Package 'mlmtools'

July 23, 2025

Title Multi-Level Model Assessment Kit

Version 1.0.2

Maintainer Laura Jamison <lj5yn@virginia.edu>

Description

Multilevel models (mixed effects models) are the statistical tool of choice for analyzing multilevel data (Searle et al, 2009). These models account for the correlated nature of observations within higher level units by adding group-level error terms that augment the singular residual error of a standard OLS regression. Multilevel and mixed effects models often require specialized data pre-processing and further post-estimation derivations and graphics to gain insight into model results. The package presented here, 'mlmtools', is a suite of pre- and post-estimation tools for multilevel models in 'R'. Package implements post-estimation tools designed to work with models estimated using 'lme4"s (Bates et al., 2014) lmer() function, which fits linear mixed effects regression models. Searle, S. R., Casella, G., & McCulloch, C. E. (2009, ISBN:978-0470009598). Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014) <doi:10.18637/jss.v067.i01>.

Depends R (>= 3.5.0)

Imports lme4,stats,ggplot2

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.2.1

NeedsCompilation no

Author Laura Jamison [aut, cre],
Jessica Mazen [aut],
Erik Ruzek [aut],
Gus Sjobeck [ctb]

Repository CRAN

Date/Publication 2022-10-26 06:22:36 UTC

Contents

`betweenPlot` *Plots Between Group Associations*

Description

Plots the between-group associations between an outcome and predictor variable.

Usage

```
betweenPlot(
  x,
  y,
  grouping,
  dataset,
  xlab = x,
  ylab = y,
  between_title = "Between-Group Association Plot",
  point_color = "gray40",
  line_color = "black",
  se = FALSE,
  full_range = FALSE,
  lty = 1,
  size = 2
)
```

Arguments

`x`	Predictor variable.
`y`	Outcome variable.
`grouping`	Grouping variable.
`dataset`	A dataset containing the predictor, outcome, and grouping variables.
`xlab`	Character vector specifying the horizontal axis label.
`ylab`	Character vector specifying the vertical axis label.

between_title	Character vector specifying the title for the between group plot.
point_color	Color for points.
line_color	Color for lines.
se	A logical value indicating whether confidence intervals should be displayed.
full_range	A logical value indicating whether the fit line should span the full range of the plot or just the data.
lty	Line type.
size	Width of fit line.

Value

Produces a plot of the between-group associations between an outcome and predictor variable.

References

Chow, S., Gilmore, R. O., Hallquist, M., Ram, N., & Brinberg, M. (2019). Introduction to multilevel model and interactions. GitHub. https://github.com/psu-psychology/r-bootcamp-2019/blob/master/talks/RBootcamp_MLMI

Examples

```
# Read in data
data(instruction)
# Produce between plot
betweenPlot(x = "mathkind", y = "mathgain", grouping = "classid",
dataset = instruction, xlab = "Kindergarten Math Score",
ylab = "Gain in Math Score")
```

---

caterpillarPlot *Caterpillar Plot*

Description

Plots empirical Bayes both point prediction and prediction intervals for each random effect parameter across all groups.

Usage

```
caterpillarPlot(
  model,
  grouping,
  title = print(grouping),
  tall = TRUE,
  grey = FALSE
)
```

Arguments

model	A given lmer model.
grouping	The name of the grouping variable of interest, as a character string.
title	The title of the plot.
tall	Logical argument specifying whether the plot should be plotted vertically or horizontally.
grey	Logical argument specifying whether the intervals should be plotted in color or greyscale.

Value

Produces a caterpillar plot.

References

Rabe-Hesketh S, Skrondal A (2012). Multilevel and Longitudinal Modeling Using Stata, Volumes I and II, Third Edition. 3 edition edition. Stata Press. ISBN 978-1-59718-108-2.

Examples

```
# Read in data
data(instruction)
# Create model
mod <- lme4::lmer(mathgain ~ (1 | classid), data = instruction)
# Produce caterpillar plot
caterpillarPlot(mod, title = "title", grouping = "classid", grey = TRUE)
```

---

center *Centers variables for mixed effects models*

Description

Centers variables using the group-mean (person-mean) centering approach for mixed-effects models, and adds these variables to the data frame.

Usage

```
center(
  dataset,
  x,
  grouping,
  type = "mean",
  standardize = FALSE,
  centerResult = FALSE
)
```

Arguments

dataset	A dataset containing the variables to be centered and the grouping variable
x	The variable or variables to be centered
grouping	The variable or variables that define the grouping structure of the data
type	a function to compute the grouping summary variable
standardize	a logical value indicating whether x should be standardized before the computaion proceeds
centerResult	a logical value indicating whether resulting grouping summary variable values should be centered at 0

Value

Creates two new variables in the data frame - a mean of the desired variable computed for each unique value in the grouping variable and a deviation score for each observation within the grouping variable that is that observation's raw score subtracted from the group mean.

References

Enders, C. & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old problem. Psychological Methods, 12(2), 121–138

Examples

```
data(instruction)
#Center student level socioeconomic status, "ses", around class mean "ses"
### To repress output: use invisible()
center(dataset = instruction, x = "ses", grouping = "classid")
#Center class-level variable teacher's mathematic prepartion,
# mathprep, around school mean "mathprep"
center(dataset = instruction, x = "mathprep", grouping = "schoolid")
```

ICCm *Computes ICC values for mixed-effects models*

Description

Computes ICC values for lme4-fitted mixed-effects models.

Usage

```
ICCm(model, re_type = c("NA"))
```

Arguments

model	A linear mixed-effects model of class lmerMod or lmerModLmerTest
re_type	A value indicating whether a model with two random effects is nested or cross-classified

Value

If re_type is "NA", the proportion of variance at the random effect is computed.

If re_type = "nested", the likeness of y scores in the same level 3 unit (the proportion of variance at Level3_factor), the likeness of y scores in the same level 2 units in the same level 3 unit (proportion of variance at Level3_factor and Level2_factor), and the likeness of level 2 units in the same level 3 unit (proportion of Level2_factor variance at Level3_factor) are computed.

If re_type = "cc", the likeness of y scores in the same C1_factor unit (correlation between outcome values of units in same C1_factor but different C2_factor), the likeness of y scores in the same C2_factor (correlation between outcome values of units in the same C2_factor but different C2_factor), and the likeness of y scores in the same C1_factor and C2_factor combination (correlation between outcome values of units in the same C2_factor and C2_factor) are computed.

References

Snijders, T. A. B. & Bosker, R. J. (2012). Multilevel Analysis (2nd Ed.). Sage Publications Ltd. Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning variation in multilevel models. Understanding statistics: statistical issues in psychology, education, and the social sciences, 1(4), 223-231.

Examples

```
# Gaussian
## Read in data
data(instruction)
## Create model
mod <- lme4::lmer(mathgain ~ (1 | classid), data = instruction)
## Estimate ICC
ICCm(mod)

# Logistic
## Read in data
data(reporting)
## Create model
mod <- lme4::glmer(mention.outliers ~ Basics + (1 | Journal), data = reporting, family = "binomial")
## Estimate ICC
ICCm(mod)
```

---

instruction *Instruction Data*

Description

Data from a study on instructional improvement across 312 classrooms.

Usage

```
data(instruction)
```

Format

A data frame with 1190 observations on the following 8 variables.

`female` Dummy variable for being female

`mathkind` Math achievement score in the spring of kindergarten

`mathgain` Gain in math achievement score from spring of kindergarten to spring of first grade

`ses` Socioeconomic status

`mathprep` First grade teacher's mathematic preparation (based on number of courses taken)

`classid` Classroom identifier

`schoolid` School identifier

`childid` Student identifier

Source

Stata Press

References

Rabe-Hesketh, Sophia, and Brian Everitt. Handbook of statistical analyses using Stata. CRC Press, 2003.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. American educational research journal, 42(2), 371-406.

Examples

```
data(instruction)
```

---

mlm_assumptions *Reports the output of testing all assumptions for a multilevel model*

Description

Reports the results from testing all assumptions of a multilevel model and provides suggestions if an assumption is not passed

Usage

```
mlm_assumptions(model)
```

Arguments

`model`	A linear mixed-effects model of class lmerMod, lmerModLmerTest, or glmerMod of type binomial.

Value

Tests the relevant assumptions of the specified multilevel model.

References

Glaser, R. E. (2006). Levene's Robust Test of Homogeneity of Variances. Encyclopedia of Statistical Sciences. 6.

Examples

```
# Read in data
data(instruction)
# Create model
mod <- lme4::lmer(mathgain ~ mathkind + (1 | classid), data = instruction)
# Evaluate assumptions
mlm_assumptions(mod)
```

`prints` *S3Methods for Printing*

Description

Prints for `mlmtools` objects

Usage

```
## S3 method for class 'center'
print(x, ...)

## S3 method for class 'ICCm'
print(x, ...)

## S3 method for class 'rsqmlm'
print(x, ...)

## S3 method for class 'varCompare'
print(x, ...)
```

Arguments

`x`	Object from `mlmtools` package
`...`	Additional arguments

Value

Prints `mlmtools` object

`reporting` *Reporting Data*

Description

Data from a study on the reporting rates of outliers with data on 2235 experiments.

Usage

```
data(reporting)
```

Format

A data frame with 2235 observations on the following 18 variables.

`Reference.Code` Bibtex reference code for the article

`year` Year of publication

`time.pulled` Year article was pulled for these data

`Type` Type of psychological journal

`Journal` Journal of publication

`authors` Authors

`article` Title of article

`original.sample.size` Original sample size from the article

`mention.outliers` Whether or not the article mentioned outliers

`final.sample.size` Final sample size from the article

`number.outliers` Number of outliers identified by the article

`Basics` Whether or not they ran basic statistics (e.g., descriptive statistics, z scores, t tests, and correlations)

`ANOVA` Whether or not they ran ANOVA

`Regression` Whether or not they ran a regression

`ChiSquare` Whether or not they ran a chi-squared test

`Nonparametric` Whether or not they ran a nonparametric test

`Modeling` Whether or not they used structural equation modeling

`BayesOther` Whether or not they used Bayes or another form of analysis

Source

GitHub

References

Valentine, K. D., Buchanan, E. M., Cunningham, A., Hopke, T., Wikowsky, A., & Wilson, H. (2021). Have psychologists increased reporting of outliers in response to the reproducibility crisis?. Social and Personality Psychology Compass, 15(5), e12591.

Examples

```
data(reporting)
```

`rsqmlm` *Calculates R-squared from lmer models*

Description

Calculates variance explained by lme4-fitted mixed-effects models.

Usage

```
rsqmlm(model, by_cluster = FALSE)
```

Arguments

`model`	A linear mixed-effects model of class lmerMod or lmerModLmerTest
`by_cluster`	Logical, if TRUE returns variance explained at each level

Value

Computes the percent variance explained by the model.

References

Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of the Royal Society Interface, 14(134), 20170213.

Examples

```
# Read in data
data(instruction)
# Center mathkind by classid
center(dataset = instruction, x = "mathkind", grouping = "classid")
# Create model
mod <- lme4::lmer(mathgain ~ classid_mathkind.cmn +
classid_mathkind.devcmn + (1 | classid), data = instruction)
# Calculate r-squared
### To repress output: use invisible()
rsq <- rsqmlm(mod)
rsq
rsq$marginal
rsq$conditional
```

varCompare *Compares variance explained for two mixed effects models*

Description

Compares variance explained by additional fixed effects for two lme4-fitted mixed-effects models.

Usage

```
varCompare(model1, model2)
```

Arguments

model1	A linear mixed-effects model of class lmerMod or lmerModLmerTest
model2	A linear mixed-effects model of class lmerMod or lmerModLmerTest

Details

Specifically, 1-(total variance for less parsimonious model/total variance for more parsimonious model).

Value

Computes the percent increase in variance explained by the less parsimonious (more complicated) model compared to the more parsimonious (less complicated) model.

References

Snijders, T. A. B. & Bosker, R. J. (2012). Multilevel Analysis (2nd Ed.). Sage Publications Ltd.

Examples

```
# Read in data
data(instruction)
# Create null model
mod0 <- lme4::lmer(mathgain ~ (1 | classid), data = instruction)
# Create model of interest
mod1 <- lme4::lmer(mathgain ~ mathkind + (1 | classid), data = instruction)
# Compare variance explained
### To repress output: use invisible()
varCompare(mod0, mod1)
```

withinPlot *Plots Within Group Associations*

Description

Plots the within-group associations between an outcome and predictor variable.

Usage

```
withinPlot(
  x,
  y,
  grouping,
  dataset,
  xlab = x,
  ylab = y,
  within_title = "Within-Group Association Plot",
  point_color = "gray40",
  line_color = "black",
  se = FALSE,
  full_range = FALSE,
  lty = 1,
  size = 2
)
```

Arguments

x	Predictor variable.
y	Outcome variable.
grouping	Grouping variable (individual may be grouping variable).
dataset	A dataset containing the predictor, outcome, and grouping variables.
xlab	Character vector specifying the horizontal axis label.
ylab	Character vector specifying the vertical axis label.
within_title	Character vector specifying the title for the within group plot.
point_color	Color for points.
line_color	Color for lines.
se	A logical value indicating whether confidence intervals should be displayed.
full_range	A logical value indicating whether the fit line should span the full range of the plot or just the data.
lty	Line type.
size	Width of fit line.

Value

Produces a plot of the within-group associations between an outcome and predictor variable.

References

Chow, S., Gilmore, R. O., Hallquist, M., Ram, N., & Brinberg, M. (2019). Introduction to multilevel model and interactions. GitHub. https://github.com/psu-psychology/r-bootcamp-2019/blob/master/talks/RBootcamp_MLMI

Examples

```
# Read in data
data(instruction)
# Create within plot
mathkind_withinPlot <- withinPlot(x = "mathkind", y = "mathgain",
grouping = "classid", dataset = instruction,
xlab = "Kindergarten Math Score", ylab = "Gain in Math Score")
```

Index