

Package ‘MCM’

September 2, 2021

Type Package

Title Estimating and Testing Intergenerational Social Mobility Effect

Version 0.1.3

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 gee, dplyr, lme4, stringr, broomExtra

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

NeedsCompilation no

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

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2021-09-02 04:30:06 UTC

R topics documented:

mcm	2
mcm_lmer	3
sim_datlmer	6
sim_moderate_het	7
Index	8

mcm

Estimate and Test Intergenerational Social Mobility using the Mobility Contrast Model 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,
  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 fit. A typical model used in studying intergenerational 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 'prior weights' to be used in the model 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 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 the Generalized Estimating Equation (GEE) be used?
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, indicating whether the generalized linear model results based on which the MCM estimates are generated should be displayed. The default is TRUE.
...	additional arguments to be passed to the function.

Value

A list containing:

model	Fitted generalized linear model of outcome on predictors. See more on function <code>glm</code> in package <code>stats</code> .
estimates	Mobility effects estimates.
se	Standard errors of the mobility effects estimates.
significance	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

*Estimate and Test Intergenerational Social Mobility Effect using the
Mobility Contrast Model for Longitudinal Data*

Description

This function fits a multilevel mobility contrast model to estimate and test intergenerational mobility effects using the mobility contrast model for 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 <code>lme4</code> 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 <code>~</code> operator and the terms, separated by <code>+</code> operators, on the right. Random-effects terms are distinguished by vertical bars (<code> </code>) separating expressions for design matrices from grouping factors. Two vertical bars (<code> </code>) can be used to specify multiple uncorrelated random effects for the same grouping variable. (Because of the way it is implemented, the <code> </code> -syntax works only for design matrices containing numeric (continuous) predictors; to fit models with independent categorical effects, see <code>dummy</code> or the <code>lmer_alt</code> function from the <code>afex</code> package.) A typical model used in studying social mobility with longitudinal data takes the form <code>response ~ origin*destination + id</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> . <code>id</code> 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 <code>data</code> , the variables are taken from <code>environment(formula)</code> , typically the environment from which the function is called.
REML	logical, indicating whether the estimates should be chosen to optimize the restricted log-likelihood (REML) criterial (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 `lmerControl` documentation in `lme4` package for details.

<code>start</code>	Inherit from <code>lme4</code> package. It is a named list of starting values for the parameters in the model.
<code>verbose</code>	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.
<code>subset</code>	optional expression selecting the subset of the rows of data to fit the model.
<code>weights</code>	an optional vector of ‘prior weights’ to be used in the fitting process. Should be <code>NULL</code> or a numeric vector.
<code>na.action</code>	a function indicating actions to be taken for 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.
<code>offset</code>	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. It 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.
<code>contrasts</code>	an optional list. The default is set as sum-to-zero contrast.
<code>devFunOnly</code>	logical; return only the deviance evaluation function.
<code>origin</code>	a character indicating the column name of origin.
<code>destination</code>	a character indicating the column name of destination.
<code>time</code>	a character indicating the time when individual was observed
<code>displayresult</code>	logical, indicating whether the original multilevel model results should be displayed. The default is <code>TRUE</code> .
<code>...</code>	additional arguments to be passed to the function.

Value

A list containing:

<code>model</code>	Fitted multilevel models. See more on function <code>glm</code> in package <code>stats</code> .
<code>estimates</code>	Mobility effects estimates.
<code>se</code>	Standard errors of the mobility effects estimates.
<code>significance</code>	Statistical significance of the the mobility effects estimates.
<code>esti_3way</code>	Mobility effects estimates for a given age.
<code>se_3way</code>	Standard errors of the mobility effects estimates for a given specific age.
<code>sig_3way</code>	Statistical significance of the mobility effects estimates for a given 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

A simulated longitudinal dataset of intergenerational social mobility.

Usage

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

Format

A data frame with 50,000 observations and 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 Cross-Sectional Data of Intergenerational Social Mobility</i>
------------------	--

Description

A simulated cross-sectional dataset of intergenerational social mobility. The mobility effects are heterogeneous (e.g., with different directions and effect sizes) and the mobility scale are moderate (e.g., the ratio of the number of nonmobile and the mobile is set at 3:2).

Usage

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

Format

A data frame with 30,000 observations and 6 variables.

response a numeric vector indicating the outcome variable

origin a numeric vector indicating parents' socioeconomic status group

destination a numeric vector indicating child' socioeconomic status group

mobility a numeric vector indicating if child's socioeconomic status differs from that of their 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, 3

sim_datlmer, 6

sim_moderate_het, 7