Package 'dynpred'

July 22, 2025

Version 0.1.2

Date 2015-07-13

Title Companion Package to ``Dynamic Prediction in Clinical Survival Analysis''

Author Hein Putter

Maintainer Hein Putter <H.Putter@lumc.nl>

Depends survival

Imports graphics, stats, utils

Suggests mstate

Description The dynpred package contains functions for dynamic prediction in survival analysis.

License GPL (>= 2)

URL http://www.msbi.nl/putter

NeedsCompilation yes

Repository CRAN

Date/Publication 2015-07-13 21:22:03

Contents

`dynpred-package` *The companion package of the book "Dynamic Prediction in Survival Analysis"*

Description

The companion package of the book "Dynamic Prediction in Survival Analysis".

Details

Package:	dynpred
Type:	Package
Version:	0.1.2
Date:	2014-11-10
License:	GPL (>= 2)

An overview of how to use the package, including the most important functions.

Author(s)

Hein Putter Maintainer: Hein Putter <H.Putter@lumc.nl>

References

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

AUC *Calculate AUC(t) curve*

Description

Calculate model-free curve of Area Under the Curve values over time, based on the dynamic/incident AUC of Heagerty and Zheng.

Usage

```
AUC(formula, data, plot = TRUE)
```

Arguments

formula	Formula for prediction model to be used as in coxph
data	Data set in which to interpret the formula
plot	Determines whether the AUC function should be plotted (if TRUE (default)) along with a lowess curve or not (if FALSE)

Value

A list with elements

AUCt	A data frame with time t in column time and AUC(t) in column AUC
AUC	The AUC(t) weighted by Y(t)-1, with Y(t) the number at risk at t; this coincides with Harrell's c-index

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

Harrell FE, Lee KL & Mark DB (1996), Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, Statistics in Medicine 15, 361-387.

Heagerty PJ & Zheng Y (2005), Survival model predictive accuracy and ROC curves, Biometrics 61, 92-105.

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
AUC(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova)
```

AUCw *Calculate dynamic AUC(t) curve*

Description

Calculate dynamic model-free curve of Area Under the Curve values over time, based on the dynamic/incident AUC of Heagerty and Zheng.

Usage

```
AUCw(formula, data, width)
```

Arguments

formula	Formula for prediction model to be used as in coxph
data	Data set in which to interpret the formula
width	Width of the window

Value

A data frame with columns

time	The time points t at which AUCw(t) changes value (either t or t+width is an event time point)
AUCw	The AUCw(t) function

and with attribute "width" given as input.

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
AUCw(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova,
  width = 2)
```

cindex *Calculate Harrell's c-index*

Description

This function calculates Harrell's c-index.

Usage

```
cindex(formula, data)
```

Arguments

formula	Formula for prediction model to be used as in coxph
data	Data set in which to interpret the formula

Value

A list with elements

concordant	The number of concordant pairs
total	The total number of pairs that can be evaluated
cindex	Harrell's c-index

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

Harrell FE, Lee KL & Mark DB (1996), Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, Statistics in Medicine 15, 361-387.

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
cindex(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova)
```

cutLM *Create landmark data set*

Description

Create landmark data set from original data, which can be either in wide or long format, see details.

Usage

```
cutLM(data, outcome, LM, horizon, covs, format = c("wide", "long"), id, rtime,
  right = TRUE)
```

Arguments

data	Data frame from which to construct landmark dataset
outcome	List with items time and status, containing character strings identifying the names of time and status variables, respectively, of the survival outcome
LM	Scalar, the value of the landmark time point
horizon	Scalar, the value of the horizon. Administrative censoring is applied at horizon.
covs	List with items fixed and varying, containing character strings specifying column names in the data containing time-fixed and time-varying covariates, respectively
format	Character string specifying whether the original data are in wide (default) or in long format
id	Character string specifying the column name in data containing the subject id; only needed if format="long"
rtime	Character string specifying the column name in data containing the (running) time variable associated with the time-varying covariate(s); only needed if format="long"
right	Boolean (default=TRUE), indicating if the intervals for the time-varying covariates are closed on the right (and open on the left) or vice versa, see cut

Details

For a given landmark time point LM, patients who have reached the event of interest (outcome) or are censored before or at LM are removed. Administrative censoring is applied at the time horizon. Time-varying covariates are evaluated at the landmark time point LM. Time-varying covariates can be specified in the varying item of the covs argument, in two ways. In the first way (data in long format) different values of time-dependent covariate(s) are stored different rows of the data, with id identifying which values belong to the same subject; the column specified through rtime then contains the time points at which the value of the covariate changes value; with right=TRUE (default), it is assumed that the covariate changes value at the time point specified in rtime (and hence is not used for prediction of an event at rtime), while with right=FALSE, it is assumed that the covariate changes value just before the time point specified in rtime. The second way (data in wide format) can only be used for a specific type of time-varying covariates, often used to model whether some other event has occurred or not, namely those that change value from 0 (event not yet occurred) to 1 (event has occurred).

Value

A landmark data set, containing the outcome and the values of time-fixed and time-varying covariates taken at the landmark time points. The value of the landmark time point is stored in column LM.

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
test0 <- data.frame(id=c(1,1,1,2,2,2),survyrs=c(2.3,2.3,2.3,2.7,2.7,2.7),
  survstat=c(1,1,1,0,0,0),age=c(76,76,76,68,68,68),gender=c(1,1,1,2,2,2),
  bp=c(80,84,88,92,90,89),bptime=c(1,2,2.2,0,1,2))
cutLM(test0, outcome=list(time="survyrs", status="survstat"),
  LM=1, horizon=2.5, covs=list(fixed=c("age","gender"),varying="bp"),
  format="long", id="id", rtime="bptime")
# Note how the previous example does not use the value of the time-varying
# covariate AT time=LM, only just before (if available). This is in line
# with the time-varying covariates being predictable.
# If you want the value of the time-varying covariate at time=LM if it
# changes value at LM, then use right=FALSE
cutLM(test0, outcome=list(time="survyrs", status="survstat"),
  LM=1, horizon=2.5, covs=list(fixed=c("age","gender"),varying="bp"),
  format="long", id="id", rtime="bptime", right=FALSE)

# An example of a time-varying covariate in wide format; recyrs and recstat
# are time and status of a (cancer) recurrence. Here it is assumed that the
# value of the time-varying covariate is 0 and changes value to 1 at recyrs.
# The status variable is not used!
test1 <- data.frame(id=1:4,survyrs=c(7.6,8.4,5.3,2.6),survstat=c(0,1,1,0),
  age=c(48,52,76,18),gender=c(1,2,2,1),recyrs=c(7.6,5.2,0.8,2.6),
  recstat=c(0,1,1,0))
cutLM(test1, outcome=list(time="survyrs", status="survstat"),
  LM=3, horizon=8, covs=list(fixed=c("id","age","gender"),varying="recyrs"))

# The same example in long format, similar to (but not the same as) the way
# one would use a time-varying covariate in long format.
test2 <- data.frame(id=c(1,2,2,3,3,4),survyrs=c(7.6,8.4,8.4,5.3,5.3,2.6),
  survstat=c(0,1,1,1,1,0),age=c(48,52,52,76,76,18),gender=c(1,2,2,2,2,1),
  rec=c(0,0,1,0,1,0),rectime=c(0,0,5.2,0,0.8,0))
cutLM(test2, outcome=list(time="survyrs", status="survstat"),
  LM=3, horizon=8, covs=list(fixed=c("age","gender"),varying="rec"),
  format="long", id="id", rtime="rectime")
```

CVcindex *Calculate cross-validated c-index*

Description

This function calculates cross-validated versions of Harrell's c-index.

Usage

```
CVcindex(formula, data, type = "single", matrix = FALSE)
```

Arguments

formula	Formula for prediction model to be used as in coxph
data	Data set in which to interpret the formula
type	One of "single", "pair" or "fullpairs". For "single" (default), the prognostic index Z_i is replaced by Z_i,(-i), for "pair", two assessments of concordance are made for each pair (i,j), one using Z_i,(-i) and Z_j,(-i), the other using Z_i,(-j) and Z_j,(-j), for "fullpairs", each of the possible pairs is left out and comparison is based on Z_i,(-i,-j) and Z_j,(-i,-j)
matrix	if TRUE, the matrix of cross-validated prognostic indices is also returned; default is FALSE

Value

A list with elements

concordant	The number of concordant pairs
total	The total number of pairs that can be evaluated
cindex	The cross-validated c-index
matrix	Matrix of cross-validated prognostic indices (only if argument matrix is TRUE

and with attribute "type" as given as input.

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

Harrell FE, Lee KL & Mark DB (1996), Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, Statistics in Medicine 15, 361-387.

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
# Real thing takes a long time, so on a smaller data set
ova2 <- ova[1:100,]
# Actual c-index
cindex(Surv(tyears,d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova2)
# Cross-validated c-indices
CVcindex(Surv(tyears,d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova2)
CVcindex(Surv(tyears,d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova2,
         type="pair")

CVcindex(Surv(tyears,d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova2,
         type="fullpairs")
```

CVPL *Calculate cross-validated log-partial likelihood (with shrinkage)*

Description

This function calculates the cross-validated log partial likelihood, with shrinkage if requested.

Usage

```
CVPL(formula, data, progress = TRUE, overall = FALSE, shrink = 1)
```

Arguments

formula	Formula for prediction model to be used as in coxph
data	Data set in which to interpret the formula
progress	if TRUE (default), progress of the cross-validation will be printed
overall	if TRUE, CVPL uses regression coefficient estimates based on the full data, for each observation i, rather than the estimates based on data minus i
shrink	Shrinkage factor; default is 1 (no shrinkage)

Value

Numeric; the cross-validated log partial likelihood

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

Verweij PJM & van Houwelingen HC (1994), Penalized likelihood in Cox regression, Statistics in Medicine 13, 2427-2436.

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
CVPL(Surv(tyears, d) ~ 1, data = ova)
CVPL(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam,
  data = ova)
CVPL(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam,
  data = ova, overall=TRUE)
```

---

deb *Debugging function*

Description

A simple but useful debugging function. It first announces the object to printed and then prints it.

Usage

```
deb(x, method = c("print", "cat"))
```

Arguments

x	The object to be printed
method	The method for printing x. Default is "print", which uses print for printing; "cat" uses cat for printing. The latter is useful for short objects (scalar and vectors), the former for more structured objects (data frames, matrices, lists etc).

Author(s)

Hein Putter <H.Putter@lumc.nl>

Examples

```
tm <- c(0.2,0.5,1,1.2,1.8,4)
ta <- 2*tm
dfr <- data.frame(time=tm, stepf=ta)
deb(dfr, method="print")
deb(tm, method="cat")
```

EBMT data *Data from the European Society for Blood and Marrow Transplantation (EBMT)*

Description

Data from the European Society for Blood and Marrow Transplantation (EBMT)

Format

A data frame of 2279 patients transplanted at the EBMT between 1985 and 1998. These data were used in Fiocco, Putter & van Houwelingen (2008) and van Houwelingen & Putter (2008). The included variables are

id Patient identification number

rec Time in days from transplantation to recovery or last follow-up

rec.s Recovery status; 1 = recovery, 0 = censored

ae Time in days from transplantation to adverse event (AE) or last follow-up

ae.s Adverse event status; 1 = adverse event, 0 = censored

recae Time in days from transplantation to both recovery and AE or last follow-up

plag.s Recovery and AE status; 1 = both recovery and AE, 0 = no recovery or no AE or censored

rel Time in days from transplantation to relapse or last follow-up

rel.s Relapse status; 1 = relapse, 0 = censored

srv Time in days from transplantation to death or last follow-up

srv.s Relapse status; 1 = dead, 0 = censored

year Year of transplantation; factor with levels "1985-1989", "1990-1994", "1995-1998"

agecl Patient age at transplant; factor with levels "<=20", "20-40", ">40"

proph Prophylaxis; factor with levels "no", "yes"

match Donor-recipient gender match; factor with levels "no gender mismatch", "gender mismatch"

Source

We gratefully acknowledge the European Society for Blood and Marrow Transplantation (EBMT) for making available these data. Disclaimer: these data were simplified for the purpose of illustration of the analysis of competing risks and multi-state models and do not reflect any real life situation. No clinical conclusions should be drawn from these data.

References

Fiocco M, Putter H, van Houwelingen HC (2008). Reduced-rank proportional hazards regression and simulation-based prediction for multi-state models. *Statistics in Medicine* **27**, 4340–4358.

van Houwelingen HC, Putter H (2008). Dynamic predicting by landmarking as an alternative for multi-state modeling: an application to acute lymphoid leukemia data. *Lifetime Data Anal* **14**, 447–463.

`evalstep` *Evaluate step function at a set of new time points*

Description

Given one or more right-continuous step functions of time, given by vector `time` and vector of matrix `stepf`, this function evaluates the step function(s) at a vector of new time points given by `newtime`. Typical application is when the step function is given by a non- or semi-parametric estimated of cumulative hazard or survival function, and the value of this function is required at a set of time points.

Usage

```
evalstep(time, stepf, newtime, subst = -Inf, to.data.frame = FALSE)
```

Arguments

`time`	A vector of time points at which the step function changes value
`stepf`	A vector (of the same length as `time`) or a matrix (with no of columns equal to the length of `time`) containing the values of the step function(s) at the time points
`newtime`	A vector of time points at which the step function(s) is/are to be evaluated
`subst`	A value that is substituted for elements of `newtime` that are smaller than the minimum of `time`. Default value is `-Inf`
`to.data.frame`	Determines whether the output is a data frame with the new time points and the values of the step function(s) (if `TRUE`) or a vector/matrix with the values of the step function(s) (if `FALSE` (default))

Details

The argument `time` should be ordered, and not contain duplicated or +/- Inf, and should be of the same length as `stepf`. There are no restrictions on ordering or duplicates of `newtime`. For elements of `newtime` that are smaller than the minimum of `time`, the value of `subst` is substituted.

Value

Either a vector/matrix containing the step function(s) evaluated at the new time points (if `to.data.frame=FALSE` (default)), or a data frame with column vectors `newtime` containing the new time points and `res` containing the step function evaluated at the new time points (if `to.data.frame=TRUE`)

Author(s)

Hein Putter <H.Putter@lumc.nl>

Examples

```
tm <- c(0.2,0.5,1,1.2,1.8,4)
ta <- 2*tm
data.frame(time=tm, stepf=ta)
evalstep(time=tm, stepf=ta, newtime=c(0,0.2,0.3,0.6,1,1.5,3,4,5,0.1), subst=0)
evalstep(time=tm, stepf=data.frame(ta=ta,ta2=1/ta),
newtime=c(0,0.2,0.3,0.6,1,1.5,3,4,5,0.1), subst=0)
```

Fwindow *Calculate dynamic "death within window" curve*

Description

Calculate dynamic "death within window" curve, in other words, one minus fixed width conditional survival curves, defined as P(T<=t+w|T>t), for a fixed window width w.

Usage

```
Fwindow(object, width, variance = TRUE, conf.level = 0.95)
```

Arguments

object	survfit object, use type="aalen"
width	Width of the window
variance	Boolean (default=TRUE); should pointwise confidence interval of the probabilities be calculated?
conf.level	The confidence level, between 0 and 1 (default=0.95)

Details

"Die within window function" with window w, Fw(t) = P(T<=t+w|T>t), evaluated at all time points t where the estimate changes value, and associated pointwise confidence intervals (if variance=TRUE).

Both estimate and pointwise lower and upper confidence intervals are based on the negative exponential of the Nelson-Aalen estimate of the cumulative hazard, so P(T<=t+w|T>t) is estimated as exp(- int_t^t+w hatH_NA(s) ds), with hatH_NA the non-parametric Nelson-Aalen estimate.

Note: in object, no event time points at or below zero allowed

Value

A data frame with columns

time	The time points t at which Fw(t) changes value (either t or t+width is an event time point)
Fw	The Fw(t) function
low	Lower end of confidence interval
up	Upper end of confidence interval

and with attribute "width" as given as input.

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(wbc1)
c0 <- coxph(Surv(tyears, d) ~ 1, data = wbc1, method="breslow")
sf0 <- survfit(c0)
Fw <- Fwindow(sf0,4)
```

NKI breast cancer clinical data

Clinical and follow-up data of breast cancer patients as collected in the Dutch Cancer Institute (NKI) in Amsterdam

---

Description

A data frame of 295 patients with breast cancer. The included variables are

patnr Patient identification number

d Survival status; 1 = death; 0 = censored

tyears Time in years until death or last follow-up

diameter Diameter of the primary tumor

posnod Number of positive lymph nodes

age Age of the patient

mlratio Estrogen level?

chemotherapy Chemotherapy used (yes/no)

hormonaltherapy Hormonal therapy used (yes/no)

typesurgery Type of surgery (excision or mastectomy)

histolgrade Histological grade (Intermediate, poorly, or well differentiated)

vasc.invasion Vascular invasion (-, +, or +/-)

crossval.clin.class ??

PICV Estrogen level?

Format

A data frame, see data.frame.

References

van't Veer LJ, Dai HY, van de Vijver MJ, He YDD, Hart AAM, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R \& Friend SH (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536.

van de Vijver MJ, He YD, van 't Veer LJ, Dai H, Hart AAM, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH \& Bernards R (2002). A gene-expression signature as a predictor of survival in breast cancer. *New England Journal of Medicine* **347**, 1999–2009.

van Houwelingen HC, Bruinsma T, Hart AAM, van't Veer LJ \& Wessels LFA (2006). Cross-validated Cox regression on microarray gene expression data. *Statistics in Medicine* **25**, 3201–3216.

`Ovarian cancer data` *Data originate from two clinical trials on the use of different combination chemotherapies, carried out in The Netherlands around 1980*

Description

A data frame of 358 patients with ovarian cancer. The included variables are

tyears Time in years until death or last follow-up

d Survival status; 1 = death; 0 = censored

Karn Karnofsky score

Broders Broders score: factor with levels "unknown", "1", "2", "3", "4"

FIGO FIGO stage; factor with levels "III", "IV"

Ascites Presence of ascires; factor with levels "unknown", "absent", "present"

Diam Diameter of the tumor; factor with levels "micr.", "<1cm", "1-2cm", "2-5cm", ">5cm"

Format

A data frame, see `data.frame`.

References

Neijt, J. P., ten Bokkel Huinink, W. W., van der Burg, M. E., van Oosterom, A. T., Vriesendorp, R., Kooyman, C. D., van Lindert, A. C., Hamerlynck, J. V., van Lent, M. & van Houwelingen, J. C. (1984), 'Randomised trial comparing two combination chemotherapy regimens (Hexa-CAF vs CHAP- 5) in advanced ovarian carcinoma', Lancet 2, 594–600.

Neijt, J. P., ten Bokkel Huinink, W. W., van der Burg, M. E., van Oosterom, A. T., Willemse, P. H., Heintz, A. P., van Lent, M., Trimbos, J. B., Bouma, J. & Vermorken, J. B. (1987), 'Randomized trial comparing two combination chemotherapy regimens (CHAP-5 vs CP) in advanced ovarian carcinoma', Journal of Clinical Oncology 5, 1157–1168.

van Houwelingen, J. C., ten Bokkel Huinink, W. W., van der Burg, M. E., van Oosterom, A. T. & Neijt, J. P. (1989), 'Predictability of the survival of patients with advanced ovarian cancer.', Journal of Clinical Oncology 7, 769–773.

pe *Calculate prediction error curve*

Description

Calculate prediction error curve.

Usage

```
pe(time, status, tsurv, survmat, tcens, censmat, FUN = c("KL", "Brier"), tout)

pecox(formula, censformula, data, censdata, FUN = c("KL", "Brier"), tout,
  CV = FALSE, progress = FALSE)
```

Arguments

time	Vector of time points in data
status	Vector of event indicators in data
tsurv	Vector of time points corresponding to the estimated survival probabilities in survmat
survmat	Matrix of estimated survival probabilities; dimension should be length of tsurv x length of time
tcens	Vector of time points corresponding to the estimated censoring probabilities in censmat
censmat	Matrix of estimated censoring probabilities; dimension should be length of tcens x length of time
FUN	The error function, either "KL" (default) for Kullback-Leibler or "Brier" for Brier score
tout	Vector of time points at which to evaluate prediction error. If missing, prediction error will be evaluated at all time points where the estimate will change value
formula	Formula for prediction model to be used as in coxph
censformula	Formula for censoring model, also to be used as in coxph
data	Data set in which to interpret formula
censdata	Data set in which to interpret censformula
CV	Boolean (default=FALSE); if TRUE, (leave-one-out) cross-validation is used for the survival probabilities
progress	Boolean (default=FALSE); if TRUE, progress is printed on screen

Details

The censformula is used to calculate inverse probability of censoring weights (IPCW).

Value

A data frame with columns

time	Event time points
Err	Prediction error of model specified by formula at these time points

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

Graf E, Schmoor C, Sauerbrei W & Schumacher M (1999), Assessment and comparison of prognostic classification schemes for survival data, Statistics in Medicine 18, 2529-2545.

Gerds & Schumacher (2006), Consistent estimation of the expected Brier score in general survival models with right-censored event times, Biometrical Journal 48, 1029-1040.

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
# Example on a subset, because the effect of CV is clearer
ova2 <- ova[1:100,]
pecox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  data = ova2, FUN="Brier", tout=seq(0,6,by=0.5))
pecox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  data = ova2, FUN="Brier", tout=seq(0,6,by=0.5), CV=TRUE, progress=TRUE)


pecox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  data = ova, FUN="Brier", tout=seq(0,6,by=0.5))
pecox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  data = ova, FUN="Brier", tout=seq(0,6,by=0.5), CV=TRUE, progress=TRUE)
```

---

pew *Calculate dynamic prediction error curve*

Description

Calculate dynamic fixed width prediction error curve.

Usage

```
pew(time, status, tsurv, survmat, tcens, censmat, width, FUN = c("KL",
  "Brier"), tout)

pewcox(formula, censformula, width, data, censdata, FUN = c("KL", "Brier"),
  tout, CV = FALSE, progress = FALSE)
```

Arguments

time	Vector of time points in data
status	Vector of event indicators in data
tsurv	Vector of time points corresponding to the estimated survival probabilities in survmat
survmat	Matrix of estimated survival probabilities; dimension should be length of tsurv x length of time
tcens	Vector of time points corresponding to the estimated censoring probabilities in censmat
censmat	Matrix of estimated censoring probabilities; dimension should be length of tcens x length of time
width	Width of the window
FUN	The error function, either "KL" (default) for Kullback-Leibler or "Brier" for Brier score
tout	Vector of time points at which to evaluate prediction error. If missing, prediction error will be evaluated at all time points where the estimate will change value
formula	Formula for prediction model to be used as in coxph
censformula	Formula for censoring model, also to be used as in coxph
data	Data set in which to interpret formula
censdata	Data set in which to interpret censformula
CV	Boolean (default=FALSE); if TRUE, (leave-one-out) cross-validation is used for the survival probabilities
progress	Boolean (default=FALSE); if TRUE, progress is printed on screen

Details

Corresponds to Equation (3.6) in van Houwelingen and Putter (2011). The censformula is used to calculate inverse probability of censoring weights (IPCW).

Value

A data frame with columns

time	Event time points
Err	Prediction error of model specified by formula at these time points

and with attribute "width" given as input.

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
# Example on a subset, because the effect of CV is clearer
ova2 <- ova[1:100,]
pewcox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  width=2, data = ova2, FUN="Brier", tout=seq(0,6,by=0.5))
pewcox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  width=2, data = ova2, FUN="Brier", tout=seq(0,6,by=0.5), CV=TRUE, progress=TRUE)


pewcox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  width=2, data = ova, FUN="Brier", tout=seq(0,6,by=0.5))
pewcox(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, Surv(tyears, 1-d) ~ 1,
  width=2, data = ova, FUN="Brier", tout=seq(0,6,by=0.5), CV=TRUE, progress=TRUE)
```

---

scatterplot *Create scatter plot with imputed survival times*

Description

Create scatter plot with imputed survival times.

Usage

```
scatterplot(formula, data, horizon, plot = TRUE, xlab)
```

Arguments

formula	Formula for prediction model to be used as in coxph
data	Data set in which to interpret the formula
horizon	The horizon, maximum value to be imputed in case of censored observations; default is 1.05 times largest event time
plot	Should the tolerance plot actually be plotted? Default is TRUE
xlab	Label for x-axis

Details

Imputation is used for censored survival times.

Value

A data frame with columns

x	Predictor (centered at zero)
imputed	(Imputed) survival time

and with attribute "horizon" (copied from input or default).

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

Royston P (2001), The lognormal distribution as a model for survival time in cancer, with an emphasis on prognostic factors, Statistica Neerlandica 55, 89-104.

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
scatterplot(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova)
```

`toleranceplot` *Create a tolerance plot*

Description

Create a tolerance plot according to the methods of Henderson, Jones & Stare (2001)

Usage

```
toleranceplot(formula, data, coverage = 0.8, horizon, plot = TRUE, xlab)
```

Arguments

`formula`	Formula for prediction model to be used as in `coxph`
`data`	Data set in which to interpret the formula
`coverage`	The coverage for the tolerance intervals (default is 0.8)
`horizon`	The horizon, maximum value to be imputed in case of censored observations; default is 1.05 times largest event time
`plot`	Should the tolerance plot actually be plotted? Default is `TRUE`
`xlab`	Label for x-axis

Details

Warnings will be issued each time the survival curve corresponding to a value of x never goes below (1-coverage)/2; these warnings may be ignored.

Value

A data frame with columns

`x`	Predictor (centered at zero)
`lower`	Lower bound of tolerance interval
`upper`	Upper bound of tolerance interval

and with attributes `"coverage"` and `"horizon"` (copied from input or default).

Author(s)

Hein Putter <H.Putter@lumc.nl>

References

Henderson R, Jones M & Stare J (2001), Accuracy of point predictions in survival analysis, Statistics in Medicine 20, 3083-3096.

van Houwelingen HC, Putter H (2012). Dynamic Prediction in Clinical Survival Analysis. Chapman & Hall.

Examples

```
data(ova)
toleranceplot(Surv(tyears, d) ~ Karn + Broders + FIGO + Ascites + Diam, data = ova)
```

`WBC follow-up data` *Data from the Benelux CML study*

Description

A data frame of 210 patients with Chronic Myeloid Leukemia from the Benelux CML study (Kluin-Nelemans et al. 1998). Data have been used in two methodological papers, de Bruijne et al. (2001) and van Houwelingen (2007), and in the book van Houwelingen \& Putter (2011), especially Chapter 8. More background is given in Appendix A.2 of van Houwelingen \& Putter (2011). Interest is in the time-dependent covariate White Blood Cell count (WBC). Data set wbc1 contains the follow-up data and time-fixed covariates, while `wbc2` contains the WBC measurements. The included variables in wbc1 are

patnr Patient identification number

tyears Time in years from randomization to death or last follow-up

d Survival status; 1 = dead, 0 = censored

sokal Clinical index based on spleen size, percentage of circulating blasts, platelet and age at diagnosis

age Age at diagnosis

Format

A data frame, see data.frame.

References

Kluin-Nelemans JC, Delannoy A, Louwagie A, le Cessie S, Hermans J, van der Burgh JF, Hagemeijer AM, van den Berghe H \& Benelux CML Study Group (1998). Randomized study on hydroxyurea alone versus hydroxyurea combined with low-dose interferon-alpha 2b for chronic myeloid leukemia. *Blood* **91**, 2713–2721.

de Bruijne MHJ, le Cessie S, Kluin-Nelemans HC \& van Houwelingen HC (2001). On the use of Cox regression in the presence of an irregularly observed time-dependent covariate. *Statistics in Medicine* **20**, 3817–3829.

van Houwelingen HC (2007). Dynamic prediction by landmarking in event history analysis. *Scandinavian Journal of Statistics* **34**, 70–85.

van Houwelingen HC, Putter H (2012). Dynamic Predicting in Clinical Survival Analysis. Chapman \& Hall.

WBC measurements data *Data from the Benelux CML study*

Description

A data frame of 210 patients with Chronic Myeloid Leukemia from the Benelux CML study (Kluin-Nelemans et al. 1998). Data have been used in two methodological papers, de Bruijne et al. (2001) and van Houwelingen (2007), and in the book van Houwelingen \& Putter (2011), especially Chapter 8. More background is given in Appendix A.2 of van Houwelingen \& Putter (2011). Interest is in the time-dependent covariate White Blood Cell count (WBC). Data set wbc1 contains the follow-up data and time-fixed covariates, while wbc2 contains the WBC measurements. The included variables in wbc2 are

patnr Patient identification number

tyears Time of WBC measurement in years from randomization

lwbc Log-transformed and standardized WBC measurement, more precisely, defined as lwbc=log10(wbc)-0.95

Format

A data frame, see data.frame.

References

Kluin-Nelemans JC, Delannoy A, Louwagie A, le Cessie S, Hermans J, van der Burgh JF, Hagemeijer AM, van den Berghe H \& Benelux CML Study Group (1998). Randomized study on hydroxyurea alone versus hydroxyurea combined with low-dose interferon-alpha 2b for chronic myeloid leukemia. *Blood* **91**, 2713–2721.

de Bruijne MHJ, le Cessie S, Kluin-Nelemans HC \& van Houwelingen HC (2001). On the use of Cox regression in the presence of an irregularly observed time-dependent covariate. *Statistics in Medicine* **20**, 3817–3829.

van Houwelingen HC (2007). Dynamic prediction by landmarking in event history analysis. *Scandinavian Journal of Statistics* **34**, 70–85.

van Houwelingen HC, Putter H (2012). Dynamic Predicting in Clinical Survival Analysis. Chapman \& Hall.

Index