Package ‘fullfact’

March 14, 2021

Type Package
Title Full Factorial Breeding Analysis
Version 1.5
Date 2021-03-07
Author Aimee Lee Houde [aut, cre]. Trevor Pitcher [aut]
Maintainer Aimee Lee Houde <aimee.lee.houde@gmail.com>
Depends R (>= 3.6)
Imports lme4, afex
VignetteBuilder knitr
Suggests knitr, rmarkdown
Description We facilitate the analysis of full factorial mating designs with mixed-effects models. There are now six vignettes containing detailed examples.
License GPL (>= 2)
NeedsCompilation no
Repository CRAN
Date/Publication 2021-03-14 15:50:18 UTC

R topics documented:

fullfact-package .................................................. 2
barMANA ......................................................... 6
boxMANA .......................................................... 7
buildBinary .......................................................... 8
buildMulti ............................................................. 9
chinook_bootL .................................................. 10
chinook_bootS .................................................. 11
chinook_jackL .................................................. 12
chinook_jackS .................................................. 13
chinook_length .................................................. 14
chinook_resampL .............................................. 15
chinook_resampS .............................................. 17
Full Factorial Breeding Analysis

Description

Full factorial breeding designs are useful for quantifying the amount of additive genetic, nonadditive genetic, and maternal variance that explain phenotypic traits. Such variance estimates are important for examining evolutionary potential. Traditionally, full factorial mating designs have been analyzed using a two-way analysis of variance, which may produce negative variance values and is not suited for unbalanced designs. Mixed-effects models do not produce negative variance values and are suited for unbalanced designs. However, extracting the variance components, calculating
significance values, and estimating confidence intervals and/or power values for the components are not straightforward using traditional analytic methods.

In this package we address these issues and facilitate the analysis of full factorial mating designs with mixed-effects models. The observed data functions extract the variance explained by random and fixed effects and provide their significance. We then calculate the additive genetic, nonadditive genetic, and maternal variance components explaining the phenotype. In particular, we integrate nonnormal error structures for estimating these components for nonnormal data types. The resampled data functions are used to produce bootstrap confidence intervals, which can then be plotted using a simple function. This package will facilitate the analyses of full factorial mating designs in R, especially for the analysis of binary, proportion, and/or count data types and for the ability to incorporate additional random and fixed effects and power analyses.

There are now six vignettes containing detailed examples: browseVignettes(package="fullfact").


Details

The DESCRIPTION file:

Package: fullfact
Type: Package
Title: Full Factorial Breeding Analysis
Version: 1.5
Date: 2021-03-07
Author: Aimee Lee Houde [aut, cre], Trevor Pitcher [aut]
Maintainer: Aimee Lee Houde <aimee.lee.houde@gmail.com>
Depends: R (>= 3.6)
Imports: lme4, afex
VignetteBuilder: knitr
Suggests: knitr, rmarkdown
Description: We facilitate the analysis of full factorial mating designs with mixed-effects models. There are now six vignettes containing detailed examples.
License: GPL (>=2)

Index of help topics:

JackGlmer Jackknife components for non-normal data
JackGlmer2 Jackknife components for non-normal data 2
JackGlmer3 Jackknife components for non-normal data 3
JackLmer Jackknife components for normal data
JackLmer2 Jackknife components for normal data 2
JackLmer3 Jackknife components for normal data 3
barMANA Bargraph of confidence intervals
boxMANA Boxplot of resampled results
buildBinary Convert to a binary data frame
buildMulti Convert to a multinomial frame
chinook_bootL Chinook salmon length, bootstrap calculations
Author(s)
Aimee Lee Houde [aut, cre], Trevor Pitcher [aut]
Maintainer: Aimee Lee Houde <aimee.lee.houde@gmail.com>

References
Traditional full factorial breeding design analysis:
Residual variance component values for generalized linear mixed-effects models:
Fixed effect variance component values for mixed-effects models:


Confidence intervals (bootstrap resampling, bias and acceleration correction, jackknife resampling):


Data sources:


Examples

data(chinook_length) #Chinook salmon offspring length

## Standard additive genetic, non-additive genetic, and maternal variance analysis

length_mod1<- observLmer(observ=chinook_length,dam="dam",sire="sire",response="length")

length_mod1

## Confidence intervals

## Bootstrap resampling of data: replicates within family

## Not run: resampRepli(dat=chinook_length,copy=c(3:8),family="family",replicate="repli", iter=1000)

## End(Not run)

#saves the files in working directory: one for each replicate and
#one final (combined) file "resamp_datR.csv"

## Import file

#length_datR<- read.csv("resamp_datR.csv")

data(chinook_resampL) #same as length_datR, 5 iterations

## Models for the resampled data: standard analysis

## Not run: length_rcomp<- resampLmer(resamp=length_datR,dam="dam",sire="sire", response="length",start=1,end=1000)

## End(Not run)

## 1. Uncorrected Bootstrap 95% confidence interval

#ciMANA(comp=length_rcomp)

data(chinook_bootL) #similar to length_rcomp, but 1,000 models

#ciMANA(comp=chinook_bootL)

## 2. Bias and accelerated corrected Bootstrap 95% confidence interval

## Jackknife resampling of data, delete-one: for acceleration estimate
## Not run: length_jack<- JackLmer(observ=chinook_length,dam="dam",sire="sire", response="length")
## End(Not run)

ciMANA(comp=length_rcomp,bias=c(0,0.7192,0.2030),accel=length_jack)
data(chinook_jackL) #similar to length_jack, but all observations

ciMANA(comp=chinook_bootL,bias=c(0,0.7192,0.2030),accel=chinook_jackL)

##3. Jackknife 95% confidence interval

ciJack(comp=length_jack,full=c(0,0.7192,0.2030,1.0404))
ciJack(comp=chinook_jackL,full=c(0,0.7192,0.2030,1.0404))

---

### Description

A simple bargraph function for confidence intervals of additive genetic, non-additive genetic, and maternal variance components. Also, plots the median for the bootstrap resampling method or mean of the pseudo-values for the jackknife resampling method.

### Usage

```r
ciMANA(ci_dat, type = "perc", bar_len = 0.1, ymax = NULL, ymin = NULL, yunit = NULL, leg = "topright", cex_ylab = 1, cex_yaxis = 1, cex_names = 1)
```

### Arguments

- `ci_dat`: Data frame of a confidence interval function.
- `type`: Default is "perc" for percentage values of variance components. Other option is "raw" for raw values of variance components.
- `bar_len`: Length of error bar in inches.
- `ymax`: Maximum value of the y-axis.
- `ymin`: Minimum value of the y-axis.
- `yunit`: Unit increment of the y-axis.
- `leg`: Position of the simple legend.
- `cex_ylab`: Magnification of the y-axis label.
- `cex_yaxis`: Magnification of the y-axis units.
- `cex_names`: Optional magnification of trait labels.
Details

Plots a bargraph with the median or mean as the top of the shaded bar and error bars covering the range of the confidence interval. Uses an object produced by any of the bootstrap resampling CI functions, i.e. \texttt{ciMANA}, \texttt{ciMANA2}, and \texttt{ciMANA3} or jackknife resampling functions, i.e. \texttt{ciJack}, \texttt{ciJack2}, and \texttt{ciJack3}. The median is plotted for bootstrap resampling and the mean of pseudo-value for jackknife resampling. Produces a simple legend. The function can plot several bar graphs grouped by \textit{label} to visualize several phenotypic traits.

Examples

```r
# Import jackknife resampling results
data(chinook_jackL) # Chinook salmon length
length_ci <- ciJack(comp=chinook_jackL,full=c(0,0.7192,0.2030,1.0404))
barMANA(ci_dat=length_ci) # default plot
barMANA(ci_dat=length_ci, bar_len=0.3, yunit=20, ymax=100, cex_ylab=1.3)

# Group length and survival together in the same plot
data(chinook_bootS) # Chinook salmon survival (bootstrap resampling)
length_ci <- ciMANA(comp=chinook_bootS,trait="length")
survival_ci <- ciMANA(comp=chinook_bootS,trait="survival")
colnames(length_ci$raw)[3]<- "median"; colnames(length_ci$percentage)[3]<- "median"
comb_bar <- list(raw=rbind(length_ci$raw,survival_ci$raw),
                  percentage=rbind(length_ci$percentage,survival_ci$percentage))
#
barMANA(ci_dat=comb_bar) # default plot
barMANA(ci_dat=comb_bar, bar_len=0.3, yunit=20, ymax=100, cex_ylab=1.3)
```

---

**boxMANA**

*Boxplot of resampled results*

**Description**

A simple boxplot function for bootstrap and jackknife resampled results of additive genetic, non-additive genetic, and maternal variance components.

**Usage**

```r
boxMANA(comp, type = "perc", ymax = NULL, ymin = NULL, yunit = NULL, leg = "topright",
         cex_ylab = 1, cex_yaxis = 1, cex_names = 1)
```

**Arguments**

- **comp**
  - Data frame of bootstrap or jackknife resampling results.

- **type**
  - Default is "perc" for percentage values of variance components. Other option is "raw" for raw values of variance components.

- **ymax**
  - Maximum value of the y-axis.
**buildBinary**

Convert to a binary data frame

**Description**

Assign a binary number (i.e. '0' or '1') to two columns containing the number of offspring. Copy information by the number of times equal to the number of offspring.

**Usage**

`buildBinary(dat, copy, one, zero)`

**Arguments**

* dat Data frame to convert.
* copy Column numbers to copy.
* one Column name of counts to assign a '1' value.
* zero Column name of counts to assign a '0' value.
**Details**

Replicate-level data should be converted to the individual-level to not underestimate phenotypic variance, which can influence genetic and maternal estimates (see Puurtinen et al. 2009).

**Value**

A converted data frame with a number of row matching the total number of individuals.

**References**


**See Also**

buildMulti

**Examples**

```r
data(chinook_survival)
chinook_survival2 <- buildBinary(dat=chinook_survival, copy=c(1:6,9), one="alive", zero="dead")
```

---

**Description**

Assign multiple numbers to multiple columns containing the number of offspring. Copy information by the number of times equal to the number of offspring.

**Usage**

`buildMulti(dat, copy, multi)`

**Arguments**

- `dat`: Data frame to convert.
- `copy`: Column numbers to copy.
- `multi`: A list containing the numbers to assign and matching column names, e.g. `list(c(2,0,1), c("two", "zero", "one")`.  

**Details**

Replicate-level data should be converted to the individual-level to not underestimate phenotypic variance, which can influence genetic and maternal estimates (see Puurtinen et al. 2009).

**Value**

A converted data frame with a number of row matching the total number of individuals.
References


See Also

buildBinary

Examples

data(chinook_survival)
chinook_survival$total<- chinook_survival$alive + chinook_survival$dead #create total column
chinook_survival3<- buildMulti(dat=chinook_survival, copy=c(1:6,9), multi=list(c(2,1,0), c("total","alive","dead"))}

---

chinook_bootL | Chinook salmon length, bootstrap calculations

Description

Bootstrap resampled Chinook salmon fork length (mm) at hatch with the amount of additive genetic, non-additive genetic, and maternal variance calculations.

Usage

data("chinook_bootL")

Format

A data frame with 1000 observations on the following 9 variables.

dam.sire, a numeric vector.
tray, a numeric vector.
sire, a numeric vector.
dam, a numeric vector.
Residual, a numeric vector.
Total, a numeric vector.
additive, a numeric vector.
maternal, a numeric vector.
nonadd, a numeric vector.

Details

Also includes the calculations for the amount of variance explained by position (tray), dam by sire, sire, dam, residual, and total.
Source

References

Examples
data(chinook_bootL)
## Extract bootstrap confidence interval
ciMANA(comp=chinook_bootL)

---

chinook_bootS Chinook salmon survival, bootstrap data

Description
Bootstrap resampled Chinook salmon binary survival to hatch (1 is alive, 0 is dead) with the amount of additive genetic, non-additive genetic, and maternal variance calculations.

Usage
data("chinook_bootS")

Format
A data frame with 1000 observations on the following 8 variables.
dam.sire, a numeric vector.
sire, a numeric vector.
dam, a numeric vector.
Residual, a numeric vector.
Total, a numeric vector.
additive, a numeric vector.
maternal, a numeric vector.
nonadd, a numeric vector.

Details
Also includes the calculations for the amount of variance explained by dam by sire, sire, dam, residual, and total.
chinook_jackL

Source


References


Examples

data(chinook_bootS)
## Extract bootstrap confidence interval
ciMANA(comp=chinook_bootS)

Description

Jackknife resampled Chinook salmon fork length (mm) at hatch with the amount of additive genetic, non-additive genetic, and maternal variance calculations. Jackknife resampling was leave-out-one.

Usage

data("chinook_jackL")

Format

A data frame with 1210 observations on the following 9 variables.

dam.sire, a numeric vector.
tray, a numeric vector.
sire, a numeric vector.
dam, a numeric vector.
Residual, a numeric vector.
Total, a numeric vector.
additive, a numeric vector.
onadd, a numeric vector.
maternal, a numeric vector.

Details

Also includes the calculations for the amount of variance explained by position (tray), dam by sire, sire, dam, residual, and total.
chinquook.jackS

Source

References

Examples

```r
data(chinook_jackL)
## Extract jackknifed confidence interval
ciJack(comp=chinook_jackL,full=c(0,0.7192,0.2030,1.0404))
```

---

**Description**

Jackknife resampled Chinook salmon survival with the amount of additive genetic, non-additive genetic, and maternal variance calculations. Jackknife resampling was leave-out-30.

**Usage**

data("chinook_jackS")

**Format**

A data frame with 1210 observations on the following 9 variables.

- **dam.sire**, a numeric vector.
- **sire**, a numeric vector.
- **dam**, a numeric vector.
- **Residual**, a numeric vector.
- **Total**, a numeric vector.
- **additive**, a numeric vector.
- **nonadd**, a numeric vector.
- **maternal**, a numeric vector.

**Details**

Also includes the calculations for the amount of variance explained by dam by sire, sire, dam, residual, and total.
Source

References

Examples
data(chinook_jackS)
## Extract jackknifed confidence interval
cijack(comp=chinook_jackS,full=c(0.6655,0.6692,0.6266,4.4166))

---

chinook_length  Chinook salmon length, raw data

Description
Raw Chinook salmon fork length (mm) at hatch for offspring produced using an 11 x 11 full factorial breeding design.

Usage
data("chinook_length")

Format
A data frame with 1210 observations on the following 8 variables.

family, a factor with levels: f1 f10 f100 f101 f102 f103 f104 f105 f106 f107 f108 f109 f11 f110 f111 f112 f113 f114 f115 f116 f117 f118 f119 f12 f120 f121 f122 f123 f124 f125 f126 f127 f128 f129 f13 f130 f131 f132 f133 f134 f135 f136 f137 f138 f139 f140 f141 f142 f143 f144 f145 f146 f147 f148 f149 f150 f151 f152 f153 f154 f155 f156 f157 f158 f159 f16 f160 f161 f162 f163 f164 f165 f166 f167 f168 f169 f17 f170 f171 f172 f173 f174 f175 f176 f177 f178 f179 f180 f181 f182 f183 f184 f185 f186 f187 f188 f189 f190 f191 f192 f193 f194 f195 f196 f197 f198 f199 
repli, a factor with levels: r1 r2
dam, a factor with levels: d1 d10 d11 d12 d13 d14 d15 d16 d17 d18 d19 d2 d20 d21 d22 d23 d24 d25 d26 d27 d28 d29 d3 d30 d31 d32 d33 d34 d35 d36 d37 d38 d39 d4 d40 d41 d42 d43 d44 d45 d46 d47 d48 d49 d5 d50 d51 d52 d53 d54 d55 d56 d57 d58 d59 d6 d60 d61 d62 d63 d64 d65 d66 d67 d68 d69 d7 d70 d71 d72 d73 d74 d75 d76 d77 d78 d79 d8 d80 d81 d82 d83 d84 d85 d86 d87 d88 d89 d9 d90 d91 d92 d93 d94 d95 d96 d97 d98 d99
sire, a factor with levels: s1 s10 s11 s12 s13 s14 s15 s16 s17 s18 s19 s2 s20 s21 s22 s23 s24 s25 s26 s27 s28 s29 s3 s30 s31 s32 s33 s34 s35 s36 s37 s38 s39 s4 s40 s41 s42 s43 s44 s45 s46 s47 s48 s49 s5 s50 s51 s52 s53 s54 s55 s56 s57 s58 s59 s6 s60 s61 s62 s63 s64 s65 s66 s67 s68 s69 s7 s70 s71 s72 s73 s74 s75 s76 s77 s78 s79 s8 s80 s81 s82 s83 s84 s85 s86 s87 s88 s89 s9 s90 s91 s92 s93 s94 s95 s96 s97 s98 s99 
tray, a factor with levels: t1 t10 t11 t12 t13 t14 t15 t16 t17 t18 t19 t2 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t3 t30 t31 t32 t33 t34 t35 t36 t37 t38 t39 t4 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t5 t50 t51 t52 t53 t54 t55 t56 t57 t58 t59 t6 t60 t61 t62 t63 t64 t65 t66 t67 t68 t69 t7 t70 t71 t72 t73 t74 t75 t76 t77 t78 t79 t8 t80 t81 t82 t83 t84 t85 t86 t87 t88 t89 t9 t90 t91 t92 t93 t94 t95 t96 t97 t98 t99
cell, a factor with levels: 1A 1B 1C 1D 2A 2B 2C 2D 3A 3B 3C 3D 4A 4B 4C 4D
length, a numeric vector.
egg_size, a numeric vector.
**Details**

Also includes family identity, family replicate, incubator position (tray and cell), and average female egg size (mm) information.

**Source**


**References**


**Examples**

```r
data(chinook_length)
## Standard additive genetic, non-additive genetic, and maternal variance analysis
length_mod1<- observLmer(observ=chinook_length,dam="dam",sire="sire",response="length")
length_mod1
```

**Description**

Bootstrap resampled Chinook salmon fork length (mm) at hatch. Number of iterations was 5.

**Usage**

```r
data("chinook_resampL")
```

**Format**

A data frame with 1210 observations on the following 30 variables.

dam1, a numeric vector
sire1, a numeric vector
tray1, a numeric vector
cell1, a numeric vector
length1, a numeric vector
egg_size1, a numeric vector
dam2, a numeric vector
sire2, a numeric vector
tray2, a numeric vector
cell2, a numeric vector
length2, a numeric vector
egg_size2, a numeric vector
dam3, a numeric vector
sire3, a numeric vector
tray3, a numeric vector
cell3, a numeric vector
length3, a numeric vector
egg_size3, a numeric vector
dam4, a numeric vector
sire4, a numeric vector
tray4, a numeric vector
cell4, a numeric vector
length4, a numeric vector
egg_size4, a numeric vector
dam5, a numeric vector
sire5, a numeric vector
tray5, a numeric vector
cell5, a numeric vector
length5, a numeric vector
egg_size5, a numeric vector

Source

Examples

data(chinook_resampL)
#the five models
length_rcompl<- resampLmer(resamp=chinook_resampL,dam="dam",sire="sire",response="length", start=1,end=5)  #full analysis should use 1,000 models
Chinook salmon survival, bootstrap resampled

Description
Bootstrap resampled Chinook salmon binary survival to hatch (1 is alive, 0 is dead). Number of iterations was 5.

Usage
data("chinook_resampS")

Format
A data frame with 36300 observations on the following 30 variables.
status1, a numeric vector
  dam1, a numeric vector
  sire1, a numeric vector
  tray1, a numeric vector
  cell1, a numeric vector
  egg_size1, a numeric vector
status2, a numeric vector
dam2, a numeric vector
sire2, a numeric vector
tray2, a numeric vector
cell2, a numeric vector
egg_size2, a numeric vector
status3, a numeric vector
dam3, a numeric vector
sire3, a numeric vector
tray3, a numeric vector
cell3, a numeric vector
egg_size3, a numeric vector
status4, a numeric vector
dam4, a numeric vector
sire4, a numeric vector
tray4, a numeric vector
cell4, a numeric vector
egg_size4, a numeric vector
status5, a numeric vector  
dam5, a numeric vector  
sire5, a numeric vector  
tray5, a numeric vector  
cell5, a numeric vector  
egg_size5, a numeric vector

Source

Examples
data(chinook_resampS)  
## Not run: survival_rcomp<- resampGlmer(resamp=chinook_resampS,dam="dam",sire="sire",response="status",fam_link=binomial(link="logit"),start=1,end=1000)  
## End(Not run)

chinook_survival  
Chinook salmon survival, raw data

Description
Raw Chinook salmon numbers alive and dead to hatching of offspring produced using an 11 x 11 full factorial breeding design.

Usage
data("chinook_survival")

Format
A data frame with 242 observations on the following 9 variables.

family, a factor with levels: f1 f10 f100 f101 f102 f103 f104 f105 f106 f107 f108 f109 f11 f110 f111 f112 f113 f114 f115 f116 f117 f118 f119 f12 f120 f121 f13 f14 f15 f16 f17 f18 f19 f2 f20 f21 f22 f23 f24 f25 f26 f27 f28 f29 f3 f30 f31 f32 f33 f34 f35 f36 f37 f38 f39 f4 f40 f41 f42 f43 f44 f45 f46 f47 f48 f49 f5 f50 f51 f52 f53 f54 f55 f56 f57 f58 f59 f6 f60 f61 f62 f63 f64 f65 f66 f67 f68 f69 f7 f70 f71 f72 f73 f74 f75 f76 f77 f78 f79 f8 f80 f81 f82 f83 f84 f85 f86 f87 f88 f89 f9 f90 f91 f92 f93 f94 f95 f96 f97 f98 f99  
repli, a factor with levels: r1 r2  
dam, a factor with levels: d1 d10 d11 d2 d3 d4 d5 d6 d7 d8 d9  
sire, a factor with levels: s1 s10 s11 s2 s3 s4 s5 s6 s7 s8 s9
ciJack

Jackknife confidence intervals

description

Extracts jackknife confidence intervals for additive genetic, non-additive genetic, and maternal variance components.

Usage

ciJack(comp, full, level = 95, rnd_r = 3, rnd_p = 1, trait = NULL)
Arguments

comp  Data frame of jackknife resampling results.
full  A vector of raw observed additive, non-additive, maternal, and total variance component values for from the full observed data set, i.e. c(additive, non-additive, maternal, total).
level Confidence level, as a percentage. Default is 95.
rnd_r Number of decimal places to round the confidence interval of raw values.
rnd_p Number of decimal places to round the confidence interval of percentage values.
trait Optional label for the phenotypic trait.

Details

Used for jackknife resampling results produced using JackLmer for normal data or JackGlmer for non-normal data. Jackknife confidence intervals, using pseudo-values are described by Efron and Tibshirani (1993). The standard errors are calculated from the pseudo-values and the Student’s t distribution is used to provide the lower and upper confidence values. For delete-d jackknife resampling, M degrees of freedom are used for producing the confidence interval (Martin et al. 2004): $M = N / d$, where $N$ is the total number of observations and $d$ is the number of deleted observations. That is, $M$ is the number of row in the jackknife resampling results. Large values of $M$, such as 1,000, can translate to the delete-d jackknife resampling method approaching bootstrap resampling expectations (Efron & Tibshirani 1993).

Value

Prints a data frame containing the lower, median, and upper values of the jackknife confidence interval for additive genetic, non-additive genetic, and maternal variance components. Values are presented as raw and percentages of the total variance value within each row.

References


See Also
ciJack2, ciJack3

Examples

data(chinook_jackL) #Chinook salmon offspring length, delete-one jackknife
ciJack(chinook_jackL,c(0,0.7192,0.2030,1.0404))
Description

Extracts jackknife confidence intervals for additive genetic, non-additive genetic, and maternal variance components. Also extracts intervals for optional position and block variance components.

Usage

`ciJack2(comp, full, level = 95, rnd_r = 3, rnd_p = 1, position = NULL, block = NULL, trait = NULL)`

Arguments

- `comp`: Data frame of jackknife resampling results.
- `full`: A vector of raw observed additive, non-additive, maternal, and total variance component values for from the full observed data set, i.e. c(additive, non-additive, maternal, total, position/block). If there is a position and a block c(..., total, position, block).
- `level`: Confidence level, as a percentage. Default is 95.
- `rnd_r`: Number of decimal places to round the confidence interval of raw values.
- `rnd_p`: Number of decimal places to round the confidence interval of percentage values.
- `position`: Optional column name containing position factor information.
- `block`: Optional column name containing block factor information.
- `trait`: Optional label for the phenotypic trait.

Details

Used for jackknife resampling results produced using `JackLmer2` for normal data or `JackGlmer2` for non-normal data. Jackknife confidence intervals, using pseudo-values are described by Efron and Tibshirani (1993). The standard errors are calculated from the pseudo-values and the Student’s $t$ distribution is used to provide the lower and upper confidence values. For delete-$d$ jackknife resampling, $M$ degrees of freedom are used for producing the confidence interval (Martin et al. 2004): $M = N / d$, where $N$ is the total number of observations and $d$ is the number of deleted observations. That is, $M$ is the number of row in the jackknife resampling results. Large values of $M$, such as 1,000, can translate to the delete-$d$ jackknife resampling method approaching bootstrap resampling expectations (Efron & Tibshirani 1993).

Value

Prints a data frame containing the lower, median, and upper values of the jackknife confidence interval for additive genetic, non-additive genetic, maternal variance components, and optional position and/or block variance components. Values are presented as raw and percentages of the total variance value within each row.
References


See Also

ciJack, ciJack3

Examples

data(chinook_jackL) #Chinook salmon offspring length, delete-one jackknife
ciJack2(chinook_jackL,position="tray",c(0,0.7192,0.2030,1.0404,0.1077))

---

ciJack3  
Jackknife confidence intervals 3

Description

Extracts jackknife confidence intervals for additive genetic, non-additive genetic, and maternal variance components. Also extracts intervals for additional fixed and/or random effects.

Usage

ciJack3(comp, full, remain = NULL, level = 95, rnd_r = 3, rnd_p = 1, trait = NULL)

Arguments

- **comp**: Data frame of jackknife resampling results
- **full**: A vector of raw observed additive, non-additive, maternal, and total variance component values for from the full observed data set, i.e. `c(additive, non-additive, maternal, total)`. Followed by any other components in the order of the vector `remain`, i.e. `c(additive, non-additive, maternal, total, component1, component2, etc.)`.
- **remain**: Vector of column names for additional effects
- **level**: Confidence level, as a percentage. Default is 95.
- **rnd_r**: Number of decimal places to round the confidence interval of raw values.
- **rnd_p**: Number of decimal places to round the confidence interval of percentage values.
- **trait**: Optional label for the phenotypic trait.
ciMANA

Details

Used for jackknife resampling results produced using JackLmer3 for normal data or JackGlmer3 for non-normal data. Jackknife confidence intervals, using pseudo-values are described by Efron and Tibshirani (1993). The standard errors are calculated from the pseudo-values and the Student’s t distribution is used to provide the lower and upper confidence values. For delete-d jackknife resampling, M degrees of freedom are used for producing the confidence interval (Martin et al. 2004): \( M = N / d \), where \( N \) is the total number of observations and \( d \) is the number of deleted observations. That is, \( M \) is the number of row in the jackknife resampling results. Large values of \( M \), such as 1,000, can translate to the delete-d jackknife resampling method approaching bootstrap resampling expectations (Efron & Tibshirani 1993).

Value

Prints a data frame containing the lower, median, and upper values of the jackknife confidence interval for additive genetic, non-additive genetic, maternal variance components, and any additional fixed effect and/or random effect variance components. Values are presented as raw and percentages of the total variance value within each row.

References


See Also

ciJack, ciJack2

Examples

data(chinook_jackL) #Chinook salmon offspring length, delete-one jackknife

ciJack3(chinook_jackL,remain=c("tray","Residual"),c(0,0.7192,0.2030,1.0404,0.1077,0.5499))

---

Description

Extracts bootstrap-t confidence intervals for additive genetic, non-additive genetic, and maternal variance components.

Usage

ciMANA(comp, level = 95, rnd_r = 3, rnd_p = 1, bias = NULL, accel = NULL, trait = NULL)
Arguments

- **comp**: Data frame of bootstrap resampling results.
- **level**: Confidence level, as a percentage. Default is 95.
- **rnd_r**: Number of decimal places to round the confidence interval of raw values.
- **rnd_p**: Number of decimal places to round the confidence interval of percentage values.
- **bias**: Optional vector of raw observed additive, non-additive, and maternal, variance component values for bias correction, i.e. c(additive, non-additive, maternal).
- **accel**: Optional data frame of jackknifed data model results for acceleration correction.
- **trait**: Optional label for the phenotypic trait.

Details

Used for bootstrap resampling results produced using `resampLmer` for normal data or `resampGlmer` for non-normal data. Bootstrap-t confidence intervals, including bias and acceleration correction methods are described by Efron and Tibshirani (1993). Jackknife data model results for acceleration correction can be produced using `JackLmer`, for normal data or `JackGlmer` for non-normal data. The 'bias fail' warning is if the bias calculation is Inf or -Inf, e.g. bias contains a zero value, so the uncorrected confidence interval is displayed.

Value

Prints a data frame containing the lower, median, and upper values of the bootstrap-t confidence interval for additive genetic, non-additive genetic, and maternal variance components. Values are presented as raw and percentages of the total variance value within each row.

References


See Also

ciMANA2, ciMANA3

Examples

```r
# Import bootstrap resampled data model results
data(chinook_bootL) # Chinook salmon offspring length

# Extract un-corrected confidence interval
ciMANA(comp=chinook_bootL)

# Extract bias corrected confidence interval
ciMANA(comp=chinook_bootL,bias=c(0,0.7192,0.2030))
# see details for 'bias' fail

# Extract bias and accelerated corrected confidence interval
# Import jackknife resampled data model results
```
ciMANA2

```r
data(chinook_jackL)
# ciMANA(comp=chinook_bootL,bias=c(0,0.7192,0.2030),accel=chinook_jackL)
#see details for 'bias' fail
```

---

### ciMANA2

**Bootstrap confidence intervals 2**

#### Description

Extracts bootstrap confidence intervals for additive genetic, non-additive genetic, and maternal variance components. Also extracts intervals for optional position and block variance components.

#### Usage

```r
ciMANA2(comp, level = 95, rnd_r = 3, rnd_p = 1, position = NULL, block = NULL, bias = NULL, accel = NULL, trait = NULL)
```

#### Arguments

- `comp` Data frame of bootstrap resampling results.
- `level` Confidence level, as a percentage. Default is 95.
- `rnd_r` Number of decimal places to round the confidence interval of raw values.
- `rnd_p` Number of decimal places to round the confidence interval of percentage values.
- `position` Optional column name containing position factor information.
- `block` Optional column name containing block factor information.
- `bias` Optional vector of raw observed additive, non-additive, maternal, position and/or block variance component values for bias correction, i.e. c(additive, non-additive, maternal, position/block). If there is a position and a block c(..., maternal, position, block).
- `accel` Optional data frame of jackknifed data model results for acceleration correction.
- `trait` Optional label for the phenotypic trait.

#### Details

Used for bootstrap resampling results produced using `resampLmer2` for normal data or `resampGlmer2` for non-normal data. Bootstrap confidence intervals, including bias and acceleration correction methods are described by Efron and Tibshirani (1993). Jackknife data model results for acceleration correction can be produced using `JackLmer2`, for normal data or `JackGlmer2` for non-normal data. The ‘bias fail’ warning is if the bias calculation is Inf or -Inf, e.g. `bias` contains a zero value, so the uncorrected confidence interval is displayed.
Value

Prints a data frame containing the lower, median, and upper values of the bootstrap-t confidence interval for additive genetic, non-additive genetic, maternal, and optional position and/or block variance components. Values are presented as raw and percentages of the total variance value within each row.

References


See Also

ciMANA, ciMANA3

Examples

# Import bootstrap resampled data model results
data(chinook_bootL) # Chinook salmon offspring length

# Extract un-corrected confidence interval
ciMANA2(comp=chinook_bootL, position="tray")

# Extract bias corrected confidence interval
ciMANA2(comp=chinook_bootL, position="tray", bias=c(0, 0.7192, 0.2030, 0.1077))
# see details for 'bias' fail

# Extract bias and accelerated corrected confidence interval
# Import jackknife resampled data model results
data(chinook_jackL)
#
ciMANA2(comp=chinook_bootL, position="tray", bias=c(0, 0.7192, 0.2030, 0.1077), accel=chinook_jackL)
# see details for 'bias' fail

---

**ciMANA3**

Bootstrapping confidence intervals 3

---

Description

Extracts bootstrap-t confidence intervals for additive genetic, non-additive genetic, and maternal variance components. Also extracts intervals for additional fixed and/or random effects.

Usage

ciMANA3(comp, level = 95, rnd_r = 3, rnd_p = 1, bias = NULL, accel = NULL, remain = NULL, trait = NULL)
ciMANA3

Arguments

- **comp**: Data frame of bootstrap resampling results.
- **level**: Confidence level, as a percentage. Default is 95.
- **rnd_r**: Number of decimal places to round the confidence interval of raw values.
- **rnd_p**: Number of decimal places to round the confidence interval of percentage values.
- **bias**: Optional vector of raw observed additive, non-additive, and maternal variance components for bias correction. Followed by any other components in the order of the vector `remain`, i.e. c(additive, non-additive, maternal, component1, component2, etc.).
- **accel**: Optional data frame of jackknifed data model results for acceleration correction.
- **remain**: Vector of column names for additional effects.
- **trait**: Optional label for the phenotypic trait.

Details

Used for bootstrap resampling results produced using `resampLmer3` for normal data or `resampGlmer3` for non-normal data. Bootstrap-\(t\) confidence intervals, including bias and acceleration correction methods are described by Efron and Tibshirani (1993). Jackknife data model results for acceleration correction can be produced using `JackLmer3`, for normal data or `JackGlmer3` for non-normal data. The 'bias fail' warning is if the bias calculation is Inf or -Inf, e.g. bias contains a zero value, so the uncorrected confidence interval is displayed.

Value

Prints a data frame containing the lower, median, and upper values of the bootstrap-\(t\) confidence interval for additive genetic, non-additive genetic, maternal, and any additional fixed effect and/or random effect variance components. Values are presented as raw and percentages of the total variance value within each row.

References


See Also

- ciMANA, ciMANA2

Examples

```r
# Import bootstrap resampled data model results
data(chinook_bootL) # Chinook salmon offspring length

# Extract un-corrected confidence interval
ciMANA3(comp=chinook_bootL, remain=c("tray", "Residual"))

# Extract bias corrected confidence interval
ciMANA3(comp=chinook_bootL, remain=c("tray", "Residual"),
```
bias=c(0,0.7192,0.2030,0.1077,0.5499))
#see details for 'bias' fail

#Extract bias and accelerated corrected confidence interval
#Import jackknife resampled data model results
data(chinook_jackL)

#ciMANA3(comp=chinook_bootL,remain=c("tray","Residual"),
bias=c(0,0.7192,0.2030,0.1077,0.5499),accel=chinook_jackL)

---

JackGlmer

*Jackknife components for non-normal data*

**Description**

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the *lmer* function of the *lme4* package. Model random effects are dam, sire, and dam by sire.

**Usage**

`JackGlmer(observ, dam, sire, response, fam_link, quasi = F, size = 1, first = NULL)`

**Arguments**

- **observ**
  - Data frame of observed data.
- **dam**
  - Column name containing dam (female) parent identity information.
- **sire**
  - Column name containing sire (male) parent identity information.
- **response**
  - Column name containing the offspring (response) phenotype values.
- **fam_link**
  - The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
- **quasi**
  - Incorporate overdispersion or quasi-error structure.
- **size**
  - Default is 1 for delete-one jackknife resampling. If `size > 1`, delete-`d` jackknife resampling occurs removing a block of size `d` equal to `size`.
- **first**
  - Number of initial sub-samples to run. Useful for examining if there is variation among sub-samples before jackknife resampling the entire data set. There can be little variation for delete-one jackknife resampling with large data sets, and delete-`d` jackknife resampling should be considered.

**Details**

Uses delete-one jackknife resampling (Efron & Tibshirani 1993, p. 141-145). For the option of delete-`d` jackknife resampling, the rows of the observed data frame are shuffled and a block of observations of size `d` is deleted sequentially. Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam
by sire variance components. The residual variance component for the `fam_links` are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

**Value**

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. The number of rows in the data frame matches the total number of observations \( N \) for delete-one jackknife resampling or \( M \) groups for delete-\( d \) jackknife resampling to the lowest integer. Each row represents a deleted single observation or deleted \( d \) observations group.

**Note**

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

**References**


**See Also**

`JackGlmer2`, `JackGlmer3`

**Examples**

data(chinook_survival) # Chinook salmon offspring survival
## Convert replicate-level recorded data to individual-level (binary) data
chinook_survival2 <- buildBinary(dat=chinook_survival,copy=c(1:6,9),one="alive",zero="dead")
#Delete-one
## Not run: survival_jack1 <- JackGlmer(observ = chinook_survival2, dam = "dam", sire = "sire", response = "status", fam_link = binomial(link = "logit"))
## End(Not run)

#Delete-d, d=30
## Not run: survival_jack1.2 <- JackGlmer(observ = chinook_survival2, dam = "dam", sire = "sire", response = "status", fam_link = binomial(link = "logit"), size = 30)
## End(Not run)

---

JackGlmer2  

**Jackknife components for non-normal data 2**

---

**Description**

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the `lmer` function of the `lme4` package. Model random effects are dam, sire, and dam by sire. Options to include one random position and/or one random block effect(s).

**Usage**

JackGlmer2(observ, dam, sire, response, fam_link, position = NULL, block = NULL, quasi = F, size = 1, first = NULL)

**Arguments**

- **observ**  
  Data frame of observed data.

- **dam**  
  Column name containing dam (female) parent identity information.

- **sire**  
  Column name containing sire (male) parent identity information.

- **response**  
  Column name containing the offspring (response) phenotype values.

- **fam_link**  
  The family and link in family(link) format. Supported options are binomial(link = "logit"), binomial(link = "probit"), poisson(link = "log"), and poisson(link = "sqrt").

- **position**  
  Optional column name containing position factor information.

- **block**  
  Optional column name containing block factor information.

- **quasi**  
  Incorporate overdispersion or quasi-error structure.

- **size**  
  Default is 1 for delete-one jackknife resampling. If size > 1, delete-d jackknife resampling occurs removing a block d equal to size.

- **first**  
  Number of initial sub-samples to run. Useful for examining if there is variation among sub-samples before jackknife resampling the entire data set. There can be little variation for delete-one jackknife resampling with large data sets, and delete-d jackknife resampling should be considered.
Details

Uses delete-one jackknife resampling (Efron & Tibshirani 1993, p. 141-145). For the option of delete-d jackknife resampling, the rows of the observed data frame are shuffled and a block of observations of size d is deleted sequentially. Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam by sire variance components. Extracts optional position and block variance components. The residual variance component for the fam_links are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for the options of position and/or block. The number of rows in the data frame matches the total number of observations (N) for delete-one jackknife resampling or M groups for delete-d jackknife resampling to the lowest integer. Each row represents a deleted single observation or deleted d observations group.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References


See Also

JackGlmer, JackGlmer3

Examples

data(chinook_survival) #Chinook salmon offspring survival

## Convert replicate-level recorded data to individual-level (binary) data

chinook_survival2<- buildBinary(dat=chinook_survival,copy=c(1:6,9),one="alive",zero="dead")

#Delete-one

## Not run: survival_jack2<- JackGlmer2(observ=chinook_survival2,dam="dam",sire="sire",


response="status",fam_link=binomial(link="logit"),position="tray")

## End(Not run)

#Delete-d, d=30

## Not run: survival_jack2.2<- JackGlmer2(observ=chinook_survival2,dam="dam",sire="sire",


response="status",fam_link=binomial(link="logit"),position="tray",size=30)

## End(Not run)

JackGlmer3

Jackknife components for non-normal data 3

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear
mixed-effect model using the lmer function of the lme4 package. Model random effects are dam,
sire, dam by sire, and any additional fixed and/or random effects.

Usage

JackGlmer3(observ, dam, sire, response, fam_link, remain, quasi = F, size = 1, first = NULL)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>observ</td>
<td>Data frame of observed data.</td>
</tr>
<tr>
<td>dam</td>
<td>Column name containing dam (female) parent identity information.</td>
</tr>
<tr>
<td>sire</td>
<td>Column name containing sire (male) parent identity information.</td>
</tr>
<tr>
<td>response</td>
<td>Column name containing the offspring (response) phenotype values.</td>
</tr>
</tbody>
</table>
| fam_link | The family and link in family(link) format. Supported options are binomial(link="logit"),


binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt"). |
| remain   | Remaining formula using lme4 package formula. |
| quasi    | Incorporate overdispersion or quasi-error structure. |
| size     | Default is 1 for delete-one jackknife resampling. If size > 1, delete-d jackknife resampling occurs removing a block d equal to size. |
Number of initial sub-samples to run. Useful for examining if there is variation among sub-samples before jackknife resampling the entire data set. There can be little variation for delete-one jackknife resampling with large data sets, and delete-$d$ jackknife resampling should be considered.

Details

Uses delete-one jackknife resampling (Efron & Tibshirani 1993, p. 141-145). For the option of delete-$d$ jackknife resampling, the rows of the observed data frame are shuffled and a block of observations of size $d$ is deleted sequentially. Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam by sire, and dam by sire variance components. Extracts any additional fixed effect and random effect variance components. The fixed-effect variance component is as a single group using the method described by Nakagawa and Schielzeth (2013). The residual variance component for the `fam_links` are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for remaining formula components. The number of rows in the data frame matches the total number of observations ($N$) for delete-one jackknife resampling or $M$ groups for delete-$d$ jackknife resampling to the lowest integer. Each row represents a deleted single observation or deleted $d$ observations group.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References


See Also

JackGlmer, JackGlmer2

Examples

data(chinook_survival) #Chinook salmon offspring survival
## Convert replicate-level recorded data to individual-level (binary) data
chinook_survival2<- buildBinary(dat=chinook_survival,copy=c(1:6,9),one="alive",zero="dead")

#Delete-one
## Not run: survival_jack3<- JackGlmer3(observ=chinook_survival2,dam="dam",sire="sire",
response="status",fam_link=binomial(link="logit"),remain="egg_size + (1|tray)")
## End(Not run)

#Delete-d, d=30
## Not run: survival_jack3.2<- JackGlmer3(observ=chinook_survival2,dam="dam",sire="sire",
response="status",fam_link=binomial(link="logit"),remain="egg_size + (1|tray)",size=30)
## End(Not run)

JackLmer

Jackknife components for normal data

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the lmer function of the lme4 package. Model random effects are dam, sire, and dam by sire.

Usage

JackLmer(observ, dam, sire, response, ml = F, size = 1, first = NULL)

Arguments

- observ: Data frame of observed data.
- dam: Column name containing dam (female) parent identity information.
- sire: Column name containing sire (male) parent identity information.
- response: Column name containing the offspring (response) phenotype values.
- ml: Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.
size

Default is 1 for delete-one jackknife resampling. If size > 1, delete-\(d\) jackknife resampling occurs removing a block \(d\) equal to size.

first

Number of initial sub-samples to run. Useful for examing if there is variation among sub-samples before jackknife resampling the entire data set. There can be little variation for delete-one jackknife resampling with large data sets, and delete-\(d\) jackknife resampling should be considered.

Details

Uses delete-one jackknife resampling (Efron & Tibshirani 1993, p. 141-145). For the option of delete-\(d\) jackknife resampling, the rows of the observed data frame are shuffled and a block of observations of size \(d\) is deleted sequentially. Extracts the dam, sire, dam by sire, and residual variance components. Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. The number of rows in the data frame matches the total number of observations (\(N\)) for delete-one jackknife resampling or \(M\) groups for delete-\(d\) jackknife resampling to the lowest integer. Each row represents a deleted single observation or deleted \(d\) observations group.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

JackLmer2, JackLmer3
Examples

```r
data(chinook_length) #Chinook salmon offspring length

#Delete-one
#length_jack1 <- JackLmer(observ=chinook_length,dam="dam",sire="sire",response="length")
length_jack1 <- JackLmer(observ=chinook_length,dam="dam",sire="sire",response="length", first=2) #first 2

#Delete-d, d=5
#length_jackD <- JackLmer(observ=chinook_length,dam="dam",sire="sire",response="length", size=5)
length_jackD <- JackLmer(observ=chinook_length,dam="dam",sire="sire",response="length", size=5,first=2) #first 2
```

---

**JackLmer2**

*Jackknife components for normal data 2*

**Description**

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the `lmer` function of the `lme4` package. Model random effects are dam, sire, and dam by sire. Options to include one random position and/or one random block effect(s).

**Usage**

```r
JackLmer2(observ, dam, sire, response, position = NULL, block = NULL, ml = F, size = 1, first = NULL)
```

**Arguments**

- `observ`: Data frame of observed data.
- `dam`: Column name containing dam (female) parent identity information.
- `sire`: Column name containing sire (male) parent identity information.
- `response`: Column name containing the offspring (response) phenotype values.
- `position`: Optional column name containing position factor information.
- `block`: Optional column name containing block factor information.
- `ml`: Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.
- `size`: Default is 1 for delete-one jackknife resampling. If `size` > 1, delete-`d` jackknife resampling occurs removing a block d equal to `size`.
- `first`: Number of initial sub-samples to run. Useful for examining if there is variation among sub-samples before jackknife resampling the entire data set. There can be little variation for delete-one jackknife resampling with large data sets, and delete-`d` jackknife resampling should be considered.
Details

Uses delete-one jackknife resampling (Efron & Tibshirani 1993, p. 141-145). For the option of delete-\(d\) jackknife resampling, the rows of the observed data frame are shuffled and a block of observations of size \(d\) is deleted sequentially. Extracts the dam, sire, dam, dam by sire, and residual variance components. Extracts optional position and block variance components. Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for the options of position and/or block. The number of rows in the data frame matches the total number of observations \(N\) for delete-one jackknife resampling or \(M\) groups for delete-\(d\) jackknife resampling to the lowest integer. Each row represents a deleted single observation or deleted \(d\) observations group.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

JackLmer, JackLmer3

Examples

data(chinook_length) #Chinook salmon offspring length

#Delete-one
#length_jack2<- JackLmer2(observ=chinook_length,dam="dam",sire="sire",response="length", position="tray")
length_jack2<- JackLmer2(observ=chinook_length,dam="dam",sire="sire",response="length", position="tray",first=2) #first 2
#Delete-d, d=5
#length_jack2.2 <- JackLmer2(observ=chinook_length,dam="dam",sire="sire",response="length", 
#position="tray",size=5)
length_jack2.2 <- JackLmer2(observ=chinook_length,dam="dam",sire="sire",response="length", 
position="tray",size=5,first=2) #first 2

## JackLmer3

### Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the `lmer` function of the `lme4` package. Model random effects are dam, sire, dam by sire, and any additional fixed and/or random effects.

#### Usage

```r
JackLmer3(observ, dam, sire, response, remain, ml = F, size = 1, first = NULL)
```

#### Arguments

- `observ`: Data frame of observed data
- `dam`: Column name containing dam (female) parent identity information.
- `sire`: Column name containing sire (male) parent identity information.
- `response`: Column name containing the offspring (response) phenotype values.
- `remain`: Remaining formula using `lme4` package format.
- `ml`: Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.
- `size`: Default is 1 for delete-one jackknife resampling. If `size > 1`, delete-d jackknife resampling occurs removing a block d equal to `size`.
- `first`: Number of initial sub-samples to run. Useful for examining if there is variation among sub-samples before jackknife resampling the entire data set. There can be little variation for delete-one jackknife resampling with large data sets, and delete-d jackknife resampling should be considered.

#### Details

Uses delete-one jackknife resampling (Efron & Tibshirani 1993, p. 141-145). For the option of delete-d jackknife resampling, the rows of the observed data frame are shuffled and a block of observations of size d is deleted sequentially. Extracts the dam, sire, dam, dam by sire, and residual variance components. Extracts any additional fixed effect and random effect variance components. The fixed-effect variance component is as a single group using the method described by Nakagawa and Schielzeth (2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).
Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for remaining formula components. The number of rows in the data frame matches the total number of observations \(N\) for delete-one jackknife resampling or \(M\) groups for delete-\(d\) jackknife resampling to the lowest integer. Each row represents a deleted single observation or deleted \(d\) observations group.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

JackLmer, JackLmer2

Examples

data(chinook_length) #Chinook salmon offspring length

#Delete-one
length_jack3<- JackLmer3(observ=chinook_length,dam="dam",sire="sire",response="length", remain="egg_size + (1|tray)")

#Delete-d, d=5
length_jack3.2<- JackLmer3(observ=chinook_length,dam="dam",sire="sire",response="length", remain="egg_size + (1|tray)",size=5)

#first 2
observGlmer

Variance components for non-normal data

Description
Extracts additive genetic, non-additive genetic, and maternal variance components from a generalized linear mixed-effect model using the `glmer` function of the `lme4` package. Model random effects are dam, sire, and dam by sire.

Usage

```r
observGlmer(observ, dam, sire, response, fam_link, quasi = F)
```

Arguments

- `observ` Data frame of observed data.
- `dam` Column name containing dam (female) parent identity information.
- `sire` Column name containing sire (male) parent identity information.
- `response` Column name containing the offspring (response) phenotype values.
- `fam_link` The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
- `quasi` Incorporate overdispersion or quasi-error structure.

Details

Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam by sire variance components. The residual variance component for the `fam_links` are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value

A list object containing the raw variance components, the variance components as a percentage of the total variance component. Also, contains the difference in AIC and BIC, and likelihood ratio test Chi-square and p-value for all random effects.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.
References


See Also

observGlmer2, observGlmer3

Examples

data(chinook_survival) #Chinook salmon offspring survival
## Convert replicate-level recorded data to individual-level (binary) data
chinook_survival2<- buildBinary(dat=chinook_survival,copy=c(2:6,9),one="alive",zero="dead")
## Not run: survival_mod1<- observGlmer(observ=chinook_survival2,dam="dam",sire="sire",
response="status",fam_link=binomial(link="logit")) #a few minutes
survival_mod1
## End(Not run)

observGlmer2

Variance components for non-normal data 2

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a generalized linear mixed-effect model using the glmer function of the lme4 package. Model random effects are dam, sire, and dam by sire. Options to include one random position and/or one random block effect(s).

Usage

observGlmer2(observ, dam, sire, response, fam_link, position = NULL, block = NULL, quasi = F)
Arguments

- **observ**: Data frame of observed data.
- **dam**: Column name containing dam (female) parent identity information.
- **sire**: Column name containing sire (male) parent identity information.
- **response**: Column name containing the offspring (response) phenotype values.
- **fam_link**: The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
- **position**: Optional column name containing position factor information.
- **block**: Optional column name containing block factor information.
- **quasi**: Incorporate overdispersion or quasi-error structure.

Details

Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam by sire variance components. Extracts optional position and block variance components. The residual variance component for the **fam_links** are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value

A list object containing the raw variance components, the variance components as a percentage of the total variance component. Also, contains the difference in AIC and BIC, and likelihood ratio test Chi-square and p-value for all random effects.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References


See Also

observGlmer, observGlmer3

Examples

```r
data(chinook_survival) # Chinook salmon offspring survival

## Convert replicate-level recorded data to individual-level (binary) data
chinook_survival2 <- buildBinary(dat=chinook_survival,copy=c(2:6,9),one="alive",zero="dead")

## Not run: survival_mod2 <- observGlmer2(observ=chinook_survival2,dam="dam",sire="sire", response="status",fam_link=binomial(link="logit"),position="tray") # a few minutes
survival_mod2

## End(Not run)
```

observGlmer3  
Variance components for non-normal data 3

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a generalized linear mixed-effect model using the `glmer` function of the `lme4` package. Model random effects are dam, sire, dam by sire, and any additional fixed and/or random effects.

Usage

```r
observGlmer3(observ, dam, sire, response, fam_link, remain, quasi = F, iter = 1000)
```

Arguments

- **observ**: Data frame of observed data.
- **dam**: Column name containing dam (female) parent identity information.
- **sire**: Column name containing sire (male) parent identity information.
- **response**: Column name containing the offspring (response) phenotype values.
- **fam_link**: The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
- **remain**: Remaining formula using `lme4` package format.
- **quasi**: Incorporate overdispersion or quasi-error structure.
- **iter**: Number of iterations for computing the parametric bootstrap significance value for any fixed effects.
Details

Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam by sire variance components. Extracts any additional fixed effect and random effect variance components. The fixed-effect variance component is as a single group using the method described by Nakagawa and Schielzeth (2013). The residual variance component for the \textit{fam_links} are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009). Significance values for any fixed effects are determined using likelihood ratio tests and a parametric bootstrap method (Bolker et al. 2009) from the \textit{mixed} function of the \textit{afex} package.

Value

A list object containing the raw variance components, the variance components as a percentage of the total variance component. Contains the difference in AIC and BIC, likelihood ratio test Chi-square and p-value for random and/or fixed effects. Also contains the parametric bootstrap Chi-square and p-value for any fixed effects.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References


See Also

\textit{observGlmer, observGlmer2}
Examples

data(chinook_survival) #Chinook salmon offspring survival
## Convert replicate-level recorded data to individual-level (binary) data
chinook_survival2<- buildBinary(dat=chinook_survival,copy=c(2:6,9),one="alive",zero="dead")
#just a few iterations for the p-value of fixed effect
## Not run: survival_mod3<- observGlmer3(observ=chinook_survival2,dam="dam",sire="sire",
response="status",fam_link=binomial(link="logit"),remain="egg_size + (1|tray)",iter=5)
survival_mod3
## End(Not run)

observLmer

Variance components for normal data

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the *lmer* function of the *lme4* package. Model random effects are dam, sire, and dam by sire.

Usage

`observLmer(observ, dam, sire, response, ml = F)`

Arguments

- `observ` Data frame of observed data.
- `dam` Column name containing dam (female) parent identity information.
- `sire` Column name containing sire (male) parent identity information.
- `response` Column name containing the offspring (response) phenotype values.
- `ml` Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.

Details

Extracts the dam, sire, dam, dam by sire, and residual variance components. Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value

A list object containing the raw variance components, the variance components as a percentage of the total variance component. Also, contains the difference in AIC and BIC, and likelihood ratio test Chi-square and p-value for all random effects.
Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

observLmer2, observLmer3

Examples

data(chinook_length) #Chinook salmon offspring length
length_mod1<- observLmer(observ=chinook_length,dam="dam",sire="sire",response="length")
length_mod1

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the lmer function of the lme4 package. Model random effects are dam, sire, and dam by sire. Options to include one random position and/or one random block effect(s).

Usage

observLmer2(observ, dam, sire, response, position = NULL, block = NULL, ml = F)
Arguments

- **observ**: Data frame of observed data.
- **dam**: Column name containing dam (female) parent identity information.
- **sire**: Column name containing sire (male) parent identity information.
- **response**: Column name containing the offspring (response) phenotype values.
- **position**: Optional column name containing position factor information.
- **block**: Optional column name containing block factor information.
- **ml**: Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.

Details

Extracts the dam, sire, dam, dam by sire, and residual variance components. Extracts optional position and block variance components. Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value

A list object containing the raw variance components, the variance components as a percentage of the total variance component. Also, contains the difference in AIC and BIC, and likelihood ratio test Chi-square and p-value for all random effects.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

observLmer2, observLmer3
Examples

data(chinook_length) #Chinook salmon offspring length
length_mod2<- observLmer2(observ=chinook_length,dam="dam",sire="sire",response="length", position="tray")
length_mod2

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the `lmer` function of the `lme4` package. Model random effects are dam, sire, dam by sire, and any additional fixed and/or random effects.

Usage

`observLmer3(observ, dam, sire, response, remain, ml = F, iter = 1000)`

Arguments

- **observ**: Data frame of observed data.
- **dam**: Column name containing dam (female) parent identity information.
- **sire**: Column name containing sire (male) parent identity information.
- **response**: Column name containing the offspring (response) phenotype values.
- **remain**: Remaining formula using `lme4` package format.
- **ml**: Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.
- **iter**: Number of iterations for computing the parametric bootstrap significance value for any fixed effects.

Details

Extracts the dam, sire, dam, dam by sire, and residual variance components. Extracts any additional fixed effect and random effect variance components. The fixed-effect variance component is as a single group using the method described by Nakagawa and Schielzeth (2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009). Significance values for any fixed effects are determined using likelihood ratio tests and a parametric bootstrap method (Bolker et al. 2009) from the `mixed` function of the `afex` package.
**Value**

A list object containing the raw variance components, the variance components as a percentage of the total variance component. Contains the difference in AIC and BIC, likelihood ratio test Chi-square and p-value for random and/or fixed effects. Also contains the parametric bootstrap Chi-square and p-value for any fixed effects.

**Note**

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

**References**


**See Also**

`observLmer, observLmer2`

**Examples**

data(chinook_length) #Chinook salmon offspring length

#just a few iterations for the p-value of fixed effect
length_mod3 <- observLmer3(observ=chinook_length,dam="dam", sire="sire", response="length", remain="egg_size + (1|tray)", iter=5)

length_mod3
Description

Extracts the power values of dam, sire, and dam by sire variance components from a generalized linear mixed-effect model using the `glmer` function of the `lme4` package.

Usage

```r
powerGlmer(varcomp, nval, fam_link, alpha = 0.05, nsim = 100, poisLog = NULL)
```

Arguments

- `varcomp`: Vector of known dam, sire, and dam by sire variance components, i.e. c(dam, sire, dam x sire).
- `nval`: Vector of known dam, sire, and offspring per family sample sizes, i.e. c(dam, sire, offspring).
- `fam_link`: The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
- `alpha`: Statistical significance value. Default is 0.05.
- `nsim`: Number of simulations. Default is 100.
- `poisLog`: The residual variance component value if using poisson(link="log").

Details

Extracts the dam, sire, dam, and dam by sire power values. The residual variance component for the `fam_links` are described by Nakagawa and Schielzeth (2010, 2013). Power values are calculated by stochastically simulation data and then calculating the proportion of significance values less than `alpha` for each component (Bolker 2008). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value

Prints a data frame with the sample sizes, variance component inputs, variance component outputs, and power values.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.
References


See Also

topGene, powerGlmer2, powerGlmer3

Examples

```r
# 100 simulations
## Not run: powerGlmer(varcomp=c(0.7930,0.1664,0.1673),nval=c(11,11,300),
## fam_link=binomial(link="logit"))
## End(Not run)
```

Description

Extracts the power values of dam, sire, and dam by sire variance components from a generalized linear mixed-effect model using the `glmer` function of the `lme4` package. Options to include one random position and/or one random block effect(s).

Usage

```r
powerGlmer2(varcomp, nval, fam_link, alpha = 0.05, nsim = 100, position = NULL,
block = NULL, poisLog = NULL)
```

Arguments

- `varcomp`: Vector of known dam, sire, dam by sire, and position and/or block variance components, i.e. c(dam, sire, dam x sire, position/block). If there is a position and a block c(..., dam x sire, position, block).
- `nval`: Vector of known dam, sire, offspring per family, and offspring per position or number of blocks sample sizes, i.e. c(dam, sire, offspring, position/block). If there is a position and a block c(..., offspring, position, block).
fam_link  The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
alpha    Statistical significance value. Default is 0.05.
nsim     Number of simulations. Default is 100.
position Optional number of positions.
block    Optional vector of dams and sires per block, e.g. c(2,2).
poisLog  The residual variance component value if using poisson(link="log").

Details
Extracts the dam, sire, dam, dam by sire, and position and/or block power values. The residual variance component for the fam_links are described by Nakagawa and Schielzeth (2010, 2013). Power values are calculated by stochastically simulation data and then calculating the proportion of significance values less than alpha for each component (Bolker 2008). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value
Prints a data frame with the sample sizes, variance component inputs, variance component outputs, and power values.

Note
The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References

See Also
powerGlmer, powerGlmer3
powerGlmer3

Examples

#100 simulations
## Not run: powerGlmer2(varcomp=c(0.7880,0.1667,0.1671,0.0037),nval=c(11,11,300,3300),
position=11,fam_link=binomial(link="logit"))
## End(Not run)

powerGlmer3  Power analysis for non-normal data 3

Description

Extracts the power values of dam, sire, and dam by sire variance components from a generalized linear mixed-effect model using the \textit{glmer} function of the \textit{lme4} package. Model can include additional fixed and/or random effects.

Usage

\texttt{powerGlmer3(var\_rand, n\_rand, design, remain, fam\_link, var\_fix = NULL, n\_fix = NULL, 
alpha = 0.05, nsim = 100, poisLog = NULL, ftest = "LR", iter = NULL)}

Arguments

- \texttt{var\_rand}  Vector of known dam, sire, dam by sire, and remaining random variance components, i.e. c(dam, sire, dam by sire, rand1, rand2, etc.).
- \texttt{n\_rand}  Vector of known dam, sire, family, and remaining random sample sizes, i.e. c(dam, sire, family, rand1, rand2, etc.).
- \texttt{design}  A data frame of the experimental design, using only integers. First three columns must contain and be named "dam", "sire", "family". Remaining columns are the random effects followed by the fixed effects. Continuous fixed effects are a column containing the values 1:nrow(design).
- \texttt{remain}  Remaining formula using \textit{lme4} package format. Must be random effects followed by fixed effects. No interactions or random slopes; formulate as intercepts in design.
- \texttt{fam\_link}  The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
- \texttt{var\_fix}  Vector of known fixed variance components, i.e. c(fix1, fix2, etc.). Continuous fixed random values are sorted to match column values.
- \texttt{n\_fix}  Vector of known fixed sample sizes, i.e. c(fix1, fix2, etc.). Continuous fixed effects must have a sample size of 1.
- \texttt{alpha}  Statistical significance value. Default is 0.05.
- \texttt{nsim}  Number of simulations. Default is 100.
- \texttt{poisLog}  The residual variance component value if using poisson(link="log").
- \texttt{ftest}  Default is "LR" for likelihood ratio test for fixed effects. Option "PB" is for parametric bootstrap.
- \texttt{iter}  Number of iterations for computing the parametric bootstrap significance value for any fixed effects.
Details

Extracts the dam, sire, dam, dam by sire, and any remaining random and fixed effects power values. The residual variance component for the `fam_links` are described by Nakagawa and Schielzeth (2010, 2013). Power values are calculated by stochastically simulation data and then calculating the proportion of significance values less than `alpha` for each component (Bolker 2008). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009). Significance values for any fixed effects are determined using likelihood ratio tests or parametric bootstrap method (Bolker et al. 2009) from the `mixed` function of the `afex` package.

Value

Prints a data frame with the sample sizes, variance component inputs, variance component outputs, and power values.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References


See Also

powerGlmer, powerGlmer2

Examples

```r
#design object: 2 remaining random effects and 1 continuous fixed effect
block=c(2,2); blocN=4; position=16; posN=20; offN=20
dam0<- stack(as.data.frame(matrix(1:(block[1]*blocN),ncol=blocN,nrow=block[1])))
sire0<- stack(as.data.frame(matrix(1:(block[2]*blocN),ncol=blocN,nrow=block[2])))
observ0<- merge(dam0,sire0, by="ind")
```
Power analysis for normal data

Description

Extracts the power values of dam, sire, and dam by sire variance components from a linear mixed-effect model using the \textit{lmer} function of the \textit{lme4} package.

Usage

\begin{verbatim}
powerLmer(varcomp, nval, alpha = 0.05, nsim = 100, ml = F)
\end{verbatim}

Arguments

\begin{itemize}
  \item \texttt{varcomp} \hspace{1cm} Vector of known dam, sire, dam by sire, and residual variance components, i.e. c(dam, sire, dam x sire, residual).
  \item \texttt{nval} \hspace{1cm} Vector of known dam, sire, and offspring per family sample sizes, i.e. c(dam, sire, offspring).
  \item \texttt{alpha} \hspace{1cm} Statistical significance value. Default is 0.05.
  \item \texttt{nsim} \hspace{1cm} Number of simulations. Default is 100.
  \item \texttt{ml} \hspace{1cm} Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.
\end{itemize}

Details

Extracts the dam, sire, dam, and dam by sire power values. Power values are calculated by stochastically simulation data and then calculating the proportion of significance values less than \texttt{alpha} for each component (Bolker 2008). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value

Prints a data frame with the sample sizes, variance component inputs, variance component outputs, and power values.
Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

powerLmer2, powerLmer3

Examples

#100 simulations
#powerLmer(varcomp=c(0.1900,0.0,0.1719,0.6315),nval=c(11,11,10))
#
#5 simulations
powerLmer(varcomp=c(0.1900,0,0.1719,0.6315),nval=c(11,11,10),nsim=5)

Description

Extracts the power values of dam, sire, and dam by sire variance components from a linear mixed-effect model using the lmer function of the lme4 package. Options to include one random position and/or one random block effect(s).

Usage

powerLmer2(varcomp, nval, alpha = 0.05, nsim = 100, position = NULL, block = NULL, ml = F)
Arguments

varcomp Vector of known dam, sire, dam by sire, residual, and position and/or block variance components, i.e. c(dam, sire, dam x sire, residual, position/block). If there is a position and a block c(..., residual, position, block).

nval Vector of known dam, sire, offspring per family, and offspring per position or number of blocks sample sizes, i.e. c(dam, sire, offspring, position/block). If there is a position and a block c(..., offspring, position, block).

alpha Statistical significance value. Default is 0.05.

nsim Number of simulations. Default is 100.

position Optional number of positions.

block Optional vector of dams and sires per block, e.g. c(2,2).

ml Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.

Details

Extracts the dam, sire, dam, dam by sire, and position and/or block power values. Power values are calculated by stochastically simulation data and then calculating the proportion of significance values less than alpha for each component (Bolker 2008). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009).

Value

Prints a data frame with the sample sizes, variance component inputs, variance component outputs, and power values.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


powerLmer3

Power analysis for normal data 3

Description

Extracts the power values of dam, sire, and dam by sire variance components from a linear mixed-effect model using the lmer function of the lme4 package. Model can include additional fixed and/or random effects.

Usage

```
powerLmer3(var_rand, n_rand, design, remain, var_fix = NULL, n_fix = NULL, alpha = 0.05, nsim = 100, ml = F, ftest = "LR", iter = NULL)
```

Arguments

- `var_rand`: Vector of known dam, sire, dam by sire, residual, and remaining random variance components, i.e. c(dam, sire, dam x sire, residual, rand1, rand2, etc.).
- `n_rand`: Vector of known dam, sire, family, and remaining random sample sizes, i.e. c(dam, sire, family, rand1, rand2, etc.).
- `design`: A data frame of the experimental design, using only integers. First three columns must contain and be named "dam", "sire", "family". Remaining columns are the random effects followed by the fixed effects. Continuous fixed effects are a column containing the values 1:nrow(design).
- `remain`: Remaining formula using lme4 package format. Must be random effects followed by fixed effects. No interactions or random slopes; formulate as intercepts in design.
- `var_fix`: Vector of known fixed variance components, i.e. c(fix1, fix2, etc.). Continuous fixed random values are sorted to match column values.
- `n_fix`: Vector of known fixed sample sizes, i.e. c(fix1, fix2, etc.). Continuous fixed effects must have a sample size of 1.

Examples

```
#100 simulations
#position only, e.g. 8 tanks
## Not run: powerLmer2(varcomp=c(0.2030,0.1798,0.5499,0.1077),nval=c(8,8,20,40),position=8)
#block only, e.g. four 2 x 2
## Not run: powerLmer2(varcomp=c(0.2030,0.1798,0.5499,0.1077),nval=c(8,8,20,4),block=c(2,2))
#position and block
## Not run: powerLmer2(varcomp=c(0.2030,0.1798,0.5499,0.1077,0.1077),nval=c(8,8,20,4,4),
#position=8,block=c(2,2))
## End(Not run)
```
alpha     Statistical significance value. Default is 0.05.
nsim      Number of simulations. Default is 100.
ml        Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.
ftest     Default is "LR" for likelihood ratio test for fixed effects. Option "PB" is for parametric bootstrap.
iter      Number of iterations for computing the parametric bootstrap significance value for any fixed effects.

Details
Extracts the dam, sire, dam, dam by sire, and any remaining random and fixed effects power values. Power values are calculated by stochastically simulation data and then calculating the proportion of significance values less than alpha for each component (Bolker 2008). Significance values for the random effects are determined using likelihood ratio tests (Bolker et al. 2009). Significance values for any fixed effects are determined using likelihood ratio tests or parametric bootstrap method (Bolker et al. 2009) from the mixed function of the afex package.

Value
Prints a data frame with the sample sizes, variance component inputs, variance component outputs, and power values.

Note
Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References

See Also
powerLmer, powerLmer2
Examples

```r
# Design object: 2 remaining random effects and 1 continuous fixed effect
block=c(2,2); blocN=4; position=16; posN=20; offN=20

dam0<- stack(as.data.frame(matrix(1:(block[1]*blocN),ncol=blocN,nrow=block[1])))
sire0<- stack(as.data.frame(matrix(1:(block[2]*blocN),ncol=blocN,nrow=block[2])))
observ0<- merge(dam0,sire0, by="ind")
levels(observ0[,1])<- 1:blocN; colnames(observ0)<- c("block","dam","sire")
observ0$family<- 1:nrow(observ0) # Add family
# Expand for offspring, observ0 x offN
observ1<- do.call("rbind", replicate(offN,observ0,simplify=FALSE))
observ1$position<- rep(1:position,each=posN)
observ1$position<- sample(observ1$position,nrow(observ1)) # Shuffle

desn<- observ1[,c(2,3,4,5,1)]; rm(observ0,observ1) # dam, sire, family, position, block

desn$egg_size<- 1:nrow(desn)

# 100 simulations
## Not run: powerLmer3(var_rand=c(0.19,0.03,0.02,0.51,0.1,0.05),n_rand=c(8,8,16,16,4),
## var_fix=0.1,n_fix=1,design=desn,remain="(1|position)+ (1|block)+ egg_size")
## End(Not run)
```

### resampFamily

**Bootstrap resample within families**

**Description**

Bootstrap resample observations grouped by family identities for a specified number of iterations to create a resampled data set.

**Usage**

```r
resampFamily(dat, copy, family, iter)
```

**Arguments**

- `dat` Data frame observed data to resample.
- `copy` Column numbers to copy.
- `family` Column name containing family identity information.
- `iter` Number of iterations for resampling.

**Details**

The resampled data can be used for producing bootstrap confidence intervals.

**Value**

Because of the large file sizes that can be produced, the resampling of each family X is saved separately as a common separated (X_resampF.csv) file in the working directory. These files are merged to create the final resampled data set (resamp_datF.csv).
resampGlmer

See Also

resampRepli

Examples

data(chinook_length) #Chinook salmon offspring length
#resampFamily(dat=chinook_length,copy=c(3:8),family="family",iter=1000)
#example with a couple iterations
#resampFamily(dat=chinook_length,copy=c(3:8),family="family",iter=2)

---

resampGlmer Bootstrap components for non-normal data

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a generalized linear mixed-effect model using the \texttt{glmer} function of the \texttt{lme4} package. Model random effects are dam, sire, and dam by sire.

Usage

resampGlmer(resamp, dam, sire, response, fam_link, start, end, quasi = F)

Arguments

resamp Data frame of bootstrap resampled data.
dam Column name containing dam (female) parent identity information.
sire Column name containing sire (male) parent identity information.
response Column name containing the offspring (response) phenotype values.
fam_link The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
start Starting model number.
end Ending model number.
quasi Incorporate overdispersion or quasi-error structure.

Details

Used for bootstrap resampled data set produced using \texttt{resampRepli} or \texttt{resampFamily}. Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam by sire variance components. The residual variance component for the \texttt{fam_links} are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).
Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. The number of rows in the data frame matches the number of iterations in the resampled data set and each row represents a model number.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References


See Also

`resampGlmer2`, `resampGlmer3`

Examples

data(chinook_resampS) #5 iterations

#survival_rcomp<- resampGlmer(resamp=survival_datR,dam="dam",sire="sire",
#response="status",fam_link=binomial(link="logit"),start=1,end=1000)
## Not run: survival_rcomp<- resampGlmer(resamp=chinook_resampS,dam="dam",sire="sire",
## response="status",fam_link=binomial(link="logit"),start=1,end=5)
## End(Not run)
Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a generalized linear mixed-effect model using the `glmer` function of the `lme4` package. Model random effects are dam, sire, and dam by sire. Options to include one random position and/or one random block effect(s).

Usage

```
resampGlmer2(resamp, dam, sire, response, fam_link, start, end, position = NULL, block = NULL, quasi = F)
```

Arguments

- `resamp`: Data frame of bootstrap resampled data.
- `dam`: Column name containing dam (female) parent identity information.
- `sire`: Column name containing sire (male) parent identity information.
- `response`: Column name containing the offspring (response) phenotype values.
- `fam_link`: The family and link in family(link) format. Supported options are `binomial(link="logit")`, `binomial(link="probit")`, `poisson(link="log")`, and `poisson(link="sqrt")`.
- `start`: Starting model number.
- `end`: Ending model number.
- `position`: Optional column name containing position factor information.
- `block`: Optional column name containing block factor information.
- `quasi`: Incorporate overdispersion or quasi-error structure.

Details

Used for bootstrap resampled data set produced using `resampRepli` or `resampFamily`. Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam by sire variance components. Extracts optional position and block variance components. The residual variance component for the `fam_links` are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for the options of position and/or block. The number of rows in the data frame matches the number of iterations in the resampled data set and each row represents a model number.
Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.

References


See Also

resampGlmer, resampGlmer3

Examples

data(chinook_resampS) #5 iterations

#survival_rcomp2<- resampGlmer2(resamp=survival_datR,dam="dam",sire="sire", #response="status",fam_link=binomial(link="logit"),position="tray",start=1,end=1000)
## Not run: survival_rcomp2<- resampGlmer2(resamp=chinook_resampS,dam="dam",sire="sire", #response="status",fam_link=binomial(link="logit"),position="tray",start=1,end=5)
## End(Not run)

resampGlmer3

Bootstrap components for non-normal data 3

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a generalized linear mixed-effect model using the glmer function of the lme4 package. Model random effects are dam, sire, dam by sire, and any additional fixed and/or random effects.
resampGlmer3(resamp, dam, sire, response, fam_link, start, end, remain, quasi = F)

Arguments

resamp: Data frame of bootstrap resampled data.
dam: Column name containing dam (female) parent identity information.
sire: Column name containing sire (male) parent identity information.
response: Column name containing the offspring (response) phenotype values.
fam_link: The family and link in family(link) format. Supported options are binomial(link="logit"), binomial(link="probit"), poisson(link="log"), and poisson(link="sqrt").
start: Starting model number.
end: Ending model number.
remain: Remaining formula using lme4 package format with # sign (see column names), e.g. fixed# + (1|random#).
quasi: Incorporate overdispersion or quasi-error structure.

Details

Used for bootstrap resampled data set produced using resampRepli or resampFamily. Laplace approximation parameter estimation is used, which is a true likelihood method (Bolker et al. 2009). For the overdispersion option, an observation-level random effect is added to the model (Atkins et al. 2013). Extracts the dam, sire, dam, and dam by sire variance components. Extracts any additional fixed effect and random effect variance components. The fixed-effect variance component is as a single group using the method described by Nakagawa and Schielzeth (2013). The residual variance component for the fam_links are described by Nakagawa and Schielzeth (2010, 2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for remaining formula components. The number of rows in the data frame matches the number of iterations in the resampled data set and each row represents a model number.

Note

The Laplace approximation is used because there were fewer disadvantages relative to penalized quasi-likelihood and Gauss-Hermite quadrature parameter estimation (Bolker et al. 2009). That is, penalized quasi-likelihood is not recommended for count responses with means less than 5 and binary responses with less than 5 successes per group. Gauss-Hermite quadrature is not recommended for more than two or three random effects because of the rapidly declining analytical speed with the increasing number of random effects.
References


See Also

resampGlmer, resampGlmer2

Examples

data(chinook_resampS) #5 iterations

#survival_rcomp3<- resampGlmer3(resamp=survival_datR,dam="dam",sire="sire", response="status",fam_link=binomial(link="logit"),remain="egg_size# + (1|tray#)", start=1,end=1000)
## Not run: survival_rcomp3<- resampGlmer3(resamp=survival_datR,dam="dam",sire="sire", response="status",fam_link=binomial(link="logit"),remain="egg_size# + (1|tray#)", start=1,end=5)
## End(Not run)

resampLmer Bootstrap components for normal data

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the lmer function of the lme4 package. Model random effects are dam, sire, and dam by sire.

Usage

resampLmer(resamp, dam, sire, response, start, end, ml = F)
resampLmer

Arguments

- **resamp**: Data frame of bootstrap resampled data.
- **dam**: Column name containing dam (female) parent identity information.
- **sire**: Column name containing sire (male) parent identity information.
- **response**: Column name containing the offspring (response) phenotype values.
- **start**: Starting model number.
- **end**: Ending model number.
- **ml**: Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.

Details

Used for bootstrap resampled data set produced using `resampRepli` or `resampFamily`. Extracts the dam, sire, dam, dam by sire, and residual variance components. Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. The number of rows in the data frame matches the number of iterations in the resampled data set and each row represents a model number.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

`resampLmer2`, `resampLmer3`
Examples

data(chinook_resampL) #5 iterations

#length_rcomp1<- resampLmer(resamp=length_datR,dam="dam",sire="sire",response="length",
#start=1,end=1000)
length_rcomp1<- resampLmer(resamp=chinook_resampL,dam="dam",sire="sire",response="length",
start=1,end=5)

resampLmer2

Bootstrap components for normal data 2

Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear
mixed-effect model using the lmer function of the lme4 package. Model random effects are dam,
sire, and dam by sire. Options to include one random position and/or one random block effect(s).

Usage

resampLmer2(resamp, dam, sire, response, start, end, position = NULL, block = NULL,
ml = F)

Arguments

resamp Data frame of bootstrap resampled data.
dam Column name containing dam (female) parent identity information.
sire Column name containing sire (male) parent identity information.
response Column name containing the offspring (response) phenotype values.
start Starting model number.
end Ending model number.
position Optional column name containing position factor information.
block Optional column name containing block factor information.
ml Default is FALSE for restricted maximum likelihood. Change to TRUE for
maximum likelihood.

Details

Used for bootstrap resampled data set produced using resampRepli or resampFamily. Extracts the
dam, sire, dam, dam by sire, and residual variance components. Extracts optional position and block
variance components. Calculates the total variance component. Calculates the additive genetic,
non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).
Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for the options of position and/or block. The number of rows in the data frame matches the number of iterations in the resampled data set and each row represents a model number.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

resampLmer, resampLmer3

Examples

data(chinook_resampL) #5 iterations

#length_rcomp2<- resampLmer2(resamp=length_datR,dam="dam",sire="sire",response="length", #start=1,end=1000,position="tray")
length_rcomp2<- resampLmer2(resamp=chinook_resampL,dam="dam",sire="sire",response="length", #start=1,end=5,position="tray")
Description

Extracts additive genetic, non-additive genetic, and maternal variance components from a linear mixed-effect model using the `lmer` function of the `lme4` package. Model random effects are dam, sire, dam by sire, and any additional fixed and/or random effects.

Usage

```r
resampLmer3(resamp, dam, sire, response, start, end, remain, ml = F)
```

Arguments

- `resamp`: Data frame of bootstrap resampled data.
- `dam`: Column name containing dam (female) parent identity information.
- `sire`: Column name containing sire (male) parent identity information.
- `response`: Column name containing the offspring (response) phenotype values.
- `start`: Starting model number.
- `end`: Ending model number.
- `remain`: Remaining formula using `lme4` package format with # sign (see column names), e.g. fixed# + (1|random#).
- `ml`: Default is FALSE for restricted maximum likelihood. Change to TRUE for maximum likelihood.

Details

Used for bootstrap resampled data set produced using `resampRepli` or `resampFamily`. Extracts the dam, sire, dam, dam by sire, and residual variance components. Extracts any additional fixed effect and random effect variance components. The fixed-effect variance component is as a single group using the method described by Nakagawa and Schielzeth (2013). Calculates the total variance component. Calculates the additive genetic, non-additive genetic, and maternal variance components (see Lynch and Walsh 1998, p. 603).

Value

A data frame with columns containing the raw variance components for dam, sire, dam by sire, residual, total, additive genetic, non-additive genetic, and maternal. Also columns containing the raw variance components for remaining formula components. The number of rows in the data frame matches the number of iterations in the resampled data set and each row represents a model number.

Note

Maximum likelihood (ML) estimates the parameters that maximize the likelihood of the observed data and has the advantage of using all the data and accounting for non-independence (Lynch and Walsh 1998, p. 779; Bolker et al. 2009). On the other hand, ML has the disadvantage of assuming that all fixed effects are known without error, producing a downward bias in the estimation of the residual variance component. This bias can be large if there are lots of fixed effects, especially if sample sizes are small. Restricted maximum likelihood (REML) has the advantage of not assuming the fixed effects are known and averages over the uncertainty, so there can be less bias in the
estimation of the residual variance component. However, REML only maximizes a portion of the likelihood to estimate the effect parameters, but is the preferred method for analyzing large data sets with complex structure.

References


See Also

resampLmer, resampLmer2

Examples

data(chinook_resampL)

#length_rcomp3<- resampLmer3(resamp=length_datR,dam="dam",sire="sire",response="length", #start=1,end=1000,remain="egg_size# + (1|tray#)")
length_rcomp3<- resampLmer3(resamp=chinook_resampL,dam="dam",sire="sire",response="length", start=1,end=5,remain="egg_size# + (1|tray#)")

resampRepli Bootstrap resample within replicates

Description

Bootstrap resample observations grouped by replicate identities within family identities for a specified number of iterations to create a resampled data set.

Usage

resampRepli(dat, copy, family, replicate, iter)

Arguments

dat Data frame observed data to resample.
copy Column numbers to copy.
family Column name containing family identity information.
replicate Column name containing replicate identity information.
iter Number of iterations for resampling.
Details

The resampled data can be used for producing bootstrap confidence intervals.

Value

Because of the large file sizes that can be produced, the resampling of each replicate Y per family X is saved separately as a common separated (X_Y_resampR.csv) file in the working directory. These files are merged to create the final resampled data set (resamp_datR.csv).

See Also

resampFamily

Examples

data(chinook_length) #Chinook salmon offspring length
#resampRepli(dat=chinook_length,copy=c(3:8),family="family",replicate="repli",iter=1000)
#example with a couple iterations
#resampRepli(dat=chinook_length,copy=c(3:8),family="family",replicate="repli",iter=2)
Index

* datasets
  chinook_bootL, 10
  chinook_bootS, 11
  chinook_jackL, 12
  chinook_jackS, 13
  chinook_length, 14
  chinook_resampL, 15
  chinook_resampS, 17
  chinook_survival, 18

* package
  fullfact-package, 2

barMANA, 6
boxMANA, 7
buildBinary, 8, 10
buildMulti, 9, 9

chinook_bootL, 10
chinook_bootS, 11
chinook_jackL, 12
chinook_jackS, 13
chinook_length, 14
chinook_resampL, 15
chinook_resampS, 17
chinook_survival, 18
ciJack, 19, 22, 23
ciJack2, 20, 21, 23
ciJack3, 20, 22, 22
ciMANA, 23, 26, 27
ciMANA2, 24, 25, 27
ciMANA3, 24, 26, 26

ciGlmer, 40, 43, 44
obsGlmer, 41, 43, 44
obsGlmer2, 41, 43, 44
obsGlmer3, 41, 43, 43
obsLmer, 45, 47, 49
obsLmer2, 46, 46, 49
obsLmer3, 46, 47, 48

fullfact (fullfact-package), 2
fullfact-package, 2

JackGlmer, 28, 32, 34
JackGlmer2, 29, 30, 34
JackGlmer3, 29, 32, 32
JackLmer, 34, 37, 39
JackLmer2, 35, 36, 39

powerGlmer, 50, 52, 54
powerGlmer2, 51, 51, 54
powerGlmer3, 51, 52, 53
powerLmer, 55, 58, 59
powerLmer2, 56, 56, 59
powerLmer3, 56, 58, 58

resampFamily, 60, 72
resampGlmer, 61, 64, 66
resampGlmer2, 62, 63, 66
resampGlmer3, 62, 64, 64
resampLmer, 66, 69, 71
resampLmer2, 67, 68, 71
resampLmer3, 67, 69, 69
resampRepli, 61, 71