Package ‘gcmrec’

April 17, 2009

Version 1.0-3

Date 2009-Feb-05

Depends R (>= 2.0.0), survrec

Title General class of models for recurrent event data

Author Juan R. González <jrgonzalez@creal.cat>, Elizabeth H. Slate <SlateEH@musc.edu>, and Edsel A Peña <pena@stat.sc.edu>

Maintainer Juan R González <jrgonzalez@creal.cat>

Description Parameters estimation of the general semiparametric model for recurrent event data proposed by Peña and Hollander.

License GPL (>= 2)

URL http://www.r-project.org

Encoding latin1

Repository CRAN

Date/Publication 2009-02-06 12:32:34

R topics documented:

`addCenTime` *Add censored time equal to 0*

Description

Add a new line to the dataframe with a censored time equal to 0 when the end of follow-up coincides to the last occurrence

Usage

```
addCenTime(datin)
```

Arguments

`datin`	Dataframe containing id, time and event variables. Another covariates are allowed

Value

A data frame with an added line (censored time equal to 0) for those subjects where the end of follow-up coincides to the last occurrence

Examples

```
library(survival)
data(bladder2)

# we compute the interocurrence time
bladder2$time<-bladder2$stop-bladder2$start

# If we execute:
#     gcmrec(Survr(id,time,event)~rx+size+number,data=bladder2,s=2060)

# We will obtain the following error message:
#   Error in Survr(id, time, event) : Data doesn't match...

# This means that we have some patients without right-censoring time. So,
# we understand that the last event coincides with the end of study.
# Consequently,we need to add a line with time 0 and status value equal
# to 0, too. To do so, we can use the function "addCenTime" as follows:

# for example:
#   bladder2[bladder2$id==12,]

#     id rx number size start stop event enum time
#  45 12  1      1    1     0    3     1    1    3
#  46 12  1      1    1     3   16     1    2   13
#  47 12  1      1    1    16   23     1    3    7
```

```
# there is no censored time for 12th patient. So, if we execute

bladderOK<-addCenTime(bladder2)

# we get

#       id rx number size start stop event enum time
#   45  12  1      1    1     0    3     1    1    3
#   46  12  1      1    1     3   16     1    2   13
#   47  12  1      1    1    16   23     1    3    7
#   471 12  1      1    1    16   23     0    3    0
```

---

gcmrec *General Class of Models for recurrent event data*

Description

Fits the parameters for the general semiparametric model for recurrent events proposed by Peña and Hollander (2004). This class of models incorporates an effective age function which encodes the changes that occur after each event occurrence such as the impact of an intervention, it allows for the modeling of the impact of accumulating event occurrences on the unit, it admits a link function in which the effect of possibly time-dependent covariates are incorporated, and it allows the incorporation of unobservable frailty components which induce dependencies among the inter-event times for each unit.

Usage

```
gcmrec(formula, data, effageData = NULL, s, Frailty = FALSE,
    alphaSeed, betaSeed, xiSeed, tol = 10^(-6), maxit = 100,
    rhoFunc = "alpha to k", typeEffage = "perfect",
    maxXi = "Newton-Raphson", se = "Information matrix",
    cancer = NULL)
```

Arguments

`formula` A formula object. If a formula object is supplied it must have a Survr object as the response on the left of the '~' operator and a term on the right. Covariates are needed.

`data` A data frame in which to interpret the variables named in the formula. This data frame must contain the variables called "id","time" and "event" for subject identification, time of interocurrence, and censored status (coded 1: event, 0:censored), respectively. Furthermore, we can have some covariates.

Alternatively, it can also be a list containing the elements "n" and "subjects". Number of subjects must be recorded in "n". The element "subject" must have the following elements: subj, k, tau, caltimes, gaptimes, intercepts, slopes, lastperrep, perrepind, effage, effagebegin, and covariate including this information:

subj: Subject number or identificator.
k: Number of recurrences (time 0 must be included).
tau: Administrative time, time of study termination.
caltimes: Calendar times at each recurrence (time 0 must be included).
gaptimes: Gap times at each recurrence (time 0 must be included).
intercepts: Intercept value for the effect after each recurrence.
slopes: Slope value for the effect after each each recurrence.
lastperrep: Element from Brown and Proschan minimal repair model.
perrepind: Element from Brown and Proschan minimal repair model.
effagebegin: Initial value for effective age.
effage: Effective age after each recurrence.
covariate: covariate value at each recurrence.

See either GeneratedData or hydraulic data sets as an example.

`effageData` List containing the information about effective age. The list must have the elements described in the option 2 of data argument. If NULL we generate these elements under perfect repair model or minimal repair one depending on the 'typeEffage' argument (see below).

`s` A selected calendar time.

`Frailty` Logical value. Is model with frailties fitted? If so parameters for General Class of Models with frailty component are estimated.

`alphaSeed` Seed value for α.

`betaSeed` Seed value for β.

`xiSeed` Seed value for ξ.

`tol` Tolerance for maximization procedures.

`maxit` Maximum number of iterations in maximization procedures.

`rhoFunc` A character string specifying the effects attributable to the accumulating event occurrences, $\rho(k;\alpha)$. Possible values are "Identity" for $\rho(k;\alpha) = 1$ or "alpha to k" for $\rho(k;\alpha) = \alpha^k$. The default is "alpha to k". Only the first words are required, e.g "Id","a". Future versions will include other functions such as Markovian model for tumor occurrences, $\rho(k;\alpha) = \alpha - k + 1$ proposed by Gail et al. (1980).

`typeEffage` Effective age function. Possible value are "perfect" or "minimal" for perfect repair model or minimal repair model, respectively. The default is "perfect". Only the first words are required, e.g "p","m"

`maxXi` Maximization method for marginal likelihood with respect to ξ. Possible values are "Newton-Raphson" for Newton-Raphson maximization procedure or "Brent" for Brent's method maximization in one dimension. The default value is "Newton-Raphson". Only the first words are required, e.g. "N","B"

`se` Standard errors of parameters. Possible values are 'Information matrix' or 'Jacknife' for inverse of the partial likelihood information matrix or jacnife estimates, respectively.

`cancer` Effective age for fitting a cancer model proposed by Gonzalez et al (2005). This variable contains the information of the effect of treatments administred after cancer relapses coded as "CR", "PR" or "SD" depending on if complete, partial, or null response (stable disease) is achieved. See lymphoma data set as an example.

Details

Estimation with frailties are implemented using expectation-maximization (EM) algorithm. In this procedure, we need to maximize the marginal likelihood with respect to ξ. This maximization is a one-dimensional maximization without derivarives. First we bracket the maximizing value, and the we obtain it using Brent's method in one dimension. When we implement this algorithm, we re-parameterize ξ using $\xi^* = log(\xi)$ to alleviate the problem of getting negative estimates for ξ. Iteration is terminated when successive values of $\xi/(1+\xi)$ differ by no more than the "tol" parameter. Maybe estimation under frailty model can be not too fast.

Value

a gcmrec object. Methods defined for gcmrec objects are provided for print and plot.

References

Pe\textasciitildena, E. and M. Hollander (2004). Mathematical Reliability: An Expository Perspective, Chapter 6. Models for Recurrent Events in Reliability and Survival Analysis, pp. 105-123. Kluwer Academic Publishers.

M. Gail, T Santner, and C Brown (1980). An analysis of comparative carfinogenesis experiments based on multiple times to tumor. *Biometrics* **36**, 255-266.

JR Gonzalez, E Peña, E Slate (2005). Modelling treatment effect after cancer relapses, with application to recurrences in indolent non-Hodgkin's lymphomas. *Stat Med*, submitted.

R. Brent. *Algorithms for Minimization Without Derivatives*. Prentice-Hall, New York, 1973.

Examples

```
###################################
## Models using different data formats
###################################

#
#    Data input as a data frame
#

#   We use the well-known bladder cancer data set from survival package

library(survival)
data(bladder)

# we compute the interocurrence time
bladder2$time<-bladder2$stop-bladder2$start
```

```
# If we execute:
#     gcmrec(Survr(id,time,event)~rx+size+number,data=bladder2,s=2060)

# We will obtain the following error message:
#   Error in Survr(id, time, event) : Data doesn't match...

# This means that we have some patients without right-censoring time. So,
# we understand that the last event coincides with the end of study.
# Consequently,we need to add a line with time 0 and status value equal
# to 0, too. To do so, we can use the function "addCenTime" as follows:

bladderOK<-addCenTime(bladder2)

# Now, we can fit the model using this new data set:

gcmrec(Survr(id,time,event)~rx+size+number,data=bladderOK,s=2060)


#
#    Data as a list. See either GeneratedData or hydraulic data
#                    sets as an example.
#

#
# We can fit the model by transforming our data in a data frame
# using "List.to.Dataframe" function:
#

data(hydraulic)
hydraulicOK<-List.to.Dataframe(hydraulic)
gcmrec(Survr(id,time,event)~covar.1+covar.2,data=hydraulicOK,s=4753)


#
# Our model allows us to incorporate effective age information
#
# To illustrate this example, we will use a simulated data set generated
# under the minimal repair model with probability of perfect repair equal to 0.6
#
# As we have the data in a list, first we need to obtain a data frame containing
# the time, event, and covariates information:
#

data(GeneratedData)
temp<-List.to.Dataframe(GeneratedData)

# then, we can fit the model incorporating the information about the effective
#        age in the effageData argument:

gcmrec(Survr(id,time,event)~covar.1+covar.2, data=temp,
                    effageData=GeneratedData, s=100)
```

```
######################################################################
##  How to fit minimal or perfect repair models, with and without frailties
######################################################################

# Model with frailties

mod.Fra<-gcmrec(Survr(id,time,event)~rx+size+number,data=bladderOK,s=2060,Frailty=TRUE)
print(mod.Fra)

# effective age function: perfect repair and minimal repair models
# (models without frailties)

data(readmission)

# perfect
mod.per<-gcmrec(Survr(id,time,event)~as.factor(dukes),data=readmission,
    s=3000,typeEffage="per")
print(mod.per)

# minimal
mod.min<-gcmrec(Survr(id,time,event)~as.factor(dukes),data=readmission,
   s=3000,typeEffage="min")
print(mod.min)

######################################################################
##  How to fit models with \rho function equal to identity
######################################################################

data(lymphoma)

gcmrec(Survr(id, time, event) ~ as.factor(distrib),
    data = lymphoma, s = 1000, Frailty = TRUE, rhoFunc = "Ident")

######################################################################
##  How to fit cancer model
######################################################################

mod.can<-gcmrec(Survr(id,time,event)~as.factor(distrib), data=lymphoma,
     s=1000, Frailty=TRUE,  cancer=lymphoma$effage)

#  standard errors can be obtained by adding se="Jacknife".
# This procedure can be very time consuming
```

`GeneratedData` *Simulated data set generated under the minimal repair model*

Description

This contains recurrent times under minimal repair model with probability of perfect repair equal to 0.6. Data are as a list (see gcmrec help).

Usage

```
data(GeneratedData)
```

---

graph.caltimes *Plot of recurrent events*

Description

Plots calendar times at succesive recurrences from a data set. Information about effective age and categories of covariates are allowed.

Usage

```
graph.caltimes(data, var = NULL, effageData = NULL, width = 2,
         lines = TRUE, sortevents = TRUE, ...)
```

Arguments

`data`	data frame containing id, time, event variables and some other covariates
`var`	categorical variable
`effageData`	effective age function information
`width`	point width
`lines`	Are horizontal lines printed? The default is TRUE
`sortevents`	Are events sorted? The default is TRUE
`...`	other graphical parameters

Examples

```
# with data in a data frame
library(survival)
data(bladder2)
bladder2$time<-bladder2$stop-bladder2$start

graph.caltimes(bladder2)

# or data in a list

data(hydraulic)
graph.caltimes(hydraulic)

# We can print some covariate as follows:
graph.caltimes(bladder2,bladder2$rx)
```

hydraulic — *hydraulic load-haul-dump (LHD) subsystems*

Description

Hydraulic load-haul-dump (LHD) subsystems used in moving ore and rock in underground mines in Sweden. The data set provides the calendar times (in hours), excluding repair or down times, of the successive failures of n = 6 such systems during the two-year development phase.

Usage

```
data(hydraulic)
```

Source

Kumar, D. and B. Klefsjo (1992). Reliability analysis of hydraulic systems of lhd machines using the power law process model. *Reliability Engineering and System Safety* 35, 217- 224.

lymphoma — *Indolent non-Hodgkin's lymphomas*

Description

This contains cancer relapses times after first treatment in patients diagnosed with low grade lymphoma

Usage

```
data(lymphoma)
```

Format

This data frame contains the following columns:

id identificator of each subject. Repeated for each recurrence

time interocurrence or censoring time

event censoring status. All event are 1 for each subject excepting last one that it is 0

enum which lymphoma

delay delay between first symptom and date of first treatment as a continuous variable

age age at diagnosis

sex gender: 1:Males 2:Females

distrib lesions involved at diagnosis (0=Single, 1=Localized, 2=More than one nodal site, 3=Generalized

effage response achieved after treatment upon relapses, coded as CR: Complete remission, PR: Partial remission or SD: stable disease or null response.

Source

JR Gonzalez, E Peña, E Slate (2005). Modelling treatment effect after cancer relapses, with application to recurrences in indolent non-Hodgkin's lymphomas. *Stat Med*, submitted.

O. Servitje, F. Gallardo, T. Estrach, et al. (2002). Primary cutaneous marginal zone B-cell lymphoma: a clinical, histopathological, immunophenotypic and molecular genetic study of 22 cases. *Br J Dermatol*, 147:1147-1158.

`plot.gcmrec` *Plot Method for an object of class 'gcmrec'.*

Description

Plots estimated baseline survival and hazard functions from an object of class 'gcmrec'.

Usage

```
## S3 method for class 'gcmrec':
plot(x, type.plot = "surv", ...)
```

Arguments

`x`	Object of class `gcmrec` (output from calling `gcmrec` function).
`type.plot`	a character string specifying the type of curve. Possible value are "hazard", or "survival". The default is "hazard". Only the first words are required, e.g "haz", "su"
`...`	Other graphical parameters

Value

Print a plot of class `gcmrec`

See Also

`print.gcmrec`

Examples

```
data(lymphoma)

mod<-gcmrec(Survr(id,time,event)~as.factor(distrib),data=lymphoma,s=1000)

# baseline survivor function

plot(mod)

# baseline hazard function
```

```
plot(mod,type="haz")
```

`print.gcmrec` *Print a Short Summary of parameter estimates of a general class of models for reccurrent event data*

Description

Prints a short summary of 'gcmrec' object

Usage

```
## S3 method for class 'gcmrec':
print(x, digits = max(options()$digits - 4, 3), ...)
```

Arguments

`x`	the result of a call to the gcmrec function
`digits`	number of digits to print
`...`	other unsued arguments

Value

x, with the invisible flag set

See Also

`summary.gcmrec`, `gcmrec`

Examples

```
data(lymphoma)
mod<-gcmrec(Survr(id,time,event)~as.factor(distrib),data=lymphoma,s=1000)
print(mod)
```

readmission *Rehospitalization colorectal cancer*

Description

This contains rehospitalization times after surgery in patients diagnosed with colorectal cancer

Usage

```
data(readmission)
```

Format

This data frame contains the following columns:

id identificator of each subject. Repeated for each recurrence

enum which readmission

t.start start of interval (0 or previous recurrence time)

t.stop recurrence or censoring time

time interocurrence or censoring time

event censoring status. All event are 1 for each subject excepting last one that it is 0

chemo Did patient receive chemotherapy? 1: No; 2:Yes

sex gender: 1:Males 2:Females

dukes Dukes' tumoral stage: 1:A-B; 2:C 3:D

charlson Comorbidity Charlson's index. Time-dependent covariate. 0: Index 0; 1: Index 1-2; 3: Index >=3

Source

González, JR., Fernandez, E., Moreno, V. et al. Gender differences in hospital readmission among colorectal cancer patients. Journal of Epidemiology and Community Health. In press, 2005.

summary.gcmrec *summary of 'gcmrec'*

Description

This function returns hazard rations (HR) and its confidence intervals

Usage

```
## S3 method for class 'gcmrec':
summary(object, level = 0.95, len = 6, d = 2, lab="hr", ...)
```

Arguments

`object`	output from a call to gcmrec.
`level`	significance level of confidence interval. Default is 95%.
`len`	the desired number of digits after the decimal point. Default of 6 digits is used.
`d`	the total field width. Default is 6.
`lab`	label of printed results.
`...`	other unusued arguments.

Details

This function calls to intervals.gcmrec

Value

Prints HR and its confidence intervals. Confidence level is allowed (level argument)

See Also

`intervals.gcmrec`

Examples

```
data(lymphoma)
mod<-gcmrec(Survr(id,time,event)~as.factor(distrib),data=lymphoma,s=1000)
summary(mod)

# confidence interval at 99

summary(mod,level=0.99)
```

Index