Package 'cohorttools'

May 8, 2026

Type Package

Title Cohort Data Analyses

Version 0.1.8

Maintainer Jari Haukka <jari.haukka@helsinki.fi>

Depends R (>= 3.6), Epi, cmprsk, ggplot2

Imports stats, survival, DiagrammeR, DiagrammeRsvg, rsvg, mgcv

Suggests knitr, rmarkdown, lattice, mstate, testthat

Description Functions to make lifetables and to calculate hazard function estimate using Poisson regression model with splines. Includes function to draw simple flowchart of cohort study. Function boxesLx() makes boxes of transition rates between states. It utilizes 'Epi' package 'Lexis' data.

License GPL-2

Encoding UTF-8

RoxygenNote 7.3.3

NeedsCompilation no

Author Jari Haukka [aut, cre] (ORCID: <https://orcid.org/0000-0003-1450-6208>)

Repository CRAN

Date/Publication 2026-04-21 05:40:03 UTC

Contents

boxesLx *Boxes plot summarizing Lexis object*

Description

Creates boxes graph describing Lexis

Usage

```
boxesLx(
  x,
  layout = "circo",
  prop.penwidth = FALSE,
  scale.Y = 1,
  rankdir = "TB",
  node.attr = "shape=box",
  edge.attr = "minlen=1",
  show.loop = FALSE,
  show.persons = FALSE,
  fontsizeN = 14,
  fontsizeL = 8,
  show.gr = TRUE
)
```

Arguments

x	Lexis object
layout	Graphviz layout "circo", "dot", "twopi" or, "neato". It determines general layout of graph.
prop.penwidth	use line width relative to incidence. If TRUE linewidths of showing transition rates beween states are relative to log of rate.
scale.Y	scale for incidence. Scale factor rates, default is 1.
rankdir	for graph, default is TB. NOTE! this works best with layout "dot"
node.attr	general node attributers. Attributes like shape, color, fillcolor, etc. for nodes. Consult Graphviz documentation for details https://www.graphviz.org/doc/info/attrs.html.
edge.attr	general edge (line) attributers. Attributes like color, arrowhead, fontcolor etc. for edges. Consult Graphviz documentation for details https://www.graphviz.org/doc/info/attrs.html
show.loop	, should loop (staying in same state be shown), default FALSE
show.persons	, should number of persons be shown (entry->exit), default FALSE
fontsizeN	font size for nodes
fontsizeL	font size for edges
show.gr	should graph be shown. If TRUE, function DiagrammeR::grViz is used to show graph.

Value

Character vector containing Graphviz script. This may used to create graph by DiagrammeR::grViz function.

Author(s)

Jari Haukka jari.haukka@helsinki.fi

Examples

```
library(DiagrammeR)
library(survival)
library(Epi)
library(mstate)
data(ebmt3)
bmt <- Lexis(exit = list(tft = rfstime/365.25),
             exit.status = factor(rfsstat, labels = c("Tx", "RD")),
                          data = ebmt3)
bmtr <- cutLexis(bmt, cut = bmt$prtime/365.25, precursor.states = "Tx",
                                          new.state = "PR")

summary(bmtr)
kk<-boxesLx(bmtr)
## Not run:
# Graph to file
gv2image(kk, file="k1", type="pdf")

## End(Not run)
boxesLx(bmtr,layout="dot",rankdir = "LR",show.loop = FALSE,show.persons = TRUE)
boxesLx(bmtr,node.attr='shape=hexagon color=navy style=filled fillcolor=lightblue',
edge.attr = ' color=steelblue arrowhead=vee fontcolor="#8801d7" ',
layout="circo",prop.penwidth=TRUE)
```

`estim.hazard` *Estimates hazard function using Poisson model*

Description

Estimates hazard function using Poisson model

Usage

```
estim.hazard(
  formula,
  data,
  time,
  status,
  breaks,
  knots,
```

```
  time.eval = breaks,
  alpha = 0.05,
  use.GAM = FALSE,
  print.GAM.summary = FALSE,
  ...
)
```

Arguments

formula	formula with Surv in LHS, NOTE! only one variable in RHS
data	data used by formula
time	time variables
status	status indicator Lowest value used as sensoring. If only one unique value detected, all are assumed events
breaks	time is splitted with these values
knots	knots for natural splines used in estimation of hazard function
time.eval	in which time points hazard function is evaluate.
alpha	significance level for confidence intervals
use.GAM	logical determining if generalized additive model (GAM) is used
print.GAM.summary	logical determining if summary of GAM is printed
...	parameters for glm

Value

Returns data frame with time and hazard function values with attribute 'estim.hazard.param' containing estimation parameters (breaks and knots)

Author(s)

Jari Haukka <jari.haukka@helsinki.fi>

Examples

```
library(survival)
tmp.hz<-estim.hazard(time=lung$time,status=lung$status)
head(tmp.hz,2)
attributes(tmp.hz)$estim.hazard.param # estimation parameters
tmp.hz2<-estim.hazard(formula=Surv(time,status)~sex,data=lung)
head(tmp.hz2,2)
```

gv2image *Function makes image from graphviz code*

Description

Function makes image from graphviz code

Usage

```
gv2image(gv, file = "gv", type = "png", engine = "dot", ...)
```

Arguments

gv	character string containing graphviz code
file	file name for image, character string
type	type of ('pdf', 'png', 'ps', 'raw','svg','webp') as character string
engine	grViz engine, defaults is 'dot'
...	parameters for rsvg_

Value

Invisible name of file created.

Author(s)

Jari Haukka <jari.haukka@helsinki.fi>

mkflowchart *Function makes flowchart in graphviz*

Description

Function makes flowchart in graphviz

Usage

```
mkflowchart(N, text.M, text.P, type = 1)
```

Arguments

N	Population sizes
text.M	Text for exclusions, length one less than N
text.P	Text for main boxes, must be same length with N
type	flowchart type (1 or 2)

Value

Character string, graphviz language

Author(s)

Jari Haukka <jari.haukka@helsinki.fi>

Examples

```
DiagrammeR::grViz(mkflowchart(N=c(743,32,20),
text.M=c("Excluded","Excluded \n other with reasons"),
text.P=c("Studies","Relevant studies","Included in final review"),type=1))
```

`mkratetable` *Function makes rate table with confidence intervals for crude incidences (rates)*

Description

Function makes rate table with confidence intervals for crude incidences (rates)

Usage

```
mkratetable(formula, data, alpha = 0.05, add.RR = FALSE, lowest.N = 0, ...)
```

Arguments

`formula`	where Surv object is on lhs and marginal variable(s) on rhs. Marginal variables should usually be factors
`data`	data.frame to be used
`alpha`	confidence level, default is 0.05
`add.RR`	should rate ratio (RR) be added
`lowest.N`	lowest frequency to be shown
`...`	additional parameter for function survival::pyears

Value

table with columns named after marginal variables and n, event, incidence, se, exact.lower95ci and exact.upper95ci variables

Note

packages survival is utilized. Frequencies lower than lowest.N replaced by 999999 Person-years scaled by default with 365.25

Author(s)

Jari Haukka <jari.haukka@helsinki.fi>

Examples

```
library(survival)
tmp.lt1<-mkratetable(Surv(time,status)~ sex,data=lung)
tmp.lt2<-mkratetable(Surv(time,status)~ sex+ph.ecog,data=lung,add.RR=TRUE,lowest.N=10)
```

---

plotcuminc *Plots cumulative incidence rates*

Description

Plots cumulative incidence rates

Usage

```
plotcuminc(ftime, fstatus, cencode, pop.length = 50, group, ...)
```

Arguments

ftime	failure time variable
fstatus	variable with distinct codes for different causes of failure and also a distinct code for censored observations
cencode	value of fstatus variable which indicates the failure time is censored.
pop.length	number of population sizes shown
group	plots will be made for each group. If missing then treated as all one group
...	additional parameters

Value

if missing group ggplot2 object or if group given named list of ggplot2 objects

Note

package cmprsk and ggplot2 are utilized

Author(s)

Jari Haukka <jari.haukka@helsinki.fi>

Examples

```
set.seed(2)
ss <- rexp(100)
gg <- factor(sample(1:3,100,replace=TRUE),1:3,c('a','b','c'))
cc <- sample(0:2,100,replace=TRUE)
print(plotcuminc(ftime=ss,fstatus=cc,cencode=0))
print(plotcuminc(ftime=ss,fstatus=cc,cencode=0,group=gg))
```

`plotratetable` *Function makes plot(s) from ratetable*

Description

Function makes plot(s) from ratetable

Usage

```
plotratetable(rt, RR = FALSE)
```

Arguments

`rt`	Rate table produced by function mkratetable
`RR`	Boolean, if TRUE rate ratios plotted

Value

ggplot object, or list if multiple variables in rate table

Examples

```
library(ggplot2)
library(survival)
tmp.lt1<-mkratetable(Surv(time,status)~ ph.ecog,data=lung,add.RR = FALSE)
plotratetable(tmp.lt1)
tmp.lt2<-mkratetable(Surv(time,status)~ sex+ph.ecog+cut(age,4),data=lung,add.RR=TRUE,lowest.N=1)
plotratetable(tmp.lt2,TRUE)
```

Index