Package ‘robmixglm’

April 3, 2020

Type Package
Title Robust Generalized Linear Models (GLM) using Mixtures
Version 1.0-3
Date 2020-04-02
Maintainer Ken Beath <ken.beath@mq.edu.au>
Contact Ken Beath <ken.beath@mq.edu.au>
Author Ken Beath [aut, cre]
Description Robust generalized linear models (GLM) using a mixture method, as described in Beath (2018) <doi:10.1080/02664763.2017.1414164>. This assumes that the data are a mixture of standard observations, being a generalised linear model, and outlier observations from an overdispersed generalized linear model. The overdispersed linear model is obtained by including a normally distributed random effect in the linear predictor of the generalized linear model.
Depends R(>= 3.2.0)
Suggests R.rsp, robustbase, lattice, forward
VignetteBuilder R.rsp
Imports fastGHQuad, stats, bbmle, MASS, VGAM, actuar, Rcpp (>= 0.12.15), methods, boot, numDeriv
LazyLoad yes
LazyData yes
NeedsCompilation yes
License GPL (>= 2)
LinkingTo Rcpp
Repository CRAN
Date/Publication 2020-04-03 05:20:11 UTC

R topics documented:

  robmixglm-package .................................................. 2
  AIC ................................................................. 4
robmixglm-package

Fits random effects meta-analysis models including robust models

Description


The robmixglm function

This is the main function that allows fitting the models. The robmixglm objects may be tested for outliers using outlierTest. The results of test.outliers may also be plotted.

Author(s)

Ken Beath <ken.beath@mq.edu.au>

References


Examples

# animal brain vs body weight
library(MASS)
data(Animals)
Animals$logbrain <- log(Animals$brain)
Animals$logbody <- log(Animals$body)
lm1 <- lm(logbrain~logbody,data=Animals)
lm2 <- robmixglm(logbrain~logbody,data=Animals)
plot(Animals$logbody,Animals$logbrain)
abline(lm1,col="red")
abline(lm2,col="green")
plot(outlierProbs(lm2))
outlierTest(lm2,showProgress=FALSE)

# Forbes data on relationship between atmospheric pressure and boiling point of water
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(100*log10(pres)~bp,data=MASS::forbes)
summary(forbes.robustmix)
plot(outlierProbs(forbes.robustmix))
outlierTest(forbes.robustmix,showProgress=FALSE)

# Hawkins' data
library(forward)
data(hawkins)
hawkins.robustmix <- robmixglm(y~x1+x2+x3+x4+x5+x6+x7+x8,
data=hawkins)
summary(hawkins.robustmix)
plot(outlierProbs(hawkins.robustmix))
outlierTest(hawkins.robustmix,showProgress=FALSE)

# diabetes
data(diabdata)
diabdata.robustmix <- robmixglm(glyhb~age+gender+bmi+waisthip+frame,
data=diabdata)
summary(diabdata.robustmix)
# this will take about 5-10 minutes
diabdata.step <- step(diabdata.robustmix,glyhb~age+gender+bmi+waisthip+frame)
summary(diabdata.step)
plot(outlierProbs(diabdata.step))
outlierTest(diabdata.step,showProgress=FALSE)

# carrot damage
library(robustbase)
data(carrots)
carrots.robustmix <- robmixglm(cbind(success,total-success)~logdose+factor(block),
family="binomial",data=carrots)
summary(carrots.robustmix)
plot(outlierProbs(carrots.robustmix))
outlierTest(carrots.robustmix,showProgress=FALSE)

# train derailment
library(forward)
data(derailme)
derailme$cYear <- derailme$Year-mean(derailme$Year)
derailme$TrainKm100 <- derailme$TrainKm*100.0
derailme.robustmix <- robmixglm(y~cYear+factor(Type),offset=log(TrainKm100),family="truncpoisson",
quadpoints=51, data=derailme)
summary(derailme.robustmix)
plot(outlierProbs(derailme.robustmix))
outlierTest(derailme.robustmix,showProgress=FALSE)
# hospital costs
hospcosts.robustmix <- robmixglm(costs~adm+age+dest+ins+loglos+sex,family="gamma",data=hospcosts)
summary(hospcosts.robustmix)
plot(outlierProbs(hospcosts.robustmix))
outlierTest(hospcosts.robustmix,showProgress=FALSE)

---

### AIC

#### AIC for `robmixglm` object

**Description**

Returns AIC for a `robmixglm` object.

**Usage**

```r
## S3 method for class 'robmixglm'
AIC(object, ..., k = 2)
```

**Arguments**

- `object` : `robmixglm` object
- `...` : additional argument; currently none is used.
- `k` : penalty per parameter

**Value**

AIC

**Author(s)**

Ken Beath

**Examples**

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
AIC(forbes.robustmix)
```
**BIC**

*Description*

Returns BIC for a robmixglm object.

**Usage**

```r
## S3 method for class 'robmixglm'
BIC(object, ...)
```

**Arguments**

- `object`: robmixglm object
- `...`: additional argument; currently none is used.

**Value**

BIC

**Author(s)**

Ken Beath

**Examples**

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
BIC(forbes.robustmix)
```

---

**coef**

*Coefficients for a robmixglm object*

**Description**

Returns coefficients for a robmixglm object. Only the coefficients for the linear part of the model are returned. Additional coefficients may be obtained using summary().

**Usage**

```r
## S3 method for class 'robmixglm'
coef(object, ...)
```
Arguments

- object: robmixglm object
- ...: additional argument; currently none is used.

Value

- coef

Author(s)

Ken Beath

Examples

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
coef(forbes.robustmix)
```

---

**Diabetes data**

Data from Heritier et al (2009), originally from Harrell (2001,p379). This data was from a study of the prevalence of cardiovascular risk factors such as obesity and diabetes for African Americans. (Willems et al, 19997) Data was available for 403 subjects screened for diabetes, reduced to 372 after removal of cases with missing data.

Usage

- diabdata

Format

A data frame with 372 observations on the following 8 variables.

- glyhb: Glycosated haemoglobin (values above 7.0 are usually taken as a positive diagnosis of diabetes)
- age: age in years
- gender: male or female
- bmi: body mass index in kg/m^2
- waisthip: ratio of waist to hip measurement
- frame: body frame, small, medium or large
- stab.glu: glucose
- location: location, Buckingham or Louisa
extractAIC

Source
Heritier et al (2009)

References

Examples

diabdata.robustmix <- robmixglm(glyhb~age+gender+bmi+waishtip+frame+location, data=diabdata)
summary(diabdata.robustmix)
diabdata.step <- step(diabdata.robustmix,glyhb~age+gender+bmi+waishtip+frame+location) summary(diabdata.step)

extractAIC  

**Extract AIC from a Fitted Model**

Description
Computes the (generalized) AIC for a fitted `robmixglm` model. Used in `step`, otherwise use `AIC`.

Usage
```
## S3 method for class 'robmixglm'
extractAIC(fit, scale, k = 2, ...)
```

Arguments
- `fit`  
fitted `robmixglm` model.
- `scale`  
ignored.
- `k`  
numeric specifying the `weight` of the *equivalent degrees of freedom (≡ edf)* part in the AIC formula.
- `...`  
further arguments (currently unused).

Author(s)
Ken Beath
fitted.robmixglm

See Also

extractAIC, step

Examples

library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=MASS::forbes)
extractAIC(forbes.robustmix)

fitted.robmixglm  Fitted values.

Description

Calculates the fitted values.

Usage

## S3 method for class 'robmixglm'
fitted(object, ...)

Arguments

object  A robmixglm object with a mixture (robust) random effects distribution.
...
Other parameters. (not used)

Value

A vector of the fitted values.

Author(s)

Ken Beath <ken.beath@mq.edu.au>

Examples

library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
BIC(forbes.robustmix)
plot(fitted(forbes.robustmix), residuals(forbes.robustmix))
### hospcosts

**Description**

Data for the analysis in Beath (2018), previously analysed in Marazzi and Yohai (2004), Cantoni and Ronchetti (2006) and Heritier et al (2009). The data is for 100 patients hospitalised at the Centre Hospitalier Universitaire Vaudois in Lausanne, Switzerland for “medical back problems” (APDRG 243).

**Usage**

hospcosts

**Format**

A data frame with 100 observations on the following 9 variables.

- **id**: patient id
- **costs**: cost of stay in Swiss francs
- **los**: length of stay in days
- **adm**: admission type, 0=planned, 1=emergency
- **ins**: insurance type, 0=regular, 1=private
- **age**: age in years
- **sex**: sex, 0=female, 1=male
- **dest**: discharge destination, 0=another health institution, 1=home
- **loglos**: log of length of stay

**Source**

Heritier et al (2009)

**References**


**Examples**

```r
hospcosts.robustmix <- robmixglm(costs~adm+age+dest+ins+loglos+sex,family="gamma",data=hospcosts)
summary(hospcosts.robustmix)
```
### logLik

**log Likelihood for robmixglm object**

**Description**

Returns log Likelihood for a robmixglm object.

**Usage**

```r
## S3 method for class 'robmixglm'
logLik(object, ...)
```

**Arguments**

- `object` robmixglm object
- `...` additional argument; currently none is used.

**Value**

The loglikelihood.

**Author(s)**

Ken Beath

**Examples**

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
logLik(forbes.robustmix)
```

### outlierProbs

**Calculate outlier probabilities for each observation.**

**Description**

For the normal mixture random effect calculates the probability that each observation is an outlier based on the posterior probability of it being an outlier.

**Usage**

```r
outlierProbs(object)
```

**Arguments**

- `object` A metaplus object with a mixture (robust) random effects distribution.
outlierTest

Details

The outlier probabilities are obtained as the posterior probabilities of each observation being an outlier based on the fitted mixture model.

Value

outlier.prob Posterior probability that each observation is an outlier

Author(s)

Ken Beath <ken.beath@mq.edu.au>

Examples

library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
outlierProbs(forbes.robustmix)

outlierTest Test for the presence of outliers.

Description

Uses the parametric bootstrap to test for the presence of outliers.

Usage

outlierTest(object, R=999, showProgress = TRUE)

Arguments

object A robmixglm object with a mixture (robust) random effects distribution.
R number of bootstrap replications
showProgress Show progress bar. Should be set to FALSE if output is sent to a file.

Details

Performs a parametric bootstrap to compare models with and without outliers.

Value

An outlierTest object.

Author(s)

Ken Beath <ken.beath@mq.edu.au>
Examples

```r
hospcosts.robustmix <- robmixglm(costs~adm+age+dest+ins+loglos+sex,family="gamma",data=hospcosts)
summary(hospcosts.robustmix)
summary(outlierTest(hospcosts.robustmix))
```

---

**plot.outlierProbs**

*Plot outlier probabilities.*

**Description**

Plots the outlier probability for each observation, from an outlierProbs object.

**Usage**

```r
## S3 method for class 'outlierProbs'
plot(x, ...)
```

**Arguments**

- `x` : outlierProbs object to be plotted
- `...` : additional parameters to plot

**Value**

Plot

**Author(s)**

Ken Beath <ken.beath@mq.edu.au>

**Examples**

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
plot(outlierProbs(forbes.robustmix))
```
**print.outlierTest**

*Print an outlierTest object*

**Description**

Print an outlierTest object.

**Usage**

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

**Arguments**

- `x` : outlierTest object
- `...` : further arguments (not currently used)

**Author(s)**

Ken Beath

**Examples**

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres, data=forbes)
summary(forbes.robustmix)
print(outlierTest(forbes.robustmix))
```

**residuals.robmixglm**

*Extract Model Residuals*

**Description**

Extracts model residuals from objects returned by modeling functions.

**Usage**

```r
## S3 method for class 'robmixglm'
residuals(object, type = c("deviance", "pearson"), ...)
```
Arguments

- **object**: an object for which the extraction of model residuals is meaningful.
- **type**: Type of residual where valid types are deviance and pearson.
- **...**: other arguments.

Value

Residuals extracted from the object `object`.

Examples

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
BIC(forbes.robustmix)
plot(fitted(forbes.robustmix),residuals(forbes.robustmix))
```

---

**robmixglm**

*Fits a Robust Generalized Linear Model and Variants*

Description

Fits robust generalized linear models and variants described in Beath (2018).

Usage

```r
robmixglm(formula,family=c("gaussian","binomial","poisson","gamma","truncpoisson"),data,offset=NULL,quadpoints=21,notrials=20,EMTol=1.0e-4,verbose=FALSE)
```

Arguments

- **formula**: Formula with columns containing outcomes on the left and covariates on the right.
- **family**: Distribution of response
- **data**: Data frame from which variables are obtained
- **offset**: Offset to be incorporated in the linear predictor.
- **quadpoints**: Number of quadrature points used in the Gauss-Hermite integration.
- **notrials**: Number of random starting values to be used for EM
- **EMTol**: Relative change in likelihood for completion of EM algorithm before switching to quasi-Newton
- **verbose**: Print out diagnostic information? This includes the likelihood and parameter estimates for each EM run.
Value

robnxglm object. This contains

call Call to function
family Family of model to be fitted
model List of model components. These are X, the response, Y, the predictors, and offset, the offset
quadpoints Number of quadrature points.
notrials Number of trials
EMtol EM tolerance used
verbose Was verbose output requested?

Author(s)

Ken Beath

References


Examples

```r
library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(100*log10(pres)~bp,data=forbes)
```

```
summary.robmixglm
summary.robmixglm

Description

Returns summary for a robmixglm object.

Usage

```r
## S3 method for class 'robmixglm'
summary(object, ...)
```

Arguments

- `object` robnxglm object
- `...` additional argument; currently none is used.

Value

summary
Author(s)

Ken Beath

Examples

library(MASS)
data(forbes)
forbes.robustmix <- robmixglm(bp~pres,data=forbes)
summary(forbes.robustmix)
Index

*Topic **datasets**
  diabdata, 6
  hospcosts, 9
*Topic **methods**
  AIC, 4
  BIC, 5
  coef, 5
  logLik, 10
  print.outlierTest, 13
  summary.robmixglm, 15
*Topic **method**
  fitted.robmixglm, 8
  outlierProbs, 10
  outlierTest, 11
  plot.outlierProbs, 12
*Topic **models**
  extractAIC, 7
  residuals.robmixglm, 13
*Topic **multivariate**
  robmixglm, 14

AIC, 4
BIC, 5
ccoef, 5
diabdata, 6
extractAIC, 7, 8
fitted.robmixglm, 8
hospcosts, 9
logLik, 10
outlierProbs, 10
outlierTest, 11
plot.outlierProbs, 12

print.outlierTest, 13
print.summary.outlierTest
  (print.outlierTest), 13
residuals.robmixglm, 13
robmixglm, 14
robmixglm-package, 2
step, 8
summary.outlierTest
  (print.outlierTest), 13
summary.robmixglm, 15