Package ‘gte’

October 13, 2022

Type Package
Title Generalized Turnbull's Estimator
Version 1.2-3
Date 2022-04-13
Author Mohammad Hossein Dehghan, Thierry Duchesne and Sophie Baillargeon
Maintainer Thierry Duchesne <thierry.duchesne@mat.ulaval.ca>
Description Generalized Turnbull's estimator proposed by Dehghan and Duchesne (2011).
License GPL-2
LazyData true
Depends survival
Suggests interval
RoxygenNote 7.1.2
NeedsCompilation yes
Repository CRAN
Date/Publication 2022-04-16 14:52:29 UTC

R topics documented:

  gte-package ......................................................... 2
  gte ................................................................. 2
  simul .............................................................. 5

Index
Description

The `gte` function computes the generalized Turnbull’s estimator proposed by Dehghan and Duchesne (2011). It is a nonparametric estimator of a conditional survival function given a vector of continuous covariates that can handle interval-censored lifetimes.

Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>gte</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>1.2-3</td>
</tr>
<tr>
<td>Date:</td>
<td>2022-04-13</td>
</tr>
<tr>
<td>License:</td>
<td>GPL-2</td>
</tr>
</tbody>
</table>

Author(s)

Mohammad Hossein Dehghan, Thierry Duchesne and Sophie Baillargeon

Maintainer: Thierry Duchesne <thierry.duchesne@mat.ulaval.ca>

References

Usage

gte(formula, data, z, h = NULL, itermax = 1e+05, tole = 5e-04)

## S3 method for class 'gte'
print(x, ...)

## S3 method for class 'gte'
plot(
x,
overlay = TRUE,
shade = TRUE,
xlab = "time",
ylab = "survival",
xleg = "bottomleft",
yleg = NULL,
...
)

Arguments

formula A formula object with the response on the left of a ~ operator, and the covariates on the right. The response must be a survival object as returned by the Surv function from the package survival (see Details).
data An optional data frame, list or environment containing the variables in the model formula. If not found in data, the variables are taken from environment(formula), typically the environment from which gte is called.
z A matrix: each row contains the values of a covariate vector at which an estimate of the conditional survival function is requested. If there is only one covariate, it can be a vector (possibly of length 1).
h A vector: the values of the bandwidth parameter $h$ for each covariate (default = equation 7 of Dehghan and Duchesne (2011)).
itermax maximal number of iterations for the algorithm (default=100000).
tole maximal distance between successive iterations tolerated before declaring convergence (default=0.0005).
x An object, produced by the gte function, to print or to plot.
... Further arguments to be passed to print.default or plot.default.
overlay A logical: Should the curves be overlayed when there is more than one estimate of the conditional survival function in the gte object $x$? (default=TRUE)
shade A logical: Should the rectangles of indeterminate NPMLE (innermost interval) be shaded? (default=TRUE)
xlab A label for the x-axis, by default xlab = "time".
ylab A label for the y-axis, by default ylab = "survival".
xleg x location for legend, "bottomleft" by default (see legend).
yleg y location for legend, NULL by default (see legend).
Details

For interval-censored data, the `Surv` function should be called with the argument `type="interval"` or `type="interval2"`. If `type="interval"`, the event argument is mandatory. Therefore, in addition to the left and right endpoints of the censoring interval (called, respectively, left and right for illustrative purpose), one would need a third variable (status) taking the value 0 for right censored data, 1 for an event at exact time, 2 for left censored data and 3 for interval censored data. The `Surv` function would be called as follows:

```
Surv(time=left, time2=right, event=status, type="interval")
```

If `type="interval2"`, the event argument cannot be given. The value of event is derived from the time and time2 argument as follows:
- if time takes the value NA, event=2 (left censored data);
- if time2 takes the value NA, event=0 (right censored data);
- if time=time2, event=1 (exact time);
- otherwise, event=3 (interval censored data).

See the help page of the `Surv` function for more details.

In the `gte` function, the data must be given through the `Surv` function but it is internally transformed in two vectors: L and R for the left and right endpoints of the censoring interval, respectively. If event=0 (right censored data), then L=time and R=Inf;
- if event=1 (exact time), then L=time and R=time;
- if event=2 (left censored data), then L=0 and R=time;
- and if event=3 (interval censored data), then L=time and R=time2;

If one has vectors L and R respecting this convention, they can be given directly to `gte` by calling `Surv` as follows:

```
Surv(L, R, type="interval2")
```

Value

- **time**: A vector: the ordered distinct values of the left and right endpoints of the censoring interval (omitting the smallest value, but always including time 0).
- **surv**: A matrix: the estimates of the conditional survival function at time `time`. The `i`th column refers to the `i`th value of the covariate vector given in `z` (row `i` of `z`).
- **intmap**: A matrix: The intervals of the potential steps in the conditional survival function, called innermost interval, over which the GTE is indeterminate. The left endpoints of the intervals are in the first row, and the right endpoints in the second. The object attribute LRin denotes whether to include each of the endpoints or not. This matrix is computed with an internal function derived from function `Aintmap` of the `interval` package.
- **surv.summary**: A summary of surv: the estimates of the conditional survival function only for the intervals of the potential steps in the function (innermost intervals). The row names describe the intervals, which are detailed in `intmap`.

Call

The function call.

Author(s)

Mohammad Hossein Dehghan, Thierry Duchesne and Sophie Baillargeon
References


See Also

\texttt{Surv}

Examples

```r
## Calling \texttt{Surv()} with type="interval2"
Fit <- gte(Surv(L, R, type="interval2") ~ Z, data=simul, z=c(10, 20))
Fit

## Calling \texttt{Surv()} with type="interval"
event <- ifelse(is.na(simul$R), 0,
    ifelse(is.na(simul$L), 2,
        ifelse(simul$R==simul$L, 1, 3)))
time <- ifelse(event==2, simul$R, simul$L)
time2 <- ifelse(event==3, simul$R, NA)
simul_event <- cbind(simul, time, time2, event)
Fit_event <- gte(Surv(time, time2, event, type="interval") ~ Z, data=simul_event, z=c(10, 20))
Fit_event

# The results are the same
all.equal(Fit_event$time, Fit$time)
all.equal(Fit_event$surv, Fit$surv)

## Plotting the results
plot(Fit, xleg="topright")
```

---

**simul**

### Simulated Data

**Description**

Simulated Interval-censored data

**Format**

A data frame with 100 observations on the following 3 variables.

- **L** the left endpoints of the censoring interval
- **R** the right endpoints of the censoring interval
- **Z** a continuous covariate
Details

The value $R = \text{NA}$ means that the observation is right censored (occurs 2 times). If $L = \text{NA}$, then the observation is left censored (occurs 26 times). An observation with $R = L$ means that the time of occurrence of the event is known exactly (occurs 3 times).

References


Examples

```r
Fit <- gte(Surv(L, R, type="interval2") ~ Z, data=simul, z=15)
plot(Fit)
```
Index

* datasets
  simul, 5
* package
  gte-package, 2

gte, 2
gte-package, 2

legend, 3

plot.gte(gte), 2
print.gte(gte), 2

simul, 5
Surv, 4, 5