Package ‘crrSC’

February 19, 2015

Title  Competing risks regression for Stratified and Clustered data
Version 1.1
Author Bingqing Zhou and Aurelien Latouche
Description Extension of cmprsk to Stratified and Clustered data. Goodness of fit test for Fine-Gray model.
Depends survival
Maintainer Aurelien Latouche <aurelien.latouche@cnam.fr>
License GPL-2
NeedsCompilation yes
Repository CRAN
Date/Publication 2013-06-23 00:24:34

R topics documented:

bce ................................................................. 1
cdata ............................................................... 2
center ............................................................. 3
crc ................................................................. 3
crcs ............................................................... 5
crrvsvs ............................................................ 8
print.crrs ....................................................... 8
psh.test ......................................................... 9

Index

bce

Breast Cancer Data

Description

Data Randomly Generated According To EI178 clinical trial
Usage
data(bce)

Format
A data frame with 200 observations and the following 6 variables.

trt  Treatment: 0=Placebo, 1=Tamoxifen
time  Event time
type  Event type. 0=censored, 1=Breast Cancer recurrence, 2=Death without recurrence
nnodes  Number of positive nodes
tsize  Tumor size
age  Age

Examples
data(bce)

------------------------------------------------------------------

| cdata | Clustered competing risks simulated data |

Description
sample of 200 observations

Usage
data(cdata)

Format

ID  Id of cluster, each cluster is of size 2
ftime  Event time
fstatus  Event type. 0=censored, 1, 2
z  a binary covariate with P(z=1)=0.5

Examples
data(cdata)
Multicenter Bone Marrow transplantation data

Description
Random sub sample of 400 patients

Usage
data(center)

Format
A data frame with 400 observations and the following 5 variables.

id  Id of transplantation center
ftime  Event time
fstatus  Event type. 0=censored, 1=Acute or Chronic GvHD, 2=Death free of GvHD
cells  source of stem cells: peripheral blood vs bone marrow
fm  female donor to male recipient match

Examples
data(center)

Competing Risks Regression for Clustered Data

Description
Regression modeling of subdistribution hazards for clustered right censored data. Failure times within the same cluster are dependent.

Usage
crrc(ftime,fstatus,cov1,cov2,tf,cluster,
cengroup,failcode=1,
cencode=0, subset,
na.action=na.omit,
gtol=1e-6,maxiter=10,init)
Arguments

- **cluster**: Clustering covariate
- **ftime**: vector of failure/censoring times
- **fstatus**: vector with a unique code for each failure type and a separate code for censored observations
- **cov1**: matrix (nobs x ncovs) of fixed covariates (either cov1, cov2, or both are required)
- **cov2**: matrix of covariates that will be multiplied by functions of time; if used, often these covariates would also appear in cov1 to give a proportional hazards effect plus a time interaction
- **tf**: functions of time. A function that takes a vector of times as an argument and returns a matrix whose jth column is the value of the time function corresponding to the jth column of cov2 evaluated at the input time vector. At time tk, the model includes the term cov2[,j]*tf(tk)[,j] as a covariate.
- **cengroup**: vector with different values for each group with a distinct censoring distribution (the censoring distribution is estimated separately within these groups). All data in one group, if missing.
- **failcode**: code of fstatus that denotes the failure type of interest
- **cencode**: code of fstatus that denotes censored observations
- **subset**: a logical vector specifying a subset of cases to include in the analysis
- **na.action**: a function specifying the action to take for any cases missing any of ftime, fstatus, cov1, cov2, cengroup, or subset.
- **gtol**: iteration stops when a function of the gradient is < gtol
- **maxiter**: maximum number of iterations in Newton algorithm (0 computes scores and var at init, but performs no iterations)
- **init**: initial values of regression parameters (default=all 0)

Details

This method extends Fine-Gray proportional hazards model for subdistribution (1999) to accommodate situations where the failure times within a cluster might be correlated since the study subjects from the same cluster share common factors. This model directly assesses the effect of covariates on the subdistribution of a particular type of failure in a competing risks setting.

Value

Returns a list of class crr, with components

- **$coef**: the estimated regression coefficients
- **$loglik**: log pseudo-likelihood evaluated at coef
- **$score**: derivatives of the log pseudo-likelihood evaluated at coef
- **$inf**: second derivatives of the log pseudo-likelihood
- **$var**: estimated variance covariance matrix of coef
- **$res**: matrix of residuals
$tftime$ vector of unique failure times
$bfjit$ jumps in the Breslow-type estimate of the underlying sub-distribution cumulative hazard (used by predict.crr())
$tfs$ the tfs matrix (output of tf(), if used)
$converged$ TRUE if the iterative algorithm converged
$call$ The call to crr
$n$ The number of observations used in fitting the model
$n.missing$ The number of observations removed from the input data due to missing values
$loglik.null$ The value of the log pseudo-likelihood when all the coefficients are 0

Author(s)
Bingqing Zhou, <bingqing.zhou@yale.edu>

References

See Also
cmprsk

Examples
#library(cmprsk)
crr(time=cdata$time, fstatus=cdata$status, cov1=cdata$z)
# Simulated clustered data set
data(cdata)
crrc(time=cdata[,1],fstatus=cdata[,2],
cov1=cdata[,3],
cluster=cdata[,4])

Description
Regression modeling of subdistribution hazards for stratified right censored data
Two types of stratification are addressed: Regularly stratified: small number of large groups (strata) of subjects Highly stratified: large number of small groups (strata) of subjects
Usage

crrs(ftime, fstatus, cov1, cov2, strata,
tf, failcode=1, cencode=0,
ctype=1,
subsets, na.action=na.omit,
gtol=1e-6, maxiter=10,init)

Arguments

strata stratification covariate
ctype 1 if estimating censoring dist within strata (regular stratification), 2 if estimating censoring dist across strata (highly stratification)
ftime vector of failure/censoring times
fstatus vector with a unique code for each failure type and a separate code for censored observations
cov1 matrix (nobs x ncovs) of fixed covariates (either cov1, cov2, or both are required)
cov2 matrix of covariates that will be multiplied by functions of time; if used, often these covariates would also appear in cov1 to give a prop hazards effect plus a time interaction
tf functions of time. A function that takes a vector of times as an argument and returns a matrix whose jth column is the value of the time function corresponding to the jth column of cov2 evaluated at the input time vector. At time tk, the model includes the term cov2[,j]*tf(tk)[,j] as a covariate.
failcode code of fstatus that denotes the failure type of interest
cencode code of fstatus that denotes censored observations
subsets a logical vector specifying a subset of cases to include in the analysis
na.action a function specifying the action to take for any cases missing any of ftime, fstatus, cov1, cov2, cengroup, or subset.
gtol iteration stops when a function of the gradient is < gtol
maxiter maximum number of iterations in Newton algorithm (0 computes scores and var at init, but performs no iterations)
init initial values of regression parameters (default=all 0)

Details

Fits the stratified extension of the Fine and Gray model (2011). This model directly assesses the effect of covariates on the subdistribution of a particular type of failure in a competing risks setting.

Value

Returns a list of class crr, with components (see crr for details)

$coef the estimated regression coefficients
$loglik log pseudo-likelihood evaluated at coef
$score$ derivitives of the log pseudo-likelihood evaluated at $coef$
$\mathbf{inf}$ -second derivitives of the log pseudo-likelihood
$\mathbf{var}$ estimated variance covariance matrix of $coef$
$\mathbf{res}$ matrix of residuals
$\mathbf{ftime}$ vector of unique failure times
$\mathbf{bfitj}$ jumps in the Breslow-type estimate of the underlying sub-distribution cumulative hazard (used by predict.crr())
$\mathbf{tfs}$ the tfs matrix (output of tf(), if used)
$\mathbf{converged}$ TRUE if the iterative algorithm converged
$\mathbf{call}$ The call to crr
$\mathbf{n}$ The number of observations used in fitting the model
$\mathbf{n.missing}$ The number of observations removed from the input data due to missing values
$\mathbf{loglik.null}$ The value of the log pseudo-likelihood when all the coefficients are 0

**Author(s)**
Bingqing Zhou, <bingqing.zhou@yale.edu>

**References**

**See Also**
cmprsk

**Examples**
```r
## using fine and gray model
crrs($ftime$=center$ftime$, $fstatus$=center$fstatus$,
cov1=cbind(center$fm$,center$cells$))
# # High Stratification: ctype=2
# Random sub-sample
data(center)
cov.test=cbind(center$fm$,center$cells$)
crrs($ftime$=center[,1],$fstatus$=center[,2],
cov1= cov.test,
strata= center$id$,ctype=2)
```
Description
for internal use

Author(s)
Bingqing Zhou

---

# print.crrs

Print method for crrs output

## Description
Prints call for crrs object

## Usage
```r
## S3 method for class 'crrs'
print(x, ...)
```

## Arguments
- `x` crr object (output from `crrs()`)  
- `...` additional arguments to `print()`

## Author(s)
B. Zhou
**psh.test**

*Goodness-of-fit test for proportional subdistribution hazards model*

**Description**

This Goodness-of-fit test proposed a modified weighted Schoenfeld residuals to test the proportionality of subdistribution hazards for the Fine and Gray model.

**Usage**

```r
psh.test(time, fstatus, z, D=c(1,1), tf=function(x) cbind(x,x^2), init)
```

**Arguments**

- `time`: vector of failure times
- `fstatus`: failure status =0 if censored
- `z`: covariates
- `D`: components of `z` that are tested for time-varying effect
- `tf`: functions of `t` for `z` being tested on the same location
- `init`: initial values of regression parameters (default=all 0)

**Details**

The proposed score test employs Schoenfeld residuals adapted to competing risks data. The form of the test is established assuming that the non-proportionality arises via time-dependent coefficients in the Fine-Gray model, similar to the test of Grambsch and Therneau.

**Value**

Returns a data.frame with percentage of cens, cause 1, Test Statistic, d.f., p-value

**Author(s)**

Bingqing Zhou, <bingqing.zhou@yale.edu>

**References**

Examples

```r
data(bce)
attach(bce)
lognodes <- log(nnodes)
Z1 <- cbind(lognodes, ts/10, age, trt)
# trt = 0 if placebo, = 0 treatment
# testing for linear time varying effect of trt
psh.test(time=time, fstatus=type, z=Z1, D=c(0,0,0,1), tf=function(x) x)
# testing for quadratic time varying effect of trt
psh.test(time=time, fstatus=type, z=Z1, D=c(0,0,0,1), tf=function(x) x^2)
# testing for log time varying effect of trt
psh.test(time=time, fstatus=type, z=Z1, D=c(0,0,0,1), tf=function(x) log(x))
# testing for both linear and quadratic time varying effect of trt
psh.test(time=time, fstatus=type, z=Z1, D=matrix(c(0,0,0,0,0,1,0,0,0,1), 4,2), tf=function(x) cbind(x,x^2))
```
Index

*Topic crrSC
  bce, 1
  cdata, 2
  center, 3
*Topic datasets
  bce, 1
  cdata, 2
  center, 3
*Topic survival
  crrc, 3
  crrs, 5

bce, 1

cdata, 2
center, 3
crrc, 3
crrfs (crrvvs), 8
crrfsvs (crrvvs), 8
crrs, 5
crrvvs, 8

print.crrs, 8
psh.test, 9