Package ‘BivRec’

November 16, 2018

Type Package

Title Bivariate Alternating Recurrent Event Data Analysis

Date 2018-10-31

Version 1.0.0

Author Sandra Castro-Pearson [aut, cre],
Chi-Hyun Lee [aut],
Chiung-Yu Huang [aut],
Xianghua Luo [ctb]

Maintainer Sandra Castro-Pearson <cast0135@umn.edu>

Description Alternating recurrent event data arise frequently in biomedical and social sciences where 2 types of events such as hospital admissions and discharge occur alternatively over time.
As such we implement a collection of non-parametric and semiparametric methods to analyze such data.
The main functions are biv.rec.fit() and biv.rec.np(). Use biv.rec.fit() for estimation of covariate effects on the two alternating event gap times (xij and yij) using semiparametric methods. The method options are ``Lee.et.al'' and ``Chang''.
Use biv.rec.np() for estimation of the joint cumulative distribution function (cdf) for the two alternating events gap times (xij and yij) as well as the marginal survival function for type I gap times (xij) and the conditional cdf of the type II gap times (yij) given an interval of type I gap times (xij) in a non-parametric fashion.
The package also provides options to simulate and visualize the data and results of analysis.

BugReports https://github.com/SandraCastroPearson/BivRec/issues

Depends R (>= 3.2.0), survival, stats

Imports MASS, stringr, utils, knitr, rmarkdown, Rcpp, graphics,
grDevices

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0

LinkingTo Rcpp

NeedsCompilation yes

Repository CRAN

Date/Publication 2018-11-16 16:10:17 UTC

R topics documented:

`biv.rec.fit` *Semi-Parametric Accelerated Failure Time Analysis of Bivariate Alternating Recurrent Event Gap Time Data*

Description

This function allows the user to evaluate covariate effects on two alternating recurrent events gap times (referred as type I and type II gap times) under the assumption that the two gap times follow accelerated failure time (AFT) models. See details for the estimation methods provided.

Usage

```
biv.rec.fit(formula, data, method, CI)
```

Arguments

`formula` A formula with six variables indicating the bivariate alternating gap time response on the left of the ~ operator and the covariates on the right. The six variables on the left must have the same length and be given as **ID + episode + xij + yij + delta_x + delta_y ~ covariates**, where

- ID: A vector of subjects' unique identifier which can be numeric or character.
- episode: A vector indicating the episode of the bivariate alternating gap time pairs, e.g.: 1, 2, ..., m_i where m_i indicates the last episode for subject i.
- xij: A vector with the lengths of the type I gap times.
- yij: A vector with the lengths of the type II gap times.
- delta_x: A vector of indicators with values
  - 0 for the last episode for subject i (m_i) if subject was censored during period xij.
  - 1 otherwise.

A subject with only one episode (m_i=1) could have a 0 if he was censored during period xi1 or 1 if he was censored during period yi1. If delta_x is not provided estimation proceeds with the assumption that no subject was censored during period xij.

- delta_y: A vector of indicators with values
  - 0 for the last episode of subject i (m_i).
  - 1 otherwise.

  A subject with only one episode (m_i=1) will have one 0.
- covariates: the names of the covariates in the form covariate_1 + ... + covariate_p.

`data` A data frame that includes all the vectors/covariates listed in the formula above.

`method` A string indicating which method to use to estimate effects of the covariates. See details.

`CI` The level to be used for confidence intervals. Must be between 0.50 and 0.99, where 0.99 would give 99% CI. The default is 0.95. CI=NULL gives point estimates without confidence intervals.

Details

Two different estimation methods are available:

- method = "Lee.et.al" (default) is a U-statistics-based smooth estimating function approach. See Lee CH, Huang C-Y, Xu G, Luo X (2017) for further details.
- method = "Chang" is a rank-based estimating function approach. See Chang (2004) for further details. Note that following the Chang method, the variances of the estimated regression coefficients are approximated using the resampling techniques developed by Parzen, Wei and Ying (1994). This approximation requires extensive computing time for a relatively small sample size. In addition, using the Chang method does not guarantee convergence for the estimation of the coefficients.

Value

A BivRec list object containing:

- **covariate.effects:** A data frame summarizing effects of the covariates including the point estimate, standard error and confidence interval.
- **formula:** The formula used to specify components of bivariate recurrent response and covariates.

References

1. Chang S-H. (2004). Estimating marginal effects in accelerated failure time models for serial sojourn times among repeated events. Lifetime Data Analysis, 10: 175-190. https://doi.org/10.1023/B:LIDA.0000030202.20842.c9
2. Lee C, Huang CY, Xu G, Luo X (2017). Semiparametric regression analysis for alternating recurrent event data. Statistics in Medicine, 37: 996-1008. https://doi.org/10.1002/sim.7563

3. Parzen MI, Wei LJ, Ying Z (1994). A resampling method based on pivotal estimating functions. Biometrika, 81: 341-350. https://doi.org/10.1093/biomet/81.2.341

Examples

```
library(BivRec)
# Simulate bivariate alternating recurrent event data
set.seed(1234)
biv.rec.data <- biv.rec.sim(nsize=150, beta1=c(0.5,0.5), beta2=c(0,-0.5), tau_c=63, set=1.1)
# Apply Lee C, Huang CY, Xu G, Luo X (2017) method using one covariate
fit.lee <- biv.rec.fit(formula = id + epi + xij + yij + d1 + d2 ~ a1,
                data=biv.rec.data, method="Lee.et.al", CI=NULL)
fit.lee$covariate.effects
## Not run:

#This is an example with longer runtime.

library(BivRec)
# Simulate bivariate alternating recurrent event data
set.seed(1234)
biv.rec.data <- biv.rec.sim(nsize=150, beta1=c(0.5,0.5), beta2=c(0,-0.5), tau_c=63, set=1.1)

# Apply Lee C, Huang CY, Xu G, Luo X (2017) method using multiple covariates
# and 99% confidence intervals.
fit.lee <- biv.rec.fit(formula = id + epi + xij + yij + d1 + d2 ~ a1 + a2,
                data=biv.rec.data, method="Lee.et.al", CI=0.99)
fit.lee$covariate.effects


## End(Not run)
# To apply Chang (2004) method use method="Chang"
```

---

biv.rec.np — *Non-Parametric Analysis of Bivariate Alternating Recurrent Event Gap Time Data*

Description

This function allows the user to apply a non-parametric method to estimate the joint cumulative distribution function (cdf) for the two alternating events gap times (xij and yij) as well as the marginal survival function for type I gap times (xij) and the conditional cdf of the type II gap times (yij) given an interval of type I gap times (xij). See Huang and Wang (2005) for more details.

Usage

```
biv.rec.np(formula, data, CI, ai, u1, u2, conditional, given.interval,
  jointplot, marginalplot, condiplot)
```

Arguments

`formula` A formula with six variables indicating the bivariate alternating gap time response on the left of the ~ operator and a 1 on the right. The six variables on the left must have the same length and be given as **ID + episode + xij + yij + delta_x + delta_y ~ 1**, where

- ID: A vector of subjects' unique identifier which can be numeric or character.
- episode: A vector indicating the episode of the bivariate alternating gap time pairs, e.g.: 1, 2, ..., m_i where m_i indicates the last episode for subject i.
- xij: A vector with the lengths of the type I gap times.
- yij: A vector with the lengths of the type II gap times.
- delta_x: An optional vector of indicators with values:
  - 0 for the last episode for subject i (m_i) if subject was censored during period xij.
  - 1 otherwise.

  A subject with only one episode (m_i = 1) could have a 0 if he was censored during period xi1 or 1 if he was censored during period yi1. If delta_x is not provided estimation proceeds with the assumption that no subject was censored during period xij.
- delta_y: A vector of indicators with values:
  - 0 for the last episode of subject i (m_i).
  - 1 otherwise.

  A subject with only one episode (m_i = 1) will have one 0.

`data` A data frame that includes all the vectors listed in the formula.

`CI` The level for confidence intervals for joint cdf plot, marginal plot and conditional cdf. Must be between 0.50 and 0.99, where 0.99 would give 99% CI. Default is 0.95.

`ai` A real non-negative function of censoring time. See details.

`u1` A vector or single number to be used for estimation of joint cdf $P(X0 \leq u1, Y0 \leq u2)$ in the non-parametric method.

`u2` A vector or single number to be used for estimation of joint cdf $P(X0 \leq u1, Y0 \leq u2)$ in the non-parametric method.

`conditional` A logical value. If TRUE, this function will calculate the conditional cdf for the type II gap time given an interval of the type I gap time and the bootstrap standard error and confidence interval at the specified confidence level. Default is FALSE.

`given.interval` A vector c(v1, v2) that must be specified if conditional = TRUE. The vector indicates an interval for the type I gap time to use for estimation of the cdf of the type II gap time given this interval. If given.interval = c(v1, v2), the function calculates $P(Y0 \leq y|v1 \leq X0 \leq v2)$. The given values v1 and v2 must be in the range of gap times in the estimated marginal survival. Valid values for these times are given in the "Time" column of the marginal survival data frame that results from biv.rec.np().

`jointplot` A logical value. If TRUE (default), this function will create a 3D plot of the joint cdf for the two gap times with pointwise large sample confidence interval at the specified confidence level.

`marginalplot` A logical value. If TRUE (default), this function will plot the marginal survival function for the type I gap times with pointwise large sample confidence interval at the specified confidence level.

`condiplot` A logical value. Can only be TRUE if conditional=TRUE. If TRUE, this function will plot the conditional cdf with bootstrap confidence interval at the specified confidence level. Default is FALSE.

Details

ai indicates a real non-negative function of censoring times to be used as weights in the non-parametric method. This variable can take on values of 1 or 2 which indicate:

- 1: the weights are simply 1 for all subjects $a(C_i) = 1$ (default).
- 2: the weight for each subject is his/her censoring time $a(C_i) = C_i$.

For further information, see Huang and Wang (2005).

Value

Plots as specified from jointplot, marginalplot, conditional and a BivRec list object containing:

- **joint.cdf:** Data frame with joint cdf and standard error for the two alternating gap times.
- **marginal.survival:** Data frame with marginal survival for the first gap time and standard error.
- **conditional.cdf:** Data frame with conditional cdf, bootstrap standard error and bootstrap confidence interval.
- **formula:** The formula used to specify components of bivariate recurrent response.
- **ai:** The function of censoring time used as weights.

References

Huang CY, Wang MC (2005). Nonparametric estimation of the bivariate recurrence time distribution. Biometrics, 61: 392-402. doi.org/10.1111/j.1541-0420.2005.00328.x

Examples

```
library(BivRec)
# Simulate bivariate alternating recurrent event data
set.seed(1234)
biv.rec.data <- biv.rec.sim(nsize=150, beta1=c(0.5,0.5), beta2=c(0,-0.5), tau_c=63, set=1.1)
# Apply the non-parametric method of Huang and Wang (2005) and
# Visualize joint cdf and marginal survival results
nonpar.result <- biv.rec.np(formula = id + epi + xij + yij + d1 + d2 ~ 1,
          data=biv.rec.data, ai=1, u1 = c(2, 5, 10, 20), u2 = c(1, 5, 10, 15),
          conditional = FALSE, given.interval=c(0, 10), jointplot=TRUE,
          marginalplot = TRUE, condiplot = FALSE)
```

```
head(nonpar.result$joint.cdf)
head(nonpar.result$marginal.survival)

## Not run:
#This is an example with longer runtime.
library(BivRec)
# Simulate bivariate alternating recurrent event data
set.seed(1234)
biv.rec.data <- biv.rec.sim(nsize=150, beta1=c(0.5,0.5), beta2=c(0,-0.5), tau_c=63, set=1.1)

# Apply the non-parametric method of Huang and Wang (2005) and Visualize all results
nonpar.result <- biv.rec.np(formula = id + epi + xij + yij + d1 + d2 ~ 1,
          data=biv.rec.data, ai=1, u1 = c(2, 5, 10, 20), u2 = c(1, 5, 10, 15),
          conditional = TRUE, given.interval=c(0, 10), jointplot=TRUE,
          marginalplot = TRUE, condiplot = TRUE)
head(nonpar.result$joint.cdf)
head(nonpar.result$marginal.survival)
head(nonpar.result$conditional.cdf)

## End(Not run)
```

`biv.rec.plot` *Bivariate Alternating Recurrent Series Plotting*

Description

This function plots bivariate recurrent event gap times.

Usage

```
biv.rec.plot(formula, data)
```

Arguments

`formula` Formula of the form **ID + episode ~ xij + yij**.

- ID: A vector of subjects' unique identifier which can be numeric or character.
- episode: A vector indicating the episode of the bivariate alternating gap time pairs, e.g.: 1, 2, ..., m_i where m_i indicates the last episode for subject i.
- xij: A vector with the lengths of the type I gap times.
- yij: A vector with the lengths of the type II gap times.

`data` A data frame that contains all the vectors listed in the formula

Examples

```
library(BivRec)
set.seed(1234)
sim.data <- biv.rec.sim(nsize=150, beta1=c(0.5,0.5), beta2=c(0,-0.5), tau_c=63, set=1.1)
biv.rec.plot(formula = id + epi ~ xij + yij, data = sim.data)
```

---

biv.rec.sim *Bivariate Recurrent Response and Covariate Data Simulation*

Description

This function simulates a series of alternating recurrent events based on simulations in Lee CH, Huang C-Y, Xu G, Luo X (2017).

Usage

```
biv.rec.sim(nsize, beta1, beta2, tau_c, set)
```

Arguments

nsize	sample size which refers to the number of subjects in the data set where each subject could have multiple episodes of events.
beta1	true coefficients for first gap time in the accelerated failure time model (AFT).
beta2	true coefficients for second gap time in the accelerated failure time model (AFT).
tau_c	maximum support of censoring time. Can take values as follows: • tau_c=63: corresponds to cr=15% and corresponding $\bar{m}$ for each scenario in tables 1 and 2 of Lee CH, Huang C-Y, Xu G, Luo X (2017). • tau_c=30: corresponds to cr=30% and corresponding $\bar{m}$ for each scenario in tables 1 and 2 of Lee CH, Huang C-Y, Xu G, Luo X (2017).
set	Simulation setting based on scenarios outlined in tables 1 and 2 in Lee CH, Huang C-Y, Xu G, Luo X (2017). Choose 1.1 (default) for scenario 1 with $\rho = 1$ in the covariance matrix of the frailty vector, 1.2 for scenario 1 with $\rho = 0.5$, 1.3 for scenario 1 with $\rho = 0$ and 2.0 for scenario 2.

Value

Data frame with alternating recurrent event data and two covariates

References

1. Lee C, Huang CY, Xu G, Luo X (2017). Semiparametric regression analysis for alternating recurrent event data. Statistics in Medicine, 37: 996-1008. https://doi.org/10.1002/sim.7563

Examples

```
library(BivRec)
set.seed(1234)
biv.rec.sim(nsize=150, beta1=c(0.5,0.5), beta2=c(0,-0.5), tau_c=63, set=1.1)
```

Index