

Package ‘pwlmm’

June 14, 2022

Type Package

Title PWIGLS for Two-Level Multivariate and Multilevel Linear Models

Version 1.1.1

Maintainer Alinne Veiga <alinne.veiga@ibge.gov.br>

Description Estimates two-level multilevel linear model and two-level multivariate linear multi-level model with weights following Probability Weighted Iterative Generalised Least Squares approach. For details see Veiga et al.(2014) <[doi:10.1111/rssc.12020](https://doi.org/10.1111/rssc.12020)>.

License GPL (>= 2)

Depends R (>= 3.5.0)

Encoding UTF-8

LazyData true

LinkingTo Rcpp, RcppEigen

Imports lme4, Matrix, Rcpp, RcppEigen

Suggests knitr, rmarkdown

VignetteBuilder knitr

RoxygenNote 7.1.2

NeedsCompilation yes

Author Alinne Veiga [aut, cre] (<<https://orcid.org/0000-0002-1611-2918>>),
João Messa [aut, trl, com]

Repository CRAN

Date/Publication 2022-06-13 22:00:08 UTC

R topics documented:

datamultv	2
dataw1	3
pwigs2	4
wmlmm	5
Index	7

`datamultv`*datamultv data*

Description

Longitudinal data set with the five occasions for each reference person. Generated from the Continuous PNAD (IBGE/Brazil) from the first quarter of 2018 until the first quarter of 2019 following households included in the sample at the first quarter of 2018.

Usage

```
data(datamultv)
```

Format

A `data.frame` with 1685 observations and the following 13 variables.

`Y` the logarithm of the monthly income in the main job for people aged 14 and over (only for people who received cash, products or goods in the main job) plus 1

`X1` age of the resident in the reference date centered around 40

`X2` indicator variable for male residents

`X3` indicator variable for residents of white color or race

`X4` the logarithm of hours actually worked in the reference week in all jobs for people aged 14 and over plus 1

`X5` years of study (people aged 5 or over) standardized for elementary school 9 YEARS SYSTEM

`PSU` level 2 identifiers, is the group identifier for this data

`STRAT` variable that identifies the strata (not needed in the command functions)

`wave` time-dummies for level 1 units

`idd` level 1 identifiers

`wj` weights corresponding to level 2 units

`w_ij` vector of weights corresponding to level 1 units, conditional to their respective level 2 unit (also longitudinal weights in the multivariate data)

`wi_j` weights corresponding to level 1 and 2 units (not needed in the command functions)

Examples

```
data(datamultv)
summary(datamultv)
```

dataw1	<i>dataw1 data</i>
--------	--------------------

Description

Crosssectional data set. A two-level data containing wave one from datamultv data. Generated from the first quarter of 2018 data of the Continuous PNAD (IBGE/Brazil).

Usage

```
data(dataw1)
```

Format

A data.frame with 337 observations and the following 13 variables.

Y the logarithm of the monthly income in the main job for people aged 14 and over (only for people who received cash, products or goods in the main job) plus 1

X1 age of the resident in the reference date centered around 40

X2 indicator variable for male residents

X3 indicator variable for residents of white color or race

X4 the logarithm of hours actually worked in the reference week in all jobs for people aged 14 and over plus 1

X5 years of study (people aged 5 or over) standardized for elementary school 9 YEARS SYSTEM

PSU level 2 identifiers, is the group identifier for this data

STRAT variable that identifies the strata (not needed in the command functions)

wave time-dummies for level 1 units

idd level 1 identifiers

wj weights corresponding to level 2 units

w_ij vector of weights corresponding to level 1 units, conditional to their respective level 2 unit (also longitudinal weights in the multivariate data)

wi_j weights corresponding to level 1 and 2 units (not needed in the command functions)

Examples

```
data(dataw1)
summary(dataw1)
```

pwigls2

Fit Weighted Linear Multilevel Model

Description

Fit a probability-weighted two-level linear model with unequal selection probabilities at each level, via IGLS algorithm.

Usage

```
pwigls2(formula, data = NULL, wj, wij)
```

Arguments

formula	a two-sided linear formula object describing both the fixed-effects and random-effects part of the model, with the response on the left of a ~ operator and the terms, separated by + operators, on the right. Random-effects terms are distinguished by vertical bars () separating expressions for design matrices from grouping factors.
data	an optional data frame containing the variables in formula. If not found in data, the variables are taken from the environment of formula (if specified as a formula) or from the parent frame (if specified as a character vector).
wj	a vector of sampling weights for level two units. Level two units are selected with inclusion probabilities. Then, sampling weights for the level two units are defined as the inverse of these probabilities.
wij	a vector of sampling weights for level one units. After selecting a level two unit, level one units belonging to them are selected with inclusion probabilities. Then, sampling weights for the level one units are defined as the inverse of these probabilities.

Details

Follows estimation process described in Pfeiffermann et al. (1998). Uses probability-weighted IGLS with scaled weights.

Value

Estimated list of estimators

References

D. Pfeiffermann; C. J. Skinner; D. J. Holmes; H. Goldstein; J. Rasbash, 2008, Weighting for Unequal Selection Probabilities in Multilevel Models *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, Vol. 60, No. 1. (1998), pp. 23-40.

Examples

```
data(dataw1)
pwigls2( Y ~ X1 + X2 + (1 | PSU), data = dataw1, wj, wi_j)
```

wmlmm	<i>Fit Weighted Multivariate Linear Multilevel Model to Longitudinal Data.</i>
-------	--

Description

Fit a two-level probability-weighted multivariate linear model with a linear error covariance matrix structure, via IGLS algorithm.

Usage

```
wmlmm(formula, data = NULL, ID3, ID2, ID1, wj, wij, type, rot = NULL)
```

Arguments

formula	a linear formula object with the response on the left of a ~ operator and the terms, separated by + operators, on the right.
data	an optional data frame containing the variables in formula. If not found in data, the variables are taken from the environment of formula (if specified as a formula) or from the parent frame (if specified as a character vector).
ID3	vector of indexes for level two units
ID2	vector of indexes for level one units.
ID1	vector of successive measurements within the same level one unit, for all units.
wj	a vector of sampling weights for level two units. Level two units are selected with inclusion probabilities. Then, sampling weights for the level two units are defined as the inverse of these probabilities.
wij	a vector of sampling weights for level one units. After selecting a level two unit, level one units belonging to them are selected with inclusion probabilities. Then, sampling weights for the level one units are defined as the inverse of these probabilities.
type	type of structure imposed in the error covariance matrix; "toep" refers to the toeplitz, "uns" refers to the unstructured and "genlin" refers to the general linear.
rot	vector of 0's and 1's related to the measurements in time when "genlin" is passed to the type argument. Use 1 if the data were collected in that specific time unit, and 0 otherwise.

Details

Follows estimation process described in Veiga et al. (2014). Uses probability-weighted IGLS with scaled weights.

Value

Estimated list of estimators

References

A. Veiga, P. W. F. Smith and J. J. Brown (2014), The use of sample weights in multivariate multilevel models with an application to income data collected by using a rotating panel survey *Journal of the Royal Statistical Society. Series C (Applied Statistics)* Vol. 63, No. 1 (JANUARY 2014), pp. 65-84 (20 pages)

Examples

```
data(datamultv)
wmlmm ( Y ~ X1 + X2, data = datamultv, PSU, idd, wave, wj, wi_j, "toep")
```

Index

* **datasets**

datamultv, 2

dataw1, 3

datamultv, 2

dataw1, 3

pwigs2, 4

wmlmm, 5