PSYC402-week-19-LME-3

Rob Davies (Lancaster University)

Targets for Week 19 – three elements

1. the capacity to understand mixed-effects models
2. the capacity to work with them practically in R
3. the capacity to present the results

We want to develop the capacity to understand mixed-effects models, to:

1. recognize where data have a multilevel structure
2. recognize where multilevel or mixed-effects models are required
3. distinguish the elements of a mixed-effects model, including fixed effects and random effects

We want to develop the capacity to understand mixed-effects models, to:

Be able to explain how

1. random effects can be understood in terms of random differences (or deviations) between groups or classes or individuals in intercepts or slopes
2. random effects can be understood in terms of variances
3. mixed-effects models work better than linear models, for multilevel structured data, because they take into account variances associated with random differences in intercepts or slopes
4. mixed-effects models work better because they allow partial-pooling of estimates

Develop the capacity to work practically in R with mixed-effects models, to:

1. be able to specify a mixed-effects model in lmer() code
2. be able to identify how the mixed-effects model code varies depending on the kinds of random effects that are assumed
3. be able to identify the elements of the output or results that come from an lmer() mixed-effects analysis
4. be able to interpret the fixed-effects estimates, consistent with the interpretation of the linear model effect coefficient estimate
5. be able to interpret the random effects estimates, variance, covariance

Develop the capacity to talk about and present the results, to:

1. be able to describe in words and summary tables the results of a mixed-effects model
2. be able to visualise the effects estimates from a mixed-effects model

Introduction to the ML study dataset – ML hypothesised:

effects of stimulus attributes: words that are shorter, learnt earlier in life, and appear frequently in the language would be easier to recognise;

effects of participant attributes: older readers would be faster and more accurate than younger readers in word recognition;

interactions between the effects of word attributes and person attributes: it was possible that better (older) readers would be faster and would be less affected by the attributes of words – they would show smaller effects of word frequency, length, age-of-acquisition in better readers' performance.

Introduction to the ML study dataset

- 39 participants were asked to respond to 160 word and 160 nonword stimuli in the lexical decision task
- Participants had to press a button: 'yes' (that is a word) or 'no'
- ML recorded response RT, as well as participant and stimulus word attribites
- subjects.behaviour.words-310114.csv holds information about (word) stimuli, participants, and responses in the ML study

When we do we need mixed-effects models? When we have repeated measures data

Many researchers conduct studies where it is not sensible to think of observations as being nested (Baayen et al., 2008): crossed random effects

- In a reading study, we may ask all individuals in a participant sample to read all words in a stimulus sample of words
- For each individual, we will have multiple observations and these observations will not be independent
  - One participant will tend to be slower or less accurate compared to another
  - Her responses may be more or less susceptible to the effects of the experimental variables
- The lowest trial-level observations can be grouped with respect to participants

When we do we need mixed-effects models? When we have repeated measures data

- In a reading study, we may ask all individuals in a participant sample to read all words in a stimulus sample of words
- For each stimulus, there are multiple observations and these observations will not be independent
  - One stimulus may prove to be more challenging to all participants compared to another, eliciting slower or less accurate responses
  - The effects of *within-items* experimental variables may be more or less prominent for responses to some stimuli than to others
- So the data can also be grouped by stimuli

Repeated measures designs and the 'language-as-fixed-effect fallacy'

- If you are doing a repeated measures study in which there are different stimuli and different subjects
- And all subjects see all stimuli
- Then you need to take into account both random variation due to differences between people and random variation due to differences between stimuli (words)

Making the data tidy – next

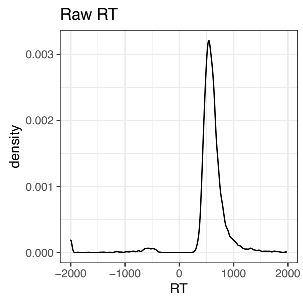
Making the data tidy

We only need to perform steps 1 and 3 of the usual *tidy data* process

1. import the data
2. restructure the data
3. transform variables

Examine the frequency distribution of RT values – Using density plots

```
ML.all %>%
ggplot(aes(x = RT)) +
geom_density() +
ggtitle("raw RT") +
theme_bw()
```



Transform the data: removing observations using filter()

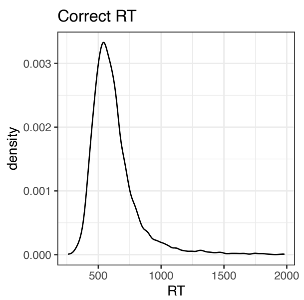
```
ML.all.correct <-
  filter(ML.all, RT >= 200)

length(ML.all$RT)
```

```
## [1] 5440
```

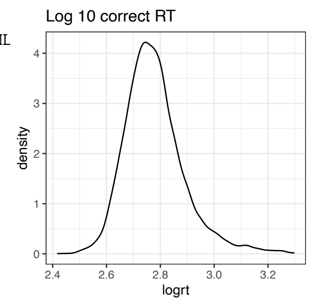
```
length(ML.all.correct$RT)
```

```
## [1] 5257
```

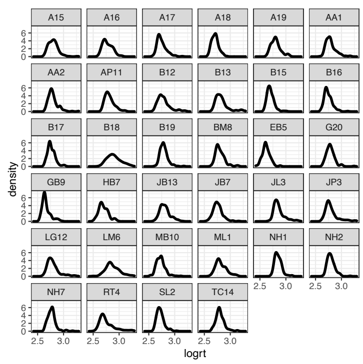


The log10() transformation of RT

```
ML.all.correct$logrt <- log10(ML
```



Use facetting in ggplot to examine log RT by person



Approximations to Linear Mixed-effects models: complete pooling or no pooling

Approximations to Linear Mixed-effects models: complete pooling

We can estimate the relationship between lexical decision RTs and word frequency using a linear model

$$Y_{ij} = \beta_0 + \beta_1 X_j + e_{ij} \quad (1)$$

- Y_{ij} is observed RT, the latency of the response made by the i participant to the j item;
- $\beta_1 X_j$ refers to the fixed effect of the explanatory variable (here, word frequency), where the frequency value X_j is different for different words j, and β_1 is the estimated coefficient of the effect;
- e_{ij} is the residual error term: differences between observed Y_{ij} and predicted values (given the model) for each response made by the i participant to the j item

The linear model can be fit in R using the lm() function

```
ML.all.correct.lm  <- lm(logrt ~
                                        LgSUBTLCD,
                                    data = ML.all.correct)
summary(ML.all.correct.lm)
```

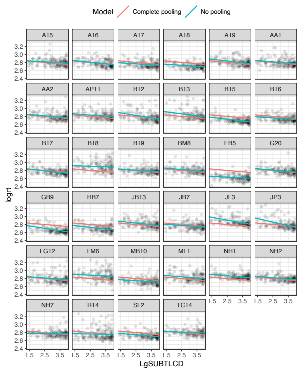

Approximations to Linear Mixed-effects models: no pooling

Alternatively, we can examine variation between participants by analysing the data for each participant's responses separately

- Fitting a linear model of the effect of word frequency on lexical decision RTs
- To estimate the intercept and the slope of the frequency effect for each participant using just that person's data

Complete-pooling compared to no-pooling estimates

- Blue lines represent estimated intercepts and frequency effect slopes calculated for each participant analysed separately no pooling
- Red lines represent estimated effects calculated over all data complete pooling



No pooling and *complete pooling* estimates are often but not always similar

- The "complete pooling" estimate is unsatisfactory because it ignores the variation between the participants: some people *are* slower than others; some people *do* show a larger frequency effect than others
- The "no pooling" estimate is unsatisfactory because it ignores the similarities between the participants
- What we need is an analytic method that can *both* estimate the overall population-level effect (here, of word frequency) and take into account the differences between sampling units (here, participants)

Mixed-effects models – next

Mixed-effects models – We can incorporate *fixed effects* due to the average intercept and the average frequency effect, as well as the *random effects*, error variance due to unexplained differences between participants in intercepts and in frequency effects

$$Y_{ij} = \gamma_0 + \gamma_1 X_j + U_{0i} + U_{1i} X_j + e_{ij} \quad (2)$$

- where the outcome Y_{ij} is related to ...
- the average intercept γ_0 and differences between i participants in the intercept U_{0i};
- the average effect of the explanatory variable frequency $\gamma_1 X_j$ and differences between i participants in the slope $U_{1i} X_j$;
- in addition to residual error variance e_{ij}.

Mixed-effects models – Further, we can take into account error variance due to unexplained differences between responses to different items in intercepts

$$Y_{ij} = \gamma_0 + \gamma_1 X_j + U_{0i} + U_{1i} X_j + W_{0j} + e_{ij} \quad (3)$$

- where the outcome Y_{ij} is related also to ...
- random differences between items in intercepts W_{0j}

In fact, in conducting mixed-effects models, we do not usually aim to examine the specific deviations

We estimate just the spread of deviations – variances – by-participants or by-items

- $var(U_{0i})$ variance of deviations by-participants from the average intercept;
- $var(U_{1i} X_j)$ variance of deviations by-participants from the average slope of the frequency effect;
- $var(W_{0j})$ variance of deviations by-items from the average intercept;
- $var(e_{ij})$ residuals, at the response level, after taking into account all other terms.

We may expect the random effects of participants or items to covary

Our specification of the random effects can incorporate terms corresponding to the covariance of deviations

- $covar(U_{0i}, U_{1i} X_j)$

Fitting a mixed-effects model using lmer()

```
ML.all.correct.lmer  <- lmer(logrt ~

                             LgSUBTLCD +

                             (LgSUBTLCD + 1|subjectID) +

                             (1|item_name),

                            data = ML.all.correct)

summary(ML.all.correct.lmer)
```


Fitting a mixed-effects model using lmer()

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: logrt ~ LgSUBTLCD + (LgSUBTLCD + 1 | subjectID) + (1 | item_name)
##    Data: ML.all.correct
##
## REML criterion at convergence: -9868.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6307 -0.6324 -0.1483  0.4340  5.6132
##
## Random effects:
##  Groups    Name        Variance  Std.Dev. Corr
##  item_name (Intercept) 0.0003268 0.01808
##  subjectID (Intercept) 0.0054212 0.07363
##            LgSUBTLCD   0.0002005 0.01416  -0.63
##  Residual              0.0084333 0.09183
## Number of obs: 5257, groups:  item_name, 160; subjectID, 34
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.887997   0.015479 186.577
## LgSUBTLCD   -0.034471   0.003693  -9.333
##
## Correlation of Fixed Effects:
##           (Intr)
## LgSUBTLCD -0.764
```

Reading the results

- `Random Effects:` information about the distribution of the model residuals, the variance, the corresponding standard deviation, and the correlation estimates associated with the random effects
- `Residual:` error variance, a distribution of deviations between the model prediction and the observed RT for each response made by a participant to a stimulus

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: logrt ~ LgSUBTLCD + (LgSUBTLCD + 1 | subjectID) + (1 |
##    Data: ML.all.correct
##
## REML criterion at convergence: -9868.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6307 -0.6324 -0.1483  0.4340  5.6132
##
## Random effects:
##  Groups    Name        Variance  Std.Dev. Corr
##  item_name (Intercept) 0.0003268 0.01808
##  subjectID (Intercept) 0.0054212 0.07363
##            LgSUBTLCD   0.0002005 0.01416  -0.63
##  Residual              0.0084333 0.09183
## Number of obs: 5257, groups:  item_name, 160; subjectID, 34
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.887997   0.015479 186.577
## LgSUBTLCD   -0.034471   0.003693  -9.333
##
## Correlation of Fixed Effects:
##           (Intr)
## LgSUBTLCD -0.764
```


Reading the results

- Plus variance terms corresponding to random differences between participants in intercepts and in the slopes of the frequency effect
- And the variance due to random differences between items in intercepts


```
## Linear mixed model fit by REML ['lmerMod']
## Formula: logrt ~ LgSUBTLCD + (LgSUBTLCD + 1 | subjectID) + (1 |
##    Data: ML.all.correct
##
## REML criterion at convergence: -9868.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6307 -0.6324 -0.1483  0.4340  5.6132
##
## Random effects:
##  Groups    Name        Variance  Std.Dev. Corr
##  item_name (Intercept) 0.0003268 0.01808
##  subjectID (Intercept) 0.0054212 0.07363
##            LgSUBTLCD   0.0002005 0.01416  -0.63
##  Residual              0.0084333 0.09183
## Number of obs: 5257, groups:  item_name, 160; subjectID, 34
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.887997   0.015479 186.577
## LgSUBTLCD   -0.034471   0.003693  -9.333
##
## Correlation of Fixed Effects:
##           (Intr)
## LgSUBTLCD -0.764
```


Reading the results

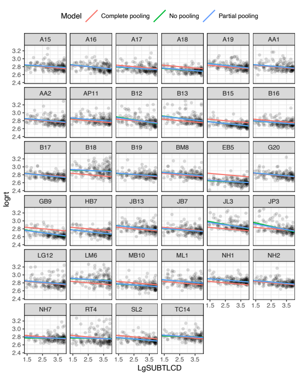
- Then we see estimates of the coefficients (of the slopes) of the fixed effects, the intercept and the slope of the `logrts \sim LgSUBTLCD` relationship
- Note that we see coefficient estimates like in a linear model summary


```
## Linear mixed model fit by REML ['lmerMod']
## Formula: logrt ~ LgSUBTLCD + (LgSUBTLCD + 1 | subjectID) + (1 |
##    Data: ML.all.correct
##
## REML criterion at convergence: -9868.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6307 -0.6324 -0.1483  0.4340  5.6132
##
## Random effects:
##  Groups    Name        Variance  Std.Dev. Corr
##  item_name (Intercept) 0.0003268 0.01808
##  subjectID (Intercept) 0.0054212 0.07363
##            LgSUBTLCD   0.0002005 0.01416  -0.63
##  Residual              0.0084333 0.09183
## Number of obs: 5257, groups:  item_name, 160; subjectID, 34
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.887997   0.015479 186.577
## LgSUBTLCD   -0.034471   0.003693  -9.333
##
## Correlation of Fixed Effects:
##           (Intr)
## LgSUBTLCD -0.764
```


What is the impact of including random effects? – Next

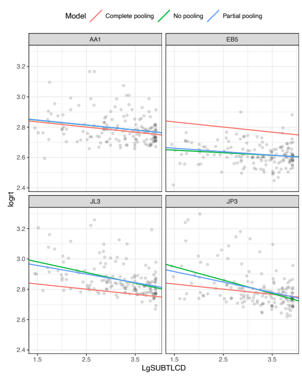
What is the impact of including random effects?

- Mixed-effects models can be understood as a method to compromise
- between ignoring the differences between groups (here, participants) as in *complete pooling*
- or focusing entirely on each group (participant) as in *no pooling*



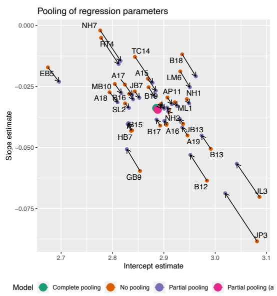
What is the impact of the incorporation of random effects?

- What happens in mixed-effects models is that we pool information
- Calculating the estimates for each participant in part based on the information we have for the whole sample (all participants, in complete pooling)
- . . . in part based on the information we have about the specific participant (in no pooling)



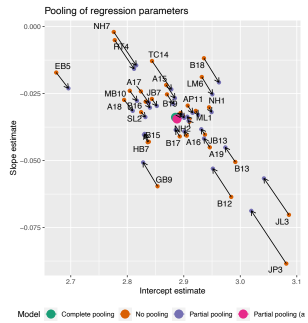
Mixed-effects models: shrinkage, regularisation

- The optimal combined estimate for a participant is termed the *Empirical Bayes* ‘estimate’ and the weighting
- Whether an estimate for a participant (in our example) is pulled (shrunk) more or less towards the overall estimate
- Will depend on the reliability of the estimate (of the intercept or the frequency effect) given by analysing that participant’s data



Mixed-effects models: shrinkage, regularisation

- What we are looking at is a form of *regularization*
- in which we use all the sources of information to take into account the variability in the data
- while not getting over-excited by extreme differences (McElreath, 2015)
- we want to see estimates pulled towards an overall average where we have little data or unreliable estimates



How does this work? Estimation

- If we knew the random effects, we could find the fixed effects estimates by minimising differences – like linear modelling
- If we knew the fixed effects – the regression coefficients – we could work out the residuals and the other random effects

How does this work? Estimation

- If we knew the random effects, we could find the fixed effects estimates by minimising differences – like linear modelling
- If we knew the fixed effects – the regression coefficients – we could work out the residuals and the other random effects

How does this work? Estimation

- If we knew the random effects, we could find the fixed effects estimates by minimising differences – like linear modelling
- If we knew the fixed effects – the regression coefficients – we could work out the residuals and the other random effects
- At the start, we know neither, but we can move between partial estimation of fixed and random effect in an *iterative approach* to *converge* on the maximum likelihood estimates of effects – when the estimates stop changing

How does this work? Estimation

- In mixed-effects models, the things that are estimated are the fixed effects (the intercept, the slope of the frequency effect, in our example)
- along with the variance and correlation terms associated with the random effects
- Strictly speaking, the partial-pooling mixed-effects ‘estimates’ of the intercept or the frequency effect for each person, are actually predictions, *Best Unbiased Linear Predictions (BLUPs)*, based on the estimates of the fixed and random effects

How do we know if including an effect helps a model to fit the data?

Researchers can compare models using the *Likelihood Ratio Test*

- The test statistic is the comparison of the likelihood of the simpler model with the more complex model
- Comparison by division $2log\frac{likelihood-complex}{likelihood-simple}$
- The likelihood ratio is compared to the χ^2 distribution for a significance test
- Assuming the null hypothesis that the simpler model is adequate
- With degrees of freedom equal to the difference in the number of parameters of the models being compared

How do we know if including an effect helps a model to fit the data?

Researchers can compare models using the *Likelihood Ratio Test*

```
ML.all.correct.lmer.REML  <- lmer(logrt ~

       LgSUBTLCD + (1|subjectID) + (1|item_name),

       data = ML.all.correct, REML = TRUE)

summary(ML.all.correct.lmer.REML)
```

- REML = TRUE – the only change to the code, requiring the change in model fitting method
- Restricted maximum likelihood used if comparing models varying in random effects

How do we know if including an effect helps a model to fit the data?

Researchers can compare models using the *Likelihood Ratio Test*

```
ML.all.correct.lmer.REML.i  <- lmer(logrt ~

       LgSUBTLCD + (1|item_name),

       data = ML.all.correct, REML = TRUE)

summary(ML.all.correct.lmer.REML.i)
```

- What if we include just the random effect of items on intercepts?

How do we know if including an effect helps a model to fit the data?

Researchers can compare models using the *Likelihood Ratio Test*

```
ML.all.correct.lmer.REML.s  <- lmer(logrt ~

       LgSUBTLCD + (1|subjectID),

       data = ML.all.correct, REML = TRUE)

summary(ML.all.correct.lmer.REML.s)
```

- What if we include just the random effect of subjects on intercepts?

How do we know if including an effect helps a model to fit the data?

Researchers can compare models using the *Likelihood Ratio Test*

```
anova(ML.all.correct.lmer.REML, ML.all.correct.lmer.REML.i, refit =
anova(ML.all.correct.lmer.REML, ML.all.correct.lmer.REML.s, refit =
```

- anova() compare listed models

How do you report a mixed-effects model?

- Explain what variables went into the analysis: say what the outcome and predictor variables were
- Report the model equation
- Report a table of coefficients: coefficient estimate; SE; t (or z); and p
- Add to that table a report of the random effects terms
- You should comment on the coefficient estimates; you may (or may not) comment on the random effects variances

Summary

1. Consolidate an understanding of how we can account for the crossed random effects of subjects or of items in data from repeated measures design studies
2. Practise fitting linear mixed-effects models incorporating random effects due to unexplained differences between subjects or between items
3. Develop an understanding of random intercepts and random slopes
4. Develop an understanding of random variances and covariances