Package 'LMMstar'

March 8, 2023

Type Package

Title Repeated Measurement Models for Discrete Times

Version 0.9.0

Date 2023-03-08

Description Companion R package for the course ``Statistical analysis of correlated and repeated measurements for health science researchers"
taught by the section of Biostatistics of the University of Copenhagen.
It implements linear mixed models where the model for the variance-covariance of the residuals is specified via patterns (compound symmetry, toeplitz, unstructured, ...).
Statistical inference for mean, variance, and correlation parameters is performed based on the observed information and a Satterthwaite approximation of the degrees of freedom.
Normalized residuals are provided to assess model misspecification.
Statistical inference can be performed for arbitrary linear or non-linear combination(s) of model coefficients.
Predictions can be computed conditional to covariates only or also to outcome values.

License GPL-3

Encoding UTF-8

URL https://github.com/bozenne/LMMstar

BugReports https://github.com/bozenne/LMMstar/issues

Depends R (>= 3.5.0), nlme

Imports copula, emmeans, ggplot2, grid, lava, Matrix, multcomp, numDeriv, parallel, pbapply, rlang, sandwich

Suggests AICcmodavg, asht, data.table, ggh4x, ggpubr, lattice, mvtnorm, lme4, lmerTest, mice, nlmeU, optimx, psych, Publish, qqtest, R.rsp, reshape2, rmcorr, testthat

VignetteBuilder R.rsp

RoxygenNote 7.2.1

Collate '0-onload.R' 'LMMstar-package.R' 'LMMstar.options.R' 'add.R' 'anova.R' 'autoplot.R' 'backtransform.R' 'baselineAdjustment.R' 'coef.R' 'confint.R' 'constrain.R' 'df.R' 'doc-data.R' 'dummy.coef.R' 'emmeans.R' 'estimate.R' 'findPatterns.R'

'fitted.R' 'formula.R' 'iid.R' 'information.R' 'levels.R'
'lmm.R' 'logLik.R' 'manifest.R' 'mlmm.R' 'model.frame.R'
'model.matrix.R' 'model.tables.R' 'moments.R' 'multcomp.R'
'nobs.R' 'partialCor.R' 'plot.R' 'precompute.R' 'predict.R'
'print.R' 'profile.R' 'proportion.R' 'ranef.R' 'rbind.R'
'remove.R' 'reparametrize.R' 'resample.R' 'residuals.R'
'sampleRem.R' 'scatterplot.R' 'score.R' 'sigma.R'
'structure-calc_Omega.R' 'structure-calc_d2Omega.R'
'structure-calc_dOmega.R' 'structure-initialization.R'
'structure-skeleton.R' 'structure.R' 'summarize.R'
'summarizeNA.R' 'summary.R' 'terms.R' 'unorderedPairs.R'
'utils.R' 'vcov.R' 'weights.R'

NeedsCompilation no

Author Brice Ozenne [aut, cre] (<https://orcid.org/0000-0001-9694-2956>),
Julie Forman [aut] (<https://orcid.org/0000-0001-7368-0869>)

Maintainer Brice Ozenne <brice.mh.ozenne@gmail.com>

Repository CRAN

Date/Publication 2023-03-08 14:30:05 UTC

R topics documented:

`LMMstar-package` *LMMstar package: repeated measurement models for discrete times*

Description

Companion R package for the course "Statistical analysis of correlated and repeated measurements for health science researchers" taught by the section of Biostatistics of the University of Copenhagen. It implements linear mixed models where the model for the variance-covariance of the residuals is specified via patterns (compound symmetry, toeplitz, unstructured, ...). Statistical inference for mean, variance, and correlation parameters is performed based on the observed information and a Satterthwaite approximation of the degrees of freedom. Normalized residuals are provided to assess model misspecification. Statistical inference can be performed for arbitrary linear or non-linear combination(s) of model coefficients. Predictions can be computed conditional to covariates only or also to outcome values.

Notations: the linear mixed model estimated by `lmm` is denoted:

$$\mathbf{Y}_i = \mathbf{X}_i\beta + \boldsymbol{\varepsilon}_i$$

where

- $\mathbf{Y} = (Y_1, \ldots, Y_m)$: vector of outcomes.
- $\mathbf{X} = (X_1, \ldots, X_p)$: design matrix (extractor: `model.matrix.lmm`).
- $\boldsymbol{\varepsilon}$: vector of residuals with 0-mean and variance Ω_i (extractor: `residuals.lmm`).
- β: estimated mean coefficients relative to X (extractor: `coef.lmm`).
- Ω: the modeled variance-covariance of the residuals with diagonal elements σ_j^2 (extractor: `sigma.lmm`).
- i indexes the cluster (level where replicates are assumed independent).
- j indexes the repetitions, e.g. the variance of ε_{ij} is σ_{ij}^2.

Covariance patterns: Ω can be parametrized as:

- `"ID"`: identity (no correlation, constant variance).
- `"IND"`: independent (no correlation, time-specific variance).

- "CS": compound symmetry (constant correlation and variance). Can also be used to specify a nested random effect structure or a block specific correlation and variance.
- "TOEPLITZ": toeplitz (lag-specific correlation, time-specific variance).
- "UN": unstructured (time-specific correlation, time-specific variance).

It possible to stratify each structure with respect to a categorical variable.

Optimizer: the default optimizer, "FS", implements a fisher scoring algorithm descent with backtracking in case of decreasing or undefined log-likelihood. It does not constrain Ω to be positive definite which may cause problem in small sample or complex models. It is possible to use other optimizer: `nlme::gls` for certain covariance patterns or `stats::optim`.

`abetaL` *Data From The Bland Altman Study (Long Format)*

Description

Extract data from a longitudinal case control study including 87 patients newly diagnosed with bipolar disorder and 44 age and sex matched healthy controls. Contains demographic data and lifestyle factors at baseline, as well as measures of psychosocial functioning at baseline and 1 year follow-up. This dataset is in the long format (i.e. one line per measurement).

- `id`: study participant.
- `sex`: male (M) or female (F).
- `age`: age in years.
- `group`: bipolar disorder (BD) or healthy control (HC).
- `episode`: whether the patient experience an affective episode during follow-up.
- `visit`: index of time at which pss, fast, and qol measurements where performed.
- `year`: time at which pss, fast, and qol measurements where performed.
- `pss`: perceived stress score.
- `fast`: functioning assessment short test.
- `qol`: WHO quality of life score.
- `educationyears`: years of education including basic school.
- `alcohol`: daily alcohol consumption.
- `missingreason`: reason of drop out or missed visit.

Usage

```
data(abetaL)
```

References

Pech, Josefine, et al. The impact of a new affective episode on psychosocial functioning, quality of life and perceived stress in newly diagnosed patients with bipolar disorder: A prospective one-year case-control study.Journal of Affective Disorders 277 (2020): 486-494.

abetaW *Data From The abeta Study (Wide Format)*

Description

Extract data from a longitudinal case control study including 87 patients newly diagnosed with bipolar disorder and 44 age and sex matched healthy controls. Contains demographic data and lifestyle factors at baseline, as well as measures of psychosocial functioning at baseline and 1 year follow-up. This dataset is in the wide format (i.e. one line per participant).

- `id`: study participant.
- `sex`: male (M) or female (F).
- `age`: age in years.
- `group`: bipolar disorder (BD) or healthy control (HC).
- `episode`: whether the patient experience an affective episode during follow-up.
- `fast0`,`fast1`: functioning assessment short test at baseline and follow-up.
- `qol0`,`qol1`: WHO quality of life score at baseline and follow-up.
- `pss0`,`pss1`: perceived stress score at baseline and follow-up.
- `educationyears`: years of education including basic school.
- `alcohol`: daily alcohol consumption.
- `missingreason`: reason of drop out or missed visit.

Usage

```
data(abetaW)
```

References

Pech, Josefine, et al. "The impact of a new affective episode on psychosocial functioning, quality of life and perceived stress in newly diagnosed patients with bipolar disorder: A prospective one-year case-control study."Journal of Affective Disorders 277 (2020): 486-494.

add *Add Columns to Output*

Description

Auxiliary function that can be used when specifying the argument `columns` (e.g. calling `confint.lmm`) to add columns.

Usage

```
add(...)
```

Arguments

...	[character vector] name of the columns to be added to the default output.

Value

A character vector

Examples

```
set.seed(10)
dW <- sampleRem(25, n.times = 1, format = "long")
e.lmm <- lmm(Y~X1, data = dW)

confint(e.lmm, columns = add("statistic"))
```

---

anova.lmm *Multivariate Tests For Linear Mixed Model*

Description

Simultaneous tests of linear combinations of the model paramaters using Wald tests or Likelihood Ratio Test (LRT).

Usage

```
## S3 method for class 'lmm'
anova(
  object,
  effects = NULL,
  robust = FALSE,
  rhs = NULL,
  df = !is.null(object$df),
  ci = TRUE,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  ...
)
```

Arguments

object	a lmm object. Only relevant for the anova function.
effects	[character or numeric matrix] Should the Wald test be computed for all variables ("all"), or only variables relative to the mean ("mean" or "fixed"), or only variables relative to the variance structure ("variance"), or only variables relative to the correlation structure ("correlation"). Can also be use to specify linear combinations of coefficients or a contrast matrix, similarly to the linfct argument of the multcomp::glht function.

robust	[logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors.
rhs	[numeric vector] the right hand side of the hypothesis. Only used when the argument effects is a matrix.
df	[logical] Should a F-distribution be used to model the distribution of the Wald statistic. Otherwise a chi-squared distribution is used.
ci	[logical] Should an estimate, standard error, confidence interval, and p-value be output for each hypothesis?
transform.sigma, transform.k, transform.rho, transform.names	are passed to the vcov method. See details section in coef.lmm.
...	Not used. For compatibility with the generic method.

Details

By default adjustment of confidence intervals and p-values for multiple comparisons is based on the distribution of the maximum-statistic. This is refered to as a single-step Dunnett multiple testing procedures in table II of Dmitrienko et al. (2013). It is performed using the multcomp package with the option `test = adjusted("single-step")` with equal degrees of freedom or by simulation using a Student's t copula with unequal degrees of freedom (more in the note of the details section of confint.Wald_lmm).

Value

A data.frame (LRT) or a list of containing the following elements (Wald):

- multivariate: data.frame containing the multivariate Wald test. The column df.num refers to the degrees of freedom for the numerator (i.e. number of hypotheses) wherease the column df.denum refers to the degrees of freedom for the denominator (i.e. Satterthwaite approximation).
- univariate: data.frame containing each univariate Wald test.
- glht: used internally to call functions from the multcomp package.
- object: list containing key information about the linear mixed model.
- vcov: variance-covariance matrix associated to each parameter of interest (i.e. hypothesis).
- iid: matrix containing the influence function relative to each parameter of interest (i.e. hypothesis).
- args: list containing argument values from the function call.

References

Dmitrienko, A. and D'Agostino, R., Sr (2013), Traditional multiplicity adjustment methods in clinical trials. Statist. Med., 32: 5172-5218. https://doi.org/10.1002/sim.5990.

See Also

summary.Wald_lmm or confint.Wald_lmm for a summary of the results.
autoplot.Wald_lmm for a graphical display of the results.
rbind.Wald_lmm for combining result across models and adjust for multiple comparisons.

Examples

```
#### simulate data in the long format ####
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")

#### fit Linear Mixed Model ####
eUN.lmm <- lmm(Y ~ visit + X1 + X2 + X5,
               repetition = ~visit|id, structure = "UN", data = dL)

#### Multivariate Wald test ####
## F-tests
anova(eUN.lmm)
anova(eUN.lmm, effects = "all")
anova(eUN.lmm, robust = TRUE, df = FALSE)
summary(anova(eUN.lmm))

## user defined F-test
summary(anova(eUN.lmm, effects = c("X1=0","X2+X5=10")))
print(anova(eUN.lmm, effects = "mean_visit"), columns = add("null"))

## chi2-tests
anova(eUN.lmm, df = FALSE)


## with standard contrast
if(require(multcomp)){
amod <- lmm(breaks ~ tension, data = warpbreaks)
e.amod <- anova(amod, effect = mcp(tension = "Tukey"))
summary(e.amod)
}

#### Likelihood ratio test ####
eUN0.lmm <- lmm(Y ~ X1 + X2, repetition = ~visit|id, structure = "UN", data = dL)
anova(eUN.lmm, eUN0.lmm)

eCS.lmm <- lmm(Y ~ X1 + X2 + X5, repetition = ~visit|id, structure = "CS", data = dL)
anova(eUN.lmm, eCS.lmm)
```

`autoplot.lmm` *Graphical Display For Linear Mixed Models*

Description

Display fitted values or residual plot for the mean, variance, and correlation structure. Can also display quantile-quantile plot relative to the normal distribution.

Usage

```
## S3 method for class 'lmm'
```

```
autoplot(
  object,
  type = "fit",
  type.residual = "normalized",
  obs.alpha = 0,
  obs.size = c(2, 0.5),
  at = NULL,
  time.var = NULL,
  color = TRUE,
  ci = TRUE,
  ci.alpha = 0.25,
  ylim = NULL,
  mean.size = c(3, 1),
  size.text = 16,
  position.errorbar = "identity",
  ...
)

## S3 method for class 'lmm'
plot(x, ...)
```

Arguments

`object`, `x`	a `lmm` object.
`type`	[character] the type of plot • `"fit"`: fitted values over repetitions. • `"qqplot"`: quantile quantile plot of the normalized residuals • `"correlation"`: residual correlation over repetitions • `"scatterplot"`: normalized residuals vs. fitted values (diagnostic for missing non-linear effects), • `"scatterplot2"`: square root of the normalized residuals vs. fitted values (diagnostic for heteroschedasticity), • `"partial"`: partial residual plot.
`type.residual`	[character] the type of residual to be used. Not relevant for `type="fit"`. By default, normalized residuals are used except when requesting a partial residual plot.
`obs.alpha`	[numeric, 0-1] When not NA, transparency parameter used to display the original data by cluster.
`obs.size`	[numeric vector of length 2] size of the point and line for the original data.
`at`	[data.frame] values for the covariates at which to evaluate the fitted values.
`time.var`	[character] x-axis variable for the plot.
`color`	[character] name of the variable in the dataset used to color the curve.
`ci`	[logical] should confidence intervals be displayed?
`ci.alpha`	[numeric, 0-1] When not NA, transparency parameter used to display the confidence intervals.

ylim	[numeric vector of length 2] the lower and higher value of the vertical axis.
mean.size	[numeric vector of length 2] size of the point and line for the mean trajectory.
size.text	[numeric, >0] size of the font used to display text.
position.errorbar	[character] relative position of the errorbars.
...	arguments passed to the predict.lmm or autoplot.residual_lmm functions.

Value

A list with two elements

- data: data used to create the graphical display.
- plot: ggplot object.

Functions

- plot(lmm): Graphical Display For Linear Mixed Models

See Also

plot.lmm for other graphical display (residual plots, partial residual plots).

Examples

```
if(require(ggplot2)){

#### simulate data in the long format ####
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")
dL$X1 <- as.factor(dL$X1)

#### fit Linear Mixed Model ####
eCS.lmm <- lmm(Y ~ visit + X1,
               repetition = ~visit|id, structure = "CS", data = dL, df = FALSE)

plot(eCS.lmm, type = "fit")
autoplot(eCS.lmm, type = "fit")$plot + facet_wrap(~X1)
plot(eCS.lmm, type = "qqplot") ## engine.qqplot = "qqtest"
plot(eCS.lmm, type = "qqplot", engine.qqplot = "qqtest")
plot(eCS.lmm, type = "correlation")
plot(eCS.lmm, type = "scatterplot")
plot(eCS.lmm, type = "scatterplot2")
plot(eCS.lmm, type = "partial", type.residual = "visit")
plot(eCS.lmm, type = "partial", type.residual = "X1")
}
```

`autoplot.partialCor` *Graphical Display For Partial Correlation*

Description

Extract and display the correlation modeled via the linear mixed model.

Usage

```
## S3 method for class 'partialCor'
autoplot(
  object,
  size.text = 16,
  limits = c(-1, 1.00001),
  low = "blue",
  mid = "white",
  high = "red",
  midpoint = 0,
  ...
)

## S3 method for class 'partialCor'
plot(x, ...)
```

Arguments

`object, x`	a `partialCor` object.
`size.text`	[numeric, >0] size of the font used to display text.
`limits`	[numeric vector of length 2] minimum and maximum value of the colorscale relative to the correlation.
`low, mid, high`	[character] color for the the colorscale relative to the correlation.
`midpoint`	[numeric] correlation value associated with the color defined by argument `mid`.
`...`	Not used. For compatibility with the generic method.

Value

A list with two elements

- `data`: data used to create the graphical display.
- `plot`: ggplot object.

Functions

- `plot(partialCor)`: Graphical Display For Partial Correlation

Examples

```
if(require(ggplot2)){
data(gastricbypassL, package = "LMMstar")

e.pCor <- partialCor(c(weight,glucagonAUC)~time, repetition = ~visit|id,
                     data = gastricbypassL)
plot(e.pCor)
}
```

`autoplot.profile_lmm` *Graphical Display of Profile Likelihood*

Description

Graphical representation of the profile likelihood from a linear mixed model

Usage

```
## S3 method for class 'profile_lmm'
autoplot(
  object,
  type = "logLik",
  quadratic = TRUE,
  ci = FALSE,
  size = c(3, 2, 1, 1),
  linetype = c("dashed", "dashed", "dashed"),
  shape = 19,
  scales = "free",
  nrow = NULL,
  ncol = NULL,
  ...
)

## S3 method for class 'profile_lmm'
plot(x, ...)
```

Arguments

`object, x`	an object of class `profile_lmm`, output of the `profile.lmm` function.
`type`	[character] Should the log-likelihood (`"logLik"`) or the ratio to the maximum likelihood (`"ratio"`) be displayed?
`quadratic`	[logical] Should a quadratic approximation of the likelihood be displayed?
`ci`	[logical] Should a 95% confidence intervals obtained from the Wald test (vertical lines) and Likelihood ratio test (horizontal line) be displayed?

size	[numeric vector of length 4] Size of the point for the MLE, width of the line representing the likelihood, width of the corresponding quadratic approximation, and width of the line representing the confidence intervals.
linetype	[integer vector of length 2] type of line used to represent the quadratic approximation of the likelihood and the confidence intervals.
shape	[integer, >0] type of point used to represent the MLE.
scales, nrow, ncol	argument passed to ggplot2::facet_wrap.
...	Not used. For compatibility with the generic method.

Value

A list with three elements

- data.fit: data containing the quadratice approximation of the log-likelihood
- data.ci: data containing the confidence intervals.
- plot: ggplot object.

Functions

- plot(profile_lmm): Display Contour of the log-Likelihood

---

autoplot.residuals_lmm

Graphical Display of the Residuals

Description

Graphical representation of the residuals from a linear mixed model. Require a long format (except for the correlation where both format are accepted) and having exported the dataset along with the residual (argument keep.data when calling residuals.lmm).

Usage

```
## S3 method for class 'residuals_lmm'
autoplot(
  object,
  type = NULL,
  type.residual = NULL,
  by.repetition = TRUE,
  engine.qqplot = "ggplot2",
  add.smooth = TRUE,
  digits.cor = 2,
  size.text = 16,
  scales = "free",
  labeller = "label_value",
```

```
  ...
)

## S3 method for class 'residuals_lmm'
plot(x, ...)
```

Arguments

object, x	an object of class residuals_lmm, output of the residuals.lmm function.
type	[character] Should a qqplot ("qqplot"), or a heatmap of the correlation between residuals ("correlation", require wide format), or a plot of residuals along the fitted values ("scatterplot", require long format) be displayed?
type.residual	[character] Type of residual for which the graphical representation should be made.
by.repetition	[logical] Should a seperate graphical display be made for each repetition.
engine.qqplot	[character] Should ggplot2 or qqtest be used to display quantile-quantile plots? Only used when argument type is "qqplot".
add.smooth	[logical] should a local smoother be used to display the mean of the residual values across the fitted values. Only relevant for when argument type is "scatterplot".
digits.cor	[integer, >0] Number of digit used to display the correlation coefficients? No correlation coefficient is displayed when set to 0. Only used when argument plot is "correlation".
size.text	[numeric, >0] Size of the font used to displayed text when using ggplot2.
scales, labeller	[character] Passed to ggplot2::facet_wrap.
...	Not used. For compatibility with the generic method.

Value

A list with two elements

- data: data used to generate the plot.
- plot: ggplot object.

Functions

- plot(residuals_lmm): Graphical Display of the Residuals

`autoplot.summarizeNA` *Graphical Display of Missing Data Pattern*

Description

Graphical representation of the possible missing data patterns in the dataset.

Usage

```
## S3 method for class 'summarizeNA'
autoplot(
  object,
  variable = NULL,
  size.text = 16,
  add.missing = " missing",
  order.pattern = NULL,
  ...
)

## S3 method for class 'summarizeNA'
plot(x, ...)
```

Arguments

`object, x`	a `summarizeNA` object, output of the `summarizeNA` function.
`variable`	[character] variable for which the missing patterns should be displayed. Only required when the argument `repetition` has been specified when calling `summarizeNA`.
`size.text`	[numeric, >0] size of the font used to display text.
`add.missing`	[logical] should the number of missing values per variable be added to the x-axis tick labels.
`order.pattern`	[numeric vector or character] in which order the missing data pattern should be displayed. Can either be a numeric vector indexing the patterns or a character refering to order the patterns per number of missing values (`"n.missing"`) or number of observations (`"frequency"`).
`...`	Not used. For compatibility with the generic method.

Value

A list with two elements

- `data`: data used to create the graphical display.
- `plot`: ggplot object.

Functions

- `plot(summarizeNA)`: Graphical Display of Missing Data Pattern

`autoplot.Wald_lmm` *Graphical Display For Linear Hypothesis Test*

Description

Graphical Display For Linear Hypothesis Test

Usage

```
## S3 method for class 'Wald_lmm'
autoplot(object, type = "forest", size.text = 16, add.args = NULL, ...)

## S3 method for class 'Wald_lmm'
plot(x, ...)
```

Arguments

`object, x`	a `Wald_lmm` object.
`type`	[character] what to display: a forest plot (`"forest"`) or a heatmap (`"heat"`).
`size.text`	[numeric, >0] size of the font used to display text.
`add.args`	[list] additional arguments used to customized the graphical display. Must be a named list. See details.
`...`	arguments passed to the confint method.

Details

Argument **add.args**: parameters specific to the forest plot:

- `color`: [logical] should the estimates be colored by global null hypothesis, e.g. when testing the effect of a 3 factor covariate, the two corresponding coefficient will have the same color. Alternatively a vector of positive integers giving the color with which each estimator should be displayed.
- `color`: [logical] should the estimates be represented by a different shape per global null hypothesis, e.g. when testing the effect of a 3 factor covariate, the two corresponding coefficient will have the same type of point. Alternatively a vector of positive integers describing the shape to be used for each estimator.
- `ci`: [logical] should confidence intervals be displayed?
- `size.estimate`: [numeric, >0] size of the dot used to display the estimates.
- `size.ci`: [numeric, >0] thickness of the line used to display the confidence intervals.
- `width.ci`: [numeric, >0] width of the line used to display the confidence intervals.
- `size.null`: [numeric, >0] thickness of the line used to display the null hypothesis.

Parameters specific to the heatmap plot:

- `limits`: [numeric vector of length 2] minimum and maximum value of the colorscale relative to the correlation.

- `low`, `mid`, `high`: [character] color for the the colorscale relative to the correlation.
- `midpoint`: [numeric] correlation value associated with the color defined by argument `mid`

Value

A list with two elements

- `data`: data used to create the graphical display.
- `plot`: ggplot object.

Functions

- `plot(Wald_lmm)`: Graphical Display For Linear Hypothesis Test

Examples

```
## From the multcomp package
if(require(datasets) && require(ggplot2)){

## only tests with 1 df
ff <- Fertility ~ Agriculture + Examination + Education + Catholic + Infant.Mortality
e.lmm <- lmm(ff, data = swiss)
e.aovlmm <- anova(e.lmm)

autoplot(e.aovlmm, type = "forest")
autoplot(e.aovlmm, type = "heat") ## 3 color gradient
autoplot(e.aovlmm, type = "heat", add.args = list(mid = NULL)) ## 2 color gradient

## test with more than 1 df
e.lmm2 <- lmm(breaks ~ tension + wool, data = warpbreaks)
e.aovlmm2 <- anova(e.lmm2)
autoplot(e.aovlmm2)
autoplot(e.aovlmm2, add.args = list(color = FALSE, shape = FALSE))
}
```

---

baselineAdjustment *Perform Baseline Adjustment*

Description

Create a new variable based on a time variable and a group variable where groups are constrained to be equal at specific timepoints.

Usage

```
baselineAdjustment(
  object,
  variable,
  repetition,
  constrain,
  new.level = NULL,
  collapse.time = NULL
)
```

Arguments

object	[data.frame] dataset
variable	[character] Column in the dataset to be constrained at specific timepoints.
repetition	[formula] Time and cluster structure, typically ~time|id. See examples below.
constrain	[vector] Levels of the time variable at which the variable is constained.
new.level	[character or numeric] Level used at the constraint. If NULL, then the first level of the variable argument is used.
collapse.time	[character] When not NULL character used to combine the time and argument variable into a new (interaction) variable.

Value

A vector of length the number of rows of the dataset.

Examples

```
data(ncgsL, package = "LMMstar")

## baseline adjustment 1
ncgsL$treat <- baselineAdjustment(ncgsL, variable = "group",
                repetition= ~ visit|id, constrain = 1)
table(treat = ncgsL$treat, visit = ncgsL$visit, group = ncgsL$group)

ncgsL$treattime <- baselineAdjustment(ncgsL, variable = "group",
                  repetition= ~ visit|id, constrain = 1, collapse.time = ".")
table(treattime = ncgsL$treattime, visit = ncgsL$visit, group = ncgsL$group)

## baseline adjustment 2
ncgsL$treat2 <- baselineAdjustment(ncgsL, variable = "group",
                new.level = "baseline",
                repetition= ~ visit|id, constrain = 1)
table(treat = ncgsL$treat2, visit = ncgsL$visit, group = ncgsL$group)

ncgsL$treattime2 <- baselineAdjustment(ncgsL, variable = "group",
                  new.level = "baseline",
                  repetition= ~ visit|id, constrain = 1, collapse.time = ".")
table(treattime = ncgsL$treattime2, visit = ncgsL$visit, group = ncgsL$group)
```

blandAltmanL *Data From The Bland Altman Study (Long Format)*

Description

Data From The Bland Altman Study where two methods to measure the peak expiratory flow rate (PEFR) where compared. This dataset is in the long format (i.e. one line per measurement).

- `id`: patient identifier.
- `replicate`: index of the measurement (first or second).
- `method`: device used to make the measurement (Wright peak flow meter or mini Wright peak flow meter).
- `pefr`: measurement (peak expiratory flow rate).

Usage

```
data(blandAltmanL)
```

References

Bland & Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet, 1986; i: 307-310.

blandAltmanW *Data From The Bland Altman Study (Wide Format)*

Description

Data From The Bland Altman Study where two methods to measure the peak expiratory flow rate (PEFR) where compared. This dataset is in the wide format (i.e. one line per patient).

- `id`: patient identifier.
- `wright1`: first measurement made with a Wright peak flow meter.
- `wright2`: second measurement made with a Wright peak flow meter.
- `mini1`: first measurement made with a mini Wright peak flow meter.
- `mini2`: second measurement made with a mini Wright peak flow meter.

Usage

```
data(blandAltmanW)
```

References

Bland & Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet, 1986; i: 307-310.

bloodpressureL *Data From The Blood Pressure Study (Long Format)*

Description

Data from a cross-over trial comparing the impact of three formulations of a drug on the blood pressure. The study was conducted on 12 male volunteers randomly divided into tree groups and receiving each of the three formulations with a wash-out period of one week.

- `id`: patient identifier.
- `sequence`: sequence of treatment .
- `treatment`: formulation of the treatment A (50 mg tablet) B (100 mg tablet) C (sustained-release formulation capsule)
- `period`: time period (in weeks).
- `duration`: duration of the drug (in hours).

Usage

```
data(bloodpressureL)
```

References

TO ADD

calciumL *Data From The Calcium Supplements Study (Long Format)*

Description

Data from a randomized study including 112 girls at age 11 investigate the effect of a calcium supplement (n=55) vs. placebo (n=57) on bone mineral density over a 2 year follow-up. The clinical question is: does a calcium supplement help to increase bone gain in adolescent women? This dataset is in the long format (i.e. one line per measurement).

- `girl`: patient identifier.
- `grp`: treatment group: calcium supplement (coded `C`) or placebo (coded `P`).
- `visit`: visit index.
- `bmd`: bone mineral density (mg/cm3).
- `time.obs`: visit time (in years).
- `time.num`: scheduled visit time (numeric variable, in years).
- `time.fac`: scheduled visit time (factor variable).

Usage

```
data(calciumL)
```

References

TO ADD

`calciumW` *Data From The Calcium Supplements Study (Wide Format)*

Description

Data from a randomized study including 112 girls at age 11 investigate the effect of a calcium supplement (n=55) vs. placebo (n=57) on bone mineral density over a 2 year follow-up. The clinical question is: does a calcium supplement help to increase bone gain in adolescent women? This dataset is in the wide format (i.e. one line per patient).

- `girl`: patient identifier
- `grp`: treatment group: calcium supplement (coded `C`) or placebo (coded `P`).
- `obstime1`: time after the start of the study at which the first visit took place (in years).
- `obstime2`: time after the start of the study at which the second visit took place (in years).
- `obstime3`: time after the start of the study at which the third visit took place (in years).
- `obstime4`: time after the start of the study at which the fourth visit took place (in years).
- `obstime5`: time after the start of the study at which the fifth visit took place (in years).
- `bmd1`: bone mineral density measured at the first visit (in mg/cm3).
- `bmd2`: bone mineral density measured at the second visit (in mg/cm3).
- `bmd3`: bone mineral density measured at the third visit (in mg/cm3).
- `bmd4`: bone mineral density measured at the fourth visit (in mg/cm3).
- `bmd5`: bone mineral density measured at the fifth visit (in mg/cm3).

Usage

```
data(calciumW)
```

References

Vonesh and Chinchilli 1997. Linear and Nonlinear models for the analysis of repeated measurement (Table 5.4.1 on page 228). New York: Marcel Dekker.

ckdL *CKD long*

Description

TODO

- id: patient identifier.
- allocation:
- sex:
- age:
- visit:
- time:
- pwv:
- aix:
- dropout:

Usage

```
data(ckdL)
```

References

TO ADD

ckdW *CKD wide*

Description

TODO

- id: patient identifier.
- allocation:
- sex:
- age:
- pwv0:
- pwv12:
- pwv24:
- aix0:
- aix12:
- aix24:
- dropout:

Usage

```
data(ckdW)
```

References

TO ADD

`coef.lmm` *Extract Coefficients From a Linear Mixed Model*

Description

Extract coefficients from a linear mixed model.

Usage

```
## S3 method for class 'lmm'
coef(
  object,
  effects = NULL,
  p = NULL,
  transform.sigma = "none",
  transform.k = "none",
  transform.rho = "none",
  transform.names = TRUE,
  ...
)
```

Arguments

`object`	a `lmm` object.
`effects`	[character] Should all coefficients be output (`"all"`), or only coefficients relative to the mean (`"mean"` or `"fixed"`), or only coefficients relative to the variance structure (`"variance"`), or only coefficients relative to the correlation structure (`"correlation"`). Can also be `"ranef"` to output random effect (only for `CS` structure).
`p`	[numeric vector] value of the model coefficients to be used. Only relevant if differs from the fitted values.
`transform.sigma`	[character] Transformation used on the variance coefficient for the reference level. One of `"none"`, `"log"`, `"square"`, `"logsquare"` - see details.
`transform.k`	[character] Transformation used on the variance coefficients relative to the other levels. One of `"none"`, `"log"`, `"square"`, `"logsquare"`, `"sd"`, `"logsd"`, `"var"`, `"logvar"` - see details.
`transform.rho`	[character] Transformation used on the correlation coefficients. One of `"none"`, `"atanh"`, `"cov"` - see details.

`transform.names`
: [logical] Should the name of the coefficients be updated to reflect the transformation that has been used?

`...`
: Not used. For compatibility with the generic method.

Details

transform.sigma:

- `"none"` ouput residual standard error.
- `"log"` ouput log-transformed residual standard error.
- `"square"` ouput residual variance.
- `"logsquare"` ouput log-transformed residual variance.

transform.k:

- `"none"` ouput ratio between the residual standard error of the current level and the reference level.
- `"log"` ouput log-transformed ratio between the residual standard errors.
- `"square"` ouput ratio between the residual variances.
- `"logsquare"` ouput log-transformed ratio between the residual variances.
- `"sd"` ouput residual standard error of the current level.
- `"logsd"` ouput residual log-transformed standard error of the current level.
- `"var"` ouput residual variance of the current level.
- `"logvar"` ouput residual log-transformed variance of the current level.

transform.rho:

- `"none"` ouput correlation coefficient.
- `"atanh"` ouput correlation coefficient after tangent hyperbolic transformation.
- `"cov"` ouput covariance coefficient.

When using a (pure) compound symmetry covariance structure (`structure = "CS"`), estimated random effects can be extracted by setting argument `effects` to `"ranef"`.

Value

A vector with the value of the model coefficients.

See Also

`confint.lmm` or `model.tables.lmm` for a data.frame containing estimates with their uncertainty.

Examples

```
## simulate data in the long format
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")

## fit linear mixed model
eUN.lmm <- lmm(Y ~ X1 + X2 + X5, repetition = ~visit|id, structure = "UN", data = dL, df = FALSE)

## output coefficients
coef(eUN.lmm)
coef(eUN.lmm, effects = "mean")
coef(eUN.lmm, transform.sigma = "none", transform.k = "none", transform.rho = "none")
```

---

confint.lmm *Statistical Inference for Linear Mixed Model*

Description

Compute confidence intervals (CIs) and p-values for the coefficients of a linear mixed model.

Usage

```
## S3 method for class 'lmm'
confint(
  object,
  parm = NULL,
  level = 0.95,
  effects = NULL,
  robust = FALSE,
  null = NULL,
  columns = NULL,
  df = NULL,
  type.information = NULL,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  backtransform = NULL,
  ...
)
```

Arguments

object	a lmm object.
parm	Not used. For compatibility with the generic method.
level	[numeric,0-1] the confidence level of the confidence intervals.

effects	[character] Should the CIs/p-values for all coefficients be output ("all"), or only for mean coefficients ("mean" or "fixed"), or only for variance coefficients ("variance"), or only for correlation coefficients ("correlation").
robust	[logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors. Not feasible for variance or correlation coefficients estimated by REML.
null	[numeric vector] the value of the null hypothesis relative to each coefficient.
columns	[character vector] Columns to be output. Can be any of "estimate", "se", "statistic", "df", "null", "lower", "upper", "p.value".
df	[logical] Should a Student's t-distribution be used to model the distribution of the coefficient. Otherwise a normal distribution is used.
type.information, transform.sigma, transform.k, transform.rho, transform.names	are passed to the vcov method. See details section in coef.lmm.
backtransform	[logical] should the variance/covariance/correlation coefficient be backtransformed?
...	Not used. For compatibility with the generic method.

Value

A data.frame containing some of the following coefficient (in rows):

- column estimate: the estimate.
- column se: the standard error.
- column statistic: the test statistic.
- column df: the degree of freedom.
- column lower: the lower bound of the confidence interval.
- column upper: the upper bound of the confidence interval.
- column null: the null hypothesis.
- column p.value: the p-value relative to the null hypothesis.

See Also

the function anova to perform inference about linear combinations of coefficients and adjust for multiple comparisons.

coef.lmm for a simpler output (e.g. only estimates).
model.tables.lmm for a more detailed output (e.g. with p-value).

Examples

```
#### simulate data in the long format ####
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")

#### fit Linear Mixed Model ####
eUN.lmm <- lmm(Y ~ X1 + X2 + X5, repetition = ~visit|id, structure = "UN", data = dL)
```

```
#### Confidence intervals ####
## based on a Student's t-distribution with transformation
confint(eUN.lmm, effects = "all")
## based on a Student's t-distribution without transformation
confint(eUN.lmm, effects = "all",
        transform.sigma = "none", transform.k = "none", transform.rho = "none")
## based on a Student's t-distribution transformation but not backtransformed
confint(eUN.lmm, effects = "all", backtransform = FALSE)
## based on a Normal distribution with transformation
confint(eUN.lmm, df = FALSE)
```

`confint.mlmm` *Confidence Intervals for Multiple Linear Mixed Model.*

Description

Compute confidence intervals for several linear mixed models.

Usage

```
## S3 method for class 'mlmm'
confint(object, parm = NULL, level = 0.95, method = NULL, ...)
```

Arguments

`object`	an `mlmm` object, output of `mlmm`.
`parm`	Not used. For compatibility with the generic method.
`level`	[numeric,0-1] the confidence level of the confidence intervals.
`method`	[character] type of adjustment for multiple comparisons: one of `"none"`, `"bonferroni"`, `"single-step"`, `"single-step2"`, or `"pool"`.
`...`	other arguments are passed to `confint.Wald_lmm`.

Details

Statistical inference following pooling is performed according to Rubin's rule whose validity requires the congeniality condition of Meng (1994).

References

Meng X. L.(1994). Multiple-imputation inferences with uncongenial sources of input. Statist. Sci.9, 538–58.

confint.Wald_lmm *Confidence Intervals for Multivariate Wald Tests*

Description

Compute confidence intervals for linear hypothesis tests, possibly with adjustment for multiple comparisons.

Usage

```
## S3 method for class 'Wald_lmm'
confint(
  object,
  parm,
  level = 0.95,
  method = NULL,
  columns = NULL,
  backtransform = NULL,
  ...
)
```

Arguments

object	a Wald_lmm object
parm	Not used. For compatibility with the generic method.
level	[numeric, 0-1] nominal coverage of the confidence intervals.
method	[character] type of adjustment for multiple comparisons, one of "none", "bonferroni", ..., "fdr", "single-step", "single-step2". Alternatively, a method for combining the estimates, one of "average", "pool.se", "pool.gls", "pool.gls1", "pool.rubin".
columns	[character vector] Columns to be output. Can be any of "estimate", "se", "statistic", "df", "null", "lower", "upper", "p.value".
backtransform	[logical] should the estimates, standard errors, and confidence intervals be back-transformed?
...	Not used. For compatibility with the generic method.

Details

Adjustment for multiple comparisons: available methods are:

- "none", "bonferroni", "single-step2"
- "holm", "hochberg", "hommel", "BH", "BY", "fdr": adjustment performed by [stats::p.adjust()], no confidence interval is computed.
- "single-step", "free", "Westfall", "Shaffer": adjustment performed by [multcomp::glht()], for all but the first method no confidence interval is computed.

Note: method "single-step" adjust for multiple comparisons using quantiles of the multivariate Student's t-distribution, assuming equal degrees of freedom in the marginal. When degrees of freedom differs between individual hypotheses, method "single-step2" is recommended. It simulates data using copula whose marginal distributions are Student's t-distribution (with possibly different degrees of freedom) and elliptical copula with parameters the estimated correlation between the test statistics (via the copula package). It then computes the frequency at which the simulated maximum exceed the observed maximum and appropriate quantile of simulated maximum for the confidence interval.

Pooling estimates: available methods are:

- "average": average estimates
- "pool.fixse": weighted average of the estimates, with weights being the inverse of the squared standard error. The uncertainty about the weights is neglected.
- "pool.se": weighted average of the estimates, with weights being the inverse of the squared standard error. The uncertainty about the weights is computed under independence of the variance parameters between models.
- "pool.gls": weighted average of the estimates, with weights being based on the variance-covariance matrix of the estimates. When this matrix is singular, its spectral decomposition is truncated when the correspodning eigenvalues are below 10^{-12}. The uncertainty about the weights is neglected.
- "pool.gls1": similar to "pool.gls" but ensure that the weights are at most 1 in absolute value by shrinking toward the average.
- "pool.rubin": average of the estimates and compute the uncertainty according to Rubin's rule (Barnard et al. 1999).

References

Barnard and Rubin, Small-sample degrees of freedom with multiple imputation. Biometrika (1999), 86(4):948-955.

CS *Compound Symmetry Structure*

Description

Variance-covariance structure where the residuals have constant variance and correlation, possibly only within certain covariate levels. Can be stratified on a categorical variable.

Usage

```
CS(formula, var.cluster, var.time, heterogeneous = TRUE, add.time)
```

Arguments

formula	formula indicating on which variable to stratify the residual variance and correlation (left hand side) and variables influencing the residual variance and correlation (right hand side).
var.cluster	[character] cluster variable.
var.time	[character] time variable.
heterogeneous	[logical] when covariates are used for the correlation structure, should correlation parameters should be specific to each level of the covariate?
add.time	not used.

Details

A typical formula would be ~1, indicating a variance constant over time and the same correlation between all pairs of times.

Value

An object of class CS that can be passed to the argument structure of the lmm function.

Examples

```
CS(~1, var.cluster = "id", var.time = "time")
CS(gender~1, var.cluster = "id", var.time = "time")
CS(list(~time,~1), var.cluster = "id", var.time = "time")
CS(list(gender~time,gender~1), var.cluster = "id", var.time = "time")
```

CUSTOM *Custom Structure*

Description

Variance-covariance structure specified by the user.

Usage

```
CUSTOM(
  formula,
  var.cluster,
  var.time,
  FCT.sigma,
  dFCT.sigma = NULL,
  d2FCT.sigma = NULL,
  init.sigma,
  FCT.rho,
  dFCT.rho = NULL,
```

```
  d2FCT.rho = NULL,
  init.rho,
  add.time
)
```

Arguments

formula	formula indicating variables influencing the residual variance and correlation (right hand side).
var.cluster	[character] cluster variable.
var.time	[character] time variable.
FCT.sigma	[function] take as argument the model parameters, time, and design matrix. Output the vector of residuals standard deviations.
dFCT.sigma	[list of vectors] list whose elements are the first derivative of argument FCT.sigma.
d2FCT.sigma	[list of vectors] list whose elements are the second derivative of argument FCT.sigma (no cross-terms).
init.sigma	[numeric vector] initial value for the variance parameters.
FCT.rho	[function] take as argument the model parameters, time, and design matrix. Output the matrix of residuals correlation.
dFCT.rho	[list of matrices] list whose elements are the first derivative of argument FCT.rho.
d2FCT.rho	[list of matrices] list whose elements are the second derivative of argument FCT.rho (no cross-terms).
init.rho	[numeric vector] initial value for the correlation parameters.
add.time	not used.

Value

An object of class CUSTOM that can be passed to the argument structure of the lmm function.

Examples

```
## Compound symmetry structure
CUSTOM(~1,
       FCT.sigma = function(p,time,X){rep(p,length(time))},
       init.sigma = c("sigma"=1),
       dFCT.sigma = function(p,time,X){list(sigma = rep(1,length(time)))},
       d2FCT.sigma = function(p,time,X){list(sigma = rep(0,length(time)))},
       FCT.rho = function(p,time,X){
           matrix(p,length(time),length(time))+diag(1-p,length(time),length(time))
       },
       init.rho = c("rho"=0.5),
       dFCT.rho = function(p,time,X){
        list(rho = matrix(1,length(time),length(time))-diag(1,length(time),length(time)))
       },
       d2FCT.rho = function(p,time,X){list(rho = matrix(0,length(time),length(time)))}
)
```

```
## 2 block structure
rho.2block <- function(p,time,X){
   n.time <- length(time)
   rho <- matrix(0, nrow = n.time, ncol = n.time)
   rho[1,2] <- rho[2,1] <- rho[4,5] <- rho[5,4] <- p["rho1"]
   rho[1,3] <- rho[3,1] <- rho[4,6] <- rho[6,4] <- p["rho2"]
   rho[2,3] <- rho[3,2] <- rho[5,6] <- rho[6,5] <- p["rho3"]
   rho[4:6,1:3] <- rho[1:3,4:6] <- p["rho4"]
   return(rho)
}
drho.2block <- function(p,time,X){
   n.time <- length(time)
   drho <- list(rho1 = matrix(0, nrow = n.time, ncol = n.time),
                rho2 = matrix(0, nrow = n.time, ncol = n.time),
                rho3 = matrix(0, nrow = n.time, ncol = n.time),
                rho4 = matrix(0, nrow = n.time, ncol = n.time))
   drho$rho1[1,2] <- drho$rho1[2,1] <- drho$rho1[4,5] <- drho$rho1[5,4] <- 1
   drho$rho2[1,3] <- drho$rho2[3,1] <- drho$rho2[4,6] <- drho$rho2[6,4] <- 1
   drho$rho3[2,3] <- drho$rho3[3,2] <- drho$rho3[5,6] <- drho$rho3[6,5] <- 1
   drho$rho4[4:6,1:3] <- drho$rho4[1:3,4:6] <- 1
   return(drho)
}
d2rho.2block <- function(p,time,X){
   n.time <- length(time)
   d2rho <- list(rho1 = matrix(0, nrow = n.time, ncol = n.time),
                 rho2 = matrix(0, nrow = n.time, ncol = n.time),
                 rho3 = matrix(0, nrow = n.time, ncol = n.time),
                 rho4 = matrix(0, nrow = n.time, ncol = n.time))
   return(d2rho)
}

CUSTOM(~variable,
       FCT.sigma = function(p,time,X){rep(p,length(time))},
       dFCT.sigma = function(p,time,X){list(sigma=rep(1,length(time)))},
       d2FCT.sigma = function(p,time,X){list(sigma=rep(0,length(time)))},
       init.sigma = c("sigma"=1),
       FCT.rho = rho.2block,
       dFCT.rho = drho.2block,
       d2FCT.rho = d2rho.2block,
       init.rho = c("rho1"=0.25,"rho2"=0.25,"rho3"=0.25,"rho4"=0.25))
```

`df.residual.lmm` *Residuals Degrees of Freedom*

Description

Residuals degrees of freedom. Computed as the sum of squared normalized residuals

Usage

```
## S3 method for class 'lmm'
df.residual(object, ...)
```

Arguments

object	a lmm object.
...	Passed to residuals.lmm.

Value

A numeric value

---

dummy.coef.lmm *Marginal Mean Values For Linear Mixed Model*

Description

Compute the marginal mean (via the emmeans package) for each combination of categorical covariates. When there is no numeric covariate, this outputs all the mean values fitted by the model.

Usage

```
## S3 method for class 'lmm'
dummy.coef(object, drop = TRUE, ...)
```

Arguments

object	a lmm object.
drop	[logical] should combinations of covariates that do no exist in the original dataset be removed?
...	arguments passed to emmeans.

Value

A data.frame containing the level for which the means have been computed (if more than one), the estimated mean (estimate), standard error (se), degree of freedom (df), and 95% confidence interval (lower and upper).

`estfun.lmm` *Extract the Score Function for Multcomp*

Description

Extract the Score Function for Multcomp. For internal use.

Usage

```
## S3 method for class 'lmm'
estfun(x, ...)
```

Arguments

`x`	a `lmm` object.
`...`	Not used. For compatibility with the generic method.

Value

A matrix containing the score function for each model parameter (columns) relative to each cluster (rows).

Examples

```
## simulate data in the long format
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")

## fit Linear Mixed Model
eUN.lmm <- lmm(Y ~ X1 + X2 + X5, repetition = ~visit|id, structure = "UN", data = dL, df = FALSE)

## test multiple linear hypotheses
if(require(multcomp)){
LMMstar.options(effects = c("mean"))
e.glht <- multcomp::glht(eUN.lmm)
e.glht$linfct
}
```

`estimate.lmm` *Delta Method for Mixed Models*

Description

Perform a first order delta method

Usage

```
## S3 method for class 'lmm'
estimate(
  x,
  f,
  df = !is.null(x$df),
  robust = FALSE,
  type.information = NULL,
  level = 0.95,
  method.numDeriv = NULL,
  average = FALSE,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  ...
)
```

Arguments

`x`	a `lmm` object.
`f`	[function] function of the model coefficient computing the parameter(s) of interest. Can accept extra-arguments.
`df`	[logical] Should degree of freedom, computed using Satterthwaite approximation, for the parameter of interest be output.
`robust`	[logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors.
`type.information`	[character] Should the expected information be used (i.e. minus the expected second derivative) or the observed inforamtion (i.e. minus the second derivative).
`level`	[numeric,0-1] the confidence level of the confidence intervals.
`method.numDeriv`	[character] method used to approximate the gradient: either `"simple"` or `"Richardson"`. Passed to `numDeriv::jacobian`.
`average`	[logical] is the estimand the average output of argument `f`? Otherwise consider each individual output of argument `f`.

transform.sigma
: [character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.

transform.k
: [character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details.

transform.rho
: [character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.

...
: extra arguments passed to f.

Examples

```
if(require(lava)){

#### Random effect ####
set.seed(10)
dL <- sampleRem(1e2, n.times = 3, format = "long")
e.lmm1 <- lmm(Y ~ X1+X2+X3, repetition = ~visit|id, structure = "CS", data = dL)
coef(e.lmm1, effects = "ranef")
e.ranef <- estimate(e.lmm1, f  = function(p){coef(e.lmm1, p = p, effects = "ranef")})
e.ranef

if(require(ggplot2)){
df.gg <- cbind(index = 1:NROW(e.ranef), e.ranef)
gg.ranef <- ggplot(df.gg, aes(x = index, y=estimate, ymin=lower, ymax = upper))
gg.ranef + geom_point() + geom_errorbar() + ylab("estimated random effect") + xlab("id")
}

#### ANCOVA via mixed model ####
set.seed(10)
d <- sampleRem(1e2, n.time = 2)
e.ANCOVA1 <- lm(Y2~Y1+X1, data = d)

if(require(reshape2)){
   dL2 <- melt(d, id.vars = c("id","Y1","X1"),  measure.vars = c("Y1","Y2"))
   e.lmm <- lmm(value ~ variable + variable:X1, data = dL2, repetition = ~variable|id)

   e.delta <- estimate(e.lmm, function(p){
       c(Y1 = p["rho(Y1,Y2)"]*p["k.Y2"],
         X1 = p["variableY2:X1"]-p["k.Y2"]*p["rho(Y1,Y2)"]*p["variableY1:X1"])
})
   ## same estimate and similar standard errors.
   e.delta
   summary(e.ANCOVA1)$coef
   ## Degrees of freedom are a bit off though
}

}
```

EXP *Exponential Structure*

Description

Variance-covariance structure where the residuals have a correlation decreasing exponentially, Can be stratified on a categorical variable.

Usage

```
EXP(formula, var.cluster, var.time, nugget = FALSE, add.time)
```

Arguments

formula	formula indicating on which variable to stratify the residual variance and correlation (left hand side) and variables influencing the residual variance and correlation (right hand side).
var.cluster	[character] cluster variable.
var.time	[character] time variable.
nugget	[logical] whether a nugget effect is present.
add.time	not used.

Details

A typical formula would be ~1, indicating a variance constant over time and correlation with exponential decrease over time.

Inspired from nlme::corExp where if K denotes the nugget effect and ρ the time effect, the correlation between two observations with a time gap dt is $exp(-\rho dt)$ when no nugget effect is present and $(1-K)exp(-\rho dt)$ when a nugget effect is assumed.

Value

An object of class EXP that can be passed to the argument structure of the lmm function.

Examples

```
EXP(var.cluster = "id", var.time = "time", add.time = TRUE)
EXP(~space, var.cluster = "id", var.time = "time", add.time = TRUE)
EXP(list(~space,~space), var.cluster = "id", var.time = "time", add.time = TRUE)
```

fitted.lmm *Predicted Mean Value For Linear Mixed Model*

Description

Predicted Mean Value For Linear Mixed Model

Usage

```
## S3 method for class 'lmm'
fitted(
  object,
  newdata = NULL,
  format = "long",
  keep.newdata = FALSE,
  impute = FALSE,
  se.impute = FALSE,
  ...
)
```

Arguments

object	a lmm object.
newdata	[data.frame] the covariate values for each cluster.
format	[character] Should the predicted mean be output relative as a vector ("long"), or as a matrix with in row the clusters and in columns the outcomes ("wide").
keep.newdata	[logical] Should the argument newdata be output along side the predicted values? The output will then be a data.frame.
impute	[logical] Should the missing data in the outcome be imputed based on covariates and other outcome values from the same cluster.
se.impute	[character] If FALSE the most likely value is imputed. Otherwise the imputed value is sampled from a normal distribution. The value of the argument determine which standard deviation is used: all uncertainty about the predicted value ("total"), only uncertainty related to the estimation of the model parameters ("estimate"), or only uncertainty related to the residual variance of the outcome ("residual"). Passed to predict.lmm.
...	Not used. For compatibility with the generic method.

Value

When format="wide", a data.frame with as many rows as clusters. When format="long" or keep.newdata==TRUE, a data.frame with as many rows as observations. Otherwise:

- if impute=FALSE a vector of length the number of row of newdata containing the fitted values (i.e. based on the covariates only).

- if impute=TRUE a vector of length the number of missing values in the outcome of newdata containing the cluster-specific conditional means (i.e. based on the covariates and outcome measurements from the same cluster).

When keep.newdata==TRUE, a dataframe with an additional column containing the fitted values (i.e. based on the covariates only). If impute=TRUE, the missing value in the outcome column are replaced by the cluster-specific conditional means (i.e. based on the covariates and outcome measurements from the same cluster).

Examples

```
#### simulate data in the long format ####
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")

#### fit Linear Mixed Model ####
eCS.lmm <- lmm(Y ~ X1 + X2 + X5, repetition = ~visit|id,
               structure = "CS", data = dL, df = FALSE)

## prediction
fitted(eCS.lmm)
fitted(eCS.lmm, newdata = data.frame(X1 = 1, X2 = 2, X5 = 3))
fitted(eCS.lmm, newdata = data.frame(X1 = 1, X2 = 2, X5 = 3), keep.newdata = TRUE)

#### fit Linear Mixed Model with missing data ####
dL2 <- dL
dL2[3,"Y"] <- NA
eCS2.lmm <- lmm(Y ~ X1 + X2 + X5, repetition = ~visit|id,
                structure = "CS", data = dL2, df = FALSE)

## most likely value to impute
fitted(eCS2.lmm, impute = TRUE)
head(fitted(eCS2.lmm, impute = TRUE, keep.newdata = TRUE))

## multiple imputation
dL2.imp1 <- data.frame(imp = "1",
    fitted(eCS2.lmm, impute = TRUE, se.impute = "total", keep.newdata = TRUE))
dL2.imp2 <- data.frame(imp = "2",
    fitted(eCS2.lmm, impute = TRUE, se.impute = "total", keep.newdata = TRUE))
head(dL2.imp1)
head(dL2.imp2)
```

---

gastricbypassL *Data From The Gastric Bypass Study (Long Format)*

Description

Data from the gastric bypass study where the bodyweight and serum glucagon (a gut hormone) were measured in 20 obese subjects prior and after gastric bypass surgery. This dataset is in the long format (i.e. one line per measurement).

- id: patient identifier.
- visit: the visit index.
- time: the time at which the visit took place.
- weight: bodyweight (in kg) measured during the visit.
- glucagonAUC: glucagon measured during the visit.

Usage

```
data(gastricbypassL)
```

References

The effect of Roux-en-Y gastric bypass surgery on the gut mucosal gene expression profile and circulating gut hormones. https://easddistribute.m-anage.com/from.storage?image=4iBH9mRQm1kfeEHULC2Cxovdly

---

gastricbypassW *Data From The Gastric Bypass Study (Wide Format)*

Description

Data from the gastric bypass study where the bodyweight and serum glucagon (a gut hormone) were measured in 20 obese subjects prior and after gastric bypass surgery. This dataset is in the wide format (i.e. one line per patient).

- id: patient identifier.
- weight1: bodyweight (in kg) 3 months before surgery.
- weight2: bodyweight (in kg) 1 week before surgery.
- weight3: bodyweight (in kg) 1 week after surgery.
- weight4: bodyweight (in kg) 3 months after surgery.
- glucagonAUC1: glucagon value 3 months before surgery.
- glucagonAUC2: glucagon value 1 week before surgery.
- glucagonAUC3: glucagon value 1 week after surgery.
- glucagonAUC4: glucagon value 3 months after surgery.

Usage

```
data(gastricbypassW)
```

References

The effect of Roux-en-Y gastric bypass surgery on the gut mucosal gene expression profile and circulating gut hormones. https://easddistribute.m-anage.com/from.storage?image=4iBH9mRQm1kfeEHULC2Cxovdly

getVarCov.lmm *Depreciated Extractor of the Residual Variance-Covariance Matrix*

Description

Depreciated extractor of the residual variance-covariance matrix.

Usage

```
## S3 method for class 'lmm'
getVarCov(obj, ...)
```

Arguments

obj	a lmm object.
...	other arguments.

Value

Nothing

See Also

sigma.lmm

ID *identity Structure*

Description

Variance-covariance structure where the residuals are independent and identically distribution. Can be stratified on a categorical variable.

Usage

```
ID(formula, var.cluster, var.time, add.time)
```

Arguments

formula	formula indicating on which variable to stratify the residual variance (left hand side).
var.cluster	[character] cluster variable.
var.time	[character] time variable.
add.time	not used.

Details

A typical formula would be ~1.

Value

An object of class IND that can be passed to the argument structure of the lmm function.

Examples

```
ID(NULL, var.cluster = "id", var.time = "time")
ID(~1, var.cluster = "id", var.time = "time")
ID(~gender, var.cluster = "id", var.time = "time")
ID(gender~1, var.cluster = "id", var.time = "time")
```

---

iid.lmm *Extract the Influence Function From a Linear Mixed Model*

Description

Extract the influence function from a linear mixed model.

Usage

```
## S3 method for class 'lmm'
iid(
  x,
  effects = "mean",
  robust = TRUE,
  type.information = NULL,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  ...
)
```

Arguments

x	a lmm object.
effects	[character] Should the variance-covariance matrix for all coefficients be output ("all"), or only for coefficients relative to the mean ("mean" or "fixed"), or only for coefficients relative to the variance structure ("variance"), or only for coefficients relative to the correlation structure ("correlation").
robust	[logical] If FALSE the influence function is rescaled to match the model-based standard errors. The correlation however will not (necessarily) match the model-based correlation.

type.information
: [character] Should the expected information be used (i.e. minus the expected second derivative) or the observed inforamtion (i.e. minus the second derivative).

transform.sigma
: [character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.

transform.k
: [character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details.

transform.rho
: [character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.

transform.names
: [logical] Should the name of the coefficients be updated to reflect the transformation that has been used?

...
: Not used. For compatibility with the generic method.

IND *Independence Structure*

Description

Variance-covariance structure where the residuals are independent. Can be stratified on a categorical variable.

Usage

```
IND(formula, var.cluster, var.time, add.time)
```

Arguments

formula
: formula indicating variables influencing the residual variance, using either as a multiplicative factor (right hand side) or stratification (left hand side) to model their effect.

var.cluster
: [character] cluster variable.

var.time
: [character] time variable.

add.time
: Should the default formula (i.e. when NULL) contain a time effect.

Details

A typical formula would be either ~1 indicating constant variance or ~time indicating a time dependent variance.

Value

An object of class IND that can be passed to the argument structure of the lmm function.

Examples

```
IND(NULL, var.cluster = "id", var.time = "time", add.time = TRUE)
IND(~1, var.cluster = "id", var.time = "time")
IND(gender~1, var.cluster = "id", var.time = "time")

IND(~time, var.cluster = "id", var.time = "time")
IND(gender~time, var.cluster = "id", var.time = "time")
IND(~time+gender, var.cluster = "id", var.time = "time")
```

`information.lmm` *Extract The Information From a Linear Mixed Model*

Description

Extract or compute the (expected) second derivative of the log-likelihood of a linear mixed model.

Usage

```
## S3 method for class 'lmm'
information(
  x,
  effects = NULL,
  data = NULL,
  p = NULL,
  indiv = FALSE,
  type.information = NULL,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  ...
)
```

Arguments

`x`	a `lmm` object.
`effects`	[character] Should the information relative to all coefficients be output (`"all"` or `"fixed"`), or only coefficients relative to the mean (`"mean"`), or only coefficients relative to the variance and correlation structure (`"variance"` or `"correlation"`).
`data`	[data.frame] dataset relative to which the information should be computed. Only relevant if differs from the dataset used to fit the model.
`p`	[numeric vector] value of the model coefficients at which to evaluate the information. Only relevant if differs from the fitted values.
`indiv`	[logical] Should the contribution of each cluster to the information be output? Otherwise output the sum of all clusters of the derivatives.

type.information
: [character] Should the expected information be computed (i.e. minus the expected second derivative) or the observed inforamtion (i.e. minus the second derivative).

transform.sigma
: [character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.

transform.k
: [character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details.

transform.rho
: [character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.

transform.names
: [logical] Should the name of the coefficients be updated to reflect the transformation that has been used?

...
: Not used. For compatibility with the generic method.

Details

For details about the arguments **transform.sigma**, **transform.k**, **transform.rho**, see the documentation of the coef.lmm function.

Value

When argument indiv is FALSE, a matrix with the value of the infroamtion relative to each pair of coefficient (in rows and columns) and each cluster (in rows). When argument indiv is TRUE, a 3-dimensional array with the value of the information relative to each pair of coefficient (dimension 2 and 3) and each cluster (dimension 1).

---

levels.lmm *Contrasts and Reference Level*

Description

Contrasts and reference level used when modeling the mean in a linear mixed modek.

Usage

```
## S3 method for class 'lmm'
levels(x)
```

Arguments

x
: an lmm object

Value

a list with two elements

- all: contrast matrix for each categorical or factor variable
- reference: reference level: one value for each categorical variable

---

`lmm` *Fit Linear Mixed Model*

Description

Fit a linear mixed model defined by a mean and a covariance structure. g

Usage

```
lmm(
  formula,
  repetition,
  structure,
  data,
  weights = NULL,
  scale.Omega = NULL,
  method.fit = NULL,
  df = NULL,
  type.information = NULL,
  trace = NULL,
  control = NULL
)
```

Arguments

`formula`	[formula] Specify the model for the mean. On the left hand side the outcome and on the right hand side the covariates affecting the mean value. E.g. Y ~ Gender + Gene.
`repetition`	[formula] Specify the structure of the data: the time/repetition variable and the grouping variable, e.g. ~ timelid.
`structure`	[character] type of covariance structure, either `"CS"` (compound symmetry) or `"UN"` (unstructured).
`data`	[data.frame] dataset (in the long format) containing the observations.
`weights`	[formula or character] variable in the dataset used to weight the log-likelihood and its derivative. Should be constant within cluster.
`scale.Omega`	[formula or character] variable in the dataset used to rescale the residual variance-covariance matrix. Should be constant within cluster.
`method.fit`	[character] Should Restricted Maximum Likelihoood (`"REML"`) or Maximum Likelihoood (`"ML"`) be used to estimate the model parameters?

df [logical] Should the degree of freedom be computed using a Satterthwaite approximation?

type.information [character] Should the expected information be computed (i.e. minus the expected second derivative) or the observed inforamtion (i.e. minus the second derivative).

trace [interger, >0] Show the progress of the execution of the function.

control [list] Control values for the optimization method. The element optimizer indicates which optimizer to use and additional argument will be pass to the optimizer.

Details

Computation time the lmm has not been developped to be a fast function as, by default, it uses REML estimation with the observed information matrix and uses a Satterthwaite approximation to compute degrees of freedom (this require to compute the third derivative of the log-likelihood which is done by numerical differentiation). The computation time can be substantially reduced by using ML estimation with the expected information matrix and no calculation of degrees of freedom: arguments method.fit="ML", type.information="expected", df=FALSE. This will, however, lead to less accurate p-values and confidence intervals in small samples.

By default, the estimation of the model parameters will be made using a Newton Raphson algorithm. This algorithm does not ensure that the residual covariance matrix is positive definite and therefore may sometimes fail. When using an unstructured pattern, i.e. structure="UN" the nlme::gls may be preferable as it can ensures positve definitness. See argument optimizer in LMMstar.options.

Argument control: when using the optimizer "FS", the following elements can be used

- init: starting values for the model parameters.
- n.iter: maximum number of interations of the optimization algorithm.
- tol.score: score value below which convergence has been reached.
- tol.param: difference in estimated parameters from two successive iterations below which convergence has been reached.
- trace: display progress of the optimization procedure.

Argument repetition: when numeric, it will be converted into a factor variable, possibly adding a leading 0 to preserve the ordering. This transformation may cause inconsistency when combining results between different lmm object. This is why the grouping variable should preferably be of type character or factor.

Value

an object of class lmm containing the estimated parameter values, the residuals, and relevant derivatives of the likelihood.

See Also

summary.lmm for a summary of the model fit.
model.tables.lmm for a data.frame containing estimates with their uncertainty.

plot.lmm for a graphical display of the model fit or diagnostic plots.
levels.lmm to display the reference level.
anova.lmm for testing linear combinations of coefficients (F-test, multiple Wald tests)
getVarCov.lmm for extracting estimated residual variance-covariance matrices.
residuals.lmm for extracting residuals or creating residual plots (e.g. qqplots).
predict.lmm for evaluating mean and variance of the outcome conditional on covariates or other outcome values.

Examples

```
#### 1- simulate data in the long format ####
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")
dL$X1 <- as.factor(dL$X1)
dL$X2 <- as.factor(dL$X2)

#### 2- fit Linear Mixed Model ####
eCS.lmm <- lmm(Y ~ X1 * X2 + X5, repetition = ~visit|id, structure = "CS", data = dL)

logLik(eCS.lmm) ## -670.9439
summary(eCS.lmm)


#### 3- estimates ####
## reference level
levels(eCS.lmm)$reference
## mean parameters
coef(eCS.lmm)
model.tables(eCS.lmm)
confint(eCS.lmm)

if(require(emmeans)){
  dummy.coef(eCS.lmm)
}

## all parameters
coef(eCS.lmm, effects = "all")
model.tables(eCS.lmm, effects = "all")
confint(eCS.lmm, effects = "all")

## variance-covariance structure
sigma(eCS.lmm)

#### 4- diagnostic plots ####
quantile(residuals(eCS.lmm))
quantile(residuals(eCS.lmm, type = "normalized"))

## Not run:
if(require(ggplot2)){
  ## investigate misspecification of the mean structure
  plot(eCS.lmm, type = "scatterplot")
  ## investigate misspecification of the variance structure
```

```
  plot(eCS.lmm, type = "scatterplot2")
  ## investigate misspecification of the correlation structure
  plot(eCS.lmm, type = "correlation")
  ## investigate misspecification of the residual distribution
  plot(eCS.lmm, type = "qqplot")
}

## End(Not run)

#### 5- statistical inference ####
anova(eCS.lmm) ## effect of each variable
anova(eCS.lmm, effects = "X11-X21=0") ## test specific coefficient
## test several hypothese with adjustment for multiple comparisons
summary(anova(eCS.lmm, effects = c("X11=0","X21=0")))

#### 6- prediction ####
## conditional on covariates
newd <- dL[1:3,]
predict(eCS.lmm, newdata = newd, keep.newdata = TRUE)
## conditional on covariates and outcome
newd <- dL[1:3,]
newd$Y[3] <- NA
predict(eCS.lmm, newdata = newd, type = "dynamic", keep.newdata = TRUE)

#### EXTRA ####
if(require(mvtnorm)){
## model for the average over m replicates
## (only works with independent replicates)
Sigma1 <- diag(1,1,1); Sigma5 <- diag(1,5,5)
n <- 100
dfW <- rbind(data.frame(id = 1:n, rep = 5, Y = rowMeans(rmvnorm(n, sigma = Sigma5))),
             data.frame(id = (n+1):(2*n), rep = 1, Y = rmvnorm(n, sigma = Sigma1)))

e.lmmW <- lmm(Y~1, data = dfW, scale.Omega=~rep, control = list(optimizer = "FS"))
e.lmm0 <- lmm(Y~1, data = dfW, control = list(optimizer = "FS"))
model.tables(e.lmmW, effects = "all")
model.tables(e.lmm0, effects = "all")
## TRUE standard error is 1

}
```

`LMMstar.options` *Global options for LMMstar package*

Description

Update or select global options for the LMMstar package.

Usage

```
LMMstar.options(..., reinitialise = FALSE)
```

Arguments

...	options to be selected or updated
reinitialise	should all the global parameters be set to their default value

Details

The options are:

- backtransform.confint [logical]: should variance/covariance/correlation estimates be back-transformed when they are transformed on the log or atanh scale. Used by `confint`.
- columns.anova [character vector]: columns to ouput when using `anova` with argument `ci=TRUE`.
- columns.confint [character vector]: columns to ouput when using `confint`.
- columns.summary [character vector]: columns to ouput when displaying the model coefficients using `summary`.
- df [logical]: should approximate degrees of freedom be computed for Wald and F-tests. Used by `lmm`, `anova`, `predict`, and `confint`.
- drop.X [logical]: should columns causing non-identifiability of the model coefficients be dropped from the design matrix. Used by `lmm`.
- effects [character]: parameters relative to which estimates, score, information should be output.
- min.df [integer]: minimum possible degree of freedom. Used by `confint`.
- method.fit [character]: objective function when fitting the Linear Mixed Model (REML or ML). Used by `lmm`.
- method.numDeriv [character]: type used to approximate the third derivative of the log-likelihood (when computing the degrees of freedom). Can be `"simple"` or `"Richardson"`. See `numDeriv::jacobian` for more details. Used by `lmm`.
- n.sampleCopula [integer]: number of samples used to compute confidence intervals and p-values adjusted for multiple comparisons via `"single-step2"`. Used by `confint.Wald_lmm`.
- optimizer [character]: method used to estimate the model parameters: can the `nlme::gls` (`"gls"`) or an algorithm combine fisher scoring for the variance parameters and generalized least squares for the mean parameters (`"FS"`).
- param.optimizer [numeric vector]: default option for the `FS` optimization routine: maximum number of gradient descent iterations (`n.iter`), maximum acceptable score value (`tol.score`), maximum acceptable change in parameter value (`tol.param`).
- precompute.moments [logical]: Should the cross terms between the residuals and design matrix be pre-computed. Useful when the number of subject is substantially larger than the number of mean paramters.
- trace [logical]: Should the progress of the execution of the `lmm` function be displayed?
- tranform.sigma, tranform.k, tranform.rho: transformation used to compute the confidence intervals/p-values for the variance and correlation parameters. See the detail section of the coef function for more information. Used by `lmm`, `anova` and `confint`.

- type.information [character]: Should the expected or observed information (`"expected"` or `"observed"`) be used to perform statistical inference? Used by `lmm`, `anova` and `confint`.

Value

A list containing the default options.

`logLik.lmm` *Extract The Log-Likelihood From a Linear Mixed Model*

Description

Extract or compute the log-likelihood of a linear mixed model.

Usage

```
## S3 method for class 'lmm'
logLik(object, data = NULL, p = NULL, indiv = FALSE, ...)
```

Arguments

`object`	a `lmm` object.
`data`	[data.frame] dataset relative to which the log-likelihood should be computed. Only relevant if differs from the dataset used to fit the model.
`p`	[numeric vector] value of the model coefficients at which to evaluate the log-likelihood. Only relevant if differs from the fitted values.
`indiv`	[logical] Should the contribution of each cluster to the log-likelihood be output? Otherwise output the sum of all clusters of the derivatives.
`...`	Not used. For compatibility with the generic method.

Details

indiv: only relevant when using maximum likelihood. Must be `FALSE` when using restricted maximum likelihood.

Value

A numeric value (total logLikelihood) or a vector of numeric values, one for each cluster (cluster specific logLikelihood).

manifest.lmm *Variables Involved in a Linear Mixed Model*

Description

Extract the variables used by the linear mixed model.

Usage

```
## S3 method for class 'lmm'
manifest(x, effects = "all", original = TRUE, ...)
```

Arguments

x	a lmm object.
effects	[character] Should all variable be output ("all"), or only those related to the outcome ("outcome"), mean ("mean"), variance ("variance"), correlation ("correlation"), time ("time"), cluster ("cluster"), strata ("strata")?
original	[logical] Should only the variables present in the original data be output? When NULL, variables internally created to fill absent variables will be added to the output. When FALSE, variables internally created are output instead of the original variable for time, cluster, and strata.
...	not used. For compatibility with the generic function

Value

A character vector

mlmm *Fit Multiple Linear Mixed Model*

Description

Fit several linear mixed models, extract relevant coefficients, and combine them into a single table.

Usage

```
mlmm(
  ...,
  data,
  by,
  contrast.rbind = NULL,
  effects = NULL,
  robust = NULL,
  df = TRUE,
```

```
  ci = TRUE,
  name.short = c(TRUE, TRUE),
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  trace = TRUE
)
```

Arguments

...	arguments passed to lmm.
data	[data.frame] dataset (in the long format) containing the observations.
by	[character] variable used to split the dataset. On each split a seperate linear mixed model is fit.
contrast.rbind	[character or numeric matrix] Contrast to be be applied to compare the groups. Argument passed to the argument effects of rbind.Wald_lmm. Right hand side can be specified via an attribute "rhs".
effects	[character or numeric matrix] Linear combinations of coefficients relative to which Wald test should be computed. Argument passed to anova.lmm. Right hand side can be specified via an attribute "rhs".
robust	[logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors. Argument passed to anova.lmm.
df	[logical] Should the degree of freedom be computed using a Satterthwaite approximation? Argument passed to anova.lmm.
ci	[logical] Should a confidence interval be output for each hypothesis? Argument passed to anova.lmm.
name.short	[logical vector of length 2] use short names for the output coefficients: omit the name of the by variable, omit the regression variable name when the same regression variable is used in all models.
transform.sigma, transform.k, transform.rho, transform.names	[character] transformation used on certain type of parameters.
trace	[interger, >0] Show the progress of the execution of the function.

Details

Grouping variable in argument repetition: when numeric, it will be converted into a factor variable, possibly adding a leading 0 to preserve the ordering. This transformation may cause inconsistency when combining results between different lmm object. This is why the grouping variable should preferably be of type character or factor.

See Also

confint.mlmm for a data.frame containing estimates with their uncertainty.
summary.mlmm for a summary of the model and estimates.
autoplot.Wald_lmm for a graphical display.

Examples

```
#### univariate regression ####
if(require(lava)){

set.seed(10)
d1 <- cbind(sim(lvm(Y~0.5*X1), 25), group = "A")
d2 <- cbind(sim(lvm(Y~0.1*X1), 100), group = "B")
d3 <- cbind(sim(lvm(Y~0.01*X1), 1000), group = "C")
d1$id <- 1:NROW(d1)
d2$id <- 1:NROW(d2)
d3$id <- 1:NROW(d3)

d <- rbind(d1,d2,d3)

e.mlmm <- mlmm(Y~X1, data = d, by = "group", effects = "X1=0")
summary(e.mlmm, method = "single-step")
summary(e.mlmm, method = "bonferroni")
summary(e.mlmm, method = "single-step2")
## summary(e.mlmm)
}

#### multivariate regression ####
set.seed(10)
dL <- sampleRem(250, n.times = 3, format = "long")

e.mlmm <- mlmm(Y~X1+X2+X6, repetition = ~visit|id, data = dL,
               by = "X4", structure = "CS")
summary(e.mlmm, method = "none")
confint(e.mlmm, method = "none")

e.mlmmX1 <- mlmm(Y~X1+X2+X6, repetition = ~visit|id, data = dL,
               by = "X4", effects = "X1=0", structure = "CS")
summary(e.mlmmX1)
summary(e.mlmmX1, method = "single-step")
```

`model.tables.lmm` *Statistical Inference for Linear Mixed Model*

Description

Export estimates, standard errors, degrees of freedom, confidence intervals (CIs) and p-values for the mean coefficients of a linear mixed model.

Usage

```
## S3 method for class 'lmm'
model.tables(x, columns, ...)
```

Arguments

x	a lmm object.
columns	[character vector] Columns to be output. Can be any of "estimate", "se", "statistic", "df", "null", "lower", "upper", "p.value".
...	arguments to be passed to the confint method. Should not contain the argument column.
method	[character] type of adjustment for multiple comparisons, one of "none", "bonferroni", ..., "fdr", "single-step", "single-step2". Alternatively, method for combining the estimates, one of "average", "pool.se", "pool.gls", "pool.rubin".

Details

This function simply calls confint with a specific value for the argument column.

---

ncgsL *Data From National Cooperative Gallstone Study (Long Format)*

Description

Data from the National Cooperative Gallstone Study (NCGS), a randomized study where the level of serum cholesterol was measured at baseline and after intake of high-dose chenondiol (750mg/day) or placebo. This dataset is in the long format (i.e. one line per measurement).

- group: treatment group (highdose or placebo).
- id: patient identifier.
- visit: visit index.
- cholest: cholesterol measurement.
- time: time after the start of the study at which the measurement has been done (in month). Treatment is given at 0+.

Usage

```
data(ncgsL)
```

References

Grundy SM, Lan SP, Lachin J. The effects of chenodiol on biliary lipids and their association with gallstone dissolution in the National Cooperative Gallstone Study (NCGS). J Clin Invest. 1984 Apr;73(4):1156-66. doi: 10.1172/JCI111301.

ncgsW *Data From National Cooperative Gallstone Study (Wide Format)*

Description

Data from the National Cooperative Gallstone Study (NCGS), a randomized study where the level of serum cholesterol was measured at baseline and after intake of high-dose chenondiol (750mg/day) or placebo. This dataset is in the wide format (i.e. one line per patient).

- `group`: treatment group (highdose or placebo).
- `id`: patient identifier.
- `cholest1`: cholesterol measurement at baseline (before treatment).
- `cholest2`: cholesterol measurement at 6 months (after treatment).
- `cholest3`: cholesterol measurement at 12 months (after treatment).
- `cholest4`: cholesterol measurement at 20 months (after treatment).
- `cholest5`: cholesterol measurement at 24 months (after treatment).

Usage

```
data(ncgsW)
```

References

Grundy SM, Lan SP, Lachin J. The effects of chenodiol on biliary lipids and their association with gallstone dissolution in the National Cooperative Gallstone Study (NCGS). J Clin Invest. 1984 Apr;73(4):1156-66. doi: 10.1172/JCI111301.

partialCor *Partial Correlation*

Description

Estimate the partial correlation based on equation 19 of Lloyd et al 2008 (`partialCor.lmm`) or explicitely modeling the correlation via a linear mixed model (`partialCor.list`, `partialCor.formula`). The first option is numerically more efficient and exact with a single observation per cluster. With multiple repetitions, what is being estimated with the first option may not be clear and the second option is therefore preferrable.

Usage

```
partialCor(object, ...)

## S3 method for class 'list'
partialCor(
  object,
  data,
  repetition = NULL,
  structure = NULL,
  by = NULL,
  effects = NULL,
  rhs = NULL,
  method = "none",
  df = NULL,
  transform.rho = NULL,
  name.short = c(TRUE, FALSE),
  ...
)

## S3 method for class 'formula'
partialCor(object, repetition, ...)

## S3 method for class 'lmm'
partialCor(object, level = 0.95, R2 = FALSE, se = TRUE, df = TRUE, ...)
```

Arguments

object	a formula with in the left hand side the variables for which the correlation should be computed and on the right hand side the adjustment set. Can also be a list of formula for outcome-specific adjustment set.
...	arguments passed to confint for partialCor.list and partialCor.formula. Not used for partialCor.lmm.
data	[data.frame] dataset containing the variables.
repetition	[formula] Specify the structure of the data: the time/repetition variable and the grouping variable, e.g. ~ timelid.
structure	[character] Specify the residual variance-covariance structure. Without repetitions, either "UN" or "CS". With repetitions, one of "UN", "PEARSON", "HLAG", "LAG", "HCS", "CS".
by	[character] variable used to stratified the correlation on.
effects	[character or matrix] type of contrast to be used for comparing the correlation parameters. One of "Dunnett", "Tukey", "Sequen", or a contrast matrix.
rhs	[numeric vector] right hand side for the comparison of correlation parameters.
method	[character] adjustment for multiple comparisons (e.g. "single-step").
df	[logical] Should a Student's t-distribution be used to model the distribution of the coefficient. Otherwise a normal distribution is used.

transform.rho	[character] scale on which perform statistical inference (e.g. "atanh")
name.short	[logical vector of length 2] use short names for the output coefficients (omit the name of the by variable, omit name of the correlation parameter)
level	[numeric,0-1] the confidence level of the confidence intervals.
R2	[logical] Should the R2 (coefficient of determination) be computed?
se	[logical] Should the uncertainty about the partial correlation be evaluated? Only relevant for partialCor.lmm.

Details

Fit a mixed model to estimate the partial correlation with the following variance-covariance pattern:

- **no repetition**: unstructure or compound symmetry structure for M observations, M being the number of variables on the left hand side (i.e. outcomes).
- **repetition**: structure for M*T observations where M being the number of variables (typically 2) and T the number of repetitions. Can be
  - "UN": unstructured (except the off-diagonal containing the correlation parameter which is constant).
  - "PEARSON": same as unstructured except it only uses a single variance parameter per variable, i.e. it assumes constant variance over repetitions.
  - "HLAG": toeplitz by block with variable and repetition specific variance.
  - "LAG": toeplitz by block, i.e. correlation depending on the gap between repetitions and specific to each variable. It assumes constant variance over repetitions.
  - "HCS": heteroschedastic compound symmetry by block, i.e. variable specific correlation constant over repetitions. A specific parameter is used for the off-diagonal crossing the variables at the same repetition (which is the marginal correlation parameter).
  - "CS": compound symmetry by block. It assumes constant variance and correlation over repetitions.

Value

A data.frame with the estimate partial correlation (rho), standard error, degree of freedom, confidence interval, and p-value (test of no correlation). When structure="CS" or structure="HCS" is used with repeated measurements, a second correlation coefficient (r) is output where the between subject variance has been removed (similar to Bland et al. 1995).

References

Bland J M, Altman D G. Statistics notes: Calculating correlation coefficients with repeated observations: Part 1—correlation within subjects BMJ 1995; 310 :446 doi:10.1136/bmj.310.6977.446 Edwards, L.J., Muller, K.E., Wolfinger, R.D., Qaqish, B.F. and Schabenberger, O. (2008), An R2 statistic for fixed effects in the linear mixed model. Statist. Med., 27: 6137-6157. https://doi.org/10.1002/sim.3429

Examples

```
#### no repetition ####

## example from ppcor::pcor
y.data <- data.frame(
  hl=c(7,15,19,15,21,22,57,15,20,18),
  disp=c(0.000,0.964,0.000,0.000,0.921,0.000,0.000,1.006,0.000,1.011),
  deg=c(9,2,3,4,1,3,1,3,6,1),
  BC=c(1.78e-02,1.05e-06,1.37e-05,7.18e-03,0.00e+00,0.00e+00,0.00e+00
,  4.48e-03,2.10e-06,0.00e+00)
)
## ppcor::pcor(y.data)

## partial correlation based on a formula
partialCor(c(hl,disp)~BC+deg, data = y.data)
partialCor(hl + disp~BC+deg, data = y.data)
## partial correlation based on a list
partialCor(list(hl~BC+deg,disp~BC+deg), data = y.data)
## via an existing model
e.lm <- lmm(hl~disp+BC+deg, data = y.data)
partialCor(e.lm)

## using a different set of covariates for outcome
partialCor(list(hl~BC+deg, disp~BC), data = y.data)

## statified correlation (using another dataset)
data(gastricbypassW, package = "LMMstar")
gastricbypassW$weight.bin <- gastricbypassW$weight1>=120
partialCor(glucagonAUC1+glucagonAUC2~1, data = gastricbypassW, by = "weight.bin")

## compared correlation between groups
partialCor(glucagonAUC1+glucagonAUC2~1, data = gastricbypassW, by = "weight.bin",
           effects = "Dunnett")

#### with repetitions ####
## Not run:
data(gastricbypassL, package = "LMMstar")
## via a mixed model
eUN.lmm <- lmm(weight ~ glucagonAUC+time, repetition =~time|id,
               data = gastricbypassL, structure = "UN")
partialCor(eUN.lmm)

## mean: variable and timepoint specific mean parameter (8)
## variance: variable and timepoint specific variance parameter (8)
## correlation: correlation parameter specific for each variable and time lag (10)
e.cor <- partialCor(weight+glucagonAUC~time, repetition =~time|id,
                    data = gastricbypassL, structure = "LAG")
e.cor
coef(attr(e.cor,"lmm"), effects = "correlation")
if(require(ggplot2)){
autoplot(e.cor)
```

```
}

## same except for the mean structure: variable specific mean parameter (2)
e.cor2 <- partialCor(weight+glucagonAUC~time, repetition =~time|id,
                     data = gastricbypassL, structure = "LAG")

## mean: variable and timepoint specific mean parameter (8)
## variance: variable specific variance parameter (2)
## correlation: correlation parameter specific for each variable and some time lag (4)
e.cor3 <- partialCor(weight+glucagonAUC~time, repetition =~time|id,
                      data = gastricbypassL, structure = "CS")
e.cor3
coef(attr(e.cor3,"lmm"), effects = "correlation")
if(require(ggplot2)){
autoplot(e.cor3)
}


## End(Not run)
```

`potassiumRepeatedL` *Data From The Potassium Intake Study (Long Format with intermediate measurements)*

Description

Data from the potassium intake study, a randomized placebo-controlled crossover study where the effect of potassium supplement (90 mmol/day) on the renin-angiostensin-aldosteron system (RAAS) was assessed. This dataset is in the long format (i.e. one line per measurement) and contains measurement over 6 timepoints for each time period.

- `id`: patient identifier.
- `sequence`: treatment group to which the patient has been randomized.
- `period`: time period.
- `treatment`: treatment during the time period.
- `time`: time within each period.
- `aldo`: ??

Usage

```
data(potassiumRepeatedL)
```

References

Dreier et al. Effect of increased potassium intake on the reninangiotensinaldosterone system and subcutaneous resistance arteries: a randomized crossover study, Nephrol Dial Transplant (2020) 110. doi: 10.1093/ndt/gfaa114

potassiumSingleL *Data From The Potassium Intake Study (Long Format)*

Description

Data from the potassium intake study, a randomized placebo-controlled crossover study where the effect of potassium supplement (90 mmol/day) on the renin-angiostensin-aldosteron system (RAAS) was assessed. This dataset is in the long format (i.e. one line per measurement).

- `id`: patient identifier.
- `sequence`: treatment group to which the patient has been randomized.
- `period`: time period.
- `treatment`: treatment during the time period.
- `auc`: area under the curve of ?? during the time period.
- `bsauc`: ??
- `aldo`: ??

Usage

```
data(potassiumSingleL)
```

References

Dreier et al. Effect of increased potassium intake on the reninangiotensinaldosterone system and subcutaneous resistance arteries: a randomized crossover study, Nephrol Dial Transplant (2020) 110. doi: 10.1093/ndt/gfaa114

potassiumSingleW *Data From The Potassium Intake Study (Wide Format)*

Description

Data from the potassium intake study, a randomized placebo-controlled crossover study where the effect of potassium supplement (90 mmol/day) on the renin-angiostensin-aldosteron system (RAAS) was assessed. This dataset is in the wide format (i.e. one line per patient).

- `id`: patient identifier.
- `sequence`: treatment group to which the patient has been randomized.
- `treatment1`: treatment during the first time period.
- `treatment2`: treatment during the second time period.
- `auc1`: area under the curve of ?? during the first time period.
- `auc2`: area under the curve of ?? during the second time period.
- `bsauc1`: ??
- `aldo1`: ??
- `aldo2`: ??

Usage

```
data(potassiumSingleW)
```

References

Dreier et al. Effect of increased potassium intake on the reninangiotensinaldosterone system and subcutaneous resistance arteries: a randomized crossover study, Nephrol Dial Transplant (2020) 110. doi: 10.1093/ndt/gfaa114

---

predict.lmm *Predicted Mean Value With Uncertainty For Linear Mixed Model*

Description

Predicted mean value conditional on covariates or on covariates and other outcome values.

Usage

```
## S3 method for class 'lmm'
predict(
  object,
  newdata,
  p = NULL,
  se = "estimation",
  df = !is.null(object$df),
  type = "static",
  level = 0.95,
  keep.newdata = FALSE,
  se.fit,
  ...
)
```

Arguments

object	a lmm object.
newdata	[data.frame] the covariate values for each cluster.
p	[numeric vector] value of the model coefficients at which to evaluate the predictions. Only relevant if differs from the fitted values.
se	[character] Type of uncertainty to be accounted for: estimation of the regression parameters ("estimation"), residual variance ("residual"), or both ("total"). Can also be NULL to not compute standard error, p-values, and confidence intervals.
df	[logical] Should a Student's t-distribution be used to model the distribution of the predicted mean. Otherwise a normal distribution is used.

type	[character] Should prediction be made conditional on the covariates only ("static") or also on outcome values at other timepoints ("dynamic"). Can also output the model term ("terms", similarly to stats::predict.lm.
level	[numeric,0-1] the confidence level of the confidence intervals.
keep.newdata	[logical] Should the argument newdata be output along side the predicted values?
se.fit	For internal use. When not missing mimic the output of predict.se. Overwrite argument se.
...	Not used. For compatibility with the generic method.

Details

Static prediction are made using the linear predictor $X\beta$ while dynamic prediction uses the conditional normal distribution of the missing outcome given the observed outcomes. So if outcome 1 is observed but not 2, prediction for outcome 2 is obtain by $X_2\beta + \sigma_{21}\sigma_{22}^{-1}(Y_1 - X_1\beta)$. In that case, the uncertainty is computed as the sum of the conditional variance $\sigma_{22} - \sigma_{21}\sigma_{22}^{-1}\sigma_{12}$ plus the uncertainty about the estimated conditional mean (obtained via delta method using numerical derivatives).

The model terms are computing by centering the design matrix around the mean value of the covariates used to fit the model. Then the centered design matrix is multiplied by the mean coefficients and columns assigned to the same variable (e.g. three level factor variable) are summed together.

Value

A data.frame with 5 columns:

- estimate: predicted mean.
- se: uncertainty about the predicted mean.
- df: degree of freedom
- lower: lower bound of the confidence interval of the predicted mean
- upper: upper bound of the confidence interval of the predicted mean

except when the argument se.fit is specified (see predict.lm for the output format).

Examples

```
## simulate data in the long format
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")

## fit Linear Mixed Model
eUN.lmm <- lmm(Y ~ visit + X1 + X2 + X5,
               repetition = ~visit|id, structure = "UN", data = dL)

## prediction
newd <- data.frame(X1 = 1, X2 = 2, X5 = 3, visit = factor(1:3, levels = 1:3))
predict(eUN.lmm, newdata = newd)
predict(eUN.lmm, newdata = newd, keep.newdata = TRUE)
```

```
predict(eUN.lmm, newdata = newd, keep.newdata = TRUE, se = "total")

## dynamic prediction
newd.d1 <- cbind(newd, Y = c(NA,NA,NA))
predict(eUN.lmm, newdata = newd.d1, keep.newdata = TRUE, type = "dynamic")
newd.d2 <- cbind(newd, Y = c(6.61,NA,NA))
predict(eUN.lmm, newdata = newd.d2, keep.newdata = TRUE, type = "dynamic")
newd.d3 <- cbind(newd, Y = c(1,NA,NA))
predict(eUN.lmm, newdata = newd.d3, keep.newdata = TRUE, type = "dynamic")
newd.d4 <- cbind(newd, Y = c(1,1,NA))
predict(eUN.lmm, newdata = newd.d4, keep.newdata = TRUE, type = "dynamic")
```

`profile.lmm` *Evaluate Contour of the Log-Likelihood*

Description

Display the (restricted) log-likelihood around Maximum Likelihood Estimate (MLE) under specific constrains.

Usage

```
## S3 method for class 'lmm'
profile(
  fitted,
  effects = NULL,
  profile.likelihood = FALSE,
  maxpts = NULL,
  conf.level = 0.95,
  trace = FALSE,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  ...
)
```

Arguments

`fitted`	a `lmm` object.
`effects`	[character vector] name of the parameters who will be constrained. Alternatively can be the type of parameters, e.g. `"mean"`, `"variance"`, `"correlation"`, or `"all"`.
`profile.likelihood`	[logical] should profile likelihood be performed? Otherwise varying one parameter at a time around the MLE while keeping the other constant).
`maxpts`	[integer, >0] number of points use to discretize the likelihood, `maxpts` points smaller than the MLE and `maxpts` points higher than the MLE.

conf.level	[numeric, 0-1] the confidence level of the confidence intervals used to decide about the range of values for each parameter.
trace	[logical] Show the progress of the execution of the function.
transform.sigma	[character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.
transform.k	[character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details.
transform.rho	[character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.
transform.names	[logical] Should the name of the coefficients be updated to reflect the transformation that has been used?
...	Not used. For compatibility with the generic method.

Details

Each parameter defined by the argument effets is treated separately:

- the confidence interval of a parameter is discretized with maxpt points,
- this parameter is set to a discretization value.
- the other parameters are either set to the (unconstrained) MLE (profile.likelihood=FALSE) or to constrained MLE (profile.likelihood=TRUE). The latter case is much more computer intensive as it implies re-running the estimation procedure.
- the (restricted) log-likelihood is evaluated.

Value

A data.frame object containing the log-likelihood for various parameter values.

Examples

```
data(gastricbypassW, package = "LMMstar")
e.lmm <- lmm(weight2 ~ weight1 + glucagonAUC1,
             data = gastricbypassW, control = list(optimizer = "FS"))

## profile logLiklihood
## Not run:
e.pro <- profile(e.lmm, effects = "all", maxpts = 10, profile.likelihood = TRUE)
head(e.pro)
plot(e.pro)

## End(Not run)

## along a single parameter axis
e.sliceNone <- profile(e.lmm, effects = "all", maxpts = 10, transform.sigma = "none")
plot(e.sliceNone)
```

```
e.sliceLog <- profile(e.lmm, effects = "all", maxpts = 10, transform.sigma = "log")
plot(e.sliceLog)
```

`proportion` *Proportion of Significant Findings*

Description

Evaluate the proportion of test above the statistical significance level

Usage

```
proportion(object, n.sample, trace, ...)
```

Arguments

`object`	`Wald_lmm` object
`n.sample`	[numeric,>=0] number of bootstrap sample used to assess the uncertainty. If 0, then only the point estimate is computed.
`trace`	[logical] shoudl the execution of the boostrap be trace.
`...`	additional arguments passed to codeconfint.Wald_lmm

Value

a data.frame with the estimated proportion (estimate column), standard error and confidence interval (when boostrap is used).

`rbind.Wald_lmm` *Linear Hypothesis Testing Across Linear Mixed Models*

Description

Linear hypothesis testing accross linear mixed model.

Usage

```
## S3 method for class 'Wald_lmm'
rbind(model, ..., effects = NULL, rhs = NULL, name = NULL, sep = ": ")
```

Arguments

model	a Wald_lmm object (output of anova applied to a lmm object)
...	possibly other Wald_lmm objects
effects	[character or numeric matrix] how to combine the left-hand side of the hypotheses. By default identity matrix but can also be "Dunnett", "Tukey", or "Sequen" (see function multcomp::contrMat from the multcomp package).
rhs	[numeric vector] the right hand side of the hypothesis. Should have the same length as the number of row of argument effects.
name	[character vector or NULL] character used to identify each model in the output. By default, use the name of the outcome of the model.
sep	[character] character used to separate the outcome and the covariate when naming the tests.

Examples

```
## simulate data
set.seed(10)
dL <- sampleRem(1e2, n.times = 3, format = "long")

## estimate mixed models
e.lmm1 <- lmm(Y ~ X1+X2+X3, repetition = ~visit|id, data = dL,
              structure = "CS", df = FALSE)
e.lmm2 <- lmm(Y ~ X1+X8+X9, repetition = ~visit|id, data = dL,
              structure = "CS", df = FALSE)

## select null hypotheses
AAA <- anova(e.lmm1, ci = TRUE, effect = c("X1|X2,X3"="X1=0","X2|X1,X3"="X2=0"))
BBB <- anova(e.lmm2, ci = TRUE, effect = c("X1|X8,X9"="X1=0"))

## combine
ZZZ <- rbind(AAA,BBB)
summary(ZZZ)
```

---

recover_data.lmm *Link to emmeans package*

Description

Link to emmeans package. Not meant for direct use.

Usage

```
## S3 method for class 'lmm'
recover_data(object, ...)

## S3 method for class 'lmm'
emm_basis(object, trms, xlev, grid, ...)
```

Arguments

object	a lmm object.
...	Not used. For compatibility with the generic method.
trms	see emmeans::emm_basis documentation
xlev	see emmeans::emm_basis documentation
grid	see emmeans::emm_basis documentation

Value

dataset or list used by the emmeans package.

---

remove *Remove Columns from Output*

Description

Auxiliary function that can be used when specifying the argument columns (e.g. calling confint.lmm) to remove columns.

Usage

```
remove(...)
```

Arguments

...	[character vector] name of the columns to be removed to the default output.

Value

A character vector

Examples

```
set.seed(10)
dW <- sampleRem(25, n.times = 1, format = "long")
e.lmm <- lmm(Y~X1, data = dW)

confint(e.lmm, columns = remove("estimate"))
```

resample *Inference via Resampling for Linear Mixed Model*

Description

Non-parametric bootstrap or permutation test for Linear Mixed Models.

Usage

```
resample(
  object,
  type,
  effects,
  n.sample = 1000,
  studentized = TRUE,
  level = 0.95,
  trace = TRUE,
  seed = NULL,
  cpus = 1
)
```

Arguments

object	a lmm object.
type	[character] should permutation test ("perm-var" or "perm-res") or non-parametric bootstrap ("boot") be used?
effects	[character vector] the variable(s) to be permuted or the effect(s) to be tested via non-parametric bootstrap.
n.sample	[integer] the number of samples used.
studentized	[logical] should a studentized boostrap or permutation test be used?
level	[numeric,0-1] the confidence level of the confidence intervals.
trace	[logical] should the execution of the resampling be traced?
seed	[integer] Random number generator (RNG) state used when starting resampling.
cpus	[integer] number of child-processes for parallel evaluation.

Details

All approach are carried at the cluster level:

- Bootstrap: sampling with replacement clusters. If a cluster is picked twice then different cluster id is used for each pick.
- Permutation: permuting covariate values between clusters (this only lead to the null hypothesis when the covariate values are constant within clusters) or assigning new outcome values by, under the null, permuting the normalized residuals, rescaling to residuals, and adding back the permuted fixed effect (any mean effect under H1 would be 0 because of the permutation if

the variance-covariance structure is correct). The later procedure originates from Oliver et al (2012).

The studentized bootstrap keep the original estimate and standard error but uses the samples to evaluates the quantiles of the distribution used to form the confidence intervals. The studentized permutation test approximate the distribution of the test statistic under the null (instead of the distribution of the estimate under the null.).

P-values for the bootstrap are computed by centering the bootstrap distribution of the estimate or test statistic around 0 and evaluating the frequency at which it takes values more extreme than the observed estimate or test statistics.

References

Oliver E. Lee and Thomas M. Braun (2012), Permutation Tests for Random Effects in Linear Mixed Models. Biometrics, Journal 68(2).

Examples

```
## Not run:

#### univariate linear regression ####
data(gastricbypassW, package = "LMMstar")
## rescale to ease optimization
gastricbypassW$weight1 <- scale(gastricbypassW$weight1)
gastricbypassW$weight2 <- scale(gastricbypassW$weight2)
gastricbypassW$glucagonAUC1 <- scale(gastricbypassW$glucagonAUC1)

e.lm <- lmm(weight2~weight1+glucagonAUC1, data = gastricbypassW)
model.tables(e.lm)

## non-parametric bootstrap
resample(e.lm, type = "boot", effects = c("weight1","glucagonAUC1"))
## permutation test
resample(e.lm, type = "perm-var", effects = "weight1")
resample(e.lm, type = "perm-var", effects = "glucagonAUC1")
## using multiple cores
resample(e.lm, type = "boot", effects = c("weight1","glucagonAUC1"), cpus = 4)

#### random intercept model ####
data(gastricbypassL, package = "LMMstar")
gastricbypassL$weight <- scale(gastricbypassL$weight)
gastricbypassL$glucagonAUC <- scale(gastricbypassL$glucagonAUC)
gastricbypassL$gender <- as.numeric(gastricbypassL$id) %% 2
gastricbypassLR <- na.omit(gastricbypassL)

eCS.lmm <- lmm(weight~glucagonAUC+gender, data = gastricbypassLR,
               repetition = ~visit|id, structure = "CS")
model.tables(eCS.lmm)

## non-parametric bootstrap
resample(eCS.lmm, type = "boot", effects = c("glucagonAUC","gender"))
## permutation test
```

```
resample(eCS.lmm, type = "perm-var", effects = "gender")
resample(eCS.lmm, type = "perm-res", effects = "glucagonAUC")

## End(Not run)
```

`residuals.lmm` *Extract The Residuals From a Linear Mixed Model*

Description

Extract or compute the residuals of a linear mixed model.

Usage

```
## S3 method for class 'lmm'
residuals(
  object,
  type = "response",
  var = NULL,
  data = NULL,
  p = NULL,
  format = "long",
  keep.data = FALSE,
  ...
)
```

Arguments

`object`	a `lmm` object.
`type`	[character] type of residual to output: raw residuals (`"response"`), Pearson residuals (`"pearson"`), normalized residuals (`"normalized"`, scaled residual `"scaled"`), or partial residuals (`"partial"` or `"partial-center"`). Can also be `"all"` to output all except partial residuals. See detail section.
`var`	[character vector] name of the variable relative to which the partial residuals should be computed.
`data`	[data.frame] dataset relative to which the residuals should be computed. Only relevant if differs from the dataset used to fit the model.
`p`	[numeric vector] value of the model coefficients at which to evaluate the residuals. Only relevant if differs from the fitted values.
`format`	[character] Should the residuals be output relative as a vector (`"long"`), or as a matrix with in row the clusters and in columns the outcomes (`"wide"`).
`keep.data`	[logical] Should the argument `data` be output along side the residuals? Only possible in the long format.
`...`	Not used. For compatibility with the generic method.

Details

The argument type defines how the residuals are computed:

- "fitted": fitted value $X_{ij}\hat{\beta}$.
- "response": raw residual, i.e. observed outcome minus fitted value $\varepsilon_{ij} = Y_{ij} - X_{ij}\hat{\beta}$.
- "pearson": each raw residual is divided by its modeled standard deviation $\varepsilon_{ij} = \frac{Y_{ij}-X_{ij}\hat{\beta}}{\sqrt{\hat{\omega}_{ij}}}$.
- "studentized": same as "pearson" but excluding the contribution of the cluster in the modeled standard deviation $\varepsilon_{ij} = \frac{Y_{ij}-X_{ij}\hat{\beta}}{\sqrt{\hat{\omega}_{ij}-\hat{q}_{ij}}}$.
- "normalized": raw residuals are multiplied, within clusters, by the inverse of the (upper) Cholesky factor of the modeled residual variance covariance matrix $\varepsilon_{ij} = (Y_i - X_i\hat{\beta})\hat{C}^{-1}$.
- "normalized2": same as "normalized" but excluding the contribution of the cluster in the modeled residual variance covariance matrix $\varepsilon_{ij} = (Y_i - X_i\hat{\beta})\hat{D}_i^{-1}$.
- "scaled": scaled residuals (see PROC MIXED in SAS). Numerically identical to "normalized" but computed by sequentially scaling and centering the residuals, to make them conditionally independent of previous residuals from the same cluster at previous repetitions.
- "partial": partial residuals ($\gamma E + \hat{\varepsilon}$). A reference level can be also be specified via the attribute "reference" to change the absolute level of the partial residuals. "partial-center": partial residuals with centered covariates ($\gamma E + \hat{\varepsilon}$ where E has been centered, i.e., has 0-mean)

where

- $X = (E, W)$ the design matrix. For partial residuals, it is split according to the variable(s) in argument var (E) and the rest (W).
- Y the outcome
- $\hat{\beta} = (\hat{\gamma}, \hat{\delta})$ the estimated mean coefficients relative to $X = (E, W)$
- $\hat{\Omega}$ the modeled variance-covariance of the residuals and $\hat{\omega}$ its diagonal elements
- $\hat{C}$ the upper Cholesky factor of $\hat{\Omega}$, i.e. upper triangular matrix satisfying $\hat{C}^t\hat{C} = \hat{\Omega}$
- $\hat{Q}_i = X_i(X^t\hat{\Omega}X)^{-1}X_i^t$ a cluster specific correction factor, approximating the contribution of cluster i to $\hat{\Omega}$. Its diagonal elements are denoted $\hat{q}_i$.
- $\hat{D}_i$ the upper Cholesky factor of $\hat{\Omega} - \hat{Q}_i$

Value

When argument format is "long" and type.oobject is "lmm", a vector containing the value of the residual realtive to each observation. It is a matrix if the argument type contains several values. When argument format is "wide" and type.oobject is "lmm", a data.frame with the value of the residual relative to each cluster (in rows) at each timepoint (in columns).

Examples

```
#### simulate data in the long format ####
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")
```

```
#### Linear Model ####
e.lm <- lmm(Y ~ visit + X1 + X2 + X6, data = dL)

## partial residuals
residuals(e.lm, type = "partial", var = "X6")
## residuals(e.lm) + dL$X6 * coef(e.lm)["X6"]
e.reslm <- residuals(e.lm, type = "partial", var = "X6", keep.data = TRUE)
plot(e.reslm)

## partial residuals with specific reference
type <- "partial"
attr(type,"reference") <- data.frame(visit=factor(2,1:3),X2=0,X6=3)
residuals(e.lm, type = type, var = "X1")
## residuals(e.lm) + dL$X1 * coef(e.lm)["X1"] + coef(e.lm)["visit2"]

## partial residuals with centered covariates
residuals(e.lm, type = "partial-center", var = "X1")
## residuals(e.lm) + (dL$X1-mean(dL$X1)) * coef(e.lm)["X1"]

#### Linear Mixed Model ####
eUN.lmm <- lmm(Y ~ visit + X1 + X2 + X5 + X6,
               repetition = ~visit|id, structure = "UN", data = dL)

## residuals
e.reslmm <- residuals(eUN.lmm, type = "normalized", keep.data = TRUE)
plot(e.reslmm)
plot(e.reslmm, type = "correlation")
plot(e.reslmm, type = "scatterplot", labeller = ggplot2::label_both)
residuals(eUN.lmm, format = "wide", type = "normalized")

## residuals and predicted values
residuals(eUN.lmm, type = "all")
residuals(eUN.lmm, type = "all", keep.data = TRUE)

## partial residuals
residuals(eUN.lmm, type = "partial", var = c("(Intercept)","X6"))
residuals(eUN.lmm, type = "partial", var = c("X6"))
```

`sampleRem` *Sample Longuitudinal Data*

Description

Sample longuitudinal data with covariates

Usage

```
sampleRem(
  n,
```

```
  n.times,
  mu = 1:n.times,
  sigma = rep(1, n.times),
  lambda = rep(1, n.times),
  beta = c(2, 1, 0, 0, 0, 1, 1, 0, 0, 0),
  gamma = matrix(0, nrow = n.times, ncol = 10),
  format = "wide",
  latent = FALSE
)
```

Arguments

n	[integer,>0] sample size
n.times	[integer,>0] number of visits (i.e. measurements per individual).
mu	[numeric vector] expected measurement value at each visit (when all covariates are fixed to 0). Must have length n.times.
sigma	[numeric vector,>0] standard error of the measurements at each visit (when all covariates are fixed to 0). Must have length n.times.
lambda	[numeric vector] covariance between the measurement at each visit and the individual latent variable. Must have length n.times.
beta	[numeric vector of length 10] regression coefficient between the covariates and the latent variable.
gamma	[numeric matrix with n.times rows and 10 columns] regression coefficient specific to each timepoint (i.e. interaction with time).
format	[character] Return the data in the wide format ("wide") or long format ("long")
latent	[logical] Should the latent variable be output?

Details

The generative model is a latent variable model where each outcome (Y_j) load on the latent variable (η) with a coefficient lambda:

$$Y_j = \mu_j + \lambda_j * \eta + \sigma_j \epsilon_j$$

The latent variable is related to the covariates ($X_1, \ldots, X_(10)$):

$$\eta = \alpha + \beta_1 X_1 + ... + \beta_{10} X_{10} + \xi$$

ϵ_j and ξ are independent random variables with standard normal distribution.

Value

a data.frame

Examples

```
set.seed(10)
dW <- sampleRem(100, n.times = 3, format = "wide")
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")
```

scatterplot *Scatterplot for Continuous Variables*

Description

Produce a matrix of plot for continuous variables: scatterplots, histograms, correlation and missing values. Inspired from the `ggpairs` function of the R package GGally.

Usage

```
scatterplot(
  data,
  formula,
  columns,
  format = NULL,
  group = NULL,
  transform = NULL,
  facet = "grid",
  alpha.point = 1,
  type.diag = "boxplot",
  bins = NULL,
  position.bar = "identity",
  linewidth.density = NULL,
  alpha.area = NULL,
  method.cor = "pearson",
  size.cor = NULL,
  digits = c(3, 2),
  display.NA = NULL,
  color = NULL,
  xlim = NULL,
  ylim = NULL,
  size.axis = NULL,
  size.legend = NULL,
  size.facet = NULL
)
```

Arguments

`data`	[data.frame] dataset containing the variables to be displayed.
`formula`	[formula] formula indicating the variables to be used (outcome~timelid). Long format only.
`columns`	[character vector] Columns whose numerical values are to be displayed. Wide format only.
`format`	[character] Is the dataset in the long (`"long"`) or wide (`"wide"`) format?
`group`	[character] optional group variable used to color the points, stratify the histogram/density and correlation.

transform	[character or function] optional transformation to be applied on the outcome.
facet	[character] whether to use ggplot:::facet_grid ("grid") or ggh4x::facet_grid2 ("grid2").
alpha.point	[numeric] the transparency level used to display the points in the scatterplot.
type.diag	[character] type of graphical display on the diagonal: "boxplot", "histogram", or "density".
bins	[character or numeric vector] algorithm or values or number of values used to create the histogram cells. When using facet="grid2" and density=TRUE a character of length two indicating the bandwith and the kernel to be used. See ggplot2::stat_density.
position.bar	[character] passed to geom_histogram (argument position). Only relevant when having multiple groups and using ggh4x::facet_grid2.
linewidth.density	[numeric,>0] width of the lines on the density plot.
alpha.area	[numeric, 0-1] the transparency level used to display the area under the density curve or histogram.
method.cor	[character] estimator of the correlation. Argument passed to stats::cor. When NA, the correlation is not displayed.
size.cor	[numeric,>0] size of the font used to display the correlation or information about missing values.
digits	[numeric of length 2] number of digits used to display the correlation or round the percentage of missing values.
display.NA	[0:2 or "only"] Should the number of missing values be displayed. When taking value 2, will also display the percentage of missing values.
color	[character vector] color used to display the values for each group.
xlim	[numeric,>0 or "common"] range of the x-axis.
ylim	[numeric,>0 or "common"] range of the y-axis.
size.axis	[numeric,>0] size of the font used to display the tick labels.
size.legend	[numeric,>0] size of the font used to display the legend. Can have a second element to control the size of the legend key.
size.facet	[numeric,>0] size of the font used to display the facets (row and column names).

Details

In the long format, the outcome variable contains the numerical values to be displayed. The time variable will be used to spit outcome and display each split separately or jointly with one other split. The identifier links the outcome values across time.

Value

a list of ggplot objects (facet="grid") or a ggplot object (facet="grid2")

Examples

```
data(gastricbypassL, package = "LMMstar")
gastricbypassL$group <- as.numeric(gastricbypassL$id) %% 3
data(gastricbypassW, package = "LMMstar")

## single group (wide or long format)
scatterplot(gastricbypassL, formula = weight~time|id)
scatterplot(gastricbypassW, columns = paste0("weight",1:4))

## Not run:
## use histogram instead of boxplot
scatterplot(gastricbypassL, formula = weight~time|id, type.diag = "hist")
scatterplot(gastricbypassL, formula = weight~time|id, type.diag = "hist", bins = 15)

## same scale
scatterplot(gastricbypassL, formula = weight~time|id,
            xlim = "common", ylim = "common")

## transform outcome
scatterplot(gastricbypassL, formula = weight~time|id, transform = "log")

## handling missing values
scatterplot(gastricbypassL, formula = glucagonAUC~time|id)

## coloring per group
scatterplot(gastricbypassL, formula = weight~time|id, group = "group")

## only display percentage of NAs
scatterplot(gastricbypassL, formula = glucagonAUC~time|id,
            display.NA = "only", group = "group")
scatterplot(gastricbypassL, formula = glucagonAUC~time|id,
            display.NA = "only", group = "group", size.legend = c(15,2))

## End(Not run)
```

---

schoolL *Simulated Data with 3-level struture (Long Format)*

Description

Simulated data a nested structure: Student/Class/School and one outcome.

- school:
- class:
- student:
- outcome:

Usage

```
data(schoolL)
```

score.lmm *Extract The Score From a Linear Mixed Model*

Description

Extract or compute the first derivative of the log-likelihood of a linear mixed model.

Usage

```
## S3 method for class 'lmm'
score(
  x,
  effects = "mean",
  data = NULL,
  p = NULL,
  indiv = FALSE,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  ...
)
```

Arguments

x	a lmm object.
effects	[character] Should the score relative to all coefficients be output ("all"), or only coefficients relative to the mean ("mean" or "fixed"), or only coefficients relative to the variance and correlation structure ("variance" or "correlation").
data	[data.frame] dataset relative to which the score should be computed. Only relevant if differs from the dataset used to fit the model.
p	[numeric vector] value of the model coefficients at which to evaluate the score. Only relevant if differs from the fitted values.
indiv	[logical] Should the contribution of each cluster to the score be output? Otherwise output the sum of all clusters of the derivatives.
transform.sigma	[character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.
transform.k	[character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details.
transform.rho	[character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.
transform.names	[logical] Should the name of the coefficients be updated to reflect the transformation that has been used?

...	Not used. For compatibility with the generic method.

Details

For details about the arguments **transform.sigma**, **transform.k**, **transform.rho**, see the documentation of the coef.lmm function.

Value

When argument indiv is FALSE, a vector with the value of the score relative to each coefficient. When argument indiv is TRUE, a matrix with the value of the score relative to each coefficient (in columns) and each cluster (in rows).

---

sigma.lmm *Extract The Residuals Variance-Covariance Matrix From a Linear Mixed Model*

Description

Extract the unique set of residuals variance-covariance matrices or the one relative to specific clusters.

Usage

```
## S3 method for class 'lmm'
sigma(
  object,
  cluster = NULL,
  p = NULL,
  chol = FALSE,
  inverse = FALSE,
  simplifies = TRUE,
  ...
)
```

Arguments

object	a lmm object.
cluster	[character, data.frame, NULL] identifier of the cluster(s) for which to extract the residual variance-covariance matrix. For new clusters, a dataset containing the information (cluster, time, strata, ...) to be used to generate the residual variance-covariance matrices. When NULL, will output complete data covariance patterns.
p	[numeric vector] value of the model coefficients at which to evaluate the residual variance-covariance matrix. Only relevant if differs from the fitted values.
chol	[logical] Output the cholesky factorization of the variance-covariance matrix.
inverse	[logical] Output the matrix inverse of the variance-covariance matrix.

simplifies	[logical] When there is only one variance-covariance matrix, output a matrix instead of a list of matrices.
...	Not used. For compatibility with the generic method.

Value

A list where each element contains a residual variance-covariance matrix. Can also be directly a matrix when argument is simplifies=TRUE and there is a single residual variance-covariance matrix.

Examples

```
## simulate data in the long format
set.seed(10)
dL <- sampleRem(100, n.times = 3, format = "long")
dL$id.fac <- paste0("id",dL$id)

## fit Linear Mixed Model
eUN.lmm <- lmm(Y ~ X1 + X2 + X5, repetition = ~visit|id.fac,
               structure = "UN", data = dL, df = FALSE)

## extract residuals variance covariance matrix
sigma(eUN.lmm) ## unique patterns
sigma(eUN.lmm, cluster = c("id1","id5")) ## existing clusters
sigma(eUN.lmm, cluster = dL[1:7,,drop=FALSE]) ## new clusters
```

---

summarize *Compute summary statistics*

Description

Compute summary statistics for multiple variables and/or multiple groups and save them in a data frame.

Usage

```
summarize(
  formula,
  data,
  na.action = stats::na.pass,
  na.rm = FALSE,
  level = 0.95,
 columns = c("observed", "missing", "mean", "sd", "min", "q1", "median", "q3", "max",
    "correlation"),
  FUN = NULL,
  which = NULL,
  skip.reference = TRUE,
  digits = NULL,
  ...
)
```

Arguments

formula	[formula] on the left hand side the outcome(s) and on the right hand side the grouping variables. E.g. Y1+Y2 ~ Gender + Gene will compute for each gender and gene the summary statistics for Y1 and for Y2. Passed to the stats::aggregate function.
data	[data.frame] dataset containing the observations.
na.action	[function] a function which indicates what should happen when the data contain 'NA' values. Passed to the stats::aggregate function.
na.rm	[logical] Should the summary statistics be computed by omitting the missing values.
level	[numeric,0-1] the confidence level of the confidence intervals.
columns	[character vector] name of the summary statistics to kept in the output. Can be any of, or a combination of: • "observed": number of observations with a measurement. • "missing": number of missing observations. When specifying a grouping variable, it will also attempt to count missing rows in the dataset. • "pc.missing": percentage missing observations. • "mean", "mean.lower" "mean.upper": mean with its confidence interval. • "median", "median.lower" "median.upper": median with its confidence interval. • "sd": standard deviation. • "q1", "q3", "IQR": 1st and 3rd quartile, interquartile range. • "min", "max": minimum and maximum observation. • "predict.lower", "predict.upper": prediction interval for normally distributed outcome. • "correlation": correlation matrix between the outcomes (when feasible, see detail section).
FUN	[function] user-defined function for computing summary statistics. It should take a vector as an argument and output a named single value or a named vector.
which	deprecated, use the argument columns instead.
skip.reference	[logical] should the summary statistics for the reference level of categorical variables be omitted?
digits	[integer, >=0] the minimum number of significant digits to be used to display the results. Passed to print.data.frame
...	additional arguments passed to argument FUN.

Details

This function is essentially an interface to the stats::aggregate function.

Confidence intervals (CI) and prediction intervals (PI) for the mean are computed via stats::t.test. Confidence intervals (CI) for the median are computed via asht::medianTest.

Correlation can be assessed when a grouping and ordering variable are given in the formula interface , e.g. Y ~ timelid.

Value

A data frame containing summary statistics (in columns) for each outcome and value of the grouping variables (rows). It has an attribute "correlation" when it was possible to compute the correlation matrix for each outcome with respect to the grouping variable.

Examples

```
## simulate data in the wide format
set.seed(10)
d <- sampleRem(1e2, n.times = 3)
d$treat <-  sample(LETTERS[1:3], NROW(d), replace=TRUE, prob=c(0.3, 0.3, 0.4) )

## add a missing value
d2 <- d
d2[1,"Y2"] <- NA

## run summarize
summarize(Y1 ~ 1, data = d)
summarize(Y1 ~ 1, data = d, FUN = quantile, p = c(0.25,0.75))
summarize(Y1+Y2 ~ X1, data = d)
summarize(treat ~ 1, skip.reference = FALSE, data = d)

summarize(Y1 ~ X1, data = d2)
summarize(Y1+Y2 ~ X1, data = d2, na.rm = TRUE)

## long format
dL <- reshape(d, idvar = "id", direction = "long",
              v.names = "Y", varying = c("Y1","Y2","Y3"))
summarize(Y ~ time + X1, data = dL)

## compute correlations (single time variable)
e.S <- summarize(Y ~ time + X1 | id, data = dL, na.rm = TRUE)
e.S
attr(e.S, "correlation")

## compute correlations (composite time variable)
dL$time2 <- dL$time == 2
dL$time3 <- dL$time == 3
e.S <- summarize(Y ~ time2 + time3 + X1 | id, data = dL, na.rm = TRUE)
e.S
attr(e.S, "correlation")
```

---

summarizeNA *Summarize missing data patterns*

Description

Summarize missing data patterns.

Usage

```
summarizeNA(
  data,
  repetition = NULL,
  sep = "",
  newnames = c("variable", "frequency", "missing.pattern", "n.missing"),
  keep.data = TRUE
)
```

Arguments

data	[data.frame] dataset containing the observations.
repetition	[formula] Specify the structure of the data when in the long format: the time/repetition variable and the grouping variable, e.g. ~ timelid. When specified the missing data pattern is specific to each variable not present in the formula.
sep	[character] character used to separate the missing data indicator (0/1) when naming the missing data patterns.
newnames	[character vector of length 4] additional column containing the variable name (only when argument repetition is used), the frequency of the missing data pattern in the dataset, the name of the missing data pattern in the dataset, and the number of missing data per pattern.
keep.data	[logical] should the indicator of missing data per variable in the original dataset per pattern be output.

Value

a data frame

See Also

autoplot.summarizeNA for a graphical display.

Examples

```
data(gastricbypassW, package = "LMMstar")
summarizeNA(gastricbypassW)
summarizeNA(gastricbypassW, keep.data = FALSE)

data(gastricbypassL, package = "LMMstar")
summarizeNA(gastricbypassL, repetition = ~time|id)

data(calciumL, package = "LMMstar")
mp <- summarizeNA(calciumL, repetition = ~visit|girl)
plot(mp, variable = "bmd")
summarizeNA(calciumL[,c("visit","girl","bmd")], repetition = ~visit|girl)

data(vasscoresW, package = "LMMstar")
summarizeNA(vasscoresW)
```

summary.lmm *Summary Output for a Linear Mixed Model*

Description

Summary output for a linear mixed model fitted with lmm.

Usage

```
## S3 method for class 'lmm'
summary(
  object,
  level = 0.95,
  robust = FALSE,
  print = TRUE,
  columns = NULL,
  digits = 3,
  digits.df = 1,
  digits.p.value = 3,
  hide.data = FALSE,
  hide.fit = FALSE,
  hide.cor = is.null(object$formula$cor),
  type.cor = NULL,
  hide.var = TRUE,
  hide.sd = FALSE,
  hide.mean = FALSE,
  ...
)
```

Arguments

object	[lmm] output of the lmm function.
level	[numeric,0-1] confidence level for the confidence intervals.
robust	[logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors.
print	[logical] should the output be printed in the console.
columns	[character vector] Columns to be output for the fixed effects. Can be any of "estimate", "se", "statistic", "df", "null", "lower", "upper", "p.value".
digits	[interger, >0] number of digits used to display estimates.
digits.df	[interger, >0] number of digits used to display degrees of freedom.
digits.p.value	[interger, >0] number of digits used to display p-values.
hide.data	[logical] should information about the dataset not be printed.
hide.fit	[logical] should information about the model fit not be printed.
hide.cor	[logical] should information about the correlation structure not be printed.

type.cor	[character] should the correlation matrix be display ("matrix") or the parameter values ("param").
hide.var	[logical] should information about the variance not be printed.
hide.sd	[logical] should information about the standard deviation not be printed.
hide.mean	[logical] should information about the mean structure not be printed.
...	not used. For compatibility with the generic function.

Value

A list containing elements displayed in the summary:

- correlation: the correlation structure.
- variance: the variance structure.
- sd: the variance structure expressed in term of standard deviations.
- mean: the mean structure.

---

summary.mlmm *Summary of Multiple Linear Mixed Models*

Description

Estimates, p-values, and confidence intevals for multiple linear mixed models.

Usage

```
## S3 method for class 'mlmm'
summary(
  object,
  digits = 3,
  method = NULL,
  print = NULL,
  hide.data = FALSE,
  hide.fit = FALSE,
  ...
)
```

Arguments

object	an mlmm object, output of mlmm.
digits	[integer,>0] number of digits used to display numeric values.
method	[character] type of adjustment for multiple comparisons: one of "none", "bonferroni", "single-step", "single-step2".
print	[logical] should the output be printed in the console. Can be a vector of length 2 where the first element refer to the global tests and the second to the individual tests.

hide.data	[logical] should information about the dataset not be printed.
hide.fit	[logical] should information about the model fit not be printed.
...	other arguments are passed to summary.Wald_lmm.

---

summary.partialCor *Summary for partial correlation*

Description

Display estimated partial correlation and associated p-values and confidence intevals.

Usage

```
## S3 method for class 'partialCor'
summary(object, digits = 3, detail = TRUE, ...)
```

Arguments

object	a partialCor object, output of partialCor.
digits	[integer,>0] number of digits used to display numeric values.
detail	[integer,>0] passed to print.confint_lmm. If above 0.5 also display when a back-transformation has been used.
...	other arguments are passed to print.confint_lmm.

---

summary.Wald_lmm *Summary of Testing for a Linear Mixed Models*

Description

Estimates, p-values, and confidence intevals for linear hypothesis testing, possibly adjusted for multiple comparisons.

Usage

```
## S3 method for class 'Wald_lmm'
summary(
  object,
  print = TRUE,
  seed = NULL,
  columns = NULL,
  legend = TRUE,
  digits = 3,
  digits.df = 1,
  digits.p.value = 3,
  sep = ": ",
  ...
)
```

Arguments

object	an Wald_lmm object, output of anova.
print	[logical] should the output be printed in the console. Can be a vector of length 2 where the first element refer to the global tests and the second to the individual tests.
seed	[integer] value that will be set before adjustment for multiple comparisons to ensure reproducible results. Can also be NULL: in such a case no seed is set.
columns	[character vector] Columns to be displayed for each null hypothesis. Can be any of "type", "estimate", "se", "statistic", "df", "null", "lower", "upper", "p.value".##'
legend	[logical] should explanations about the content of the table be displayed.
digits	[interger, >0] number of digits used to display estimates.
digits.df	[interger, >0] number of digits used to display degrees of freedom.
digits.p.value	[interger, >0] number of digits used to display p-values.
sep	[character] character string used to separate the type of test (e.g. mean, variance) and the name of the test.
...	arguments method, level, and backtransform passed to confint.Wald_lmm

Details

By default adjustment for multiple comparisons via a single step max-test adjustment, either using the multcomp package (equal degrees of freedom) or the copula package (unequal degrees of freedom). When multiple multivariate Wald tests are performed, adjustment for multiple comparisons for the univariate Wald tests is performed within each multivariate Wald test. The number of tests ajusted for equal the first degree of freedom of the multivariate Wald statistic.

Adding the value "type" in argument "columns" ensures that the type of parameter that is being test (mean, variance, correlation) is output.

Value

NULL

---

swabsL *Data From The SWABS Study (Long Format)*

Description

Data from the swabs study, where the pneumococcus was studied in 18 families with different space available for the household. This dataset is in the long format (i.e. one line per measurement).

- crowding: space available in the household.
- family: family serial number
- name: type of family member.
- swabs: number of times the swab measurement was positive.

Usage

```
data(swabsL)
```

References

TODO

---

swabsW *Data From The SWABS Study (Wide Format)*

Description

Data from the swabs study, where the pneumococcus was studied in 18 families with different space available for the household. This dataset is in the wide format (i.e. one line per patient).

- `crowding`: space available in the household.
- `family`: family serial number
- `mother`: number of times the swab measurement was positive for the mother.
- `father`: number of times the swab measurement was positive for the father.
- `child1`: number of times the swab measurement was positive for the first child.
- `child2`: number of times the swab measurement was positive for the second child.
- `child3`: number of times the swab measurement was positive for the third child.

Usage

```
data(swabsW)
```

References

Grundy SM, Lan SP, Lachin J. The effects of chenodiol on biliary lipids and their association with gallstone dissolution in the National Cooperative Gallstone Study (SWABS). J Clin Invest. 1984 Apr;73(4):1156-66. doi: 10.1172/JCI111301.

terms.lmm *Model Terms For Linear Mixed Models*

Description

Model terms for linear mixed models. Used by `multcomp::glht`.

Usage

```
## S3 method for class 'lmm'
terms(x, ...)
```

Arguments

`x`	a `lmm` object
`...`	not used, for compatibility with the generic method.

Value

An object of class `terms` giving a symbolic representation of the mean structure.

TOEPLITZ *Toeplitz Structure*

Description

Variance-covariance structure for stationnary processes. Can be stratified on a categorical variable.

Usage

```
TOEPLITZ(formula, var.cluster, var.time, heterogeneous = "LAG", add.time)
```

Arguments

`formula`	formula indicating on which variable to stratify the residual variance and correlation (left hand side) and variables influencing the residual variance and correlation (right hand side).
`var.cluster`	[character] cluster variable.
`var.time`	[character] time variable.
`heterogeneous`	[character] degree of flexibility of the correlation structure within covariate (`"UN"`,`"LAG"`,`"CS"`). Will also affect the variance structure when not explicit.
`add.time`	Should the default formula (i.e. when `NULL`) contain a time effect.

Details

formula: there can only be at most one covariate for the correlation structure. A typical formula would be ~1, indicating a variance constant over time and a correlation specific to each gap time.

heterogeneous: for a binary covariate the correlation matrix can be decomposed into four blocs: A, B, B, C. A correspond the correlation within level 0 of the covariate, C within level 1, and B between level 0 and 1. Different correlation structures can be specified:

- "UN": unstructured matrix except for the diagonal elements of C which are constrained to be equal.
- "LAG": Toeplitz structure within A, B, and C, i.e. correlation specific to each time lag and covariate level.
- "CS": block-specific value except for C which has a different value for its diagonal elements.

Value

An object of class TOEPLITZ that can be passed to the argument structure of the lmm function.

Examples

```
## no covariate
TOEPLITZ(~1, var.cluster = "id", var.time = "time")
TOEPLITZ(gender~1, var.cluster = "id", var.time = "time")
TOEPLITZ(list(~time,~1), var.cluster = "id", var.time = "time")
TOEPLITZ(list(gender~time,gender~1), var.cluster = "id", var.time = "time")

## with covariates
TOEPLITZ(~1, var.cluster = "id", heterogeneous = "UN",
         var.time = "time", add.time = c("time","side"))
TOEPLITZ(~1, var.cluster = "id", heterogeneous = "LAG",
         var.time = "time", add.time = c("time","side"))
TOEPLITZ(~1, var.cluster = "id", heterogeneous = "CS",
         var.time = "time", add.time = c("time","side"))
TOEPLITZ(gender~1, var.cluster = "id", heterogeneous = "CS",
         var.time = "time", add.time = c("time","side"))
```

UN *Unstructured Structure*

Description

Variance-covariance structure where the residuals have time-specific variance and correlation. Can be stratified on a categorical variable.

Usage

```
UN(formula, var.cluster, var.time, add.time)
```

Arguments

formula	formula indicating on which variable to stratify the covariance structure.
var.cluster	[character] cluster variable.
var.time	[character] time variable.
add.time	Should the default formula (i.e. when NULL) contain a time effect.

Details

A typical formula would be ~1, indicating a time-specific variance parameter and a correlation parameter specific to each pair of times.

Value

An object of class UN that can be passed to the argument structure of the lmm function.

Examples

```
UN(NULL, var.cluster = "id", var.time = "time", add.time = TRUE)
UN(~gender, var.cluster = "id", var.time = "time", add.time = TRUE)
```

---

vasscoresL *Data From The VAS Study (Long Format)*

Description

Data from the VAS Study, a randomized controlled clinial trial assessing the healing effect of topical zink sulfate on epidermal wound. The study includes 30 heatlhy volunteers with induced wounds on each buttock which where subsequently treated with a different treatment for each wound. Then the VAS-score (pain sensation on a 0-100mm visual analogue scale) was assessed after each treatment application and summarized by area under the curve. This dataset is in the long format (i.e. one line per measurement).

- id: patient identifier.
- group: treatment group to which the patient has been randomized.
- treat.num:
- vas: VAS-score relative to the wound.
- treatment: Treatment used on the wound. A: active treatment (zink shower gel), B: placebo treatment (shower gel without zink), C: control treatment (demineralized water).

Usage

```
data(vasscoresL)
```

References

TODO

vasscoresW *Data From The VAS Study (Wide Format)*

Description

Data from the VAS Study, a randomized controlled clinial trial assessing the healing effect of topical zink sulfate on epidermal wound. The study includes 30 heatlhy volunteers with induced wounds on each buttock which where subsequently treated with a different treatment for each wound. Then the VAS-score (pain sensation on a 0-100mm visual analogue scale) was assessed after each treatment application and summarized by area under the curve. This dataset is in the wide format (i.e. one line per patient).

- `id`: patient identifier.
- `group`: treatment group to which the patient has been randomized.
- `vasA`: VAS-score when using a zink shower gel.
- `vasB`: VAS-score when using a placebo treatment (shower gel without zink).
- `vasC`: VAS-score when using a control treatment with demineralized water.

Usage

```
data(vasscoresW)
```

References

TODO

vcov.lmm *Extract The Variance-Covariance Matrix From a Linear Mixed Model*

Description

Extract the variance-covariance matrix of the model coefficients of a linear mixed model.

Usage

```
## S3 method for class 'lmm'
vcov(
  object,
  effects = "mean",
  robust = FALSE,
  df = FALSE,
  strata = NULL,
  data = NULL,
  p = NULL,
```

```
  type.information = NULL,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  ...
)
```

Arguments

object	a lmm object.
effects	[character] Should the variance-covariance matrix for all coefficients be output ("all"), or only for coefficients relative to the mean ("mean" or "fixed"), or only for coefficients relative to the variance structure ("variance"), or only for coefficients relative to the correlation structure ("correlation").
robust	[logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors. Not feasible for variance or correlation coefficients estimated by REML.
df	[logical] Should degree of freedom, computed using Satterthwaite approximation, for the model parameters be output.
strata	[character vector] When not NULL, only output the variance-covariance matrix for the estimated parameters relative to specific levels of the variable used to stratify the mean and covariance structure.
data	[data.frame] dataset relative to which the information should be computed. Only relevant if differs from the dataset used to fit the model.
p	[numeric vector] value of the model coefficients at which to evaluate the information. Only relevant if differs from the fitted values.
type.information	[character] Should the expected information be used (i.e. minus the expected second derivative) or the observed inforamtion (i.e. minus the second derivative).
transform.sigma	[character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.
transform.k	[character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details.
transform.rho	[character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.
transform.names	[logical] Should the name of the coefficients be updated to reflect the transformation that has been used?
...	Not used. For compatibility with the generic method.

Details

For details about the arguments **transform.sigma**, **transform.k**, **transform.rho**, see the documentation of the coef.lmm function.

Value

A matrix with an attribute `"df"` when argument df is set to `TRUE`.

`vitaminL` *Data From The Vitamin Study (Long Format)*

Description

Data from the vitamin Study, a randomized study where the growth of guinea pigs was monitored before and after intake of vitamin E/placebo. The weight of each guinea pig was recorded at the end of week 1, 3, 4, 5, 6, and 7. Vitamin E/placebo is given at the beginning of week 5. This dataset is in the long format (i.e. one line per measurement).

- `group`: treatment group: vitamin or placebo.
- `animal`: identifier
- `weight1`: weight (in g) of the pig at the end of week 1 (before treatment).
- `weight3`: weight (in g) of the pig at the end of week 3 (before treatment).
- `weight4`: weight (in g) of the pig at the end of week 4 (before treatment).
- `weight5`: weight (in g) of the pig at the end of week 5 (after treatment).
- `weight6`: weight (in g) of the pig at the end of week 6 (after treatment).
- `weight7`: weight (in g) of the pig at the end of week 7 (after treatment).

Usage

```
data(vitaminL)
```

References

Crowder and Hand (1990, p. 27) Analysis of Repeated Measures.

vitaminW *Data From The Vitamin Study (Wide Format)*

Description

Data from the vitamin Study, a randomized study where the growth of guinea pigs was monitored before and after intake of vitamin E/placebo. The weight of each guinea pig was recorded at the end of week 1, 3, 4, 5, 6, and 7. Vitamin E/placebo is given at the beginning of week 5. This dataset is in the wide format (i.e. one line per patient).

- `group`: treatment group: vitamin or placebo.
- `animal`: identifier
- `weight1`: weight (in g) of the pig at the end of week 1 (before treatment).
- `weight3`: weight (in g) of the pig at the end of week 3 (before treatment).
- `weight4`: weight (in g) of the pig at the end of week 4 (before treatment).
- `weight5`: weight (in g) of the pig at the end of week 5 (after treatment).
- `weight6`: weight (in g) of the pig at the end of week 6 (after treatment).
- `weight7`: weight (in g) of the pig at the end of week 7 (after treatment).

Usage

```
data(vitaminW)
```

References

TODO

weights.Wald_lmm *Extract Weights Used to Pool Estimates*

Description

Extract weights used to pool estimates.

Usage

```
## S3 method for class 'Wald_lmm'
weights(object, method, ...)
```

Arguments

`object`	a `Wald_lmm` object, output of `anova.lmm`, or `rbind.lmm`, or `mlmm`.
`method`	[character] method for combining the estimates, one of `"average"`, `"pool.se"`, `"pool.gls"`, `"pool.rubin"`.
`...`	Not used. For compatibility with the generic method.

Value

a numeric vector whose elements sum to 1.

Examples

```
set.seed(10)
dL <- sampleRem(250, n.times = 3, format = "long")

e.mlmm <- mlmm(Y~X1+X2+X6, repetition = ~visit|id, data = dL,
               by = "X4", effects = "X1=0", structure = "CS")
weights(e.mlmm, method = "average")
weights(e.mlmm, method = "pool.fixse")
weights(e.mlmm, method = "pool.se")
weights(e.mlmm, method = "pool.gls")
```

Index