Course Business

• Midterm project due next Wednesday at 1:30 PM
  • Please submit on CourseWeb

• Next week’s class:
  • Continue categorical outcomes
  • Discuss current use of mixed-effects models in the literature

• Two datasets on CourseWeb for Week 8
  • We’ll work with alcohol.csv first
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
  - Implementation in R
  - Parameter Interpretation for Logit Models
    - Main effects
    - Confidence intervals
    - Interactions
- Coding the Dependent Variable
- Other Families
Distributed Practice!

- Tzipi has collected a measure of frequency of alcohol use as a function of marital status (single, married, or divorced) in several different US cities. The `head()` of this dataframe, `alcohol`, is as follows:

<table>
<thead>
<tr>
<th>SubjectID</th>
<th>City</th>
<th>MaritalStatus</th>
<th>WeeklyDrinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>S0001</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>3.770059</td>
</tr>
<tr>
<td>S0002</td>
<td>Pittsburgh</td>
<td>Divorced</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0003</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>7.016863</td>
</tr>
<tr>
<td>S0004</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>2.894105</td>
</tr>
<tr>
<td>S0005</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0006</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

- Complete the `tapply()` statement to show Tzipi the average (mean) weekly alcohol use as a function of marital status:

```r
> tapply(alcohol, ,
```

(a) `av()`

(b) `mean()`

(c) `mean()`
Distributed Practice!

- Tzipi has collected a measure of frequency of alcohol use as a function of marital status (single, married, or divorced) in several different US cities. The `head()` of this dataframe, `alcohol`, is as follows:

- Complete the `tapply()` statement to show Tzipi the average (mean) weekly alcohol use as a function of marital status:
  - `tapply(alcohol$WeeklyDrinks,`
Distributed Practice!

- Tzipi has collected a measure of frequency of alcohol use as a function of marital status (single, married, or divorced) in several different US cities. The `head()` of this dataframe, `alcohol`, is as follows:

<table>
<thead>
<tr>
<th>SubjectID</th>
<th>City</th>
<th>MaritalStatus</th>
<th>WeeklyDrinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>S0001</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>3.77059</td>
</tr>
<tr>
<td>S0002</td>
<td>Pittsburgh</td>
<td>Divorced</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0003</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>7.016863</td>
</tr>
<tr>
<td>S0004</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>2.894105</td>
</tr>
<tr>
<td>S0005</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0006</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

- Complete the `tapply()` statement to show Tzipi the average (mean) weekly alcohol use as a function of marital status:

```r
  tapply(alcohol$WeeklyDrinks, alcohol$MaritalStatus,
```
Distributed Practice!

- Tzipi has collected a measure of frequency of alcohol use as a function of marital status (single, married, or divorced) in several different US cities. The `head()` of this dataframe, `alcohol`, is as follows:

<table>
<thead>
<tr>
<th>SubjectID</th>
<th>City</th>
<th>MaritalStatus</th>
<th>WeeklyDrinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>S0001</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>3.770059</td>
</tr>
<tr>
<td>S0002</td>
<td>Pittsburgh</td>
<td>Divorced</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0003</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>7.016863</td>
</tr>
<tr>
<td>S0004</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>2.894105</td>
</tr>
<tr>
<td>S0005</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0006</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

- Complete the `tapply()` statement to show Tzipi the average (mean) weekly alcohol use as a function of marital status:

```r
tapply(alcohol$WeeklyDrinks, alcohol$MaritalStatus, mean)
```
Distributed Practice!

- Tzipi has collected a measure of frequency of alcohol use as a function of marital status (single, married, or divorced) in several different US cities. The `head()` of this dataframe, `alcohol`, is as follows:

<table>
<thead>
<tr>
<th>SubjectID</th>
<th>City</th>
<th>MaritalStatus</th>
<th>WeeklyDrinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>S0001</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>3.770059</td>
</tr>
<tr>
<td>S0002</td>
<td>Pittsburgh</td>
<td>Divorced</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0003</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>7.016863</td>
</tr>
<tr>
<td>S0004</td>
<td>Pittsburgh</td>
<td>Single</td>
<td>2.894105</td>
</tr>
<tr>
<td>S0005</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
<tr>
<td>S0006</td>
<td>Pittsburgh</td>
<td>Married</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

- Complete the `tapply()` statement to show Tzipi the average (mean) weekly alcohol use as a function of marital status:
Distributed Practice!

- Deshawn is looking at some R code sent by a collaborator for a study of threat detection (as measured by response time). The R code sets the following contrasts:

<table>
<thead>
<tr>
<th></th>
<th>[,1]</th>
<th>[,2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DissociativePTSD</td>
<td>0.33</td>
<td>0.5</td>
</tr>
<tr>
<td>NonDissociativePTSD</td>
<td>0.33</td>
<td>-0.5</td>
</tr>
<tr>
<td>NoPTSD</td>
<td>-0.67</td>
<td>0.0</td>
</tr>
</tbody>
</table>

- What comparison is performed by the first contrast? And what about the second?
Distributed Practice!

- Deshawn is looking at some R code sent by a collaborator for a study of threat detection (as measured by response time). The R code sets the following contrasts:

<table>
<thead>
<tr>
<th></th>
<th>[,1]</th>
<th>[,2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DissociativePTSD</td>
<td>0.3333333</td>
<td>0.5</td>
</tr>
<tr>
<td>NonDissociativePTSD</td>
<td>0.3333333</td>
<td>-0.5</td>
</tr>
<tr>
<td>NoPTSD</td>
<td>-0.6666667</td>
<td>0.0</td>
</tr>
</tbody>
</table>

- What comparison is performed by the first contrast? And what about the second?
  - 1\textsuperscript{st} contrast: Compares PTSD vs. no PTSD
  - 2\textsuperscript{nd} contrast: Compares dissociative PTSD to non-dissociative PTSD
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
  - Implementation in R
  - Parameter Interpretation for Logit Models
    - Main effects
    - Confidence intervals
    - Interactions
  - Coding the Dependent Variable
  - Other Families
Cued Recall

• Main week 8 dataset: cuedrecall.csv

• Cued recall task:
  • **Study phase:** See *pairs* of words
    • WOLF--PUPPY

• **Test phase:** See the first word, have to type in the second
  • WOLF--___?____
Categorical Outcomes
Categorical Outcomes

WARNING

MATH AHEAD
This Week’s Dataset

• Main week 8 dataset: cuedrecall.csv

• Cued recall task:
  • **Study phase**: See *pairs* of words
    • WOLF--PUPPY

• **Test phase**: See the first word, have to type in the second
  • WOLF--___?____
CYLINDER—CAN
CAREER—JOB
EXPERT—PROFESSOR
GAME—MONOPOLY
CYLINDER — ___?_____
EXPERT — ____?______
“Over Proportions” Approach

• On each trial, only 2 possible outcomes: target is recalled (a “hit”) or it’s forgotten (a “miss”)

• “Over proportions” approach: Calculate the proportion (or percentage) of targets recalled correctly for each subject & in each condition
  • Use that as our DV in an ANOVA or linear regression

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scott</td>
<td>0.81</td>
</tr>
<tr>
<td>Tessa</td>
<td>0.71</td>
</tr>
<tr>
<td>Mike</td>
<td>0.75</td>
</tr>
</tbody>
</table>
Problems with “Over Proportions”

• Suppose we do a regression on percentages and end up with the following model:
  \[
  \text{Percent Recalled} = 51\% + 10\% \times \text{StudyTime (per pair, in seconds)}
  \]

• If we study the word pairs for 9 seconds each, what percent of pairs does the model predict we’ll recall?
  • 141% – impossible!

• Proportions have to be between 0 and 1, but ANOVA/linear regression assume infinite tails.
Problems with “Over Proportions”

- I don’t care about predicting values! I just want to test which variables have a **significant effect**

- e.g., Does study time have a **significant effect** on whether you’ll get a “passing grade”?

**PREDICTIONS:**
**STUDY TIME = 2 s.**

- Recall 0-69%: No Pass 0.42
- Recall 70-100%: Pass 0.58

**PREDICTIONS:**
**STUDY TIME = 5 s.**

- Recall >100%: ???? 0.35
- Recall 0-69%: No Pass 0.1
- Recall 70-100%: Pass 0.55

???
Problems with “Over Proportions”

- I don’t care about predicting values! I just want to test which variables have a **significant effect**

- e.g., Does study time have a **significant effect** on whether you’ll get a “passing grade”?

- Problem: Our model assigns probability to things that can never happen
  - Means we’re underestimating the probabilities of everything that *can* happen
Solutions?

- Transform the proportions
  - e.g. arcsine transformation: \( \text{asin}(\sqrt{p}) \)
  - Still possible to predict impossible values; just happens less often
  - Kind of a kludge: “Arcsine of the square root of a proportion” doesn’t have any real-world meaning

- Even if we found a good transformation...
- Calculating a proportion over all of the items means we lose the item information!

```r
> tapply(alcohol$DoesDrink, alcohol$MaritalStatus, mean)
  Divorced  Married  Single
0.5978261 0.6416309 0.6413502
```
Solutions?

• *Transform* the proportions
  • e.g. arcsine transformation: \( \text{asin}(\sqrt{p}) \)
  • Still *possible* to predict impossible values; just happens less often
  • Kind of a kludge: “Arcsine of the square root of a proportion” doesn’t have any real-world meaning

• Even if we found a good transformation…
  • Calculating a proportion over all of the items means we *lose* the item information!

• What we’d really like is to model the actual task—each pair is either recalled or not
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
    - Introduction to Generalized LMEMs
    - Implementation in R
    - Parameter Interpretation for Logit Models
      - Main effects
      - Confidence intervals
      - Interactions
    - Coding the Dependent Variable
    - Other Families
Generalized Linear Mixed Effects Models

• With our mixed effect models, we’ve been predicting the outcome of particular trials/observations

\[ \text{RT} = \text{Intercept} + \text{Study Time} + \text{Subject} + \text{Item} \]

• But, those were for \textit{normally distributed} DVs like RT
Generalized Linear Mixed Effects Models

• With our mixed effect models, we’ve been predicting the outcome of particular trials/observations
• But, those were for normally distributed DVs
• Here, we have just 2 possible outcomes per trial
  • Clearly not a normal distribution
  • But maybe we can model this with a different distribution

Recalled or Not? = Intercept + Study Time + Subject + Item
**Binomial Distribution**

- Distribution of outcomes when one of two events (a “hit”) occurs with probability $p$

- Examples:
  - Word pair recalled or not
  - Person diagnosed with depression or not
  - High school student decides to attend college or not
  - Speaker produces active sentence or passive sentence

---

don't you like me?

- yes
- no
Generalized Linear Mixed Effects Models

- We can model recall as a binomial variable
- But, we need a way to link the linear model to 1 of 2 binomial outcomes
- Won’t work to model the probability of a hit
  - Probability bounded between 0 and 1, but linear predictor can take on any value
Never Always Tell Me the Odds

• What about the **odds** of recalling an item?
  \[
  \frac{p(\text{recalled})}{p(\text{forgotten})} = \frac{p(\text{recalled})}{1-p(\text{recalled})}
  \]

• If the probability of recall is .67, what are odds?
  • \( .67/(1-.67) = .67/.33 \approx 2 \)

• Some other odds:
  • Odds of being right-handed: \( \approx .9/.1 = 9 \)
  • Odds of identical twins: \( 1/375 \approx .003 \)
    • Odds are < 1 if the event **doesn’t happen** more often than it **does happen**
Never Always Tell Me the Odds

• What about the **odds** of recalling an item?

\[
\frac{p(\text{recalled})}{p(\text{forgotten})} = \frac{p(\text{recalled})}{1-p(\text{recalled})}
\]

• If the probability of recall is .67, what are odds?
  • \(.67/(1-.67) = .67/.33 \approx 2\)

• Some other odds:
  • Odds of being right-handed: \(\approx .9/.1 = 9\)
  • Odds of identical twins: \(1/375 \approx .003\)
  • Odds of having five fingers per hand: \(\approx 500/1\)
Never Always Tell Me the Odds

• What about the **odds** of recalling an item?

\[
\frac{p(\text{recalled})}{p(\text{forgotten})} = \frac{p(\text{recalled})}{1-p(\text{recalled})}
\]

• Try converting these probabilities into odds
  • Probability of graduating high school in the US: .92
  • Probability of a coin flip being tails: .51
  • Probability of depression sometime in your life: .17
  • Probability of detecting a gorilla walking through a crowd of people: .50
Never Always Tell Me the Odds

- What about the **odds** of recalling an item?
  \[
  \frac{p(\text{recalled})}{p(\text{forgotten})} = \frac{p(\text{recalled})}{1 - p(\text{recalled})}
  \]

- Try converting these probabilities into odds
  - Probability of graduating high school in the US: .92
    - \( \approx 11.5 \) odds you’ll graduate
  - Probability of a coin flip being tails: .51
    - \( \approx 1.04 \)
  - Probability of depression sometime in your life: .17
    - \( \approx 0.20 \)
  - Probability of detecting a gorilla walking through a crowd of people: .50
    - \( = 1.00 \)
Never Always Tell Me the Odds

• What about the odds of recalling an item?
  \[
  \frac{p(\text{recalled})}{p(\text{forgotten})} = \frac{p(\text{recalled})}{1-p(\text{recalled})}
  \]

• Using the odds in our model would be somewhat better than probabilities
  • Odds have no upper bound
    • Can have 1,000,000:1 odds!
  • But, still a lower bound at 0
Logit

• Now, let’s take the logarithm of the odds
  • Specifically, the natural log (sometimes written as \(\ln\))
  • The natural log is what we get by default from \(\log()\) in R (and in most other programming languages, too)

\[
\text{log odds} = \log\left(\frac{p(\text{recalled})}{1-p(\text{recalled})}\right)
\]

• The log odds or logit
Logit

- Now, let’s take the logarithm of the odds

\[
\text{log odds} = \log \left( \frac{p(\text{recalled})}{1 - p(\text{recalled})} \right)
\]

- The log odds or logit
- If the probability of recall is 0.2, what are the log odds of recall?
  - \( \log(0.2/(1-.2)) \)
  - \( \log(0.2/0.8) \)
  - \( \log(0.25) \)
  - -1.39
As probability of hit approaches 0, log odds approach negative infinity. No lower bound.

As probability of hit approaches 1, log odds approach infinity. No upper bound.

If probability of hit is .5 (even odds), log odds are zero.

Probabilities equidistant from .5 have log odds with the same absolute value (-1.39 and 1.39).
Logit

• Now, let’s take the logarithm of the odds

\[
\text{log odds} = \log \left( \frac{p(\text{hit})}{1-p(\text{hit})} \right)
\]

• What are the log odds when…
  • …the probability of correctly translating a word from English to Klingon is 50%?
  • …the probability that your cause of death will be a heart attack is 29%?
  • …the probability that a particular square foot of the Earth’s surface is covered with water is 71%?
Logit

Now, let’s take the logarithm of the odds

$$\text{log odds} = \log \left( \frac{p(\text{hit})}{1-p(\text{hit})} \right)$$

What are the log odds when…

• …the probability of correctly translating a word from English to Klingon is 50%?
  • 0

• …the probability that your cause of death will be a heart attack is 29%?
  • -0.90

• …the probability that a particular square foot of the Earth’s surface is covered with water is 71%?
  • 0.90
Generalized Linear Mixed Effects Models

• To make predictions about a \textit{binomial} distribution, we’ll be predicting the \textit{log odds} of a hit
  • No upper or lower bound
  • \textbf{Link function} is the \textit{logit}

\[
\log \left[ \frac{p(\text{hit})}{1-p(\text{hit})} \right] = \text{Intercept} + \text{Study Time} + \text{Subject} + \text{Item}
\]

• “\textit{Generalized linear mixed effect models}” when we use a \textbf{link function} to relate the model to a distribution other than the normal
  • Before, our link function was just the \textit{identity}
**Week 8: Categorical Outcomes**

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
    - Implementation in R
    - Parameter Interpretation for Logit Models
      - Main effects
      - Confidence intervals
      - Interactions
    - Coding the Dependent Variable
    - Other Families
For generalized linear mixed effects models, we use `glmer()`.
Part of `lme4`, so you already have it!
\textit{glmer()} \newline
\begin{itemize}
  \item \texttt{glmer()} syntax identical to \texttt{lmer()} except we add \texttt{family=binomial} argument to indicate which distribution we want
  \item Generic example:
  \begin{itemize}
    \item \texttt{glmer(DV \sim 1 + Variables + (1+Variables|RandomEffect), data=mydataframe, family=binomial)}
  \end{itemize}
\end{itemize}
cuedrecall.csv

- Let’s model our cued recall data with `glmer()`
- 120 **Subjects**, all see the same 36 **WordPairs**
- **AssocStrength** (property of **WordPairs**):
  - Two words have **Low** or **High** relation in meaning
    - VIKING—HELMET = high associative strength
    - VIKING—COLLEGE = low associative strength
- **Study Strategy** (property of **Subjects**):
  - **Maintenance** rehearsal: Repeat it over & over
  - **Elaborative** rehearsal: Relate the two words
- These are both categorical variables! How should we code them?
  - 2 x 2 design where we’re interested in the **main effect** of elaborative rehearsal (averaging over assoc. strength) & vice versa
  - Hint: We expect High **AssocStrength** & Elaborative **Strategy** to be better
Let’s model our cued recall data with `glmer()`

- 120 Subjects, all see the same 36 WordPairs
- AssocStrength (property of WordPairs):
  - Two words have **Low** or **High** relation in meaning
    - VIKING—HELMET = high associative strength
    - VIKING—COLLEGE = low associative strength
- Study Strategy (property of Subjects):
  - Maintenance rehearsal: Repeat it over & over
  - Elaborative rehearsal: Relate the two words
- These are both categorical variables! How should we code them?
  - `contrasts(cuedrecall$AssocStrength) <- ????
  - `contrasts(cuedrecall$Strategy) <- ???
Let’s model our cued recall data with `glmer()`

120 Subjects, all see the same 36 WordPairs

**AssocStrength** (property of **WordPairs**):
- Two words have **Low** or **High** relation in meaning
  - VIKING—HELMET = high associative strength
  - VIKING—COLLEGE = low associative strength

**Study Strategy** (property of **Subjects**):
- **Maintenance** rehearsal: Repeat it over & over
- **Elaborative** rehearsal: Relate the two words

These are both categorical variables! How should we code them?
- `contrasts(cuedrecall$AssocStrength) <- c(???, ???)`
- `contrasts(cuedrecall$Strategy) <- c(???, ???)`
cuedrecall.csv

- Let’s model our cued recall data with `glmer()`
- 120 Subjects, all see the same 36 WordPairs
- AssocStrength (property of WordPairs):
  - Two words have **Low** or **High** relation in meaning
    - VIKING—HELMET = high associative strength
    - VIKING—COLLEGE = low associative strength
- Study Strategy (property of Subjects):
  - **Maintenance** rehearsal: Repeat it over & over
  - **Elaborative** rehearsal: Relate the two words
- These are both categorical variables! How should we code them?
  - `contrasts(cuedrecall$AssocStrength) <- c(0.5, -0.5)`
  - `contrasts(cuedrecall$Strategy) <- c(0.5, -0.5)`
Let’s model our cued recall data with `glmer()`

- 120 Subjects, all see the same 36 WordPairs
- AssocStrength (property of WordPairs):
  - Two words have Low or High relation in meaning
    - VIKING—HELMET = high associative strength
    - VIKING—COLLEGE = low associative strength
- Study Strategy (property of Subjects):
  - Maintenance rehearsal: Repeat it over & over
  - Elaborative rehearsal: Relate the two words

Model with maximal random effects structure:

```r
model1 <- glmer(Recalled ~ 1 + AssocStrength * Strategy + (1 + AssocStrength|Subject) + (1 + Strategy|WordPair),
                data=cuedrecall, family=binomial)
```

Random slope of AssocStrength by subjects because it’s a within-subjects variable. AssocStrength effect could be different for each subject.
cuedrecall.csv

- Let’s model our cued recall data with `glmer()`
- 120 **Subjects**, all see the same 36 **WordPairs**
- **AssocStrength** (property of **WordPairs**):
  - Two words have **Low** or **High** relation in meaning
    - **VIKING**—**HELMET** = high associative strength
    - **VIKING**—**COLLEGE** = low associative strength
- Study **Strategy** (property of **Subjects**):
  - **Maintenance** rehearsal: Repeat it over & over
  - **Elaborative** rehearsal: Relate the two words

- Model with maximal random effects structure:
  - `model1 <- glmer(Recalled ~ 1 + AssocStrength * Strategy + (1 + AssocStrength|Subject) + (1 + Strategy|WordPair), data=cuedrecall, family=binomial)`

  No random slope of Strategy subjects because it’s **between-subjects**. Each subject has only 1 strategy. We can’t calculate a strategy effect separately for each subject.
Can You Spot the Differences?

What an lmer() model looked like...

```r
> summary(model1)
Linear mixed model fit by REML ['lmerMod']
Formula: RT ~ 1 + PrevTrials + FontSize + (1 | Subject) + (1 | Item)
Data: Stroop

REML criterion at convergence: 19708.1

Scaled residuals:
Min      1Q  Median      3Q     Max
-2.2998 -0.4836 -0.0889  0.3132  11.1219

Random effects:
Groups   Name     Variance  Std.Dev.
Subject  (Intercept)  1262.7  35.53
Item     (Intercept)  109.1   10.44
Residual               50705.5 225.18
Number of obs: 1440, groups: Subject, 60; Item, 6

Fixed effects:
                         Estimate Std. Error t value
(Intercept)              968.5671  17.6221    54.96
PrevTrials              -17.9604   0.8583   -20.93
FontSize                 12.7588   0.2309    55.26

Correlation of Fixed Effects:
                         (Intr) PrvTrl PrevTrials
(Intercept)         1.0000  0.551  -0.551
PrevTrials           0.5510  1.000
FontSize            -0.6680 -0.014  1.000
```
Can You Spot the Differences?

Fit by Laplace estimation (don’t need to worry about REML vs ML)

Wald z test: p values automatically given by Laplace estimation, don’t need lmerTest()

Binomial family with logit link

No residual error variance. Trial outcome can only be “recalled” or “forgotten,” so each prediction is either correct or incorrect.
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
- Implementation in R
  - Parameter Interpretation for Logit Models
    - Main effects
    - Confidence intervals
    - Interactions
  - Coding the Dependent Variable
  - Other Families
Parameter Interpretation

- Effect of **AssocStrength** has a positive sign...

| Fixed effects:          | Estimate | Std. Error | z value | Pr(>|z|) |
|-------------------------|----------|------------|---------|----------|
| (Intercept)             | 0.50485  | 0.04939    | 10.221  | < 2e-16  |
| AssocStrength1          | 0.32199  | 0.09649    | 3.337   | 0.000847 |
| Strategy1               | 0.72851  | 0.07735    | 9.418   | < 2e-16  |
| AssocStrength1:Strategy1| -0.48515 | 0.14880    | -3.261  | 0.001112 |
**Parameter Interpretation**

- Results are always framed in terms of what predicts **hits**
- `glmer`'s rule:
  - If a numerical variable, 0s are considered **misses** and 1s are considered **hits**
  - If a two-level categorical variable, the **first** category is considered a **miss** and the **second** is a **hit**
- Could use `relevel()` to reorder
  - “Forgotten” listed first, so it’s the “miss”
  - “Remembered” listed second, so it’s the “hit”
- So, + effect of associative strength means **better** recall
Parameter Interpretation

- Effect of **AssocStrength** has a positive sign…
  - It has a positive effect on recall

| Fixed effects:                  | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------------------------|----------|------------|---------|----------|
| (Intercept)                     | 0.50485  | 0.04939    | 10.221  | <2e-16   **
| AssocStrength1                  | 0.32199  | 0.09649    | 3.337   | 0.000847 ***
| Strategy1                       | 0.72851  | 0.07735    | 9.418   | <2e-16   ***
| AssocStrength1:Strategy1        | -0.48515 | 0.14880    | -3.261  | 0.001112 **

- But how should we interpret the parameter estimates?
Logarithm Review

- \( \log(10) = 2.30 \) because \( e^{2.30} = 10 \)
- “The power to which we raise \( e \) (≈ 2.72) to get 10.”
- Natural log (now standard meaning of \( \log \))
- What are…?
  - \( \log(1) \) \( \Rightarrow \) \( e^{0} = 1 \) \( \Rightarrow \log(1) = 0 \)
  - \( \log(4) \) \( \Rightarrow \) \( e^{1.39} = 4 \) \( \Rightarrow \log(4) = 1.39 \)
  - \( \log(0.25) \) \( \Rightarrow \) \( e^{-1.39} = 0.25 = \frac{1}{4} \) \( \Rightarrow \log(0.25) = -1.39 \)
  - Multiply 2 * 3, then take the log
    - Find \( \log(2) \) and \( \log(3) \), then add them
  - Things that are \textit{multiplications} become \textit{additive} in log world!
  - Because \( e^{a} \cdot e^{b} = e^{a+b} \)

\[ \begin{align*}
\log(1) &= 0 \\
\log(4) &= 1.39 \\
\log(0.25) &= -1.39 \\
\end{align*} \]
exp()

• Help! Get me out of log world!
• We can undo $\log()$ with $\exp()$
  • $\exp(3)$ means “Raise e to the exponent of 3”
  • $\exp(\log(3))$
    • Find “the power to which we raise $e$ to get 3” and then “raise $e$ to that power” (giving us 3)
• Log World turned multiplication into addition; $\exp()$ turns additions back into multiplications
  • $\exp(2+3) = \exp(2) * \exp(3)$
Parameter Interpretation

• Our model is all about logits (log odds)

| Fixed effects:         | Estimate | Std. Error | z value | Pr(>|z|) |
|------------------------|----------|------------|---------|----------|
| (Intercept)            | 0.50485  | 0.04939    | 10.221  | < 2e-16  *** |
| AssocStrength1         | 0.32199  | 0.09649    | 3.337   | 0.000847 *** |
| Strategy1              | 0.72851  | 0.07735    | 9.418   | < 2e-16  *** |
| AssocStrength1:Strategy1 | -0.48515 | 0.14880    | -3.261  | 0.001112 ** |

• What is average performance here?
  • **0.50** logits
  • One statistically correct way to interpret the model
    … but not easy to understand in real-world terms
Parameter Interpretation

- Let’s go from log odds back to regular odds
  - \( \exp() \)

- Average odds of recall are 1.65
  - So, not quite 2:1
Parameter Interpretation

• Our model is all about logits (log odds)

| Fixed effects:                  | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------------------------|----------|------------|---------|----------|
| (Intercept)                     | 0.50485  | 0.04939    | 10.221  | < 2e-16  *** |
| AssocStrength1                  | 0.32199  | 0.09649    | 3.337   | 0.000847 *** |
| Strategy1                       | 0.72851  | 0.07735    | 9.418   | < 2e-16  *** |
| AssocStrength1:Strategy1        | -0.48515 | 0.14880    | -3.261  | 0.001112 ** |

• What about the effect of study strategy?
• On average, difference between elaborative and maintenance rehearsal = **0.73** logits
Parameter Interpretation

• Let’s go from log odds back to regular odds
  • \( \exp() \)

• Effects that were additive in log odds become multiplicative in odds
  • Elaborative rehearsal increases odds by 2.08 times

• When we study COFFEE-TEA with maintenance rehearsal, our odds of recall are 3:1. What if we use elaborative rehearsal?
  • Initial odds of 3 x 2.08 increase = 6.24 (NOT 5.08!)
Parameter Interpretation

ODDS ARE NOT PROBABILITIES
ODDS ARE NOT PROBABILITIES
ODDS ARE NOT PROBABILITIES
...even if you use exp()
Parameter Interpretation

• Our model is all about logits (log odds)

| Fixed effects:                  | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------------------------|----------|------------|---------|----------|
| (Intercept)                     | 0.50485  | 0.04939    | 10.221  | < 2e-16  *** |
| AssocStrength1                  | 0.32199  | 0.09649    | 3.337   | 0.000847 *** |
| Strategy1                       | 0.72851  | 0.07735    | 9.418   | < 2e-16  *** |
| AssocStrength1:Strategy1        | -0.48515 | 0.14880    | -3.261  | 0.001112 ** |

• Describe the effect of associative strength
Parameter Interpretation

• Our model is all about logits (log odds)

| Fixed effects                  | Estimate | Std. Error | z value | Pr(>|z|) |
|--------------------------------|----------|------------|---------|----------|
| (Intercept)                    | 0.50485  | 0.04939    | 10.221  | < 2e-16  *** |
| AssocStrength1                 | 0.32199  | 0.09649    | 3.337   | 0.000847 *** |
| Strategy1                      | 0.72851  | 0.07735    | 9.418   | < 2e-16  *** |
| AssocStrength1:Strategy1       | -0.48515 | 0.14880    | -3.261  | 0.001112 ** |

• Describe the effect of associative strength
  • \( \exp(0.32) = 1.38 \)
  • High associative strength increases the odds of recall by 1.38 times
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
  - Implementation in R
- Parameter Interpretation for Logit Models
  - Main effects
  - Confidence intervals
  - Interactions
- Coding the Dependent Variable
- Other Families
Confidence Intervals

- Both our estimates and standard errors are in terms of log odds.
  - Thus, so is our confidence interval.

- 95% confidence interval for AssocStrength effect in terms of log odds:
  - Estimate +/- (1.96 * standard error)
  - 0.32 +/- (1.96 * .10)
  - 0.32 +/- .20
  - [0.12, 0.52]

- Estimate is 0.32 change in logits. 95% CI around that estimate is [0.12, 0.52]
Confidence Intervals

• Both our estimates and standard errors are in terms of log odds
  • Thus, so is our confidence interval

• 95% confidence interval for AssocStrength effect in terms of log odds
  • Estimate is 0.32 change in logits. 95% CI around that estimate is [0.12, 0.52]

• But, log odds hard to understand. Let’s use exp() to turn the endpoints of the confidence interval into odds
  • 95% CI is \( \exp(c(0.12, 0.52)) = [1.13, 1.68] \)
Confidence Intervals

• For confidence intervals around \textit{log odds}
  • As usual, we care about whether the confidence interval contains 0
  • Adding or subtracting 0 to the log odds doesn’t change it. It’s the null effect.
  • So, we’re interested in whether the estimate of the effect significantly differs from 0.

• When we transform to the \textit{odds}
  • Now, we care about whether the CI contains 1
  • Remember, effects on odds are multiplicative. \textit{Multiplying} by 1 is the null effect we test against.
  • A CI that contains 0 in log odds will always contain 1 when we transform to odds (and vice versa).
Confidence Intervals

- Compute the 95% confidence interval for **Strategy** effect in terms of **log odds**

- Then, convert it a CI on the **odds**
Confidence Intervals

- Compute the 95% confidence interval for the **Strategy** effect in terms of **log odds**
  - Estimate +/- (1.96 * standard error)
  - 0.73 +/- (1.96 * .08)
  - 0.73 +/- .16
  - [0.57, 0.89]
- Then, convert it a CI on the **odds**

| Fixed effects: | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------|----------|------------|---------|----------|
| (Intercept)   | 0.50485  | 0.04939    | 10.221  | < 2e-16  *** |
| AssocStrength1| 0.32199  | 0.09649    | 3.337   | 0.000847 *** |
| Strategy1     | 0.72851  | 0.07735    | 9.418   | < 2e-16  *** |
| AssocStrength1:Strategy1 | -0.48515 | 0.14880    | -3.261  | 0.001112 ** |
Confidence Intervals

• Compute the 95% confidence interval for Strategy effect in terms of log odds
• Estimate +/- (1.96 * standard error)
• 0.73 +/- (1.96 * .08)
• 0.73 +/- .16
• [0.57, 0.89]
• Then, convert it a CI on the odds
• \( \exp(\text{c}(0.57, 0.89)) = [1.77, 2.44] \)
Asymmetric Confidence Intervals

- Confidence interval for Strategy effect:
  - Our estimate is 2.08…
    - Compare the distance to 1.77 vs. the distance to 2.44
  - Confidence intervals are numerically \textit{asymmetric} once turned back into odds
Asymmetric Confidence Intervals

- We’re more certain about the odds for smaller/lower logits
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
  - Implementation in R
- Parameter Interpretation for Logit Models
  - Main effects
  - Confidence intervals
- Interactions
- Coding the Dependent Variable
- Other Families
Interactions

- Associative strength has a + effect on recall
- Elaborative strategy has a + effect on recall
- But, their interaction has a - coefficient
- Interpretation?:
  - “With elaborative rehearsal, associative strength matters less”
  - “If pair has high associative strength, it matters less how you study it”
    (another way of saying the same thing)

Fixed effects:

|                | Estimate | Std. Error | z value | Pr(>|z|) |
|----------------|----------|------------|---------|----------|
| (Intercept)    | 0.50485  | 0.04939    | 10.221  | < 2e-16  *** |
| AssocStrength1 | 0.32199  | 0.09649    | 3.337   | 0.000847 *** |
| Strategy1      | 0.72851  | 0.07735    | 9.418   | < 2e-16  *** |
| AssocStrength1:Strategy1 | -0.48515 | 0.14880    | -3.261  | 0.001112 ** |
Interactions

• We now understand the **sign** of the interaction
• What about the specific numeric **estimate**?
  • What does -0.48515 mean in this context?

| AssocStrength1:Strategy1 | -0.48515 | 0.14880 | -3.261 | 0.001112 ** |

• Descriptive stats: Log odds in each condition
  • Not something you have to do when running your own model—this is just to understand where the numbers come from
• High associative strength pair:
  • Elaborative rehearsal -> Increase of ≈ 0.49 logits
• Low associative strength pair:
  • Elaborative rehearsal -> Increase of ≈ 0.97 logits
Interactions

- Low associative strength pair:
  - Elaborative rehearsal -> Increase of 0.97 logits

- High associative strength pair:
  - Elaborative rehearsal -> Increase of 0.49 logits

We can compute a difference in log odds:

\[ 0.49 - 0.97 = -0.48 \]

Or an odds ratio in terms of the odds:

\[ \frac{\exp(0.49)}{\exp(0.97)} = \exp(-0.48) = 0.62 \]
Interactions

- Low associative strength pair:
  - Elaborative rehearsal -> Increase of 0.97 logits
- High associative strength pair:
  - Elaborative rehearsal -> Increase of 0.49 logits
- An odds ratio in terms of the odds:
  \[
  \frac{\exp(0.49)}{\exp(0.97)} = \exp(-0.48) = 0.62
  \]
- “For high associative strength items, the difference [or ratio] between elaborative versus maintenance rehearsal was only 0.62 times what it was for low associative strength items.”
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
  - Implementation in R
- Parameter Interpretation for Logit Models
  - Main effects
  - Confidence intervals
  - Interactions
- Coding the Dependent Variable
- Other Families
Coding the Dependent Variable

• So far, positive numbers in the results meant better recall

| Fixed effects: | Estimate | Std. Error | z value | Pr(>|z|) |
|----------------|----------|------------|---------|----------|
| (Intercept)    | 0.50485  | 0.04939    | 10.221  | < 2e-16  *** |
| AssocStrength1 | 0.32199  | 0.09649    | 3.337   | 0.000847 *** |
| Strategy1      | 0.72851  | 0.07735    | 9.418   | < 2e-16  *** |
| AssocStrength1:Strategy1 | -0.48515 | 0.14880    | -3.261  | 0.001112 ** |

• That’s because we treat **correct recall** as a 1 ("hit") and an **error** as a 0 ("miss")
• We’re looking at things that **predict recall**

```
> contrasts(cuedrecall$Recalled)

Remembered
Forgotten    0
Remembered   1
```
Coding the Dependent Variable

I don’t trust these results. What if we’d coded it the other way, with “forgotten” as 1 and “remembered” as 0? Things might be totally different!

• This is also a totally plausible coding scheme
  • Variable that tracks whether you forgot something!
• Let’s see if Evil Scott is right:
  • Step 1: Create a new variable that codes things the way Evil Scott wants

• Step 2: Re-run the model
• Step 3: ???
• Step 4: PROFIT!
Coding the Dependent Variable

This is also a totally plausible coding scheme
• Variable that tracks whether you forgot something!
• Let’s see if Evil Scott is right:
  • Step 1: Create a new variable that codes things the way Evil Scott wants
    • `cuedrecall$Forgotten <- ifelse(cuedrecall$Recalled == 'Forgotten', 1, 0)`
  • Step 2: Re-run the model
  • Step 3: ???
  • Step 4: PROFIT!

I don’t trust these results. What if we’d coded it the other way, with “forgotten” as 1 and “remembered” as 0? Things might be totally different!
**Hits and Misses**

Let’s try running our model with the new coding:

- All we’ve done is flip the signs
  - Anything that increases remembering decreases forgetting (and vice versa)
  - Remember how logits equally distant from even odds have the same absolute value?
  - Won’t affect pattern of significance
- Conclusion: What we code as 1 vs 0 doesn’t affect our conclusions (good!!)
  - Choose the coding that makes sense for your research question. Do you want to talk about “what predicts graduation” or “what predicts dropping out”?
Week 8: Categorical Outcomes

- Distributed Practice
- Generalized Linear Mixed Effects Models
  - Problems with “Over Proportions”
  - Introduction to Generalized LMEMs
  - Implementation in R
- Parameter Interpretation for Logit Models
  - Main effects
  - Confidence intervals
  - Interactions
- Coding the Dependent Variable
- Other Families
One More Thing…

- \texttt{glmer()} supports other non-normal distributions

- \texttt{family=poisson}
  - For count data
    - Example: Number of solutions you brainstormed for a problem
  - Counts range from 0 to positive infinity
  - Link is \texttt{log(count)}