

Package ‘PResiduals’

August 22, 2014

Type Package

Title Probability scale residuals and residual correlations.

Version 0.2-0

Date 2014-08-22

Author Charles Dupont, Jeffrey Horner, Chun Li, Bryan Shepherd

Maintainer Charles Dupont <charles.dupont@vanderbilt.edu>

Description Computes probability scale residuals and residual correlations for continuous, ordinal, binary, count, and time-to-event data (currently only implements ordinal data).

Imports Formula, rms, sandwich, actuar

Suggests MASS, survival

License GPL (>= 2)

Collate 'GKGamma.R' 'PResiduals-package.R' 'pgumbel.R' 'newPolr.R' 'diagn.R' 'cobot.R' 'cocobot.R' 'presid.R' 'print.cobot.R' 'print.cocobot.R'

NeedsCompilation no

Repository CRAN

Date/Publication 2014-08-22 22:30:38

R topics documented:

PResiduals-package	2
cobot	2
cocobot	4
diagn	5
GKGamma	6
presid	7
PResidData	9
print.cobot	10
print.cocobot	10

Index**11**

PResiduals-package *Computes probability scale residuals and residual correlations.*

Description

This package outputs probability scale residuals from multiple models and computes residual correlation. Probability scale residual can be computed for continuous, ordinal, binary, count, and time-to-event data (although the current implementation is only for ordinal variables). Plots of probability scale residuals can be useful for model diagnostics. Residual correlation can be used to test for conditional independence between multiple types of variables.

Author(s)

Bryan Shepherd <bryan.shepherd@vanderbilt.edu>

Chun Li <chun.li@vanderbilt.edu>

Charles Dupont <dupontct@vanderbilt.edu>

cobot *Conditional ordinal by ordinal tests for association.*

Description

cobot tests for independence between two ordered categorical variables, X and Y conditional on other variables, Z . The basic approach involves fitting models of X on Z and Y on Z and determining whether there is any remaining information between X and Y . This is done by computing one of 3 test statistics. T1 compares empirical distribution of X and Y with the joint fitted distribution of X and Y under independence conditional on Z . T2 computes the correlation between ordinal (probability scale) residuals from both models and tests the null of no residual correlation. T3 evaluates the concordance–discordance of data drawn from the joint fitted distribution of X and Y under conditional independence with the empirical distribution. Details are given in *Li C and Shepherd BE, Test of association between two ordinal variables while adjusting for covariates. Journal of the American Statistical Association 2010, 105:612-620.*

Usage

```
cobot(formula, link = c("logit", "probit", "cloglog", "cauchit"),
      link.x = link, link.y = link, data, subset, na.action = na.fail,
      fisher = FALSE, conf.int = 0.95)
```

Arguments

formula	an object of class <code>Formula</code> (or one that can be coerced to that class): a symbolic description of the model to be fitted. The details of model specification are given under ‘Details’.
link	The link family to be used for ordinal models of both X and Y . Defaults to ‘logit’. Other options are ‘probit’, ‘cloglog’, and ‘cauchit’.
link.x	The link function to be used for a model of the first ordered variable. Defaults to value of link.
link.y	The link function to be used for a model of the second variable. Defaults to value of link.
data	an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>cobot</code> is called.
subset	an optional vector specifying a subset of observations to be used in the fitting process.
na.action	how NAs are treated.
fisher	logical; if TRUE, Fisher transformation and delta method a used to compute p value for the test statistic based on correlation of residuals.
conf.int	numeric specifying confidence interval coverage.

Details

formula is specified as $X \mid Y \sim Z$. This indicates that models of $X \sim Z$ and $Y \sim Z$ will be fit. The null hypothesis to be tested is $H_0 : X$ independant of Y conditional on Z .

Note that T2 can be thought of as an adjust rank correlation. (Li C and Shepherd BE, *A new residual for ordinal outcomes. Biometrika* 2012; 99:473-480)

Value

object of ‘cobot’ class.

Author(s)

Charles Dupont <charles.dupont@vanderbilt.edu>

Chun Li <chun.li@vanderbilt.edu>

Bryan Shepherd <bryan.shepherd@vanderbilt.edu>

References

Li C and Shepherd BE, Test of association between two ordinal variables while adjusting for covariates. *Journal of the American Statistical Association* 2010, 105:612-620.

Li C and Shepherd BE, A new residual for ordinal outcomes. *Biometrika* 2012; 99:473-480

See Also

[Formula](#), [as.data.frame](#)

Examples

```
data(PResidData)
cocobot(x|y~z, data=PResidData)
```

cocobot

Conditional continuous by ordinal tests for association.

Description

cocobot tests for independence between an ordered categorical variable, X , and a continuous variable, Y , conditional on other variables, Z . The basic approach involves fitting an ordinal model of X on Z , a linear model of Y on Z , and then determining whether there is any residual information between X and Y . This is done by computing residuals for both models, calculating their correlation, and testing the null of no residual correlation. This procedure is analogous to test statistic T2 in `cobot`. Two test statistics (correlations) are currently output. The first is the correlation between probability-scale residuals (PResid). The second is the correlation between the observed-minus-expected residual for the continuous outcome model and a latent variable residual for the ordinal model.

Usage

```
cocobot(formula, data, link = c("logit", "probit", "cloglog", "cauchit"),
        subset, na.action = getOption("na.action"), emp = TRUE, fisher = FALSE,
        conf.int = 0.95)
```

Arguments

formula	an object of class Formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The details of model specification are given under ‘Details’.
link	The link family to be used for the ordinal model of X on Z . Defaults to ‘logit’. Other options are ‘probit’, ‘cloglog’, and ‘cauchit’.
data	an optional data frame, list or environment (or object coercible by as.data.frame to a data frame) containing the variables in the model. If not found in <code>data</code> , the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>cocobot</code> is called.
subset	an optional vector specifying a subset of observations to be used in the fitting process.
na.action	action to take when NA present in data.
emp	logical indicating whether the residuals from the model of Y on Z are computed based on the assumption of normality (FALSE) or empirically (TRUE).

fisher	logical indicating whether to apply fisher transformation to compute confidence intervals and p-values for the correlation.
conf.int	numeric specifying confidence interval coverage.

Details

Formula is specified as $X \mid Y \sim Z$. This indicates that models of $X \sim Z$ and $Y \sim Z$ will be fit. The null hypothesis to be tested is $H_0 : X$ independent of Y conditional on Z . The ordinal variable, X , must precede the \mid and be a factor variable, and Y must be continuous.

Value

object of 'cocobot' class.

References

Li C and Shepherd BE (2012) A new residual for ordinal outcomes. *Biometrika*. **99**: 473–480.

Shepherd BE, Li C, Liu Q (submitted) Probability-scale residuals for continuous, discrete, and censored data.

Examples

```
data(PResidData)
cocobot(y|w ~ z, data=PResidData)
```

diagn	<i>Extract or construct a diagonal matrix.</i>
-------	--

Description

This works like [diag](#) except when x is a single integer value. If x is a single integer value then it assumes that you want a 1 by 1 matrix with the value set to x

Usage

```
diagn(x = 1, nrow = length(x), ncol = nrow)
```

Arguments

x	a matrix, vector or 1D array, or missing.
nrow, ncol	optional dimensions for the result when x is not a matrix.

Value

matrix with diagonal elements set to x

Author(s)

Chun Li

See Also[diag](#)**Examples**

```
diag(5)
diagn(5)
```

 GKGamma

Goodman-Kruskal's γ

Description

Computes Goodman-Kruskal's γ

Usage

```
GKGamma(M)
```

Arguments

M a matrix

Value

scon concordance

sdis disconcordance

gamma a real number between -1 and 1. calculated as $\text{gamma} = \frac{\text{scon} - \text{sdis}}{\text{scon} + \text{sdis}}$

Author(s)

Chun Li

References

Goodman LA, Kruskal WH (1954) Measures of association for cross classifications, Journal of the American Statistical Association, 49, 732-764.

presid

Probability-scale Residual

Description

presid Calculates the probability-scale residual for various model function objects. Currently supported models include `glm` (Poisson, binomial, and gaussian families), `lm` in the **stats** library, `survreg` (Weibull, exponential, gaussian, logistic, and lognormal distributions) and `coxph` in the **survival** library, and `polr` and `glm.nb` in the **MASS** library.

Usage

```
presid(object, ...)
```

Arguments

object	The model object for which the probability-scale residual is calculated
...	Additional arguments passed to methods

Details

Probability-scale residual is $P(Y^* < y) - P(Y^* > y)$ where y is the observed outcome and Y^* is a random variable from the fitted distribution.

Value

The probability scale residual for the model

Author(s)

Charles Dupont <charles.dupont@vanderbilt.edu>

Chun Li <chun.li@vanderbilt.edu>

Bryan Shepherd <bryan.shepherd@vanderbilt.edu>

References

Shepherd BE, Li C, Liu Q. Probability-scale residuals for continuous, discrete, and censored data. Submitted.

Li C and Shepherd BE, A new residual for ordinal outcomes. *Biometrika* 2012; 99:473-480

Examples

```

library(survival)
library(stats)

set.seed(100)
n <- 1000
x <- rnorm(n)
t <- rweibull(n, shape=1/3, scale=exp(x))
c <- rexp(n, 1/3)
y <- pmin(t, c)
d <- ifelse(t<=c, 1, 0)

mod.survreg <- survreg(Surv(y, d) ~ x, dist="weibull")
summary(presid(mod.survreg))
plot(x, presid(mod.survreg))

##### example for proportional hazards model
n <- 1000
x <- rnorm(n)
beta0 <- 1
beta1 <- 0.5
t <- rexp(n, rate = exp(beta0 + beta1*x))
c <- rexp(n, rate=1)
y <- ifelse(t<=c, t, c)
delta <- as.integer(t<=c)

mod.coxph <- coxph(Surv(y, delta) ~ x)
presid <- presid(mod.coxph)
plot(x, presid, cex=0.4, col=delta+2)

#### example for Negative Binomial regression
library(MASS)

n <- 1000
beta0 <- 1
beta1 <- 0.5
x <- runif(n, min=-3, max=3)
y <- rnbinom(n, mu=exp(beta0 + beta1*x), size=3)

mod.glm.nb <- glm.nb(y~x)
presid <- presid(mod.glm.nb)
summary(presid)
plot(x, presid, cex=0.4)

##### example for proportional odds model
library(MASS)

n <- 1000
x <- rnorm(n)
y <- numeric(n)
alpha = c(-1, 0, 1, 2)
beta <- 1

```



```
py <- (1 + exp(- outer(alpha, beta*x, "+"))) ^ (-1)
aa = runif(n)
for(i in 1:n)
  y[i] = sum(aa[i] > py[,i])
y <- as.factor(y)

mod.polr <- polr(y~x, method="logistic")
summary(mod.polr)
presid <- presid(mod.polr)
summary(presid)
plot(x, presid, cex=0.4)
```

PResidData

PResidual Example Dataset

Description

An Example Dataset for use with PResidual's cobot

Usage

```
data(PResidData)
```

Format

A data frame with 500 observations on the following 3 variables.

x an ordered factor with levels 1 < 2 < 3 < 4 < 5

y an ordered factor with levels 1 < 2 < 3 < 4

z a numeric vector

w a numeric vector

Examples

```
data(PResidData)
```

print.cobot *cobot class print method*

Description

cobot class print method

Usage

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

Arguments

x cobot object
... arguments passed to print.default

Author(s)

Charles Dupont

print.cocobot *cocobot class print method*

Description

cocobot class print method

Usage

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

Arguments

x cocobot object
... arguments passed to print.default

Author(s)

Charles Dupont

Index

- *Topic **array**
 - diag, [5](#)
- *Topic **datasets**
 - PResidData, [9](#)
- *Topic **package**
 - PResiduals-package, [2](#)
- *Topic **print**
 - print.cobot, [10](#)
 - print.cocobot, [10](#)

- as.data.frame, [3, 4](#)

- cobot, [2](#)
- cocobot, [4](#)
- coxph, [7](#)

- diag, [5, 6](#)
- diagn, [5](#)

- Formula, [3, 4](#)

- GKGamma, [6](#)
- glm, [7](#)
- glm.nb, [7](#)

- lm, [7](#)

- polr, [7](#)
- presid, [7](#)
- PResidData, [9](#)
- PResiduals (PResiduals-package), [2](#)
- PResiduals-package, [2](#)
- print.cobot, [10](#)
- print.cocobot, [10](#)

- survreg, [7](#)