Package 'intccr'

May 9, 2019

Type Package

Title Semiparametric Competing Risks Regression under Interval Censoring

Version 1.1.4

Author Giorgos Bakoyannis <gbakogia@iu.edu>, Jun Park <jp84@iu.edu>

Maintainer Jun Park <jp84@iu.edu>

Description Semiparametric regression models on the cumulative incidence function with interval-censored competing risks data as described in Bakoyannis, Yu, & Yiannoutsos (2017) <doi:10.1002/sim.7350>. The main function fits the proportional subdistribution hazards model (Fine-Gray model), the proportional odds model, and other models that belong to the class of semiparametric generalized odds rate transformation models.

Date 2019-05-09

Imports alabama (>= 2015.3.1), doParallel, foreach, parallel, numDeriv, stats, utils

Depends R (>= 2.14.0)

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2019-05-09 20:10:11 UTC

R topics documented:

bssmle 2
bssmle_se 3
ciregic 4
dataprep 7
longdata 8
naive_b 8

predict.ciregic 9
pseudo.HIV.long 10
simdata 11
summary.ciregic 12
Surv2 13
vcov.ciregic 14
vcov.summary.ciregic 15
waldtest 16

Index **18**

`bssmle` *B-spline Sieve Maximum Likelihood Estimation*

Description

Routine that performs B-spline sieve maximum likelihood estimation with linear and nonlinear inequality constraints

Usage

```
bssmle(formula, data, alpha, k = 1)
```

Arguments

`formula`	a formula object relating survival object `Surv2(v, u, event)` to a set of covariates
`data`	a data frame that includes the variables named in the formula argument
`alpha`	$\alpha = (\alpha 1, \alpha 2)$ contains parameters that define the link functions from class of generalized odds-rate transformation models. The components $\alpha 1$ and $\alpha 2$ should both be ≥ 0. If $\alpha 1 = 0$, the user assumes the proportional subdistribution hazards model or the Fine-Gray model for the event type 1. If $\alpha 2 = 1$, the user assumes the proportional odds model for the event type 2.
`k`	a parameter that controls the number of knots in the B-spline with $0.5 \leq$k≤ 1

Details

The function `bssmle` performs B-spline sieve maximum likelihood estimation.

Value

The function `bssmle` returns a list of components:

`beta`	a vector of the estimated coefficients for the B-splines
`varnames`	a vector containing variable names
`alpha`	a vector of the link function parameters
`loglikelihood`	a loglikelihood of the fitted model

convergence	an indicator of convegence
tms	a vector of the minimum and maximum observation times
Bv	a list containing the B-splines basis functions evaluated at v

Author(s)

Giorgos Bakoyannis, <gbakogia at iu dot edu>

Jun Park, <jp84 at iu dot edu>

Examples

```
est.simdata <- intccr:::bssmle(Surv2(v, u, c) ~ z1 + z2, data = simdata, alpha = c(1, 1))
newdata <- intccr::dataprep(data = longdata, ID = "id", time = "t",
                            event = "c", Z = c("z1", "z2"))
est.longdata <- intccr:::bssmle(Surv2(v, u, c) ~ z1 + z2, data = newdata, alpha = c(1, 1))
```

---

bssmle_se *Bootstrap varince-covariance estimation*

Description

Bootstrap varince estimation for the estimated regression coefficients

Usage

```
bssmle_se(formula, data, alpha, k = 1, do.par, nboot)
```

Arguments

formula	a formula object relating survival object Surv2(v, u, event) to a set of covariates
data	a data frame that includes the variables named in the formula argument
alpha	$\alpha = (\alpha 1, \alpha 2)$ contains parameters that define the link functions from class of generalized odds-rate transformation models. The components $\alpha 1$ and $\alpha 2$ should both be ≥ 0. If $\alpha 1 = 0$, the user assumes the proportional subdistribution hazards model or the Fine-Gray model for the event type 1. If $\alpha 2 = 1$, the user assumes the proportional odds model for the event type 2.
k	a parameter that controls the number of knots in the B-spline with $0.5 \leq$k≤ 1
do.par	using parallel computing for bootstrap calculation. If do.par = TRUE, parallel computing will be used during the bootstrap estimation of the variance-covariance matrix for the regression parameter estimates.
nboot	a number of bootstrap samples for estimating variances and covariances of the estimated regression coefficients. If nboot = 0, the function ciregic does dot perform bootstrap estimation of the variance matrix of the regression parameter estimates and returns NA in the place of the estimated variance matrix of the regression parameter estimates.

Details

The function bssmle_se estimates bootstrap standard errors for the estimated regression coefficients from the function bssmle.

Value

The function bssmle_se returns a list of components:

notconverged	a list of number of bootstrap samples not converged
numboot	a number of bootstrap converged
Sigma	an estimated bootstrap variance-covariance matrix of the estimated regression coefficients

Author(s)

Giorgos Bakoyannis, <gbakogia at iu dot edu>

Jun Park, <jp84 at iu dot edu>

Examples

```
attach(simdata)
est.vcov <- intccr:::bssmle_se(formula = Surv2(v, u, c) ~ z1 + z2, data = simdata,
                               alpha = c(1, 1), do.par = FALSE, nboot = 1)
```

---

ciregic *Competing Risks Regression with Interval-Censored Data*

Description

The function ciregic performs semiparametric regression on cumulative incidence function with interval-censored competing risks data. It fits the proportional subdistribution hazards model (Fine-Gray model), the proportional odds model, and other models that belong to the class of semiparametric generalized odds rate transformation models.

Usage

```
ciregic(formula, data, alpha, k = 1, do.par, nboot)
```

Arguments

formula	a formula object relating the survival object Surv2(v, u, event) to a set of covariates
data	a data frame that includes the variables named in the formula argument

alpha	$\alpha = (\alpha 1, \alpha 2)$ contains parameters that define the link functions from class of generalized odds-rate transformation models. The components $\alpha 1$ and $\alpha 2$ should both be ≥ 0. If $\alpha 1 = 0$, the user assumes the proportional subdistribution hazards model or the Fine-Gray model for the event type 1. If $\alpha 2 = 1$, the user assumes the proportional odds model for the event type 2.
k	a parameter that controls the number of knots in the B-spline with $0.5 \leq$k≤ 1
do.par	an option to use parallel computing for bootstrap. If `do.par = TRUE`, parallel computing will be used during the bootstrap estimation of the variance-covariance matrix for the regression parameter estimates.
nboot	a number of bootstrap samples for estimating variances and covariances of the estimated regression coefficients. If `nboot = 0`, the function `ciregic` does not perform bootstrap estimation of the variance-covariance matrix of the regression parameter estimates and returns `NA` in the place of the estimated variance-covariance matrix of the regression parameter estimates.

Details

The formula for the model has the form of `response ~ predictors`. The response in the formula is a `Surv2(v, u, event)` object where `v` is the last observation time prior to the event, `u` is the first observation time after the event, and `event` is the event or censoring indicator. `event` should include 0, 1 or 2, denoting right-censoring, event type 1 and 2, respectively. If `event=0` (i.e. right-censored observation) then `u` is not included in any calculation as it corresponds to ∞. The user can provide any value in `u` for the right-censored cases, even `NA`. The function `ciregic` fits models that belong to the class of generalized odds rate transformation models which includes the proportional subdistribution hazards or the Fine-Gray model and the proportional odds model. The parameter $\alpha = (\alpha 1, \alpha 2)$ defines the link function/model to be fitted for event 1 and 2, respectively. A value of `0` corresponds to the Fine-Gray model and a value of 1 corresponds to the proportional odds model. For example, if $\alpha = (0, 1)$ then the function `ciregic` fits the Fine-Gray model for the event type 1 and the proportional odds model for the event type 2.

Value

The function `ciregic` provides an object of class `ciregic` with components:

varnames	a vector containing variable names
coefficients	a vector of the regression coefficient estimates
gamma	a vector of the estimated coefficients for the B-splines
vcov	a variance-covariance matrix of the estimated regression coefficients
alpha	a vector of the link function parameters
loglikelihood	a loglikelihood of the fitted model
convergence	an indicator of convegence
tms	a vector of the minimum and maximum observation times
Bv	a list containing the B-splines basis functions evaluated at `v`
numboot	a number of converged bootstrap
notconverged	a list of number of bootstrap samples not converged
call	a matched call

Author(s)

Giorgos Bakoyannis, <gbakogia at iu dot edu>

Jun Park, <jp84 at iu dot edu>

References

Bakoyannis, G., Yu, M., and Yiannoutsos C. T. (2017). Semiparametric regression on cumulative incidence function with interval-censored competing risks data. *Statistics in Medicine*, **36**:3683-3707.

Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, **94**:496-509.

See Also

summary.ciregic for the summarized results and predict.ciregic for value of the predicted cumulative incidence functions. coef and vcov are the generic functions. dataprep for reshaping data from a long format to a suitable format to be used in the function ciregic.

Examples

```
## Set seed in order to have reproducibility of the bootstrap standard error estimate
set.seed(1234)

## Reshaping data from a long format to a suitable format
newdata <- dataprep(data = longdata, ID = "id", time = "t",
                    event = "c", Z = c("z1", "z2"))
## Estimation of regression parameters only. No bootstrap variance estimation.
## with 'newdata'
fit <- ciregic(formula = Surv2(v, u, c) ~ z1 + z2, data = newdata,
               alpha = c(1, 1), nboot = 0, do.par = FALSE)
fit

## Estimation of regression parameters only. No bootstrap variance estimation.
## with 'simdata'
fit.simdata <- ciregic(formula = Surv2(v, u, c) ~ z1 + z2, data = simdata,
                       alpha = c(1, 1), nboot = 0, do.par = FALSE)
fit.simdata

## Not run:
## Bootstrap variance estimation based on 50 replications
fit <- ciregic(formula = Surv2(v, u, c) ~ z1 + z2, data = newdata,
               alpha = c(1, 1), nboot = 50, do.par = FALSE)

## End(Not run)
## Note that the user can use parallel computing to decrease
## the computation time of the bootstrap variance-covariance
## estimation (e.g. nboot = 50)

## Summarize semiparametric regression model
summary(fit)
```

```
## Predict and draw plot the cumulative incidence function evaluated at z1 = 1 and z2 = 0.5
t <- seq(from = 0, to = 2.8, by = 2.8 / 99)
pred <- predict(object = fit, covp = c(1, 0.5), times = t)
pred
plot(pred$t, pred$cif1, type = "l", ylim = c(0, 1))
points(pred$t, pred$cif2, type = "l", col = 2)
```

---

dataprep *Data manipulation*

Description

The function dataprep reshapes data from a long format to a ready-to-use format to be used directly in the function ciregic.

Usage

```
dataprep(data, ID, time, event, Z)
```

Arguments

data	a data frame that includes the variables named in the ID, time, event, and z arguments
ID	a variable indicating individuals' ID
time	a variable indicating observed time points
event	a vector of event indicator. If an observation is righ-censored, event = 0; otherwise, event = 1 or event = 2, where 1 represents the event type 1, and 2 represents the event type 2. The current version of package only allows two event types.
Z	a vector of variables indicating name of covariates

Details

The function dataprep provides a ready-to-use data format that can be directly used in the function ciregic. The returned data frame consists of id, v, u, c, and covariates as columns. The v and u indicate time window with the last observation time before the event and the first observation after the event. The c represents a type of event, for example, c = 1 for the event type 1, c = 2 for the event type 2, and c = 0 for the right-censored. Individuals who have only one time record with right-censored event will be omitted because its time interval is (0, Inf), and the lower bound v will be replaced by zero, for example (0, v], if individuals are not right-censored and have only one time record.

Value

a data frame

Author(s)

Giorgos Bakoyannis, <gbakogia at iu dot edu>

Jun Park, <jp84 at iu dot edu>

Examples

```
library(intccr)
dataprep(data = longdata, ID = "id", time = "t", event = "c", Z = c("z1", "z2"))
```

---

longdata *Simulated long format data*

Description

Data containing subject id, a series of time points, cause of failure, and covariates with 200 observations.

Usage

```
longdata
```

Format

A data frame with 868 rows and 5 variables.

Examples

```
head(longdata, n = 5)
```

---

naive_b *Initial values for the sieve maximum likelihood estimation*

Description

The function naive_b provides a vector of initial values for the B-spline sieve maximum likelihood estimation.

Usage

```
naive_b(data, v, u, c, q, k = 1)
```

Arguments

data	a data frame that includes the variables named in each argument
v	the last observation time prior to the failure.
u	the first observation time after the failure.
c	an indicator of event type. If an observation is righ-censored, `event = 0`; otherwise, `event = 1` or `event = 2`, where 1 represents the event type 1, and 2 represents the event type 2. The current version of package only allows for two event types.
q	a dimension of design matrix.
k	a parameter that controls the number of knots in the B-spline with $0.5 \leq$k≤ 1

Details

The function `naive_b` provides initial values for the optimization procedure.

Value

Initial values of B-spline estimation

b	a vector of the initial values to be used in the optimization process

Author(s)

Giorgos Bakoyannis, <gbakogia at iu dot edu>

Jun Park, <jp84 at iu dot edu>

Examples

```
attach(simdata)
intccr:::naive_b(data = simdata, v = v, u = u, c = c, q = 2)
```

---

predict.ciregic *Covariate-Specific Cumulative Incidence Prediction*

Description

`predict` method for class `ciregic`. It provides the predicted cumulative incidence function for a given covariate pattern and timepoint(s).

Usage

```
## S3 method for class 'ciregic'
predict(object, covp, times, ...)
```

Arguments

object	an object of class ciregic, which is a result of a call to ciregic
covp	a desired values for covariates
times	time points that user wants to predict value of cumulative incidence function
...	further arguments

Details

predict.ciregic returns the predicted cumulative incidence function for a given covariate pattern and timepoint(s).

Value

The function predict.ciregic returns a list of predicted values of the model from object.

t	time points
cif1	the predicted value of cumulative incidence function for the event type 1
cif2	the predicted value of cumulative incidence function for the event type 2

See Also

The fitted semiparametric regression on cumulative incidence function with interval-censored competing risks data ciregic and summary of the fitted semiparametric regression model summary.ciregic

Examples

```
## Continuing the ciregic(...) example

pfit <- predict(object = fit, covp = c(1, 0.5), times = c(0.1, 0.15, 0.5, 0.7))
pfit
mint <- fit$tms[1]
maxt <- fit$tms[2]
pfit1 <- predict(object = fit, covp = c(1, 0.5),
                 times = seq(mint, maxt, by = (maxt-mint)/99))
plot(pfit1$t, pfit1$cif1, ylim = c(0, 1), type = "l")
lines(pfit1$t, pfit1$cif2, ylim = c(0, 1), lty = 2, col = 2)
```

---

pseudo.HIV.long *Artificial HIV dataset*

Description

Artificial dataset that was simulated to resemble the HIV study on loss to HIV care and death in sub-Saharan Africa, that was presented in Bakoyannis, Yu, & Yiannoutsos (2017). It contains subject id, observation times, cause of failure, and covariates.

Usage

```
pseudo.HIV.long
```

Format

A data frame with 22710 rows and 6 variables.

References

Bakoyannis, G., Yu, M., and Yiannoutsos C. T. (2017). Semiparametric regression on cumulative incidence function with interval-censored competing risks data. *Statistics in Medicine*, **36**:3683-3707.

Examples

```
head(pseudo.HIV.long, n = 20)
```

---

simdata *Simulated data*

Description

Data containing the last time point prior to the event, the first time point after the event, cause of failure, and covariates with 200 observations.

Usage

```
simdata
```

Format

A data frame with 200 rows and 5 variables.

Examples

```
head(simdata, n = 5)
```

summary.ciregic *Summary of* ciregic

Description

summary method for class ciregic

Usage

```
## S3 method for class 'ciregic'
summary(object, ...)
```

Arguments

object	an object of class ciregic, which is a result of a call to ciregic
...	further arguments

Details

The function summary.ciregic returns the coefficients, bootstrap standard errors, and etc. Additionally, 'significance star' is included.

Value

The function summary.ciregic returns a list of summary statistics of the model from object.

varnames	a vector containing variable names
coefficients	a vector of the regression coefficient estimates
se	a bootstrap standard error of the coefficients
z	z value of the estimated coefficients
p	p value of the estimated coefficients
call	a matched call

See Also

The fitted semiparametric regression on cumulative incidence function with interval-censored competing risks data ciregic and values of the predicted cumulative incidence functions predict.ciregic

Examples

```
## Continuing the ciregic(...) example

sfit <- summary(fit)
sfit
```

`Surv2` *Creating data frame*

Description

The function `Surv2` generates the survival object to be treated as the response from `ciregic`.

Usage

```
Surv2(v, u, event)
```

Arguments

`v`	the last observation time prior to the failure; $0 \leq v \leq u$.
`u`	the first observation time after the failure; $u \geq 0$.
`event`	an indicator of event type. If an observation is righ-censored, `event = 0`; otherwise, `event = 1` or `event = 2`, where `1` represents the event type 1, and `2` represents the event type 2. The current version of package only allows for two event types.

Details

The function `Surv2` provides a response data frame which is used in the function `ciregic`.

Value

data frame

Author(s)

Giorgos Bakoyannis, <gbakogia at iu dot edu>

Jun Park, <jp84 at iu dot edu>

Examples

```
attach(simdata)
Surv2(v, u, c)
```

vcov.ciregic *Variance-covariance matrix of* ciregic

Description

vcov method for class ciregic

Usage

```
## S3 method for class 'ciregic'
vcov(object, ...)
```

Arguments

object	an object of class ciregic, which is a result of a call to ciregic
...	further arguments

Details

The function vcov returns the variance-covariance matrix of the fitted semiparametric regression model.

Value

The estimated bootstrap variance-covariance matrix

See Also

The fitted semiparametric regression on cumulative incidence function with interval-censored competing risks data ciregic, summary of the fitted semiparametric regression model summary.ciregic, and values of predicted cumulative incidence functions predict.ciregic

Examples

```
## Continuing the ciregic(...) example

vcov(fit)
```

`vcov.summary.ciregic` *Variance-covariance matrix of* `summary.ciregic`

Description

`vcov` method for class `summary.ciregic`

Usage

```
## S3 method for class 'summary.ciregic'
vcov(object, ...)
```

Arguments

`object`	an object of class `summary.ciregic`, which is a result of a call to `ciregic`
`...`	further arguments

Details

The `vcov` returns the variance-covariance matrix of the fitted semiparametric regression model.

Value

The estimated bootstrap variance-covariance matrix

See Also

The fitted semiparametric regression on cumulative incidence function with interval-censored competing risks data `ciregic`, summary of the fitted semiparametric regression model `summary.ciregic`, and values of the predicted cumulative incidence functions `predict.ciregic`

Examples

```
## Continuing the ciregic(...) example

vcov(summary(fit))
```

`waldtest` *Wald test for* `ciregic`

Description

`waldtest` for class `ciregic`. It provides the result of Wald test for the fitted model from the function `ciregic`.

Usage

```
waldtest(obj1, obj2 = NULL, ...)
```

Arguments

`obj1`	an object of the fitted model in `ciregic`
`obj2`	an object of the fitted model in `ciregic`, the default is `NULL`
`...`	further arguments

Details

The function `waldtest.ciregic` returns a result of Wald test.

Value

The function `waldtest` returns an output table of Wald test of the model from `object`.

`varnames.full`	a variable name of a vector of variables names in the full model
`varnames.nested`	a variable name of a vector of variables names in the nested model
`vcov`	the estimated bootstrap variance-covariance matrix for overall Wald test
`vcov.event1`	the estimated bootstrap variance-covariance matrix for cause-specific Wald test (event type 1)
`vcov.event2`	the estimated bootstrap variance-covariance matrix for cause-specific Wald test (event type 2)
`table`	a table including test statistic, degrees of freedom, and p-value

See Also

The fitted semiparametric regression on cumulative incidence function with interval-censored competing risks data `ciregic`

Examples

```
## Continuing the ciregic(...) example
## Not run:
## Bootstrap variance estimation based on 5 replications
set.seed(12345)
fit <- ciregic(formula = Surv2(v, u, c) ~ z1 + z2, data = newdata,
               alpha = c(1, 1), nboot = 5, do.par = FALSE)
waldtest(obj1 = fit)
set.seed(12345)
fit.nested <- ciregic(formula = Surv2(v, u, c) ~ z2, data = newdata,
               alpha = c(1, 1), nboot = 5, do.par = FALSE)
waldtest(obj1 = fit, obj2 = fit.nested)

## End(Not run)
```

Index

∗Topic **Surv2**
    Surv2, 13
∗Topic **bssmle_se**
    bssmle_se, 3
∗Topic **bssmle**
    bssmle, 2
∗Topic **ciregic**
    ciregic, 4
∗Topic **dataprep**
    dataprep, 7
∗Topic **datasets**
    longdata, 8
    pseudo.HIV.long, 10
    simdata, 11
∗Topic **naive_b**
    naive_b, 8

bssmle, 2
bssmle_se, 3

ciregic, 4, *10*, *12*, *14–16*

dataprep, *6*, 7

longdata, 8

naive_b, 8

predict.ciregic, *6*, 9, *12*, *14, 15*
pseudo.HIV.long, 10

simdata, 11
summary.ciregic, *6*, *10*, *12*, 12, *14, 15*
Surv2, 13

vcov.ciregic, 14
vcov.summary.ciregic, 15

waldtest, 16