Multivariate Statistics and Methodology using R [PSYL11054]

Block on Linear Mixed Models

Lab Session 4: Multiple levels and nested and crossed effects, random-effects structure

packages needed: `lme4`, `lattice` (lattice comes with R; it needs to be loaded using `library()` but there is no need to install it)

Examples in previous lab sessions have highlighted changes over time (vocabulary study, arthritis study, fertility and climate data). Now we are going to focus on an alternative setup for repeated measurements where individuals are studied under a series of related conditions. We will use experiments from psycholinguistics (Lab 4) and cognitive psychology (Lab 5) as examples.

1. Nested and crossed effects

An important advantage of the `lme4` package in R is that it offers mixed-effects modelling with crossed random effects. For illustration, we will use a (constructed) dataset taken from Raaijmakers et al. (1999) (see their Table 2).[1] A lexical decision task was performed where subjects had to decide whether each stimulus was a word. The interval between the presentation of a distractor stimulus and the target word was varied to see if this affected reaction time in the classification task. This interval is referred to as stimulus onset asynchrony (SOA). 8 words were shown to each participant under either a long or short SOA. Load the data

1.1 Load data and examine design

```
> # Load lexical decision task data
> ldt <- read.table("raai.dat", header=T)
```

Question 1 Use `table()` to explore the design of the experiment. How many times does each participant see each stimulus item W1–W8? Are the ways of the data (subject, SOA, and item) nested or crossed?

Tip 1 Pass one or two columns from a data from to table to build a contingency table.

```
> # Examine experimental design
> table(ldt$SOA)
> table(ldt[,c('Subject','Item')])
> table(ldt[,c('Subject', 'SOA')])
> table(ldt[,c('Item', 'SOA')])
```

Tip 2 **Nested** means a group at one level of the hierarchy only appears in one of the higher-order levels. **Crossed** designs mean that each of the lower-order group is observed within each of the higher-order groups.

solution:

[1] This data is also discussed in Baayen (2008) (Section 7.2.1, note that the wording and data frames differ in different (pre)prints of the book) as well as Baayen et al. (2008, p. 401).

```
> table(ldt$SOA)
long short
   32    32
```

=> The factor SOA has two levels, long and short. There are 32 observations for each SOA condition.

```
> table(ldt[,c('Subject','Item')])
       Item
Subject W1 W2 W3 W4 W5 W6 W7 W8
     S1  1  1  1  1  1  1  1  1
     S2  1  1  1  1  1  1  1  1
     S3  1  1  1  1  1  1  1  1
     S4  1  1  1  1  1  1  1  1
     S5  1  1  1  1  1  1  1  1
     S6  1  1  1  1  1  1  1  1
     S7  1  1  1  1  1  1  1  1
     S8  1  1  1  1  1  1  1  1
```

=> Each subject responds to each word (item) once.

```
> table(ldt[,c('Subject', 'SOA')])
       SOA
Subject long short
     S1    4     4
     S2    4     4
     S3    4     4
     S4    4     4
     S5    4     4
     S6    4     4
     S7    4     4
     S8    4     4
```

=> Each subject responds an equal number of times to items in the two SOA conditions.

```
> table(ldt[,c('Item', 'SOA')])
> table(ldt[,c('Item', 'SOA')])
    SOA
Item long short
  W1    0     8
  W2    0     8
  W3    0     8
  W4    0     8
  W5    8     0
  W6    8     0
  W7    8     0
  W8    8     0
```

=> Notably, the items are NESTED under SOA: items 1 through 4 are always used in the short condition, and items 5 through 8 in the long condition.

solution:
Subject and item are CROSSED in this design. Subject and SOA are also CROSSED. The items, however, are NESTED under SOA.
(This kind of design is known as a split-plot design.)

1.2 Variance component models

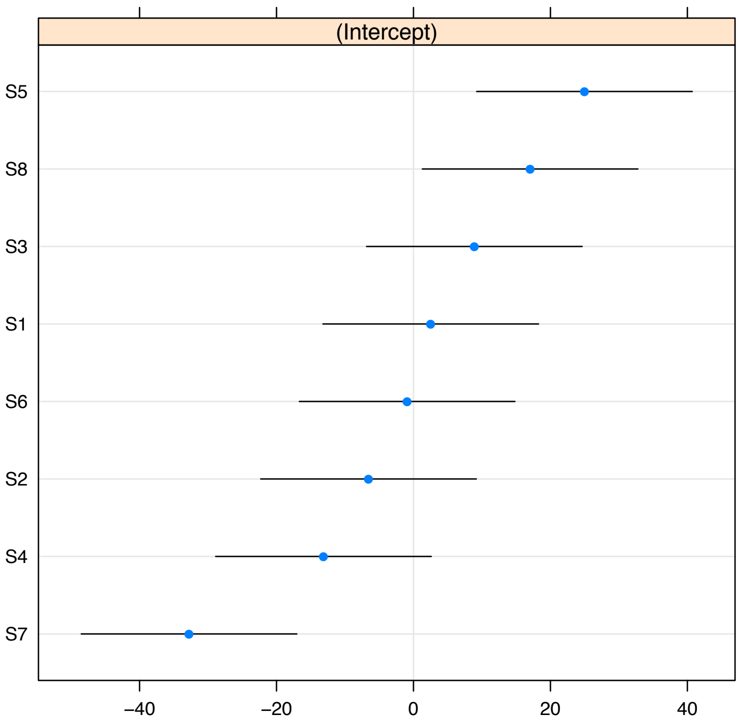
Because the stimulus words can be seen as samples from a larger population of possible stimuli, it is useful to model them as random effects along with the subjects. We start with a model that estimates the effect of SOA and reaction time while accounting for differences between subjects.

```
> library(lme4)

> # Model with subject as a random effect
> ldt.M1 <- lmer(RT ~ 1 + SOA + (1 | Subject), data=ldt)
```

The subject intercepts from the model can be visualized with a "caterpillar plot" for conditional modes (lecture 3, slide 56; lecture 4, slide 33):

```
> # Dot plot of random effects for subject
> print(dotplot(ranef(ldt.M1, condVar=T))[['Subject']])
```



The `condVar=T` requests the standard errors for each subject's parameter estimate. The name of the group to be plotted goes between the `[[' ']]`.
(In `lme4` versions < 1.0+ the `condVar` argument was named `postVar`.)

Question 2 What is the difference in reaction times between the long and short SOA condition?

solution:

```
> str(ldt$SOA)
```

tells us that SOA is a factor with 2 levels "long" and "short" (long is the reference)

```
> fixef(ldt.M1)
(Intercept)     SOAshort
```

```
540.90625    22.40625
```

tells us that reaction times for the short SOA are 22.4 ms longer than for the long SOA

Question 3 Model `ldt.M1` fits an intercept for each subject as a random effect. Build another model, called `ldt.M2`, that also fits an intercept for each item as a random effect.
Tip Note that the determination of nested or crossed random effects is done for you by the `lmer` programme.

solution:

```
ldt.M2 <- lmer(RT ~ 1 + SOA + (1 | Subject) + (1 | Item),
data=ldt)
```

Compare the fit of both models using the likelihood ratio test

```
> # Compare fit of model 1 and 2
> anova(ldt.M1, ldt.M2)
```

NOT IN STUDENT HANDOUT:

```
Data: ldt
Models:
ldt.M1: RT ~ 1 + SOA + (1 | Subject)
ldt.M2: RT ~ 1 + SOA + (1 | Subject) + (1 | Item)
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
ldt.M1  4 613.23 621.87 -302.62   605.23
ldt.M2  5 582.12 592.92 -286.06   572.12 33.109      1  8.712e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 4 Does model 2 offer a better fit compared to the likelihood ratio test?
solution:
Yes, model 2 offers a better fit, logLik $\Delta\chi^2(1) = 33.109$, $p < .001$.
Note that `logLik` values are typically negative, and relatively larger values (close to zero) are indicative of a better fit (lecture 3), `-286.06 > -302.62`.

Question 5 Which model has the better fit according to AIC?
solution:
Lower values of AIC are better (lecture 3).
Model 2 provides the better fit, `582.12 < 613.23`.

Question 6 How does the estimate of the difference in RT between the long and short SOA from model 2 compare to model 1?
solution:
It's identical:

```
> fixef(ldt.M2)['SOAshort']
SOAshort
22.40625
```

1.3 Between subject variability in effect of SOA

Like fitting a different slope for each subject to represent change over time in a longitudinal study, individual differences in treatment effects or experimental conditions can also be examined with linear mixed-effects models. For example, do subjects differ in how quickly they react under short versus long SOA?

```
> # Varying slope of SOA for Subject
> ldt.M3 <- lmer(RT ~ 1 + SOA + (1 + SOA | Subject) + (1 |
Item), data=ldt)
```

Question 7 Do subjects differ in how changes in SOA affect their reaction times?

solution:

```
> anova(ldt.M2, ldt.M3)
Data: ldt
Models:
ldt.M2: RT ~ 1 + SOA + (1 | Subject) + (1 | Item)
ldt.M3: RT ~ 1 + SOA + (1 + SOA | Subject) + (1 | Item)
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
ldt.M2  5 582.12 592.92 -286.06   572.12
ldt.M3  7 557.59 572.70 -271.80   543.59 28.534      2  6.368e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Yes, subjects differ in how changes in SOA affect their reaction times because model 3 offers a better fit, logLik $\Delta\chi^2(2) = 28.534$, $p < .001$.

Question 8 Visualize these individual differences. Technically, produce a caterpillar plot of the by-subject random slopes for SOA.
solution:
`print(dotplot(ranef(ldt.M3, condVar=T))[['Subject']][2])`
- for slopes only
OR (simply applying the code from p. 3 above)
`print(dotplot(ranef(ldt.M3, condVar=T))[['Subject']])`
- for both intercept and slopes

Question 9 Based on this incremental model building, what is your conclusion with regard to the SOA effect?
Tip keyword is significance of fixed effects
solution: The fixed effect of SOA is only significant in model 1, $t > 2$; it does not generalize once additional sources of variation in the data are controlled for ($t < 2$ in models 2 and 3).

2. Random-effects structure & zero-correlation parameter model

It has been suggested that linear mixed-effects models generalize best when they include the maximal random effects structure justified by the design (Barr, Levy, Scheepers, & Tily, 2013), see Lecture 4 as of slide 37.

Question 10 What is the maximal random effects structure justified by the design for the lexical decision data in data frame `ldt`?
Tip Random slopes are only appropriate for within-subjects and within-items factors

solution: it is the structure of model `ldt.M3: (1 + SOA | Subject) + (1 | Item)`; it is not sensible to include by-item random slopes as the items are nested under SOA

Let's take another look at model `ldt.M3`. The by-subject random effects for slopes and intercepts are paired observations. Therefore, the model specification that we used here allows for these two random variables (intercept and slope) to be correlated. The estimate of this correlation is a parameter of the mixed effects model.

Question 11 How large is this correlation?
Tip You will find it in the 'summary' output of model `ldt.M3`

solution: -0.81

```
print(summary(ldt.M3),cor=F)
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Subject  (Intercept) 855.9    29.26
          SOAshort    491.8    22.18    -0.81
 Item     (Intercept) 449.4    21.20
 Residual             100.2    10.01
Number of obs: 64, groups:  Subject, 8; Item, 8
```

A question that arises is whether there is reliable evidence that this correlation parameter is different from zero. We can test this by specifying a zero-correlation parameter model, i.e., we are forcing the correlation between by-subject random intercept and slope to be zero.
- See slides 46-48 in Lecture 4, from slide 47:

- Syntax for uncorrelated random intercept and slope
  - `1 + A + (1 | subject) + (0 + A | subject)`
  - `More convenient: A + (A || subject)`

Question 12 Change the random-effects structure of model `ldt.M3` such that it turns into a zero-correlation parameter model; call this new model `ldt.M3.zcp`.
a) Use the above syntax

solution:

```
ldt.M3.zcp <- lmer(RT ~ 1 + SOA + (1 + SOA || Subject) + (1 |
Item), data=ldt)

> print(summary(ldt.M3.zcp),cor=F)
…
Random effects:
 Groups    Name        Variance Std.Dev. Corr
 Subject   (Intercept) 146.7    12.11
 Subject.1 SOAlong     709.2    26.63
           SOAshort    155.8    12.48    0.56
 Item      (Intercept) 449.4    21.20
 Residual              100.2    10.01
Number of obs: 64, groups:  Subject, 8; Item, 8
…
```

b) Well, if you use above syntax, there is still a correlation in the model output, how annoying. As I mentioned in the lecture, there are (currently) some constraints on using this syntax. A detailed account is provided here: http://rpubs.com/Reinhold/22193. Specifically, 'SOA' is a categorical variable and therefore coded as a factor. The solution to the problem is to turn 'SOA' into a numeric variable. If you don't know how to convert a factor to a numeric variable, ask Google. Then try again to specify a zero-correlation parameter model.

solution:

```
ldt$SOA_num <- as.numeric(ldt$SOA) # not sure why this works,
and as.numeric(levels(ldt$SOA)) doesn't

ldt.M3.zcp <- lmer(RT ~ 1 + SOA_num + (1 + SOA_num || Subject)
+ (1 | Item), data=ldt)

print(summary(ldt.M3.zcp),cor=F)
…
Random effects:
 Groups     Name        Variance Std.Dev.
 Subject    (Intercept) 2060.2   45.39
 Subject.1  SOA_num      419.6   20.48
 Item       (Intercept)  449.1   21.19
 Residual                102.8   10.14
Number of obs: 64, groups:  Subject, 8; Item, 8
```

Question 13 Perform a likelihood ratio test to check whether forcing the correlation parameter to be zero significantly decreases the goodness of fit.
Tip The critical model for this comparison are models `ldt.M3.zcp` and `ldt.M3`

```
> anova(ldt.M3.zcp,ldt.M3)
refitting model(s) with ML (instead of REML)
Data: ldt
Models:
ldt.M3.zcp: RT ~ 1 + SOA_num + ((1 | Subject) + (0 + SOA_num |
Subject)) +
ldt.M3.zcp:     (1 | Item)
ldt.M3: RT ~ 1 + SOA + (1 + SOA | Subject) + (1 | Item)
           Df    AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
ldt.M3.zcp  6 570.24 583.2 -279.12   558.24
ldt.M3      7 557.59 572.7 -271.80   543.59 14.653      1  0.0001292
***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 14
Match the R output obtained when answering Question 13 to a sentence of the following form:
According to the likelihood ratio test, `ldt.M3.zcp` offers an equal/ worse/ better fit (make a pick) than `ldt.M3`, logLik $\Delta\chi 2$(x) = xx.y, p = .xxx.
Tips
- (x) is difference in degrees of freedoms between the two models
- xx.y is the Chi square value

solution:
According to the likelihood ratio test, `ldt.M3.zcp` offers a worse fit than `ldt.M3`, logLik $\Delta\chi 2(1) = 14.6$, $p < .001$.

3. References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: University Press.
http://www.sfs.uni-tuebingen.de/~hbaayen/publications/baayenCUPstats.pdf

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412. doi:10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278. doi:10.1016/j.jml.2012.11.001

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41*(3), 416-426. doi:10.1006/jmla.1999.2650