Package 'PAsso'

Type Package

Title Assessing the Partial Association Between Ordinal Variables

Description An implementation of the unified framework for assessing partial
association between ordinal variables after adjusting a set of covariates
(Dungang Liu, Shaobo Li, Yan Yu and Irini Moustaki (2020). Accepted by JASA).
This package provides a set of tools to quantify partial association, conduct
hypothesis testing for partial association, visualize partial regression models,
and diagnose the specifications of each fitted model. This framework is
based on the surrogate approach de-
scribed in Liu and Zhang (2017) (<doi:10.1080/01621459.2017.1292915>).

Version 0.1.8

Maintainer Xiaorui (Jeremy) Zhu <zhuxiaorui1989@gmail.com>

URL GitHub: https://github.com/XiaoruiZhu/PAsso

BugReports https://github.com/XiaoruiZhu/PAsso/issues

Depends R (>= 3.5.0), stats (>= 3.5.0), ggplot2 (>= 2.2.1), dplyr

License GPL (>= 2)

Encoding UTF-8

LazyData true

Imports VGAM, copBasic, pcaPP (>= 1.9-73), methods, doParallel (>=
1.0.11), foreach (>= 1.4.8), tidyverse, MASS (>= 7.3-51.0),
GGally, goftest, gridExtra, utils (>= 3.5.3), progress (>=
1.2.0), plotly, copula

Suggests faraway, ordinal, rms, testthat, mgcv, PResiduals, knitr,
rmarkdown

LinkingTo Rcpp

RoxygenNote 7.1.0

NeedsCompilation yes

Author Xiaorui (Jeremy) Zhu [aut, cre],
Shaobo Li [aut],
Yuejie Chen [ctb],
Dungang Liu [ctb]

R topics documented:

ANES2016 *The 2016 ANES Time Series Study with pre-election interview*

Description

A subset of 2,188 participants of the 2016 American National Election Time Series Study, which was to track the enduring social trend and record the political moment of 2016 (DeBell, 2018). This study consisted of two surveys with same population. The pre-election interview was during the weeks before the 2016 general election, including 4,271 respondents in total. The post-election interview is the re-interview during the weeks after the election, including 3,649 respondents (662 respondents did not complete post-interviews).

Usage

```
data(ANES2016)
```

Format

A data frame with 2188 rows and 10 variables.

- age Respondent's age in years.

- edu.year Respondent's education year, which is mapped from education level: MS=8, HSdrop=11, HS=12, Coll=14, CCdeg=15, BAdeg=17, MAdeg=19.

- education Respondent's education level.

- income.num Respondent's family income in thousands: an numerical variable. It is median value of the range of each income level

- income Respondent's family income level:
  '(01) 01. Under $5,000' = 5,
  '(02) 02. $5,000-$9,999',
  '(03) 03. $10,000-$12,499',
  '(04) 04. $12,500-$14,999',
  '(05) 05. $15,000-$17,499',
  '(06) 06. $17,500-$19,999',
  '(07) 07. $20,000-$22,499',
  '(08) 08. $22,500-$24,999',
  '(09) 09. $25,000-$27,499',
  '(10) 10. $27,500-$29,999',
  '(11) 11. $30,000-$34,999',
  '(12) 12. $35,000-$39,999',
  '(13) 13. $40,000-$44,999',
  '(14) 14. $45,000-$49,999',
  '(15) 15. $50,000-$54,999',
  '(16) 16. $55,000-$59,999',
  '(17) 17. $60,000-$64,999',
  '(18) 18. $65,000-$69,999',
  '(19) 19. $70,000-$74,999',
  '(20) 20. $75,000-$79,999',
  '(21) 21. $80,000-$89,999',
  '(22) 22. $90,000-$99,999',
  '(23) 23. $100,000-$109,999',
  '(24) 24. $110,000-$124,999',
  '(25) 25. $125,000-$149,999',
  '(26) 26. $150,000-$174,999',
  '(27) 27. $175,000-$249,999',
  '(28) 28. $250,000 or more' = 250.
- PID Party identification: a numeric variable with value from 1 to 7 representing strong Democrat, strDem < weak Democrat, weakDem < independent Democrat, indDem < independent independent indind < independent Republican, indRep < weak Republican, weakRep < strong Republican, strRep.
- selfLR The respondent' self-placement about own left-right in 7 ordinal levels (from extremely liberal to extremely conservative). extLib: extremely liberal, Lib: liberal, sliLib: slightly liberal, Mod: moderate, sliCon: slightly conservative, Con: conservative, extCon: extremely conservative. extLib < Lib < sliLib < Mod < sliCon < Con < extCon.
- TrumpLR The respondent's opinion about Donald Trump's left-right placement (same scale as selfLR).
- ClinLR The respondent's opinion about Hilary Clinton's left-right placement (same scale as selfLR).
- PreVote The respondent's voting preference between Donald Trump and Hilary Clinton two months preceeding the November election (Pre-election interview). It is a factor with levels HillaryClinton and DonaldTrump.

- PreVote.num Recode the PreVote to numeric values, 'HillaryClinton'=0, 'DonaldTrump'=1.
- WeightforPreVote Pre-election weight of a respondent.

Details

The Pre-election preference is recorded as "PreVote" and the "PreVote.num" is the numeric of it. Observations with missing values, or "No thought" responses have been removed. Respondents expressing a voting preference other than Clinton or Trump have been removed.

References

DeBell, Matthew, Jon A. Krosnick, Katie Gera, David S. Yeager, and Michael P. McDonald. The turnout gap in surveys: Explanations and solutions. *Sociological Methods & Research*, 2018: 0049124118769085.

Enamorado, Ted, Benjamin Fifield, and Kosuke Imai. *User's guide and codebook for the ANES 2016 time series voter validation supplemental data*. Tech. rep., American National Election Studies, 2018.

Examples

```
head(ANES2016)
```

---

diagnostic.plot   *Residual-based diagnostic plots*

Description

A set of visualization tools for the diagnostic of the fitted model in the partial association analysis. It can provides a plot matrix including Q-Q plots, residual-vs-fitted plots, residual-vs-covariate plots of all the fitted models. This function also support the direct diagnostic of the cumulative link regression model in the class of clm, glm, lrm, orm, polr. Currently, vglm is not supported.

Usage

```
diagnostic.plot(object, ...)

## Default S3 method:
diagnostic.plot(object, ...)

## S3 method for class 'resid'
diagnostic.plot(object, output = c("qq", "fitted", "covariate"), ...)

## S3 method for class 'PAsso'
diagnostic.plot(
  object,
  output = c("qq", "fitted", "covariate"),
  model_id = NULL,
```

```
  x_name = NULL,
  ...
)

## S3 method for class 'glm'
diagnostic.plot(
  object,
  output = c("qq", "fitted", "covariate"),
  x = NULL,
  fit = NULL,
  distribution = qnorm,
  ncol = NULL,
  alpha = 1,
  xlab = NULL,
  color = "#444444",
  shape = 19,
  size = 2,
  qqpoint.color = "#444444",
  qqpoint.shape = 19,
  qqpoint.size = 2,
  qqline.color = "#888888",
  qqline.linetype = "dashed",
  qqline.size = 1,
  smooth = TRUE,
  smooth.color = "red",
  smooth.linetype = 1,
  smooth.size = 1,
  fill = NULL,
  resp_name = NULL,
  ...
)

## S3 method for class 'clm'
diagnostic.plot(object, output = c("qq", "fitted", "covariate"), ...)

## S3 method for class 'lrm'
diagnostic.plot(object, output = c("qq", "fitted", "covariate"), ...)

## S3 method for class 'orm'
diagnostic.plot(object, output = c("qq", "fitted", "covariate"), ...)

## S3 method for class 'polr'
diagnostic.plot(object, output = c("qq", "fitted", "covariate"), ...)
```

Arguments

`object`	The object in the support classes (This function is mainly designed for PAsso).
`...`	Additional optional arguments can be passed onto [ggplot](#) for drawing various

	plots.
output	A character string specifying what type of output to plot. Default is "qq" which produces a plot matrix with quantile-quantile plots of the residuals. "fitted" produces a plot matrix between residuals and all corresponding fitted responses. "covariates" produces a plot matrix between residuals and corresponding covariate.
model_id	A number refers to the index of the model that needs to be diagnosed. If NULL, all models will be diagnosed.
x_name	A string refers to the covariate name that needs to be diagnosed. If NULL, all adjustments will be diagnosed.
x	A vector giving the covariate values to use for residual-by- covariate plots (i.e., when output = "covariate").
fit	The fitted model from which the residuals were extracted. (Only required if output = "fitted" and object inherits from class "resid".)
distribution	Function that computes the quantiles for the reference distribution to use in the quantile-quantile plot. Default is qnorm which is only appropriate for models using a probit link function. When jitter.scale = "probability", the reference distribution is always U(-0.5, 0.5). (Only required if object inherits from class "resid".)
ncol	Integer specifying the number of columns to use for the plot layout (if requesting multiple plots). Default is NULL.
alpha	A single values in the interval [0, 1] controlling the opacity alpha of the plotted points. Only used when nsim > 1.
xlab	Character string giving the text to use for the x-axis label in residual-by-covariate plots. Default is NULL.
color	Character string or integer specifying what color to use for the points in the residual vs fitted value/covariate plot. Default is "black".
shape	Integer or single character specifying a symbol to be used for plotting the points in the residual vs fitted value/covariate plot.
size	Numeric value specifying the size to use for the points in the residual vs fitted value/covariate plot.
qqpoint.color	Character string or integer specifying what color to use for the points in the quantile-quantile plot.
qqpoint.shape	Integer or single character specifying a symbol to be used for plotting the points in the quantile-quantile plot.
qqpoint.size	Numeric value specifying the size to use for the points in the quantile-quantile plot.
qqline.color	Character string or integer specifying what color to use for the points in the quantile-quantile plot.
qqline.linetype	
	Integer or character string (e.g., "dashed") specifying the type of line to use in the quantile-quantile plot.
qqline.size	Numeric value specifying the thickness of the line in the quantile-quantile plot.

smooth	Logical indicating whether or not too add a nonparametric smooth to certain plots. Default is TRUE.
smooth.color	Character string or integer specifying what color to use for the nonparametric smooth.
smooth.linetype	
	Integer or character string (e.g., ″dashed″) specifying the type of line to use for the nonparametric smooth.
smooth.size	Numeric value specifying the thickness of the line for the nonparametric smooth.
fill	Character string or integer specifying the color to use to fill the boxplots for residual-by-covariate plots when x is of class ″factor″. Default is NULL which colors the boxplots according to the factor levels.
resp_name	Character string to specify the response name that will be displayed in the figure.

Value

A ″ggplot″ object for supported models. For class "PAsso", it returns a plot in ″gtable″ object that combines diagnostic plots of all responses.

A "ggplot" object based on the input residuals.

A "ggplot" object based on the input residuals.

A plot in "gtable" object that combines diagnostic plots of all responses.

A "ggplot" object based on the residuals generated from glm object.

A "ggplot" object based on the residuals generated from clm object.

A "ggplot" object based on the residuals generated from lrm object.

A "ggplot" object based on the residuals generated from orm object.

A "ggplot" object based on the residuals generated from polr object.

Examples

```
# Import data for partial association analysis
data(″ANES2016″)
ANES2016$PreVote.num <- as.factor(ANES2016$PreVote.num)

PAsso_3v <- PAsso(responses = c(″PreVote.num″, ″PID″, ″selfLR″),
                  adjustments = c(″income.num″, ″age″, ″edu.year″),
                  data = ANES2016, uni.model = ″probit″,
                  method = c(″kendall″),
                  resids.type = ″surrogate″, jitter = ″latent″)

diag_p1 <- diagnostic.plot(object = PAsso_3v, output = ″qq″)
diag_p2 <- diagnostic.plot(object = PAsso_3v, output = ″fitted″)
diag_p3 <- diagnostic.plot(object = PAsso_3v, output = ″covariate″)

# Simply diagnose a model
# Fit cumulative link models

fit1 <- ordinal::clm(PreVote.num ~ income.num + age + edu.year, data = ANES2016, link = ″logit″)
```

```
# diagnostic.plot
plot_qq_1 <- diagnostic.plot(object = fit1, output = "qq")
plot_fit_1 <- diagnostic.plot(object = fit1, output = "fitted")
plot_cov_1 <- diagnostic.plot(object = fit1, output = "covariate")
```

PAsso   *Partial association analysis between ordinal responses after adjusting*
                             for a set of covariates

Description

This function is mainly designed for conducting the partial association analysis. It provides two
ways of using:

1. A user-friendly way: only need "responses", "adjustments", and "data". All the rest of the
argument will be setted as default (see Arguments for details of default).

2. An advanced way: user can input a list of fitted models by "fitted.models", then the "responses"
and "adjustments" are not necessary. Supported class of cumulative link models in clm, glm, lrm,
orm, polr, vglm, .

It generates an object that has partial association matrix, marginal association, and some attributes:
"arguments" saves c(association, method, resids.type). "responses" contains the names of response
variables. The attribute "adjustments" contains the names of covariates. The "summary" function
of "PAsso" class of object provides marginal association ' matrix for comparison purpose.

Usage

```
PAsso(
  responses,
  adjustments,
  data,
  uni.model = c("probit", "logit", "acat"),
  models = NULL,
  method = c("kendall", "pearson", "wolfsigma"),
  resids.type = c("surrogate", "sign", "general", "deviance"),
  jitter = c("latent", "uniform"),
  jitter.uniform.scale = c("probability", "response"),
  fitted.models = NULL,
  n_draws = 20,
  association = "partial",
  ...
)
```

Arguments

responses   A string vector that specifies response variables. It requires to be equal or greater
                     than two variables in the data frame.

adjustments	A string vector specifies covariates/confounders that need to be adjusted.
data	A data.frame including responses and adjustments.
uni.model	A character string specifying the universal model setting for all responses. Default ″logit″ refers to cumulative logit model. ″probit″ refers to cumulative probit model. ″acat″ fits an adjacent categories regression model.
models	A string vector contains default link functions of fitting models with respect to each response variable. If ″models″ is missing or has any one of the model unspecified, ″uni.model″ is used to specify same models for all responses automatically. But, this argument has higher priority than the ″uni.model″ as long as the length of ″models″ equals to the number of ″responses″.
method	A string argument to specify correlation coefficient method. Three choices c(″kendall″,″pearson″,″wolfsigma″). The default is ″kendall″
resids.type	A character string specifying which type of residuals to generate Current options are ″latent″ and ″uniform″. Default is ″latent″.
	surrogate surrogate residuals (Liu and Zhang, 2017);
	sign sign-based residuals (Li and Shepherd, 2010, 2012);
	general generalized residuals (Franses and Paap, 2001);
	deviance deviance residuals (-2*loglik).
	Although ″sign″, ″general″, and ″deviance″ are provided in this package, these residuals are problematic for partial association analysis between ordinal response (more discussions see Liu, Dungang, Li, Shaobo, Yu, Yan, and Moustaki, Irini.(2020))
jitter	A character string specifying how to generate surrogate residuals. Current options are ″latent″ and ″uniform″. Default is ″latent″.
	latent surrogate residuals.
	uniform sign-based residuals.
jitter.uniform.scale	
	A character string specifying the scale on which to perform the jittering whenever jitter = ″uniform″. More details: PAsso::residuals.
fitted.models	A list contains all the models (S3 objects) you want to assess for the partial association between ordinal responses after adjusting for a set of covariates covariates. All of these models should be applied to the same dataset, having same covariates, same sample size etc. The models in this list can be an object of class [clm](), [glm](), [lrm](), [orm](), [polr](), [vglm]().
n_draws	A number to specify draws of surrogate residuls such that the partial correlation coefficients are calculated repeatedly. The final correlation coefficients are the average of all partial correlation coefficients. It is the ″nsim″ argument in ″residuals()″ function.
association	An default argument to specify the partial association. Leave this further development of package such that other association analyses can be embedded.
...	Additional optional arguments.

Value

An object of class "PAsso" is a list containing at least the following components. It contains the partial correlation matrix and multiple repeats if n_draws > 1. This object has "arguments" attribute saved as c(association, method, resids.type), "responses" attribute, and "adjustments" attribute. The list contains:

corr The estimated correlation matrix(average of rep_MatCorr) of partial association after adjusting confounders;

rep_corr The replications of estimated correlation matrix;

rep_SRs The replications of surrogate residuals if partial association is applied;

fitted.models The list stores all fitted.models;

data The data.frame of original dataset;

mods_n The sample size of each fitted model;

cor_func The correlation function after assign different method;

Marg_corr The marginal association matrix.

References

Liu, Dungang, Li, Shaobo, Yu, Yan, and Moustaki, Irini. Assessing partial association between ordinal variables: quantification, visualization, and hypothesis testing, *Journal of the American Statistical Association*, Revision under review.

Liu, Dungang and Zhang, Heping. Residuals and Diagnostics for Ordinal Regression Models: A Surrogate Approach. *Journal of the American Statistical Association*. http://www.tandfonline.com/doi/abs/10.1080/01621459.2017.1292915?journalCode=uasa20

Li, Chun, and Bryan E. Shepherd. "Test of association between two ordinal variables while adjusting for covariates." *Journal of the American Statistical Association* 105, no. 490 (2010): 612-620. https://doi.org/10.1198/jasa.2010.tm09386

Li, Chun, and Bryan E. Shepherd. "A new residual for ordinal outcomes." *Biometrika* 99, no. 2 (2012): 473-480. https://doi.org/10.1093/biomet/asr073

Franses, Philip Hans, and Richard Paap. Quantitative models in marketing research. Cambridge University Press, 2001. https://pdfs.semanticscholar.org/dad0/820f287a8cf5a4e8039549e35fc111fd86e5.pdf

Examples

```
###########################################################
# User-friendly way of using
###########################################################
library(MASS)

# Import ANES2016 data in "PAsso"
data(ANES2016)

# User-friendly way of using: Parial association analysis
PAsso_1 <- PAsso(responses = c("PreVote.num", "PID"),
```

```
                  adjustments = c("income.num", "age", "edu.year"),
                  data = ANES2016,
                  method = c("kendall"))

print(PAsso_1, digits = 4)
summary(PAsso_1, digits = 4)


############################################################
# Advanced way of using
############################################################

fit.vote<- glm(PreVote.num ~ income.num+ age + edu.year, data = ANES2016,
                family = binomial(link = "probit"))
fit.PID<- polr(as.factor(PID) ~ income.num+age+edu.year, data = ANES2016,
                method="probit", Hess = TRUE)

PAsso_adv1 <- PAsso(fitted.models=list(fit.vote, fit.PID),
                    method = c("kendall"),
                    resids.type = "surrogate")

print(PAsso_adv1, digits = 4)
summary(PAsso_adv1, digits = 4)
```

| plot | *A matrix of partial regression plots between responses after adjust-ments* |

Description

A plot matrix to display the results of partial association analyses. Upper-triangle contains scatter-plot matrix between each pair of response variables. Lower-triangle contains the partial correlation coefficients adjusted by covariates.

Usage

```
## S3 method for class 'PAsso'
plot(x, color = "#444444", ...)
```

Arguments

x	the object in "PAsso" class that is generated by "PAsso" or "test".
color	The color of points.
...	Additional optional arguments to be passed onto.

Details

A pairwise plot matrix reveals the partial association between ordinal variables. All the plots are based on surrogate residuals generated from "resides" function. Graphics are designed based on ggplot2 and "GGally".

Value

A "GGally" object.

Examples

```
data(ANES2016)

summary(ANES2016)

PAsso_2v <- PAsso(responses = c("PreVote.num", "PID"),
                  adjustments = c("income.num", "age", "edu.year"),
                  data = ANES2016)

plot(PAsso_2v)
```

---

plot3D   *3-D P-P plot for the inspection of the partial association analysis*

Description

A list of 3-D P-P plots for the inspection of the partial association analysis. Each plot is the 3-D P-P plot from an empirical copula trained from the surrogate residuals of a pair of responses.

Usage

```
plot3D(object, y1, y2, ...)

## Default S3 method:
plot3D(object, y1, y2, ...)

## S3 method for class 'PAsso'
plot3D(object, y1, y2, ...)
```

Arguments

object	A PAsso class of object.
y1	A string to specify the first response for the 3D plot.
y2	A string to specify the second response for the 3D plot. If either one of the y1 or y2 is missing. The plot3D will draw 3D plots for all pairs of responses.
...	Additional optional arguments.

Details

All the plots are based on surrogate residuals generated from "resides" function in sure. Graphics are designed based on PAsso and "plotly".

Value

If response y1 or y2 is not specified, a list of ″plotly″ objects includes all pairs of responses will be returned (with name "response 1 v.s. response 2" etc.). If responses y1 and y2 are specified, returns a 3D plot as ″plotly″ object.

Examples

```
# Did not run this to save time
# data("ANES2016")
# PAsso_3v <- PAsso(responses = c("PreVote.num", "PID", "selfLR"),
#                   adjustments = c("income.num", "age", "edu.year"),
#                   data = ANES2016)

# plot3D(PAsso_3v, y1="PID", y2="selfLR")
```

---

print   *Print partial association matrix*

Description

Print partial association matrix

Usage

```
## S3 method for class 'PAsso'
print(x, digits = max(2, getOption("digits") - 2), ...)

## S3 method for class 'PAsso.test'
print(x, digits = max(3L, getOption("digits") - 3L), ...)
```

Arguments

x	A PAsso object for printing out results.
digits	A default number to specify decimal digit values.
...	Additional optional arguments.

Value

Print partial association matrix of a PAsso object

Examples

```
# See PAsso for the example.
```

---

residuals   *Extract Model Residuals*

Description

A generic function to simulate surrogate residuals for cumulative link regression models using the latent method described in Liu and Zhang (2017).

It also support the sign-based residuals (Li and Shepherd, 2010), generalized residuals (Franses and Paap, 2001), and deviance residuals for cumulative link regression models.

Usage

```
## S3 method for class 'clm'
residuals(
  object,
  type = c("surrogate", "sign", "general", "deviance"),
  jitter = c("latent", "uniform"),
  jitter.uniform.scale = c("probability", "response"),
  nsim = 1L,
  ...
)

## S3 method for class 'lrm'
residuals(
  object,
  type = c("surrogate", "sign", "general", "deviance"),
  jitter = c("latent", "uniform"),
  jitter.uniform.scale = c("probability", "response"),
  nsim = 1L,
  ...
)

## S3 method for class 'orm'
residuals(
  object,
  type = c("surrogate", "sign", "general", "deviance"),
  jitter = c("latent", "uniform"),
  jitter.uniform.scale = c("probability", "response"),
  nsim = 1L,
  ...
)

## S3 method for class 'polr'
residuals(
  object,
  type = c("surrogate", "sign", "general", "deviance"),
  jitter = c("latent", "uniform"),
```

```
    jitter.uniform.scale = c("probability", "response"),
    nsim = 1L,
    ...
)

## S3 method for class 'vglm'
residuals(
  object,
  type = c("surrogate", "sign", "general", "deviance"),
  jitter = c("latent", "uniform"),
  jitter.uniform.scale = c("probability", "response"),
  nsim = 1L,
  ...
)

## S3 method for class 'ord'
residuals(
  object,
  type = c("surrogate", "sign", "general", "deviance", "pearson", "working",
    "response", "partial"),
  jitter = c("latent", "uniform"),
  jitter.uniform.scale = c("probability", "response"),
  nsim = 1L,
  ...
)

## S3 method for class 'PAsso'
residuals(object, draw_id = 1, ...)
```

Arguments

`object`	An object of class PAsso.
`type`	The type of residuals which should be returned. The alternatives are: "surrogate" (default), "sign", "general", and "deviance". Can be abbreviated.
	surrogate surrogate residuals (Liu and Zhang, 2017);
	sign sign-based residuals;
	general generalized residuals (Franses and Paap, 2001);
	deviance deviance residuals (-2*loglik).
`jitter`	A character string specifying which method to use to generate the surrogate response values. Current options are `"latent"` and `"uniform"`. Default is `"latent"`.
	latent latent approach;
	uniform jittering uniform approach.
`jitter.uniform.scale`	
	A character string specifying the scale on which to perform the jittering whenever `jitter = "uniform"`. Current options are `"response"` and `"probability"`. Default is `"response"`.

nsim	An integer specifying the number of replicates to use. Default is `1L` meaning one simulation only of residuals.
...	Additional optional arguments.
draw_id	A number refers to the i-th draw of residuals.

Value

A numeric vector of class `c("numeric","resids")` containing the simulated surrogate residuals. Additionally, if nsim > 1, then the result will contain the attributes:

draws A matrix with `nsim` columns, one for each is a replicate of the surrogate residuals. Note, they correspond to the original ordering of the data;

draws_id A matrix with `nsim` columns. Each column contains the observation number each surrogate residuals corresponds to in `draws`. (This is used for plotting purposes.)

A matrix of class `c("matrix","resids")` containing the simulated surrogate residuals used for the partial association analysis in `PAsso`. Additionally, if rep_num > 1 in `PAsso`, then the result will contain the attributes:

draws An array contains all draws of residuals.

Note

Surrogate response values require sampling from a continuous distribution; consequently, the result will be different with every call to `surrogate`. The internal functions used for sampling from truncated distributions are based on modified versions of [rtrunc](#) and [qtrunc](#).

For `"glm"` objects, only the `binomial()` family is supported.

References

Liu, Dungang and Zhang, Heping. Residuals and Diagnostics for Ordinal Regression Models: A Surrogate Approach. *Journal of the American Statistical Association*. http://www.tandfonline.com/doi/abs/10.1080/01621459.2017.1292915?journalCode=uasa20

Li, C., & Shepherd, B. E. (2010). Test of association between two ordinal variables while adjusting for covariates. *Journal of the American Statistical Association*, 105(490), 612-620. https://doi.org/10.1198/jasa.2010.tm09386

Franses, Philip Hans, and Richard Paap. Quantitative models in marketing research. Cambridge University Press, 2001. https://pdfs.semanticscholar.org/dad0/820f287a8cf5a4e8039549e35fc111fd86e5.pdf

Examples

```
# Generate data from a quadratic probit model
set.seed(101)
n <- 2000
x <- runif(n, min = -3, max = 6)
z <- 10 + 3 * x - 1 * x^2 + rnorm(n)
y <- ifelse(z <= 0, yes = 0, no = 1)
```

```
# Scatterplot matrix
pairs(~ x + y + z)

# Misspecified mean structure
fm1 <- glm(y ~ x, family = binomial(link = "probit"))
diagnostic.plot(fm1)

# Correctly specified mean structure
fm2 <- glm(y ~ x + I(x ^ 2), family = binomial(link = "probit"))
diagnostic.plot(fm2)

# Load data
data("ANES2016")
PAsso_1 <- PAsso(responses = c("PreVote.num", "PID"),
                 adjustments = c("income.num", "age", "edu.year"),
                 data = ANES2016)

# Compute residuals
res1 <- residuals(PAsso_1)
```

---

summary   *Summary of partial association analysis*

Description

This function summarizes the partial association analysis by providing partial association matrix, marginal association matrix, and a matrix of the models' coefficients. The partial correlation coefficient matrix displays the partial association between each pair of responses after adjusting the covariates. While the marginal coefficient matrix displays association before the adjustment.

Usage

```
## S3 method for class 'PAsso'
summary(object, digits = max(3L, getOption("digits") - 2L), ...)
```

Arguments

object	A PAsso object to draw the summarized results, which includes partial association matrix and a matrix of the models' coefficients.
digits	A default number to specify decimal digit values.
...	Additional optional arguments.

Value

For a PAsso object, print its partial association matrix, marginal association matrix, and a matrix of the models' coefficients.

Examples

```
# See PAsso for the example.
```

---

surrogate *Surrogate response*

Description

Simulate surrogate response values for cumulative link regression models using the latent method described in Liu and Zhang (2017).

Usage

```
surrogate(
  object,
  method = c("latent", "uniform"),
  jitter.uniform.scale = c("probability", "response"),
  nsim = 1L,
  ...
)
```

Arguments

object	An object of class [clm](), [glm lrm](), [orm](), [polr](), or [vglm]().
method	Character string specifying which method to use to generate the surrogate response values. Current options are "latent" and "uniform". Default is "latent".
jitter.uniform.scale	
	Character string specifying the scale on which to perform the jittering whenever method = "uniform". Current options are "response" and "probability". Default is "response".
nsim	Integer specifying the number of bootstrap replicates to use. Default is 1L meaning no bootstrap samples.
...	Additional optional arguments. (Currently ignored.)

Value

A numeric vector of class c("numeric","surrogate") containing the simulated surrogate response values. Additionally, if nsim > 1, then the result will contain the attributes:

boot_reps A matrix with nsim columns, one for each bootstrap replicate of the surrogate values. Note, these are random and do not correspond to the original ordering of the data;

boot_id A matrix with nsim columns. Each column contains the observation number each surrogate value corresponds to in boot_reps. (This is used for plotting purposes.)

Note

Surrogate response values require sampling from a continuous distribution; consequently, the result will be different with every call to surrogate. The internal functions used for sampling from truncated distributions are based on modified versions of rtrunc and qtrunc.

For "glm" objects, only the binomial() family is supported.

References

Liu, Dungang and Zhang, Heping. Residuals and Diagnostics for Ordinal Regression Models: A Surrogate Approach. *Journal of the American Statistical Association.* http://www.tandfonline.com/doi/abs/10.1080/01621459.2017.1292915?journalCode=uasa20

Nadarajah, Saralees and Kotz, Samuel. R Programs for Truncated Distributions. *Journal of Statistical Software, Code Snippet*, 16(2), 1-8, 2006. https://www.jstatsoft.org/v016/c02.

Examples

```
# Generate data from a quadratic probit model
set.seed(101)
n <- 2000
x <- runif(n, min = -3, max = 6)
z <- 10 + 3*x - 1*x^2 + rnorm(n)
y <- ifelse(z <= 0, yes = 0, no = 1)

# Scatterplot matrix
pairs(~ x + y + z)

# Setup for side-by-side plots
par(mfrow = c(1, 2))

# Misspecified mean structure
fm1 <- glm(y ~ x, family = binomial(link = "probit"))
s1 <- surrogate(fm1)
scatter.smooth(x, s1 - fm1$linear.predictors,
               main = "Misspecified model",
               ylab = "Surrogate residual",
               lpars = list(lwd = 3, col = "red2"))
abline(h = 0, lty = 2, col = "blue2")

# Correctly specified mean structure
fm2 <- glm(y ~ x + I(x ^ 2), family = binomial(link = "probit"))
s2 <- surrogate(fm2)
scatter.smooth(x, s2 - fm2$linear.predictors,
               main = "Correctly specified model",
               ylab = "Surrogate residual",
               lpars = list(lwd = 3, col = "red2"))
abline(h = 0, lty = 2, col = "blue2")


dev.off() # reset to defaults once finish
```

---

test   *Hypothesis testing of the partial association coefficients*

Description

This function use bootstrapping to conduct hypothesis testing for the partial association coefficients. It directly applies onto the "PAsso" class of object generated by "PAsso".

Usage

```
test(object, bootstrap_rep = 300, H0 = 0, parallel = FALSE)
```

Arguments

object	An object of "PAsso" class, which is generated by "PAsso" function.
bootstrap_rep	The number of bootstrap replications. It may be slow.
H0	null hypothesis of partial correlation coefficient.
parallel	logical argument whether conduct parallel for bootstrapping partial association.

Examples

```
# Import ANES2016 data in "PAsso"
data(ANES2016)
# Parial association:
PAsso_2v <- PAsso(responses = c("PreVote.num", "PID"),
                  adjustments = c("income.num", "age", "edu.year"),
                  data = ANES2016)

summary(PAsso_2v, digits=4)

PAsso_2v_test <- test(object = PAsso_2v, bootstrap_rep=20, H0=0, parallel=FALSE)
PAsso_2v_test
```

Index