Package ‘qvcalc’
February 20, 2015

Version 0.8-9
Date 2015-02-03
Title Quasi Variances for Factor Effects in Statistical Models
Author David Firth
Maintainer David Firth <d.firth@warwick.ac.uk>
URL http://warwick.ac.uk/qvcalc
Description Functions to compute quasi variances and associated measures of approximation error
Suggests relimp, MASS
License GPL-2 | GPL-3
NeedsCompilation no
Repository CRAN
Date/Publication 2015-02-03 23:21:51

R topics documented:

indentPrint ................................................................. 1
plot.qv ................................................................. 2
qvcalc ................................................................. 4
worstErrors ............................................................ 6

Index

indentPrint                   Print with Line Indentation

Description

Same as print, but adds a specified amount of white space at the start of each printed line

Usage

indentPrint(object, indent=4, ...)


plot.qv

Description

Provides visualization of estimated contrasts using intervals based on quasi standard errors.

Usage

```r
## S3 method for class 'qv'
plot(x, intervalWidth = 2, ylab = "estimate",
     xlab = x$factorname, ylim = NULL,
     main = "Intervals based on quasi standard errors",
     levelNames = NULL, ...)
```

Arguments

- `x`: an object of class "qv", typically the result of calling `qvcalc`
- `intervalWidth`: the half-width, in quasi standard errors, of the plotted intervals
- `ylab`: as for `plot.default`
- `xlab`: as for `plot.default`
- `ylim`: as for `plot.default`
- `main`: as for `plot.default`
- `levelNames`: labels to be used on the x axis for the levels of the factor whose effect is plotted
- `...`: other arguments understood by `plot`
Details

If levelNames is unspecified, the row names of x$qvframe will be used.

Value

invisible(x)

Author(s)

David Firth, <d.firth@warwick.ac.uk>

References


See Also

qvcalc

Examples

```r
# Overdispersed Poisson loglinear model for ship damage data
# from McCullagh and Nelder (1989), Sec 6.3.2
library(MASS)
data(ships)
ships$year <- as.factor(ships$year)
ships$period <- as.factor(ships$period)
shipmodel <- glm(formula = incidents ~ type + year + period,
                 family = quasipoisson,
                 data = ships, subset = (service > 0), offset = log(service))
shiptype.qvs <- qvcalc(shipmodel, "type")
summary(shiptype.qvs, digits=4)
plot(shiptype.qvs)
```
qvcalc  Quasi Variances for Model Coefficients

Description
Computes a set of quasi variances (and corresponding quasi standard errors) for estimated model coefficients relating to the levels of a categorical (i.e., factor) explanatory variable. For details of the method see Firth (2000), Firth (2003) or Firth and de Menezes (2004). Quasi variances generalize and improve the accuracy of “floating absolute risk” (Easton et al., 1991).

Usage
qvcalc(object, factornname = NULL, coef.indices = NULL, labels = NULL, dispersion = NULL, estimates = NULL, modelcall = NULL)

Arguments
- **object**: A model (of class lm, glm, etc.), or the covariance (sub)matrix for the estimates of interest, or an object of class Btabilities
- **factornname**: Either NULL, or a character vector of length 1
- **coef.indices**: Either NULL, or a numeric vector of length at least 3
- **labels**: An optional vector of row names for the qvframe component of the result (redundant if object is a model)
- **dispersion**: a scalar for the covariance matrix, to cope with overdispersion for example
- **estimates**: an optional vector of estimated coefficients (redundant if object is a model)
- **modelcall**: optional, the call expression for the model of interest (redundant if object is a model)

Details
If object is a model, then at least one of factornname or coef.indices must be non-NULL. The value of coef.indices, if non-NULL, determines which rows and columns of the model’s variance-covariance matrix to use. If coef.indices contains a zero, an extra row and column are included at the indicated position, to represent the zero variances and covariances associated with a reference level. If coef.indices is NULL, then factornname should be the name of a factor effect in the model, and is used in order to extract the necessary variance-covariance estimates.

Ordinarily the quasi variances are positive and so their square roots (the quasi standard errors) exist and can be used in plots, etc.

Occasionally one (and only one) of the quasi variances is negative, and so the corresponding quasi standard error does not exist (it appears as NaN). This is fairly rare in applications, and when it occurs it is because the factor of interest is strongly correlated with one or more other predictors in the model. It is not an indication that quasi variances are inaccurate. An example is shown below using data from the car package: the quasi variance approximation is exact (since type has only 3
levels), and there is a negative quasi variance. The quasi variances remain perfectly valid (they can be used to obtain inference on any contrast), but it makes no sense to plot ‘comparison intervals’ in the usual way since one of the quasi standard errors is not a real number.

Value
A list of class qv, with components

covmat the full variance-covariance matrix for the estimated coefficients corresponding to the factor of interest
qvframe a data frame with variables estimate, SE, quasiSE and quasiVar, the last two being a quasi standard error and quasi-variance for each level of the factor of interest
relerrs relative errors for approximating the standard errors of all simple contrasts
factorname the factor name if given
coef.indices the coefficient indices if given
modelcall if object is a model, object$call; otherwise NULL

Author(s)
David Firth, <d.firth@warwick.ac.uk>

References

See Also
worstErrors, plot.qv

Examples

```r
## Overdispersed Poisson loglinear model for ship damage data
## from McCullagh and Nelder (1989), Sec 6.3.2
library(MASS)
data(ships)
ships$year <- as.factor(ships$year)
```
```
ships$period <- as.factor(ships$period)
shipmodel <- glm(formula = incidents ~ type + year + period,
  family = quasipoisson,
  data = ships, subset = (service > 0), offset = log(service))
shiptype.qvs <- qvcalc(shipmodel, "type")
summary(shiptype.qvs, digits = 4)
plot(shiptype.qvs)

## Quasi-variance summary for "ability" estimates in a Bradley-Terry model
## Requires the "BradleyTerry2" package
## Not run:
library(BradleyTerry2)
example(baseball)
baseball.qv <- qvcalc(BTabilities(baseballModel2))
print(baseball.qv)
plot(baseball.qv, levelNames = c("Bal", "Bos", "Cle", "Det", "Mil", "NY", "Tor"))

## End(Not run)

## Example of a negative quasi variance
## Requires the "car" package
## Not run:
library(car)
data(Prestige)
attach(Prestige)
mymodel <- lm(prestige ~ type + education)
library(qvcalc)
type.qvs <- qvcalc(mymodel, "type")
## Warning message:
## In sqrt(qv): NaNs produced
summary(type.qvs)
## Model call: lm(formula = prestige ~ type + education)
## Factor name: type
##       estimate   SE  quasiSE  quasiVar
## bc  0.0000000 0.000000 2.874361 8.261952
## prof 6.142444 4.258961 3.142737 9.876793
## wc -5.458495 2.690667 NaN 1.022262
## Worst relative errors in SEs of simple contrasts (%): 0 0
## Worst relative errors over *all* contrasts (%): 0 0
plot(type.qvs)
## Error in plot.qv(type.qvs): No comparison intervals available,
## since one of the quasi variances is negative. See ?qvcalc for more.

## End(Not run)
```
**worstErrors**

**Description**
Computes the worst relative error, among all contrasts, for the standard error as derived from a set of quasi variances. For details of the method see Menezes (1999) or Firth and Menezes (2004).

**Usage**

`worstErrors(qv.object)`

**Arguments**

- `qv.object` An object of class `qv`

**Value**
A numeric vector of length 2, the worst negative relative error and the worst positive relative error.

**Author(s)**
David Firth, <d.firth@warwick.ac.uk>

**References**

**See Also**
`qvcalc`

**Examples**

```r
## Overdispersed Poisson loglinear model for ship damage data
## from McCullagh and Nelder (1989), Sec 6.3.2
library(MASS)
data(ships)
ships$year <- as.factor(ships$year)
ships$period <- as.factor(ships$period)
shipmodel <- glm(formula = incidents ~ type + year + period,
                 family = quasipoisson,
                 data = ships, subset = (service > 0), offset = log(service))
shiptype.qvs <- qvcalc(shipmodel, "type")
summary(shiptype.qvs, digits = 4)
worstErrors(shiptype.qvs)
```
Index

*Topic IO
  indentPrint, 1
*Topic hplot
  plot.qv, 2
*Topic models
  plot.qv, 2
  qvcalc, 4
  worstErrors, 6
*Topic regression
  qvcalc, 4
  worstErrors, 6

indentPrint, 1
plot.default, 2
plot.qv, 2, 5
print, 7, 2
print.qv(qvcalc), 4
qvcalc, 2, 3, 4, 7
summary.qv(qvcalc), 4
worstErrors, 5, 6