSeasonality, 29
<pre>createExposureOutcome(), 31</pre>	
<pre>createFitSccsModelArgs,9</pre>	runSccsAnalyses, <i>9</i> , <i>12</i> , 29, <i>35</i>
createGetDbSccsDataArgs, 10	
createSccsAnalysis, 9, 11	saveExposureOutcomeList, 32
<pre>createSccsAnalysis(), 31</pre>	saveSccsAnalysisList, 32
createSccsIntervalData, 3, 11, 12, 18	saveSccsData, 33
<pre>createSccsIntervalData(), 33, 34</pre>	saveSccsData(), 34
createSccsSimulationSettings, 13, 35	saveSccsIntervalData, 33
createScriIntervalData, <i>11</i> , 14	saveSccsIntervalData(), 34
createSeasonalityCovariateSettings, 15	SccsData, 5, 8, 12, 15, 17, 19, 22, 25, 28, 33
<pre>createSeasonalityCovariateSettings(),</pre>	SccsData (SccsData-class), 34
12	SccsData-class, 34
createSimulationRiskWindow, 14, 16	SccsIntervalData, 3, 13, 15, 18, 25, 33
createStudyPopulation, 17, 19	SccsIntervalData
createStudyPopulation(), 12, 15, 26-28	(SccsIntervalData-class), 34
cyclicSplineDesign, 18	SccsIntervalData-class, 34
Cyclops::createControl, 18	show, SccsData-method (SccsData-class),
Cyclops::createPrior, 18	34
	show, SccsIntervalData-method
DatabaseConnector::createConnectionDetails()	
20, 30	simulateSccsData, 35
fitSccsModel, 11, 18, 22, 26, 29	summarizeSccsAnalyses, 35
11t3ccsrioue1, 11, 10, 22, 20, 29	summary, SccsData-method
getAttritionTable, 19	(SccsData-class), 34
getDbSccsData, 11, 12, 15, 17, 19, 28	summary, SccsIntervalData-method
getDbSccsData(), 33, 34	(SccsIntervalData-class), 34
getModel, 22	
ggplot2::ggsave(), 26–29	
00F	
isSccsData, 23	
isSccsIntervalData, 23	