

Package ‘CompetingRiskFrailty’

April 17, 2009

Version 2.0

Title Competing Risk Model with Frailties for Right Censored Survival Data

Author Pavel Khomski <pkhomski@web.de>

Description The package offers a fitting of smooth varying coefficients in a competing risks modelling of hazards as well as estimating of the frailties (or unobserved heterogenities) for clustered observations. Nonparametric penalized spline (p-spline) fitting of smooth covariates effects is proposed. As a spline basis truncated polynomial functions are chosen. The frailties are also fitted (via the EM-algoritghm) in a flexible way using a penalized mixture of gamma distributions.

Maintainer Pavel Khomski <pkhomski@web.de>

License GPL (>= 2)

Repository CRAN

Date/Publication 2007-11-26 10:20:45

R topics documented:

CompetingRiskFrailtySurv	2
CompetingRiskFrailtySurvfitControl	3
CompetingRiskFrailtySurvfitCreate	4
CompetingRiskFrailtySurvfitObject	6
plot.CompetingRiskFrailtySurvfit	7
print.CompetingRiskFrailtySurvfit	8
unemployed.comp.risk	9

Index	11
--------------	-----------

`CompetingRiskFrailtySurv`*Creates a Competing Risk Model Object*

Description

Creates an object of class 'CompetingRiskFrailtySurv' for using as a response in the model formula in the call to the 'CompetingRiskFrailtySurvfitCreate' function.

Usage

```
CompetingRiskFrailtySurv(ID=ID, surv.time=surv.time, status=status)
```

Arguments

<code>ID</code>	identification of clustered observations.
<code>surv.time</code>	time to follow up for right censored data.
<code>status</code>	a data frame or a matrix with columns indicating the resulting event (or competing risk) and coded as (0,1)-indicator vectors, 1 for event of the specified type, and 0 if such an event has not been observed.

Details

The function handles only right censored data. All arguments must be numeric vectors. NA's can be supplied but will not be supported by the function 'CompetingRiskFrailtySurvfitCreate' in the further analysis.

Value

An object of class 'CompetingRiskFrailtySurv'.

Author(s)

Pavel Khomski <pkhomski@wiwi.uni-bielefeld.de>

References

Kauermann G. and Khomski P. (2006): *Full time or part time reemployment: a competing risk model with frailties and smooth effects using a penalty based approach*, to appear.

See Also

[CompetingRiskFrailtySurvfitCreate](#)

Examples

```

data(unemployed.comp.risk)
ID<-unemployed.comp.risk$id
surv.time<-unemployed.comp.risk$urv.time
status<-unemployed.comp.risk[,3:4]
my.Surv<-CompetingRiskFrailtySurv(ID=ID, surv.time=surv.time, status=status)

```

CompetingRiskFrailtySurvfitControl

Control Values for Fitting of Competing Risks with Frailties Model

Description

The values supplied in the call of 'CompetingRiskFrailtySurvfitCreate' function will replace the defaults, and a list with all possible arguments is returned. The returned list is used as the 'control' argument to the 'CompetingRiskFrailtySurvfitCreate' function.

Usage

```

CompetingRiskFrailtySurvfitControl(niter.EM=50, niter.epoch = 2, tol.epoch = 1e-08,
                                   tol.frailty = 1e-06, print.penalty.mixture=TRUE,
                                   print.estimate=FALSE, print.log.lik=TRUE, ...)

```

Arguments

niter.EM	maximum number of the (outer) EM-iterations.
niter.epoch	maximum number of the (inner) iterations in optimization for varying coefficients parameters theta and penalty parameters of their random parts, within an EM-iteration.
tol.epoch	tolerance for the convergence criterion for the fixed and random parameters of the varying coefficients.
tol.variance	tolerance for the convergence criterion for the penalty values of varying coefficients.
tol.frailty	tolerance for the the convergence criterion for the frailty terms.
print.penalty.mixture	logical value for printing the value of the penalty parameter from the specified grid of values.
print.EM	logical value for printing the current number of an EM-iteration.
print.estimate	logical value for printing the estimates of the fixed parameters theta and penalties of varying coefficients after the last EM-iteration.
print.log.lik	logical value for printing the marginal log likelihood of the model in each EM-iteration.

... other parameters which can only be: 'num.knots' for the number of spline knots for survival time. If specified it is a vector of integer values of the length equal to the number of competing risks. If not specified, the optimal values will be defined internally.

Details

The defaults or user specified values are applied as the 'control' argument in the call of the 'CompetingRiskFrailtySurvfitCreate' function. It can be an (empty) list object or a call to the 'CompetingRiskFrailtySurvfitControl' function itself, whether or not with supplied arguments to be changed from their default values.

Value

a list with components for each of the possible arguments.

Author(s)

Pavel Khomski <pkhomski@wiwi.uni-bielefeld.de>

References

Kauerman G. and Khomski P. (2006). *Full time or part time reemployment: a competing risk model with frailties and smooth effects using a penalty based approach*, to appear.

See Also

[CompetingRiskFrailtySurvfitCreate](#)

CompetingRiskFrailtySurvfitCreate

Competing Risk Model with Frailties and Smooth Varying Coefficients

Description

This function offers a strategy of fitting the competing risks models with frailties for estimating a hazard function. The effects of covariates are modelled as varying coefficients. For flexible smoothing of the effects truncated polynomial splines are employed. The frailties (or unobserved heterogeneities) of the (clustered) observations are also modelled in a flexible way by employing a mixture of gamma distributions. The estimation routine is based on the EM-algorithm.

Usage

```
CompetingRiskFrailtySurvfitCreate(formula=formula, data=data, na.action=na.fail,  
                                control=control, risk.names=risk.names)
```

Arguments

<code>formula</code>	a symbolic description of the model to be fit. The details of the model specification are given below.
<code>data</code>	an optional data frame containing variables in the model. If not supplied the handling of missing values will be accordingly to the value of <code>'na.action'</code> .
<code>na.action</code>	an optional parameter for handling with missing data.
<code>control</code>	an optional list argument with control parameters needed for optimization routine. If not specified, default values from <code>'CompetingRiskFrailtySurvfitControl'</code> will be taken; otherwise the values, which are defined by user, will overwrite the defaults.
<code>risk.names</code>	a character vector containing the names of competing risks. Its elements should be in the order, in which the columns describing the resulting states (or risks) have been supplied as a <code>'status'</code> parameter in the call of the <code>'CompetingRiskFrailtySurv'</code> function.

Details

A formula argument can be specified in different ways. The response or left part of the formula can be either an object of class `'CompetingRiskFrailtySurv'` or it can be a call to the `'CompetingRiskFrailtySurv'` function with its arguments specified in it. The right part may be left unspecified (without `'~'`) or it can be of `'~1'` type. In this case no covariates are supplied to the model, hence only baseline hazard will be adopted. In the case with covariates these must be supplied explicitly through their names in a standard way. In all cases the `'formula'` argument must be supplied.

If factor variables are to be considered, their reference categories should be specified before applied for optimization.

A `'control'` argument can be absent, in this case the defaults from the `'CompetingRiskFrailtySurvfitControl'` will be taken; or it can be either a list with some or all components of it specified by their names with according values, or it can be a call to the `'CompetingRiskFrailtySurvfitControl'` function with parameters specified in it.

The most of computation and optimization is implemented in the "internal" function `'CompetingRiskFrailtyOptim'`, which is in the body of `'CompetingRiskFrailtySurvfitCreate'`.

The working spline bases are ones consisted from truncated polynomials. Other alternative could be B-splines, but has not been implemented yet and might be considered in the following update of the package.

The computation procedure can take considerable time, depending on data supplied. All evaluations are written in R-code, so no external program code like C or Fortran has been used. We advice initially to set the number of iterations not to be too large, and eventually increase it if needed.

Value

An object of class `'CompetingRiskFrailtySurvfit'`. It has methods for `'print'` and `'plot'`.

Author(s)

Pavel Khomski <pkhomski@wiwi.uni-bielefeld.de>

References

Kauermann G. and Khomski P. (2006). *Full time or part time reemployment: a competing risk model with frailties and smooth effects using a penalty based approach*, to appear.

See Also

[CompetingRiskFrailtySurv](#), [CompetingRiskFrailtySurvfitControl](#), [CompetingRiskFrailtySurv](#)

Examples

```
#data(unemployed.comp.risk)
#ID<-unemployed.comp.risk$id
#surv.time<-unemployed.comp.risk$urv.time
#status<-unemployed.comp.risk[,3:4]
#my.Surv<-CompetingRiskFrailtySurv(ID=ID, surv.time=surv.time, status=status)
#my.Survfit<-CompetingRiskFrailtySurvfitCreate(my.Surv~nationality+gender+age+training+unive
#                                     data=unemployed.comp.risk, risk.names=c("full.
#evaluation takes ca. 25 minuts
```

CompetingRiskFrailtySurvfitObject

Fitted Object

Description

An object of the class 'CompetingRiskFrailtySurvfit' returned as a list by the 'CompetingRiskFrailtySurvfitCreate' function and representing a fitted competing risks hazard model with frailties and varying coefficients. It has methods for generic functions print and plot.

Value

L	number of comepting risks.
M.list	a list with elements specifying the number of chosen gamma distributions of frailties for each competing risk.
fixed.coef.optim	named vector of estimates of fixed parameters, beta, of the model.
random.coef.optim	named vector of estimates of random parameters, u, of the model.
penalty.varying.optim	named vector of penalty values for the random parts of the varying coefficients.
penalty.weights.optim	optimal value of the penalty parameter of the frailty mixture density.
grid.frame	a frame of grid values for plotting of the varying coefficients.
varying.list	a frame containing the values of varying coefficients.
deviation.list	a list containing the values of deviation for varying coefficients.

`frailty.list` a list of frailty estimates, each component is a vector according to the competing risk.
`mixture.weights` a vector of estimated mixture weights for the components of the mixture frailty distribution.
`aic.optim` an optimal aic value of the model.
`df.weights.optim` an optimal value of the degrees of freedom for mixture weights.
`log.lik.margin.optim` an optimal value of the marginal log likelihood of the model.
`p` number of covariates; for factors including their categories (excluding reference category).
`factor.names` covariates' names; for factors the names of categories (excluding reference category).
`risk.names` a character vector with the names of competing risks.

Author(s)

Pavel Khomski <pkhomski@wiwi.uni-bielefeld.de>

References

Kauermann G. and Khomski P. (2006). *Full time or part time reemployment: a competing risk model with frailties and smooth effects using a penalty based approach*, to appear.

See Also

[CompetingRiskFrailtySurv](#), [CompetingRiskFrailtySurvfitCreate](#)

plot.CompetingRiskFrailtySurvfit
Plots the Smoothed Varying Coefficients

Description

'plot.CompetingRiskFrailtySurvfit' is a proposed function for plotting of smooth components of the model, one plot for each baseline and covariate (or factor level) effect, for each competing risk.

Usage

```
## S3 method for class 'CompetingRiskFrailtySurvfit':
plot(x, ...)
```

Arguments

`x` object of class 'CompetingRiskFrailtySurvfit'.
`...` additional plot parameters.

Details

All plots will be made in the device, which is specified by the user. One can use the object components for producing his own plots.

Author(s)

Pavel Khomski <pkhomski@wiwi.uni-bielefeld.de>

References

Kauermann G. and Khomski P. (2006). *Full time or part time reemployment: a competing risk model with frailties and smooth effects using a penalty based approach*, to appear.

See Also

`print.CompetingRiskFrailtySurvfit`

Examples

```
#pdf(file="myplot.pdf",h=12,w=12)
#layout(matrix(1:16,ncol=4,byrow=T),widths=c(4,4,4,4, 4,4,4,4, 4,4,4,4, 4,4,4,4))
#par(mar=c(2,2,3,1.5))
#plot(my.Survfit)
#dev.off()
```

```
print.CompetingRiskFrailtySurvfit
      Prints an Object of class 'CompetingRiskFrailtySurvfit'
```

Description

Prints estimates of fixed parameters of smoothing varying coefficients, penalty values for their random parts as well as mixture weights of frailty mixture distribution, their degrees of freedom and optimal AIC- and marginal log likelihood value of the model.

Usage

```
## S3 method for class 'CompetingRiskFrailtySurvfit':
print(x, ...)
```

Arguments

`x` object of class 'CompetingRiskFrailtySurvfit'.
`...` additional parameters for print.

Author(s)

Pavel Khomski <pkhomski@wiwi.uni-bielefeld.de>

References

Kauermann G. and Khomski P. (2006). *Full time or part time reemployment: a competing risk model with frailties and smooth effects using a penalty based approach*, to appear.

See Also

`plot.CompetingRiskFrailtySurvfit`

Examples

```
#print.CompetingRiskFrailtySurvfit(my.Survfit)
```

```
unemployed.comp.risk
```

Sample of German Unemployed

Description

This data give the time spent in the unemployment spell and two competing risks: full time and part time reemployment as possible resulting states for leaving that spell for a sample of german unemployed individuals. The full data have been recieved from the *Deutsches Institut fuer Wirtschaftsforschung*, s. www.diw.de and were analysed in the larger study.

Usage

```
data(unemployed.comp.risk)
```

Format

A data frame containing 500 sampled observations. Description of columns:

- id Observations or spells are clustered according to the identification number of the person.
- surv.time Time spent in the unemployment spell, measured in months. The spells which lasted more than 36 months have been truncated on 36 months and denoted as censored.
- full.time binary indicator for leaving an unemployment spell resulting in a full time job.
- part.time binary indicator for leaving an unemployment spell resulting in a part time job.
- nationality Nationality of the unemployed person, german vs. others.
- gender Gender of the person.
- age Age has been categorized in young (till 25), middleage and old (over 50).
- training Whether an individual received a professional training.
- university Whether an individual has a university degree or comparable.

Source

Socio Economic Panel (SOEP), s. www.diw.de/deutsch/sop.

References

Kauermann G. and Khomski P. (2006). Additiv two way hazards model with varying coefficients, in press.

Index

*Topic **datasets**

unemployed.comp.risk, 9

*Topic **hplot**

plot.CompetingRiskFrailtySurvfit,
7

*Topic **list**

CompetingRiskFrailtySurvfitControl,
2

*Topic **misc**

CompetingRiskFrailtySurv, 1
CompetingRiskFrailtySurvfitObject,
6

*Topic **nonparametric**

CompetingRiskFrailtySurvfitCreate,
4

*Topic **print**

print.CompetingRiskFrailtySurvfit,
8

*Topic **smooth**

CompetingRiskFrailtySurvfitCreate,
4

*Topic **survival**

CompetingRiskFrailtySurvfitCreate,
4

CompetingRiskFrailtySurv, 1, 5, 7

CompetingRiskFrailtySurvfitControl,
2, 5

CompetingRiskFrailtySurvfitCreate,
2, 4, 4, 7

CompetingRiskFrailtySurvfitObject,
5, 6

plot.CompetingRiskFrailtySurvfit,
7, 8

print.CompetingRiskFrailtySurvfit,
7, 8

unemployed.comp.risk, 9