Package 'reReg'

October 22, 2018

Title Recurrent Event Regression

Version 1.1.6

Description A collection of regression models for recurrent event process and failure time data. Available methods include these from Xu et al. (2017) <doi:10.1080/01621459.2016.1173557>, Lin et al. (2000) <doi:10.1111/1467-9868.00259>, Wang et al. (2001) <doi:10.1198/016214501753209031>, Ghosh and Lin (2003) <doi:10.1111/j.0006-341X.2003.00102.x>, and Huang and Wang (2004) <doi:10.1198/016214504000001033>.

Depends R (>= 3.4.0)

License GPL (>= 3)

Encoding UTF-8

LazyData true

URL http://github.com/stc04003/reReg

BugReports http://github.com/stc04003/reReg/issues

Imports BB, nleqslv, SQUAREM, survival, plyr, ggplot2, purrr, dplyr, tidyr, MASS, methods, tibble

RoxygenNote 6.1.0

NeedsCompilation yes

Suggests frailtypack

Author Sy Han (Steven) Chiou [aut, cre], Chiung-Yu Huang [aut]

Maintainer Sy Han (Steven) Chiou <schiou@utdallas.edu>

Repository CRAN

Date/Publication 2018-10-22 19:40:03 UTC

R topics documented:

`reReg-package` *reReg: Recurrent Event Regression*

Description

The package provides easy access to fit regression models to recurrent event data. The available implementations allow users to explore recurrent data through event plot and cumulative sample mean function plot, fit semiparametric regression models under different assumptions, and simulate recurrent event data.

Author(s)

Maintainer: Sy Han (Steven) Chiou <schiou@utdallas.edu>

Authors:

- Chiung-Yu Huang <ChiungYu.Huang@ucsf.edu>

References

Xu, G., Chiou, S.H., Huang, C.-Y., Wang, M.-C. and Yan, J. (2017). Joint Scale-change Models for Recurrent Events and Failure Time. *Journal of the American Statistical Association*, **112**(518): 796–805.

Lin, D., Wei, L., Yang, I. and Ying, Z. (2000). Semiparametric Regression for the Mean and Rate Functions of Recurrent Events. *Journal of the Royal Statistical Society: Series B (Methodological)*, **62**: 711–730.

Wang, M.-C., Qin, J., and Chiang, C.-T. (2001). Analyzing Recurrent Event Data with Informative Censoring. *Journal of the American Statistical Association*, **96**(455): 1057–1065.

Ghosh, D. and Lin, D.Y. (2003). Semiparametric Analysis of Recurrent Events Data in the Presence of Dependent Censoring. *Biometrics*, **59**: 877–885.

Huang, C.-Y. and Wang, M.-C. (2004). Joint Modeling and Estimation for Recurrent Event Processes and Failure Time Data. *Journal of the American Statistical Association*, **99**(468): 1153–1165.

See Also

Useful links:

- http://github.com/stc04003/reReg
- Report bugs at http://github.com/stc04003/reReg/issues

plot.reReg *Plotting Baseline Cumulative Rate Function and Baseline Cumulative Hazard Function*

Description

Plot the baseline cumulative rate function and the baseline cumulative hazard function (if applicable) for an reReg object.

Usage

```
## S3 method for class 'reReg'
plot(x, baseline = c("both", "rate", "hazard"),
  smooth = FALSE, control = list(), ...)
```

Arguments

x	an object of class reReg, usually returned by the reReg function.
baseline	a character string specifying which baseline function to plot. If baseline = "both" (default), both the baseline cumulative rate and baseline cumulative hazard function will be plotted in separate panels within the same display; if baseline = "rate", only the baseline cumulative rate function will be plotted; if baseline = "hazard", only the baseline cumulative hazard function will be plotted.
smooth	an optional logical value indicating whether to add a smooth curve (*loess* smooth).
control	a list of control parameters. See **Details**.
...	graphical parameters to be passed to methods. These include xlab, ylab and main.

Details

The argument control consists of options with argument defaults to a list with the following values:

xlab customizable x-label, default value is "Time".

ylab customizable y-label, default value is empty, e.g., "".

main customizable title, default value are "Baseline cumulative rate and hazard function" when baseline = "both", "Baseline cumulative rate function" when baseline = "rate", and "Baseline cumulative hazard function" when baseline = "hazard".

Value

A ggplot object.

See Also

reReg

Examples

```
data(readmission, package = "frailtypack")
fit <- reReg(reSurv(t.stop, id, event, death) ~ sex + chemo,
             data = subset(readmission, id < 50))
plot(fit)
plot(fit, baseline = "rate")
plot(fit, baseline = "rate", xlab = "Time (days)")
```

`plot.reSurv` *Produce Event Plot or Cumulative Sample Mean Function Plot*

Description

Plot the event plot or the cumulative sample mean (CSM) function for an `reSurv` object.

Usage

```
## S3 method for class 'reSurv'
plot(x, CSM = FALSE, order = TRUE, control = list(),
  ...)
```

Arguments

`x`	an object of class `reSurv`, usually returned by the `reSurv` function.
`CSM`	an optional logical value indicating whether the cumulative sample mean (CSM) function will be plotted instead of the event plot (default).
`order`	an optional logical value indicating whether the event plot (when `CSM = FALSE`) will be sorted by the terminal times.
`control`	a list of control parameters. See **Details**.
`...`	graphical parameters to be passed to methods. These include `xlab`, `ylab` and `main`.

Details

The argument `control` consists of options with argument defaults to a list with the following values:

xlab customizable x-label, default value is "Time".

ylab customizable y-label, default value is "Subject" for event plot and "Cumulative mean" for CSM plot.

main customizable title, the default value is "Recurrent event plot" when `CSM = FALSE` and "Sample cumulative mean function plot" when `CSM = TRUE`.

terminal.name customizable label for terminal event, default value is "Terminal event".

recurrent.name customizable legend title for recurrent event, default value is "Recurrent events".

recurrent.types customizable label for recurrent event type, default value is `NULL`.

alpha between 0 and 1, controls the transparency of points.

The `xlab`, `ylab` and `main` parameters can also be passed down without specifying a `control` list.

Value

A ggplot object.

See Also

reSurv

Examples

```
data(readmission, package = "frailtypack")
reObj <- with(subset(readmission, id <= 10), reSurv(t.stop, id, event, death))

## Event plots:
## Default labels
plot(reObj)
plot(reObj, order = FALSE)
## User specified labels
plot(reObj, control = list(xlab = "User xlab", ylab = "User ylab", main = "User title"))

## With multiple hypothetical event types
set.seed(1)
reObj2 <- with(readmission, reSurv(t.stop, id, event * sample(1:3, 861, TRUE), death))
plot(reObj2)

## CSM plots
plot(reObj, CSM = TRUE)
```

---

plotCSM *Produce Cumulative Sample Mean Function Plots*

Description

Plot the cumulative sample mean function (CSM) for an reSurv object. The function is similar to plot.reSurv but with more flexible options.

Usage

```
plotCSM(formula, data, onePanel = FALSE, adjrisk = TRUE,
  control = list(), ...)
```

Arguments

formula	a formula object, with the response on the left of a "~" operator, and the predictors on the right. The response must be a recurrent event survival object as returned by function reSurv.
data	an optional data frame in which to interpret the variables occurring in the "formula".

onePanel	an optional logical value indicating whether cumulative sample means (CSM) will be plotted in the same panel. This is useful when comparing CSM from different groups.
adjrisk	an optional logical value indicating whether risk set will be adjusted. See **Details**.
control	a list of control parameters.
...	graphical parameters to be passed to methods. These include xlab, ylab and main.

Details

When adjrisk = TRUE, the plotCSM is equivalent to the Nelson-Aalen estimator for the intensity function of the recurrent event process. When adjrisk = FALSE, the plotCSM does not adjust for the risk set and assumes all subjects remain at risk after the last observed recurrent event. This is known as the survivor rate function. The argument control consists of options with argument defaults to a list with the following values:

xlab customizable x-label, default value is "Time".

ylab customizable y-label, default value is "Cumulative mean".

main customizable title, default value is "Sample cumulative mean function plot".

The xlab, ylab and main parameters can also be passed down without specifying a control list.

Value

A ggplot object.

See Also

reSurv, plot.reSurv

Examples

```
data(readmission, package = "frailtypack")
plotCSM(reSurv(t.stop, id, event, death) ~ 1, data = readmission)
plotCSM(reSurv(t.stop, id, event, death) ~ sex, data = readmission)
plotCSM(reSurv(t.stop, id, event, death) ~ sex, data = readmission, onePanel = TRUE)
```

---

plotEvents *Produce Event Plots*

Description

Plot the event plot for an reSurv object. The function is similar to plot.reSurve but with more flexible options.

Usage

```
plotEvents(formula, data, order = TRUE, control = list(), ...)
```

Arguments

formula	a formula object, with the response on the left of a "~" operator, and the predictors on the right. The response must be a recurrent event survival object as returned by function reSurv.
data	an optional data frame in which to interpret the variables occurring in the "formula".
order	an optional logical value indicating whether the event plot will be sorted by the terminal times.
control	a list of control parameters.
...	graphical parameters to be passed to methods. These include xlab, ylab and main.

Details

The argument control consists of options with argument defaults to a list with the following values:

xlab customizable x-label, default value is "Time".

ylab customizable y-label, default value is "Subject".

main customizable title, default value is "Recurrent event plot".

terminal.name customizable label for terminal event, default value is "Terminal event".

recurrent.name customizable legend title for recurrent event, default value is "Recurrent events".

recurrent.types customizable label for recurrent event type, default value is NULL.

alpha between 0 and 1, controls the transparency of points.

The xlab, ylab and main parameters can also be passed down without specifying a control list. See **Examples**.

Value

A ggplot object.

See Also

reSurv, plot.reSurv

Examples

```
data(readmission, package = "frailtypack")
plotEvents(reSurv(t.stop, id, event, death) ~ 1, data = readmission)

## Separate plots by gender
plotEvents(reSurv(t.stop, id, event, death) ~ sex, data = readmission)

## Separate plots by gender and chemo type
```

```
plotEvents(reSurv(t.stop, id, event, death) ~ sex + chemo, data = readmission)

## With multiple hypothetical event types
plotEvents(reSurv(t.stop, id, event * sample(1:3, 861, TRUE), death) ~
  sex + chemo, data = readmission)
```

`plotHaz` *Produce the Baseline Cumulative Hazard Function for the Censoring Time*

Description

Plot the baseline cumulative hazard function for an `reReg` object.

Usage

```
plotHaz(x, smooth = FALSE, control = list(), ...)
```

Arguments

`x`	an object of class `reReg`, usually returned by the `reReg` function.
`smooth`	an optional logical value indicating whether the *loess* smoothing will be applied.
`control`	a list of control parameters.
`...`	graphical parameters to be passed to methods. These include `xlab`, `ylab` and `main`.

Details

The argument `control` consists of options with argument defaults to a list with the following values:

xlab customizable x-label, default value is "Time".

ylab customizable y-label, default value is empty, e.g., "".

main customizable title, default value is "Baseline cumulative hazard function".

These arguments can also be passed down without specifying a `control` list. See **Examples**.

Value

A `ggplot` object.

See Also

`reReg` `plot.reReg`

Examples

```
## readmission data
data(readmission, package = "frailtypack")
set.seed(123)
fit <- reReg(reSurv(t.stop, id, event, death) ~ sex + chemo,
             data = subset(readmission, id < 50),
             method = "am.XCHWY", se = "resampling", B = 20)
## Plot both the baseline cumulative rate and hazard function
plot(fit)
## Plot baseline cumulative hazard function
plotHaz(fit)
## Plot with user-specified labels
plotHaz(fit, control = list(xlab = "User xlab", ylab = "User ylab", main = "User title"))
```

---

plotRate *Plotting the Baseline Cumulative Rate Function for the Recurrent Event Process*

Description

Plot the baseline rate function for an reReg object.

Usage

```
plotRate(x, smooth = FALSE, control = list(), ...)
```

Arguments

x	an object of class reReg, usually returned by the reReg function.
smooth	an optional logical value indicating whether the *loess* smoothing will be applied.
control	a list of control parameters.
...	graphical parameters to be passed to methods. These include xlab, ylab and main.

Details

The argument control consists of options with argument defaults to a list with the following values:

xlab customizable x-label, default value is "Time".

ylab customizable y-label, default value is empty, e.g., "".

main customizable title, default value is "Baseline cumulative rate function".

These arguments can also be passed down without specifying a control list. See **Examples**.

Value

A ggplot object.

See Also

reReg plot.reReg

Examples

```
## readmission data
data(readmission, package = "frailtypack")
set.seed(123)
fit <- reReg(reSurv(t.stop, id, event, death) ~ sex + chemo,
                   data = subset(readmission, id < 50),
                   method = "am.XCHWY", se = "resampling", B = 20)
## Plot both the baseline cumulative rate and hazard function
plot(fit)
## Plot baseline cumulative rate function
plotRate(fit)
## Plot with user-specified labels
plotRate(fit, xlab = "User xlab", ylab = "User ylab", main = "User title")
plotRate(fit, control = list(xlab = "User xlab", ylab = "User ylab", main = "User title"))
```

---

reReg *Fits Semiparametric Regression Models for Recurrent Event Data*

Description

Fits a semiparametric regression model for the recurrent event data. The rate function of the underlying process for the recurrent event process can be specified as a Cox-type model, an accelerated mean model, or a generalized scale-change model. See details for model specifications.

Usage

```
reReg(formula, data, B = 200, method = c("cox.LWYY", "cox.GL",
  "cox.HW", "am.GL", "am.XCHWY", "sc.XCYH"), se = c("NULL", "bootstrap",
  "resampling"), contrasts = NULL, control = list())
```

Arguments

formula	a formula object, with the response on the left of a "~" operator, and the predictors on the right. The response must be a recurrent event survival object as returned by function reSurv.
data	an optional data frame in which to interpret the variables occurring in the "formula".
B	a numeric value specifies the number of resampling for variance estimation. When B = 0, variance estimation will not be performed.
method	a character string specifying the underlying model. See **Details**.
se	a character string specifying the method for standard error estimation. See **Details**.
contrasts	an optional list.
control	a list of control parameters.

Details

Suppose the recurrent event process and the failure events are observed in the time interval $t \in [0, \tau]$, for some constant τ. We formulate the rate function, $\lambda(t)$, for the recurrent event process and the hazard function, $h(t)$, for the censoring time under the following model specifications:

Cox-type model:

$$\lambda(t) = Z\lambda_0(t)e^{X^\top \alpha}, h(t) = Zh_0(t)e^{X^\top \beta},$$

Accelerated mean model:

$$\lambda(t) = Z\lambda_0(te^{X^\top \alpha})e^{X^\top \alpha}, h(t) = Zh_0(te^{X^\top \beta})e^{X^\top \beta},$$

Scale-change model:

$$\lambda(t) = Z\lambda_0(te^{X^\top \alpha})e^{X^\top \beta},$$

where $\lambda_0(t)$ is the baseline rate function, $h_0(t)$ is the baseline hazard function, X is a n by p covariate matrix and α, Z is an unobserved shared frailty variable, and β are unknown p-dimensional regression parameters.

The `reReg` function fits models with the following available methods:

`method = "cox.LWYY"` assumes the Cox-type model with `Z = 1` and requires independent censoring. The returned result is equivalent to that from `coxph`. See reference Lin et al. (2000).

`method = "cox.HW"` assumes the Cox-type model with unspecified `Z`, thus accommodate informative censoring. See the references See reference Wang, Qin and Chiang (2001) and Huang and Wang (2004).

`method = "am.GL"` assumes the accelerated mean model with `Z = 1` and requires independent censoring. See the reference Ghosh and Lin (2003).

`method = "am.XCHWY"` assumes the accelerated mean model with unspecified `Z`, thus accommodate informative censoring. See the reference Xu et al. (2017).

`method = "sc.XCYH"` assumes the generalized scale-change model, and includes the methods `"cox.HW"` and `"am.XCHWY"` as special cases. Informative censoring is accounted for through the unspecified frailty variable `Z`. The methods also provide a hypothesis test of these submodels.

The available methods for variance estimation are:

`NULL` variance estimation will not be performed. This is equivalent to setting `B = 0`.

`"resampling"` performs the efficient resampling-based sandwich estimator that works with methods `"cox.HW"`, `"am.XCHWY"` and `"sc.XCYH"`.

`"bootstrap"` works with all fitting methods.

The `control` list consists of the following parameters:

`tol` absolute error tolerance.

`a0, b0` initial guesses used for root search.

`solver` the equation solver used for root search. The available options are `BB::BBsolve`, `BB::dfsane`, `BB:BBoptim`, and `optim`.

`parallel` an logical value indicating whether parallel computation will be applied when `se = "bootstrap"` is called.

`parCl` an integer value specifying the number of CPU cores to be used when `parallel = TRUE`. The default value is half the CPU cores on the current host.

References

Xu, G., Chiou, S.H., Huang, C.-Y., Wang, M.-C. and Yan, J. (2017). Joint Scale-change Models for Recurrent Events and Failure Time. *Journal of the American Statistical Association*, **112**(518): 796–805.

Lin, D., Wei, L., Yang, I. and Ying, Z. (2000). Semiparametric Regression for the Mean and Rate Functions of Recurrent Events. *Journal of the Royal Statistical Society: Series B (Methodological)*, **62**: 711–730.

Wang, M.-C., Qin, J., and Chiang, C.-T. (2001). Analyzing Recurrent Event Data with Informative Censoring. *Journal of the American Statistical Association*, **96**(455): 1057–1065.

Ghosh, D. and Lin, D.Y. (2003). Semiparametric Analysis of Recurrent Events Data in the Presence of Dependent Censoring. *Biometrics*, **59**: 877–885.

Huang, C.-Y. and Wang, M.-C. (2004). Joint Modeling and Estimation for Recurrent Event Processes and Failure Time Data. *Journal of the American Statistical Association*, **99**(468): 1153–1165.

See Also

`reSurv`, `simDat`

Examples

```
## readmission data
data(readmission, package = "frailtypack")
set.seed(123)
## Accelerated Mean Model
(fit <- reReg(reSurv(t.stop, id, event, death) ~ sex + chemo,
              data = subset(readmission, id < 50),
              method = "am.XCHWY", se = "resampling", B = 20))
summary(fit)

## Generalized Scale-Change Model
set.seed(123)
(fit <- reReg(reSurv(t.stop, id, event, death) ~ sex + chemo,
              data = subset(readmission, id < 50),
              method = "sc.XCYH", se = "resampling", B = 20))
summary(fit)
```

reSurv *Create an* reSurv *Object*

Description

Create a recurrent event survival object, used as a response variable in reReg.

Usage

```
reSurv(time1, time2, id, event, status, origin = 0)

is.reSurv(x)
```

Arguments

time1	when "time2" is provided, this vector is treated as the starting time for the gap time between two successive recurrent events. In the absence of "time2", this is the observation time of recurrence on calendar time scale, in which, the time corresponds to the time since entry/inclusion in the study.
time2	an optional vector for ending time for the gap time between two successive recurrent events.
id	subject's id.
event	a binary vector used as the recurrent event indicator. event = 1 for recurrent times.
status	a binary vector used as the status indicator for the terminal event. status = 0 for censored times.
origin	a numerical vector indicating the time origin of subjects. When origin is a scalar, reSurv assumes all subjects have the same origin. Otherwise, origin needs to be a numerical vector, with length equals to the number of subjects. In this case, each element corresponds to different origins for different subjects. This argument is only needed when "time2" is missing.
x	an reSurv object.

Examples

```
data(readmission, package = "frailtypack")
attach(readmission)
reSurv(t.stop, id, event, death)
reSurv(t.start, t.stop, id, event, death)
detach(readmission)
```

simDat *Function to generate simulated data*

Description

The function simDat generates simulated recurrent event data from either a Cox-type model, an accelerated mean model, or a scale-change model. The censoring time could be either independent (given covariates) or informative. The simulated data is used for illustration.

Usage

```
simDat(n, a, b, indCen = TRUE, type = c("cox", "am", "sc"), tau = 60,
  summary = FALSE)
```

Arguments

n	number of observation.
a	a numeric vector of parameter of length 2.
b	a numeric vector of parameter of length 2.
indCen	a logical value indicating whether the censoring assumption is imposed. When indCen = TRUE, we set $Z = 1$. Otherwise, Z is generated from a gamma distribution with mean 1 and variance 0.25 (e.g., rgamma(1, 4, 4)). See **Details**.
type	a character string specifying the underlying model. See **Details**
tau	a numeric value specifying the maximum observation time.
summary	a logical value indicating whether a brief data summary will be printed.

Details

The function simDat generates simulated recurrent event data under different scenarios based on the following assumptions. See **Details** in reReg for a more complete model assumptions.

type = "cox" generates recurrent event data from a Cox-type model with

$$\lambda(t) = Z\lambda_0(t)e^{X^\top a}, h(t) = Zh_0(t)e^{X^\top b}.$$

type = "am" generates recurrent event data from an accelerated mean model with

$$\lambda(t) = Z\lambda_0(te^{X^\top a})e^{X^\top a}, h(t) = Zh_0(te^{X^\top b})e^{X^\top b}.$$

type = "sc" generates recurrent event data from a generalized scale-change model with

$$\lambda(t) = Z\lambda_0(te^{X^\top a})e^{X^\top b}, h(t) = Zh_0(te^{X^\top a})e^{X^\top b}.$$

Let D be the informative failure time with the above hazard function. An non-informative failure time, C, is generated separately from an exponential distribution with mean 80. The observed follow-up time is then taken to be $min(D, C, \tau)$. We further assume

$$\lambda_0(t) = \frac{2}{1+t}, h_0(t) = \frac{1}{8(1+t)}.$$

Two covariates are considered; `x1` follows a Bernoulli distribution with probability 0.5 and `x2` follows a standard normal distribution.

See Also

`reReg`

Examples

```
set.seed(123)
simDat(200, c(-1, 1), c(-1, 1), summary = TRUE)
```

Index