Package ‘kmi’

May 27, 2019

Version 0.5.5

Title Kaplan-Meier Multiple Imputation for the Analysis of Cumulative Incidence Functions in the Competing Risks Setting

Author Arthur Allignol <arthur.allignol@gmail.com>

Maintainer Arthur Allignol <arthur.allignol@gmail.com>

Imports mitools,survival,stats

Description Performs a Kaplan-Meier multiple imputation to recover the missing potential censoring information from competing risks events, so that standard right-censored methods could be applied to the imputed data sets to perform analyses of the cumulative incidence functions (Allignol and Beyersmann, 2010 <doi:10.1093/biostatistics/kxq018>).

License GPL (>= 2)

URL https://github.com/aallignol/kmi

BugReports https://github.com/aallignol/kmi/issues

NeedsCompilation no

Repository CRAN

Date/Publication 2019-05-27 21:10:13 UTC

R topics documented:

cox.kmi ................................................................. 2
icu.pneu .............................................................. 3
kmi ................................................................. 4
print.cox.kmi ......................................................... 7
print.summary.cox.kmi ............................................... 7
summary.cox.kmi ..................................................... 8

Index 10

1
**cox.kmi**

---

**Cox proportional hazards model applied to imputed data sets**

---

**Description**

This function fits Cox proportional hazards models to each imputed data set to estimate the regression coefficients in a proportional subdistribution hazards model, and pools the results.

**Usage**

```r
cox.kmi(formula, imp.data, df.complete = Inf, ...)
```

**Arguments**

- `formula` A formula object, with the response on the left of a `~` operator, and the terms on the right. The response must be a survival object as returned by the `Surv` function.
- `imp.data` An object of class `kmi`.
- `df.complete` Complete data degrees of freedom.
- `...` Further arguments for the `coxph` function.

**Details**

Fits a Cox proportional hazards model on each imputed data set to estimate the regression coefficients in a proportional subdistribution hazards model, and pools the results, using the `MIcombine` function of the mitools package.

**Value**

An object of class `cox.kmi` including the following components:

- `coefficients` Pooled regression coefficient estimates
- `variance` Pooled variance estimate
- `nimp` Number of multiple imputations
- `df` degrees of freedom
- `call` The matched call
- `individual.fit` A list of coxph objects. One for each imputed data set.

**Author(s)**

Arthur Allignol, <arthur.allignol@gmail.com>

**See Also**

- `coxph`, `MIcombine`, `print.cox.kmi`, `summary.cox.kmi`
Examples

data(icu.pneu)

if (require(survival)) {
  set.seed(1313)
  imp.dat <- kmi(Surv(start, stop, status) ~ 1, data = icu.pneu,
                 etype = event, id = id, failcode = 2, nimp = 5)

  fit.kmi <- cox.kmi(Surv(start, stop, event == 2) ~ pneu, imp.dat)
  summary(fit.kmi)

  ### Now using the censoring-complete data
  fit <- coxph(Surv(start, adm.cens.exit, event == 2) ~ pneu, icu.pneu)
  summary(fit)

  ## estimation of the censoring distribution adjusted on covariates
  dat.cova <- kmi(Surv(start, stop, status) ~ age + sex,
                  data = icu.pneu, etype = event, id = id,
                  failcode = 2, nimp = 5)

  fit.kmi2 <- cox.kmi(Surv(start, adm.cens.exit, event == 2) ~ pneu + age,
                      dat.cova)
  summary(fit.kmi2)
}

icu.pneu

**Hospital acquired pneumonia in ICU**

Description

This data set is a random sample drawn from the SIR-3 study that aimed at analysing the effect of nosocomial infections on the length of ICU stay. Patients were included in the study if they had stayed at least 1 day in the unit. The sample includes information to assess the effect of nosocomial pneumonia on the length of stay. The endpoint is either discharge alive from the ICU or dead in the unit. These data are censoring complete as the censoring time is known for all patients.

Usage

data(icu.pneu)

Format

A data frame with 1421 observations on the following 8 variables.
id  Individual patient id.
start  Start of the observation time.
stop  Failure time.
status  Censoring status. 0 if the observation is censored, 1 otherwise.
event  Event type. 2 is death in ICU, 3 is discharge alive
pneu  Nosocomial pneumonia indicator.
ad.m.cens.exit  Exit times for patients discharged alive are replaced by their administrative censoring times.
age  Age at inclusion
sex  Sex. F for female and M for male

Source

References

Examples
data(icu.pneu)

kmi  Kaplan-Meier Multiple Imputation for Competing Risks

Description
The function performs a non parametric multiple imputation that aims at recovering the missing potential censoring times from competing events.

Usage
kmi(formula, data, id = NULL, etype, failcode = 1, nimp = 10,
epsilon = 1, bootstrap = FALSE, nboot = 10)
Arguments

- **formula**: A formula object, that must have a `Surv` object on the left of a `~` operator. Covariates could be added on the right hand side of the formula. They will be used to model the censoring distribution. See Details.

- **data**: A data.frame in which to interpret the variables given in the formula, `etype` and `id`. It is mandatory.

- **id**: Used to identify individual subjects when one subject can have several rows of data, e.g., with time-dependent covariates. Set to `NULL` when there is only one raw of data per subject.

- **etype**: Variable specifying the type of competing event. When `status == 1` in formula, `etype` describes the type of event, otherwise, for censored observation, `(status == 0)`, the value of `etype` is ignored.

- **failcode**: Indicates the failure cause of interest. Imputation will be performed on the other competing events. Default is 1.

- **nimp**: Number of multiple imputation. Default is 10.

- **epsilon**: When the last time is an event, a censoring time equal to `max(time) + epsilon` is added. By default, `epsilon` is set to 1.

- **bootstrap**: Logical. Whether to estimate the censoring distribution using bootstrap samples. Default is FALSE.

- **nboot**: If `bootstrap` is set to TRUE, `nboot` determines the number of bootstrap samples.

Details

It was shown that if censoring times are observed for all individuals, methods for standard right-censored survival data can be used to analyse cumulative incidence functions from competing risks (Fine and Gray 1999). Therefore the idea proposed by Ruan and Gray (2008) is to impute potential censoring times for individuals who have failed from the competing events. The censoring times are imputed from the conditional Kaplan-Meier estimator of the censoring distribution.

Estimation of the censoring distribution may be improved through bootstrapping. Estimation might also be improved fitting a model for the censoring distribution. When covariates are given, a proportional hazards model on the hazard of censoring is fit. The censoring times are then imputed from the estimated model.

The competing risks model formulation in `formula` mimics the one in `survfit`.

Value

An object of class `kmi` with the following components:

- **imputed.data**: A list of matrices giving the imputed times in the first column and imputed event type in the second column. The event status for imputed times take value 0 (censored).

- **original.data**: The original data set

- **info**: Gives the names of the time and event indicator column in the original data set.

- **call**: The matched call.
Warning

When a proportional hazards model is fit for modelling the censoring distribution, the censoring times are imputed from the imputed model. When there is missing covariate information for the prediction, mean imputation is used.

Note

This multiple imputation technique does not work for left-truncated data.

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

References


See Also

`icu.pneu, cox.kmi, Surv, survfit`

Examples

data(icu.pneu)

if (require(survival)) {

  dat <- kmi(Surv(start, stop, status) ~ 1, data = icu.pneu,
             etype = event, id = id, failcode = 2, nimp = 5)

  ## another way to specify the formula if there is no status variable
  icu.pneu$ev <- icu.pneu$event
  icu.pneu$ev[icu.pneu$status == 0] <- 0

  dat <- kmi(Surv(start, stop, ev != 0) ~ 1, data = icu.pneu,
             etype = ev, id = id, failcode = 2, nimp = 5)

  ## with covariates to model the censoring distribution
  dat.cova <- kmi(Surv(start, stop, status) ~ age + sex,
                  data = icu.pneu, etype = event, id = id,
                  failcode = 2, nimp = 5)
}
print.cox.kmi

**Print method for cox.kmi objects**

### Description

Print method for cox.kmi objects.

### Usage

```r
## S3 method for class 'cox.kmi'
print(x, print.ind = FALSE, ...)
```

### Arguments

- **x** An object of class cox.kmi.
- **print.ind** A logical specifying whether to print the results of the analyses performed on each imputed data set. By default, only the pooled estimates are printed.
- **...** Further arguments

### Value

No value returned

### Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

### See Also

cox.kmi, summary.cox.kmi

---

print.summary.cox.kmi

**Print method for summary.cox.kmi objects**

### Description

Print method for summary.cox.kmi objects.

### Usage

```r
## S3 method for class 'summary.cox.kmi'
print(x,
      digits = max(getOption("digits") - 3, 3),
      signif.stars = getOption("show.signif.stars"),
      print.ind = FALSE, ...)
```
Arguments

- \( x \) An object of class \texttt{summary.cox.kmi}.
- \texttt{digits} Significant digits to print.
- \texttt{signif.stars} Logical. If \texttt{TRUE}, 'significance stars' are printed for each coefficient.
- \texttt{print.ind} Logical specifying whether to print a summary of the models fitted on each imputed data set. Default is \texttt{FALSE}.
- \( \ldots \) Further arguments

Value

No value returned

Author(s)

Arthur Allignol, \texttt{arthur.allignol@gmail.com}

See Also

\texttt{summary.cox.kmi}

---

\texttt{summary.cox.kmi} \hspace{1cm} \textit{Summary method for cox.kmi objects}

Description

Provides a summary of the fitted model.

Usage

```r
## S3 method for class 'cox.kmi'
summary(object, conf.int = 0.95, scale = 1, \ldots)
```

Arguments

- \texttt{object} An object of class \texttt{cox.kmi}.
- \texttt{conf.int} Level of the confidence intervals. Default is 0.95.
- \texttt{scale} Vector of scale factors for the coefficients, default to 1. The confidence limits are for the risk change associated with one scale unit.
- \( \ldots \) Further arguments
Value

An object of class `summary.cox.kmi` with the following components:

- **call**: The matched call
- **coefficients**: A matrix with 5 columns including the regression coefficients, subdistribution hazard ratios, standard-errors, t-statistics and corresponding two-sided p-values.
- **conf.int**: A matrix with 4 columns that consists of the subdistribution hazard ratios, \( \exp(-\text{coef}) \) and the lower and upper bounds of the confidence interval.
- **individual.fit**: A list of `summary.coxph` objects for each imputed data set

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

See Also

`cox.kmi`, `print.summary.cox.kmi`, `summary.coxph`
Index

*Topic **datasets**
icu.pneu, 3

*Topic **methods**
summary.cox.kmi, 8

*Topic **models**
cox.kmi, 2

*Topic **print**
  print.cox.kmi, 7
  print.summary.cox.kmi, 7

*Topic **regression**
  cox.kmi, 2

*Topic **survival**
  cox.kmi, 2
  kmi, 4
  cox.kmi, 2, 6, 7, 9
  coxph, 2
  icu.pneu, 3, 6
  kmi, 4
  MICombine, 2

  print.cox.kmi, 2, 7
  print.summary.cox.kmi, 7, 9
  summary.cox.kmi, 2, 7, 8, 8
  summary.coxph, 9
  Surv, 2, 6
  survfit, 5, 6