

Package ‘qvcalc’

January 2, 2012

Version 0.8-7

Date 2010-06-02

Title Quasi variances for factor effects in statistical models

Author David Firth

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

URL <http://www.warwick.ac.uk/go/qvcalc>, <http://www.warwick.ac.uk/go/dfirth>

Description Functions to compute quasi variances and associated measures of approximation error

Suggests relimp

License GPL (>= 2)

Repository CRAN

Date/Publication 2010-06-02 13:17:30

R topics documented:

indentPrint	2
plot.qv	2
qvcalc	4
worstErrors	7
Index	9

indentPrint	<i>Print with Line Indentation</i>
-------------	------------------------------------

Description

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

Usage

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

Arguments

object	any printable object
indent	a non-negative integer, the number of spaces to insert
...	other arguments to pass to <code>print</code>

Value

object is returned invisibly

Author(s)

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

Examples

```
indentPrint("this indented by 10 spaces", indent=10)
```

plot.qv	<i>Plot method for objects of class qv</i>
---------	--

Description

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

Usage

```
## S3 method for class 'qv'  
plot(x, intervalWidth = 2, ylab = "estimate", xlab = x$factorname, ylim = NULL, main = "Intervals based o
```

Arguments

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

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

- Easton, D. F, Peto, J. and Babiker, A. G. A. G. (1991) Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Statistics in Medicine* **10**, 1025–1035.
- Firth, D. (2000) Quasi-variances in Xlisp-Stat and on the web. *Journal of Statistical Software* **5.4**, 1–13. At <http://www.jstatsoft.org>
- Firth, D. (2003) Overcoming the reference category problem in the presentation of statistical models. *Sociological Methodology* **33**, 1–18.
- Firth, D. and Mezezes, R. X. de (2004) Quasi-variances. *Biometrika* **91**, 65–80.
- McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models*. London: Chapman and Hall.
- Menezes, R. X. (1999) More useful standard errors for group and factor effects in generalized linear models. *D.Phil. Thesis*, Department of Statistics, University of Oxford.

See Also

[qvcalc](#)

Examples

```
## 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 Menezes (2004). Quasi variances generalize and improve the accuracy of “floating absolute risk” (Easton et al., 1991).

Usage

```
qvcalc(object, factorname=NULL, coef.indices = NULL,
  labels = NULL, dispersion = NULL,
  estimates=NULL, modelcall=NULL)
```

Arguments

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

Details

If `object` is a model, then at least one of `factorname` 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 `factorname` 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

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

Author(s)

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

References

- Easton, D. F, Peto, J. and Babiker, A. G. A. G. (1991) Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Statistics in Medicine* **10**, 1025–1035.
- Firth, D. (2000) Quasi-variances in Xlisp-Stat and on the web. *Journal of Statistical Software* **5.4**, 1–13. At <http://www.jstatsoft.org>
- Firth, D. (2003) Overcoming the reference category problem in the presentation of statistical models. *Sociological Methodology* **33**, 1–18.
- Firth, D. and Mezezes, R. X. de (2004) Quasi-variances. *Biometrika* **91**, 65–80.

McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models*. London: Chapman and Hall.

Menezes, R. X. (1999) More useful standard errors for group and factor effects in generalized linear models. *D.Phil. Thesis*, Department of Statistics, University of Oxford.

See Also

[worstErrors](#), [plot.qv](#)

Examples

```
## 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)
plot(baseball.qv <- qvcalc(BTabilities(baseballModel2)))

## 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:
## NaNs produced in: sqrt(qv)
summary(type.qvs)
## Model call: lm(formula = prestige ~ type + education)
## Factor name: type
##      estimate      SE quasiSE quasiVar
## bc    0.000000 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

Accuracy of a Quasi-variance Approximation

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

Firth, D. and Mezezes, R. X. de (2004) Quasi-variances. *Biometrika* **91**, 69–80.

McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models*. London: Chapman and Hall.

Menezes, R. X. (1999) More useful standard errors for group and factor effects in generalized linear models. *D.Phil. Thesis*, Department of Statistics, University of Oxford.

See Also

[qvcalc](#)

Examples

```
## 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, 2

*Topic **hplot**

plot.qv, 2

*Topic **models**

plot.qv, 2

qvcalc, 4

worstErrors, 7

*Topic **regression**

qvcalc, 4

worstErrors, 7

indentPrint, 2

plot.default, 3

plot.qv, 2, 6

print, 2

print.qv (qvcalc), 4

qvcalc, 3, 4, 7

summary.qv (qvcalc), 4

worstErrors, 6, 7