

# Package ‘LMMstar’

August 31, 2022

**Type** Package

**Title** Repeated Measurement Models for Discrete Times

**Version** 0.7.6

**Date** 2022-08-31

**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, unstructured, ...). Statistical inference for mean, variance, and correlation parameters is performed based on the observed information and a Satterthwaite 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, lava, Matrix, multcomp, numDeriv, sandwich

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

**VignetteBuilder** R.rsp

**RoxygenNote** 7.2.0

**Collate** '0-onload.R' 'LMMstar-package.R' 'LMMstar.options.R' 'anova.R' 'autoplot.R' 'baselineAdjustment.R' 'coef.lmm.R' 'coef.mlmm.R' 'confint.anova\_lmm.R' 'confint.lmm.R' 'confint.mlmm.R' 'constrain.R' 'df.R' 'doc-data.R' 'dummy.coef.R' 'emmeans.R'

'estimate.lmm.R' 'estimate.mlmm.R' 'findPatterns.R' 'fitted.R'  
 'formula.R' 'iid.R' 'information.R' 'levels.R' 'lmm.R'  
 'logLik.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.anova\_lmm.R' 'print.lmm.R' 'print.mlmm.R' 'profile.R'  
 'ranef.R' 'rbind.anova\_lmm.R' 'reparametrize.R' 'residuals.R'  
 'sampleRem.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' 'summary.anova\_lmm.R'  
 'summary.lmm.R' 'summary.mlmm.R' 'terms.R'  
 'transformSummaryTable.R' 'unorderedPairs.R' 'utils.R' 'vcov.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** 2022-08-31 13:40:02 UTC

## R topics documented:

|                              |    |
|------------------------------|----|
| LMMstar-package . . . . .    | 4  |
| abetaL . . . . .             | 5  |
| abetaW . . . . .             | 6  |
| anova . . . . .              | 6  |
| autoplot . . . . .           | 9  |
| baselineAdjustment . . . . . | 10 |
| blandAltmanL . . . . .       | 11 |
| blandAltmanW . . . . .       | 12 |
| bloodpressureL . . . . .     | 13 |
| calciumL . . . . .           | 13 |
| calciumW . . . . .           | 14 |
| ckdL . . . . .               | 15 |
| ckdW . . . . .               | 15 |
| coef . . . . .               | 16 |
| confint . . . . .            | 18 |
| confint.anova_lmm . . . . .  | 20 |
| confint.mlmm . . . . .       | 21 |
| CS . . . . .                 | 22 |
| CUSTOM . . . . .             | 23 |
| dummy.coef.lmm . . . . .     | 25 |
| estfun . . . . .             | 26 |
| estimate.lmm . . . . .       | 27 |
| fitted.lmm . . . . .         | 29 |
| gastricbypassL . . . . .     | 30 |

|                                 |    |
|---------------------------------|----|
| gastricbypassW . . . . .        | 31 |
| getVarCov.lmm . . . . .         | 32 |
| ID . . . . .                    | 32 |
| iid.lmm . . . . .               | 33 |
| IND . . . . .                   | 34 |
| information . . . . .           | 35 |
| levels.lmm . . . . .            | 36 |
| lmm . . . . .                   | 37 |
| LMMstar.options . . . . .       | 40 |
| LMMstar2emmeans . . . . .       | 42 |
| logLik . . . . .                | 42 |
| mlmm . . . . .                  | 43 |
| model.tables . . . . .          | 45 |
| ncgsL . . . . .                 | 45 |
| ncgsW . . . . .                 | 46 |
| partialCor . . . . .            | 46 |
| plot . . . . .                  | 48 |
| potassiumRepeatedL . . . . .    | 49 |
| potassiumSingleL . . . . .      | 50 |
| potassiumSingleW . . . . .      | 50 |
| predict.lmm . . . . .           | 51 |
| profile . . . . .               | 53 |
| rbind.anova_lmm . . . . .       | 55 |
| residuals . . . . .             | 56 |
| sampleRem . . . . .             | 59 |
| schoolL . . . . .               | 60 |
| score . . . . .                 | 61 |
| sigma . . . . .                 | 62 |
| summarize . . . . .             | 63 |
| summary . . . . .               | 65 |
| summary.anova_lmm . . . . .     | 67 |
| summary.mlmm . . . . .          | 68 |
| swabsL . . . . .                | 69 |
| swabsW . . . . .                | 69 |
| terms.lmm . . . . .             | 70 |
| transformSummaryTable . . . . . | 70 |
| UN . . . . .                    | 71 |
| vasscoresL . . . . .            | 71 |
| vasscoresW . . . . .            | 72 |
| vcov . . . . .                  | 73 |
| vitaminL . . . . .              | 74 |
| vitaminW . . . . .              | 75 |

## 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, unstructured, ...). Statistical inference for mean, variance, and correlation parameters is performed based on the observed information and a Satterthwaite 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:

$$\mathbf{Y}_i = \mathbf{X}_i\beta + \varepsilon_i$$

where

- $\mathbf{Y}$ : vector of outcome.
- $\mathbf{X}$ : design matrix (extractor: `model.matrix.lmm`).
- $\varepsilon$ : vector of residuals with 0-mean and variance  $\Omega$  (extractor: `residuals.lmm`).
- $\beta$ : estimated mean coefficients relative to  $X$  (extractor: `coef.lmm`).
- $\Omega$ : the modeled variance-covariance of the residuals with diagonal elements  $\omega$  (extractor: `getVarCov.lmm`).
- $i$  indexes the cluster (level where replicates are assumed independent).
- $j$  indexes the repetitions, e.g. the variance of  $\varepsilon_{ij}$  is  $\omega_{ij}$ .

**Covariance patterns:**  $\Omega$  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.
- "UN": unstructured (time-specific correlation, time-specific variable).

It possible to stratify the last two structure with respect to a categorical variable.

**Optimizer:** the default optimizer is `nls::gls` which is restricted to certain covariance patterns. To use the other covariance patterns switch to the optimizer "FS". This may fail for complex covariance patterns in small samples since  $\Omega$  is not constrained to be positive definite.

---

`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` M/F.
- `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 M/F
- 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.

---

 anova

*Multivariate Wald Tests For Linear Mixed Model*


---

### Description

Simultaneous tests of linear combinations of the model paramaters using Wald tests.

**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] 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, similarly to the <code>linfct</code> argument of the <code>multcomp::glht</code> 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 <code>effects</code> 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 a confidence interval be output for each hypothesis?  |
| transform.sigma, transform.k, transform.rho, transform.names | are passed to the <code>vcov</code> method. See details section in <a href="#">coef.lmm</a> .  |
| ...  | Not used. For compatibility with the generic method.   |

**Details**

By default confidence intervals and p-values are adjusted based on the distribution of the maximum-statistic. This is referred to as a single-step Dunnett multiple testing procedures in table II of Dmitrienko et al. (2013) and is performed using the `multcomp` package with the option `test = adjusted("single-step")`.

**Value**

A list of matrices containing the following columns:

- `null`: null hypothesis

- `statistic`: value of the test statistic
- `df.num`: degrees of freedom for the numerator (i.e. number of hypotheses)
- `df.denom`: degrees of freedom for the denominator (i.e. Satterthwaite approximation)
- `p.value`: p-value.

as well as an attribute contrast containing the contrast matrix encoding the linear combinations of coefficients (in columns) for each hypothesis (in rows).

## 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.anova\\_lmm](#) for a summary of the results.

## 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)

## chi-2 test
summary(anova(eUN.lmm, df = FALSE))

## F-test
anova(eUN.lmm)
summary(anova(eUN.lmm, effects = "all"))
anova(eUN.lmm, effects = c("X1=0", "X2+X5=10"))

## another example
if(require(multcomp)){
  amod <- lmm(breaks ~ tension, data = warpbreaks)
  e.glht <- glht(amod, linfct = mcp(tension = "Tukey"))
  summary(e.glht, test = Chisqtest()) ## 0.000742

  e.amod <- anova(amod, effect = mcp(tension = "Tukey"))
  summary(e.amod)
}
```



**Description**

Graphical Display For Linear Mixed Models

**Usage**

```
## S3 method for class 'lmm'
autoplot(
  object,
  obs.alpha = 0,
  obs.size = c(2, 0.5),
  at = NULL,
  time.var = NULL,
  color = TRUE,
  ci = TRUE,
  ci.alpha = NA,
  plot = TRUE,
  mean.size = c(3, 1),
  size.text = 16,
  position.errorbar = "identity",
  ...
)
```

**Arguments**

|                                |  |
|--------------------------------|--|
| <code>object</code>            | a lmm object.  |
| <code>obs.alpha</code>         | [numeric, 0-1] When not NA, transparency parameter used to display the original data by cluster. |
| <code>obs.size</code>          | [numeric vector of length 2] size of the point and line for the original data.                   |
| <code>at</code>                | [data.frame] values for the covariates at which to evaluate the fitted values.                   |
| <code>time.var</code>          | [character] x-axis variable for the plot.  |
| <code>color</code>             | [character] name of the variable in the dataset used to color the curve.                         |
| <code>ci</code>                | [logical] should confidence intervals be displayed?  |
| <code>ci.alpha</code>          | [numeric, 0-1] When not NA, transparency parameter used to display the confidence intervals.     |
| <code>plot</code>              | [logical] should the plot be displayed?  |
| <code>mean.size</code>         | [numeric vector of length 2] size of the point and line for the mean trajectory.                 |
| <code>size.text</code>         | [numeric, >0] size of the font used to displayed text when using ggplot2.                        |
| <code>position.errorbar</code> | [character] relative position of the errorbars.  |
| <code>...</code>               | arguments passed to the predict method.  |

**Value**

A list with two elements

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

---

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 constrained.  |
| 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)

e1.lmm <- suppressWarnings(lmm(cholest~visit*treat,
                              data=ncgsL, repetition= ~ visit|id,
                              structure = "CS"))
e1bis.lmm <- suppressWarnings(lmm(cholest~treattime,
                                 data=ncgsL, repetition= ~ visit|id,
                                 structure = "CS"))

## 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)

e2.lmm <- suppressWarnings(lmm(cholest~visit*treat2,
                              data=ncgsL, repetition= ~ visit|id,
                              structure = "CS"))
e2bis.lmm <- suppressWarnings(lmm(cholest~treattime2,
                                 data=ncgsL, repetition= ~ visit|id,
                                 structure = "CS"))

```

**Description**

Data From The Bland Altman Study where two methods to measure the peak expiratory flow rate (PEFR) were 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) were 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 three 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/cm<sup>3</sup>)
- `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/cm<sup>3</sup>).
- bmd2 Bone mineral density measured at the second visit (in mg/cm<sup>3</sup>).
- bmd3 Bone mineral density measured at the third visit (in mg/cm<sup>3</sup>).
- bmd4 Bone mineral density measured at the fourth visit (in mg/cm<sup>3</sup>).
- bmd5 Bone mineral density measured at the fifth visit (in mg/cm<sup>3</sup>).

**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

*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"). |
| 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" output residual standard error.
- "log" output log-transformed residual standard error.
- "square" output residual variance.
- "logsquare" output log-transformed residual variance.

### transform.k:

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

### transform.rho:

- "none" output correlation coefficient.
- "atanh" output correlation coefficient after tangent hyperbolic transformation.
- "cov" output 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.

**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

*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", "partial.R".  |
| 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 <a href="#">coef.lmm</a> .   |
| backtransform  | [logical] should the variance/covariance/correlation coefficient be backtransformed?   |
| ...  | Not used. For compatibility with the generic method.   |

## Value

A data.frame containing for each 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.

## 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)

## based on a Student's t-distribution with transformation
confint(eUN.lmm)
## based on a Student's t-distribution without transformation
```

```
confint(eUN.lmm, transform.sigma = "none", transform.k = "none", transform.rho = "none")
## based on a Normal distribution with transformation
confint(eUN.lmm, df = FALSE)
```

---

confint.anova\_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 'anova_lmm'
confint(
  object,
  parm,
  level = 0.95,
  method = NULL,
  columns = NULL,
  simplify = TRUE,
  ...
)
```

## Arguments

|          |   |
|----------|---|
| object   | a anova_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", "single-step", "single-step2".                          |
| columns  | [character vector] Columns to be output. Can be any of "estimate", "se", "statistic", "df", "null", "lower", "upper", "p.value", "partial.R". |
| simplify | [logical] Return a data.frame instead of a list containing a data.frame when possible.  |
| ...      | Not used. For compatibility with the generic method.  |

## Details

Method "single-step" adjust for multiple comparisons using quantiles of the multivariate Student's t-distribution, assuming equal degrees of freedom in the marginal. This is performed by the multcomp package.

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. This is performed by the copula package.

---

 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 <a href="#">confint.anova_lmm</a> .   |

### 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.

**Description**

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

**Usage**

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

**Arguments**

|                            |  |
|----------------------------|--|
| <code>formula</code>       | formula indicating on which variable to stratify the residual variance and correlation (left hand side) and variables influencing the residual variance (right hand side). |
| <code>var.cluster</code>   | [character] cluster variable.  |
| <code>var.time</code>      | [character] time variable.   |
| <code>heterogeneous</code> | [logical] when covariates are used for the correlation structure, should correlation parameters should be specific to each level of the covariate?                         |
| <code>add.time</code>      | 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 Imm 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))

```

---

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 not 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

---

`estfun`*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

|                  |  |
|------------------|--|
| <code>x</code>   | a lmm object.  |
| <code>...</code> | 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 = TRUE,
  robust = FALSE,
  type.information = NULL,
  level = 0.95,
  method.numDeriv = NULL,
  average = FALSE,
  transform.sigma = "none",
  transform.k = "none",
  transform.rho = "none",
  ...
)
```

## 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 information (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
}
}

```

---

|            |  |
|------------|--|
| fitted.lmm | <i>Predicted Mean Value For Linear Mixed Model</i> |
|------------|--|

---

## 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 | <i>Data From The Gastric Bypass Study (Wide Format)</i> |
|----------------|---|

---

### 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 | <i>Deprecated Extractor of the Residual Variance-Covariance Matrix</i> |
|---------------|--|

---

**Description**

Deprecated 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 | <i>identity Structure</i> |
|----|---------------------------|

---

**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(
  object,
  effects = "mean",
  robust = TRUE,
  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] 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.  |

|                               |   |
|-------------------------------|---|
| <code>type.information</code> | [character] Should the expected information be used (i.e. minus the expected second derivative) or the observed information (i.e. minus the second derivative).                       |
| <code>transform.sigma</code>  | [character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.                                       |
| <code>transform.k</code>      | [character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details. |
| <code>transform.rho</code>    | [character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.   |
| <code>transform.names</code>  | [logical] Should the name of the coefficients be updated to reflect the transformation that has been used?  |
| <code>...</code>              | 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**

|                          |   |
|--------------------------|---|
| <code>formula</code>     | 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. |
| <code>var.cluster</code> | [character] cluster variable.   |
| <code>var.time</code>    | [character] time variable.  |
| <code>add.time</code>    | 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

---

*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 information (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](#) function.

### Value

When argument `indiv` is `FALSE`, a matrix with the value of the information 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 model.

### 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

---

 Imm

*Fit Linear Mixed Model*


---

**Description**

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

**Usage**

```
Imm(
  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. $Y \sim \text{Gender} + \text{Gene}$ . |
| repetition  | [formula] Specify the structure of the data: the time/repetition variable and the grouping variable, e.g. $\sim \text{timelid}$ .   |
| 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 Likelihood ("REML") or Maximum Likelihood ("ML") be used to estimate the model parameters?  |

|                               |  |
|-------------------------------|--|
| <code>df</code>               | [logical] Should the degree of freedom be computed using a Satterthwaite approximation?  |
| <code>type.information</code> | [character] Should the expected information be computed (i.e. minus the expected second derivative) or the observed information (i.e. minus the second derivative).              |
| <code>trace</code>            | [integer, >0] Show the progress of the execution of the function.  |
| <code>control</code>          | [list] Control values for the optimization method. The element <code>optimizer</code> indicates which optimizer to use and additional arguments will be passed to the optimizer. |

## Details

**Computation time** the `lmm` has not been developed 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 requires 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 the `nls::gls` function. 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 iterations 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.

## 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 hypotheses 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 output when using `anova` with argument `ci=TRUE`.
- `columns.confint` [character vector]: columns to output when using `confint`.
- `columns.summary` [character vector]: columns to output 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.anova_lmm`.
- `optimizer` [character]: method used to estimate the model parameters: can be `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 parameters.
- `trace` [logical]: Should the progress of the execution of the `lmm` function be displayed?
- `transform.sigma`, `transform.k`, `transform.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.

---

LMMstar2emmeans      *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.

---

logLik      *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****transform:**

- 0 means no transformation i.e. output standard error, ratio of standard errors, and correlations.
- 1 means log/atanh transformation i.e. output log(standard error), log(ratio of standard errors), and atanh(correlations).
- 2 output variance coefficients and correlations.

**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).

---

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, effects = NULL, robust = FALSE, df = TRUE, ci = TRUE)
```

**Arguments**

|                      |  |
|----------------------|--|
| ...                  | arguments passed to <code>lmm</code> .   |
| <code>data</code>    | [data.frame] dataset (in the long format) containing the observations.   |
| <code>by</code>      | [character] variable used to split the dataset. On each split a separate linear mixed model is fit.  |
| <code>effects</code> | [character] Linear combinations of coefficients relative to which Wald test should be computed.  |
| <code>robust</code>  | [logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors. Argument passed to <code>anova.lmm</code> . |
| <code>df</code>      | [logical] Should the degree of freedom be computed using a Satterthwaite approximation?  |
| <code>ci</code>      | [logical] Should a confidence interval be output for each hypothesis?  |

**Examples**

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

  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.lmmX1, method = "single-step")
```

---

model.tables

*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, ...)
```

### Arguments

|     |  |
|-----|--|
| x   | a lmm object.  |
| ... | arguments to be passed to the <code>confint</code> method. Should not contain the argument column. |

### 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 chenodiol (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 between two variables where the adjustment set may differ between variables.

**Usage**

```
partialCor(formula, data, repetition)
```

**Arguments**

|            |   |
|------------|---|
| formula    | 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. |
| 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.  |

**Details**

Fit a mixed model to estimate the partial correlation which can be time consuming.

**Examples**

```
#### bivariate (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)
## estimate
##           hl      disp      deg      BC
## hl      1.0000000 -0.6720863 -0.6161163 0.1148459
## disp -0.6720863  1.0000000 -0.7215522 0.2855420
## deg  -0.6161163 -0.7215522  1.0000000 0.6940953
## BC    0.1148459  0.2855420  0.6940953 1.0000000

## $p.value
##           hl      disp      deg      BC
## hl      0.0000000 0.06789202 0.10383620 0.78654997
## disp 0.06789202 0.00000000 0.04332869 0.49299871
## deg  0.10383620 0.04332869 0.00000000 0.05615021
## BC   0.78654997 0.49299871 0.05615021 0.00000000

set.seed(10)
y.data$gender <- factor(rbinom(10, size = 1, prob = 0.5), labels = c("F","M"))

partialCor(c(hl,disp)~BC+deg, data = y.data)
partialCor(hl + disp~BC+deg, data = y.data)

partialCor(list(hl~BC+deg, disp~BC), data = y.data)
partialCor(list(hl~BC+deg+gender, disp~1), data = y.data)

#### bivariate (with repetition) ####
data(gastricbypassL, package = "LMMstar")
```

```
partialCor(weight+glucagonAUC~time, data = gastricbypassL)
partialCor(weight+glucagonAUC~time, repetition =~time|id, data = gastricbypassL)
```

plot

*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'
plot(
  x,
  type = "fit",
  type.residual = "normalized",
  by.time = TRUE,
  ci = TRUE,
  plot = TRUE,
  ci.alpha = 0.2,
  mean.size = c(3, 1),
  size.text = 16,
  ...
)
```

**Arguments**

|               |  |
|---------------|--|
| x             | a lmm object.  |
| type          | [character] the type of plot: "fit", "qqplot", "correlation", "scatterplot", "scatterplot2", "partial".  |
| 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.                      |
| by.time       | [logical] should a separate plot be made at each repetition or a single plot over all repetitions be used? Only relevant for type="qqplot", type="scatterplot", and type="scatterplot2". |
| ci            | [logical] should confidence intervals be displayed?  |
| plot          | [logical] should the plot be displayed?  |
| ci.alpha      | [numeric, 0-1] Transparency parameter used to display the confidence intervals.  |
| 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 displayed text when using ggplot2.  |
| ...           | additional argument passed to residuals.lmm or autoplot.lmm.   |



**Details**

Call `autoplot.lmm` when `codetype=="fit"` and `residuals.lmm` for the other types.

**Value**

A list with two elements

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

---

|                    |  |
|--------------------|--|
| potassiumRepeatedL | <i>Data From The Potassium Intake Study (Long Format with intermediate measurements)</i> |
|--------------------|--|

---

**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-angiotensin-aldosterone 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-angiotensin-aldosterone 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-angiotensin-aldosterone 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 | <i>Predicted Mean Value With Uncertainty For Linear Mixed Model</i> |
|-------------|---|

---

**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

---

*Display 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,
  plot = TRUE,
  ci = FALSE,
  size = c(3, 2, 1, 1),
  linetype = c("dashed", "dashed", "dashed"),
  shape = 19,
  scales = "free",
  nrow = NULL,
  ncol = NULL,
  transform.sigma = NULL,
  transform.k = NULL,
  transform.rho = NULL,
  transform.names = TRUE,
  ...
)

```

**Arguments**

|                                 |   |
|---------------------------------|---|
| <code>fitted</code>             | a <code>lmm</code> object.  |
| <code>effects</code>            | [character vector] name of the parameters who will be constrained. Alternatively can be the type of parameters, e.g. "mean", "variance", "correlation", or "all".   |
| <code>profile.likelihood</code> | [logical] should profile likelihood be performed? Otherwise varying one parameter at a time around the MLE while keeping the other constant).   |
| <code>maxpts</code>             | [integer, >0] number of points use to discretize the likelihood, <code>maxpts</code> points smaller than the MLE and <code>maxpts</code> points higher than the MLE.  |
| <code>conf.level</code>         | [numeric, 0-1] the confidence level of the confidence intervals used to decide about the range of values for each parameter.  |
| <code>trace</code>              | [logical] Show the progress of the execution of the function.   |
| <code>plot</code>               | [logical] Should a graphical representation of the results be provided?   |
| <code>ci</code>                 | [logical] Should a 95% confidence intervals obtained from the Wald test (vertical lines) and Likelihood ratio test (horizontal line) be displayed?  |
| <code>size</code>               | [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. |
| <code>linetype</code>           | [integer vector of length 2] type of line used to represent the quadratic approximation of the likelihood and the confidence intervals.   |
| <code>shape</code>              | [integer, >0] type of point used to represent the MLE.  |
| <code>scales, nrow, ncol</code> | argument passed to <code>ggplot2::facet_wrap</code> .   |
| <code>transform.sigma</code>    | [character] Transformation used on the variance coefficient for the reference level. One of "none", "log", "square", "logsquare" - see details.   |
| <code>transform.k</code>        | [character] Transformation used on the variance coefficients relative to the other levels. One of "none", "log", "square", "logsquare", "sd", "logsd", "var", "logvar" - see details.                                       |
| <code>transform.rho</code>      | [character] Transformation used on the correlation coefficients. One of "none", "atanh", "cov" - see details.   |
| <code>transform.names</code>    | [logical] Should the name of the coefficients be updated to reflect the transformation that has been used?  |
| <code>...</code>                | Not used. For compatibility with the generic method.  |

**Details**

Each parameter defined by the argument `effects` 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. When argument `plot` is `TRUE`, also contain an attribute `"plot"` containing a `ggplot` object.

### Examples

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

## profile logLikelihood
## Not run:
profile(e.lmm, effects = "all", maxpts = 10, profile.likelihood = TRUE)

## End(Not run)

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

---

rbind.anova\_lmm

---

*Linear Hypothesis Testing Across Linear Mixed Models*


---

### Description

Linear hypothesis testing across linear mixed model.

### Usage

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

### Arguments

|                    |   |
|--------------------|---|
| <code>model</code> | a <code>anova_lmm</code> object (output of <code>anova</code> applied to a <code>lmm</code> object)   |
| <code>...</code>   | possibly other <code>anova_lmm</code> objects   |
| <code>name</code>  | [character vector or <code>NULL</code> ] character used to identify each model in the output. By default, use the name of the outcome of the model. |
| <code>sep</code>   | [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)
```

---

residuals

---

*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",
  format = "long",
  data = NULL,
  p = NULL,
  keep.data = FALSE,
  var = NULL,
  plot = "none",
  engine.qqplot = "ggplot2",
  add.smooth = TRUE,
  digit.cor = 2,
  size.text = 16,
  scales = "free",
  ...
)
```



**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. |
| 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").   |
| 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.   |
| keep.data     | [logical] Should the argument data be output along side the residuals? Only possible in the long format.  |
| var           | [character vector] name of the variable relative to which the partial residuals should be computed.   |
| plot          | [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?   |
| engine.qqplot | [character] Should ggplot2 or qqtest be used to display quantile-quantile plots? Only used when argument plot 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 plot="scatterplot".  |
| digit.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        | [character] Passed to ggplot2::facet_wrap.  |
| ...           | Not used. For compatibility with the generic method.  |

**Details**

The argument type defines how the residuals are computed:

- "fitted": fitted value  $X_{ij}\hat{\beta}$ .
- "raw": 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 (lower) 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).
- "partial": partial residuals ( $\gamma E + \hat{\varepsilon}$ ). A reference level can 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 lower Cholesky factor of  $\hat{\Omega}$ , i.e.  $\hat{C}\hat{C}^t = \hat{\Omega}$
- $\hat{Q}_i = X_i(X_i^t\hat{\Omega}X_i)^{-1}X_i^t$  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 lower Cholesky factor of  $\hat{\Omega} - \hat{Q}_i$

## Value

When argument format is "long" and type.object is "lmm", a vector containing the value of the residual relative to each observation. It is a matrix if the argument type contains several values. When argument format is "wide" and type.object 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 + X5, data = dL)

## partial residuals
residuals(e.lm, type = "partial", var = "X1")
## residuals(e.lm) + dL$X1 * coef(e.lm)["X1"]
residuals(e.lm, type = "partial", var = "X1", keep.data = TRUE)

## partial residuals
type <- "partial"
attr(type, "reference") <- data.frame(visit=factor(2,1:3), X2=0, X5=0)
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
residuals(eUN.lmm, format = "long", type = c("normalized", "pearson"))
residuals(eUN.lmm, format = "wide", plot = "correlation")
residuals(eUN.lmm, format = "wide", type = "normalized")
residuals(eUN.lmm, format = "wide", type = "scaled")

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

---

sampleRem

*Sample Longitudinal Data*


---

## Description

Sample longitudinal 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 ( $\eta$ ) with a coefficient lambda:

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

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

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

$\epsilon_j$  and  $\xi$  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")
```

---

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 | <i>Extract The Score From a Linear Mixed Model</i> |
|-------|--|

---

### 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` 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 | <i>Extract The Residuals Variance-Covariance Matrix From a Linear Mixed Model</i> |
|-------|---|

---

### 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,
  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 <code>NULL</code> , 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.  |
| 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 `simplified=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. $Y1+Y2 \sim \text{Gender} + \text{Gene}$ will compute for each gender and gene the summary statistics for Y1 and for Y2. Passed to the <code>stats::aggregate</code> 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 <code>stats::aggregate</code> 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: <ul style="list-style-type: none"> <li>• "observed": number of observations with a measurement.</li> <li>• "missing": number of observations with a missing value.</li> <li>• "mean", "mean.lower" "mean.upper": mean with its confidence interval.</li> <li>• "median", "median.lower" "median.upper": median with its confidence interval.</li> <li>• "sd": standard deviation.</li> <li>• "q1", "q3", "IQR": 1st and 3rd quartile, interquartile range.</li> <li>• "min", "max": minimum and maximum observation.</li> <li>• "predict.lower", "predict.upper": prediction interval for normally distributed outcome.</li> <li>• "correlation": correlation matrix between the outcomes (when feasible, see detail section).</li> </ul> |
| 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 <code>print.data.frame</code>   |
| ...            | 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.  $Y \sim \text{timelid}$ .



**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")
```

---

summary

*Summary Output for a Linear Mixed Model*


---

**Description**

Summary output for a linear mixed model fitted with `lmm`. This is a modified version of the `nlme::summary.gls` function.

**Usage**

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

**Arguments**

|                        |  |
|------------------------|--|
| <code>object</code>    | [lmm] output of the <code>lmm</code> function.   |
| <code>digit</code>     | [integer,>0] number of digit used to display numeric values.   |
| <code>level</code>     | [numeric,0-1] confidence level for the confidence intervals.   |
| <code>type.cor</code>  | [character] should the correlation matrix be display (" <code>matrix</code> ") or the parameter values (" <code>param</code> ").   |
| <code>robust</code>    | [logical] Should robust standard errors (aka sandwich estimator) be output instead of the model-based standard errors.   |
| <code>print</code>     | [logical] should the output be printed in the console.   |
| <code>columns</code>   | [character vector] Columns to be output for the fixed effects. Can be any of " <code>estimate</code> ", " <code>se</code> ", " <code>statistic</code> ", " <code>df</code> ", " <code>null</code> ", " <code>lower</code> ", " <code>upper</code> ", " <code>p.value</code> ". |
| <code>hide.data</code> | [logical] should information about the dataset not be printed.   |
| <code>hide.fit</code>  | [logical] should information about the model fit not be printed.   |
| <code>hide.cor</code>  | [logical] should information about the correlation structure not be printed.   |
| <code>hide.var</code>  | [logical] should information about the variance not be printed.  |
| <code>hide.sd</code>   | [logical] should information about the standard deviation not be printed.  |
| <code>hide.mean</code> | [logical] should information about the mean structure not be printed.  |
| <code>...</code>       | 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.anova\_lmm      *Summary of Testing for a Linear Mixed Models*

---

## Description

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

## Usage

```
## S3 method for class 'anova_lmm'
summary(
  object,
  method = NULL,
  transform = NULL,
  level = 0.95,
  print = TRUE,
  seed = NULL,
  columns = NULL,
  digits = max(3L, getOption("digits") - 2L),
  digits.p.value = max(3L, getOption("digits") - 2L),
  ...
)
```

## Arguments

|                |   |
|----------------|---|
| object         | an anova_lmm object, output of anova.   |
| method         | [character] type of adjustment for multiple comparisons: one of "none", "bonferroni", "single-step", "single-step2".  |
| transform      | [function] function to backtransform the estimates, standard errors, null hypothesis, and the associated confidence intervals (e.g. exp if the outcomes have been log-transformed). |
| level          | [numeric 0-1] level of the confidence intervals.  |
| 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 "estimate", "se", "statistic", "df", "null", "lower", "upper", "p.value".                        |
| digits         | [integer] number of digits used to display estimates.   |
| digits.p.value | [integer] number of digits used to display p-values.  |
| ...            | Not used. For compatibility with the generic method.  |

## 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).

---

 summary.mlmm

*Summary of Multiple Linear Mixed Models*


---

## Description

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

## Usage

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

## Arguments

|           |  |
|-----------|--|
| object    | an mlmm object, output of mlmm.  |
| 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 <a href="#">summary.anova_lm</a> .   |

---

`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 | <i>Model Terms For Linear Mixed Models</i> |
|-----------|--|

---

**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.

---

|                       |  |
|-----------------------|--|
| transformSummaryTable | <i>Apply Transformation to Summary Table</i> |
|-----------------------|--|

---

**Description**

Update summary table according to a transformation, e.g. log-transformation. P-values are left unchanged but estimates, standard errors, and confidence intervals are updated.

**Usage**

```
transformSummaryTable(object, transform = NULL)
```

**Arguments**

|           |   |
|-----------|---|
| object    | A data.frame with columns estimate, se, lower, upper. |
| transform | the name of a transformation or a function.           |

**Value**

a data.frame

---

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

|                          |  |
|--------------------------|--|
| <code>formula</code>     | formula indicating on which variable to stratify the covariance structure. |
| <code>var.cluster</code> | [character] cluster variable.  |
| <code>var.time</code>    | [character] time variable.   |
| <code>add.time</code>    | 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 clinical trial assessing the healing effect of topical zinc sulfate on epidermal wound. The study includes 30 healthy volunteers with induced wounds on each buttock which were 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 (zinc shower gel), B: placebo treatment (shower gel without zinc), 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 clinical trial assessing the healing effect of topical zinc sulfate on epidermal wound. The study includes 30 healthy volunteers with induced wounds on each buttock which were 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 zinc shower gel.
- vasB VAS-score when using a placebo treatment (shower gel without zinc).
- vasC VAS-score when using a control treatment with demineralized water.

**Usage**

```
data(vasscoresW)
```

**References**

TODO



**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**

|                      |  |
|----------------------|--|
| <code>object</code>  | a <code>lmm</code> object.   |
| <code>effects</code> | [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"). |
| <code>robust</code>  | [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.  |
| <code>df</code>      | [logical] Should degree of freedom, computed using Satterthwaite approximation, for the model parameters be output.  |
| <code>strata</code>  | [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.  |
| <code>data</code>    | [data.frame] dataset relative to which the information should be computed. Only relevant if differs from the dataset used to fit the model.  |
| <code>p</code>       | [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 information (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](#) 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

# Index

## \* data

abetaL, 5  
abetaW, 6  
blandAltmanL, 11  
blandAltmanW, 12  
bloodpressureL, 13  
calciumL, 13  
calciumW, 14  
ckdL, 15  
ckdW, 15  
gastricbypassL, 30  
gastricbypassW, 31  
ncgsL, 45  
ncgsW, 46  
potassiumRepeatedL, 49  
potassiumSingleL, 50  
potassiumSingleW, 50  
schoolL, 60  
swabsL, 69  
swabsW, 69  
vasscoresL, 71  
vasscoresW, 72  
vitaminL, 74  
vitaminW, 75

abetaL, 5  
abetaW, 6  
anova, 6  
anova.lmm, 38  
autoplot, 9  
autoplot.lmm, 49

baselineAdjustment, 10  
blandAltmanL, 11  
blandAltmanW, 12  
bloodpressureL, 13

calciumL, 13  
calciumW, 14  
ckdL, 15

ckdW, 15  
coef, 16, 36, 62, 74  
coef.lmm, 4, 7, 19  
confint, 18, 45  
confint.anova.lmm, 20, 21  
confint.mlmm, 21  
CS, 22  
CUSTOM, 23

dummy.coef.lmm, 25

emm\_basis.lmm (LMMstar2emmeans), 42  
estfun, 26  
estimate.lmm, 27

fitted.lmm, 29

gastricbypassL, 30  
gastricbypassW, 31  
getVarCov.lmm, 4, 32, 38

ID, 32  
iid.lmm, 33  
IND, 34  
information, 35

levels.lmm, 36, 38  
lmm, 4, 37, 44  
LMMstar-package, 4  
LMMstar.options, 38, 40  
LMMstar2emmeans, 42  
logLik, 42

mlmm, 43  
model.tables, 45  
model.tables.lmm, 38

ncgsL, 45  
ncgsW, 46

partialCor, 46

plot, 48  
plot.lmm, 38  
potassiumRepeatedL, 49  
potassiumSingleL, 50  
potassiumSingleW, 50  
predict.lmm, 38, 51  
profile, 53  
  
rbind.anova\_lmm, 55  
recover\_data.lmm (LMMstar2emmeans), 42  
residuals, 56  
residuals.lmm, 4, 38, 49  
  
sampleRem, 59  
schoolL, 60  
score, 61  
sigma, 62  
sigma.lmm, 32  
summarize, 63  
summary, 65  
summary.anova\_lmm, 8, 67, 68  
summary.lmm, 38  
summary.mlmm, 68  
swabsL, 69  
swabsW, 69  
  
terms.lmm, 70  
transformSummaryTable, 70  
  
UN, 71  
  
vasscoresL, 71  
vasscoresW, 72  
vcov, 73  
vitaminL, 74  
vitaminW, 75