

Random Effects ANOVA

Alexander McLain

Contents

1 Example one: HIV data

1

1 Example one: HIV data

First, let's read in the data and look at the variables.

```
library(tidyverse)
BP_data_wide <- read.csv("BP_data.csv")
BP_data_wide
```

ID	BP1	BP2	BP3	BP4	BP5	BP6	BP7	BP8	BP9	BP10
1	130	136	124	121	128	138	130	126	136	133
2	123	116	117	120	117	113	111	124	110	116
3	144	139	136	142	142	133	140	138	140	139
4	128	133	127	124	132	130	134	124	131	128
5	128	128	122	120	121	123	132	130	126	132

```
BP_data <- BP_data_wide %>% pivot_longer(cols = starts_with("BP"),
                                         names_to = "Time",
                                         names_prefix = "BP",
                                         values_to = "BP",
                                         values_drop_na = TRUE)
```

To fit the random effects ANOVA model we're going to use the `lmer` function in the `lme4` package. This function is very similar to the `glms` function except we won't use the `correlation` statement or `weights` statements much.

Here, the random effects are put in the `formula` portion of the model.

```
library(lme4)
?lmer

## Fit Linear Mixed-Effects Models
##
## Description:
##
##     Fit a linear mixed-effects model (LMM) to data, via REML or
##     maximum likelihood.
##
## Usage:
##
##     lmer(formula, data = NULL, REML = TRUE, control = lmerControl(),
```

```

##          start = NULL, verbose = 0L, subset, weights, na.action,
##          offset, contrasts = NULL, devFunOnly = FALSE)
##
## Arguments:
##
## formula: a two-sided linear formula object describing both the
##          fixed-effects and random-effects part of the model, with the
##          response on the left of a '~' operator and the terms,
##          separated by '+' operators, on the right. Random-effects
##          terms are distinguished by vertical bars ('|') separating
##          expressions for design matrices from grouping factors. Two
##          vertical bars ('||') can be used to specify multiple
##          uncorrelated random effects for the same grouping variable.
##          (Because of the way it is implemented, the '||'-syntax works
##          only for design matrices containing numeric (continuous)
##          predictors; to fit models with independent categorical
##          effects, see 'dummy' or the 'lmer_alt' function from the
##          'afex' package.)
##
## data: an optional data frame containing the variables named in
##        'formula'. By default the variables are taken from the
##        environment from which 'lmer' is called. While 'data' is
##        optional, the package authors strongly recommend its use,
##        especially when later applying methods such as 'update' and
##        'drop1' to the fitted model (such methods are not guaranteed
##        to work properly if 'data' is omitted). If 'data' is
##        omitted, variables will be taken from the environment of
##        'formula' (if specified as a formula) or from the parent
##        frame (if specified as a character vector).
##
## REML: logical scalar - Should the estimates be chosen to optimize
##        the REML criterion (as opposed to the log-likelihood)?
##
## control: a list (of correct class, resulting from 'lmerControl()' or
##            'glmerControl()' respectively) containing control parameters,
##            including the nonlinear optimizer to be used and parameters
##            to be passed through to the nonlinear optimizer, see the
##            '*lmerControl' documentation for details.
##
## start: a named 'list' of starting values for the parameters in the
##         model. For 'lmer' this can be a numeric vector or a list
##         with one component named '"theta"'.
##
## verbose: integer scalar. If '> 0' verbose output is generated during
##          the optimization of the parameter estimates. If '> 1'
##          verbose output is generated during the individual penalized
##          iteratively reweighted least squares (PIRLS) steps.
##
## subset: an optional expression indicating the subset of the rows of
##          'data' that should be used in the fit. This can be a logical
##          vector, or a numeric vector indicating which observation
##          numbers are to be included, or a character vector of the row
##          names to be included. All observations are included by
##          default.

```

```

##
## weights: an optional vector of 'prior weights' to be used in the
##          fitting process. Should be 'NULL' or a numeric vector.
##          Prior 'weights' are not normalized or standardized in any
##          way. In particular, the diagonal of the residual covariance
##          matrix is the squared residual standard deviation parameter
##          'sigma' times the vector of inverse 'weights'. Therefore, if
##          the 'weights' have relatively large magnitudes, then in order
##          to compensate, the 'sigma' parameter will also need to have a
##          relatively large magnitude.
##
## na.action: a function that indicates what should happen when the data
##             contain 'NA's. The default action ('na.omit', inherited from
##             the 'factory fresh' value of 'getOption("na.action")') strips
##             any observations with any missing values in any variables.
##
## offset: this can be used to specify an a priori known component to
##           be included in the linear predictor during fitting. This
##           should be 'NULL' or a numeric vector of length equal to the
##           number of cases. One or more 'offset' terms can be included
##           in the formula instead or as well, and if more than one is
##           specified their sum is used. See 'model.offset'.
##
## contrasts: an optional list. See the 'contrasts.arg' of
##             'model.matrix.default'.
##
## devFunOnly: logical - return only the deviance evaluation function.
##             Note that because the deviance function operates on variables
##             stored in its environment, it may not return exactly the
##             same values on subsequent calls (but the results should
##             always be within machine tolerance).

```

```

library(lmerTest) # Added to get p-values
RE_form <- BP ~ 1|ID
re_anova <- lmer( formula = RE_form , data = BP_data)
summary(re_anova)

```

```

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: RE_form
## Data: BP_data
##
## REML criterion at convergence: 301.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.09154 -0.72981 -0.00817  0.68723  1.79650
##
## Random effects:
##  Groups   Name                Variance Std.Dev.
##  ID       (Intercept) 64.14      8.009
##  Residual                19.12     4.372
## Number of obs: 50, groups: ID, 5
##
## Fixed effects:

```

```
##           Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 128.300      3.635    4.000    35.3 3.84e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

round( confint(re_anova), 2)
```

	2.5 %	97.5 %
.sig01	4.12	15.78
.sigma	3.60	5.46
(Intercept)	120.48	136.12