Package ‘genSurv’

November 9, 2015

Encoding UTF-8
Type Package
Depends R (>= 2.12.0)
Suggests survival
Title Generating Multi-State Survival Data
Version 1.0.3
Date 2015-11-09
Author Artur Araujo, Luis Meira-Machado
<lmachado@math.uminho.pt> and Susana Faria
<sfaria@math.uminho.pt>
Maintainer Artur Araujo <artur.stat@gmail.com>
Description Generation of survival data with one (binary)
time-dependent covariate. Generation of survival data arising
from a progressive illness-death model.
License GPL (>= 2)
LazyLoad yes
LazyData yes
NeedsCompilation yes
Repository CRAN
Date/Publication 2015-11-09 15:27:19

R topics documented:

  genSurv-package .......................................................... 2
  as.CMM ........................................................................... 3
  as.TDCM ........................................................................... 4
  as.THMM ......................................................................... 5
  genCMM .......................................................................... 6
  genCPHM ......................................................................... 8
  genTDCM ......................................................................... 9
  genTHMM ...................................................................... 11
Description

The genSurv software permits to generate data with one binary time-dependent covariate and data stemming from a progressive illness-death model.

Details

Package: genSurv
Type: Package
Version: 1.0.3
Date: 2015-11-09
License: GPL (>= 2)
LazyLoad: yes
LazyData: yes

Author(s)

Artur Araújo, Luís Meira-Machado <lmachado@math.uminho.pt>
and Susana Faria <sfaria@math.uminho.pt>
Maintainer: Artur Araújo <artur.stat@gmail.com>

References

as.CMM


---

**as.CMM**

*Coerce to an object of class CMM*

---

**Description**

Function to coerce objects of class TDCM and THMM to objects of class CMM.

**Usage**

```r
as.CMM(x)
is.CMM(x)
```

**Arguments**

- `x` Any R object.

**Value**

An object with two classes one being data.frame and the other CMM.

**Author(s)**

Artur Araújo, Luís Meira Machado and Susana Faria

**References**


**See Also**

as.TDCM, as.THMM, genCMM, genTDCM, genTHMM.
Examples

# generate TDCM data
tdcmdata <- genTDCM(n=100, dist="exponential", corr=0, dist.par=c(1,1),
model.cens="uniform", cens.par=1, beta=c(-3,2), lambda=10)

# coerce TDCM data to CMM data
cmmdata0 <- as.CMM(tdcmdata)
head(cmmdata0, n=20L)

# generate THMM data
thmmdata <- genTHMM( n=100, model.cens="uniform", cens.par=80, beta= c(0.09,0.08,-0.09),
covar=80, rate= c(0.05,0.04,0.05) )

# coerce THMM data to CMM data
thmmdata1 <- as.CMM(thmmdata)
head(thmmdata1, n=20L)

---

as.TDCM  
Coerce to an object of class TDCM

Description

Function to coerce objects of class CMM and THMM to objects of class TDCM.

Usage

as.TDCM(x)
is.TDCM(x)

Arguments

x  
Any R object.

Value

An object with two classes one being data.frame and the other TDCM.

Author(s)

Artur Araújo, Luís Meira Machado and Susana Faria

References


See Also

as.CMM, as.THMM, genCMM, genTDCM, genTHMM.

Examples

# generate CMM data
cmmdata <- genCMM( n=1000, model.cens="uniform", cens.par=2.5, beta=c(2,1,-1),
covar=10, rate=c(1,5,1,5,1,5) )

# coerce CMM data to TDCM data
tdcmdat0 <- as.TDCM(cmmdata)
head(tdcmdat0, n=20)

# generate THMM data
thmmdata <- genTHMM( n=100, model.cens="uniform", cens.par=80, beta= c(0.09,0.08,-0.09),
covar=80, rate= c(0.05,0.04,0.05) )

# coerce THMM data to TDCM data
tdcmdat1 <- as.TDCM(thmmdata)
head(tdcmdat1, n=20)

as.THMM  Coerce to an object of class THMM

Description

Function to coerce objects of class CMM and TDCM to objects of class THMM.

Usage

as.THMM(x)
is.THMM(x)

Arguments

x  Any R object.

Value

An object with two classes one being data.frame and the other THMM.
Author(s)

Artur Araújo, Luís Meira Machado and Susana Faria

References


See Also

as.CMM, as.TDCM, genCMM, genTDCM, genTHMM.

Examples

```r
# generate CMM data
cmmdata <- genCMM(n=1000, model.cens="uniform", cens.par=2.5, beta=c(2,1,-1),
covar=10, rate=c(1,5,1,5,1,5)
)

# coerce CMM data to THMM data
thmmdato <- as.THMM(cmmdata)
head(thmmdato, n=20)

# generate TDCM data
tdcmdato <- genTDCM(n=100, dist="exponential", corr=0, dist.par=c(1,1),
model.cens="uniform", cens.par=1, beta=c(-3,2), lambda=10)

# coerce TDCM data to THMM data
thmmdat1 <- as.THMM(tdcmdato)
head(thmmdat1, n=20)
```

Description

Generation of Cox Markov data from an illness-death model.

Usage

```r
genCMM(n, model.cens, cens.par, beta, covar, rate)
```
Arguments

- **n**: Sample size.
- **model.cens**: Model for censorship. Possible values are "uniform" and "exponential".
- **cens.par**: Parameter for the censorship distribution. Must be greater than 0.
- **beta**: Vector of three regression parameters for the three transitions: (beta_12, beta_13, beta_23).
- **covar**: Parameter for generating the time-fixed covariate. An uniform distribution is used.
- **rate**: Vector of dimension six: (shape1, scale1, shape2, scale2, shape3, scale3). A Weibull baseline hazard function is assumed (with two parameters) for each transition (see details below).

Details

The Weibull distribution with shape parameter $\lambda$ and scale parameter $\theta$ has hazard function given by:

$$\alpha(t) = \lambda \theta t^{\theta - 1}$$

Value

An object with two classes, `data.frame` and `CMM`. The data structure used for generating survival data from the Cox Markov Model (CMM) is similar as for the time-dependent Cox model (TDCM). In this case the data structure has one more variable representing the transition (variable `trans`). `trans=1` denotes the transition from State 1 to State 3 (without observing the intermediate event; State 2); `trans=2` denotes the transition from State 1 to State 2; and `trans=3` denotes the transition from State 2 to State 3 (absorbing).

Author(s)

Artur Araújo, Luís Meira Machado and Susana Faria

References


See Also

genCPHM, genTDCM, genTHMM.
Examples

cmmdata <- genCMM( n=1000, model.cens="uniform", cens.par=2.5, beta=c(2,1,-1),
covar=10, rate=c(1,5,1,5,1,5) )
head(cmmdata, n=20L)
library(survival)
fit_13<-coxph(Surv(start,stop,event)-covariate, data=cmmdata, subset=c(trans==1))
fit_13
fit_12<-coxph(Surv(start,stop,event)-covariate, data=cmmdata, subset=c(trans==2))
fit_12
fit_23<-coxph(Surv(start,stop,event)-covariate, data=cmmdata, subset=c(trans==3))
fit_23

Description

Generation of survival data from a Cox Proportional Hazard Model.

Usage

genCPHM(n, model.cens, cens.par, beta, covar)

Arguments

n Sample size.
model.cens Model for censorship. Possible values are "uniform" and "exponential".
cens.par Parameter for the censorship distribution. Must be greater than 0.
beta Regression parameter for the time-fixed covariate.
covar Parameter for generating the time-fixed covariate. An uniform distribution is used.

Value

An object with two classes, data.frame and CPHM.

Author(s)

Artur Araújo, Luís Meira Machado and Susana Faria

References


See Also
genCMM, genTDCM, genTHMM.
Examples

cphmdata <- genCphM(n=1000, model.cens="exponential", cens.par=2, beta= 2, covar=1)
head(cphmdata, n=20L)
library(survival)
fit<-coxph(Surv(time,status)~covariate,data=cphmdata)
summary(fit)

Description

Generating data from a Cox model with time-dependent covariates.

Usage

genTDCM(n, dist, corr, dist.par, model.cens, cens.par, beta, lambda)

Arguments

n Sample size.
dist Bivariate distribution assumed for generating the two covariates (time-fixed and
time-dependent). Possible bivariate distributions are "exponential" and "weibull"
(see details below).
corr Correlation parameter. Possible values for the bivariate exponential distribution
are between -1 and 1 (0 for independency). Any value between 0 (not included)
and 1 (1 for independency) is accepted for the bivariate weibull distribution.
dist.par Vector of parameters for the allowed distributions. Two (scale) parameters for
the bivariate exponential distribution and four (2 shape parameters and 2 scale
parameters) for the bivariate weibull distribution: (shape1, scale1, shape2, scale2).
See details below.
model.cens Model for censorship. Possible values are "uniform" and "exponential".
cens.par Parameter for the censorship distribution. Must be greater than 0.
beta Vector of two regression parameters for the two covariates.
lambda Parameter for an exponential distribution. An exponential distribution is assumed
for the baseline hazard function.

Details

The bivariate exponential distribution, also known as Farlie-Gumbel-Morgenstern distribution is
given by

\[ F(x, y) = F_1(x)F_2(y)[1 + \alpha(1 - F_1(x))(1 - F_2(y))] \]

for \( x \geq 0 \) and \( y \geq 0 \). Where the marginal distribution functions \( F_1 \) and \( F_2 \) are exponential with
scale parameters \( \theta_1 \) and \( \theta_2 \) and correlation parameter \( \alpha, -1 \leq \alpha \leq 1 \).
The bivariate Weibull distribution with two-parameter marginal distributions. Its survival function is given by

\[ S(x, y) = P(X > x, Y > y) = e^{-\left(\frac{x}{\theta_1})^{\beta_1} + (\frac{y}{\theta_2})^{\beta_2}\right)^{\delta}} \]

Where \( 0 < \delta \leq 1 \) and each marginal distribution has shape parameter \( \beta_i \) and a scale parameter \( \theta_i \), \( i = 1, 2 \).

Value

An object with two classes, data.frame and TDCM. To accommodate time-dependent effects, we used a counting process data-structure, introduced by Andersen and Gill (1982). In this data-structure, apart the time-fixed covariates (named covariate), an individual’s survival data is expressed by three variables: start, stop and event. Individuals without change in the time-dependent covariate (named tdcov) are represented by only one line of data, whereas patients with a change in the time-dependent covariate must be represented by two lines. For these patients, the first line represents the time period until the change in the time-dependent covariate; the second line represents the time period that passes from that change to the end of the follow-up. For each line of data, variables start and stop mark the time interval (start, stop) for the data, while event is an indicator variable taking on value 1 if there was a death at time stop, and 0 otherwise. More details about this data-structure can be found in papers by (Meira-Machado et al., 2009).

Author(s)

Artur Araújo, Luís Meira Machado and Susana Faria

References


See Also

gencmm, gencphm, genTHMM.
**Examples**

tdcmdata <- genTD CM(n=1000, dist="weibull", corr=0.8, dist.par=c(2,3,2,3),
model.cens="uniform", cens.par=2.5, beta=c(-3.3,4), lambda=1)
head(tdcmdata, n=20L)
library(survival)
fit1<-coxph(Surv(start,stop,event)~tdcov+covariate,data=tdcmdata)
summary(fit1)

tdcmdata2 <- genTD CM(n=1000, dist="exponential", corr=0, dist.par=c(1,1),
model.cens="uniform", cens.par=1, beta=c(-3,2), lambda=0.5)
head(tdcmdata2, n=20L)
fit2<-coxph(Surv(start,stop,event)~tdcov+covariate,data=tdcmdata2)
summary(fit2)

---

**Description**

Generation of survival data from a time-homogeneous Markov model.

**Usage**

genTHMM(n, model.cens, cens.par, beta, covar, rate)

**Arguments**

- **n**  
  Sample size.

- **model.cens**  
  Model for censorship. Possible values are "uniform" and "exponential".

- **cens.par**  
  Parameter for the censorship distribution. Must be greater than 0.

- **beta**  
  Vector of three regression parameters for the three transitions: (beta_12, beta_13, beta_23).

- **covar**  
  Parameter for generating the time-fixed covariate. An uniform distribution is used.

- **rate**  
  Vector of dimension three. We assume an exponential baseline hazard function with constant hazard rate for each transition.

**Value**

An object with two classes, `data.frame` and `THMM`. For generating survival data from the THMM model, the counting process data structure must contain the following variables: id, time, state, covariate. Each patient is identified by id. The variable time represents time for each interval of follow-up while variable state denotes the state of the individual. Variable covariate is the (time-fixed) covariate to be studied in the regression model. Individuals without change in the time dependent covariate are represented by two lines of data, whereas patients with a change in the time-dependent covariate must be represented by three lines.
Author(s)
Artur Araújo, Luís Meira Machado and Susana Faria

References

See Also
genCMM, genTDCM, genCPHM.

Examples
thmmdata <- genTHMM( n=100, model.cens="uniform", cens.par=80, beta= c(0.09,0.08,-0.09), covar=80, rate= c(0.05,0.04,0.05) )
head(thmmdata, n=20L)
Index

*Topic datagen
   genCMM, 6
   genCPHM, 8
   genTDCM, 9
   genTHMM, 11
*Topic manip
   as.CMM, 3
   as.TDCM, 4
   as.THMM, 5
*Topic methods
   as.CMM, 3
   as.TDCM, 4
   as.THMM, 5
*Topic package
   genSurv-package, 2
*Topic survival
   as.CMM, 3
   as.TDCM, 4
   as.THMM, 5
   genCMM, 6
   genCPHM, 8
   genTDCM, 9
   genTHMM, 11

   as.CMM, 3, 5, 6
   as.TDCM, 3, 4, 6
   as.THMM, 3, 5, 5
   genCMM, 3, 5, 6, 6, 8, 10, 12
   genCPHM, 7, 8, 10, 12
   genSurv-package, 2
   genTDCM, 3, 5–8, 9, 12
   genTHMM, 3, 5–8, 10, 11

   is.CMM (as.CMM), 3
   is.TDCM (as.TDCM), 4
   is.THMM (as.THMM), 5