Package ‘intercure’

January 12, 2016

Type Package
Title Cure Rate Estimators for Interval Censored Data
Version 0.1.0
Date 2016-01-02
URL https://github.com/jbrettas/intercure
BugReports https://github.com/jbrettas/intercure/issues
Maintainer Julio Brettas <jbrettas@ime.usp.br>
Description Implementations of semiparametric cure rate estimators for interval censored data in R. The algorithms are based on the promotion time and frailty models, all for interval censoring. For the frailty model, there is also a implementation contemplating clustered data.
License GPL-2 | GPL-3
LazyData TRUE
Imports foreach, survival, MASS, stats4, Matrix, iterators, parallel
Suggests testthat, knitr, rmarkdown, doParallel
Depends R (>= 2.10)
RoxygenNote 5.0.1
VignetteBuilder knitr
NeedsCompilation no
Author Julio Brettas [aut, cre]
Repository CRAN
Date/Publication 2016-01-12 08:12:06

R topics documented:

intercure .......................................................... 2
inter_bch ............................................................ 2
inter_frailty .......................................................... 3
inter_frailty_cl ....................................................... 5
inter_bch

sim_bch ................................................................. 6
sim_frailty ............................................................... 7
sim_frailty_cl ......................................................... 8

Index

intercure  intercure: Cure rate regression models for interval censored data.

Description

The intercure package provides two main functions: inter_bch and inter_frailty. These are essential algorithms for estimating the cure fraction with promotion time and frailty model, respectively. The inter_frailty_cl function provides support for analysing clustered datasets using the frailty model. For generating datasets based on these two models, the package provides the sim_bch, sim_frailty and sim_frailty_cl functions, the last providing clustered datasets.

intercure functions

inter_bch, inter_frailty, inter_frailty_cl, sim_bch, sim_frailty, sim_frailty_cl

 inter_bch  Fits promotion time cure rate model for interval censored data

Description

inter_bch returns a list with the estimated parameters par and their asymptotic covariance matrix mcov. The list also contains a dummy variable stop_c assuming 0 if algorithm converged and 1 if a stop criteria ended the process.

Usage

inter_bch(dataset, left, right, cov, sigma = 10, crit_theta = 0.001, crit_p = 0.005, max_n = 100, output_files = FALSE)

Arguments

dataset  Dataset used to fit the model.
left  Vector containing the last check times before event.
right  Vector containing the first check times after event.
cov  String vector containing the column names to be used on the cure rate predictor.
sigma  Parameter for the primal-dual interior-point algorithm used on the maximization process. Default value set to 10.
crit_theta  The effects minimum error for convergence purposes.
crit_p  Minimum error of the non-parametric cumulative distribution function.
max_n  Maximum number of iterations of the ECM algorithm.
output_files  Boolean indicating if text outputs for the estimates and variances should be generated.
**inter_frailty**

**Value**

The `inter_bch` function returns an list containing the following outputs:

- **par**: estimates of theta parameters.
- **mcov**: estimates for the asymptotic covariance matrix of theta parameters.
- **stop_c**: stop criteria indicator assuming 1 when process is stopped for a non-convergence criteria. Assumes 0 when convergence is reached.

**Examples**

```r
set.seed(3)
sample_set <- sim_bch(80)

## few iterations just to check how to use the function
inter_bch(sample_set, sample_set$l, sample_set$r, c("xi1","xi2"), max_n = 5)

## precise estimate (computationally intensive)
## Not run:
inter_bch(sample_set, sample_set$l, sample_set$r, c("xi1","xi2"))

## End(Not run)
```

---

**inter_frailty**

*Fits cure rate frailty model for interval censored data*

**Description**

`inter_frailty` returns a list with the estimated parameters `par` and their asymptotic covariance matrix `mcov`. The list also contains a dummy variable `stop_c` assuming 0 if algorithm converged and 1 if a stop criteria ended the process.

**Usage**

```r
inter_frailty(dataset, left, right, delta, cov_theta, cov_beta, M, b = 0.001, tol = 0.001, max_n = 100, par_cl = NULL, burn_in = 30, output_files = FALSE)
```

**Arguments**

- **dataset**: Dataset used to fit the model.
- **left**: Vector containing the last check times before event.
- **right**: Vector containing the first check times after event.
- **delta**: Flag vector indicating failure inside interval.
cov_theta  String vector containing the column names to be used on the cure rate predictor.
cov_beta   String vector containing the column names to be used on the predictor associated with the hazard function.
M          Number of replicates generated by each iteration on the ANDA (Asymptotic Normal Data Augmentation) algorithm.
b          Parameter for initial theta and beta variances.
tol        Numeric for tolerance of convergence.
max_n      Maximum number of algorithm’s iterations without the burn in.
par_cl     Registered SOCK cluster for parallel process. If NULL (default) the program loops are executed sequentially.
burn_in    Number of burn in iterations.
output_files Boolean indicating if text outputs for the estimates and variances should be generated.

Value

The `inter_frailty` function returns an list containing the following outputs:

- **par**: estimates of theta and beta parameters.
- **mcov**: estimates for the covariance matrix of theta and beta parameters.
- **stop_c**: stop criteria indicator assuming 1 when process is stopped for a non-convergence criteria. Assumes 0 when convergence is reached.

Examples

```r
## few iterations just to check how to use the function
set.seed(3)
sample_set <- sim_frailty(80)
inter_frailty(sample_set, sample_set$L, sample_set$R, sample_set$delta,
c("xi1","xi2"), c("xi1","xi2"), M = 10, max_n = 3, burn_in = 0)

## precise estimate (computationally intensive)
## Not run:
inter_frailty(sample_set, sample_set$L, sample_set$R, sample_set$delta,
c("xi1"), c("xi2"), M = 50, max_n = 50, burn_in = 10)

## End(Not run)
```
inter_frailty_cl returns a list with the estimated parameters `par` and their covariance matrix `mcov`. The list also contains a dummy variable `stop_c` assuming 0 if algorithm converged and 1 if a stop criteria ended the process.

**Usage**

```r
inter_frailty_cl(dataset, left, right, delta, cov_theta, cov_beta, grp, M, b = 0.001, tol = 0.001, max_n = 100, par_cl = NULL, burn_in = 50, output_files = FALSE)
```

**Arguments**

- **dataset**: Dataset used to fit the model.
- **left**: Vector containing the last check times before event.
- **right**: Vector containing the first check times after event.
- **delta**: Flag vector indicating failure inside interval.
- **cov_theta**: String vector containing the column names to be used on the cure rate predictor.
- **cov_beta**: String vector containing the column names to be used on the predictor associated with the hazard function.
- **grp**: Vector containing cluster identifier (numeric or string).
- **M**: Number of replicates generated by each iteration on the ANDA (Asymptotic Normal Data Augmentation) algorithm.
- **b**: Parameter for initial theta and beta variances.
- **tol**: Numeric for tolerance of convergence.
- **max_n**: Maximum number of algorithm’s iterations without the burn in.
- **par_cl**: Registered SOCK cluster for parallel process. If NULL (default) the program loops are executed sequentially.
- **burn_in**: Number of burn in iterations.
- **output_files**: Boolean indicating if text outputs for the estimates and variances should be generated.

**Value**

The `inter_frailty_cl` function returns an list containing the following outputs:

- **par**: estimates of theta and beta parameters.
- **mcov**: estimates for the covariance matrix of theta and beta parameters.
- **stop_c**: stop criteria indicator assuming 1 when process is stopped for a non-convergence criteria. Assumes 0 when convergence is reached.
Examples

```r
set.seed(3)
sample_set <- sim_frailty_cl(80, nclus = 3)

## few iterations just to check how to use the function
inter_frailty_cl(sample_set, sample_set$l, sample_set$r, sample_set$delta,
c("xi1"), c("xi2"), grp = sample_set$clus, M = 20,
max_n = 2, burn_in = 0)

## precise estimate (computationally intensive)
## Not run:
inter_frailty_cl(sample_set, sample_set$l, sample_set$r, sample_set$delta,
c("xi1", "xi2"), c("xi1", "xi2"), grp = sample_set$clus, M = 20,
max_n = 30, burn_in = 10)

## End(Not run)
```

### sim_bch

Generates a interval censored dataset using promotion time cure rate model

**Description**

`sim_bch_data` returns a dataset generated by the cure rate frailty model.

**Usage**

```r
sim_bch(N, theta = c(1, 0.5, 0), lambda = 1, A = 5, B = 15,
prob = 0.5)
```

**Arguments**

- `N` : Size of the sample to be generated.
- `theta` : Three parameters associated with the cure linear predictor.
- `lambda` : Rate parameter for the exponential distributed latent variables.
- `A` : A positive number representing a fixed right censoring.
- `B` : A positive number which multiplies an uniform random variable, defining another right censoring case.
- `prob` : Probability that individual presents treatment T1 (baseline is T0).

**Value**

A generated dataset with columns: Z, the actual event time; L, the leftmost limit of the censored interval; R, the rightmost limit of the censored interval; delta, the failure indicator; xi1, the treatment covariate assuming 1 with probability prob and 0 otherwise; xi2, second variable generated by a standard normal distribution.
**Description**

`sim_frailty` returns a dataset generated by the cure rate frailty model.

**Usage**

```r
sim_frailty(nL theta = c(M1L 1L 0)L beta = c(0L 0NU)L a = UL b = 1UL
prob = 0NU)
```

**Arguments**

- `n` Size of the sample to be generated.
- `theta` Three parameters associated with the cure linear predictor.
- `beta` Two parameters associated with the hazard function.
- `a` A positive number representing a fixed right censoring.
- `b` A positive number which multiplies an uniform random variable, defining another right censoring case.
- `prob` Probability that individual presents treatment T1 (baseline is T0).

**Value**

A generated dataset with columns: `Z`, the actual event time; `L`, the leftmost limit of the censored interval; `R`, the rightmost limit of the censored interval; `delta`, the failure indicator; `xi1`, the treatment covariate assuming 1 with probability `prob` and 0 otherwise; `xi2`, second variable generated by a standard normal distribution.

**Examples**

```r
sim_frailty(20)
```
sim_frailty_cl

Generates a interval censored clustered dataset using frailty cure rate model

Description

sim_frailty returns a dataset generated by the cure rate frailty model.

Usage

sim_frailty_cl(N, theta = c(-1, 1, 0), beta = c(0, 0.5), A = 5, B = 15, prob = 0.5, nclus = 2, w = exp(-0.5))

Arguments

N Size of the sample to be generated.
theta Three parameters associated with the cure linear predictor.
beta Two parameters associated with the hazard function.
A A positive number representing a fixed right censoring.
B A positive number which multiplies an uniform random variable, defining another right censoring case.
prob Probability that individual presents treatment T1 (baseline is T0).
nclus Number of clusters to generate with balanced sizes.
w Shape and rate parameters value for the Gamma distribution with mean 1.

Value

A generated dataset with columns: Z, the actual event time; L, the leftmost limit of the censored interval; R, the rightmost limit of the censored interval; delta, the failure indicator; xi1, the treatment covariate assuming 1 with probability prob and 0 otherwise; xi2, second variable generated by a standard normal distribution; clus, representing the cluster id for each observation.

Examples

sim_frailty_cl(50)
Index

inter_bch, 2
inter_frailty, 3
inter_frailty_cl, 5
intercure, 2
intercure-package (intercure), 2

sim_bch, 6
sim_frailty, 7
sim_frailty_cl, 8