Lesson 6, Part 1: Linear Mixed Effects Models
This Lesson’s Goals

Learn about linear mixed effects models (LMEM)

Make figures for data for LMEMs

Run some preliminary LMEMs in R

Summarise results in an R Markdown document
This Lesson’s Goals

Learn about linear mixed effects models (LMEM)

Make figures for data for LMEMs

Run some preliminary LMEMs in R

Summarise results in an R Markdown document
End of Lesson 5 Questions

But aren’t percentages really just summarized count data?

But we had to drop a bunch of Union states, isn’t that a problem?

But Alabama was missing one Democrat data point, isn’t it not balanced?

But what about the variance for ‘year’, shouldn’t we try and account for that too?

generalized linear mixed effects models
\[ y_i = a + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

**How do I add factors for random variance?**  
(i.e. things we’re not directly testing)
\[ y_i = a + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]

random effect

s = state
\[ y_i = a + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]

random intercept

s = state

\[ y_i = a + a_s + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]
\[ y_i = a + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

\[ y_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

random intercept #1
random intercept #2

s = state
y = year
In this paper we tested the effect of time on weight. A total of 50 baby chicks were included in the study.

<table>
<thead>
<tr>
<th>ANOVA language</th>
<th>LMEM language</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>weight</strong></td>
<td>dependent variable</td>
</tr>
<tr>
<td><strong>time</strong></td>
<td>independent variable</td>
</tr>
<tr>
<td><strong>baby chick</strong></td>
<td>error variable</td>
</tr>
</tbody>
</table>

mixed effects
$y_i = a + a_s + a_y + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i$

$y_i = $ specific $y$ value

$a = $ intercept

$a_s = $ random intercept #1 for specific level

$a_y = $ random intercept #2 for specific level

$b_1 = $ slope of first variable

$x_{1i} = $ specific $x$ value for first variable

$b_2 = $ slope of second variable

$x_{2i} = $ specific $x$ value for second variable

$b_3 = $ slope of third variable (interaction)

$e_i = $ random variance or the residual
\[ y_i = a + a_s + a_y + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- Democrat
- Republican
\[ y_i = a + a_S + a_y + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- **Democrat**
- **Republican**

![Graph showing percentage of votes for incumbent by country in Civil War and party of incumbent.](image)
\[ y_i = a + a_s + a_y + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- Democrat
- Republican
$$y_i = a + a_s + a_y + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i$$

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- Democrat
- Republican
\( y_i = a + a_{s} + a_{y} + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \)

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- **Democrat**
- **Republican**

**Civil War country**

- Union
- Confederacy
\[ y_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- Democrat
- Republican

Civil War country:
- Union
- Confederacy
\[ y_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

**Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent**

- **Democrat**
- **Republican**

<table>
<thead>
<tr>
<th>Year</th>
<th>Union</th>
<th>Confederacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1864</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\[ y_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- Democrat
- Republican
But, at the end of the last lesson we said this was bad, because it was a percentage of a count?
\[ y_i = a + a_s + a_y + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]

\[ \log[p/(1-p)]_i = a + a_s + a_y + b_1x_{1i} + b_2x_{2i} + b_3x_{1i}x_{2i} + e_i \]

logistic regression \(\rightarrow\) generalized linear mixed effects model
R Code (Part 1)
lme4

$$y_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i$$

lmer(perc_votes_incumbent ~ incumbent_party * civil_war
     + (1|state)
     + (1|year))

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>55.164</td>
<td>4.483</td>
<td>12.305</td>
</tr>
<tr>
<td>incumbent_partyRepublican</td>
<td>-6.773</td>
<td>6.340</td>
<td>-0.989</td>
</tr>
<tr>
<td>civil_warConfederacy</td>
<td>-8.990</td>
<td>1.444</td>
<td>-6.226</td>
</tr>
<tr>
<td>incumbent_partyRepublican:civil_warConfederacy</td>
<td>18.231</td>
<td>2.036</td>
<td>8.955</td>
</tr>
</tbody>
</table>
\[
y_i = a + a_s + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i
\]

```
lmer(perc_votes_incumbent ~ incumbent_party * civil_war + (1|state) + (1|year))
```

<table>
<thead>
<tr>
<th></th>
<th>incumbent_partyRepublican</th>
<th>civil_warConfederacy</th>
<th>incumbent_partyRepublican:civil_warConfederacy</th>
<th>(Intercept)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>-6.272727</td>
<td>-8.98988</td>
<td></td>
<td>55.16364</td>
</tr>
<tr>
<td>Arkansas</td>
<td>-6.272727</td>
<td>-8.98988</td>
<td></td>
<td>55.16364</td>
</tr>
<tr>
<td>Connecticut</td>
<td>-6.272727</td>
<td>-8.98988</td>
<td></td>
<td>55.16364</td>
</tr>
<tr>
<td>Delaware</td>
<td>-6.272727</td>
<td>-8.98988</td>
<td></td>
<td>55.16364</td>
</tr>
<tr>
<td>Florida</td>
<td>-6.272727</td>
<td>-8.98988</td>
<td></td>
<td>55.16364</td>
</tr>
<tr>
<td>Georgia</td>
<td>-6.272727</td>
<td>-8.98988</td>
<td></td>
<td>55.16364</td>
</tr>
<tr>
<td>1964</td>
<td>62.33514</td>
<td>-8.98988</td>
<td></td>
<td>18.23079</td>
</tr>
<tr>
<td>1972</td>
<td>65.93855</td>
<td>-8.98988</td>
<td></td>
<td>18.23079</td>
</tr>
<tr>
<td>1980</td>
<td>48.65702</td>
<td>-8.98988</td>
<td></td>
<td>18.23079</td>
</tr>
<tr>
<td>1984</td>
<td>60.95054</td>
<td>-8.98988</td>
<td></td>
<td>18.23079</td>
</tr>
<tr>
<td>1992</td>
<td>41.19320</td>
<td>-8.98988</td>
<td></td>
<td>18.23079</td>
</tr>
<tr>
<td>1996</td>
<td>54.63202</td>
<td>-8.98988</td>
<td></td>
<td>18.23079</td>
</tr>
</tbody>
</table>
But, in the ANOVA we accounted for the fact that a variable could be within- or between-subject?
Math (Part 2)
\[ y_i = a + a_S + a_y + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

- \( y_i \) = state
- \( a \), \( a_S \), \( a_y \), \( b_1 \), \( b_2 \), \( b_3 \), \( e_i \), and \( x_{1i} \), \( x_{2i} \) are variables in the model.
- The terms in square boxes are highlighted as random slopes.
\[ y_i = a + a_s + a_y + (b_{s1} + b_1) x_{1i} + (b_{y1} + b_2) x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

- \( y_i \): specific y value
- \( x_{1i} \): x value for variable #1
- \( x_{2i} \): x value for variable #2
- \( a \): intercept
- \( a_s \): random intercept #1
- \( a_y \): random intercept #2
- \( b_{s1} \): slope of r.e. #1
- \( b_1 \): slope of variable #1
- \( b_{y1} \): slope of r.e. #2
- \( b_2 \): slope of variable #2
- \( b_3 \): slope of variable #3
- \( e_i \): random variance
\[ y_i = a + a_s + a_y + (b_{s1}+b_1)x_{1i} + (b_{y1}+b_2)x_{2i} + b_3 x_{1i} x_{2i} + e_i \]

Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent

- Democrat
- Republican

Civil War country

- Union
- Confederacy

Rhode Island
\[ y_i = a + a_s + a_y + (b_{s1} + b_1)x_{1i} + (b_{y1} + b_2)x_{2i} + b_3x_{1i}x_{2i} + e_i \]

**Percentage of Votes for Incumbent by Country in Civil War and Party of Incumbent**

- **Democrat**
- **Republican**

![Graph showing percentage of votes for incumbent by country in Civil War and party of incumbent.](image-url)
R Code (Part 2)
\texttt{lme4}

\[ y_i = a + a_s + a_y + (b_{s1}+b_1)x_{1i} + (b_{y1}+b_2)x_{2i} + b_3x_{1i}x_{2i} + e_i \]

\texttt{lmer(perc\_votes\_incumbent ~ incumbent\_party * civil\_war + (1+incumbent\_party|state) + (1+civil\_war|year))}

\textbf{Fixed effects:}

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>55.164</td>
<td>5.591</td>
<td>9.866</td>
</tr>
<tr>
<td>incumbent_partyRepublican</td>
<td>-6.273</td>
<td>8.012</td>
<td>-0.783</td>
</tr>
<tr>
<td>civil_warConfederacy</td>
<td>-9.155</td>
<td>4.183</td>
<td>-2.189</td>
</tr>
<tr>
<td>incumbent_partyRepublican:civil_warConfederacy</td>
<td>18.396</td>
<td>6.189</td>
<td>2.972</td>
</tr>
</tbody>
</table>
\[ y_i = a + a_s + a_y + (b_{s1} + b_1)x_{1i} + (b_{y1} + b_2)x_{2i} + b_3x_{1i}x_{2i} + e_i \]

\[
\text{lmer(perc_votes_incumbent ~ incumbent_party * civil_war} \\
+ (1+incumbent_party|state) \\
+ (1+civil_war|year))
\]
### LMEM with only random intercepts

**Fixed effects:**

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>55.164</td>
<td>4.483</td>
<td>12.305</td>
</tr>
<tr>
<td>incumbent_partyRepublican</td>
<td>-6.273</td>
<td>6.340</td>
<td>-0.989</td>
</tr>
<tr>
<td>civil_warConfederacy</td>
<td>-8.990</td>
<td>1.444</td>
<td>-6.226</td>
</tr>
<tr>
<td>incumbent_partyRepublican:civil_warConfederacy</td>
<td>18.231</td>
<td>2.036</td>
<td>8.955</td>
</tr>
</tbody>
</table>

---

### LMEM with only random intercepts and slopes

**Fixed effects:**

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>55.164</td>
<td>5.591</td>
<td>9.866</td>
</tr>
<tr>
<td>incumbent_partyRepublican</td>
<td>-6.273</td>
<td>8.012</td>
<td>-0.783</td>
</tr>
<tr>
<td>civil_warConfederacy</td>
<td>-9.155</td>
<td>4.183</td>
<td>-2.189</td>
</tr>
<tr>
<td>incumbent_partyRepublican:civil_warConfederacy</td>
<td>18.396</td>
<td>6.189</td>
<td>2.972</td>
</tr>
</tbody>
</table>
Data set: Stroop Task

source: real students!
Say the color of the ink *not* the written word.

blue
Say the color of the ink *not* the written word.

blue
Say the color of the ink *not* the written word.

**blue**  
word = ink color  
congruent trial

**blue**  
word ≠ ink color  
incongruent trial
Data set: Stroop Task

Congruency: Are responses to incongruent trials less accurate and slower than to congruent trials?

Experiment half: Are responses more accurate and faster in the second half of the experiment than the first half of the experiment?

Congruency x Experiment half: Is there an interaction between these variables?

accuracy (logistic)
logit \( p_i \) = accuracy
\( x_1 \) = congruency
\( x_2 \) = experiment half
\( r_1 \) = subject
\( r_2 \) = item

reaction times (linear)
\( y_i \) = reaction times
\( x_1 \) = congruency
\( x_2 \) = experiment half
\( r_1 \) = subject
\( r_2 \) = item

source: real students!
```r
library(dplyr)
data_clean = data_results
```
```r
data_clean = data_results %>%
    rename(trial_number = SimpleRTBLock.TrialNr.)
```
```r
data_clean = data_results %>%
  rename(trial_number = SimpleRTBBlock.TrialNr.) %>%
  rename(congruency = Congruency) %>%
  rename(correct_response = StroopItem.CRESP.) %>%
  rename(given_response = StroopItem.RESP.) %>%
  rename(accuracy = StroopItem.ACC.) %>%
  rename(rt = StroopItem.RT.)
```
```r
data_clean = data_results %>%
  rename(trial_number = SimpleRTBBlock.TrialNr.) %>%
  rename(congruency = Congruency) %>%
  rename(correct_response = StroopItem.CRESP.) %>%
  rename(given_response = StroopItem.RESP.) %>%
  rename(accuracy = StroopItem.ACC.) %>%
  rename(rt = StroopItem.RT.) %>%
  select(subject_id, block, item, trial_number, congruency, correct_response, given_response, accuracy, rt)
```
RColorBrewer

cols = brewer.pal()
RColorBrewer

cols = brewer.pal(5)

call to make palette

number of colors
RColorBrewer

cols = brewer.pal(5, "PuOr")
RColorBrewer

cols = brewer.pal(5, "PuOr")

> cols
[1] "#E66101" "#FDB863" "#F7F7F7" "#B2ABD2" "#5E3C99"
RColorBrewer

cols = brewer.pal(5, "PuOr")
col_con = cols[1]

> cols
[1] "#E66101"  "#FDB863"  "#F7F7F7"  "#B2ABD2"  "#5E3C99"
RColorBrewer

cols = brewer.pal(5, "PuOr")
col_con = cols[1]
col_incon = cols[5]

> cols
[1] "#E66101" "#FDB863" "#F7F7F7" "#B2ABD2" "#5E3C99"