

Package ‘glmperm’

February 14, 2012

Type Package

Title Inference in Generalized Linear Models

Version 1.0-3

Date 2010-08-16

Author Wiebke Werft and Douglas M. Potter

Maintainer Wiebke Werft <w.werft@dkfz.de>

Description A permutation test is used for inference in generalized linear models (glm). The procedure is useful when parameter estimates in ordinary glm fail to converge or are unreliable due to small sample size.

License GPL-3

Suggests coin

Repository CRAN

Date/Publication 2010-08-16 16:12:13

R topics documented:

glmperm-package	2
glm.perm	3
prr.test	4
summary.prr.test	7

Index	8
--------------	----------

Description

A permutation Likelihood Ratio test is used for inference in generalized linear models. The procedure is useful when parameter estimates fail to converge or are unreliable due to small sample size.

Details

Package: glmperm
Type: Package
Version: 1.0-2
Date: 2009-31-03
License: GPL-3

The function `pr.test` requires that the user identifies one independent variable for which a test of the null hypothesis (that the regression coefficient is zero) is to be performed. This variable of interest is supplied as a character string `'var'`; the full model is supplied as a formula.

Author(s)

Douglas M. Potter (original version for logistic regression models), Wiebke Werft (extension for generalized linear models)

Maintainer: Wiebke Werft <w.werft@dkfz.de>

References

Potter D.M. (2005) *A permutation test for inference in logistic regression with small- and moderate-sized datasets*. *Statistics in Medicine*, 24:693-708.

Examples

```
### Example 1 ###
nobs<-40
x1<-rnorm(nobs)
x2<-rnorm(nobs)
xx<-cbind(x1,x2)
x0<-rnorm(nobs)+x1+x2
y<-x0+x1+x2+2*rnorm(nobs)
y<-ifelse(y>0,1,0)
LRtest1 <- pr.test(y~x0+x1+x2,var="x0", family=binomial())
summary(LRtest1)

### Example 2 ###
```

```
utils::data(anorexia, package="MASS")
LRtest <- prr.test(Postwt ~ Prewt + Treat, var="Prewt", data = anorexia, family = gaussian())
summary(LRtest)
```

glm.perm *Deviance and Dispersion Factors for a Generalized Linear Model with Permuted Variables*

Description

glm.perm permutes the residuals from a linear regression of the independent variable of interest on the remaining independent variables, and then computes the deviance and dispersion factor from a generalized linear model of these residuals and the other independent variables on the dependent variable. (This function is intended to be used only by 'prr.test'.)

Usage

```
glm.perm(y, x, Family)
```

Arguments

y	the dependent response variable; a vector.
x	a matrix of observations; the rows correspond to the sample size; the columns correspond to the covariables. x must include a column of 1s for the intercept and a column called "resid" of the residuals from a linear regression of the variable of interest 'var' on the remaining independent variables. All other columns are identical to the remaining independent variables without 'var'. This matrix is automatically generated by 'prr.test'.
Family	Usually, this is the glm family used in 'prr.test' and is transferred to 'glm.perm' automatically.

Value

'glm.perm' returns a matrix. The first column contains the deviance and the second column contains the dispersion factor as described above. The number of rows is the number of Monte Carlo replicates 'nrep' used in 'prr.test'.

Author(s)

Douglas M. Potter (original version for logistic regression models), Wiebke Werft (extension for generalized linear models)

References

Potter D.M. (2005) *A permutation test for inference in logistic regression with small- and moderate-sized datasets*. *Statistics in Medicine*, 24:693-708.

Examples

```

nobs<-40
x1<-rnorm(nobs)
x2<-rnorm(nobs)
xx<-cbind(x1,x2)
xint<-rep(1,nobs)

x0<-rnorm(nobs)+x1+x2
y<-x0+x1+x2+2*rnorm(nobs)
y<-ifelse(y>0,1,0)
resid<-lm(x0~x1+x2)$residuals
x <- cbind(xint,xx,resid)
glm.perm(y,x, Family=binomial())
### with Family=binomial() the dispersion factor is always 1
glm.perm(y,x, Family=quasibinomial())
### with Family=quasibinomial() the dispersion factor is estimated from the data

### example for Family=poisson()
utils::data(treepipit, package="coin")
y <- treepipit$counts
nobs <- length(y)
xint<-rep(1,nobs)
resid <- lm(cbpiles~coverstorey ,data=treepipit)$residuals
x <- cbind(xint, treepipit$coverstorey, resid)
glm.perm(y,x, Family=poisson())
### with Family=poisson() the dispersion factor is always 1
glm.perm(y,x, Family=quasipoisson())
### with Family=quasipoisson() the dispersion factor is estimated from the data

```

pr.test

*Inference in Generalized Linear Models***Description**

A permutation of regressor residuals (pr) test which is used for inference in Generalized Linear Models. The null hypothesis that the regression coefficient of the covariate of interest is zero is tested. The procedure is useful when parameter estimates fail to converge or are unreliable due to small sample size.

Usage

```
pr.test(formula, var, family = gaussian, data, nrep = 1000, seed = 12345, weights,
subset, na.action, start = NULL, etastart, mustart, offset, control = glm.control(...),
model = TRUE, method = "glm.fit", x = FALSE, y = TRUE, contrasts = NULL, ...)
```

Arguments

formula a symbolic description of the model to be fitted.

var	the independent variable about which inference is to be made; a character string identical to the one in the formula
family	a description of the error distribution and link function to be used in the model. This can be a character string naming a family function, a family function or the result of a call to a family function. (See family for details of family functions.)
data	an optional data frame, list or environment containing the variables in the model.
nrep	number of Monte Carlo replicates to be used for the permutation test
seed	used seed. Default is seed=12345.
weights	see glm
subset	see glm
na.action	see glm
start	see glm
etastart	see glm
mustart	see glm
offset	see glm
control	see glm
model	see glm
method	see glm
x	see glm
y	see glm
contrasts	see glm
...	see glm

Details

The prr test replaces the independent variable of interest with its residuals from a linear regression on the other independent variables. The test statistic for the permutation test is the p-value based on a likelihood ratio test for a generalized linear model. Thus, the permutation p-value is the fraction of the permutations that have a likelihood-based p-value less than or equal to that for the unpermuted data. Because the p-values for permutations will be a discrete set and round-off errors may occur, prr.test also investigates the sensitivity of the permutation p-value to small variations in the unpermuted likelihood-based p-value.

Missing values are allowed, and observations with them are eliminated.

prr.test calls glm.perm.

Value

fit1	glm.fit of the full model.
fit2	glm.fit of the model without variable 'var'.
call	the original function call.
formula	the full model formula.

seed	used in set.seed(seed).
fit1deviance	the deviance of the full model given in fit1. This value is used for calculating 'LRstat'.
fit2deviance	the deviance of the model without the variable 'var' given in 'fit2'. This value is used for calculating 'LRstat'.
Dispersion	dispersion factor of the full model. For family binomial and poisson this is set to be 1. This value is used for calculating 'LRstat'. The estimated dispersion factor 'estimated.Dispersion' may differ serverely. Then the use of family quasibinomial and quasipoisson might be more appropriate.
estimated.Dispersion	the estimated dispersion factor based on the data.
LRstat	observed Likelihood Ratio test statistics based on the original data (LRstat=abs(fit1deviance-fit2deviance)/Dispersion).
p.value.obs	observed p-value of the Likelihood Ratio test statistics 'LRstat' calculated via a Chi-square distribution with one degree of freedom.
p.value.perm	permutation p-value, fraction of permutation based p-values less than or equal to the observed p-value.
p.value.perm.se	estimated standard errors associated with permutation p-values
nobs	number of observations used
var	the independent variable which is tested for association with the response variable.

Author(s)

Douglas M. Potter (original version for logistic regression models), Wiebke Werft (extension for generalized linear models)

References

Potter D.M. (2005) *A permutation test for inference in logistic regression with small- and moderate-sized datasets*. *Statistics in Medicine*, 24:693-708.

Examples

```
##### EXAMPLE 1 #####
## 40 observations, 3 independent variables, family=binomial()
nobs<-40
x1<-rnorm(nobs)
x2<-rnorm(nobs)
xx<-cbind(x1,x2)
x0<-rnorm(nobs)+x1+x2
y<-x0+x1+x2+2*rnorm(nobs)
y<-ifelse(y>0,1,0)
LRtest1 <- prr.test(y~x0+x1+x2,var="x0", family=binomial())
summary(LRtest1)
```

```
##### EXAMPLE 2 #####
## treepipit data from library(coin) for family=poisson()
utils::data(treepipit, package="coin")
LRtest2 <- prr.test(counts ~ cbpiles + coverstorey, var="cbpiles", data=treepipit, family=poisson())
summary(LRtest2)

##### EXAMPLE 3 #####
## the same as example 2 but with family=quasipoisson()
utils::data(treepipit, package="coin")
LRtest3 <- prr.test(counts ~ cbpiles + coverstorey, var="cbpiles", data=treepipit, family=quasipoisson())
summary(LRtest3)
```

summary.prr.test *Summary of the prr.test results.*

Description

summary is a generic function used to produce result summaries of objects returned by prr.test

Usage

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

Arguments

object	a result object of the call prr.test.
digits	how many digits the values of the results should have.
...	... additional arguments affecting the summary produced.

Details

summary.prr.test prints an overview of the main results of the permutation Likelihood Ratio Test of the prr.test.

Author(s)

Wiebke Werft

Examples

```
utils::data(anorexia, package="MASS")
LRtest <- prr.test(Postwt ~ Prewt + Treat, var="Prewt", data = anorexia, family = gaussian())
summary(LRtest)
```

Index

`glm`, 5
`glm.perm`, 3
`glmperm` (`glmperm-package`), 2
`glmperm-package`, 2

`pr.r.test`, 4

`summary.pr.r.test`, 7