Course Business

• One new dataset on CourseWeb for this week

• Another add-on package you’ll probably want to install for today’s class: emmeans

• Midterm project due on CourseWeb at time of class on October 24th
  • Can run a candidate article by me if you’re not sure if it’s appropriate
**Distributed Practice!**

- Louis is a positive psychologist interested in peoples’ subjective feeling of well-being (SWB). He interviews 200 people in each of 50 countries about their SWB as well as other characteristics of their life. Louis runs the following model to relate SWB to the number of close friends that people have:
  ```r
  model1 <- lmer(SWB ~ 1 + NumFriends + (1|Country),
                  data=swb)
  ```
  However, Louis realizes that the relationship between social support and SWB might vary across countries—e.g., in more collectivistic vs. more individualistic societies. Create a `model2` that allows the relationship between `NumFriends` and `SWB` to vary across countries.
Distributed Practice!

- Louis is a positive psychologist interested in peoples’ subjective feeling of well-being (SWB). He interviews 200 people in each of 50 countries about their SWB as well as other characteristics of their life. Louis runs the following model to relate SWB to the number of close friends that people have:

  ```r
  model1 <- lmer(SWB ~ 1 + NumFriends + (1|Country), data=swb)
  ```

However, Louis realizes that the relationship between social support and SWB might vary across countries—e.g., in more collectivistic vs. more individualistic societies. Create a `model2` that allows the relationship between `NumFriends` and `SWB` to vary across countries.

- `model2 <- lmer(SWB ~ 1 + NumFriends + (1+NumFriends|Country), data=swb)`

Tip: Can read `|` as “different for each”. We want a **different** intercept (1) and `NumFriends` effect **for each** country.
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
- Simple Effects vs. Main Effects
- Post-hoc Comparisons
- Unbalanced Factors
Failures to Converge

• Remember how \texttt{lmer()} has to search for the correct parameter estimates?
• With more complex random structures, possible that we \textbf{fail to converge} on a specific estimate or hit \textbf{false convergence}

Warning messages:
1: In optwrap(optimizer, devfun, getStart(start, rho$lower, rho$pp), : convergence code 1 from bobyqa: bobyqa -- maximum number of function evaluations exceeded
2: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : unable to evaluate scaled gradient
3: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model failed to converge: degenerate Hessian with 4 negative eigenvalues

• Still returns a model, but we don’t want to trust it
Failures to Converge

- `lmer()` searched & searched but wasn’t able to figure out the right parameter estimates
- Most common with complex random effects structures
- Why does this happen?
  - Not enough data relative to the questions we’re trying to ask about it
  - Random effects structure in our `lmer` model is genuinely more complex than structure of the world

As we add subject slopes, combinatorial explosion of correlations! Is pattern of individual differences really this complex?
Failures to Converge: Solutions

• “There is no perfect solution, but there are accepted ones.” (Matuschek et al., 2017, p. 306)
Failures to Converge: Solutions

• Increase number of attempts
  • Default is 10,000
  • `model.Full <- lmer(RT ~ 1 + YearsOfStudy * WordFreq + (1 + WordFreq|Subject) + (1 + YearsOfStudy|Item), data=naming, control=lmerControl(optCtrl=list(maxfun=20000)))`

• But, if the model is too complex relative to data, spending more time doesn’t help
  • Often *not* the problem, not the solution
  • Helpful with some “failures to converge” but not “false convergence”
Failures to Converge: Solutions

- Centering or applying effects coding will often help with convergence
- A “bonus”—often, this is what you already want for your research questions
- We’ll see what this is later today 😊
Failures to Converge: Solutions

- Model may be more complicated than needed to explain data (it’s overparameterized)
  - Solution: Make it simpler
- Example: May have lots of correlation parameters, but world isn’t that complex
  - Not a complex pattern of people who show small baseline diffs, large word frequency effects, & small trial-to-trial changes

<table>
<thead>
<tr>
<th>Random effects:</th>
<th>Variance</th>
<th>Std.Dev.</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups Subject</td>
<td>(Intercept)</td>
<td>7.020e+02</td>
<td>2.649e+01</td>
</tr>
<tr>
<td></td>
<td>WordFreq</td>
<td>1.076e+04</td>
<td>1.037e+02</td>
</tr>
<tr>
<td></td>
<td>PrevTrialRT</td>
<td>7.222e+02</td>
<td>2.687e+01</td>
</tr>
<tr>
<td></td>
<td>WordFreq:PrevTrialRT</td>
<td>6.659e+02</td>
<td>2.581e+01</td>
</tr>
</tbody>
</table>

- Indicator: Correlations close to (or at) +1 or -1
Failures to Converge: Solutions

• We can often simply a model by removing unnecessary random effects
  • E.g., remove **correlation parameters**
  • Results in **near-maximal model**
  • Essentially no (or **no**) effect on tests of fixed effects (Barr et al., 2013)
• Often not of theoretical interest (except in some individual-difference studies)
• “Model tonsillectomy”

Syntax – use `||`:
• `model.NoCorr <- lmer(RT ~ YearsOfStudy * WordFreq + (1+WordFreq||Subject) + (1+YearsOfStudy||Item), data=naming)`
Failures to Converge: Solutions

- How do we know if this “tonsillectomy” is justified?
  - i.e., is the model without correlation parameters a significantly poorer fit?
  - Use `anova()` to compare the two models
  - We don’t want to exclude something that matters, so use a *liberal* criterion for including random effects (e.g., $p < .20$)
Failures to Converge: Solutions

- If it still doesn’t converge, also possible to test whether random slopes matter and only includes ones that significantly contribute
  - In experimental contexts: **Items** typically vary less than subjects, so first test **item slopes**
- This results in the **parsimonious model** or “**maximal random effects structure justified by the data**”
  - Little increase in Type I error compared to maximal model and often higher power (Matuschek et al., 2017)
Failures to Converge: Solutions

• For experimental data, a good idea to at least start with maximal model, even if we reduce it to a more parsimonious model
• Bad to neglect an important random effect
• Like having a repeated-measures design, but not running repeated-measures ANOVA

“Neither the [maximal] nor the [minimal] linear mixed models are appropriate for most repeated measures analysis. […] We can usually find middle ground, a covariance model that adequately accounts for correlation but is more parsimonious than the maximal model. Doing so allows us full control over [T]ype I error rates without needlessly sacrificing power.”
(Stroup, 2012, p. 185)
Failures to Converge: Solutions

• For experimental data, a good idea to at least start with maximal model, even if we reduce it to a more parsimonious model
  • Bad to neglect an important random effect
  • Like having a repeated-measures design, but not running repeated-measures ANOVA

• Be clear which model you are reporting and why you chose it
Failures to Converge: Solutions

• Collect more data!
  • Failure to converge can simply be a sign that we are asking too much of a small dataset
  • Just because we want the data to be able to answer this question doesn’t mean it can
Failures to Converge Due to Scaling

- I tried adding PrevTrialRT (in ms) to the naming.csv model...
  - `model.PrevTrialRT <- lmer(RT ~ YearsOfStudy * WordFreq * PrevTrialRT + (1 + WordFreq|Subject) + (1+YearsOfStudy|Item), data=naming)

  Warning message:
  Some predictor variables are on very different scales: consider rescaling

  2: In `checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, 
  Model is nearly unidentifiable: large eigenvalue ratio
  - Rescale variables?

- Algorithm doesn’t work as well if the variables are on very different scales

  YearsOfStudy measured in years    BIG effect of 1-unit change
  PrevTrialRT measured in msec     TINY effect of 1-unit change

- Latter effect basically just gets “rounded out”
Failures to Converge Due to Scaling

• I tried adding `PrevTrialRT` (in ms) to the `naming.csv` model...
  • `model.PrevTrialRT <- lmer(RT ~ YearsOfStudy * WordFreq * PrevTrialRT + (1 + WordFreq|Subject) + (1+YearsOfStudy|Item), data=naming)`

```
Warning message:
Some predictor variables are on very different scales: consider rescaling
```

2: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidentifiable: large eigenvalue ratio
- Rescale variables?

• Algorithm doesn’t work as well if the variables are on very different scales
• Simple solution: Change a variable to be a different scale
• e.g., `naming$PrevTrialRTInSeconds <- naming$PrevTrialRT / 1000`
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
- Simple Effects vs. Main Effects
- Post-hoc Comparisons
- Unbalanced Factors
The Big Picture

- Choices about how to code variables (esp. categorical ones)
  - Allow us to answer *different questions* about the data

- In most cases, multiple statistically valid ways to code
  - But, important that we actually perform the test that corresponds to *what we say we want to know*
Week 6 Sample Data: aphasia.csv

- Task: Decide whether a picture matches a sentence; measure RT
Week 6 Sample Data: aphasia.csv

- Task: Decide whether a picture matches a sentence; measure RT

“The dog was chased by the man.”
Week 6 Sample Data: aphasia.csv

- Task: Decide whether a picture matches a sentence; measure RT
- Each Item: Unique sentence w/ a unique picture
  - No picture or sentence repeats

“The dog was chased by the man.”

“The bee stung the man.”
Week 6 Sample Data: aphasia.csv

- Task: Decide whether a picture matches a sentence; measure RT
- Each Item: Unique sentence w/ a unique picture
- 16 people with aphasia and 16 healthy controls (SubjectType)
- All participants see the same sentences, which vary in SentenceLength (in words) and SentenceType (Active or Passive)
  - Active (more common): “Man bites dog.”
  - Passive: “The dog was bitten by the man.”
- Which variable(s) are between-subjects?
- Which variable(s) are within-subjects?
  - *Hint: Imagine we had only 1 subject. If we could still test the effect of a variable, it’s within-subjects.*
Week 6 Sample Data: aphasia.csv

- Task: Decide whether a picture matches a sentence; measure RT
- Each Item: Unique sentence w/ a unique picture
- 16 people with aphasia and 16 healthy controls (SubjectType)
- All participants see the same sentences, which vary in SentenceLength (in words) and SentenceType (Active or Passive)
  - Active (more common): “Man bites dog.”
  - Passive: “The dog was bitten by the man.”
- Which variable(s) are between-subjects?
  - SubjectType
- Which variable(s) are within-subjects?
  - SentenceLength and SentenceType
Week 6 Sample Data: aphasia.csv

- Task: Decide whether a picture matches a sentence; measure RT
- Each **Item**: Unique sentence w/ a unique picture
- 16 people with aphasia and 16 healthy controls (**SubjectType**)
- All participants see the same sentences, which vary in **SentenceLength** (in words) and **SentenceType** (Active or Passive)
  - Active (more common): “Man bites dog.”
  - Passive: “The dog was bitten by the man.”
- Which variable(s) are between-items?
- Which variable(s) are within-items?
  - **Hint**: Imagine we had only 1 sentence. If we could still test the effect of a variable, it’s within-items.
Week 6 Sample Data: aphasia.csv

- Task: Decide whether a picture matches a sentence; measure RT
- Each Item: Unique sentence w/ a unique picture
- 16 people with aphasia and 16 healthy controls (SubjectType)
- All participants see the same sentences, which vary in SentenceLength (in words) and SentenceType (Active or Passive)
  - Active (more common): “Man bites dog.”
  - Passive: “The dog was bitten by the man.”
- Which variable(s) are between-items?
  - SentenceLength and SentenceType
- Which variable(s) are within-items?
  - SubjectType
Interpreting Intercepts

• Let’s examine whether sentence length affects RT in this picture-verification task
  • At this stage, we don’t care about SentenceType or SubjectType
    • Common that spreadsheet contains extra, irrelevant columns
• `lengthModel.Maximal <- lmer(RT ~ 1 + SentenceLength +`  
  SUBJECT RANDOM EFFECTS
  ITEM RANDOM EFFECTS
  data = aphasia)

• Hint #1: Remember that we decided SentenceLength is a within-subjects variable

• Hint #2: Could there be a different effect of SentenceLength for each subject?
Interpreting Intercepts

- Let’s examine whether sentence length affects RT in this picture-verification task

  - At this stage, we don’t care about SentenceType or SubjectType
    - Common that spreadsheet contains extra, irrelevant columns

- `lengthModel.Maximal <- lmer(RT ~ 1 + SentenceLength + (1 + SentenceLength|Subject) + data = aphasia)`

- Sentence length is manipulated within subjects (each subject sees several different sentence lengths)
  - Possible to calculate each subject’s personal SentenceLength effect (slope)—sentence length could matter more for some people than others
  - Include random slope
Interpreting Intercepts

• Let’s examine whether **sentence length** affects RT in this picture-verification task
  • At this stage, we don’t care about **SentenceType** or **SubjectType**
    • Common that spreadsheet contains extra, irrelevant columns
  • `lengthModel.Maximal <- lmer(RT ~ 1 + SentenceLength + (1 + SentenceLength|Subject) + (1|Item), data = aphasia)`

• **Hint #1**: Remember that we decided **SentenceLength** is a between-items variable—it only differs between one sentence and another

• **Hint #2**: Could we compute the slope of a regression line relating **SentenceLength** to **RT** if we selected only a single sentence?
Interpreting Intercepts

• Let’s examine whether sentence length affects RT in this picture-verification task
  • At this stage, we don’t care about SentenceType or SubjectType
    • Common that spreadsheet contains extra, irrelevant columns
  • `lengthModel.Maximal <- lmer(RT ~ 1 + SentenceLength + (1 + SentenceLength|Subject) + (1|Item), data = aphasia)`

• Sentence length is manipulated between items (each sentence has only one length)
  • Not possible to calculate a SentenceLength effect (slope) using just one sentence
  • Don’t include random slope
Interpreting Intercepts

- Let’s examine whether **sentence length** affects RT in this picture-verification task
  - At this stage, we don’t care about `SentenceType` or `SubjectType`
    - Common that spreadsheet contains extra, irrelevant columns
  - `lengthModel.Maximal <- lmer(RT ~ 1 + SentenceLength + (1 + SentenceLength|Subject) + (1|Item), data = aphasia)`

- The statistician's cheer:
  - When I say “within,” you say “random slope!”
**Interpreting Intercepts**

- Let’s examine whether **sentence length** affects RT in this picture-verification task
  - At this stage, we don’t care about *SentenceType* or *SubjectType*
    - Common that spreadsheet contains extra, irrelevant columns
- Results:

  | Fixed effects: | Estimate | Std. Error | df | t value | Pr(>|t|) |
  |----------------|----------|------------|----|---------|--------|
  | (Intercept)    | 1215.43  | 373.96     | 30.39 | 3.250   | 0.00282 ** |
  | SentenceLength | 87.86    | 36.82      | 30.46 | 2.386   | 0.02345 *  |

\[ y = 1215 + 88 \times \text{SentenceLength} \]

- Intercept: RT is **1215** ms when sentence length is 0
- Sentence length effect: +**88** ms for each word
  - But, sentence length 0 is *impossible*. Odd to talk about.
Interpreting Intercepts

• Let’s change the model so that 0 means something
Mean Centering

• Mean sentence length is **10.00**

• Imagine we subtracted **this mean** length from **each sentence length**

<table>
<thead>
<tr>
<th>Original</th>
<th>Mean Subtracted</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traffic jam:</strong> 7</td>
<td><strong>-3</strong></td>
</tr>
<tr>
<td><strong>Chess club:</strong> 10</td>
<td><strong>0</strong></td>
</tr>
<tr>
<td><strong>Panther:</strong> 11</td>
<td><strong>1</strong></td>
</tr>
</tbody>
</table>

• **New zero** represents **mean** length
• “Mean centering”
Mean Centering

- **New zero** represents **mean** length
- “**Mean centering**”
Centering—How to Do It

- **First**, create a new variable:
  - `aphasia$SentenceLength.cen <- scale(aphasia$SentenceLength, center=TRUE, scale=FALSE)[,1]`

- **Then**, use the new variable in your model:
  - `lengthModel.cen.Maximal <- lmer(RT ~ 1 + SentenceLength.cen + (1 + SentenceLength.cen|Subject) + (1|Item), data = aphasia)`
Centering—Results

• Old model:

  \[
  y = \text{Intercept} + \text{Sentence Length} \times \text{Sentence Length} \]

  \[
  y = 2094 + 88 \times \text{Sentence Length}
  \]

  Correlation of sentence length effect with intercept is now almost 0. Indicates that we centered correctly.

• New model:

  \[
  y = \text{Intercept} + \text{Sentence Length} \times \text{Sentence Length} \]

  \[
  y = 2094 + 88 \times \text{Sentence Length}
  \]

  Intercept: RT is \textbf{2094} ms at mean sentence length

  Sentence length effect: \textbf{+88} ms for each add’l word
Centering—Results

- Intercept: RT is \textbf{2094} ms at mean sentence length
- Sentence length effect: +\textbf{88} ms for each add’l word
Centering—Results

- Both regression equations apply only to plausible sentence lengths
  - With raw sentence length, can’t have a sentence length less than 0 (no such thing as a negative # of words!)
  - With centered sentence length, can’t have a sentence length less than -10 (0 minus the mean of 10)
Which Do You Like Better?

UNCENTERED
- Good if zero is *meaningful*
- Years of study abroad, number of previous trials, number of missed classes...

CENTERED
- Good if zero is *not meaningful* or *not observed*
- Reduced correlation w/ intercept also helps with convergence (esp. in binomial models)
Other Alternatives

• It would also be possible to make 0 correspond to some other sensible/useful value
  • e.g., 0 could be the *shortest* sentence length in our set of items
  • `aphasia$SentenceLength2 <- aphasia$SentenceLength - min(aphasia$SentenceLength)`
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
  - Simple Effects vs. Main Effects
  - Post-hoc Comparisons
  - Unbalanced Factors
**Terminology**

- **Factor**: A categorical variable
  - Variables where we get counts in our R summary
  - `as.factor()` makes things categorical if they aren’t already

- **Levels**: The individual categories within a factor
  - “Active” versus “Passive”
  - “Aphasia” versus “Healthy control”
  - whether experimental or observational
Terminology

- **Factorial Design**: A design where each combination of levels appears.
- Common in experimental (and quasi-experimental) contexts!
**Terminology**

- **Factorial Design**: A design where each combination of levels appears
  - Common in experimental (and quasi-experimental) contexts!

- **Cell**: One individual combination
Introduction to Contrast Coding

• So far, we’ve been writing regression equations with numbers

\[
\text{RT} = \text{Intercept} + Y_{100} + Y_{200}
\]

• But what about active vs passive sentence?

\[
\text{RT} = \text{Intercept} + Y_{100}
\]
Introduction to Contrast Coding

• But what about active vs passive sentence?

\[ RT = \text{Intercept} + Y_{100} \]

• R’s “secret decoder wheel” assigns numerical coding schemes:
  • Variable with 2 categories (this week): Only one comparison needed
  • Variables with more categories: Multiple contrasts

Active sentence: 0
Passive sentence: 1
Introduction to Contrast Coding

• But what about *active vs passive* sentence?

\[ RT = \text{Intercept} + V_{100} \]

• R’s “secret decoder wheel” assigns numerical coding schemes

• See the current codes:
  • `contrasts(aphasia$SentenceType)`
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
  - Simple Effects vs. Main Effects
  - Post-hoc Comparisons
  - Unbalanced Factors
Treatment Coding (Dummy Coding)

- R’s default system
  - One **baseline/reference** level (category) is coded as 0
  - The other (the **treatment**) is coded as 1
- Default ordering is alphabetical: First level is 0, second is 1
  - We’ll see how to change this soon

- `contrasts(aphasia$SentenceType)`

<table>
<thead>
<tr>
<th></th>
<th>Passive</th>
<th>Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>0</td>
<td>Passive</td>
</tr>
<tr>
<td>Passive</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Active coded as 0
Passive coded as 1
Treatment Coding (Dummy Coding)

- Let’s do a model that just examines the effect of sentence type in this task:
  ```r
  SentenceTypeModel <- lmer(RT ~ 1 + SentenceType +
                           (1 + SentenceType|Subject) +
                           (1|Item),
                           data = aphasia)
  ```

- **Hint**: SentenceType varies within-subjects, but only between items
Treatment Coding (Dummy Coding)

- Let’s do a model that just examines the effect of sentence type in this task:
- `SentenceTypeModel <- lmer(RT ~ 1 + SentenceType + (1 + SentenceType|Subject) + (1|Item), data = aphasia)`
**Treatment Coding (Dummy Coding)**

- Let’s think about what the model looks like for each of our two conditions:

  \[
  RT = \gamma_{000} + \gamma_{100} \times \text{SentenceType}
  \]

  - **Active Sentences**
  
  \[
  RT = \gamma_{000} + \gamma_{100} \times \text{SentenceType}
  \]

  - **Passive Sentences**
  
  \[
  RT = \gamma_{000} + \gamma_{100} \times \text{SentenceType}
  \]
Treatment Coding (Dummy Coding)

Let’s think about what the model looks like for each of our two conditions:

- **Active Sentences**
  \[ RT = \gamma_{000} + \gamma_{100} \]
  \[ ? \]

- **Passive Sentences**
  \[ RT = \gamma_{000} + \gamma_{100} \]
  \[ ? \]
Treatment Coding (Dummy Coding)

- Let's think about what the model looks like for each of our two conditions:

\[ \text{Active Sentences} \quad RT = \gamma_{000} + \gamma_{100} \]

\[ \text{Passive Sentences} \quad RT = \gamma_{000} + \gamma_{100} \]
Treatment Coding (Dummy Coding)

• Let’s think about what the model looks like for each of our two conditions:

Active Sentences: \[ RT = \gamma_{000} \]

Intercept is just the mean RT for active sentences

Passive Sentences: \[ RT = \gamma_{000} + \gamma_{100} \]

1
Treatment Coding (Dummy Coding)

- Let’s think about what the model looks like for each of our two conditions:

Active Sentences

\[ RT = \gamma_{000} \]

Intercept is just the mean RT for active sentences

Passive Sentences

\[ RT = \gamma_{000} + \gamma_{100} \]

Sentence Type effect is the difference in RT between passive & active sentences

What is the difference between the equations for the two sentence types?
Treatment Coding Results

Intercept: RT for active sentences is 1758 ms

SentenceType: RT difference between conditions is 672 ms

• Treatment coding makes one level the baseline and compares everything to that
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
      - How to Change Codes
      - Interactions
  - Effects Coding
  - Simple Effects vs. Main Effects
  - Post-hoc Comparisons
  - Unbalanced Factors
Changing Codes

- We should think about adding SubjectType to the model. Let’s check the codes:
  - `contrasts(aphasia$SubjectType)`

- But, Control is really the baseline category here
- Assign new codes by using `<-`
  - `contrasts(aphasia$SubjectType) <- c(1,0)`
- New codes are in the order you see above & with `summary()`
Changing Codes

• Need to set codes **before** you run the model!
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
  - Interactions
- Effects Coding
- Simple Effects vs. Main Effects
- Post-hoc Comparisons
- Unbalanced Factors
Treatment Coding: Two Variables

- Now, we’d like SentenceType and SubjectType to interact:
- `Model.Maximal <- lmer(RT ~ 1 + SentenceType * SubjectType + (1 + SentenceType|Subject) + (1 + SubjectType|Item), data = aphasia)`

- **Hint #1:** Remember that we can include a random slope by subjects for **within**-subjects variables but **not** for **between**-subjects variables

- **Hint #2:** Does each subject see more than one SentenceType? Is each subject more than one SubjectType?
Treatment Coding: Two Variables

• Now, we’d like SentenceType and SubjectType to interact:
  
  Model.Maximal <- lmer(RT ~ 
  1 + SentenceType * SubjectType + 
  (1 + SentenceType|Subject) + 
  (1 + SubjectType|Item), 
  data = aphasia)

• **Hint #1:** Remember that we can include a random slope by items for within-items variables but not for between-items variables

• **Hint #2:** Is each item presented as more than one SentenceType? Is each item presented to more than one SubjectType?
Treatment Coding: Two Variables

• Now, we’d like SentenceType and SubjectType to interact:
• Model.Maximal <- lmer(RT ~ 1 + SentenceType * SubjectType + (1 + SentenceType|Subject) + (1 + SubjectType|Item), data = aphasia)
Treatment Coding: Two Variables

- Our design now has four cells:

  - Active, Control Subj.
  - Passive, Control Subj.
  - Active, Aphasics
  - Passive, Aphasics

\[
\begin{align*}
\text{RT} &= y_{000} + y_{100}\text{SentenceType} + y_{200}\text{SubjectType} + y_{1200}\text{SentenceTypeSubjectType} \\
\text{RT} &= y_{000} + y_{100}\text{SentenceType} + y_{200}\text{SubjectType} + y_{1200}\text{SentenceTypeSubjectType} \\
\text{RT} &= y_{000} + y_{100}\text{SentenceType} + y_{200}\text{SubjectType} + y_{1200}\text{SentenceTypeSubjectType} \\
\text{RT} &= y_{000} + y_{100}\text{SentenceType} + y_{200}\text{SubjectType} + y_{1200}\text{SentenceTypeSubjectType}
\end{align*}
\]
Treatment Coding: Two Variables

• Our design now has four cells:

Active, Control Subj.

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

Passive, Control Subj.

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

Active, Aphasics

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

Passive, Aphasics

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]
Treatment Coding: Two Variables

- Our design now has four cells:

Active, Control Subj.

\[ RT = \gamma_{000} \]

Passive, Control Subj.

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

Active, Aphatics

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

Passive, Aphatics

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

Intercept is the RT when all variables at their baseline: active sentence type, healthy control subject.
Treatment Coding: Two Variables

- Our design now has four cells:

  \[ RT = \gamma_{000} \]

  \[ RT = \gamma_{000} + \gamma_{100} \]

  Intercept is the RT when all variables at their baseline: active sentence type, healthy control subject.

  \[ RT = \gamma_{000} + \gamma_{100} \]

  SentenceType: Passive vs active difference for baseline healthy controls.

  \[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

  \[ \gamma_{200} = 1 + \gamma_{1200} = 0 \]

  \[ \gamma_{200} = 1 + \gamma_{1200} = 0 \]

  \[ \gamma_{200} = 1 + \gamma_{1200} = 1 \]

  Active, Control Subj.

  Passive, Control Subj.

  Active, Aphasics

  Passive, Aphasics

  SentenceType + SentenceTypeSubjectType

  SubjectType

  \[ \gamma_{200} \]
Treatment Coding: Two Variables

- Our design now has four cells:

\[ RT = \gamma_{000} \]

Active, Control Subj.

\[ RT = \gamma_{000} + \gamma_{100} \]

Sentence Type: Passive vs active difference for baseline healthy controls

\[ RT = \gamma_{000} + \gamma_{200} \]

Subject Type: Aphasia vs control difference for baseline active sentences

\[ RT = \gamma_{000} + \gamma_{100} \gamma_{200} \gamma_{1200} \]

Sentence Type + Subject Type + Sentence Type × Subject Type

Intercept is the RT when all variables at their baseline: active sentence type, healthy control subject.
**Treatment Coding: Two Variables**

- Our design now has four cells:
  - **Active, Control Subj.**
    - \( RT = \gamma_{000} \)
    - Intercept is the RT when all variables at their baseline: active sentence type, healthy control subject
  - **Passive, Control Subj.**
    - \( RT = \gamma_{000} + \gamma_{100} \)
    - SentenceType: Passive vs active difference for baseline healthy controls
  - **Active, Aphasics**
    - \( RT = \gamma_{000} + \gamma_{200} \)
    - SubjectType: Aphasia vs control difference for baseline active sentences
  - **Passive, Aphasics**
    - \( RT = \gamma_{000} + \gamma_{100} + \gamma_{200} \)
    - If no special effect of passive sentence and aphasia, we’d just have these two effects
Treatment Coding: Two Variables

- Our design now has four cells:

<table>
<thead>
<tr>
<th>Group</th>
<th>RT Formula</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active, Control Subj.</td>
<td>$RT = \gamma_{000}$</td>
<td>Intercept is the RT when all variables at their baseline: active sentence type, healthy control subject</td>
</tr>
<tr>
<td>Passive, Control Subj.</td>
<td>$RT = \gamma_{000} + \gamma_{100}$</td>
<td>SentenceType: Passive vs active difference for baseline healthy controls</td>
</tr>
<tr>
<td>Active, Aphasics</td>
<td>$RT = \gamma_{000} + \gamma_{200}$</td>
<td>SubjectType: Aphasia vs control difference for baseline active sentences</td>
</tr>
<tr>
<td>Passive, Aphasics</td>
<td>$RT = \gamma_{000} + \gamma_{100} + \gamma_{200}$</td>
<td>If no special effect of passive sentence and aphasia, we’d just have these two effects</td>
</tr>
</tbody>
</table>

If no special effect of passive sentence and aphasia, we’d just have these two effects.
Treatment Coding: Two Variables

- Our design now has four cells:

<table>
<thead>
<tr>
<th>Active, Control Subj.</th>
<th>RT = $\gamma_{000}$</th>
<th>Intercept is the RT when all variables at their baseline: active sentence type, healthy control subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive, Control Subj.</td>
<td>RT = $\gamma_{000} + \gamma_{100}$</td>
<td>SentenceType: Passive vs active difference for baseline healthy controls</td>
</tr>
<tr>
<td>Active, Aphasics</td>
<td>RT = $\gamma_{000} + \gamma_{200}$</td>
<td>SubjectType: Aphasia vs control difference for baseline active sentences</td>
</tr>
<tr>
<td>Passive, Aphasics</td>
<td>RT = $\gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200}$</td>
<td></td>
</tr>
</tbody>
</table>
Treatment Coding: Two Variables

- Our design now has four cells:

Active, Control Subj.

\[ RT = \gamma_{000} \]

Intercept is the RT when all variables at their baseline: active sentence type, healthy control subject

Passive, Control Subj.

\[ RT = \gamma_{000} + \gamma_{100} \]

SentenceType: Passive vs active difference for baseline healthy controls

Active, Aphasics

\[ RT = \gamma_{000} + \gamma_{200} \]

SubjectType: Aphasia vs control difference for baseline active sentences

Passive, Aphasics

\[ RT = \gamma_{000} + \gamma_{100} + \gamma_{200} + \gamma_{1200} \]

Interaction: Special effect of aphasia and passive sentence
**Treatment Coding: Model Results**

- **Intercept:** RT for healthy controls, active voice sentences
- Significant RT difference for passive sentences (among healthy controls)
- Not a significant RT difference for aphasics (among active sentences)
- Significant special effect of aphasia + passive sentence
Treatment Coding: Model Results

Even though the SubjectType effect is not significant here, we would *not* want to remove it from the model. It doesn’t make sense to include the interaction without the lower-order terms—the interaction is defined by what’s different from the two simple effects alone.

### Random effects:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>(Intercept)</td>
<td>47858.30</td>
<td>218.765</td>
<td></td>
</tr>
<tr>
<td>Subject</td>
<td>(Intercept)</td>
<td>43287.45</td>
<td>208.056</td>
<td>-1.00</td>
</tr>
<tr>
<td>Item</td>
<td>SubjectType1</td>
<td>55.88</td>
<td>7.475</td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>SentenceType1</td>
<td>1443.33</td>
<td>37.991</td>
<td>0.27</td>
</tr>
<tr>
<td>Residual</td>
<td>Item</td>
<td>6850.06</td>
<td>82.765</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sentence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of obs:</td>
<td>960, groups:</td>
<td>Item, 32; Subject, 30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Fixed effects:

|          | Estimate | Std. Error | df  | t value | Pr(>|t|) |
|----------|----------|------------|-----|---------|---------|
| (Intercept) | 1716.01  | 76.85      | 57.43| 22.330  | < 2e-16 *** |
| SentenceType1 | 577.17   | 78.33      | 30.94| 7.368   | 2.73e-08 *** |
| SubjectType1  | 84.52    | 76.37      | 28.03| 1.107   | 0.278   |
| SentenceType1:SubjectType1 | 188.75  | 17.71      | 28.82| 10.659  | 1.64e-11 *** |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Correlation of Fixed Effects:**

- Intr
- SntnT1
- SbjcT1
- SntncT1
- ST1

- SntncT1:ST1 -0.008 -0.258 0.165

Intercept: RT for healthy controls, active voice sentences

Significant RT difference for passive sentences (among healthy controls)

Not a significant RT difference for aphasics (among active sentences)

Significant special effect of aphasia + passive sentence
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
  - Simple Effects vs. Main Effects
  - Post-hoc Comparisons
  - Unbalanced Factors
Effects Coding (Sum Coding)

• So far, the intercept at 0 has referred to a particular baseline level

• Remember centering?
  • When we centered, we made the intercept at 0 correspond to the overall mean
Effects Coding (Sum Coding)

• We can apply centering to a factor, too

• SentenceType has:
  • 480 “Active” observations (currently 0)
  • 480 “Passive”s (currently 1)

• Mean of 0.5

• Subtracting the mean from each code gives us a new set of codes

<table>
<thead>
<tr>
<th></th>
<th>Subtract 0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>0</td>
</tr>
<tr>
<td>Passive</td>
<td>1</td>
</tr>
</tbody>
</table>
Effects Coding (Sum Coding)

• We can apply centering to a factor, too

• SentenceType has:
  • 480 “Active” observations (currently 0)
  • 480 “Passive”s (currently 1)

• Mean of 0.5

• Subtracting the mean from each code gives us a new set of codes

<table>
<thead>
<tr>
<th>Active</th>
<th>0</th>
<th>Subtract 0.5</th>
<th>Active</th>
<th>-0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive</td>
<td>1</td>
<td></td>
<td>Passive</td>
<td>0.5</td>
</tr>
</tbody>
</table>

• Effects coding (a/k/a sum coding): -0.5, 0.5
Effects Coding (Sum Coding)

• Apply effects coding (-0.5, 0.5) to our two sentence types:

Active Sentences

\[ \text{RT} = \gamma_{000} + \gamma_{100} \]

Passive Sentences

\[ \text{RT} = \gamma_{000} + \gamma_{100} \]
Effects Coding (Sum Coding)

- Apply effects coding (-0.5, 0.5) to our two sentence types:

\[ RT = \gamma_{000} + \gamma_{100} \]

Imagine subtracting the equations.

The difference between the equations for the two conditions is equal to what?

- Active Sentences: \[ RT = \gamma_{000} + \gamma_{100} \]
  - \( \gamma_{100} \) = -0.5

- Passive Sentences: \[ RT = \gamma_{000} + \gamma_{100} \]
  - \( \gamma_{100} \) = 0.5
Effects Coding (Sum Coding)

- Apply effects coding (-0.5, 0.5) to our two sentence types:

  \[ RT = \gamma_{000} + \gamma_{100} \times \text{SentenceType} \]

  The equations differ by 1 \( \gamma_{100} \)

  SentenceType effect is (still) the difference between conditions

  Intercept is always present. It’s now the \textit{mean} RT across all conditions.
Effects Coding (Sum Coding)

- Let’s apply effects coding to our aphasia data
- Old codes:
  - SENTENCETYPE
    - Active: 0
    - Passive: 1

  - SUBJECTTYPE
    - Aphasia: 1
    - Control: 0

- New codes:
  - contrasts(aphasia$SentenceType) <- c(-0.5,0.5)
  - contrasts(aphasia$SubjectType) <- c(0.5,-0.5)

- Rerun the model:
  - EffectsCoding.Maximal <- lmer(RT ~ 1 + SentenceType * SubjectType + (1 + SentenceType|Subject) + (1 + SubjectType|Item), data = aphasia)
**Effects Coding: Model Results**

Intercept: Now mean RT overall

Significant overall RT difference for passive vs active sentences (across all subjects)

Significant overall RT difference for aphasics (across all sentence types)

Significant special effect of aphasia + passive sentence

No correlation w/ intercept--we’ve successfully centered
Effects Coding: Sign Changes

• We picked one condition to be \(-0.5\) and one to be \(0.5\)
  • `contrasts(aphasia$SentenceType) <- c(-0.5,0.5)`
  • Here, Active was \(-0.5\) and Passive was \(0.5\)

• Should we worry that this affects our results?
  • Let’s try it the other way and see if we get something else
  • `contrasts(aphasia$SentenceType) <- c(0.5, -0.5)`
  • Then, re-run the model
Effects Coding: Sign Changes

• Active is $-0.5$, Passive is $0.5$:

<table>
<thead>
<tr>
<th>Fixed effects:</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2094.04</td>
<td>54.57</td>
<td>38.37</td>
</tr>
<tr>
<td>SentenceType1</td>
<td>671.54</td>
<td>76.53</td>
<td>8.78</td>
</tr>
<tr>
<td>SubjectType1</td>
<td>178.90</td>
<td>78.32</td>
<td>2.28</td>
</tr>
<tr>
<td>SentenceType1:SubjectType1</td>
<td>188.75</td>
<td>17.71</td>
<td>10.66</td>
</tr>
</tbody>
</table>

“RT 671 ms longer for Passive than for Active”

• Active is $0.5$, Passive is $-0.5$:

<table>
<thead>
<tr>
<th>Fixed effects:</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2094.04</td>
<td>54.57</td>
<td>38.37</td>
</tr>
<tr>
<td>SentenceType1</td>
<td>-671.54</td>
<td>76.53</td>
<td>-8.78</td>
</tr>
<tr>
<td>SubjectType1</td>
<td>178.90</td>
<td>78.32</td>
<td>2.28</td>
</tr>
<tr>
<td>SentenceType1:SubjectType1</td>
<td>-188.75</td>
<td>17.71</td>
<td>-10.66</td>
</tr>
</tbody>
</table>

“RT 671 ms shorter for Active than for Passive”

• Flipping the signs of the code just changes the sign of the results
  • Doesn’t affect absolute value or significance
  • Choose whichever makes sense for your question:
    • “Passive is slower than Active” vs “Active is faster than Passive”
**Effects Coding: Why -0.5 & 0.5?**

- **Passive**
  - Contrast Code: \( \frac{1}{2} \)
  - Interpretation: 1 unit change in contrast IS the difference between sentence types

- **Active**
  - Contrast Code: \( -\frac{1}{2} \)
  - Interpretation: 1 unit change in contrast IS only half the difference between levels
**Effects Coding: Why -0.5 & 0.5?**

- What if we used \((-1, 1)\) instead?
- **Doesn't** affect significance test
- **Does** make it harder to interpret the estimate
  - Parameter estimate is only half of the actual difference in means

<table>
<thead>
<tr>
<th>Fixed effects:</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2094.04</td>
<td>54.57</td>
<td>38.37</td>
</tr>
<tr>
<td>SubjectType1</td>
<td>178.90</td>
<td>78.32</td>
<td>2.28</td>
</tr>
<tr>
<td>SentenceType1</td>
<td><strong>671.54</strong></td>
<td><strong>76.53</strong></td>
<td><strong>8.78</strong></td>
</tr>
<tr>
<td>SubjectType1:SentenceType1</td>
<td><strong>188.75</strong></td>
<td><strong>17.71</strong></td>
<td><strong>10.66</strong></td>
</tr>
</tbody>
</table>

**Effects Coding: Why -0.5 & 0.5?**

<table>
<thead>
<tr>
<th>Fixed effects:</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2094.039</td>
<td>54.572</td>
<td>38.37</td>
</tr>
<tr>
<td>SubjectType1</td>
<td>178.897</td>
<td>78.322</td>
<td>2.28</td>
</tr>
<tr>
<td>SentenceType1</td>
<td><strong>335.771</strong></td>
<td><strong>38.263</strong></td>
<td><strong>8.78</strong></td>
</tr>
<tr>
<td>SubjectType1:SentenceType1</td>
<td><strong>94.378</strong></td>
<td><strong>8.854</strong></td>
<td><strong>10.66</strong></td>
</tr>
</tbody>
</table>
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
  - Simple Effects vs. Main Effects
  - Post-hoc Comparisons
  - Unbalanced Factors
**Simple vs. Main Effects**

- Treatment coding and effects coding also change our interpretation of the non-intercept effects:

- **Treatment coding** (of SentenceType):
  
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>1716.01</td>
<td>76.85</td>
<td>22.330</td>
</tr>
<tr>
<td>SubjectType1</td>
<td>84.52</td>
<td>76.37</td>
<td>1.107</td>
</tr>
</tbody>
</table>

  - Non-significant RT difference for aphasics (among active sentences)

- Effect of SubjectType *within the baseline level* of SentenceType
  
  - “Simple effect” – not a “real” main effect

- **Effects coding** (of SentenceType):
  
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2094.04</td>
<td>54.57</td>
<td>38.37</td>
</tr>
<tr>
<td>SubjectType1</td>
<td>178.90</td>
<td>78.32</td>
<td>2.28</td>
</tr>
</tbody>
</table>

  - Significant RT difference for aphasics (across all sentence types)

- Overall effect of SubjectType averaged across sentence types
  
  - “Main effect”
Simple vs. Main Effects

- Again, both of these are, in principle, reasonable questions to ask…

- In factorial designs, traditional to talk about the **main effects** averaged across other variables
  - “Main effect of aphasia,” “Overall effect of priming,” “Overall effect of study strategy,” “Main effect of ambiguity”…
  - **If you want to talk about main effects in this way, don’t use treatment / dummy coding!**

- In other designs, treatment coding may be the most appropriate!
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
- Simple Effects vs. Main Effects
- Post-hoc Comparisons
- Unbalanced Factors
Post-hoc Comparisons

• The three estimates from the model are enough to fully describe differences among conditions
• With simple effects:

<table>
<thead>
<tr>
<th>SubjectType</th>
<th>SentenceType</th>
<th>RT (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aphasia</td>
<td>Active</td>
<td>1716</td>
</tr>
<tr>
<td>Control</td>
<td>Passive</td>
<td>2293</td>
</tr>
</tbody>
</table>

PASSIVE, CONTROL
RT ≈ 2293 ms

ACTIVE, CONTROL
RT ≈ 1716 ms

Passive simple effect +577 ms
Post-hoc Comparisons

- The three estimates from the model are enough to *fully describe* differences among conditions.
- With simple effects:

<table>
<thead>
<tr>
<th>Subject Type</th>
<th>Sentence Type</th>
<th>ACTIVE, APHASIA</th>
<th>Passive, CONTROL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aphasia</td>
<td>Active</td>
<td>RT ≈ 1801 ms</td>
<td>+577 ms</td>
</tr>
<tr>
<td>Aphasia</td>
<td>Passive</td>
<td>RT ≈ 2293 ms</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>Active</td>
<td>RT ≈ 1716 ms</td>
<td>+85 ms</td>
</tr>
</tbody>
</table>

[Diagram showing the comparisons and reaction times]
**Post-hoc Comparisons**

- The three estimates from the model are enough to *fully describe* differences among conditions.
- With simple effects:

  - **ACTIVE, APHASIA**
    - RT ≈ 1801 ms
    - Aphasia simple effect +85 ms
  - **ACTIVE, CONTROL**
    - RT ≈ 1716 ms
    - Passive simple effect +577 ms
  - **PASSIVE, APHASIA**
    - RT ≈ 2547 ms
    - Interaction effect +189 ms
  - **PASSIVE, CONTROL**
    - RT ≈ 2293 ms
    - +85 ms
Post-hoc Comparisons

- But, sometimes we want to compare individual combinations (e.g., people w/ aphasia seeing active vs passive sentences)
  - i.e., individual cells
    - ACTIVE, APHASIA
      - RT ≈ 1801 ms
    - PASSIVE, APHASIA
      - RT ≈ 2547 ms
    - ACTIVE, CONTROL
      - RT ≈ 1716 ms
    - PASSIVE, CONTROL
      - RT ≈ 2293 ms
Post-hoc Comparisons: Tukey Test

• But, sometimes we want to compare individual combinations (e.g., people w/ aphasia seeing active vs passive sentences)
  • i.e., individual cells

• \texttt{emmeans(model.Maximal, pairwise~SentenceType*SubjectType)}
  
• Requires \texttt{emmeans} package to be loaded
  • \texttt{library(emmeans)}

• Which two cells don’t significantly differ?

• Uses Tukey test to correct for multiple comparisons (we’ll discuss more next week)

<table>
<thead>
<tr>
<th>contrast</th>
<th>estimate</th>
<th>SE</th>
<th>df</th>
<th>t.ratio</th>
<th>p.value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active, Aphasia - Passive, Aphasia</td>
<td>-765.91793</td>
<td>75.72157</td>
<td>31.01</td>
<td>-10.115</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Active, Aphasia - Active, Control</td>
<td>84.52172</td>
<td>76.36908</td>
<td>28.03</td>
<td>1.107</td>
<td>0.6884</td>
</tr>
<tr>
<td>Active, Aphasia - Passive, Control</td>
<td>-492.64500</td>
<td>109.50199</td>
<td>58.18</td>
<td>-4.499</td>
<td>0.0002</td>
</tr>
<tr>
<td>Passive, Aphasia - Active, Control</td>
<td>850.43965</td>
<td>109.50199</td>
<td>58.18</td>
<td>7.766</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Passive, Aphasia - Passive, Control</td>
<td>273.27293</td>
<td>81.19964</td>
<td>28.03</td>
<td>3.365</td>
<td>0.0113</td>
</tr>
<tr>
<td>Active, Control - Passive, Control</td>
<td>-577.16673</td>
<td>78.33003</td>
<td>30.94</td>
<td>-7.368</td>
<td>&lt; .0001</td>
</tr>
</tbody>
</table>

Comparisons of each pair of cells
Post-hoc Comparisons: Cell Means

- `emmeans` also returns estimated means and std. errors for each cell of the design
  - Great for descriptives write-up
  - *Estimated* means controlling for random effects (esp. relevant when dealing with unbalanced data)

<table>
<thead>
<tr>
<th>SentenceType</th>
<th>SubjectType</th>
<th>emmean</th>
<th>SE</th>
<th>df</th>
<th>lower.CL</th>
<th>upper.CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Aphasia</td>
<td>1800.528</td>
<td>75.52878</td>
<td>57.27</td>
<td>1649.300</td>
<td>1951.756</td>
</tr>
<tr>
<td>Passive</td>
<td>Aphasia</td>
<td>2566.446</td>
<td>78.00750</td>
<td>56.74</td>
<td>2410.224</td>
<td>2722.669</td>
</tr>
<tr>
<td>Active</td>
<td>Control</td>
<td>1716.007</td>
<td>76.84736</td>
<td>57.43</td>
<td>1562.147</td>
<td>1869.866</td>
</tr>
<tr>
<td>Passive</td>
<td>Control</td>
<td>2293.173</td>
<td>79.28486</td>
<td>57.13</td>
<td>2134.416</td>
<td>2451.931</td>
</tr>
</tbody>
</table>
Post-hoc Comparisons: Cell Means

• Also possible to test whether each of these estimated cell means significantly differs from 0
  • `ls_means(model.Maximal)`
  • Silly in case of RTs, but could be relevant for some other DVs (e.g., preference)

| Least Squares Means table: | Estimate | Std. Error | df | t value | lower | upper | Pr(>|t|) |
|---------------------------|----------|------------|----|---------|-------|-------|---------|
| SubjectTypeAphasia        | 2183.487 | 66.794     | 49.5| 32.690  | 2049.295 | 2317.680 | < 2.2e-16 *** |
| SubjectTypeControl        | 2004.590 | 67.542     | 50.5| 29.679  | 1868.964 | 2140.216 | < 2.2e-16 *** |
| SentenceTypeActive        | 1758.267 | 65.932     | 52.9| 26.668  | 1626.021 | 1890.513 | < 2.2e-16 *** |
| SentenceTypePassive       | 2429.810 | 67.359     | 54.5| 36.072  | 2294.788 | 2564.831 | < 2.2e-16 *** |
| SubjectTypeAphasia:SentenceTypeActive | 1800.528 | 75.529 | 57.3 | 23.839 | 1649.300 | 1951.756 | < 2.2e-16 *** |
| SubjectTypeControl:SentenceTypeActive | 1716.007 | 76.847 | 57.4 | 22.330 | 1562.147 | 1869.866 | < 2.2e-16 *** |
| SubjectTypeAphasia:SentenceTypePassive | 2566.446 | 78.007 | 56.7 | 32.900 | 2410.224 | 2722.669 | < 2.2e-16 *** |
| SubjectTypeControl:SentenceTypePassive | 2293.173 | 79.285 | 57.1 | 28.923 | 2134.416 | 2451.931 | < 2.2e-16 *** |
Post-hoc Comparisons: Marginal Means

- `emmeans` can also give us marginal means:
  - `emmeans(model.Maximal, pairwise~SubjectType)`

Now, include just one variable (for which we want marginal means)

Effect of one variable *averaging over* the other
- e.g., aphasic participants (averaging over all sentence types) vs. controls (averaging over all sentence types)
- These are what *main effects* are testing

<table>
<thead>
<tr>
<th>SentenceType</th>
<th>emmean</th>
<th>SE</th>
<th>df</th>
<th>lower.CL</th>
<th>upper.CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>1758.267</td>
<td>65.93176</td>
<td>52.94</td>
<td>1626.021</td>
<td>1890.513</td>
</tr>
<tr>
<td>Passive</td>
<td>2429.810</td>
<td>67.35936</td>
<td>54.45</td>
<td>2294.788</td>
<td>2564.831</td>
</tr>
</tbody>
</table>

Results are averaged over the levels of: SubjectType
Degrees-of-freedom method: kenward-roger
Confidence level used: 0.95
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
- Simple Effects vs. Main Effects
- Post-hoc Comparisons
- Unbalanced Factors
**Unbalanced Factors**

- Sometimes, we may have **differing numbers of observations** per level
- Possible reasons:
  - Some categories **naturally** more common
    - e.g., college majors
  - Categories may be equally common in the population, but we have **sampling error**
    - e.g., ended up 60% female participants, 40% male
  - Study was **designed** so that some conditions are more common
    - e.g., more “control” subjects than “intervention” subjects
  - We wanted equal numbers of observations, but lost some because of **errors or exclusion criteria**
    - e.g., data loss due to computer problems
    - Dropping subjects below a minimum level of performance
Weighted Coding

• “For the average student, does course size predict probability of graduation?”
  • Random sample of 200 Pitt undergrads
  • 5 are student athletes and 195 are not

• How can we make the intercept reflect the “average student”?
  • We could try to apply effects coding to the StudentAthlete variable by centering around the mean and getting $(0.5, -0.5)$, but…
Weighted Coding

• An intercept at 0 would no longer correspond to the overall mean

• As a scale, this would be totally unbalanced

• To fix balance, we need to assign a heavier weight to Athlete

But “not athlete” is actually far more common
Weighted Coding

- Change codes so the mean is 0
- `c(.975, -.025)`
- `contr.helmert.weighted()` function in my `psycholcing` package will calculate this
**Weighted Coding**

- **Weighted coding:** Change the codes so that the mean is 0 again
  - Used when the imbalance reflects something *real*
  - Like Type II sums of squares

- “For the average student, does course size predict graduation rates?”
  - Average student is *not* a student athlete, and our answer to the question about an “average student” should reflect this!
Unweighted Coding

• Oops! Our experiment loaded up the wrong image for one of our Passive sentences (“Groceries”)
  • It may have been sabotaged

• UsableItem column is No for this item

• First, can we remove this from our data?
• Some possibilities:
  • aphasia <- aphasia[aphasia$UsableItem == 'Yes', ]
  • aphasia2 <- aphasia[aphasia$UsableItem != 'No', ]
  • aphasia2 <- subset(aphasia, UsableItem == 'Yes')
  • etc.
Unweighted Coding

- Oops! Our experiment loaded up the wrong image for one of our Passive sentences ("Groceries")
- Now, there’s an imbalance, but it’s an accident and not meaningful
- In fact, we’d like to get rid of it!

> summary(aphasia2)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Item</th>
<th>SubjectType</th>
<th>SentenceType</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Astronaut</td>
<td>Aphasia:465</td>
<td>Active:480</td>
</tr>
<tr>
<td>S10</td>
<td>Bear</td>
<td>Control:465</td>
<td>Passive:450</td>
</tr>
<tr>
<td>S11</td>
<td>Boy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S12</td>
<td>Breakfast</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S13</td>
<td>Burglar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S14</td>
<td>Cheese</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Other)</td>
<td></td>
<td>(Other) :744</td>
<td>(Other) :750</td>
</tr>
</tbody>
</table>
Unweighted Coding

• Oops! Our experiment loaded up the wrong image for one of our Passive sentences ("Groceries")

• Now, there’s an imbalance, but it’s an accident and not meaningful
  • In fact, we’d like to get rid of it!

• Retain the (-0.5, 0.5) codes
  • Weights the two conditions equally—because the imbalance isn’t meaningful
  • Like Type III sums of squares
  • Probably what you want for factorial experiments
**Unbalanced Factors: Summary**

- **Weighted coding:** Change the codes so that the mean is 0
  - Use when the imbalance reflects something *real*
  - Can be done with `contr.helmert.weighted()`

Mean across each individual:

- **Unweighted coding:** Keep the codes as -0.5 and 0.5
  - Use when the imbalance is an *accident* that we want to eliminate

Mean of the two levels:

(with perfectly balanced factors, these two methods are identical)
Week 6: Main Effects & Simple Effects

- Convergence Failures
- Centering Continuous Variables
- Categorical Variables with 2 Categories
  - Treatment Coding
    - What it Means
    - How to Change Codes
    - Interactions
  - Effects Coding
  - Simple Effects vs. Main Effects
  - Post-hoc Comparisons
  - Unbalanced Factors