Package ‘MBRM’

December 22, 2025

Type Package

Title Mixed Regression Models with Generalized Log-Gamma Random Effects

Version 0.1.1

Author Lizandra C. Fabio [aut],
Vanessa Barros [aut],
Cristian Lobos [aut],
Jalmar M. F. Carrasco [aut, cre]

Maintainer Jalmar M. F. Carrasco `<carrasco.jalmar@ufba.br>`

Description
Multivariate distribution derived from a Bernoulli mixed model under a marginal approach, incorporating a non-normal random intercept whose distribution is assumed to follow a generalized log-gamma (GLG) specification under a particular parameter setting. Estimation is performed by maximizing the log-likelihood using numerical optimization techniques (Lizandra C. Fabio, Vanessa Barros, Cristian Lobos, Jalmar M. F. Carrasco, Marginal multivariate approach: A novel strategy for handling correlated binary outcomes, 2025, under submission).

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

LinkingTo Rcpp

Imports Rcpp, stats, Formula, tibble, dplyr, ggplot2

Depends R (>= 3.5)

NeedsCompilation yes

Repository CRAN

Date/Publication 2025-12-22 19:50:07 UTC

Contents

`Arthritis1` *Arthritis1 Dataset*

Description

This dataset contains binary response data from a longitudinal study on rheumatoid arthritis. The data include repeated measurements on 38 individuals, each with 5 time points.

Usage

`Arthritis1`

Format

A data frame with 190 rows and 6 variables:

Ind Subject identifier (integer from 1 to 38)

y Binary response variable (0 or 1)

Sex Sex indicator (0 = Female, 1 = Male)

Age Age indicator (0 = <=55 years 1 = > 55 uears)

Treatment Treatment group (0 = Placebo, 1 = Auranofin)

Time Time indicator (0 = baseline, 1 = follow-up)

Details

The original data was collected to evaluate the effect of a treatment over time on arthritis severity. Covariates include demographic and clinical variables, and the response is binary (presence/absence of symptoms).

Source

Derived from the original dataset included in:
Arthritis.txt an internal clinical dataset used in Bernoulli-GLG modeling study.

References

Fitzmaurice, G. M. and Lipsitz, S. (1995). A model for binary time series data with serial odds ratio patterns. *Journal of the Royal Statistical Society: Series B*, 44, 51–61.

Examples

```
data(Arthritis1)
head(Arthritis1)

# Fit the Bernoulli-LGG model
fit <- MRMfit(y ~ Sex + Age + Treatment + Time, data = Arthritis1)
summary(fit)
```

`envelope.MRM` *Compute simulation envelopes for MRM model*

Description

This function computes the envelopes simulation of the randomized quantile residuals for objects of class `MRM`.

Usage

```
envelope.MRM(object, R = 100, ...)
```

Arguments

`object`	An object of class `MRM`, typically returned from `MRMfit`.
`R`	Integer. Number of replications to simulate the envelopes (default is 100).
`...`	Additional arguments passed to `MRMfit` when `envelope = TRUE`.

Value

A QQ-plot with envelope.

Examples

```
data(Arthritis1)
fit <- MRMfit(y ~ Sex + Age + Treatment + Time, data = Arthritis1)
envelope.MRM(fit)
```

`MRMfit` *Fit Mixed Regression Model with Log-Gamma Random Effects*

Description

This function fits a mixed regression model for binary outcomes with random effects following a generalized log-gamma distribution. The estimation is performed by maximizing a custom log-likelihood using numerical optimization via `optim`.

Usage

```
MRMfit(formula, data, hessian = TRUE, method = "BFGS", ...)
```

Arguments

`formula`	A symbolic description of the model to be fitted, e.g., `y ~ x1 + x2`. The response variable must be binary (0/1).
`data`	A data frame containing the variables in the model. The data must include an `Ind` column indicating cluster or subject IDs for the random effects.
`hessian`	Logical. Should a numerically differentiated Hessian matrix be returned?
`method`	Optimization method to be used in `optim`. One of `"Nelder-Mead"`, `"BFGS"`, `"CG"`, `"L-BFGS-B"`, `"SANN"`, or `"Brent"`. Default is `"BFGS"`.
`...`	Additional arguments passed to `optim`, such as `control`, `lower`, or `upper` (when supported by the chosen method).

Value

An object of class `"MRM"` containing:

`call`	The matched function call.
`formula`	The model formula.
`coefficients`	Estimated fixed effects coefficients.
`scale`	Estimated scale parameter for the random effects distribution.
`loglik`	Maximized log-likelihood value.
`n`	Number of clusters or subjects.
`m`	Vector with the number of observations per cluster.
`ep`	Estimated standard errors of the parameters.
`iter`	Number of iterations used by the optimizer.
`method`	Optimization method used.
`data`	The original data frame used.

Examples

```
# Simulated data
data1 <- rMRM(n = 50, m = rep(3, 50),
              theta = c(0.8, 1, -1, 1),
              X = cbind(x1 = rnorm(150), x2 = rnorm(150)))

# Fit using BFGS (default)
fit1 <- MRMfit(y ~ x1 + x2, data = data1)
summary(fit1)

# Fit using L-BFGS-B with bounds
fit2 <- MRMfit(y ~ x1 + x2, data = data1,
               method = "L-BFGS-B",
               lower = c(1e-5, rep(-Inf, 3)),
               upper = rep(Inf, 4),
               control = list(factr = 1e7))
summary(fit2)
```

`residuals.MRM` *Compute the randomized quantile residuals for MRM model*

Description

This function computes the randomized quantile residuals for objects of class `MRM`.

Usage

```
## S3 method for class 'MRM'
residuals(object, ...)
```

Arguments

`object`	An object of class `MRM`, typically returned from `MRMfit`.
`...`	Additional arguments passed to `MRMfit`.

Value

A numeric vector of randomized quantile residuals.

Examples

```
data(Arthritis1)
fit <- MRMfit(y ~ Sex + Age + Treatment + Time, data = Arthritis1)
summary(residuals(fit))
```

`rMRM` *Simulate Data from a Mixed Regression Model with GLG Random Effects*

Description

This function simulates clustered binary response data from a mixed regression model with random effects. The model allows for different link functions and random effect distributions.

Usage

```
rMRM(n, m, theta, X)
```

Arguments

`n`	Integer. Number of clusters or subjects.
`m`	Integer vector of length `n`. Each element indicates the number of observations per cluster.
`theta`	Numeric vector. The first element is the scale or dispersion parameter for the random effects, and the remaining values are the fixed effects coefficients, including the intercept.
`X`	A data frame or matrix of covariates with `n` rows (one per cluster). Should not include the intercept.

Value

A `tibble` containing the simulated dataset with the following columns:

Ind Cluster or subject ID (integer from 1 to `n`).

y Binary response variable (0 or 1).

x1, x2, ... Covariates as defined in `X`, repeated according to cluster size.

The output also has an attribute `"proportions"` indicating the proportions of 0's and 1's in `y`.

Examples

```
set.seed(123)
n <- 500
m <- rep(3,n)
theta <- c(0.5,1,-2,1)
set.seed(123)
x1 <- runif(sum(m))
x2 <- rnorm(sum(m))
X <- cbind(x1,x2)
set.seed(456)
data1 <- rMRM(n,m,theta,X)
head(data1)
```

`toenail` *toenail Dataset*

Description

A data frame with 1908 observations on the following 5 variables.

Usage

```
toenail
```

Format

A data frame with 1908 rows and 4 variables:

Ind Subject identifier (integer from 1 to 38)

y Binary response is the severity of infection, 0 (not severe) and 1(severe).

treatment Treatment group (0 = Treatment A, 1 = Treatment B)

month a numeric vector giving the time of the visit (not exactly monthly intervals hence not round numbers)

visit a numeric vector giving the number of the visit

Details

The Toenail data discussed in Molenberghs and Verbeke (2010) and Lesaffre and Spiessens (2001 come from a multicenter study comparing two oral treatments (coded as A and B) for toenail (Dermatophyte Onychomycosis - TDO) infection, involved patients evaluated at seven visits, i.e. on weeks 0, 4, 8, 12, 24, 36 and 48. This study was evaluated on 294 patients comprising 1908 measurements. The binary outcome was infection severity, coded as 0 (no and mild) and 1 (moderate and severe). The patients have not been treated prior to the first visit so this should be regarded as the baseline.

References

Molenberghs G, Verbeke G (2010). Models for Discrete Longitudinal data. Springer, New York.

Lesaffre, E. and Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random-effects model: An example. Journal of the Royal Statistical Society, Series C, 50, 325-335.

Examples

```
data(toenail)
head(toenail)

# Fit the Bernoulli-LGG model
fit <- MRMfit(y ~ treatment + month + treatment:month, data = toenail)
summary(fit)
```

Index