Package 'regplot'

October 14, 2022

Type Package

Title Enhanced Regression Nomogram Plot

Version 1.1

Date 2020-07-01

Description A function to plot a regression nomogram of regression
objects. Covariate distributions are superimposed on nomogram scales and the plot
can be animated to allow on-the-fly changes to distribution representation and to
enable outcome calculation.

License GPL-2

Imports vioplot, sm, beanplot, survival, graphics, stats, lme4

RoxygenNote 7.0.2

NeedsCompilation no

Author Roger Marshall [aut, cre]

Maintainer Roger Marshall <rj.marshall@auckland.ac.nz>

Repository CRAN

Date/Publication 2020-06-26 05:30:02 UTC

R topics documented:

---

regplot	*Plots a regression nomogram showing covariate distribution*

Description

regplot plots enhanced regression nomograms. Covariate distributions are superimposed on nomogram scales and the plot can be animated to allow on-the-fly changes to distribution representation and to enable interactive outcome calculation.

Usage

```
regplot(
  reg,
  plots = c("density", "boxes"),
  center = TRUE,
  observation = NULL,
  title = NULL,
  points = FALSE,
  failtime = NULL,
  prfail = NULL,
  baseS = NULL,
  odds = FALSE,
  nsamp = 10000,
  showP = TRUE,
  rank = NULL,
  subticks = FALSE,
  interval = NULL,
  clickable = FALSE,
  ...
)
```

Arguments

reg	An R regression object from a regression command (see Details, for allowed regressions)
plots	Specifies type of covariate plot. Default plots=c("density","boxes") specifies density plots for numeric covariates and boxes for factors (see Details for other options).
center	If TRUE the mean values of continuous variables and reference categories of factors are aligned vertically. Otherwise continuous distributions are vertically aligned at zero together with reference categories of factors.
observation	To superimpose an observation, shown in (default) red. If TRUE superimposes an observation that is first row of the data used to build reg. Otherwise it may be a specified as any row of reg data or as a dataframe conforming to the structure of the regression data. FALSE omits any superposition.
title	A string title for the plot. If omitted the regression object name and class are output.
points	If FALSE the regression scores of each βx contribution are shown. Otherwise contributions are represented by a 0-100 "points" scale.
failtime	For survival models only, otherwise ignored. Used to specify cutoff times for risk probabilities or for quantiles of survival time. For the former failtime=c(5,10), for example, specifies two probability scales for survival to 5 and 10 time units while failtime=c("50%","10%") specifies scales for 50% and 10% quantiles. If failtime is omitted or NULL, a probability scale for a cutoff that is the median of the time variable is adopted. .
prfail	For survival models, otherwise ignored. If TRUE the probability scale is of failure before failtime, otherwise after failtime.

baseS	For coxph and cph regressions only. If non-null, it specifies the baseline survival probability, for a non-centered model, corresponding to value(s) of `failtime`. If `NULL` the baseline probability is established from the regression object `reg`. Specifying `baseS` can be used coerce alternative baselines.
odds	For probability outcomes, the nomogram scale is of odds rather than probability.
nsamp	The size of a random sub-sample of data to plot covariate distributions (as plotting huge data may be slow and graphical precision, beyond a certain point, unnecessary).
showP	Whether P-value asterisk codes are to be displayed. For factors, the code for the most highly significant level is shown.
rank	Positions the nomogram scales by importance, top down. Two options: `rank="range"` is by the range of the βx's, and `rank="sd"` is by the standard deviation of the βx's. If `NULL` nomogram scales are arranged by order of main effects in the formula, and with interactions at top of the page.
subticks	Puts minor tick marks on axes, where possible.
interval	Draws 95% confidence and prediction intervals. Values `"confidence"`, or `"prediction"`, place intervals on a calculated outcome for a specified `observation` (if `observation` is non `NULL`). A value `"coefficients"` draws confidence intervals on βx for some values of x.
clickable	`TRUE` if the graphic is active for on-the-fly mouse input (see Details).
...	Additional graphics control parameters for font sizes, colours, layout (see Details).

Details

Creates a nomogram representation of a fitted regression. The regression object `reg` can be of different types from the `stats`, `survival` , `rms`, `MASS` and `lme4` libraries. Specifically models generated by the commands: `glm`, `Glm`, `lm`, `ols`, `lrm`, `survreg`, `psm`, `coxph`, `cph`, `glm.nb`, `polr` or mixed model regressions `lmer`, `glmer`, and `glmer.nb`. For `glm`, `Glm` and `glmer` the supported family/link pairings are: gaussian/identity, binomial/logit, quasibinomial/logit, poisson/log and quasipoisson/log. For ordinal regression, using `polr`, logit and probit models are supported. For `survreg` and `psm` the distribution may be lognormal, gaussian, weibull, exponential or loglogistic. For `glm.nb` (from package `MASS`) and `glmer.nb` only log-link is allowed.

The plot can be made active for mouse input if `clickable=TRUE` so allowing on-the-fly changes to distribution plot type (frequency boxes, bars, spikes, box plot, density, empirical cdf, violin and bean plots). These options are presented by a temporary heading menu bar. Individual plots may be selected. Also values of `observation` (if non-null) can be changed by clicking new values, effectively making `regplot` a interactive regression calculator.

The `plots` parameter specifies initial plot types. It is length 2. The first item specifies a plot type for non-factor variables as one of: `"no plot"`, `"density"`, `"boxes"`, `"spikes"`, `"ecdf"`, `"bars"`, `"boxplot"`, `"violin"` or `"bean"`. The second item, is for factors and is one of: `"no plot"`, `"boxes"`, `"bars"` or `"spikes"`.

The graphic shows a scale for all main effects in the regression formula. Interactions are shown by separate nomogram scales. Factor-by-factor interactions are considered as factors and displayed with factor combinations. Factor-by-numeric interactions are displayed for the scale of the numeric

variable(s) and separate scale for each factor level. Numeric-by-numeric interactions are shown with respect to the interaction product scale.

For random effects models (`lmer` and `glmer`) an additional random effects scale is included.

If models are stratified, by a `strata()` (or `strat()` for `rms` models) in the model formula, the behaviour differs depending on the model class. For survival models each stratum has its own outcome scale, otherwise it is included as a term in the linear score with a its own nomogram scale.

If a model formula includes a function (e.g `log()` or a spline `rcs()`) a thumbnail plot of the shape of the transformation is placed on the right of the nomogram scale. It can be toggled on and off by clicking on it (if `clickable=TRUE`).

Additional ... parameters may include items to control the look of the plot if users wish to change default settings: `dencol` colour fill of density plots and other representations of numeric data, `boxcol` fill of factor/logical frequency boxes, `obscol` colour of superimposed observation, `spkcol` colour of spikes. Also font sizes can be set: `cexscales` for font size of points and nomogram scales, `cexvars` for variable names, `cexcats` for category and variable values. To label scales immediately adjacent to the scale (not on the left) use `leftlabel=FALSE`. To draw dotted vertical lines to show more clearly score contributions to an observation use `droplines=TRUE`.

Value

If `points=TRUE`, an object is returned that is a list of dataframes, each frame giving points per covariate, and the last on the list a total points-to-outcome look-up table.

Author(s)

Roger Marshall <rj.marshall@auckland.ac.nz> School of Population Health, The University of Auckland, New Zealand

Examples

```
## 1.  Simulation
n <-500
X <- cbind(rnorm(n, sd = 1),rnorm(n, sd = 0.5))
## make  outcome Y  with intercept 10 + random variation
Y <- 10 + X %*% c(0.2, 0.1) + rnorm(n, sd = 1)
D <- as.data.frame(cbind( Y,X)); colnames(D) <- c("Y","x1","x2")
model <- lm( Y ~ x1 + x2, data = D)
regplot(model, observation = D[1,], interval = "confidence")
## 2 Survival model for pbc data
library(survival)
data(pbc)
pbccox <-  coxph(formula = Surv(time,status==2) ~ age
                 + cut(bili,breaks=c(-Inf, 2, 4, Inf)) + sex
                 + copper +as.factor(stage),data=pbc)
regplot(pbccox,observation=pbc[1,], clickable=TRUE,
        points=TRUE, rank="sd",failtime = c(730,1825))
```

Index