Package 'riskgaps'

September 16, 2021

Title Absolute and relative risk estimation in the presence of outcome ascertainment gaps and competing risks

Version 0.0.2

Date 2021-09-16

Description A package which includes methods to account for observation gaps in the time-to-event analyses of cancer risk.

Imports stats, utils, numDeriv, survival, lubridate

Depends R (>= 3.5.0)

License GPL-2

NeedsCompilation yes

Author Danping Liu [aut],
Emily Wu [aut],
Joanna Shih [aut],
Cari Kitahara [aut],
Li Cheung [aut, cre]

Maintainer Li Cheung <li.cheung@nih.gov>

R topics documented:

riskgaps-package *Absolute and relative risk estimation in the presence of outcome ascertainment gaps and competing risks*

Description

A package which includes methods to account for observation gaps in the time-to-event analyses of cancer risk.

Details

Imperfect coverage by cancer registries can lead to an underreporting of cancers and a resulting bias in risk estimates. For example, in the U.S. Radiologic Technologist (USRT) cohort, where cancer outcomes are ascertained through 43 state registries, gaps in observation can arise as individuals move in and out of geographic areas covered by the registries. Moreover, the exact periods of non-observation may be unknown due to incomplete reporting of residential histories. This package includes a two-step procedure, which reduces bias and improves efficiency, for estimating relative and absolute risk in this setting. First, using a mover stayer model fitted to individuals' known residential history, we obtain individual posterior probabilities of residing outside the coverage area each year. Second, we incorporate these probabilities in the survival data likelihood for competing risks to account for unobserved events.

Author(s)

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung

data1 *Data for examples*

Description

Data for the examples.

Details

The data contains the event status, time to event, covariates, and residence variables for 1000 subjects.

Examples

```
data(data1, package="riskgaps")

# Display some of the data
data1[1:5, ]
```

data2 *Data for examples*

Description

Data for the examples.

Details

The data contains the event status, times to event, covariates, and residence variables for 1000 subjects.

Examples

```
data(data2, package="riskgaps")

# Display some of the data
data2[1:5, ]
```

gcuminc1 *Absolute risk*

Description

Compute absolute risk without competing risks

Usage

```
gcuminc1(parms, var_cov, newdata, times)
```

Arguments

parms	Vector of parameter estimates (log(shape), intercept, covariates).
var_cov	Variance-covariance matrix for parms.
newdata	NULL or matrix/data frame of covariates. If not NULL, then the columns must be in the same order as the covariates in parms.
times	Vector of times to compute the absolute risks at.

Details

The objects parms and var_cov should be from noCompetingRisks.

Value

See the returned object from getAbsRisk.

Author(s)

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung

References

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung (2021) Absolute and relative risk estimation in the presence of outcome ascertainment gaps and competing risks.

See Also

gcuminc2, getAbsRisk

Examples

```
options(warn=0)
data(data1, package="riskgaps")

time.var        <- "y"
event.var       <- "d"
residence.vars  <- as.character(2000:2019)
residence.dates <- 0:19
cov.vars        <- c("x1", "x2")

ret <- noCompetingRisks(data1, time.var, event.var, residence.vars, residence.dates,
                        cov.vars=cov.vars)

newdata <- matrix(c(1, 2, 2, 1, 2, 2), ncol=2, byrow=TRUE)
gcuminc1(ret$cox.estimates, ret$cox.cov, newdata, 15)
```

---

gcuminc2 *Absolute risk*

Description

Compute absolute risk with two competing events

Usage

```
gcuminc2(parms, var_cov, newdata, times)
```

Arguments

parms	Vector of parameter estimates (log(shape), intercept, covariates).
var_cov	Variance-covariance matrix for parms.
newdata	NULL or matrix/data frame of covariates. If not NULL, then the columns must be in the same order as the covariates in parms.
times	Vector of times to compute the absolute risks at.

Details

The objects parms and var_cov should be from twoCompetingRisks.

Value

See the returned object from getAbsRisk.

Author(s)

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung

References

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung (2021) Absolute and relative risk estimation in the presence of outcome ascertainment gaps and competing risks.

See Also

gcuminc1, getAbsRisk

Examples

```
options(warn=0)
data(data2, package="riskgaps")

time.var        <- "y"
event1.var      <- "d1"
event2.var      <- "d2"
residence.vars  <- as.character(2000:2019)
residence.dates <- 0:19
cov.vars        <- c("x1", "x2")

ret <- twoCompetingRisks(data2, time.var, event1.var, event2.var, residence.vars, residence.dates,
                         cov.vars=cov.vars)

newdata <- matrix(c(1, 2, 2, 1, 2, 2), ncol=2, byrow=TRUE)
gcuminc2(ret$cox.estimates, ret$cox.cov, newdata, 15)
```

---

getAbsRisk *Absolute risk*

Description

Compute absolute risk

Usage

```
getAbsRisk(obj, newdata, risk.times=NULL)
```

Arguments

obj	Return object from noCompetingRisks or twoCompetingRisks.
newdata	Data frame containing subjects to compute absolute risks for. This data must contain column names that were used as covariates in the call to noCompetingRisks or twoCompetingRisks.
risk.times	NULL or vector of times to compute the absolute risks at. If NULL, then the risk times used in the call to noCompetingRisks or twoCompetingRisks will be used.

Details

This function is more efficient than gcuminc1 and gcuminc2, especially when all of the covariates are categorical and there are is a large number of subjects.

Value

A list with names absRisk, absRisk.se, and times.

- absRisk Matrix of absolute risk estimates for each subject and risk time. The columns of the matrix correspond to the times.
- absRisk.se Matrix of standard errors for absRisk.
- times Vector of risk times.

Author(s)

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung

References

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung (2021) Absolute and relative risk estimation in the presence of outcome ascertainment gaps and competing risks.

See Also

noCompetingRisks, twoCompetingRisks, gcuminc1, gcuminc2

Examples

```
options(warn=0)
data(data1, package="riskgaps")

time.var        <- "y"
event1.var      <- "d"
residence.vars  <- as.character(2000:2019)
residence.dates <- 0:19
cov.vars        <- c("x1", "x2")

# In this example, absolute risks are not computed in the call to noCompetingRisks.
ret  <- noCompetingRisks(data1, time.var, event1.var, residence.vars, residence.dates,
                         cov.vars=cov.vars, op=list(abs.risk.compute=FALSE))
risk <- getAbsRisk(ret, data1, risk.times=1:5)
risk$absRisk[1:3, ]
```

---

noCompetingRisks *noCompetingRisks*

Description

Survival analysis without competing risks

Usage

```
noCompetingRisks(data, time.var, event.var, residence.vars, residence.dates,
                 cov.vars=NULL, op=NULL)
```

Arguments

data	Data frame containing all variables.
time.var	Name of the time to event variable in data.
event.var	Name of the event variable in data. This variable must be coded as 0-1.
residence.vars	Names of the residence variables in data. These variables must be coded as 0, 1 or NA
residence.dates	Character vector of the form "yyyy-mm-dd" or a numeric vector corresponding to residence.vars. This vector must have the same length and order of residence.vars.
cov.vars	Names of the adjusted (numeric) covariates in data. The default is NULL.
op	List of options (see details). The default is NULL.

Details

See the reference for details of this method. Samples with missing values in any of time.var, event.var or cov.vars will be removed before the analysis. Currently only numeric covariates are allowed, so any covariates that are character or factors should be converted into dummy variables.

options list:

Option	Description	Default
abs.risk.compute	TRUE or FALSE to compute absolute risk	TRUE
abs.risk.times	NULL or vector of times to compute absolute risk	quartiles of event times
em.eps	stopping tolerance in EM alg	1e-6
em.maxiter	max number of iterations in EM alg	100000
ms.p_00	Mover-Stayer parameter	0.1
ms.p_11	Mover-Stayer parameter	0.4
ms.pi_0	Mover-Stayer parameter	0.9
ms.pi_1	Mover-Stayer parameter	0.9
ms.r_0	Mover-Stayer parameter	0.3
optim.init	NULL or vector of initial estimates for optim in the form (log(shape), intercept, cov.vars)	estimated from Weibull regression
print	0-2 to print information	1

Value

A list with names cox.estimates, cox.cov, data, and possibly abs.risk if the option abs.risk.compute was set to TRUE.

- abs.risk See the returned object from getAbsRisk.
- cox.estimates Vector of estimated parameters from the call to optim. This vector is in the form (log(shape),intercept,cov.vars).
- cox.cov The covariance matrix for cox.estimates.
- data The data frame used in the analysis.

Author(s)

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung

References

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung (2021) Absolute and relative risk estimation in the presence of outcome ascertainment gaps and competing risks.

See Also

twoCompetingRisks, getAbsRisk

Examples

```
options(warn=0)
data(data1, package="riskgaps")

time.var        <- "y"
event.var       <- "d"
residence.vars  <- as.character(2000:2019)
residence.dates <- 0:19
cov.vars        <- c("x1", "x2")

ret <- noCompetingRisks(data1, time.var, event.var, residence.vars, residence.dates,
                        cov.vars=cov.vars)
names(ret)
```

---

twoCompetingRisks *twoCompetingRisks*

Description

Survival analysis with two competing risks

Usage

```
twoCompetingRisks(data, time.var, event1.var, event2.var, residence.vars, residence.dates,
                        cov.vars=NULL, op=NULL)
```

Arguments

data	Data frame containing all variables.
time.var	Name of the time to event variable in data.
event1.var	Name of the first event variable in data. This variable must be coded as 0-1.
event2.var	Name of the second event variable in data. This variable must be coded as 0-1.
residence.vars	Names of the residence variables in data. These variables must be coded as 0, 1 or NA
residence.dates	Character vector of the form "yyyy-mm-dd" or a numeric vector corresponding to residence.vars. This vector must have the same length and order of residence.vars.

cov.vars	Names of the adjusted (numeric) covariates in data. The default is NULL.
op	List of options (see details). The default is NULL.

Details

See the reference for details of this method. Samples with missing values in any of time.var, event1.var, event2.var, or cov.vars will be removed before the analysis. Currently only numeric covariates are allowed, so any covariates that are character or factors should be converted into dummy variables.

options list:

Option	Description	Default
abs.risk.compute	TRUE or FALSE to compute absolute risk	TRUE
abs.risk.times	NULL or vector of times to compute absolute risk	quartiles of event times
em.eps	stopping tolerance in EM alg	1e-6
em.maxiter	max number of iterations in EM alg	100000
ms.p_00	Mover-Stayer parameter	0.1
ms.p_11	Mover-Stayer parameter	0.4
ms.pi_0	Mover-Stayer parameter	0.9
ms.pi_1	Mover-Stayer parameter	0.9
ms.r_0	Mover-Stayer parameter	0.3
optim.init	NULL or vector of initial estimates for optim in the form (log(shape), intercept, cov.vars)	estimated from Weibull regression
print	0-2 to print information	1

Value

A list with names cox.estimates, cox.cov, data, and possibly abs.risk if the option abs.risk.compute was set to TRUE.

- abs.risk See the returned object from getAbsRisk.
- cox.estimates Vector of estimated parameters from the call to optim. This vector is in the form (log(shape),intercept,cov.vars).
- cox.cov The covariance matrix for cox.estimates.
- data The data frame used in the analysis.

Author(s)

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung

References

Danping Liu, Emily Wu, Joanna H. Shih, Cari Kitahara, and Li C. Cheung (2021) Absolute and relative risk estimation in the presence of outcome ascertainment gaps and competing risks.

See Also

noCompetingRisks, getAbsRisk

Examples

```
options(warn=0)
data(data2, package="riskgaps")

time.var        <- "y"
event1.var      <- "d1"
event2.var      <- "d2"
residence.vars  <- as.character(2000:2019)
residence.dates <- 0:19
cov.vars        <- c("x1", "x2")

ret <- twoCompetingRisks(data2, time.var, event1.var, event2.var, residence.vars, residence.dates,
                         cov.vars=cov.vars)
names(ret)
```

Index