Package 'wally'

May 8, 2026

Type Package

Title The Wally Calibration Plot for Risk Prediction Models

Version 1.0.10

Author Paul F Blanche <paulfblanche@gmail.com>, Thomas A. Gerds
<tag@biostat.ku.dk>

Maintainer Paul F. Blanche <paulfblanche@gmail.com>

Description A prediction model is calibrated if, roughly, for any percentage x we can expect that x subjects out of 100 experience the event among all subjects that have a predicted risk of x%. A calibration plot provides a simple, yet useful, way of assessing the calibration assumption. The Wally plot consists of a sequence of usual calibration plots. Among the plots contained within the sequence, one is the actual calibration plot which has been obtained from the data and the others are obtained from similar simulated data under the calibration assumption. It provides the investigator with a direct visual understanding of the shape and sampling variability that are common under the calibration assumption. The original calibration plot from the data is included randomly among the simulated calibration plots, similarly to a police lineup. If the original calibration plot is not easily identified then the calibration assumption is not contradicted by the data. The method handles the common situations in which the data contain censored observations and occurrences of competing events.

Depends R (>= 2.9.0), prodlim (>= 1.4.9)

Imports stats, riskRegression (>= 1.0.8), data.table (>= 1.10.4)

Suggests testthat, survival

License GPL (>= 2)

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2019-10-30 14:40:02 UTC

Contents

`divat` *divat data*

Description

Extracted data from a french population based cohort (DIVAT cohort). The dataset includes followup information on kidney graft failure outcome and predicted 5-year risks based on based on the subject specific information which includes age, gender, cardiovascular and diabetes histories, monitoring of the evolution of the kidney function measured via serum creatinine and relevant characteristics of his or her kidney donor. Graft failure is defined as either death with functioning kidney graft or return to dialysis. The prediction model from which the predictions have been computed has been previously fitted using an independent training sample from the DIVAT data. Details about data and modeling can be found in Fournier et al. (2016).

Format

A subsample consisting of 1300 observations on the following 3 variables.

pi 5-year risk prediction of kidney graft failure.

status 0=censored, 1=kidney graft failure

time time to event (i.e., time to kidney graft failure or loss of follow-up)

References

Fournier, M. C., Foucher, Y., Blanche, P., Buron, F., Giral, M., & Dantan, E. (2016). A joint model for longitudinal and time-to-event data to better assess the specific role of donor and recipient factors on long-term kidney transplantation outcomes. European journal of epidemiology, 31(5), 469-479.

Examples

```
data(divat)
```

`threecity` *threecity data*

Description

Extracted data from a french population based cohort (Three-City cohort). The dataset includes followup information on dementia outcome and predicted 5-year risks based on based on the subject specific information which includes age, gender, education level and cognitive decline measured by a psychometric test (Mini Mental State Examination). The prediction model from which the predictions have been computed has been fitted on independent training data from the Paquid cohort, another french population based cohort with similar design (see Reference Blanche et al. 2015 for details) .

Format

A subsample consisting of 2000 observations on the following 3 variables.

pi 5-year absolute risk predictions of dementia.

status 0=censored, 1=dementia, 2=death dementia free

time time to event (i.e., time to either dementia, death dementia free or loss of follow-up)

Source

Web-appendix of Blanche et al. (2015).

References

Blanche, P., Proust-Lima, C., Loubere, L., Berr, C., Dartigues, J. F., Jacqmin-Gadda, H. (2015). Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. Biometrics, 71(1), 102-113.

Examples

```
data(threecity)
```

`wallyPlot` *Wally plots to assess calibration of a risk or survival prediction*

Description

##' Wally plots to assess calibration of a risk or survival prediction

Usage

```
wallyPlot(object, time, formula, data, cause = 1, q = 10, ylim,
  hanging = FALSE, seed = NULL, mar = c(4.1, 4.1, 2, 2),
  colbox = "red", verbose = TRUE, col = c("grey90", "grey30"),
  xlab = "Risk groups", labels = "quantiles.labels", ...)
```

Arguments

`object`	Probabilistic survival predictions or probabilistic event risk predictions evaluated at `time` for the subjects in `data`. Either given in form of a numeric vector of probabilistic predictions or as an object which has a `predictRisk` method
`time`	Time interest for evaluating calibration of the predictions.
`formula`	A survival or event history formula. The left hand side is used to compute the expected event status. If `formula` is `missing`, try to extract a formula from the first element in object.

data	A data frame in which to validate the prediction models and to fit the censoring model. If data is missing, try to extract a data set from the first element in object.
cause	For competing risks settings the cause of interest.
q	The number of quantiles. Defaults to 10.
ylim	Limits of y-axis. If missing the function tries to find appropriate limits based on the simulated and real data.
hanging	If TRUE, hang bars corresponding to observed frequencies at the value of the corresponding prediction.
seed	A seed value to make results reproducible.
mar	Plot margins passed to par.
colbox	Color of the box which identifies the real data calibration plot.
verbose	If TRUE warn about missing formula and data.
col	Colour of the bars.
xlab	Label for x-axis
labels	Label below the bars. Either "quantiles" or "quantiles.label"
...	Further arguments passed to the subroutine wallyCalPlot and if hanging is TRUE also to subroutine lines.

Value

List of simulated and real data.

Author(s)

Paul F. Blanche <paulfblanche@gmail.com> and Thomas A. Gerds <tag@biostat.ku.dk>

References

Blanche P, Gerds T A, Ekstrom C T (2017). The Wally plot approach to assess the calibration of clinical prediction models, submitted.

Examples

```
# Survival setting
library(prodlim)
library(data.table)
library(survival)
set.seed(180)
d = SimSurv(180)
f = coxph(Surv(time,status)~X1+X2,data=d,x=TRUE)
## Not run:
wallyPlot(f,
          time=4,
          q=10,
          data=d,
          formula=Surv(time,status)~1)
```

```
 wallyPlot(f,
          time=4,
          q=10,
          hanging=TRUE,
          data=d,
          formula=Surv(time,status)~1)

## End(Not run)

# Competing risks setting
library(prodlim)
library(survival)
library(riskRegression)
set.seed(180)
d2 = SimCompRisk(180)
f2 = CSC(Hist(time,event)~X1+X2,data=d2)
## Not run:
wallyPlot(f2,
          time=5,
          q=3,
          hanging=TRUE,
          data=d2,
          formula=Hist(time,event)~1)


## End(Not run)

# Reproduce Wally plots presented in Blanche et al. (2017)
## Not run:
data(threecity)
wallyPlot(threecity$pi,
time=5,
hanging=TRUE,
formula=Hist(time,status)~1,
data=threecity,
ylim=c(-.1,.25),
seed= 511,
hline.lwd=3,
mar=c(1.01, 4.1, 1.15, 2))

## End(Not run)

## Not run:
data(divat)
wallyPlot(divat$pi,
time=5,
hanging=TRUE,
formula=Hist(time,status)~1,
data=divat,
ylim=c(-.1,.60),
seed= 123459,
hline.lwd=3,
mar=c(1.01, 4.1, 1.15, 2))
```

```
## End(Not run)
```

Index