Package ‘gof’
March 15, 2020

Type Package
Title Model Diagnostics Based on Cumulative Residuals
Version 1.0.1
Date 2020-03-15
Author Klaus K. Holst
Maintainer Klaus. K. Holst <klaus@holst.it>
Description Implementation of model-checking techniques for generalized linear
models and linear structural equation models based on cumulative residuals
Depends R (>= 3.3), lava (>= 1.6.3)
Imports methods, mets (>= 1.0.0), Rcpp (>= 1.0.0),
Suggests futile.logger, testthat (>= 0.11), numDeriv
License GPL (>= 2)
LinkingTo Rcpp, RcppArmadillo
LazyLoad yes
NeedsCompilation yes
SystemRequirements C++11
ByteCompile yes
RcppModules gofmod
Encoding UTF-8
RoxygenNote 7.0.2
Repository CRAN
Date/Publication 2020-03-15 22:20:08 UTC

R topics documented:

gof-package .......................................................... 2
cumres.glm .......................................................... 2
cumres.lvmfit ...................................................... 4
plot.cumres ......................................................... 6
surgunit ............................................................ 7
Index 8

gof-package  Model-diagnostics based on cumulative residuals

Description
Model-diagnostics based on cumulative residuals

Author(s)
Klaus K. Holst <klaus@holst.it>

References

See Also
cox.aalen in the timereg-package for similar GoF-methods for survival-data.

Examples
example(cumres)

cumres.glm  Calculates GoF statistics based on cumulative residual processes

Description
Given the generalized linear models model

\[ g(E(Y_i|X_{i1}, ..., X_{ik})) = \sum_{i=1}^{k} \beta_j X_{ij} \]

the cumres-function calculates the observed cumulative sum of residual process, cumulating the residuals, \(e_i\), by the jth covariate:

\[ W_j(t) = n^{-1/2} \sum_{i=1}^{n} 1\{X_{ij} < t\} e_i \]

and Sup and L2 test statistics are calculated via simulation from the asymptotic distribution of the cumulative residual process under the null (Lin et al., 2002).
Usage

```r
## S3 method for class 'glm'
cumres(
  model,
  variable = c("predicted", colnames(model.matrix(model))),
  data = data.frame(model.matrix(model)),
  R = 1000,
  b = 0,
  plots = min(R, 50),
...
)
```

Arguments

- `model`: Model object (`lm` or `glm`)
- `variable`: List of variable to order the residuals after
- `data`: data.frame used to fit model (complete cases)
- `R`: Number of samples used in simulation
- `b`: Moving average bandwidth (0 corresponds to infinity = standard cumulated residuals)
- `plots`: Number of realizations to save for use in the plot-routine
- `...`: additional arguments

Value

Returns an object of class `cumres`.

Note

Currently linear (normal), logistic and poisson regression models with canonical links are supported.

Author(s)

Klaus K. Holst

References


See Also

- `cox.aalen` in the timereg-package for similar GoF-methods for survival-data.
Examples

```r
cumres.lvmfit

Examples

```

```
sim1 <- function(n=100, f=function(x1,x2) {10+x1+x2^2}, sd=1, seed=1) {
  if (!is.null(seed))
    set.seed(seed)
  x1 <- rnorm(n);
  x2 <- rnorm(n)
  X <- cbind(1,x1,x2)
  y <- f(x1,x2) + rnorm(n, sd=sd)
  d <- data.frame(y,x1,x2)
  return(d)
}

```r
d <- sim1(100); l <- lm(y ~ x1 + x2, d)
system.time(g <- cumres(l, R=100, plots=50))
g
plot(g)
g1 <- cumres(l, c("y"), R=100, plots=50)
g1
g2 <- cumres(l, c("y"), R=100, plots=50, b=0.5)
g2
```

---

cumres.lvmfit  
*Cumulative residual processes for structural equation models*

Description

Calculates GoF statistics based on cumulative residual processes for structural equation models fitted with the lava package.

Usage

```r
## S3 method for class 'lvmfit'
cumres(
  model,
  y,
  x,
  full = FALSE,
  data = model.frame(model),
  p,
  R = 1000,
  b = 0,
  plots = min(R, 50),
  ...)
```


Arguments

- **model**: lvm object
- **y**: A formula specifying the association to be checked. Alternatively the outcome specified as a function or a string with the name of the outcome in the model.
- **x**: Predictor. A function, vector or character
- **full**: If FALSE the prediction, Pr, of the variable that are ordered after is only calculated based on the conditional distribution given covariates. If TRUE the conditional expectation is based on the largest set of covariates and endogenous variables such that the residual and Pr are uncorrelated.
- **data**: data.frame (default is the model.frame of the model)
- **p**: Optional parameter vector
- **R**: Number of processes to simulate
- **b**: Moving average parameter
- **plots**: Number of processes to save for use with the plot method
- **...**: Additional arguments parsed on to lower-level functions

Details

With y and x given as functions the user can decide which variables to use in the prediction of the outcome and predictor (use the `predict` method as below).

Value

Returns a `cumres` object with associated plot, print, coef methods

Author(s)

Klaus K. Holst

References


Examples

```r
library(lava)
m <- lvm(list(c(y1,y2,y3)~eta,eta~x)); latent(m) <- ~eta
## simulate some data with non-linear covariate effect
functional(m,eta~x) <- function(x) 0.3*x^2
d <- sim(m,100)
e <- estimate(m,d)
## Checking the functional form of eta on x
g <- cumres(e,eta~x,R=1000)
```
plot(g)

x <- function(p) predict(e,x~y2+y3,p=p)[,"eta"]
## Checking the functional form of y1 on eta
cumres(e,y1~eta,R=1000)

g <- cumres(e,"y1",x=x,R=1000)
plot(g)

plot.cumres

Plot cumulative residuals from a 'cumres' object

Description

plot displays the observed cumulative residual process with realizations under the null.

Usage

## S3 method for class 'cumres'
plot(
x, 
idx = seq_along(x$W),
col = "purple",
col.alpha = 0.3, 
legend = c("type1", "type2", "none"), 
xlab, 
ylab, 
vs = TRUE, 
ylim = NULL, 
title, 
...
)

Arguments

x Object produced by the function cumres.
idx vector of numbers (or variable names) indicating which processes from the x to plot.
col Color of the sample processes. By setting this parameter to "none" or NULL no realizations will be drawn. The number of realizations is determined by the cumres-object.
col.alpha Transparency-level of plotted sample processes
legend Type of legend where "type1" gives p-values of GoF-tests and "type2" gives usual type of legends.
xlab Optional label of x-axis
Optional label of y-axis
Label of predictor
Range of y axis
Main title
Additional arguments passed to the plot-routine.

Author(s)
Klaus K. Holst

Examples

```r
n <- 500; x <- abs(rnorm(n,sd=0.2))+0.01; y <- sqrt(x) + rnorm(n,sd=0.2)
l <- lm(y ~ x)
g <- cumres(l, R=500)
plot(g, idx=1, legend="type2")
```

---

Surgical Unit Data

Surgical Unit Data used in the paper by Lin et al. (2002). Survival time and covariates for 54 patients undergoing liver surgery.

Source

References
Index

*Topic hplot
  plot.cumres, 6
*Topic models
  cumres.glm, 2
  cumres.lvmfit, 4
*Topic package
  gof-package, 2
*Topic regression
  cumres.glm, 2
  cumres.lvmfit, 4
  plot.cumres, 6

cox.aalen, 2, 3
cumres (cumres.glm), 2
cumres.glm, 2
cumres.lvmfit, 4

gof (gof-package), 2
gof-package, 2

plot.cumres, 6

surgunit, 7