

Package ‘MCM’

October 12, 2022

Type Package

Title Estimating and Testing Intergenerational Social Mobility Effect

Version 0.1.7

Maintainer Jiahui Xu <jpx5053@psu.edu>

Description Estimate and test inter-generational social mobility effect on an outcome with cross-sectional or longitudinal data.

Imports survey, gee, dplyr, lme4, stringr, parameters

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 7.2.1

NeedsCompilation no

Author Jiahui Xu [aut, cre],
Liyong Luo [aut]

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2022-09-26 15:10:08 UTC

R topics documented:

| | |
|----------------------------|----------|
| mcm | 2 |
| mcm_lmer | 4 |
| sim_datlmer | 6 |
| sim_moderate_het | 7 |
| Index | 8 |

mcm

*Estimate and Test Inter-generational Social Mobility Effect on an Outcome***Description**

This function implements the mobility contrast model designed for estimating and testing inter-generational mobility effect on an outcome.

Usage

```
mcm(
  formula,
  data,
  weights = 1,
  na.action = na.omit,
  origin,
  destination,
  family = gaussian(),
  contrasts = NULL,
  gee = FALSE,
  id = NULL,
  corstr = "exchangeable",
  displayresult = TRUE,
  ...
)
```

Arguments

| | |
|-----------|--|
| formula | an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. A typical model used in studying social mobility takes the form <code>response ~ origin*destination</code> , where <code>response</code> is the numeric response vector and <code>origin</code> (<code>destination</code>) is a vector indicating the origin (<code>destination</code>). The specification of <code>origin*destination</code> indicates the cross of <code>origin</code> and <code>destination</code> , which is the same as <code>origin + destination + origin:destination</code> where <code>origin:destination</code> indicates the interaction of <code>origin</code> and <code>destination</code> . |
| data | an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in <code>data</code> , the variables are taken from <code>environment(formula)</code> , typically the environment from which the function is called. |
| weights | an optional vector of unit-level sampling weights to be used in analysis. Should be <code>NULL</code> or a numeric vector. |
| na.action | a function which indicates what should happen when the data contain NAs. The default is set by the <code>na.action</code> setting in options and is <code>na.fail</code> if that is unset. |

| | |
|---------------|--|
| origin | a character indicating the column name of origin. |
| destination | a character indicating the column name of destination. |
| family | a character string, a function or the result of a call to a family function describing the error distribution and link function to be used in the model. |
| contrasts | an optional list. The default is set as sum-to-zero contrast. |
| gee | logical. Should gee be used in estimating the model? |
| id | a vector which identifies the clusters, which is required while gee is used. The length of id should be the same as the number of observations. Data are assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula. |
| corstr | a character string specifying the correlation structure. The following are permitted: "independence", "fixed", "stat_M_dep", "non_stat_M_dep", "exchangeable", "AR-M" and "unstructured". |
| displayresult | logical. Should model results be displayed after estimation. The default is TRUE. |
| ... | additional arguments to be passed to the function. |

Value

A list containing:

| | |
|--------------------|---|
| model | Fitted generalized models of outcome on predictors. See more on function <code>glm</code> in package <code>stats</code> . |
| origin_main | Estimated main effects of origin. |
| destination_main | Estimated main effects of destination. |
| mobility_estimates | Estimated mobility effects. |
| mobility_se | Standard errors of the estimated mobility effects. |
| mobility_sig | Statistical significance of the the estimated mobility effects. |

Examples

```
library(MCM)
data('sim_moderate_het')
mcm(response ~ origin * destination, data = sim_moderate_het,
      origin = "origin", destination="destination")
```

| | |
|----------|--|
| mcm_lmer | <i>Estimate and Test Inter-generational Mobility Effect with Longitudinal Data</i> |
|----------|--|

Description

This function fits a multilevel mobility contrast model to estimate and test inter-generational mobility effect on an outcome in longitudinal data.

Usage

```
mcm_lmer(
  formula,
  data = NULL,
  REML = TRUE,
  control = lme4::lmerControl(),
  start = NULL,
  verbose = 0L,
  subset,
  weights,
  na.action,
  offset,
  contrasts = NULL,
  devFunOnly = FALSE,
  origin = NULL,
  destination = NULL,
  time = NULL,
  displayresult = TRUE,
  ...
)
```

Arguments

| | |
|---------|---|
| formula | Inherit the function form from lme4 package. It is 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. Two vertical bars () can be used to specify multiple uncorrelated random effects for the same grouping variable. (Because of the way it is implemented, the -syntax works only for design matrices containing numeric (continuous) predictors; to fit models with independent categorical effects, see dummy or the lmer_alt function from the afex package.) A typical model used in studying social mobility with longitudinal data takes the form response ~ origin*destination + id, where response is the numeric response vector and origin (destination) is a vector indicating the origin (destination). The specification of origin*destination indicates the cross of origin and destination, which is the same as origin + |
|---------|---|

| | |
|---------------|---|
| | destination + origin:destination where origin:destination indicates the interaction of origin and destination. id is a identifier for the clusters. |
| data | an optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which the function is called. |
| REML | logical. Should the estimates be chosen to optimize the restricted log-likelihood (REML) criterion (as opposed to the log-likelihood)? |
| control | Inherit from <code>lme4</code> package. It is a list (of correct class, resulting from <code>lmerControl()</code> or <code>glmerControl()</code> respectively) containing control parameters, including the nonlinear optimizer to be used and parameters to be passed through to the nonlinear optimizer, see the <code>lmerControl</code> documentation in <code>lme4</code> package for details. |
| start | Inherit from <code>lme4</code> package. It is a named list of starting values for the parameters in the model. |
| verbose | Inherit from <code>lme4</code> package. It is an integer scalar. If > 0 verbose output is generated during the optimization of the parameter estimates. If > 1 verbose output is generated during the individual penalized iteratively reweighted least squares (PIRLS) steps. |
| subset | optional expression selecting the subset of the rows of data to fit the model. |
| weights | an optional vector of 'prior weights' to be used in the fitting process. Should be <code>NULL</code> or a numeric vector. |
| na.action | a function which indicates what should happen when the data contain NAs. The default is set by the <code>na.action</code> setting in <code>options</code> and is <code>na.fail</code> if that is unset. |
| offset | Inherit from <code>lme4</code> package. This can be used to specify an a priori known component to be included in the linear predictor during fitting. This should be <code>NULL</code> or a numeric vector of length equal to the number of cases. One or more offset terms can be included in the formula instead or as well, and if more than one is specified their sum is used. |
| contrasts | an optional list. The default is set as sum-to-zero contrast. |
| devFunOnly | logical - return only the deviance evaluation function. |
| origin | a character indicating the column name of origin. |
| destination | a character indicating the column name of destination. |
| time | a character indicating the time when individual was observed |
| displayresult | logical. Should model results be displayed after estimation. The default is <code>TRUE</code> . |
| ... | additional arguments to be passed to the function. |

Value

A list containing:

| | |
|-------|---|
| model | Fitted generalized models of outcome on predictors. See more on function <code>glm</code> in package <code>stats</code> . |
|-------|---|

| | |
|--------------|--|
| estimates | Estimated mobility effects. |
| se | Standard errors of the estimated mobility effects. |
| significance | Statistical significance of the the estimated mobility effects. |
| esti_3way | Estimated mobility effects conditional on specific age. |
| se_3way | Standard errors of the estimated mobility effects conditional specific age. |
| sig_3way | Statistical significance of the the estimated mobility effects conditional on age. |

Examples

```
library(MCM)
library(lme4)
data("sim_datlmer")
fit_mcm_lmer <- mcm_lmer(yij ~ origin*destination*age +
  (1|id), data = sim_datlmer,
  origin = "origin",
  destination = "destination",
  time = "age")
```

sim_datlmer

Simulated Data Studying Social Mobility (Longitudinal)

Description

This is a simulated data used to study social mobility under longitudinal setting.

Usage

```
data("sim_datlmer")
```

Format

A data frame with 50000 observations on the following 14 variables.

id an ordered factor with levels
 obs a numeric vector
 eij a numeric vector
 origin a factor with levels 1 2 3
 destination a factor with levels 1 2 3
 origin_1 a numeric vector
 origin_2 a numeric vector
 origin_3 a numeric vector
 destination_1 a numeric vector

destination_2 a numeric vector
destination_3 a numeric vector
yij a numeric vector
age a numeric vector
dir a factor with levels 0 1 2

Examples

```
data(sim_datlmer)  
## maybe str(sim_datlmer) ; plot(sim_datlmer) ...
```

| | |
|------------------|--|
| sim_moderate_het | <i>Simulated Data Studying Social Mobility (Cross-Sectional)</i> |
|------------------|--|

Description

This is a simulated data used to study social mobility. In this dataset, it is assumed that there exists a moderate social mobility.

Usage

```
data("sim_moderate_het")
```

Format

A data frame with 30,000 observations on the following 6 variables.

response a numeric vector indicating the outcome variable
origin a numeric vector indicating parents' socioeconomic status
destination a numeric vector indicating child' socioeconomic status
mobility a numeric vector indicating if child's socioeconomic status is different from that of parents'.
upmob a numeric vector indicating child' socioeconomic status is higher than that of parents'.
downmob a numeric vector indicating child' socioeconomic status is lower than that of parents'.

Examples

```
data(sim_moderate_het)  
head(sim_moderate_het)
```

Index

* datasets

sim_datlmer, 6

sim_moderate_het, 7

mcm, 2

mcm_lmer, 4

sim_datlmer, 6

sim_moderate_het, 7