Package 'curesurv'

September 17, 2024

Type Package

Title Mixture and Non Mixture Parametric Cure Models to Estimate Cure Indicators

Version 0.1.1

Author Juste Goungounga [aut, cre] (<https://orcid.org/0000-0002-9039-2639>), Judith Breaud [aut] (<https://orcid.org/0009-0001-1432-6883>), Olayide Boussari [aut] (<https://orcid.org/0000-0002-0343-853X>), Valerie Jooste [aut] (<https://orcid.org/0000-0002-9902-0700>)

Maintainer Juste Goungounga <juste.goungounga@ehesp.fr>

Description Fits a variety of cure models using excess hazard modeling methodology such as the mixture model proposed by Phillips et al. (2002) <doi:10.1002/sim.1101> The Weibull distribution is used to represent the survival function of the uncured patients; Fits also non-mixture cure model such as the time-to-null excess hazard model proposed by Boussari et al. (2020) <doi:10.1111/biom.13361>.

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.3.1

Depends R (>= 3.5), stringr, survival

Imports numDeriv, stats, randtoolbox, bbmle, optimx, Formula, Deriv, statmod

Suggests testthat, knitr, rmarkdown, xhaz, survexp.fr

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2024-09-17 16:20:46 UTC

Contents

`AIC.curesurv` *Akaike's An Information Criterion for cure models*

Description

Calculates the Akaike's "An Information Criterion" for fitted models from `curesurv`

Usage

```
## S3 method for class 'curesurv'
AIC(object, ..., k = 2)
```

Arguments

`object`	a fitted model object obtained from `curesurv`
`...`	optionally more fitted model objects obtained from `curesurv`.
`k`	numeric, the penalty per parameter to be used; the default k = 2 is the classical AIC.

Details

When comparing models fitted by maximum likelihood to the same data, the smaller the AIC, the better the fit.

However in our case, one should be careful when comparing the AIC. Specifically, when one implements a mixture cure model with `curesurv` without correcting the rate table (`pophaz.alpha=FALSE`), one is not obligated to specify `cumpophaz`. However, you cannot compare a model where `cumpophaz` is not specified with a model where `cumpophaz` is specified. If one wants to compare different models using AIC, one should always specify `cumpophaz` when using the `curesurv` function.

Value

the value corresponds to the AIC calculated from the log-likelihood of the fitted model if just one object is provided. If multiple objects are provided, a data.frame with columns corresponding to the objects and row representing the AIC

Examples

```
library("curesurv")
library("survival")

 testiscancer$age_crmin <- (testiscancer$age- min(testiscancer$age)) /
              sd(testiscancer$age)

fit_m1_ad_tneh <- curesurv(Surv(time_obs, event) ~ z_tau(age_crmin) +
                           z_alpha(age_crmin),
                           pophaz = "ehazard",
                           cumpophaz = "cumehazard",
                           model = "nmixture", dist = "tneh",
                           link_tau = "linear",
                           data = testiscancer,
                           method_opt = "L-BFGS-B")

 AIC(fit_m1_ad_tneh)
```

`anova.curesurv` *anova.curesurv function for likelihood-ratio test of two nested models from curesurv function*

Description

This function computes an analysis of deviance table for two excess hazard models fitted using the curesurv R package.

Usage

```
## S3 method for class 'curesurv'
anova(object, ..., test = "LRT")
```

Arguments

`object`	An object of class curesurv.
`...`	Additional object of class curesurv.
`test`	A character string. Computes the likelihood-ratio test for value `"LRT"`. In case the two models are the same, but one with the correction of mortality tables and one without, the likelihood ratio test is computed for value `"LRT_alpha"` These are the only tests available for now.

Value

An object of class anova inheriting from class matrix. The different columns contain respectively the degrees of freedom and the log-likelihood values of the two nested models, the degree of freedom of the chi-square statistic, the chi-square statistic, and the p-value of the likelihood ratio test.

Note

The comparison between two or more models by anova or more excess hazard models will only be valid if they are fitted to the same dataset, and if the compared models are nested. This may be a problem if there are missing values.

Examples

```
library("curesurv")
library("survival")

testiscancer$age_crmin <- (testiscancer$age - min(testiscancer$age)) / sd(testiscancer$age)

fit_m0 <- curesurv(Surv(time_obs, event) ~ 1 | 1,
                           pophaz = "ehazard",
                           cumpophaz = "cumehazard",
                           model = "nmixture", dist = "tneh",
                           link_tau = "linear",
                           data = testiscancer,
                           method_opt = "L-BFGS-B")

fit_m1 <- curesurv(Surv(time_obs, event) ~ age_crmin | 1,
                           pophaz = "ehazard",
                           cumpophaz = "cumehazard",
                           model = "nmixture", dist = "tneh",
                           link_tau = "linear",
                           data = testiscancer,
                           method_opt = "L-BFGS-B")

anova(fit_m0, fit_m1)
```

`cumLexc_mul` *cumLexc_mul function*

Description

returns the cumulative excess hazard for an TNEH model in case of parametrization of log the of the time to null excess hazard as function to fit the data

Usage

```
cumLexc_mul(z_tau, z_alpha, x, theta)
```

Arguments

`z_tau`	covariates depending on tau
`z_alpha`	covariates depending on alpha
`x`	time value
`theta`	of the coefficient of tneh parameters

Value

An object of class numeric containing the cumulative excess hazard with the same length as the time.

`curesurv` *Fitting cure models using curesurv*

Description

Fits the non-mixture cure model proposed by Boussari et *al.* (2020), or mixture cure model such as proposed by De Angelis et *al.* (1999) with the possibility to correct the background mortality as proposed by Phillips et *al.* (2002) in the net survival framework.

Non-mixture cure model:

The Boussari model:
This model allows for direct estimation of time-to-null-excess-hazard which can be interpreted as time-to-cure. The parametrization offers various link functions for the covariates effects on the time-to-null-excess-hazard: $\tau(z_k) = g(\tau_0 + z_k\tau_k)$. If `link_tau=linear`, then g is the identity function. If `link_tau=loglinear` then g is the exponential function. In this model, the cure proportion is expressed as: $\pi(z;\theta) = \exp(-g(\tau_0 + z_k\tau_k)\text{Beta}((\alpha_0 + Z_k\alpha_k), \beta))$.

Mixture cure model:

The user can choose the survival function modeling the uncured patients net survival among Weibull (default) and exponentiated Weibull. The parametrization for weibull distribution is $S_u(t) = (\exp\{-\lambda * (t)^\gamma\})^{\exp(\{\delta Z\})}$. The related hazard function is expressed as:
$lambda_u(t) =$
$gamma$
$lambdat^{gamma-1}$
$exp($
$deltaz)$ The net survival and the excess hazard functions can be respectively expressed as $S_E(t) = \pi(z;\beta)+(1-\pi(z;\beta))S_u(t)$. and $\lambda_E(t) = \frac{(1-\pi(z;\beta))fu(t)}{\pi(z;\beta)+(1-\pi(z;\beta))S_u(t)}$, with $\pi(z;\beta) = \frac{1}{(1+exp(-[\beta_0+Z\beta]))}$.

Correction of background mortality:

Usually, in the net survival framework the expected hazard is directly obtained from life tables. However some patients in cancer registries can have some factors impacting their expected mortality rates (such as comorbidities, deprivation) that are not always accounted #' for in the available life tables, and there is a need to account for this problem. The correction proposed by Phillips et *al* (2002) assumes that $\lambda_{exp}(t,z) = \alpha\lambda_{pop}(t,z_k)$ with $\lambda_{exp}(t,z)$ the patient expected hazard and $\lambda_{pop}(t,z_k)$ the population hazard obtained from life table.

Usage

```
curesurv(
  formula,
  data,
  pophaz = NULL,
```

```
  cumpophaz = NULL,
  pophaz.alpha = FALSE,
  model = "nmixture",
  dist = "weib",
  link_tau = "linear",
  ncoor_des = NULL,
  init = NULL,
  maxit_opt = 10000,
  gradient = FALSE,
  hessian_varcov = TRUE,
  optim_func = "optim",
  optimizer = "optim",
  method_opt = "L-BFGS-B",
  trace = 0,
  nvalues = 10,
  iter_eps = 1e-08,
  optim_fixed = NULL,
  clustertype = NULL,
  nproc = 1,
  subset,
  na.action,
  sign_delta,
  ...
)
```

Arguments

formula	a formula object of the Surv function with the response on the left of a ~ operator and the terms on the right. The response must be a survival object as returned by the Surv function (time in first and status in second).
data	a data frame in which to interpret the variables named in the formula
pophaz	corresponds to the name of the column in the data representing the values of the population instantaneous mortality rates. If the pophaz argument is not specified, overall survival is fitted.
cumpophaz	corresponds to the name of the column in the data representing the values of the instantaneous population cumulative mortality rates. If not specified, the model cannot be compared with model with pophaz.alpha = TRUE using AIC.
pophaz.alpha	to be specified if user want an excess hazard model with correction of mortality rates by a scale parameter
model	To fit a mixture model, specify model = "mixture". To fit Time-To-Null Excess Hazard model the argument is model = "tneh".
dist	For mixture model, it corresponds to the function used to fit the uncured patients survival. By default, ("weib") is used. Another option is the exponentiated Weibull function ("eweib"). For non-mixture models, this argument corresponds to the name of the model. By default, ("tneh") is used to fit the time to null excess hazard model proposed by Boussari et al..

link_tau	must be specified only for model ="tneh". Default is linear link ("linear"). Another link is loglinear ("loglinear").
ncoor_des	if null, the initial parameters are defaults. If else, the initials parameters are obtained via coordinates descent algorithms
init	a list containing the vector of initial values theta_init, the vector of upper bounds theta_upper and the vector of the lower bounds theta_lower for the parameters to estimate. For each elements of the list, give the name of the covariate followed by the vector of the fixed initials values
maxit_opt	option for maximum of iteration in optimization function
gradient	True if optimization process requires gradient to be provided
hessian_varcov	TRUE if user wants variance covariance matrix using hessian function
optim_func	specify which function to be used for optimization purposes.
optimizer	only use this argument when optim_func="bbmle"
method_opt	optimization method used in optim function. The default algorithm is "L-BFGS-B".
trace	Non-negative integer corresponding to the trace argument as in optim
nvalues	number of set of initial values when using multiple initials values
iter_eps	this parameter only works when ncoor_des = "iter"; It allows to run coordinates descent algorithm until the stooping criteria equal at least to the specified value.
optim_fixed	to specify with parameter to not estimated in the estimation process
clustertype	related to cluster type in marqLevAlg package
nproc	number of processors for parallel computing as in marqLevAlg
subset	an expression indicating which subset of the data should be used in the modeling. All observations are included by default
na.action	as in the coxph function, a missing-data filter function.
sign_delta	only used for mixture cure rate models to specify if the effects or minus the effects of covariates acting on uncured survival to be considered. Default will be sign_delta = "1". The alternative is sign_delta = "-1".
...	additional parameters such z_alpha, and z_tau. For more details, use the help function.

Value

An object of class curesurv. This object is a list containing the following components:

iter_coords	number of iterations performed to obtain initial values of the parameters in tneh model only
coefficients	estimates found for the model
estimates	estimates in the appropriate scale for the model

loglik	corresponds to the log-likelihood computed; if only the pophaz is provided, the log-likelihood doesn't correspond to the total log-likelihood. The part of the cumulative population hazard is a constant and is dropped for the computation as presented in Esteve et *al.* (1990); The total log-likekihood is calculed if the user specifies a column name equal expected cumulative mortality (cumpophaz)
iteractions	the number iterations attained to estimate the parameters of the related model
evaluations	the number of times the log-likelihood function was evaluated until to reach the convergence
convergence	an integer code as in optim when L-BFGS-B method is used in optim.
message	a character string returned by the optimizer
varcov	the variance covariance matrix of the parameters estimated
varcov_star	the variance covariance matrix of the coefficients of the model of interest
std_err	the standard errors of the estimated parameters
std_err_star	the standard errors of the coefficients of the model of interest
AIC	the Akaike information criteria from the model of interest
n.events	the number of events in the dataset. Events are considered
n.obs	the number of observations in the dataset.
model	if fitted model is a mixture model, it returns "mixture". If fitted model is Time-To-Null Excess Hazard model, it returns "nmixture".
Terms	the representation of the terms in the model
pophaz.alpha	logical value to indicate if fitted cure model requires correction of mortality rates by a scale parameter
pophaz	corresponds to the the population instantaneous mortality rates.
cumpophaz	corresponds to the population cumulative mortality rates.
frailtyhp	a booleen to be specified if a frailty correction is needed for the population hazard.
dist	For mixture model, it corresponds to the function used to fit the uncured patients survival. By default, ("weib") is used. Another option is the exponentiated Weibull function ("eweib"). For non-mixture models, this argument corresponds to the name of the model. By default, ("tneh") is used to fit the time to null excess hazard model proposed by Boussari et al.
xmax	maximum follow-up time to evaluate the TTC
z_tau	Covariates acting on parameter tau in non mixture cure model tneh
link_tau	returned only for model ="tneh"; returned by default is "linear" or "loglinear" for linear or loglinear link function of covariates acting on tau parameter.
z_alpha	Covariates acting on parameter alpha in non mixture cure model tneh
z_c	Covariates acting on cure fraction in mixture cure model
z_ucured	covariates acting on survival of uncured in mixture cure model
z_pcured	Covariates acting on cure fraction in mixture cure model
z_ucured	covariates acting on survival of uncured in mixture cure model
data	the dataset used to run the model
call	the function call based on model
formula	the formula as a formula object

Note

Note that all these models can be fitted in the overall survival setting.

`time` is OBLIGATORY in years

Author(s)

Juste Goungounga, Judith Breaud, Olayide Boussari, Valerie Jooste

References

Boussari O, Bordes L, Romain G, Colonna M, Bossard N, Remontet L, Jooste V. Modeling excess hazard with time-to-cure as a parameter. Biometrics. 2021 Dec;77(4):1289-1302. doi: 10.1111/biom.13361. Epub 2020 Sep 12. PMID: 32869288. (pubmed)

Boussari O, Romain G, Remontet L, Bossard N, Mounier M, Bouvier AM, Binquet C, Colonna M, Jooste V. A new approach to estimate time-to-cure from cancer registries data. Cancer Epidemiol. 2018 Apr;53:72-80. doi: 10.1016/j.canep.2018.01.013. Epub 2018 Feb 4. PMID: 29414635. (pubmed)

Phillips N, Coldman A, McBride ML. Estimating cancer prevalence using mixture models for cancer survival. Stat Med. 2002 May 15;21(9):1257-70. doi: 10.1002/sim.1101. PMID: 12111877. (pubmed)

De Angelis R, Capocaccia R, Hakulinen T, Soderman B, Verdecchia A. Mixture models for cancer survival analysis: application to population-based data with covariates. Stat Med. 1999 Feb 28;18(4):441-54. doi: 10.1002/(sici)1097-0258(19990228)18:4<441::aid-sim23>3.0.co;2-m. PMID: 10070685. (pubmed)

Botta L, Caffo O, Dreassi E, Pizzoli S, Quaglio F, Rugge M, Valsecchi MG. A new cure model that corrects for increased risk of non-cancer death: analysis of reliability and robustness, and application to real-life data. BMC Med Res Methodol. 2023 Mar 25;23(1):70. doi: 10.1186/s12874-023-01876-x. PMID: N/A. (pubmed)

See Also

`predict.curesurv()`, `print.curesurv()`, `browseVignettes("curesurv")`

Examples

```
library("curesurv")
library("survival")



# Net survival setting
# Mixture cure model with Weibull function for the uncured patients survival:
# no covariate

theta_init2 <- rep(0, 3)
theta_lower2 <- c(-Inf,-Inf,-Inf)
theta_upper2 <- c(Inf, Inf, Inf)
```

```
fit_m0_ml <- curesurv(Surv(time_obs, event) ~ 1 | 1,
             pophaz = "ehazard",
             cumpophaz = "cumehazard",
             model = "mixture", dist = "weib",
             data = testiscancer,
             init = list(theta_init = theta_init2,
             theta_lower = theta_lower2,
             theta_upper = theta_upper2),
             method_opt = "L-BFGS-B")
fit_m0_ml



# Mixture cure model with Weibull function for the uncured patients survival:
#standardized age as covariate


fit_m2_ml <- curesurv(Surv(time_obs, event) ~ age_cr | age_cr,
                  pophaz = "ehazard",
                  cumpophaz = "cumehazard",
                  model = "mixture", dist = "weib",
                  data = testiscancer,
                  method_opt = "L-BFGS-B")

 fit_m2_ml



## Non mixture cure model
### TNEH Null model
#### loglinear effect of covariates on time-to-null excess hazard

theta_init2 <- rep(0, 3)
theta_lower2 <- c(-Inf,-Inf,-Inf)
theta_upper2 <- c(Inf, Inf, Inf)

fit_m0_mult_tneh <- curesurv(Surv(time_obs, event) ~ 1,
                         pophaz = "ehazard",
                         cumpophaz = "cumehazard",
                         model = "nmixture",
                         dist = "tneh", link_tau = "loglinear",
                         data = testiscancer,
                         init = list(theta_init = theta_init2,
                                     theta_lower = theta_lower2,
                                     theta_upper = theta_upper2),
                         method_opt = "L-BFGS-B")


fit_m0_mult_tneh

#### Additive parametrization
theta_init2 <- c(1, 6, 6)
```

```
theta_lower2 <- c(0,1,0)
theta_upper2 <- c(Inf, Inf, Inf)

fit_m0_ad_tneh <- curesurv(Surv(time_obs, event) ~ 1,
                           pophaz = "ehazard",
                           cumpophaz = "cumehazard",
                           model = "nmixture",
                           dist = "tneh", link_tau = "linear",
                           data = testiscancer,
                           init = list(theta_init = theta_init2,
                                       theta_lower = theta_lower2,
                                       theta_upper = theta_upper2),
                           method_opt = "L-BFGS-B")



 fit_m0_ad_tneh

#### Additive parametrization, with covariates
fit_m1_ad_tneh <- curesurv(Surv(time_obs, event) ~ z_alpha(age_cr) +
                           z_tau(age_cr),
                           pophaz = "ehazard",
                           cumpophaz = "cumehazard",
                           model = "nmixture",
                           dist = "tneh", link_tau = "linear",
                           data = testiscancer,
                           method_opt = "L-BFGS-B")



 fit_m1_ad_tneh
```

`dataweib` *Simulated data with vital status information from Weibull mixture cure model*

Description

Simulated data

Usage

```
data(dataweib)
```

Format

This dataset contains the following variables:

age Age at diagnosis

age_cr centered and scaled age at diagnosis

age_classe "<45", "45_59" and ">=60" age groups

sexe "male", "female" gender groups

stage "<0", "1" , "2" and "3" for stage I-IV groups

time_obs Follow-up time (years)

event Vital status

cumehazard individual cumulative expected hazard

ehazard individual instantaneous expected hazard

Examples

```
data(dataweib)
summary(dataweib)
```

---

pancreas_data *Simulated pancreas data with vital status information*

Description

Simulated data

Usage

```
data(pancreas_data)
```

Format

This dataset contains the following variables:

age Age at diagnosis

age_cr centered and scaled age at diagnosis

age_classe "<45", "45_59" and ">=60" age groups

time_obs Follow-up time (years)

event Vital status

cumehazard individual cumulative expected hazard

ehazard individual instantaneous expected hazard

Examples

```
data(pancreas_data)
summary(pancreas_data)
```

plot.predCuresurv *plot method for curesurv prediction objects*

Description

Produces figures of (excess) hazard, (net) survival and probability P(t) of being cured at a given time t after diagnosis knowing that he/she was alive up to time t.

Usage

```
## S3 method for class 'predCuresurv'
plot(
  x,
  fun = "all",
  conf.int = FALSE,
  conf.type = c("log", "log-log", "plain"),
  legend.out = TRUE,
  xlab = "Time since diagnosis",
  ylab.haz = "excess hazard",
  ylab.surv = "net survival",
  ylab.ptcure = "P(t)",
  ylab.cumhaz = "cumulative excess hazard",
  ylab.logcumhaz = "logarithm of cumulative excess hazard",
  col.haz = "black",
  col.surv = "black",
  col.ptcure = "black",
  col.cumhaz = "black",
  col.logcumhaz = "black",
  col.tau = "red",
  col.ttc = "green4",
  col.p95 = "black",
  col.pi = "blue",
  lty.surv = 1,
  lty.haz = 1,
  lty.ptcure = 1,
  lty.cumhaz = 1,
  lty.logcumhaz = 1,
  lty.pi = 2,
  lty.tau = 2,
  lty.ttc = 3,
  lty.p95 = 4,
  lty.ic = 5,
  lwd.main = 1,
  lwd.sub = 1,
  lwd.ic = 1,
  ...
)
```

Arguments

x	result of the predCuresurv function
fun	in "haz" or "surv" or "pt_cure", "cumhaz", "logcumhaz", the plot produced is that of (excess) hazard, or that of (net) survival, or that of the probability P(t) of being cured at a given time t after diagnosis knowing that he/she was alive up to time t is provided, or that of cumulative hazard or that of the logarithm of the cumulative hazard; if fun = "all", the plots of the three first indicators are produced.
conf.int	an argument expected to be TRUE if the confidence intervals of the related-indicator specified by the argument "fun" are needed. The default option is FALSE. Confidence intervals are not available for fun="cumhaz" and fun="logcumhaz"
conf.type	One of "plain", "log", "log-log". The first option causes the standard intervals curve +- k *se(curve), where k is determined from conf.int. The log option calculates intervals based on log(curve). The log-log option bases the intervals on the log(-log(curve)).
legend.out	an argument deciding the place of the legend if fun="all". The default value is TRUE and forces most of the legend on the empty bottom-right plot slot. If value is FALSE, the legend will be printed entirely in each subplot.
xlab	label for the x-axis of the plot.
ylab.haz	optional label for the y-axis of the plot of excess hazard
ylab.surv	optional label for the y-axis of the plot of net survival
ylab.ptcure	optional label for the y-axis of the plot of the probability P(t) of being cured at a given time t after diagnosis knowing that he/she was alive up to time t
ylab.cumhaz	optional label for the y-axis of the plot of cumulative excess hazard
ylab.logcumhaz	optional label for the y-axis of the plot of logarithm of cumulative excess hazard
col.haz	optional argument to specify the color of curve of the excess hazard
col.surv	optional argument to specify the color of curve of the net survival
col.ptcure	optional argument to specify the color of curve of probability P(t) of being cured at a given time t after diagnosis knowing that he/she was alive up to time t.
col.cumhaz	optional argument to specify the color of curve of cumulative excess hazard
col.logcumhaz	optional argument to specify the color of curve of the logarithm of cumulative excess hazard
col.tau	optional argument to specify the color of curve of time-to-null excess hazard
col.ttc	optional argument to specify the color of curve of time-to-cure
col.p95	optional argument to specify the color for the line highlighting ϵ when $P(t) \geq 1 - \epsilon$
col.pi	optional argument to specify the color of cure proportion
lty.surv	stands for line types for net survival
lty.haz	stands for line types for excess hazard
lty.ptcure	stands for line types for probability P(t) of being cured at a given time t after diagnosis knowing that he/she was alive up to time t.

lty.cumhaz	stands for line types for cumulative excess hazard
lty.logcumhaz	stands for line types for logarithm cumulative excess hazard
lty.pi	stands for line types for cure proportion
lty.tau	stands for line types for time-to-null excess hazard
lty.ttc	stands for line types for time-to-cure
lty.p95	stands for line types for the line highlighting ϵ when $P(t) \geq 1 - \epsilon$
lty.ic	stands for line types for confidence intervals
lwd.main	line width for the main line (haz, surv, pt_cure, cumhaz, logcumhaz)
lwd.sub	line width for the additionnal lines (ttc, p95, tau...)
lwd.ic	line width for the confidence intervals lines
...	additional options as in the classical plot method.
ylab	optional label for the y-axis of the plot. Depending to the curve of interest (hazard, survival, probability of being cured at a given time t, or all),the argument must be named ylab.haz, ylab.surv, ylab.ptcure. If missing some default labels are provided depending on the curve of interest. This name can be found in the data.frame from the result of the predict.curesurv function.

Value

No value is returned.

Author(s)

Juste Goungounga, Judith Breaud, Eugenie Blandin, Olayide Boussari, Valerie Jooste

See Also

predict.curesurv(), print.curesurv(), curesurv(), browseVignettes("curesurv")

Examples

```
library("curesurv")
library("survival")

 testiscancer$age_crmin <- (testiscancer$age- min(testiscancer$age)) /
              sd(testiscancer$age)

fit_m1_ad_tneh <- curesurv(Surv(time_obs, event) ~ z_tau(age_crmin) +
                           z_alpha(age_crmin),
                           pophaz = "ehazard",
                           cumpophaz = "cumehazard",
                           model = "nmixture", dist = "tneh",
                           link_tau = "linear",
                           data = testiscancer,
                           method_opt = "L-BFGS-B")
```

```
 fit_m1_ad_tneh


#'  #mean of age
 newdata1 <- with(testiscancer,
 expand.grid(event = 0, age_crmin = mean(age_crmin), time_obs  = seq(0.001,10,0.1)))

 pred_agemean <- predict(object = fit_m1_ad_tneh, newdata = newdata1)


 #max of age
 newdata2 <- with(testiscancer,
 expand.grid(event = 0,
 age_crmin = max(age_crmin),
  time_obs  = seq(0.001,10,0.1)))

 pred_agemax <- predict(object = fit_m1_ad_tneh, newdata = newdata2)

   # predictions at time 2 years and  of age

   newdata3 <- with(testiscancer,
      expand.grid(event = 0,
      age_crmin = seq(min(testiscancer$age_crmin),max(testiscancer$age_crmin), 0.1),
      time_obs  = 2))

   pred_age_val <- predict(object = fit_m1_ad_tneh, newdata = newdata3)

 #plot of 3 indicators for mean age

 plot(pred_agemean, fun="all")


 #plot of net survival for mean and maximum age (comparison)

oldpar <- par(no.readonly = TRUE)

par(mfrow = c(2, 2),
    cex = 1.0)
plot(pred_agemax$time,
    pred_agemax$ex_haz,
    type = "l",
    lty = 1,
    lwd = 2,
    xlab = "Time since diagnosis",
    ylab = "excess hazard")
lines(pred_agemean$time,
     pred_agemean$ex_haz,
     type = "l",
     lty = 2,
     lwd = 2)

legend("topright",
```

```
      horiz = FALSE,
      legend = c("hE(t) age.max = 79.9", "hE(t) age.mean = 50.8"),
      col = c("black", "black"),
      lty = c(1, 2, 1, 1, 2, 2))
grid()

plot(pred_agemax$time,
    pred_agemax$netsurv,
    type = "l",
    lty = 1,
    lwd = 2,
    ylim = c(0, 1),
    xlab = "Time since diagnosis",
    ylab = "net survival")
lines(pred_agemean$time,
     pred_agemean$netsurv,
     type = "l",
     lty = 2,
     lwd = 2)
legend("bottomleft",
       horiz = FALSE,
       legend = c("Sn(t) age.max = 79.9", "Sn(t) age.mean = 50.8"),
       col = c("black", "black"),
      lty = c(1, 2, 1, 1, 2, 2))
grid()

plot(pred_agemax$time,
    pred_agemax$pt_cure,
    type = "l",
    lty = 1,
    lwd = 2,
    ylim = c(0, 1), xlim = c(0,30),
    xlab = "Time since diagnosis",
    ylab = "probability of being cured P(t)")

lines(pred_agemean$time,
     pred_agemean$pt_cure,
     type = "l",
     lty = 2,
     lwd = 2)


abline(v = pred_agemean$tau[1],
      lty = 2,
      lwd = 2,
      col = "blue")
abline(v = pred_agemean$TTC[1],
       lty = 2,
       lwd = 2,
       col = "red")
abline(v = pred_agemax$tau[1],
       lty = 1,
       lwd = 2,
```

```
       col = "blue")
abline(v = pred_agemax$TTC[1],
       lty = 1,
       lwd = 2,
      col = "red")
grid()

legend("bottomright",
       horiz = FALSE,
       legend = c("P(t) age.max = 79.9",
                  "P(t) age.mean = 50.8",
                  "TNEH age.max = 79.9",
                  "TTC age.max = 79.9",
                  "TNEH age.mean = 50.8",
                  "TTC age.mean = 50.8"),
      col = c("black", "black", "blue", "red", "blue", "red"),
      lty = c(1, 2, 1, 1, 2, 2))


 val_age <- seq(min(testiscancer$age_crmin),
                max(testiscancer$age_crmin), 0.1) * sd(testiscancer$age) +
                min(testiscancer$age)


 pred_age_val <- predict(object = fit_m1_ad_tneh, newdata = newdata3)


par(mfrow=c(2,2))
 plot(val_age,
     pred_age_val$ex_haz, type = "l",
     lty=1, lwd=2,
     xlab = "age",
     ylab = "excess hazard")
grid()

 plot(val_age,
     pred_age_val$netsurv, type = "l", lty=1,
     lwd=2, xlab = "age", ylab = "net survival")
     grid()

 plot(val_age,
     pred_age_val$pt_cure, type = "l", lty=1, lwd=2,
     xlab = "age",
     ylab = "P(t)")
     grid()
par(oldpar)
```

`predict.curesurv` *Prediction for a curesurv cure model*

Description

return predicted (excess) hazard, (net) survival, cure fraction and time to null excess hazard or time to cure.

Usage

```
## S3 method for class 'curesurv'
predict(
  object,
  newdata = NULL,
  xmax = 10^9,
  level = 0.975,
  epsilon = 0.05,
  sign_delta = 1,
  ...
)
```

Arguments

object	Output from curesurv function
newdata	the new data to be specified for predictions; If else, predictions are made using the data provided during the estimation step in order to obtain the output from curesurv function.
xmax	maximum time at which Time-to-Cure is evaluated numerically.
level	$1 - \frac{\alpha}{2}$-order quantile of a normal distribution for the confidence intervals
epsilon	value fixed by user to estimate the TTC $Pi(t) \geq 1 - \epsilon$. By default epsilon = 0.05.
sign_delta	sign of effect of delta on covariates acting on survival function, positive by default "sign_delta = 1" and alternative is "sign_delta = -1"
...	additional parameters

Value

An object of class c("pred_curesurv", "data.frame"). This object is a list containing the following components:

time	time in the input new data
ex_haz	predicted excess hazard at the time provided in the new data
netsurv	predicted net survival at the time provided in the new data
pt_cure	probability to be cured
tau	time to null in model TNEH when object corresponds to the results from Boussari model or its extension.
netsurv_tau	pi or net survival at time tau when object corresponds to the results from Boussari model or its extension.
time_to_cure_ttc	time to cure (TTC)

Author(s)

Juste Goungounga, Judith Breaud, Olayide Boussari, Valerie Jooste

References

Boussari O, Bordes L, Romain G, Colonna M, Bossard N, Remontet L, Jooste V. Modeling excess hazard with time-to-cure as a parameter. Biometrics. 2021 Dec;77(4):1289-1302. doi: 10.1111/biom.13361. Epub 2020 Sep 12. PMID: 32869288. (pubmed)

Boussari O, Romain G, Remontet L, Bossard N, Mounier M, Bouvier AM, Binquet C, Colonna M, Jooste V. A new approach to estimate time-to-cure from cancer registries data. Cancer Epidemiol. 2018 Apr;53:72-80. doi: 10.1016/j.canep.2018.01.013. Epub 2018 Feb 4. PMID: 29414635. (pubmed)

Phillips N, Coldman A, McBride ML. Estimating cancer prevalence using mixture models for cancer survival. Stat Med. 2002 May 15;21(9):1257-70. doi: 10.1002/sim.1101. PMID: 12111877. (pubmed)

De Angelis R, Capocaccia R, Hakulinen T, Soderman B, Verdecchia A. Mixture models for cancer survival analysis: application to population-based data with covariates. Stat Med. 1999 Feb 28;18(4):441-54. doi: 10.1002/(sici)1097-0258(19990228)18:4<441::aid-sim23>3.0.co;2-m. PMID: 10070685. (pubmed)

See Also

`print.curesurv()`, `curesurv()`, `browseVignettes("curesurv")`

Examples

```
library("curesurv")
library("survival")

fit_m2_ml <- curesurv(Surv(time_obs, event) ~ age_cr|age_cr,
                  pophaz = "ehazard",
                  cumpophaz = "cumehazard",
                  model = "mixture",
                  data = pancreas_data,
                  method_opt = "L-BFGS-B")

 fit_m2_ml

 newdata <- pancreas_data[2,]

 predict(object = fit_m2_ml, newdata = newdata)

## Non mixture cure model
### TNEH model

#### Additive parametrization

testiscancer$age_crmin <- (testiscancer$age- min(testiscancer$age)) /
              sd(testiscancer$age)
```

```
fit_m1_ad_tneh <- curesurv(Surv(time_obs, event) ~ z_tau(age_crmin) +
                           z_alpha(age_crmin),
                           pophaz = "ehazard",
                           cumpophaz = "cumehazard",
                           model = "nmixture", dist = "tneh",
                           link_tau = "linear",
                           data = testiscancer,
                           method_opt = "L-BFGS-B")

 fit_m1_ad_tneh

 predict(object = fit_m1_ad_tneh, newdata = testiscancer[3:6,])

 #mean of age
 newdata1 <- with(testiscancer,
 expand.grid(event = 0, age_crmin = mean(age_crmin), time_obs  = seq(0.001,10,0.1)))

 pred_agemean <- predict(object = fit_m1_ad_tneh, newdata = newdata1)


 #max of age
 newdata2 <- with(testiscancer,
 expand.grid(event = 0,
 age_crmin = max(age_crmin),
  time_obs  = seq(0.001,10,0.1)))

 pred_agemax <- predict(object = fit_m1_ad_tneh, newdata = newdata2)
 head(pred_agemax)
```

`print.curesurv` *print a curesurv object*

Description

Print an object of class "curesurv"

Usage

```
## S3 method for class 'curesurv'
print(x, digits = max(1L, getOption("digits") - 3L), signif.stars = FALSE, ...)
```

Arguments

`x`	an object of class "curesurv".
`digits`	minimum number of significant digits to be used for most numbers.

signif.stars	logical; if TRUE, P-values are additionally encoded visually as "significance stars" in order to help scanning of long coefficient tables.
...	additional options

Value

an object of class "curesurv" representing the fit. See curesurv for details.

Author(s)

Juste Goungounga, Judith Breaud, Eugenie Blandin, Olayide Boussari, Valerie Jooste

References

Boussari O, Bordes L, Romain G, Colonna M, Bossard N, Remontet L, Jooste V. Modeling excess hazard with time-to-cure as a parameter. Biometrics. 2020 Aug 31. doi: 10.1111/biom.13361. Epub ahead of print. PMID: 32869288. (pubmed)

Phillips N, Coldman A, McBride ML. Estimating cancer prevalence using mixture models for cancer survival. Stat Med. 2002 May 15;21(9):1257-70. doi: 10.1002/sim.1101. PMID: 12111877. (pubmed)

De Angelis R, Capocaccia R, Hakulinen T, Soderman B, Verdecchia A. Mixture models for cancer survival analysis: application to population-based data with covariates. Stat Med. 1999 Feb 28;18(4):441-54. doi: 10.1002/(sici)1097-0258(19990228)18:4<441::aid-sim23>3.0.co;2-m. PMID: 10070685. (pubmed)

See Also

predict.curesurv(), curesurv(), browseVignettes("curesurv")

Examples

```
library("curesurv")
library("survival")



# overall survival setting
# Mixture cure model with Weibull function for the uncured patients survival:
# no covariate



fit_ml0 <- curesurv(Surv(time_obs, event) ~ 1 | 1,
             model = "mixture", dist = "weib",
             data = testiscancer,
             method_opt = "L-BFGS-B")


print(fit_ml0)
```

summary.curesurv *summary for a curesurv cure model*

Description

summary an object of class "curesurv"

Usage

```
## S3 method for class 'curesurv'
summary(
  object,
  digits = max(1L, getOption("digits") - 3L),
  signif.stars = FALSE,
  ...
)
```

Arguments

object	an object of class "curesurv".
digits	minimum number of significant digits to be used for most numbers.
signif.stars	logical; if TRUE, P-values are additionally encoded visually as "significance stars" in order to help scanning of long coefficient tables.
...	additional options

Value

an object of class "curesurv" representing the fit. See curesurv for details.

Author(s)

Juste Goungounga, Judith Breaud, Eugenie Blandin, Olayide Boussari, Valerie Jooste

References

Boussari O, Bordes L, Romain G, Colonna M, Bossard N, Remontet L, Jooste V. Modeling excess hazard with time-to-cure as a parameter. Biometrics. 2020 Aug 31. doi: 10.1111/biom.13361. Epub ahead of print. PMID: 32869288. (pubmed)

Phillips N, Coldman A, McBride ML. Estimating cancer prevalence using mixture models for cancer survival. Stat Med. 2002 May 15;21(9):1257-70. doi: 10.1002/sim.1101. PMID: 12111877. (pubmed)

De Angelis R, Capocaccia R, Hakulinen T, Soderman B, Verdecchia A. Mixture models for cancer survival analysis: application to population-based data with covariates. Stat Med. 1999 Feb 28;18(4):441-54. doi: 10.1002/(sici)1097-0258(19990228)18:4<441::aid-sim23>3.0.co;2-m. PMID: 10070685. (pubmed)

See Also

predict.curesurv(), curesurv(), browseVignettes("curesurv")

Examples

```
library("curesurv")
library("survival")



# overall survival setting
# Mixture cure model with Weibull function for the uncured patients survival:
# no covariate



fit_ml0 <- curesurv(Surv(time_obs, event) ~ 1 | 1,
             model = "mixture", dist = "weib",
             data = testiscancer,
             method_opt = "L-BFGS-B")


 summary(fit_ml0)
```

---

testiscancer *Simulated testis cancer data using a cure model*

Description

Simulated dataset of 2000 individuals as in Boussari et *al.* (2020), following setting 1 sub-scenario design.

Usage

```
data(testiscancer)
```

Format

This dataset contains the following variables:

age Age at diagnosis

age_cr centered and scaled age at diagnosis

age_classe "<40", "40_65" and ">=65" age groups

time_obs Follow-up time (years)

event Vital status

cumehazard individual cumulative expected hazard

ehazard individual instantaneous expected hazard

weisurvpop individual expected survival

Examples

```
data(testiscancer)
summary(testiscancer)
```

`z_alpha` *z_alpha function identifying variables acting on alpha parameter*

Description

variables adjusted on alpha parameter in non-mixture cure model with "tneh" specified for the distribution.

Usage

```
z_alpha(x)
```

Arguments

x	a simple formula.

Value

the variable x

Author(s)

Juste Goungounga, Judith Breaud, Olayide Boussari, Gaelle Romain, Valerie Jooste

References

Boussari O, Bordes L, Romain G, Colonna M, Bossard N, Remontet L, Jooste V. Modeling excess hazard with time-to-cure as a parameter. Biometrics. 2020 Aug 31. doi: 10.1111/biom.13361. Epub ahead of print. PMID: 32869288. (pubmed)

`z_tau` *z_tau function identifying variables acting on tau parameter*

Description

variables adjusted on tau parameter in non-mixture cure model with "tneh" specified for the distribution.

Usage

```
z_tau(x)
```

Arguments

x	the name of the column in the dataset representing the variable that will act on tau parameter of the "tneh" model

Value

the variable x

Author(s)

Juste Goungounga, Judith Breaud, Eugenie Blandin, Olayide Boussari, Valerie Jooste

References

Boussari O, Bordes L, Romain G, Colonna M, Bossard N, Remontet L, Jooste V. Modeling excess hazard with time-to-cure as a parameter. Biometrics. 2020 Aug 31. doi: 10.1111/biom.13361. Epub ahead of print. PMID: 32869288. (pubmed)

Index