Package ‘Cprob’

May 22, 2018

Title The Conditional Probability Function of a Competing Event

Version 1.4.1

Author Arthur Allignol

Description Permits to estimate the conditional probability function of a competing event, and to fit, using the temporal process regression or the pseudo-value approach, a proportional-odds model to the conditional probability function (or other models by specifying another link function). See <doi:10.1111/j.1467-9876.2010.00729.x>.

Maintainer Arthur Allignol <arthur.allignol@gmail.com>

Depends prodlim

Imports tpr, geepack, lgtel, graphics, stats, lattice

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2018-05-22 19:34:56 UTC

R topics documented:

cpf ................................................................. 2
cpfpo ............................................................. 4
lines.cpf ......................................................... 5
mgus ................................................................. 6
plot.cpf .......................................................... 7
predict.cpf .................................................... 8
print.cpf ......................................................... 9
print.cpfpo ..................................................... 10
pseudocpf ....................................................... 11
summary.cpf .................................................. 12
summary.pseudocpf ......................................... 13
xyplot.cpfpo ................................................... 14

Index 16
Conditional Probability Function of a Competing Event

**Description**

This function computes estimates of the conditional probability function of a competing event and its variance. It also tests equality of conditional probability functions in two samples.

**Usage**

```r
cpf(formula, data, subset, na.action, conf.int = 0.95, failcode)
```

**Arguments**

- `formula`: A formula object that has a `Hist` object on the left of a `~` operator, and if desired, terms separated by `+` on the right. Note that any subsetting, i.e., `data$var` or `data[, "var"]`, is invalid for this function.
- `data`: A data frame in which the variables in the formula can be interpreted.
- `subset`: Expression identifying a subset of the data to be used for conditional probability estimation.
- `na.action`: A missing-data filter function, applied to the model frame, after any `subset` argument has been used. Default option is `options()$na.action`.
- `conf.int`: Level for pointwise two-sided confidence intervals. Default is 0.95.
- `failcode`: Failure code of the event of interest. Default is the smallest event type provided in the data.

**Details**

The conditional probability function is defined as the probability of having failed due to one competing event (the event of interest), given that no other event has previously occurred (Pepe, 1993).

The `cpf` function aims at estimating this quantity along with its variance at each event times. It also computes a test of equality of conditional probability curves in two samples (and only in two samples).

Of note, if there is more than 2 competing events, the failure types that are not of interest are aggregated into one competing event.

**Value**

`cpf` returns an object of class `cpf` with components

- `cp`: Estimates of the conditional probability function given at all event times
- `var`: Variance estimates
- `time`: Event times
- `lower`: Lower confidence limit for the conditional probability curve
cpf

upper  Upper confidence limit for the conditional probability curve
n.risk  Number of individuals at risk just before t
n.event  A matrix giving the number of events of interest at time t in the first column, and the number of competing events at time t in the second column
n.lost  Number of censored observations at time t
size.strata  Displays the size of each strata
X  Gives covariate's name and labels
strata  Gives the covariate labels that will be used by default for plotting the conditional probability curves, for example.
call  Call that produced the object
z  Test statistic
p  p value of the test
failcode  Same as in function call

Author(s)
Arthur Allignol, <arthur.allignol@gmail.com>

References
M.S. Pepe and M. Mori, Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? Statistics in Medicine, 12(8):737–751.

See Also
Hist, print.cpf, summary.cpf

Examples

data(mgus)

CP <- cpf(Hist(time, ev), data = mgus)
CP

## With age dichotomised according to its median
mgus$AGE <- ifelse(mgus$age < 64, 0, 1)
CP <- cpf(Hist(time, ev)~AGE, data = mgus)
CP
summary(CP)

## Conditional probability of the competing event
CP.death <- cpf(Hist(time, ev), data = mgus, failcode = 2)
CP.death
Proportional-odds Model for the Conditional Probability Function

Description

This function uses the Temporal Process Regression framework to fit a proportional-odds model to the conditional probability function.

Usage

cpfpo(formula, data, subset, na.action, failcode, tis, w, ...)

Arguments

- `formula`: A formula object whose response, on the left of a ~ operator, is a Hist object, and the terms on the right of ~
- `data`: A data.frame in which to interpret the variable names in the formula and subset
- `subset`: Expression specifying that only a subset of the data set should be used
- `na.action`: A missing data filter function applied to the model.frame, after any subset argument has been used. Default is `options()$na.action`
- `failcode`: Integer specifying the code for the event of interest
- `tis`: Vector of timepoints on which the model is fitted
- `w`: Vector of weights. Should be of the same length as `tis`. Default is `rep(1, length(tis))`
- `...`: Further arguments for `tpr`

Details

The conditional probability function of a competing event is the probability of having failed due to one risk (the event of interest) given that no other failure has previously occurred.

The `cpfpo` function fits a proportional-odds model for the conditional probability function within the Temporal Process Regression framework, which is a marginal mean model, where the mean of a response $Y(t)$ at time $t$ is specified conditionally on a vector of covariates $Z$ and a time-dependent stratification factor $S(t)$

$$E\{Y(t)|Z, S(t) = 1\} = g^{-1}\{\beta(t)'Z\}$$

This approach enables the application of standard binary regression models in continuous time.

The regression model is fitted using the `tpr` package. See `tpr` for further details.

Value

cpfpo returns an object of class `cpfpo` and `tpr`. See `tpr` for further details.

Note

As the returned value is also a `tpr` object, all the methods defined in the `tpr` package are available.
lines.cpf

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

References


See Also

tpr, print.cpfpo, xyplot.cpfpo

Examples

data(mgus)
mgus$A <- ifelse(mgus$age < 64, 0, 1)

## fit the model for 2 covariates
fit.cpfpo <- cpfpo(Hist(time, ev)~factor(A) + creat,
                   data = mgus, tis=seq(10, 30, 0.3),
                   w=rep(1,67))

## and plot the odds-ratios
if(require("lattice")) {
  xyplot(fit.cpfpo, scales = list(relation = "free"), layout = c(3, 1))
}

lines.cpf

Lines method for `cpf` objects

Description

Lines method for cpf objects

Usage

## S3 method for class `cpf`
lines(x, conf.int = FALSE, mark.time = FALSE, mark = 3,
       col = 1, lty, ci.lty = 3, ...)
Arguments

x  An object of class cpf
conf.int Logical. Whether to add a pointwise confidence interval
mark.time Controls the labelling of the curves. If set to TRUE, then the curves are marked at each occurrence of a competing event
mark Mark parameter which will be used to label the curves. Same as pch
col A vector of colours
lty A vector specifying the line types for the curves
ci.lty A vector specifying the line type for the confidence intervals
... Other arguments

Value

No value returned

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

See Also

plot.cpf

---

mgus  

Monoclonal Gammapathy of Undetermined Significance

Description

All 241 patients diagnosed with monoclonal gammopathy of undetermined significance at the Mayo Clinic before Jan. 1, 1971, and followed forward until 1992. The interest lies in the possible transformation of mgus towards a cancer of the plasma cells, death from other causes acting as a competing risk.

Usage

data(mgus)

Format

A data frame with 241 observations and the following 10 variables.

id  Patient identification number
time Event time
ev  Event type. 0=censored, 1=cancer of the plasma cells, 2=other
age  Age at diagnostic of mgus
sex Sex. 1=male, 2=female
y.diag Calendar year of diagnosis
albu Albumine level at mgus diagnosis
creat Serum creatinine level at mgus diagnosis
hb Hemoglobine level at mgus diagnosis
size Size of the monoclonal protein peak at mgus diagnosis

Source

Examples
data(mgus)

plot.cpf  Plot method for cpf objects

Description
A plot of conditional probability curves is produced, one for each strata. Pointwise confidence intervals and legend can also be displayed

Usage
## S3 method for class 'cpf'
plot(x, conf.int = FALSE, mark.time = FALSE, mark = 3,
col = 1, lty, ci.lty = 3, xlim, ylim = c(0, 1),
xlabs = "", ylab = "", bty = "n", legend = TRUE,
curvlab = NULL, legend.pos = NULL,
legend.bty = "n", ...)

Arguments

x An object of class cpf
conf.int A logical indicating whether to plot a pointwise confidence interval. Default is FALSE
mark.time Controls the labelling of the curves. If set to TRUE, then the curves are marked at each occurrence of a competing event
mark Mark parameter which will be used to label the curve. The same as pch
col Vector of colours for the curves
lty Vector of integers specifying the line types
ci.lty A vector of integer controlling the line types for the pointwise confidence intervals
predict.cpf

xlim  x-axis limits for the plot area. Default is c(0, 1)
ylim  y-axis limits for the plot area
xlab  x-axis label
ylab  y-axis label
bty   see `par`
legend Whether or nor draw a legend. Default is TRUE
curvlab Text for legend
legend.pos Position for the legend. Default is the upper left corner
legend.bty Box type. See `legend`
...    Further arguments for plot

Value
No value returned

Author(s)
Arthur Allignol, <arthur.allignol@gmail.com>

See Also
`par`, `cpf`

Examples
```r
data(mgus)
mgus$A <- ifelse(mgus$age < 64, 0, 1)

fit <- cpf(Hist(time, ev)~A, mgus)

plot(fit, curvlab=c("Age < 64", "Age >= 64"),
     main = "Conditional Probability of Cancer", xlab = "Years")
```

predict.cpf  Conditional Probability Estimates at chosen timepoints

Description
This function displays estimates of the conditional probability function at timepoints chosen by the user

Usage
```r
## S3 method for class 'cpf'
predict(object, timepoints, ...)
```
Arguments

- object: An object of class `cpf`
- timepoints: Vector of timepoints
- ...: Not used

Value

A data.frame with the following columns:

- time: The given timepoints
- cp: Conditional probability estimates
- var: Variance estimates
- lower: Lower confidence limit
- upper: Upper confidence limit
- n.risk: Number of individuals at risk just before the timepoints
- group: Group, if any

Author(s)

Arthur Allgnol, <arthur.allignol@gmail.com>

See Also

cpf

Examples

data(mgus)
mgus$A <- ifelse(mgus$age < 64, 0, 1)

test <- cpf(Hist(time, ev)~A, data = mgus)
predict(test, c(10, 20))
Arguments

x       A cpf object

...    Further arguments to the print method. Not used here.

Value

No value returned.

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

See Also

cpf

print.cpfpo  Print Method for cpfpo objects

Description

A print method for an object of class cpfpo. It displays the results of test for non-significant effects, along with the results of the test for time-independent effects.

Usage

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

Arguments

x       An object of class cpfpo

...    Not used

Value

An invisible list with 2 components

sig       Results of the test on non-significant effects
tdep      Results of the test on the time-dependence of the effects

See tpr.test

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

See Also

tpr.test, cpfpo
Pseudocpf

Pseudo values for the conditional probability function

Description

The function computes pseudo values and then fit a proportional-odds model to the conditional probability function using GEE.

Usage

pseudocpf(formula, data, id, subset, na.action, timepoints, failcode = 1, ...)

Arguments

- **formula**: A formula object, whose terms are on the right of a ~ operator and the response, a Hist object, on the left.
- **data**: A data frame in which to interpret the formula.
- **id**: Individual patient id.
- **subset**: Expression specifying that only a subset of the data set should be used.
- **na.action**: A missing data filter funtion applied to the model.frame, after any subset argument has been used. Default is options($)na.action.
- **timepoints**: Time points at which to compute the pseudo values.
- **failcode**: Integer that specifies which event is of interest.
- **...**: Other arguments for the geese function.

Details

The regression model is fitted using a method based on the pseudo-values from a jackknife statistic constructed from the conditional probability curve. Then a GEE model is used on the pseudovalues to obtain the odds-ratios.

Value

Returns an object of class pseudocpf containing the following components:

- **fit**: A geese object.
- **pseudo**: The pseudo values computed at the specified time points.
- **timepoints**: Same as in the function call.
- **call**: The matched call.

Note

Besides the estimated regression coefficients, the function returns the computed pseudo-values, so that one can fit a different model, e.g., with a different link function.
Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

References


See Also

`geese`, `summary.pseudocpf`

Examples

```r
data(mgus)
cutoffs <- quantile(mgus$time, probs = seq(0, 1, 0.05))[-1]

### with fancy variance estimation
fit1 <- pseudocpf(Hist(time, ev) ~ age + creat, mgus, id = id, timepoints = cutoffs, corstr = "independence", scale.value = TRUE)
summary(fit1)

### with jackknife variance estimation
fit2 <- pseudocpf(Hist(time, ev) ~ age + creat, mgus, id = id, timepoints = cutoffs, corstr = "independence", scale.value = TRUE, jack = TRUE)
summary(fit2)
```

**summary.cpf**

*Summary method for cpf*

**Description**

Provides a summary of a cpf object.

**Usage**

```r
# S3 method for class 'cpf'
summary(object, ...)
# S3 method for class 'summary.cpf'
print(x, ...)
```
Arguments

- **object**: An object of class `cpf`
- **x**: An object of class `summary.cpf`
- **...**: Not used

Value

A list with the following components:

- **est**: A list of data.frame according to the covariate number of levels. Each data.frame contains conditional probability estimates, variance estimates, number of individual at risk at each event times, number of events for the event of interest, number of competing events, and the event times.
- **call**: Call that produced the `cpf` object
- **X**: Covariate’s name and levels
- **z, p**: Test statistic and p value

Author(s)

arthur allignol, <arthur.allignol@gmail.com>

See Also

cpf

summary.pseudocpf  
**Summary method for pseudocpf objects**

Description

Summary method for pseudocpf objects

Usage

```r
## S3 method for class 'pseudocpf'
summary(object, conf.int = 0.95, scale = 1, ...)
```

Arguments

- **object**: An object of class `pseudocpf`
- **conf.int**: Level for the two-sided confidence intervals
- **scale**: Vector of scale factors for the coefficients, defaults to 1. The confidence limits are for the risk change associated with one scale unit
- **...**: Other arguments
Value

Returns an object of class `summary.pseudocpf` that includes the following components:

- **call**: The matched call
- **coefficients**: A matrix with 5 columns including the regression coefficients, odds-ratios, standard-errors, wald statistics and corresponding two-sided p-values
- **conf.int**: A matrix with 4 columns that consists of the odds-ratios, \( \exp(-\text{coef}) \) and the lower and upper bounds of the confidence interval

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

See Also

- `pseudocpf`

xyplot.cpfpo

'xyplot' method for object of class 'cpfpo'

Description

This function plots the results of the proportional-odds model fitted to the conditional probability function using `cpfpo`.

Usage

```r
## S3 method for class 'cpfpo'
xyplot(x, data = NULL, conf.int = TRUE, level = 0.95,
       odds = TRUE, intercept = TRUE, ylab, xlab, lty = c(1, 3, 3),
       col = c(1, 1, 1), ...)```

Arguments

- **x**: An object of class `cpfpo`
- **data**: *Useless. Can be kept to NULL*
- **conf.int**: Logical. Whether to plot pointwise confidence intervals. Default is `TRUE`
- **level**: Level of the pointwise confidence interval. Default is `0.95`
- **odds**: If set to `TRUE`, the odds-ratios are displayed. Otherwise their logs are plotted. Default is `TRUE`
- **intercept**: Logical. Controls whether the intercept should also be displayed. Default is `TRUE`
- **ylab**: Label for the y-axis
- **xlab**: Label for the x-axis
Value

The function returns a `trellis` object. See `xyplot` for further details.

Author(s)

Arthur Allignol, <arthur.allignol@gmail.com>

See Also

cpfpo, xyplot
Index

*Topic **datasets**
  mgus, 6

*Topic **hplot**
  lines.cpf, 5
  plot.cpf, 7
  xyplot.cpfpo, 14

*Topic **print**
  print.cpf, 9
  print.cpfpo, 10
  summary.cpf, 12
  summary.pseudocpf, 13

*Topic **regression**
  cpfpo, 4
  pseudocpf, 11

*Topic **survival**
  cpf, 2
  cpfpo, 4
  lines.cpf, 5
  mgus, 6
  plot.cpf, 7
  predict.cpf, 8
  pseudocpf, 11
  summary.cpf, 12
  summary.pseudocpf, 13
  xyplot.cpfpo, 14
  [.cpf (cpf), 2
  cpf, 2, 8–10, 13
  cpfpo, 4, 10, 14, 15
  geese, 11, 12
  Hist, 3
  legend, 8
  lines.cpf, 5
  mgus, 6
  par, 8
  plot.cpf, 6, 7
  predict.cpf, 8
  print.cpf, 3, 9
  print.cpfpo, 5, 10
  print.summary.cpf (summary.cpf), 12
  print.summary.pseudocpf (summary.pseudocpf), 13
  pseudocpf, 11, 14
  summary.cpf, 3, 12
  summary.pseudocpf, 12, 13
  tpr, 4, 5
  tpr.test, 10
  xyplot, 15
  xyplot.cpfpo, 5, 14

16