Package 'glmmrr'

September 24, 2025

Type Package

Title Logistic Mixed-Effects Model with Pseudo-
Observations for Estimating Risk Ratios in Clustered Binary Data Analysis

Version 2.1-1

Date 2025-09-24

Maintainer Hisashi Noma <noma@ism.ac.jp>

Description

Logistic mixed-effects model has been widely used as a multilevel statistical model for analyz-
ing clustered binary outcome data (e.g., longitudinal studies, cluster-randomised tri-
als, and multi-center clinical trials). However, the resultant odds ratio estimator can only be inter-
preted as an approximation of the risk ratio estimator for low-frequency events; it cannot be di-
rectly interpreted as an effect measure. To overcome this issue, the modified Poisson regres-
sion analysis and its extention to GEE methodology has been widely applied in recent clini-
cal and epidemiological studies, but these estimating equation-based semiparametric meth-
ods cannot be straightforwardly extended to mixed-effects mod-
els. Noma and Gosho (2025) <doi:10.1002/sim.70280> proposed a new method to provide con-
sistent risk ratio estimator on a multilevel statistical model framework using logistic mixed-
effects model incorporating pseudo-observations. The advantage of the new method is it is im-
plementable using a standard statistical software of GLMM only through modify-
ing dataset. This package involves computational functions for implementing the risk ratio esti-
mation method through multilevel modelling framework.

Depends R (>= 3.5.0)

Imports stats, glmmML, doSNOW, doParallel

License GPL-3

Encoding UTF-8

LazyData true

Contents

---

glmmrr-package   *The 'glmmrr' package*

Description

Logistic mixed-effects models are widely used as multilevel statistical models for analyzing clustered binary outcome data (e.g., longitudinal studies, cluster-randomized trials, and multi-center clinical trials). However, the resulting odds ratio can only be interpreted as an approximation of the risk ratio for rare events, and cannot be directly used as an effect measure. To address this limitation, modified Poisson regression and its extension to GEE methodology have been widely applied in recent clinical and epidemiological research. Nonetheless, these estimating-equation–based semi-parametric methods cannot be readily extended to mixed-effects models. Noma and Gosho (2025) proposed a new approach that yields consistent risk ratio estimators within a multilevel framework by incorporating pseudo-observations into logistic mixed-effects models. A key advantage of this method is that it can be implemented with standard GLMM software simply by modifying the dataset. This package provides computational tools to perform risk ratio estimation in a multilevel modeling framework using this novel approach.

References

Noma, H. and Gosho, M. (2025). Logistic mixed-effects model analysis with pseudo-observations for estimating risk ratios in clustered binary data analysis. *Statistics in Medicine* **44**: e70280. doi:10.1002/sim.70280

---

adpdt   *Creating a dataset for fitting the logistic mixed-effects model for estimating risk ratios via adding pseudo-observations*

Description

Creates a dataset for fitting logistic mixed-effects models to estimate risk ratios by adding pseudo-observations to the original dataset object.

Usage

```
adpdt(y, data=NULL)
```

Arguments

y   The outcome variable (should be coded as 0, 1 contained in the dataset objecet.

data   The dataset object.

Value

A data frame object added the pseudo-observations for fitting the logistic mixed-effects model for estimating risk ratios.

References

Noma, H. and Gosho, M. (2025). Logistic mixed-effects model analysis with pseudo-observations for estimating risk ratios in clustered binary data analysis. *Statistics in Medicine* **44**: e70280. doi:10.1002/sim.70280

Examples

```
library(glmmML)

data(resp)

resp.t <- adpdt(y, resp)  # Adding pseudo-observations to the original dataset

gmm2 <- glmmML(y ~ treat + baseline + center + sex + age, data=resp.t, family =binomial,
 cluster=id, method="ghq")  # Logistic mixed model with pseudo-observations

scoef(gmm2, eform=TRUE)
# 95%CIs and P-values are incorrect (based on the naive model variances)
```

---

cboot   *Perform the cluser-level bootstrap resampling*

Description

Cluster-level bootstrap resamling is performed from the corresponding dataset.

Usage

```
cboot(cluster, data=NULL)
```

Arguments

cluster   The variable specifying clusters contained in the dataset objecet.

data   The dataset object.

Value

A data frame object obtained by cluster-level resamling. The cluster variable after resamling was named as cluster.b.

Examples

```
data(resp)
cboot(id, resp)
```

---

mch   *A cluster-randomised trial dataset for the maternal and child health*
                             handbook

Description

A cluster-randomised trial dataset with binomial outcome.

- `ID`: ID variable of participants.

- `SOUM`: ID variable of soums (involving 18 soums).

- `x`: Binary variable specifying intervention groups (1=Intervention, 0=Control).

- `mage`: Mother's age.

- `medu`: Mother's education (1=uneducated, 2=elementary, 3=incomplete secondary, 4=complete secondary, 5=incomplete high, 6=high (completed collage or university)).

- `mmarry`: Mother's marital status (1=single, 2=married/cohabitating, 3=separated/divorce, 4=windowed/other).

- `mprig1`: First pregnancy (1=Yes, 2=No).

- `height`: Mother's height.

- `weight`: Mother's weight.

- `time`: Travel time from mother's home to antenatal care clinic.

- `Y`: Outcome variable: Number of antenatal visits.

- `y`: Outcome variable: Whether the number of antenatal visits is >= 6 (0 or 1).

- `ses`: Quintile groups by the social-economic index (= 1, 2, 3, 4, 5).

Usage

```
data(mch)
```

Format

A data frame with 500 participants with 18 soums.

References

Mori, R., Yonemoto, N., Noma, H., et al. (2015). The Maternal and Child Health (MCH) handbook in Mongolia: a cluster-randomized, controlled trial. *PloS One* **10**: e0119772.

---

resp   *A longitudinal clinical trial dataset for a respiratory disease*

Description

A longitudinal clinical trial dataset with binomial outcome.

- `id`: ID variable of participants.
- `center`: Numerical variable specifiying center (involving two centers).
- `treat`: Binary variable specifying treatment groups (1=Active, 0=Placebo).
- `age`: Age.
- `sex`: Gender (1=Male, 0=Female).
- `baseline`: Resporatory status at baseline (0 or 1).
- `visit`: Visit times (=1,2,3,4).
- `y`: Outcome variable for respiratory status (0 or 1).

Usage

```
data(resp)
```

Format

A data frame with 111 patients.

References

Berry, D. A. (1989). *Statistical Methodology in the Pharmaceutical Sciences.* New York: CRC Press.

---

sboot   *Computation of the bootstrap confidence intervals and P-values*

Description

Bootstrap confidence intervals and P-values are calculated by coefficient estimates obtained by bootstrap.

Usage

```
sboot(x, eform=FALSE, conf.level=0.95, digit=3)
```

Arguments

x	An array-type object involving coefficient estimates obtained by bootstrap.
eform	A logical value that specify whether the outcome should be transformed by exponential function (default: `FALSE`)
conf.level	Confidence level for calculating confidence intervals (default: 0.95)
digit	Number of decimal places in the output (default: 3).

Value

Results of bootstrap inferences of the regression coefficients.

- `CL`: Lower limits of confidence intervals.
- `CU`: Upper limits of confidence intervals.
- `P-value`: P-values for the coefficient tests.

Examples

```
R1 <- data.frame(rnorm(100),rnorm(100,sd=3),runif(100))

sboot(R1)
sboot(R1,eform=TRUE)
```

---

scoef   *Computation of the ordinary confidence intervals and P-values for the*
                               logistic mixed-effects model using the model variance estimator

---

Description

Confidence intervals and P-values for logistic mixed-effects models can be obtained using the standard model-based variance estimators. By directly inputting the output objects from glmmML, inference results can be computed efficiently. Wald-type confidence intervals and P-values are calculated based on the asymptotic normal approximation. The resulting coefficients and confidence limits can be transformed to the exponential scale by specifying `eform`.

Usage

```
scoef(x, eform=FALSE, conf.level=0.95, digit=3)
```

Arguments

x	An output object of `glmmML`.
eform	A logical value that specify whether the outcome should be transformed by exponential function (default: `FALSE`)
conf.level	Confidence level for calculating confidence intervals (default: 0.95)
digit	Number of decimal places in the output (default: 3).

Value

Results of inferences of the regression coefficients using the ordinary model variance estimators.

- `coef`: Coefficient estimates; transformed to the exponential scale if `eform=TRUE`.
- `SE`: Standard error estimates for `coef`.
- `CL`: Lower limits of confidence intervals.
- `CU`: Upper limits of confidence intervals.
- `P-value`: P-values for the coefficient tests.

Examples

```
library(glmmML)

data(resp)

gmm1 <- glmmML(y ~ treat + baseline + center + sex + age, data=resp, family =binomial,
cluster=id, method="ghq")  # Logistic mixed model

resp.t <- adpdt(y, resp)  # Adding pseudo-observations to the original dataset

gmm2 <- glmmML(y ~ treat + baseline + center + sex + age, data=resp.t, family =binomial,
cluster=id, method="ghq")  # Logistic mixed model with pseudo-observations

scoef(gmm1, eform=TRUE)
scoef(gmm2, eform=TRUE)
# 95%CIs and P-values are incorrect (based on the naive model variances)
```

Index