Package 'mexhaz'

March 20, 2024

Type Package

Title Mixed Effect Excess Hazard Models

Version 2.5

Date 2024-03-20

Author Hadrien Charvat, Aurelien Belot

Maintainer Hadrien Charvat <h.charvat.ef@juntendo.ac.jp>

Description Fit flexible (excess) hazard regression models with the possibility of including non-proportional effects of covariables and of adding a random effect at the cluster level (corresponding to a shared frailty). A detailed description of the package functionalities is provided in Charvat and Belot (2021) <doi:10.18637/jss.v098.i14>.

Depends survival

Imports statmod, MASS, numDeriv, splines, lamW

Suggests utils, rstpm2

License GPL (>= 2)

NeedsCompilation yes

Repository CRAN

Date/Publication 2024-03-20 11:50:02 UTC

R topics documented:

1

mexhaz–package   *Mixed effect parametric excess hazard models*

Description

Fit an (excess) hazard regression model using different shapes for the baseline hazard (Weibull, piecewise constant, exponential of a B-spline of degree 1 to 3, exponential of a restricted cubic spline), with the possibility to include time-dependent effects of variable(s) and a random effect defined at the cluster level. Follow-up time can be entered in the right-censored or counting process input style. The latter allows the modelling of survival data with delayed entries. The time-dependent effect of a covariable is modelled by adding interaction terms between the covariable and a function of time of the same class as the one used for the baseline hazard (in particular, with the same knots for piecewise constant hazards; and with the same degree and the same knots for B-spline or restricted cubic spline functions). The random effect is assumed to be normally distributed with mean 0 and standard deviation sigma. The optimisation process uses adaptive Gaussian quadrature to calculate the cluster-specific marginal likelihoods. The logarithm of the full marginal likelihood, defined as the sum of the logarithms of the cluster-specific marginal likelihoods, is then maximised using an optimisation routine such as `nlm` (default) or `optim`. Functions to compute and plot the predicted (excess) hazard and (net) survival are provided. In the case of a random intercept model, three types of predictions can be obtained : marginal (population-averaged), cluster-specific posterior (cluster-averaged), or conditional (for a given quantile of the ditribution of the random effect).

Author(s)

Hadrien Charvat, Aurelien Belot

References

Charvat H, Remontet L, Bossard N, Roche L, Dejardin O, Rachet B, Launoy G, Belot A; CENSUR Working Survival Group. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Stat Med 2016;35(18):3066-3084 (doi: 10.1002/sim.6881)

Charvat H, Belot A. An R package for fitting flexible hazard-based regression models for overall and excess mortality with a random effect. J Stat Softw 2021;98(14):1-36 (doi: 10.18637/jss.v098.i14)

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib_mix <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull",
expected="popmrate", verbose=0, random="clust")


## Examples of syntax for various models (not run)

## Fit of a fixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution and with effects of age (agecr),
## deprivation index (depindex) and sex (IsexH) using the optim
## procedure and the BFGS method (see help of optim).

# Mod_weib <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="weibull",
# expected="popmrate", verbose=1000, fnoptim="optim",
# method="BFGS")


## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a cubic B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

# Mod_bs3_2mix <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
# degree=3, knots=c(1,5), expected="popmrate", random="clust",
# verbose=1000)


## Fit of a fixed-effect overall hazard model, with the baseline hazard
## described by a piecewise constant function with the following vector
## of knots (defining the endpoints of the intervals on which the hazard
## is constant): (1,3,5,8), and with effects of age (agecr), deprivation
## index (depindex) and sex (IsexH)

# Mod_pw <- mexhaz(formula=Surv(time=timesurv, event=vstat)~
# agecr+depindex+IsexH, data= simdatn1, base="pw.cst", knots=c(1,3,5,8),
# verbose=1000)


## Fit of a fixed-effect excess hazard model, with the baseline hazard
## described by a cubic B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

# Mod_bs3_2 <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
```

```
# degree=3, knots=c(1,5), expected="popmrate", verbose=1000)


## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a cubic B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

# Mod_bs3_2mix <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
# degree=3, knots=c(1,5), expected="popmrate", random="clust",
# verbose=1000)


## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a cubic B-spline with two knots at 1 and 5 year, with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)
## and with a time-dependent effect for age (agecr) and sex (IsexH).

# Mod_bs3_2mixnph <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+depindex+IsexH + nph(agecr+IsexH), data=simdatn1,
# base="exp.bs", degree=3, knots=c(1,5), expected="popmrate",
# random="clust", verbose=1000)
```

---

adjsurv	*Computation of direct adjusted survival estimates based on a mexhaz model*

Description

Function for computing direct adjusted survival estimates from a model fitted with the mexhaz. It can be used to obtain direct adjusted survival estimates for one or two populations. In the latter case, survival difference estimates are also computed. Corresponding variance estimates are based on the Delta Method (based on the assumption of multivariate normality of the model parameter estimates). When the model includes a random effect, three types of predictions can be made: (i) marginal predictions (obtained by integration over the random effect distribution), (ii) cluster-specific posterior predictions for an existing cluster, or (iii) conditional predictions for a given quantile of the random effect distribution (by default, for the median value, that is, 0).

Usage

```
adjsurv(object, time.pts, data, data.0 = NULL, weights = NULL,
marginal = TRUE, quant.rdm = 0.5, cluster = NULL, quant.rdm.0 = 0.5,
cluster.0 = NULL, level = 0.95, dataset = NULL)
```

Arguments

object   an object of class mexhaz, corresponding to a hazard-based regression model fitted with the mexhaz function.

time.pts	a vector of numerical values representing the time points at which predictions are requested. Time values greater than the maximum follow-up time on which the model estimation was based are discarded.
data	a data.frame containing the values of the covariates of the population for which direct adjusted estimates are to be calculated.
data.0	an optional data.frame containing the values of the covariates of a second population for which direct adjusted estimates can also be calculated (and compared with those of the first population). The default value is set to NULL.
weights	optional argument specifying the weights to be associated with each row of data (and data.0). the default value is set to NULL which corresponds to attributing to each row of the dataset(s) a weight equal to one over the total number of rows.
marginal	logical value controlling the type of predictions returned by the function when the model includes a random intercept. When TRUE, marginal predictions are computed. The marginal survival is obtained by integrating the predicted survival over the distribution of the random effect. The marginal hazard rate is obtained as the opposite of the marginal time derivative of the survival divided by the marginal survival. When FALSE (default value), cluster-specific posterior predictions or conditional predictions are calculated depending on the value of the cluster argument.
quant.rdm	numerical value (between 0 and 1) specifying the quantile of the random effect distribution that should be used when requesting conditional predictions. The default value is set to 0.5 (corresponding to the median, that is a value of the random effect of 0). This argument is ignored if the model is a fixed effect model, when the marginal argument is set to TRUE, or the cluster argument is not NULL.
cluster	a single value corresponding to the name of the cluster for which posterior predictions should be calculated. These predictions are obtained by integrating over the cluster-specific posterior distribution of the random effect and thus require the original dataset. The dataset can either be provided as part of the mexhaz object given as argument or by specifying the name of the dataset in the dataset argument (see below). The cluster argument is not used if the model is a fixed effect model. The default value is NULL: this corresponds to marginal predictions (if marginal is set to TRUE, the preferred option), or to conditional predictions for a given quantile (by default, the median) of the distribution of the random effect (if marginal is set to FALSE).
quant.rdm.0	random effect distribution quantile value to be used with data.0 (see argument quant.rdm for details).
cluster.0	cluster value to be used with data.0 (see argument cluster for details).
level	a number in (0,1) specifying the level of confidence for computing the confidence intervals of the hazard and the survival. By default, level=0.95.
dataset	original dataset used to fit the mexhaz object given as argument to the function. This argument is only necessary if cluster-specific posterior predictions are requested (and if the dataset is not already provided in the mexhaz object). The default value is set to NULL.

Value

An object of class resMexhaz that can be used by the function plot.resMexhaz to produce graphics of the direct adjusted survival curve. It contains the following elements:

results   a data.frame consisting of: the time points at which the direct adjusted survival values have been calculated; the direct ajusted survival values with their confidence limits for population data; the direct ajusted survival values with their confidence limits for population data.0; the direct adjusted survival difference estimates with their confidence limits.

type   type of results returned by the function. The value is used by plot.resMexhaz and lines.resMexhaz, and set to "as" (adjusted survival).

multiobs   value used by plot.resMexhaz and lines.resMexhaz, and set to FALSE (computation of the adjusted survival at several time points for one vector of covariates).

ci.method   method used to compute confidence limits. Currently set to "delta" as only the Delta Method is implemented.

level   level of confidence used to compute confidence limits.

Author(s)

Hadrien Charvat, Aurelien Belot

References

Charvat H, Remontet L, Bossard N, Roche L, Dejardin O, Rachet B, Launoy G, Belot A; CENSUR Working Survival Group. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Stat Med 2016;35:3066-3084 (doi: 10.1002/sim.6881)

Skrondal A, Rabe-Hesketh S. Prediction in multilevel generalized linear models. J R Stat Soc A Stat Soc 2009;172(3):659-687 (doi: 10.1111/j.1467-985X.2009.00587.x).

See Also

plot.resMexhaz, lines.resMexhaz

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs1_2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=1, knots=c(1,5), verbose=0)

## Direct adjusted survival for the simdatn1 population
```

```
DAS_Modbs1_2 <- adjsurv(Mod_bs1_2, time.pts=seq(1,10),
data=simdatn1)
```

---

fixef   *Method for extracting fixed effects*

Description

This is a generic function.

Usage

```
fixef(x, ...)
```

Arguments

x   a fitted object from which fixed effects can be extracted.

...   may be required by some methods using this generic function.

Value

see the documentation for the specific class.

See Also

[ranef](#)

Examples

```
## See the specific class documentation.
```

---

fixef.mexhaz   *Method for extracting fixed effects from a mexhaz object*

Description

Display a vector containing the fixed effects of a mexhaz model.

Usage

```
## S3 method for class 'mexhaz'
fixef(x, ...)
```

Arguments

x   an object of class mexhaz.

...   not used.

Value

A vector containing the fixed effect estimates.

See Also

mexhaz

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib_mix <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull",
expected="popmrate", verbose=0, random="clust")

fixef(Mod_weib_mix)
```

---

lines.predMexhaz   *Lines method for a predMexhaz object*

Description

Function for adding to an already existing graphical window the predicted (excess) hazard or (net) survival based on a predMexhaz object.

Usage

```
## S3 method for class 'predMexhaz'
lines(x, which = c("surv", "hazard"), conf.int =
TRUE, lty.pe = "solid", lty.ci = "dashed", ...)
```

Arguments

x   an object of class predMexhaz, corresponding to (excess) hazard and (net) survival predictions based on a survival model fitted with the mexhaz function. Predictions can be obtained for multiple times for one vector of covariables ("multitime") or for several vectors of covariables at one time point ("multiobs"). The lines() function only applies to the "multitime" type of predictions.

which	type of curve to be plotted. Selection can be made between "surv" (default value) for the (net) survival curve and "hazard" for the (excess) hazard.
conf.int	logical values allowing the user to decide whether to plot the confidence limits of the survival (or hazard).
lty.pe	type of line used for drawing the hazard/survival estimate.
lty.ci	type of line used for drawing the confidence limits.
...	additional parameters that are directly passed to the lines function. These parameters, if not already used internally by the lines.predMexhaz function, will apply simultaneously to the point estimate and confidence limit curves.

See Also

[predict.mexhaz](#), [plot.predMexhaz](#)

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs2_2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=2, knots=c(1,5), verbose=0)

## Prediction at several time points for one vector of covariates
Pred_Modbs2_2A <- predict(Mod_bs2_2, time.pts=seq(0,10,by=0.1),
data.val=data.frame(agecr=0,depindex=0.5,IsexH=1))

plot(Pred_Modbs2_2A, which="hazard", col="red")
lines(Pred_Modbs2_2A, which="hazard", conf.int=FALSE)
```

---

lines.resMexhaz	*Lines method for a resMexhaz object*

Description

Function for adding to an already existing graphical window various quantities calculated from a mexhaz model.

Usage

```
## S3 method for class 'resMexhaz'
lines(x, conf.int = TRUE, lty.pe = "solid", lty.ci =
"blank", col.ci = "blue", alpha.col.ci = 0.25, ...)
```

Arguments

x	an object of class resMexhaz, corresponding to various predictions based on a survival model fitted with the mexhaz function. Predictions can be obtained for multiple times for one vector of covariables ("multitime") or for several vectors of covariables at one time point ("multiobs"). The plot() function only applies to the "multitime" type of predictions.
conf.int	logical value allowing the user to decide whether to plot the confidence limits of the survival (or hazard).
lty.pe	type of line used for drawing the hazard/survival estimate.
lty.ci	type of line used for drawing the confidence limits.
col.ci	color used to fill in the polygon defined by the confidence limits.
alpha.col.ci	parameter used internally by the rgb() function to set the color transparency.
...	additional parameters that are directly passed to the lines function. These parameters will apply simultaneously to the point estimate and confidence limit curves.

See Also

[adjsurv](#), [riskfunc](#), [plot.resMexhaz](#)

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=2, knots=c(1,5), verbose=0)

## Relative risk (ratio of cumulative incidence curves) for men versus
## women for two different values of age

RR_Modbs2a <- riskfunc(Mod_bs2, time.pts=seq(0,10, by=0.1),
data=data.frame(agecr=0.1, IsexH=1, depindex=0),
data.0=data.frame(agecr=0.1, IsexH=0, depindex=0), conf.int="delta",
type="rr")

RR_Modbs2b <- riskfunc(Mod_bs2, time.pts=seq(0,10, by=0.1),
data=data.frame(agecr=-0.1, IsexH=1, depindex=0),
data.0=data.frame(agecr=-0.1, IsexH=0, depindex=0), conf.int="delta",
type="rr")


plot(RR_Modbs2a)
lines(RR_Modbs2a, col.ci="green")
```

```
abline(h=1, lty="dashed")
```

| mexhaz | *mexhaz function* |

Description

Fit an (excess) hazard regression model using different shapes for the baseline hazard (Weibull, piecewise constant, exponential of a B-spline of degree 1 to 3, exponential of a restricted cubic spline), with the possibility to include time-dependent effects of variable(s) and a random effect defined at the cluster level. The function accepts right-censored and counting process input styles for the follow-up time. The latter allows the modelling of survival data with delayed entries. The time-dependent effect of a covariable is modelled by adding interaction terms between the covariable and a function of time of the same class as the one used for the baseline hazard (in particular, with the same knots for piecewise constant hazards; and with the same degree and the same knots for B-spline or restricted cubic spline functions). The random effect is assumed to be normally distributed with mean 0 and standard deviation sigma. The optimisation process uses adaptive Gaussian quadrature to calculate the cluster-specific marginal likelihoods. The logarithm of the full marginal likelihood, defined as the sum of the logarithms of the cluster-specific marginal likelihoods, is then maximised using an optimisation routine such as nlm (default) or optim.

Usage

```
mexhaz(formula, data, expected = NULL, base = c("weibull",
"exp.bs", "exp.ns", "pw.cst"), degree = 3, knots = NULL,
bound = NULL, n.gleg = 20, init = NULL, random = NULL,
n.aghq = 10, fnoptim = c("nlm", "optim"), verbose = 0,
method = "Nelder-Mead", iterlim = 10000, numHess = FALSE,
print.level = 1, exactGradHess = TRUE, gradtol =
ifelse(exactGradHess, 1e-8, 1e-6), testInit = TRUE,
keep.data = FALSE, ...)
```

Arguments

formula
: a formula object, with the response on the left of the ~ operator, and the linear predictor on the right. The response must be of the form Surv(time, event) for right censored data or Surv(time, time2, event) for counting process style data. The linear predictor accepts a special instruction nph() for specifying variables for which a time-dependent effect should be modelled (if several variables are modelled with time-dependent effects, separate these variables inside the nph() instruction with a + sign).

  In case time takes the value 0 for some observations when data are entered in the right censored style, it is assumed that these observations refer to events (or censoring) that occurred on the first day of follow-up. Consequently, a value of 1/730.5 (half a day) is substituted in order to make computations possible. However, it should be stressed that this is just a convention and that it does not

	make much sense if the time scale is not expressed in years. We therefore advise the analyst to deal with 0 time values during the dataset preparation stage.
data	a `data.frame` containing the variables referred to in the `formula`, as well as in the `expected` and `random` arguments if these arguments are used.
expected	name of the variable (must be given in quotes) representing the population (i.e., expected) hazard. By default, `expected=NULL`, which means that the function estimates the overall hazard (and not the excess hazard).
base	functional form that should be used to model the baseline hazard. Selection can be made between the following options: `"weibull"` for a Weibull hazard, `"exp.bs"` for a hazard described by the exponential of a B-spline (only B-splines of degree 1, 2 or 3 are accepted), `"exp.ns"` for a hazard described by the exponential of a restricted cubic spline (also called 'natural spline'), `"pw.cst"` for a piecewise constant hazard. By default, `base="weibull"`.

For the Weibull hazard model, the cumulative hazard is given by the following relationship:

H(t,x,z) = lambda*exp(x'b)*t^(rho*exp(z'g))

where lambda and rho are the parameters of the Weibull baseline hazard, x represent variables with proportional effect (with corresponding regression coefficients 'b') and z represent variables with time-dependent effects (with corresponding regression coefficients 'g'). The `mexhaz()` function does not estimate directly lambda and rho but their logarithms (in the output of the function, these are named respectively 'logLambda' and 'logRho').

For the spline-based hazards, it should be noted that the B-spline and restricted cubic spline bases created internally in the `mexhaz()` function are identical to those obtained by the use of, respectively, the `bs()` and `ns()` functions from the `splines` package.

degree	if `base="exp.bs"`, degree represents the degree of the B-spline used. Only integer values between 1 and 3 are accepted, and 3 is the default.
knots	if `base="exp.bs"` or `"exp.ns"`, knots is the vector of interior knots of the spline. If `base="pw.cst"`, knots is the vector defining the endpoints of the time intervals on which the hazard is assumed to be constant. By default, `knots=NULL` (that is, it produces a B-spline with no interior knots if `base="exp.bs"`, a linear B-spline with no interior knots if `base="exp.ns"`, or a constant hazard over the whole follow-up period if `base="pw.cst"`).
bound	a vector of two numerical values corresponding to the boundary knots of the spline functions. If `base="exp.bs"` or `base="exp.ns"`, computation of the B-spline basis requires that boundary knots be given. The bound argument allows the user to specify these boundary knots. If `base="exp.bs"`, the interval defined by the boundary knots must at least include the interval c(0,max(time)) (otherwise, there could be problems with ill-conditioned bases). If `base="exp.ns"`, the boundary knots correspond to the knots beyond which the spline is constrained to be linear (in that case, the boundary knots can be values contained in c(0,max(time))). By default, the boundary knots are set to c(0,max(time)).
n.gleg	if `base="exp.bs"` and degree is equal to 2 or 3, or if `base="exp.ns"`, the cumulative hazard is computed via Gauss-Legendre quadrature and `n.gleg` is the number of quadrature nodes to be used to compute the cumulative hazard. By default, `n.gleg=20`.

init   vector of initial values. By default init=NULL and the initial values are internally set to the following values:

for the baseline hazard:

if exactGradHess=TRUE (except for the excess hazard random effect model for which this argument is ignored), the intercept is set to 0.5*log(Number of events/Person-years of observation) and all other parameters set to 0. In case of failed convergence, and if the testInit argument is set to TRUE (default), several trials are run with an adaptation of the value of the intercept.

if exactGradHess=FALSE, the following values are used:

- if base="weibull", the logarithm of the scale and shape parameters is set to 0.1;

- if base="exp.bs", the parameters of the B-spline are all set to -1;

- if base="exp.ns", the parameters of the restricted cubic spline are all set to -1;

- if base="pw.cst", the logarithm of the piecewise-constant hazards are set to -1;

the parameters describing the effects of the covariables are all set to 0;

the parameter representing the standard deviation of the random effect is set to 0.1. In case of failed convergence, if exactGradHess=TRUE and if testInit=TRUE, several trials are run with an adaptation of the value of this random effect parameter.

random   name of the variable to be entered as a random effect (must be given between quotes), representing the cluster membership. By default, random=NULL which means that the function fits a fixed effects model.

n.aghq   number of quadrature points used for estimating the cluster-specific marginal likelihoods by adaptive Gauss-Hermite quadrature. By default, n.aghq=10.

fnoptim   name of the R optimisation procedure used to maximise the likelihood. Selection can be made between "nlm" (by default) and "optim". Note: if exactGradHess=TRUE, this argument will be ignored (fnoptim will be set automatically to "nlm").

verbose   integer parameter representing the frequency at which the current state of the optimisation process is displayed. Internally, an 'evaluation' is defined as an estimation of the log-likelihood for a given vector of parameters. This means that the number of evaluations is increased each time the optimisation procedure updates the value of any of the parameters to be estimated. If verbose=n (with n an integer), the function will display the current values of the parameters, the log-likelihood and the time elapsed every n evaluations. If verbose=0 (default), nothing is displayed.

method   if fnoptim="optim", method represents the optimisation method to be used by optim. By default, method="Nelder-Mead". This parameter is not used if fnoptim="nlm".

iterlim   if fnoptim="nlm", iterlim represents the maximum number of iterations before the nlm optimisation procedure is terminated. By default, iterlim is set to 10000. This parameter is not used if fnoptim="optim" (in this case, the maximum number of iterations must be given as part of a list of control parameters via the control argument: see the help page of optim for further details).

numHess	logical value allowing the user to choose between the Hessian returned by the optimization algorithm (default) or the Hessian estimated by the `hessian` function from the `numDeriv` package. The latter might be more accurate but its estimation is more time-consuming. We suggest to use the default Hessian estimation procedure during model building and estimate the `numDeriv`-based Hessian only on the final model. Note: if exactGradHess=TRUE, this argument is ignored.
print.level	this argument is only used if fnoptim="nlm". It determines the level of printing during the optimisation process. The default value (for the `mexhaz` function) is set to '1' which means that details on the initial and final step of the optimisation procedure are printed (see the help page of `nlm` for further details).
exactGradHess	logical value allowing the user to decide whether maximisation of the likelihood should be based on the analytic gradient and Hessian computed internally (default, corresponding to exactGradHess=TRUE). In that case, optimisation is performed with the `nlm` function. Note: even if set to TRUE, this argument is ignored when the user wants to fit an excess hazard model including a random effect because in that case, there is no simple way to obtain the analytic gradient and Hessian.
gradtol	this argument is only used if fnoptim="nlm". It corresponds to the tolerance at which the scaled gradient is considered close enough to zero to terminate the algorithm. The default value depends on the value of the argument exactGradHess.
testInit	this argument is used only when exactGradHess=TRUE and when the model is not an excess hazard random effect model. It instructs the `mexhaz` function to try several vectors of initial values in case optimization was not successful with the default (or user-defined) initial values. Because optimization based on the analytical gradient and Hessian is usually fast, this simple and empirical procedure proves useful to increase the probability of convergence in cases when it is difficult to specify appropriate initial values. Only the initial values for the intercept and for the parameter corresponding to the random effect are modified.
keep.data	logical argument determining whether the dataset should be kept in the object returned by the function: this can be useful in certain contexts (e.g., to calculate cluster-specific posterior predictions from a random intercept model) but might create unnecessarily voluminous objects. The default value is set to FALSE.
...	represents additional parameters directly passed to `nlm` or `optim` to control the optimisation process.

Value

An object of class `mexhaz` containing the following elements:

dataset	name of the dataset used to fit the model.
call	function call on which the model is based.
formula	formula part of the call.
expected	name of the variable corresponding to the expected hazard (takes the value NA for standard, i.e., 'non-excess' hazard models).
xlevels	information concerning the levels of the categorical variables used in the model (used by `predict.mexhaz`).

n.obs.tot	total number of observations in the dataset.
n.obs	number of observations used to fit the model (after exclusion of missing values).
n.events	number of events (after exclusion of missing values).
n.clust	number of clusters.
n.time.0	number of observations for which the observed follow-up time was equal to 0 (only for right censored type data).
base	function used to model the baseline hazard.
max.time	maximal observed time in the dataset.
boundary.knots	vector of boundary values used to define the B-spline (or natural spline) bases.
degree	degree of the B-spline used to model the logarithm of the baseline hazard.
knots	vector of interior knots used to define the B-spline (or natural spline) bases.
names.ph	names of the covariables with a proportional effect.
random	name of the variable defining cluster membership (set to NA in the case of a purely fixed effects model).
init	a vector containing the initial values of the parameters.
coefficients	a vector containing the parameter estimates.
std.errors	a vector containing the standard errors of the parameter estimates.
vcov	the variance-covariance matrix of the estimated parameters.
gradient	the gradient of the log-likelihood function evaluated at the estimated parameters.
hessian	the Hessian of the log-likelihood function evaluated at the estimated parameters.
mu.hat	a data.frame containing the estimated cluster-specific random effects (shrinkage estimators).
var.mu.hat	the covariance matrix of the cluster-specific shrinkage estimators.
vcov.fix.mu.hat	
	a matrix containing the covariances between the fixed effect and the cluster-specific shrinkage estimators. More specifically, the i-th line of the matrix represents the covariances between the shrinkage estimator of the i-th cluster and the fixed effect estimates. This matrix is used by the function predict.mexhaz to make cluster-specific predictions.
data	original dataset used to fit the model (if keep.data was set to TRUE).
n.par	number of estimated parameters.
n.gleg	number of Gauss-Legendre quadrature points used to calculate the cumulative (excess) hazard (only relevant if a B-spline of degree 2 or 3 or a cubic restricted spline was used to model the logarithm of the baseline hazard).
n.aghq	number of adaptive Gauss-Hermite quadrature points used to calculate the cluster-specific marginal likelihoods (only relevant if a multi-level model is fitted).
fnoptim	name of the R optimisation procedure used to maximise the likelihood.
method	optimisation method used by optim.
code	code (integer) indicating the status of the optimisation process (this code has a different meaning for nlm and for optim: see their respective help page for details).

loglik	value of the log-likelihood at the end of the optimisation procedure.
iter	number of iterations used in the optimisation process.
eval	number of evaluations used in the optimisation process.
time.elapsed	total time required to reach convergence.

Author(s)

Hadrien Charvat, Aurelien Belot

References

Charvat H, Remontet L, Bossard N, Roche L, Dejardin O, Rachet B, Launoy G, Belot A; CENSUR Working Survival Group. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Stat Med 2016;35(18):3066-3084 (doi: 10.1002/sim.6881)

Charvat H, Belot A. An R package for fitting flexible hazard-based regression models for overall and excess mortality with a random effect. J Stat Softw 2021;98(14):1-36 (doi: 10.18637/jss.v098.i14)

See Also

fixef.mexhaz, predict.mexhaz, print.mexhaz, ranef.mexhaz, summary.mexhaz, update.mexhaz, vcov.mexhaz

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib_mix <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull",
expected="popmrate", verbose=0, random="clust")


## More complex examples (not run)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a cubic B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

# Mod_bs3_2mix_nph <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+depindex+IsexH+nph(agecr), data=simdatn1,
# base="exp.bs", degree=3, knots=c(1,5), expected="popmrate",
# random="clust", verbose=1000)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a restricted cubic spline with two knots at the
## tertiles of the distribution of follow-up times for individuals who
```

```
## presented the event and with effects of age (agecr) and sex (IsexH)

# Mod_ns3_2mix_nph <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+nph(agecr), data=simdatn1, base="exp.ns", degree=3,
# knots=quantile(simdatn1[simdatn1$vstat==1,]$timesurv, probs=c(1:2/3)),
# expected="popmrate", random="clust", verbose=1000)
```

---

plot.predMexhaz   *Plot method for a predMexhaz object*

Description

Function for plotting the predicted (excess) hazard or (net) survival based on a predMexhaz object.

Usage

```
## S3 method for class 'predMexhaz'
plot(x, which = c("surv", "hazard"), conf.int =
TRUE, lty.pe = "solid", lty.ci = "dashed", ...)
```

Arguments

x   an object of class predMexhaz, corresponding to (excess) hazard and (net) sur-
             vival predictions based on a survival model fitted with the mexhaz function.
             Predictions can be obtained for multiple times for one vector of covariables
             ("multitime") or for several vectors of covariables at one time point ("multiobs").
             The plot() function only applies to the "multitime" type of predictions.

which   type of curve to be plotted. Selection can be made between "surv" (default
             value) for the (net) survival curve and "hazard" for the (excess) hazard.

conf.int   logical value allowing the user to decide whether to plot the confidence limits of
             the survival (or hazard).

lty.pe   type of line used for drawing the hazard/survival estimate.

lty.ci   type of line used for drawing the confidence limits.

...   additional parameters that are directly passed to the plot function. These pa-
             rameters will apply simultaneously to the point estimate and confidence limit
             curves.

See Also

[predict.mexhaz](), [points.predMexhaz](), [lines.predMexhaz]()

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs2_2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=2, knots=c(1,5), verbose=0)

## Prediction at several time points for one vector of covariates
Pred_Modbs2_2A <- predict(Mod_bs2_2, time.pts=seq(0,10,by=0.1),
data.val=data.frame(agecr=0,depindex=0.5,IsexH=1))

plot(Pred_Modbs2_2A, which="hazard")
```

| plot.resMexhaz | *Plot method for a resMexhaz object* |

Description

Function for plotting various quantities calculated from a mexhaz model.

Usage

```
## S3 method for class 'resMexhaz'
plot(x, conf.int = TRUE, lty.pe = "solid", lty.ci =
"blank", col.ci = "blue", alpha.col.ci = 0.25, ylim = NULL, ...)
```

Arguments

x	an object of class resMexhaz, corresponding to various predictions based on a survival model fitted with the mexhaz function. Predictions can be obtained for multiple times for one vector of covariables ("multitime") or for several vectors of covariables at one time point ("multiobs"). The plot() function only applies to the "multitime" type of predictions.
conf.int	logical value allowing the user to decide whether to plot the confidence limits of the survival (or hazard).
lty.pe	type of line used for drawing the hazard/survival estimate.
lty.ci	type of line used for drawing the confidence limits.
col.ci	color used to fill in the polygon defined by the confidence limits.
alpha.col.ci	parameter used internally by the rgb() function to set the color transparency.
ylim	when set to NULL (default value), the range of the y-axis is based on the range of the confidence limits of the quantity to be plotted.

... additional parameters that are directly passed to the plot function. These parameters will apply simultaneously to the point estimate and confidence limit curves.

See Also

[adjsurv](), [riskfunc](), [lines.resMexhaz]()

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=2, knots=c(1,5), verbose=0)

## Relative risk (ratio of cumulative incidence curves) for men versus
## women

RR_Modbs2 <- riskfunc(Mod_bs2, time.pts=seq(0,10, by=0.1),
data=data.frame(agecr=0.1, IsexH=1, depindex=0),
data.0=data.frame(agecr=0.1, IsexH=0, depindex=0), conf.int="delta",
type="rr")

plot(RR_Modbs2)
abline(h=1, lty="dashed")
```

---

points.predMexhaz   *Points method for a predMexhaz object*

Description

Function for adding to an already existing graphical window the predicted (excess) hazard or (net) survival based on a predMexhaz object.

Usage

```
## S3 method for class 'predMexhaz'
points(x, which = c("surv", "hazard"), conf.int =
TRUE, lty.pe = "solid", lty.ci = "dashed", ...)
```

Arguments

x	an object of class predMexhaz, corresponding to (excess) hazard and (net) survival predictions based on a survival model fitted with the mexhaz function. Predictions can be obtained for multiple times for one vector of covariables (″multitime″) or for several vectors of covariables at one time point (″multiobs″). The points() function only applies to the ″multitime″ type of predictions.
which	type of curve to be plotted. Selection can be made between ″surv″ (default value) for the (net) survival curve and ″hazard″ for the (excess) hazard.
conf.int	logical values allowing the user to decide whether to plot the confidence limits of the survival (or hazard).
lty.pe	type of line used for drawing the hazard/survival estimate (used when 'type="l"').
lty.ci	type of line used for drawing the confidence limits (used when 'type="l"').
...	additional parameters that are directly passed to the points function. These parameters, if not already used internally by the points.predMexhaz function, will apply simultaneously to the point estimate and confidence limit curves.

See Also

[predict.mexhaz](#), [plot.predMexhaz](#)

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs2_2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=2, knots=c(1,5), verbose=0)

## Prediction at several time points for one vector of covariates
Pred_Modbs2_2A <- predict(Mod_bs2_2, time.pts=seq(0,10,by=0.1),
data.val=data.frame(agecr=0,depindex=0.5,IsexH=1))

plot(Pred_Modbs2_2A, which="hazard", col="red")
points(Pred_Modbs2_2A, which="hazard", type="l", conf.int=FALSE)
```

---

predict.mexhaz   *Predictions based on a mexhaz model*

Description

Function for predicting the (excess) hazard and the corresponding (net) survival from a model fitted with the `mexhaz` function for a particular vector of covariates. If the survival model was fitted with an expected hazard (excess hazard model), the estimates obtained are excess hazard and net survival estimates. Corresponding variance estimates are based on the Delta Method or Monte Carlo simulation (based on the assumption of multivariate normality of the model parameter estimates). This function allows the computation of the hazard and the survival at one time point for several vectors of covariates or for one vector of covariates at several time points. When the model includes a random effect, three types of predictions can be made: (i) marginal predictions (obtained by integration over the random effect distribution), (ii) cluster-specific posterior predictions for an existing cluster, or (iii) conditional predictions for a given quantile of the random effect distribution (by default, for the median value, that is, 0).

Usage

```
## S3 method for class 'mexhaz'
predict(object, time.pts, data.val = data.frame(.NotUsed = NA),
marginal = FALSE, quant.rdm = 0.5, cluster = NULL, conf.int = c("delta", "simul", "none"),
level = 0.95, delta.type.h = c("log", "plain"), delta.type.s = c("log-log",
"log", "plain"), nb.sim = 10000, keep.sim = FALSE, include.gradient =
FALSE, dataset = NULL, ...)
```

Arguments

`object`	an object of class `mexhaz`, corresponding to a hazard-based regression model fitted with the `mexhaz` function.
`time.pts`	a vector of numerical values representing the time points at which predictions are requested. Time values greater than the maximum follow-up time on which the model estimation was based are discarded.
`data.val`	a `data.frame` containing the values of the covariates at which predictions should be calculated.
`marginal`	logical value controlling the type of predictions returned by the function when the model includes a random intercept. When `TRUE`, marginal predictions are computed. The marginal survival is obtained by integrating the predicted survival over the distribution of the random effect. The marginal hazard rate is obtained as the opposite of the marginal time derivative of the survival divided by the marginal survival. When `FALSE` (default value), cluster-specific posterior predictions or conditional predictions are calculated depending on the value of the `cluster` argument.
`quant.rdm`	numerical value (between 0 and 1) specifying the quantile of the random effect distribution that should be used when requesting conditional predictions. The default value is set to 0.5 (corresponding to the median, that is a value of the random effect of 0). This argument is ignored if the model is a fixed effect model, when the `marginal` argument is set to `TRUE`, or the `cluster` argument is not `NULL`.
`cluster`	a single value corresponding to the name of the cluster for which posterior predictions should be calculated. These predictions are obtained by integrating over

	the cluster-specific posterior distribution of the random effect and thus require the original dataset. The dataset can either be provided as part of the mexhaz object given as argument or by specifying the name of the dataset in the dataset argument (see below). The cluster argument is not used if the model is a fixed effect model. The default value is NULL: this corresponds to marginal predictions (if marginal is set to TRUE, the preferred option), or to conditional predictions for a given quantile (by default, the median) of the distribution of the random effect (if marginal is set to FALSE).
conf.int	method to be used to compute confidence limits. Selection can be made between the following options: "delta" for the Delta Method (default value); "simul" for Monte Carlo simulations (can be time-consuming, especially for models using B-splines for the logarithm of the baseline hazard); "none" indicates absence of confidence limits estimation.
level	a number in (0,1) specifying the level of confidence for computing the confidence intervals of the hazard and the survival. The default value is set to 0.95.
delta.type.h	type of confidence limits for the hazard when using the Delta Method. With the default value ("log"), the confidence limits are based on a Wald-type confidence interval for the logarithm of the hazard, otherwise they are based directly on a Wald-type CI for the hazard.
delta.type.s	type of confidence limits for the survival when using the Delta Method. With the default value ("log-log"), the confidence limits are based on a Wald-type confidence interval for the logarithm of the cumulative hazard; when the argument is set to "log", they are based on a Wald-type CI for the cumulative hazard; otherwise they are based directly on a Wald-type CI for the survival.
nb.sim	integer value representing the number of simulations used to estimate the confidence limits for the (excess) hazard and the (net) survival. This argument is used only if conf.int is set to "simul".
keep.sim	logical value determining if the simulated hazard and survival values should be returned (only used when conf.int is set to "simul"). These simulated values can be used by the riskfunc function to compute simulation-based confidence intervals for hazard ratios / risk ratios. The default value is set to FALSE.
include.gradient	
	logical value allowing the function to return the components of the gradient of the logarithm of the hazard and of the logarithm of the cumulative hazard for each prediction. This argument is used only if conf.int is set to "delta". The default value is FALSE.
dataset	original dataset used to fit the mexhaz object given as argument to the function. This argument is only necessary if cluster-specific posterior predictions are requested (and if the dataset is not already provided in the mexhaz object). The default value is set to NULL.
...	for potential additional parameters.

Value

An object of class predMexhaz that can be used by the function plot.predMexhaz to produce graphics of the (excess) hazard and the (net) survival. It contains the following elements:

call	the mexhaz function call on which the model is based.
results	a data.frame consisting of: the time points at which the (excess) hazard and the (net) survival have been calculated; the values of the covariates used to estimate the (excess) hazard and the (net) survival; the name of the cluster (when cluster-specific posterior predictions are requested), or the quantile and its value (when conditional predictions are requested); the (excess) hazard values with their confidence limits; and the (net) survival values with their confidence limits.
variances	a data.frame consisting of two columns: the variance of the logarithm of the (excess) hazard and the variance of the (excess) cumulative hazard for each time points or each vector of covariates. The object variances is produced only when conf.int is set to "delta".
grad.loghaz	a data.frame consisting of the components of the gradient of the logarithm of the (excess) hazard for each time points or each vector of covariates. The number of columns corresponds to the number of model parameters. This object is produced only when conf.int is set to "delta" and include.gradient to TRUE.
grad.logcum	a data.frame consisting of the components of the gradient of the logarithm of the (excess) cumulative hazard for each time points or each vector of covariates. The number of columns corresponds to the number of model parameters. This object is produced only when conf.int is set to "delta" and include.gradient to TRUE.
vcov	a matrix corresponding to the covariance matrix used to compute the confidence intervals.
type	this item can take the value multitime (computation of the hazard and the survival at at several time points for one vector of covariates) or multiobs (computation of the hazard and the survival at at one time point for several vectors of covariates). It is used by plot.predMexhaz and points.predMexhaz.
type.me	the type of predictions produced in case of a model including a random intercept. Can take the values marginal (marginal predictions), posterior (cluster-specific posterior predictions), or conditional (conditional prediction for a given quantile of the distribution of the random intercept).
ci.method	the method used to compute confidence limits.
level	level of confidence used to compute confidence limits.
delta.type	type of confidence limits for the hazard and the survival when using the Delta Method.
nb.sim	number of simulations used to estimate the confidence limits when ci.method is set to "simul".
sim.haz	matrix containing the simulated hazards (each column representing a simulated vector of values); only returned when keep.sim is set to TRUE.
sim.surv	matrix containing the simulated survival probabilities (each column representing a simulated vector of values); only returned when keep.sim is set to TRUE.

Author(s)

Hadrien Charvat, Aurelien Belot

References

Charvat H, Remontet L, Bossard N, Roche L, Dejardin O, Rachet B, Launoy G, Belot A; CENSUR Working Survival Group. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Stat Med 2016;35:3066-3084 (doi: 10.1002/sim.6881).

Skrondal A, Rabe-Hesketh S. Prediction in multilevel generalized linear models. J R Stat Soc A Stat Soc 2009;172(3):659-687 (doi: 10.1111/j.1467-985X.2009.00587.x).

See Also

`print.predMexhaz`, `plot.predMexhaz`, `points.predMexhaz`, `lines.predMexhaz`

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and
## with effects of age (agecr), deprivation index (depindex) and sex
## (IsexH)

Mod_bs1_2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=1, knots=c(1,5), verbose=0)

## Prediction at several time points for one vector of covariates
Pred_Modbs1_2A <- predict(Mod_bs1_2, time.pts=seq(0.1,10,by=0.1),
data.val=data.frame(agecr=0,depindex=0.5,IsexH=1))

## Prediction for several vectors of covariates at one time point
Pred_Modbs1_2B <- predict(Mod_bs1_2, time.pts=10,
data.val=data.frame(agecr=c(-0.2,-0.1,0), depindex=c(0.5,0.5,0.5),
IsexH=c(1,1,1)))

## Prediction for all individuals of the study population
## at one time point

Pred_Modbs1_2C <- predict(Mod_bs1_2, time.pts=10,
data.val=simdatn1)


# Example of cluster-specific posterior prediction (not run)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a cubic B-spline with two knots at 1 and 5 year

# Mod_bs3_2mix <- mexhaz(formula=Surv(time=timesurv,
# event=vstat)~agecr+IsexH, data=simdatn1, base="exp.bs", degree=3,
# knots=c(1,5), expected="popmrate", random="clust", verbose=1000)
```

```
## Posterior predictions at several time points for an individual
## in cluster 15 with a specific vector of covariates
# Pred_Modbs3_2A <- predict(Mod_bs3_2mix,
# time.pts=seq(0.1,10,by=0.1), data.val=data.frame(agecr=0.2, IsexH=1),
# cluster=15)
```

---

print.mexhaz   *Print method for a mexhaz object*

Description

Display the model call as well as the values of the estimated model parameters.

Usage

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

Arguments

x   an object of class mexhaz.

...   represents additional parameters directly passed to print.

See Also

mexhaz, summary.mexhaz

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib_mix <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull",
expected="popmrate", verbose=0, random="clust")

print(Mod_weib_mix)
```

print.predMexhaz   *Print method for a predMexhaz object*

Description

Display the first lines of the `data.frame` containing the predictions provided by the `predMexhaz` function.

Usage

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

Arguments

x   an object of class `predMexhaz`.

...   represents additional parameters directly passed to `print`.

See Also

predict.mexhaz

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs1_2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=1, knots=c(1,5), verbose=0)

## Prediction at several time points for one vector of covariates
Pred_Modbs1_2A <- predict(Mod_bs1_2, time.pts=seq(0.1,10,by=0.1),
data.val=data.frame(agecr=0,depindex=0.5,IsexH=1), conf.int="delta")

print(Pred_Modbs1_2A)
```

---

print.summary.mexhaz　　*Printing method for a summary.mexhaz object*

Description

Display the model call, the values of the estimated model parameters, as well as the corresponding hazard ratios (only for proportional effects).

Usage

```
## S3 method for class 'summary.mexhaz'
print(x, ...)
```

Arguments

x　　　　　　　an object of class summary.mexhaz.

...　　　　　　represents additional parameters directly passed to print.

See Also

[mexhaz](#), [print.mexhaz](#), [summary.mexhaz](#)

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull", verbose=0)

summary(Mod_weib)
```

---

ranef　　　　　　　　*Method for extracting random effects*

Description

This is a generic function.

Usage

```
ranef(x, ...)
```

Arguments

x	a fitted object from which random effects can be extracted.
...	might be used by some methods using this generic function.

Value

see the documentation for the specific class.

See Also

[fixef](#)

Examples

```
## See the specific class documentation.
```

---

ranef.mexhaz	*Method for extracting random effects from a mexhaz object*

Description

Display a data frame containing the cluster-specific random effects with their standard errors.

Usage

```
## S3 method for class 'mexhaz'
ranef(x, ...)
```

Arguments

x	an object of class mexhaz.
...	not used.

Value

A data frame with three columns containing the cluster names, the random effect estimates, and their standard errors.

See Also

[mexhaz](#)

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib_mix <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull",
expected="popmrate", verbose=0, random="clust")

ranef(Mod_weib_mix)
```

---

riskfunc	*Computation of hazard ratio and risk ratio estimates based on a mexhaz model*

Description

Function for computing hazard ratio and risk ratio (ratio of cumulative probabilities of failure) estimates from a model fitted with the mexhaz function. Corresponding confidence intervals are based on the Delta Method or Monte Carlo simulation (based on the assumption of multivariate normality of the model parameter estimates). This function allows the computation of estimates at one time point for several vectors of covariates or for one vector of covariates at several time points. When the model includes a random effect, three types of predictions can be made: (i) marginal predictions (obtained by integration over the random effect distribution), (ii) cluster-specific posterior predictions for an existing cluster, or (iii) conditional predictions for a given quantile of the random effect distribution (by default, for the median value, that is, 0).

Usage

```
riskfunc(object, time.pts, data, data.0, marginal = TRUE, quant.rdm = 0.5,
cluster = NULL, quant.rdm.0 = 0.5, cluster.0 = NULL, type = c("hr", "rr"),
conf.int = c("delta", "simul"), level = 0.95, nb.sim = 10000, seed = NULL,
dataset = NULL)
```

Arguments

object	an object of class mexhaz, corresponding to a hazard-based regression model fitted with the mexhaz function.
time.pts	a vector of numerical values representing the time points at which predictions are requested. Time values greater than the maximum follow-up time on which the model estimation was based are discarded.
data	a data.frame containing the values of the covariates of the population for which hazard ratios or risk ratios are to be calculated.

data.0   a data.frame containing the values of the covariates of the reference popula-
                tion. Each row of data.0 is used as the reference for the corresponding row of
                data.

marginal   logical value controlling the type of predictions returned by the function when
                the model includes a random intercept. When TRUE, marginal predictions are
                computed. The marginal survival is obtained by integrating the predicted sur-
                vival over the distribution of the random effect. When FALSE (default value),
                conditional predictions depending on the value of the cluster argument are
                calculated.

quant.rdm   numerical value (between 0 and 1) specifying the quantile of the random effect
                distribution that should be used when requesting conditional predictions. The
                default value is set to 0.5 (corresponding to the median, that is a value of the
                random effect of 0). This argument is ignored if the model is a fixed effect
                model, when the marginal argument is set to TRUE, or the cluster argument is
                not NULL.

cluster   a single value corresponding to the name of the cluster for which posterior pre-
                dictions should be calculated. These predictions are obtained by integrating over
                the cluster-specific posterior distribution of the random effect and thus require
                the original dataset. The dataset can either be provided as part of the mexhaz ob-
                ject given as argument or by specifying the name of the dataset in the dataset
                argument (see below). The cluster argument is not used if the model is a fixed
                effect model. The default value is NULL: this corresponds to marginal predictions
                (if marginal is set to TRUE, the preferred option), or to conditional predictions
                for a given quantile (by default, the median) of the distribution of the random
                effect (if marginal is set to FALSE).

quant.rdm.0   random effect distribution quantile value to be used with data.0 (see argument
                quant.rdm for details).

cluster.0   cluster value to be used with data.0 (see argument cluster for details).

type   argument specifying the type of predictions to be calculated. Selection can be
                made between "hr" (hazard ratio) and "rr" (risk ratio, i.e., ratio of cumulative
                failure probabilities).

conf.int   method to be used to compute confidence limits. Selection can be made between
                the following options: "delta" for the Delta Method (default value); "simul"
                for Monte Carlo simulations (can be time-consuming, especially for models us-
                ing B-splines for the logarithm of the baseline hazard).

level   a number in (0,1) specifying the level of confidence for computing the confi-
                dence intervals of the hazard and the survival. By default, this argument is set to
                0.95.

nb.sim   integer value representing the number of simulations used to estimate the con-
                fidence limits for the (excess) hazard and the (net) survival. This argument is
                used only if conf.int is set to "simul".

seed   argument allowing the user to set a random seed for simulations (only relevant
                when conf.int is set to "simul"). The default value is set to NULL in which
                case a random seed is automatically generated inside the function.

dataset   original dataset used to fit the mexhaz object given as argument to the function. This argument is only necessary if cluster-specific posterior predictions are requested (and if the dataset is not already provided in the mexhaz object). The default value is set to NULL.

Value

An object of class resMexhaz that can be used by the function plot.resMexhaz to produce graphics of the hazard ratio or risk ratio curve. It contains the following elements:

results   a data.frame consisting of: the time points at which values have been calculated; the hazard ratio / risk ratio values with their confidence limits.

type   type of results returned by the function. The value is used by plot.resMexhaz and lines.resMexhaz, and can take the values "hr" (hazard ratio) or "rr" (risk ratio).

multiobs   value used by plot.resMexhaz and lines.resMexhaz, and set to FALSE when estimates are computed at several time points for one vector of covariates.

ci.method   method used to compute confidence limits.

level   level of confidence used to compute confidence limits.

Author(s)

Hadrien Charvat, Aurelien Belot

References

Charvat H, Remontet L, Bossard N, Roche L, Dejardin O, Rachet B, Launoy G, Belot A; CENSUR Working Survival Group. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Stat Med 2016;35:3066-3084 (doi: 10.1002/sim.6881)

Skrondal A, Rabe-Hesketh S. Prediction in multilevel generalized linear models. J R Stat Soc A Stat Soc 2009;172(3):659-687 (doi: 10.1111/j.1467-985X.2009.00587.x).

See Also

[plot.resMexhaz](), [lines.resMexhaz]()

Examples

```
data(simdatn1)

## Fit of a fixed-effect hazard model, with the baseline hazard
## described by a linear B-spline with two knots at 1 and 5 year and with
## effects of age (agecr), deprivation index (depindex) and sex (IsexH)

Mod_bs1_2 <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~agecr+depindex+IsexH, data=simdatn1, base="exp.bs",
degree=1, knots=c(1,5), verbose=0)
```

```
## Risk ratio along time for agecr=0.2 compared to agecr=0.1

RR_Modbs1_2 <- riskfunc(Mod_bs1_2, time.pts=seq(0,10,le=101),
data=data.frame(agecr=0.2,depindex=0,IsexH=1),
data.0=data.frame(agecr=0.1,depindex=0,IsexH=1),type="rr",
conf.int="delta")

plot(RR_Modbs1_2)
```

---

simdatn1   *Simulated dataset*

Description

The simdatn1 dataset has 4000 rows and 8 columns.

Format

This dataset contains the following columns:

age Age at diagnosis (continuous).

agecr Centred and rescaled age variable (age-70)/100.

depindex Deprivation index (continuous).

IsexH Sex (0 = Female, 1 = Male).

clust ID number of the cluster.

vstat Vital status (0 = Alive, 1 = Dead).

timesurv Follow-up time (years).

popmrate Population (expected) mortality rate at the time of censoring.

---

simdatn2   *Simulated dataset*

Description

The simdatn2 dataset has 4000 rows and 8 columns. The IsexH variable is simulated with a non-proportional effect.

Format

This dataset contains the following columns:

age Age at diagnosis (continuous).

agecr Centred and rescaled age variable (age-70)/100.

depindex Deprivation index (continuous).

IsexH Sex (0 = Female, 1 = Male).

clust ID number of the cluster.

vstat Vital status (0 = Alive, 1 = Dead).

timesurv Follow-up time (years).

popmrate Population (expected) mortality rate at the time of censoring.

---

summary.mexhaz *Summary method for a mexhaz object*

Description

Produces a summary of a mexhaz object.

Usage

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

Arguments

object	an object of class mexhaz.
...	represents additional parameters directly passed to summary.

See Also

mexhaz, print.mexhaz, print.summary.mexhaz

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull", verbose=0)

summary(Mod_weib)
```

---

update.mexhaz   *Update of a mexhaz model*

Description

Function allowing the user to update an existing mexhaz model. All the arguments of the model can be updated. If the argument 'init' is not provided, the function uses the estimated values of the existing model as starting values for the corresponding parameters of the new model.

Usage

```
## S3 method for class 'mexhaz'
update(object, formula, data, expected = NULL, base = NULL,
degree = 3, knots = NULL, bound = NULL, n.gleg = 20, init = NULL,
random = NULL, n.aghq = 10, fnoptim = c("nlm", "optim"),
verbose = 0, method = "Nelder-Mead", iterlim = 10000, numHess = FALSE,
print.level = 1, exactGradHess = TRUE, gradtol =
ifelse(exactGradHess, 1e-8, 1e-6), envir = parent.frame(), ...)
```

Arguments

object	an object of class mexhaz, corresponding to a survival model fitted with the mexhaz function.
formula	a formula object, with the response on the left of the ~ operator, and the linear predictor on the right. The response must be of the form Surv(time, event) for right censored data or Surv(time, time2, event) for counting process style data. The linear predictor accepts a special instruction nph() for specifying variables for which a time-dependent effect should be modelled (if several variables are modelled with time-dependent effects, separate these variables inside the nph() instruction with a + sign).
	In case time takes the value 0 for some observations when data are entered in the right censored style, it is assumed that these observations refer to events (or censoring) that occurred on the first day of follow-up. Consequently, a value of 1/730.5 (half a day) is substituted in order to make computations possible. However, it should be stressed that this is just a convention and that it does not make much sense if the time scale is not expressed in years. We therefore advise the analyst to deal with 0 time values during the dataset preparation stage.
	See update.formula for more details.
data	a data.frame containing the variables referred to in the formula, as well as in the expected and random arguments if these arguments are used.
expected	name of the variable (must be given in quotes) representing the population (i.e., expected) hazard. By default, expected=NULL, which means that the function estimates the overall hazard (and not the excess hazard).
base	functional form that should be used to model the baseline hazard. Selection can be made between the following options: "weibull" for a Weibull hazard, "exp.bs" for a hazard described by the exponential of a B-spline (only

B-splines of degree 1, 2 or 3 are accepted), `"exp.ns"` for a hazard described by the exponential of a restricted cubic B-spline (also called 'natural spline'), `"pw.cst"` for a piecewise constant hazard. By default, base=NULL.

degree   if base=`"exp.bs"`, degree represents the degree of the B-spline used. Only integer values between 1 and 3 are accepted, and 3 is the default.

knots   if base=`"exp.bs"` or `"exp.ns"`, knots is the vector of interior knots of the spline. If base=`"pw.cst"`, knots is the vector defining the endpoints of the time intervals on which the hazard is assumed to be constant. By default, knots=NULL (that is, it produces a B-spline with no interior knots if base=`"exp.bs"`, a linear B-spline with no interior knots if base=`"exp.ns"`, or a constant hazard over the whole follow-up period if base=`"pw.cst"`).

bound   If base=`"exp.bs"` or base=`"exp.ns"`, computation of the B-spline basis requires that boundary knots be given. The bound argument allows the user to specify these boundary knots. If base=`"exp.bs"`, the interval defined by the boundary knots must at least include the interval c(0,max(time)) (otherwise, there could be problems with ill-conditioned bases). If base=`"exp.ns"`, the boundary knots correspond to the knots beyond which the spline is contrained to be linear (in that case, the boundary knots can be values contained in c(0,max(time))). By default, the boundary knots are set to c(0,max(time)).

n.gleg   if base=`"exp.bs"` and degree is equal to 2 or 3, or if base=`"exp.ns"`, the cumulative hazard is computed via Gauss-Legendre quadrature and `n.gleg` is the number of quadrature nodes to be used to compute the cumulative hazard. By default, `n.gleg=20`.

init   vector of initial values. By default init=NULL and the initial values are internally set to the following values:

for the baseline hazard:

if exactGradHess=TRUE (except for the excess hazard random effect model for which this argument is ignored), the intercept is set to 0.5*log(Number of events/Person-years of observation) and all other parameters set to 0. In case of failed convergence, several trials are run with an adaptation of the value of the intercept.

if exactGradHess=FALSE, the following values are used:

- if base=`"weibull"`, the scale and shape parameters are set to 0.1;

- if base=`"exp.bs"`, the parameters of the B-spline are all set to -1;

- if base=`"exp.ns"`, the parameters of the restricted cubic B-spline are all set to -1;

- if base=`"pw.cst"`, the logarithm of the piecewise-constant hazards are set to -1;

the parameters describing the effects of the covariates are all set to 0;

the parameter representing the standard deviation of the random effect is set to 0.1. (if exactGradHess=TRUE, several trials are run with an adaptation of the value in case of failed convergence).

random   name of the variable to be entered as a random effect (must be given between quotes), representing the cluster membership. By default, random=NULL which means that the function fits a fixed effects model.

n.aghq	number of quadrature points used for estimating the cluster-specific marginal likelihoods by adaptive Gauss-Hermite quadrature. By default, n.aghq=10.
fnoptim	name of the R optimisation procedure used to maximise the likelihood. Selection can be made between "nlm" (by default) and "optim". Note: if exactGradHess=TRUE, this argument will be ignored (fnoptim will be set automatically to "nlm").
verbose	integer parameter representing the frequency at which the current state of the optimisation process is displayed. Internally, an 'evaluation' is defined as an estimation of the log-likelihood for a given vector of parameters. This means that the number of evaluations is increased each time the optimisation procedure updates the value of any of the parameters to be estimated. If verbose=n (with n an integer), the function will display the current values of the parameters, the log-likelihood and the time elapsed every n evaluations. If verbose=0, nothing is displayed.
method	if fnoptim="optim", method represents the optimisation method to be used by optim. By default, method="Nelder-Mead". This parameter is not used if fnoptim="nlm".
iterlim	if fnoptim="nlm", iterlim represents the maximum number of iterations before the nlm optimisation procedure is terminated. By default, iterlim is set to 10000. This parameter is not used if fnoptim="optim" (in this case, the maximum number of iterations must be given as part of a list of control parameters via the control argument: see the help page of optim for further details).
numHess	logical value allowing the user to choose between the Hessian returned by the optimization algorithm (default) or the Hessian estimated by the hessian function from the numDeriv package. The latter might be more accurate but its estimation is more time-consuming. We suggest to use the default Hessian estimation procedure during model building and estimate the numDeriv-based Hessian only on the final model. Note: if exactGradHess=TRUE, this argument is ignored.
print.level	this argument is only used if fnoptim="nlm". It determines the level of printing during the optimisation process. The default value (for the mexhaz function) is set to '1' which means that details on the initial and final step of the optimisation procedure are printed (see the help page of nlm for further details).
exactGradHess	logical value allowing the user to decide whether maximisation of the likelihood should be based on the analytic gradient and Hessian computed internally (default, corresponding to exactGradHess=TRUE). In that case, optimisation is performed with the nlm function. Note: even if set to TRUE, this argument is ignored when the user wants to fit an excess hazard model including a random effect because in that case, there is no simple way to obtain the analytic gradient and Hessian.
gradtol	this argument is only used if fnoptim="nlm". It corresponds to the tolerance at which the scaled gradient is considered close enough to zero to terminate the algorithm. The default value depends on the value of the argument exactGradHess.
envir	environment in which the objects' names given as arguments to the updated model are to be found.
...	represents additional parameters directly passed to nlm or optim to control the optimisation process.

Value

An object of class mexhaz. See [mexhaz](#) for more details.

Author(s)

Hadrien Charvat, Aurelien Belot

References

Charvat H, Remontet L, Bossard N, Roche L, Dejardin O, Rachet B, Launoy G, Belot A; CENSUR Working Survival Group. A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates. Stat Med 2016;35:3066-3084 (doi: 10.1002/sim.6881)

See Also

[mexhaz](#)

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull", verbose=0)

## Add an effect of gender
Mod_weib_2 <- update(Mod_weib, formula=~.+IsexH)
```

---

vcov.mexhaz   *Method for extracting the covariance matrix of the fixed effects*

Description

Display a the covariance matrix of the fixed effects of a mexhaz model.

Usage

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

Arguments

object	an object of class mexhaz.
...	not used.

Value

The estimated covariance matrix of the fixed effects.

See Also

[mexhaz](#)

Examples

```
data(simdatn1)

## Fit of a mixed-effect excess hazard model, with the baseline hazard
## described by a Weibull distribution (without covariables)

Mod_weib_mix <- mexhaz(formula=Surv(time=timesurv,
event=vstat)~1, data=simdatn1, base="weibull",
expected="popmrate", verbose=0, random="clust")

vcov(Mod_weib_mix)
```

Index