Course Business

- Two new datasets on CourseWeb: vocab.csv and bpd.csv

- Midterm grades & feedback now posted!
  - Great work!
  - One small point: Consider distribution of DV (normal/Gaussian, binomial, Poisson)

- Three lectures to go
  - Today: Specialized designs
  - 3/22 & 4/5: Troubleshooting & R help
  - 3/29: No class
Course Business

• Last two weeks of the course will be final project presentations
  • For each, 10 min. + 5 min. for questions

• Sign up on CourseWeb for your preferred day (Apr. 12\textsuperscript{th} or 19\textsuperscript{th})
  • Assignments → Final Project Presentation
  • Submit a response that is just your preferred date
    • I will assign them first come, first serve
    • Assignment will be returned to you on CourseWeb after the presentation with your grade from the presentation
  • Can also view the rubric for how the presentations will be scored

• Final paper rubric also now available (due 19\textsuperscript{th})
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
  - Growth Curve Analysis
    - Main Effect
    - Random Slopes
    - Other Variables
    - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
Distributed Practice

- Elika is running an experiment in which subjects envision themselves in a number of hypothetical dating scenarios (items) and rate their relationship satisfaction in that scenario. Elika is interested both in features of the scenarios & how those features may interact with the participants’ gender. Her initial model, with a maximal random effects structure, is:

\[
\text{model1 <- lmer(Rating ~ 1 + SubjectGender * PhysicalIntimacy * EmotionalIntimacy + (1 + PhysicalIntimacy * EmotionalIntimacy|Subject) + (1 + SubjectGender|Item), data=dating)}
\]

- Unfortunately, this model did not converge. What could Elika do as her next step?
Elika is running an experiment in which subjects envision themselves in a number of hypothetical dating scenarios (items) and rate their relationship satisfaction in that scenario. Elika is interested both in features of the scenarios & how those features may interact with the participants’ gender. Her initial model, with a maximal random effects structure, is:

```r
model1 <- lmer(Rating ~ 1 + SubjectGender * PhysicalIntimacy * EmotionalIntimacy + (1 + PhysicalIntimacy * EmotionalIntimacy || Subject) + (1 + SubjectGender || Item), data=dating)
```

Unfortunately, this model did not converge. What could Elika do as her next step?

- Try taking out the correlation parameters by using `||` instead of `|`
Distributed Practice

- Elika is running an experiment in which subjects envision themselves in a number of hypothetical dating scenarios (items) and rate their relationship satisfaction in that scenario. Elika is interested both in features of the scenarios & how those features may interact with the participants’ gender. Her initial model, with a maximal random effects structure, is:

  ```r
  model1 <- lmer(Rating ~ 1 + SubjectGender * PhysicalIntimacy * EmotionalIntimacy + (1 + PhysicalIntimacy * EmotionalIntimacy || Subject) + (1|Item), data=dating)
  ```

- Unfortunately, this model did not converge. What could Elika do as her next step?
  - Variance across items is often smaller. Try removing the slope by items
    - And, use `anova()` to compare that model with only random intercepts to verify that item slope does not contribute to model fit
Distributed Practice

- Elika is running an experiment in which subjects envision themselves in a number of hypothetical dating scenarios (items) and rate their relationship satisfaction in that scenario. Elika is interested both in features of the scenarios & how those features may interact with the participants’ gender. Her initial model, with a maximal random effects structure, is:

  ```r
  model1 <- lmer(Rating ~ 1 + SubjectGender * PhysicalIntimacy * EmotionalIntimacy + (1 + PhysicalIntimacy * EmotionalIntimacy || Subject) + (1|Item), data=dating)
  ```

- Unfortunately, this model did not converge. What could Elika do as her next step?
  - Could just add more iterations (but this probably won’t be helpful)
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
  - Growth Curve Analysis
    - Main Effect
    - Random Slopes
    - Other Variables
    - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
Vocabulary Size at 2 Years Old

LEVEL 2
Sampled NEIGHBORHOODS

LEVEL 1
Sampled CHILDREN

• What kind of random-effects structure is this?
Vocabulary Size at 2 Years Old

- Level 2: Sampled neighborhoods
- Level 1: Sampled children

What kind of random-effects structure is this?
- Two levels of nesting – sample neighborhoods, then sample children inside each neighborhood
Now imagine we observed each child several different times
  • e.g., every month over the course of a year
vocab.csv: Vocabulary Size in 2nd Year of Life

- This is just another level of nesting
  - Sample neighborhoods
  - Sample children within each neighborhood
  - Sample time points within each child
Week 11: Advanced Designs

### Distributed Practice

### Longitudinal Data

- Overview
  - Growth Curve Analysis
    - Main Effect
    - Random Slopes
    - Other Variables
    - Quadratic & Higher Degrees

### Signal Detection Theory

- Why Do We Need SDT?
- Sensitivity vs. Response Bias
- Implementation
- SDT & Other Independent Variables
- Logit vs. Probit
Time as a Predictor Variable

- We can add the effect of **Time** to our model
  - Here: Months since the study started
  - Nothing “special” about time as a predictor

- `model1 <- lmer(VocabWords ~ 1 + Time + (1|Child) + (1|Neighborhood), data=vocab)`

- We need to account for the nested random effects structure … can you add appropriate **random intercepts**?
  - *Tip #1:* There are *two* levels of nesting here
  - *Tip #2:* Individual observations are nested within children, and children are nested with neighborhoods
  - *Tip #3:* Include both **Child** and **Neighborhood** differences
Time as a Predictor Variable

• We can add the effect of **Time** to our model
  • Here: Months since the study started
  • Nothing “special” about time as a predictor

```r
model1 <- lmer(VocabWords ~ 1 + Time + (1|Child) + (1|Neighborhood), data=vocab)
```

<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>4.0656</td>
<td>5.2226</td>
<td>0.78</td>
</tr>
<tr>
<td>Time</td>
<td>54.7802</td>
<td>0.9305</td>
<td>58.87</td>
</tr>
</tbody>
</table>

• Gain of about ~55 words per month
Time as a Predictor Variable

- Not necessary to have every time point represented
- Dependent variable should be on same scale across time points for this to be meaningful
- Time units don’t matter as long as they’re consistent
  - Could be hours, days, years …
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
    - Growth Curve Analysis
      - Main Effect
      - Random Slopes
      - Other Variables
      - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
• So far, we assume the same growth rate for all kids
  • Almost certainly not true!
• At level 2, we’re sampling kids both with different starting points (intercepts) and growth rates (slopes)
Longitudinal Data: Random Slopes

**RANDOM INTERCEPTS MODEL**

- Kids vary in starting point, but all acquire vocabulary at the same rate over this period.

**WITH RANDOM SLOPES**

- Allows rate of vocab acquisition to vary across kids (as well as intercept).
Can you update the model to allow the \textit{Time} effect to be \textbf{different for each} Child?

- Tip 1: This involves some type of random slope...
- Tip 2: We want a random slope of \textit{Time} by Child
Longitudinal Data: Random Slopes

- \texttt{model.Slope <- lmer(VocabWords ~ 1 + Time + (1+Time|Child) + (1|Neighborhood), data=vocab)}

In fact, LOTS of variability in the Time slope

SD is 20 words!

Mean slope is 53 words/mo, but some kids might have a slope of 73 or 33
Longitudinal Data: Random Slopes

• Would also be possible to have a random slope of Time by Neighborhood
  • If there’s clustering of growth rates at the neighborhood level

• `model.TwoSlopes <- lmer(VocabWords ~ 1 + Time + (1+Time|Child) + (1+Time|Neighborhood), data=vocab)`

• Is this are any evidence for this clustering?
  • `anova(model.Slope, model.TwoSlopes)`

|                   | DF  | AIC  | BIC  | logLik deviance Chisq Chi Df Pr(>Chisq) |
|-------------------|-----|------|------|--------------------------|----------------------|
| model.Slope       | 7   | 10253| 10271| -519.6                  | 10239                |
| model.TwoSlopes   | 9   | 10257| 10300| -519.3                  | 10239 0.6114         | 2 0.7366 n.s. |

\[ \chi^2(2) = 0.61 \]
\[ p = 0.74 \]
Week 11: Advanced Designs

- Distributed Practice
  - Longitudinal Data
    - Overview
      - Growth Curve Analysis
        - Main Effect
        - Random Slopes
      - Other Variables
      - Quadratic & Higher Degrees
  - Signal Detection Theory
    - Why Do We Need SDT?
    - Sensitivity vs. Response Bias
    - Implementation
    - SDT & Other Independent Variables
    - Logit vs. Probit
Other Variables

- We may want to include other variables in a longitudinal model:
  - Do parents frequently read picture books to the child?

- **Time + Reading**
  - Effect of Reading *invariant across time*
  - Can only affect the intercept (parallel lines)

- **Time * Reading**
  - Effect of Reading *varies with time*
  - Can affect intercept & slope
Other Variables: Results

- Model results with the interaction:

<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>5.020</td>
<td>3.120</td>
<td>1.609</td>
</tr>
<tr>
<td>Time</td>
<td>50.040</td>
<td>2.146</td>
<td>23.313</td>
</tr>
<tr>
<td>ReadingYes</td>
<td>3.908</td>
<td>4.230</td>
<td>0.924</td>
</tr>
<tr>
<td>Time:ReadingYes</td>
<td>6.703</td>
<td>3.049</td>
<td>2.198</td>
</tr>
</tbody>
</table>

Parental reading doesn’t affect vocab at time 0

But, results in faster vocab growth (amplifies + Time effect)

Growth rate for “No” group: 50.040 words / month

Growth rate for “Yes” group: 50.040 + 6.703 = 56.743 words / month

e.g., Huttenlocher et al., 1991
Other Variables

- Can be either:
  - **Time-Invariant Predictor:** *Same* across all time points within a subject
    - e.g., race/ethnicity
    - Level 2 or Level 3 variables
  - **Time-Varying Predictor:** Varies even within a subject, from one time point to another
    - e.g., hours of sleep
    - Level-1 variable
- Since R automatically figures out what’s a level-1 vs. level-2 variable, we don’t have to do anything special for either kind of variable
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
    - Growth Curve Analysis
      - Main Effect
      - Random Slopes
    - Other Variables
      - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
Growth Curve Analysis

• We’ve been assuming a linear effect of time

• But, it looks like vocab growth may accelerate
  • Growth between 2 mo. and 4 mo. is much smaller than growth between 6 mo. and 8 mo.
  • Suggests a curve / quadratic equation
Growth Curve Analysis

• Add quadratic effect (Time$^2$):
  
  \[
  \text{model.poly} \leftarrow \text{lmer}(\text{VocabWords} \sim 1 + \text{poly}(\text{Time}, \text{degree}=2, \text{raw}=\text{TRUE}) + (1 + \text{poly}(\text{Time}, \text{degree}=2, \text{raw}=\text{TRUE})|\text{Child}) + (1|\text{Neighborhood}), \text{data} = \text{vocab})
  \]

  • \text{degree=2} because we want Time$^2$
  • \text{raw=TRUE} to keep the original scale of the variables (time measured in months)

• \text{poly()} automatically adds lower-order terms as well
  • i.e., the linear term (Time)
# Growth Curve Analysis: Results

## Results:

<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>39.3533</td>
<td>2.4170</td>
<td>16.28</td>
</tr>
<tr>
<td>poly(Time, degree = 2, raw = TRUE)1</td>
<td>10.6299</td>
<td>0.6677</td>
<td>15.92</td>
</tr>
<tr>
<td>poly(Time, degree = 2, raw = TRUE)2</td>
<td>6.9948</td>
<td>0.2145</td>
<td>32.61</td>
</tr>
</tbody>
</table>

- Implied equation (approximate):
  - VocabWords = 40 + 11*Time + 7*Time²

- What are predicted values if…
  - Time=0?
  - Time=1?
  - Time=2?
Growth Curve Analysis: Results

• Results:

<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>39.3533</td>
<td>2.4170</td>
<td>16.28</td>
</tr>
<tr>
<td>poly(Time, degree = 2, raw = TRUE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10.6299</td>
<td>0.6677</td>
<td>15.92</td>
</tr>
<tr>
<td>2</td>
<td>6.9948</td>
<td>0.2145</td>
<td>32.61</td>
</tr>
</tbody>
</table>

• Implied equation (approximate):
  • VocabWords = 40 + 11*Time + 7*Time^2

• What are predicted values if...
  • Time=0?  VocabWords = 40 + (11*0) + (7*0^2) = 40
  • Time=1?  VocabWords = 40 + (11*1) + (7*1^2) = 58
  • Time=2?  VocabWords = 40 + (11*2) + (7*2^2) = 90
  • Vocab growth is accelerating (larger change from time 1 to time 2 than from time 0 to time 1)
Growth Curve Analysis
• Could go up to even higher *degrees* (Time$^3$, Time$^4$…)
  • *degree*=3 if highest exponent is 3
• Degree minus 1 = Number of **bends** in the curve
Growth Curve Analysis

- **Maximum** degree of polynomial: # of time points minus 1
  - Example: 2 time points perfectly fit by a line (degree 1). Nothing left for a quadratic term to explain.

- But, don’t want to **overfit**
  - Probably not the case that the real underlying (population) trajectory has 6 bends in it

- What degree **should** we include?
  - Theoretical considerations
  - If comparing conditions, look at mean trajectory across conditions (Mirman et al., 2008)
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
    - Growth Curve Analysis
      - Main Effect
      - Random Slopes
      - Other Variables
      - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
Tasks With Categorical Decisions

- las gatos
  - (1) Grammatical
  - (4) Ungrammatical

- The cop saw the spy with the binoculars.

- In analyzing these decisions, need to consider both overall preference for certain categories & judgments of individual items.
Study:

POTATO
SLEEP
RACCOON
WITCH
NAPKIN
BINDER
• Test:
• SLEEP
• POTATO
• BINDER
• WITCH
• RACCOON
• NAPKIN
In early memory experiments, all test probes were previously studied items. No way to distinguish a person who actually remembers everything from a person who’s realized these are ALL “old” items.
Adding “lure” items helps make the task less obvious.
But still have to interpret response to lures.
Did this person circle 50% of studied items because they remember seeing those words … or because they circled 50% of everything?
Signal Detection Theory

- For analyzing **categorical judgments**
  - Part **method for analyzing** judgments
  - Part **theory** about how people **make judgments**
- Originally developed for psychophysics
- **Purpose:**
  - Better metric properties than ANOVA on proportions (*logistic regression has already taken care of this*)
  - Distinguish **sensitivity** from **response bias**
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
    - Growth Curve Analysis
      - Main Effect
      - Random Slopes
    - Other Variables
      - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
Sensitivity vs. Response Bias

Knowing which answers are C and which aren't

Response bias

“If you’re not sure, guess C”
Sensitivity vs. Response Bias

- Imagine asking groups of second-language learners of English to judge grammaticality...
### Sensitivity vs. Response Bias

- Imagine asking groups of second-language learners of English to judge grammaticality...

<table>
<thead>
<tr>
<th>Without Intervention</th>
<th>ACCURACY</th>
<th>SAID “GRAMMATICAL”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical condition</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Ungrammatical cond.</td>
<td>20%</td>
<td>80%</td>
</tr>
</tbody>
</table>

People just judge 80% of sentences grammatical in both conditions.

This is all **response bias**—no evidence that they are **sensitive** to whether particular sentences are grammatical or not.
Similarly, an intervention could shift response bias without actually increasing sensitivity.

<table>
<thead>
<tr>
<th></th>
<th>ACCURACY</th>
<th>SAID “GRAMMATICAL”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Without Intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammatical condition</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Ungrammatical cond.</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td><strong>With Intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammatical condition</td>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>Ungrammatical cond.</td>
<td>40%</td>
<td>60%</td>
</tr>
</tbody>
</table>
**Sensitivity vs. Response Bias**

- Proportion accuracy would be misleading.
- We want an analysis that tests both subjects’ sensitivity and their response bias.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Without Intervention</th>
<th>With Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grammatical condition</strong></td>
<td>80%</td>
<td>60%</td>
</tr>
<tr>
<td><strong>Ungrammatical cond.</strong></td>
<td>20%</td>
<td>40%</td>
</tr>
<tr>
<td><strong>ACCURACY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAID “GRAMMATICAL”</td>
<td>80%</td>
<td>60%</td>
</tr>
</tbody>
</table>
Sensitivity vs. Response Bias

- Comparison to “chance” get at a similar idea
  - But, that assumes all responses equally likely
- Many experiments do balance frequency of intended responses
- But even so, bias can differ for many reasons
  - Relative frequency in experiment
  - Prior frequency in the world (“no disease” less common than “disease”)
  - Motivational factors (e.g., one error “less bad” than another)
- Not bad to have a response bias—we just need to account for it in our analysis!
Sensitivity vs. Response Bias: Examples

- We present radiologists with 20 X-rays. Half of the X-rays show lung disease and half show healthy lungs. For each X-ray, the radiologist has to judge whether lung disease is present.

- In this study, how can we define...
  - Response bias?
  - Sensitivity?
Sensitivity vs. Response Bias: Examples

• We present radiologists with 20 X-rays. Half of the X-rays show lung disease and half show healthy lungs. For each X-ray, the radiologist has to judge whether lung disease is present.

• In this study, how can we define...
  • Response bias?
    • Overall propensity to judge that lung disease is present
  • Sensitivity?
    • Does the radiologist diagnose the patient with lung disease more in the cases where the patient actually has lung disease?
Sensitivity vs. Response Bias: Examples

- We present undergraduates with a series of moral dilemmas in which they have to imagine deciding between saving 1 person’s life and saving several people’s lives. The dependent measure is how often people make the utilitarian choice to save several people. Some scenarios are less personal, and we hypothesize that people will make more utilitarian choices in these scenarios.

- In this study, how can we define...
  - Response bias?
  - Sensitivity?
Sensitivity vs. Response Bias: Examples

• We present undergraduates with a series of moral dilemmas in which they have to imagine deciding between saving 1 person’s life and saving several people’s lives. The dependent measure is how often people make the utilitarian choice to save several people. Some scenarios are less personal, and we hypothesize that people will make more utilitarian choices in these scenarios.

• In this study, how can we define…
  • Response bias?
    • Overall frequency of utilitarian judgments
  • Sensitivity?
    • Do people make more of the utilitarian judgments when the scenario is less personal?
Sensitivity vs. Response Bias: Examples

- We ask college students studying French to proofread a set of 40 French sentences, all of which contain a subject/verb agreement error. The dependent measure is whether or not the student judge the sentence as containing a subject/verb agreement error (i.e., “error” or “no error”).
- In this study, how can we define…
  - Response bias?
  - Sensitivity?
Sensitivity vs. Response Bias: Examples

• We ask college students studying French to proofread a set of 40 French sentences, all of which contain a subject/verb agreement error. The dependent measure is whether or not the student judge the sentence as containing a subject/verb agreement error (i.e., “error” or “no error”).

• In this study, how can we define…
  • Response bias?
  • Sensitivity?

Trick question!! This is like the memory test that contains only “old” items. Because the test *only* contains errors, there’s no way to tell whether a participant’s response is driven by their general bias to report errors or by noticing the error in this specific sentence. We *cannot* separate response bias from sensitivity here. Unfortunately, this limits the conclusions we can draw from this task.
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
    - Growth Curve Analysis
      - Main Effect
      - Random Slopes
      - Other Variables
      - Quadratic & Higher Degrees
  - Signal Detection Theory
    - Why Do We Need SDT?
    - Sensitivity vs. Response Bias
      - Implementation
      - SDT & Other Independent Variables
      - Logit vs. Probit
Example Study: Fraundorf, Watson, & Benjamin (2010)

Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys.

Finally, the British spotted one of the monkeys in Malaysia and planted a radio tag on it.
The British scientists spotted the endangered monkey and tagged it.
The French scientists spotted the endangered monkey and tagged it.
SDT & Mixed Effects Models

- Traditional logistic regression model:

  Correct $\sim 1 + \text{ProbeType}$

Accuracy confounds sensitivity and response bias
- Accuracy might differ across probe types just because of bias to respond true
SDT & Mixed Effects Models

- Traditional logistic regression model:
  \[
  \text{Correct} \sim 1 + \text{ProbeType}
  \]

- Signal detection model:
  \[
  \text{JudgmentMade} \sim 1 + \text{ProbeType}
  \]
Respond correctly or incorrectly?

True statement or False statement?
SDT & Mixed Effects Models

- SDT model:

  Said “TRUE”

  =

  Intercept

  +

  Probe Type is TRUE

JudgmentMade ~ 1 + ProbeType

w/ effects coding…

Baseline rate of responding TRUE.

Does item being true make you more likely to say TRUE?

Overall response bias

Sensitivity
SDT & Mixed Effects Models

- SDT model:

  Said “TRUE” = Intercept
  JudgmentMade ~ 1 + ProbeType

  w/ effects coding…

  Baseline rate of responding TRUE.
  Does item being true make you more likely to say TRUE?

Results

Overall response bias
Sensitivity
Now You Try It!

- bpd.csv
- Clinical trainees evaluating learning to diagnose borderline personal disorder (BPD). Each trainee sees 60 cases—half with BPD and half without—and makes a diagnosis for each.

- Potentially relevant columns:
  - **JudgedBPD**: Trainees’ judgment of BPD (1 yes, 0 no)
  - **HasBPD**: Whether the person in the case actually has BPD—as diagnosed by expert ("Y" or "N")
  - **Accuracy**: Was the trainees’ judgment correct? (1 yes, 0 no)
Now You Try It!

- If our memory experiment SDT analysis involved a model formula like this:

  \[ \text{JudgmentMade} \sim 1 + \text{ProbeType} + (1|\text{Subject}) \]

- Can you run a SDT model on the bpd data?
  - \textit{Tip 1:} Apply effects coding (-0.5 and 0.5) to the predictor variable!
  - \textit{Tip 2:} Should this be an \texttt{lmer} model or a \texttt{glmer} model?
Now You Try It!

- If our memory experiment SDT analysis involved a model formula like this:

  \[
  \text{JudgmentMade} \sim 1 + \text{ProbeType} + (1|\text{Subject})
  \]

- Can you run a SDT model on the `bpd` data?
  - `contrasts(bpd$HasBPD) <- c(-0.5, 0.5)`
  - `model1 <- glmer(\text{JudgedBPD} \sim 1 + \text{HasBPD} + (1|\text{Subject}), \text{family}=\text{binomial}, \text{data}=\text{bpd})`
Now You Try It!

Intercept: Overall tendency to judge people as having BPD or not
  • Response bias (here, not significant)

HasBPD: Do we get more “has BPD” judgments when the person actually has BPD?
  • Sensitivity (significant!)
Now You Try It!

Our model of the random effects is that trainees differ only in their intercept:

- They differ only in response bias … not in sensitivity

Can we also allow the sensitivity to be different for each trainee?
Now You Try It!

```r
model2 <- glmer(JudgedBPD ~ 1 + HasBPD + (1 + HasBPD|Trainee), family=binomial, data=bpd)
```

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>Trainee</td>
<td>(Intercept)</td>
<td>0.07678</td>
<td>0.27710</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HasBPD1</td>
<td>0.00124</td>
<td>0.03521</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Number of obs: 3600, groups: Trainee, 60

Fixed effects:

| Estimator     | Estimate | Std. Error | t value | Pr(>|t|) |
|---------------|----------|------------|---------|---------|
| (Intercept)   | 0.04985  | 0.04951    | 1.007   | 0.314   |
| HasBPD1       | 0.71426  | 0.06860    | 10.411  | <2e-16  | ***   |
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
  - Growth Curve Analysis
    - Main Effect
    - Random Slopes
    - Other Variables
    - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys.

Finally, the **British** spotted one of the monkeys in Malaysia and planted a radio tag on it.

---

*Emphasized or not?*

We now have an additional independent variable.
Signal detection model with another independent variable:

```
my.model <- glmer(
  JudgmentMade ~ 1 + ProbeType*Emphasis
  + (1|Trainee),
  family=binomial,
  data=memory)
```

JUDGED “TRUE” OR JUDGED “FALSE”
SDT & Other Independent Variables

- SDT model:
  - Said “TRUE”
  - Intercept
  - Probe Type is TRUE
  - Contrastive Emphasis
  - Emphasis x TRUE

\[ \text{Said “TRUE”} \]
\[ = \]
\[ \text{Intercept} \]
\[ + \]
\[ \text{Probe Type is TRUE} \]
\[ + \]
\[ \text{Contrastive Emphasis} \]
\[ + \]
\[ \text{Emphasis x TRUE} \]

w/ effects coding…

Baseline rate of responding TRUE.

- Does item being true make you more likely to say TRUE?
- Does contrastive emphasis change overall rate of saying TRUE?
- Does emphasis especially increase TRUE responses to true items?

Overall response bias

Overall sensitivity

Effect on bias

Effect on sensitivity
SDT & Other Independent Variables

- **SDT model:**

  - **Said “TRUE”**
  - **Probe Type is TRUE**
  - **Contrastive Emphasis**
  - **Emphasis x TRUE**

  Results:

  - Baseline rate of responding TRUE.
  - Does item being true make you more likely to say TRUE?
  - Does contrastive emphasis change overall rate of saying TRUE?
  - Does emphasis especially increase TRUE responses to true items?

  Overall response bias
  Overall sensitivity
  Effect on bias
  Effect on sensitivity

w/ effects coding…

- When & how do people avoid ambiguity in what they say?
- Task: Read sentences & repeat back from memory

Ambiguous sentence start: “The coach knew you…”
  - “The coach knew you since freshman year.” (knowing you)
  - “The coach knew you missed practice.” (knowing a fact)
- “The coach knew that you…”
  - “that” is optional but clarifies it’s a knowing-a-fact sentence
- Dependent measure: Do people say “that” here?

- Are people sensitive to diff. from unambiguous case?:
  - “The coach knew I…”
    - Knowing-a-person sentence would be “The coach knew me.”
- Also vary whether instructions emphasize being clear
SDT & Other Independent Variables

- SDT model:

  - **Said “that”**
  - **Intercept**
  - **Ambiguity**
  - **Instructions**
  - **Instructions x Ambiguity**

  \[ \text{Baseline rate of including “that”} \]

  \[ \text{Overall response bias} \]

  \[ \text{Overall sensitivity} \]

  \[ \text{Effect on bias} \]

  \[ \text{Effect on sensitivity} \]

  + w/ effects coding…

- Do people say “that” more for you (unambig.) than for I (ambig.)?
- Are people told to avoid ambiguity?
- Do instructions especially increase use of “that” for ambiguous items?
**SDT & Other Independent Variables**

- **SDT model:**
  
  **Said “that”**

  w/ effects coding…

  Baseline rate of including “that”

  **Intercept**

  Do people say “that” more for you (unambig.) than for I (ambig.)

  **Ambiguity**

  Are people told to avoid ambiguity?

  **Instructions**

  Do instructions especially increase use of “that” for ambiguous items?

  **Instructions x Ambiguity**

  Overall response bias

  Overall sensitivity

  Effect on bias

  Effect on sensitivity

- People NOT sensitive to whether what they’re saying is grammatically ambiguous

- Effect of emphasizing clarity is that people just add extra “that”s everywhere (whether actually needed or not)
  - Case where a change in response bias tells us something interesting about what people are doing

- Response bias is NOT just something we want to avoid / get rid of
  - Can be theoretically interesting

- Our measure of sensitivity in the SDT model is independent of response bias, so OK to look at sensitivity even if there is a response bias effect
Back to Our BPD Data…

- We’re concerned that there may be a **Gender** bias in diagnoses of BPD (e.g., Bjorklund, 2009; Skodol & Bender, 2003)

- Can you test whether **Gender** affects **response bias** and/or **sensitivity** in your model?
  - Don’t forget to apply effects coding (-0.5 and 0.5) to **Gender**
  - Which gender do we think will get more BPD diagnoses?
Back to Our BPD Data...

- We’re concerned that there may be a **Gender** bias in diagnoses of BPD (e.g., Bjorklund, 2009; Skodol & Bender, 2003)

- Can you test whether **Gender** affects response bias and/or sensitivity in your model?
  - `contrasts(bpd$Gender) <- c(0.5, -0.5)`
  - `model3 <- glmer(JudgedBPD ~ 1 + HasBPD*Gender + (1+HasBPD*Gender|Trainee), family=binomial, data=bpd)`
Back to Our BPD Data…

| Fixed effects: | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------|----------|------------|---------|----------|
| (Intercept)   | 0.0472447| 0.0519269  | 0.910   | 0.363    |
| HasBPD1       | 0.7448683| 0.0713804  | 10.435  | <2e-16 ***|
| Gender1       | 0.7947636| 0.0782717  | 10.154  | <2e-16 ***|
| HasBPD1:Gender1 | -0.0009666| 0.1459559  | -0.007  | 0.995    |

Intercept: Overall tendency to judge people as having BPD or not
- Response bias (here, not significant)

HasBPD: Do we get more “has BPD” judgments when the person actually has BPD?
- Sensitivity (significant!)

Gender: An effect of BPD on “has BPD” judgments, regardless of whether the person has BPD
- This an effect of gender on response bias!

Gender:HasBPD: Is “has BPD” larger for one gender?
- No – no effect of gender on sensitivity
Summary:
- No overall response bias to judge people as having BPD or not
- Trainees have some ability to discern which people have BPD and which don’t
- Overall bias to diagnosis more women with BPD, but doesn’t affect sensitivity to the symptoms in making the diagnosis
Week 11: Advanced Designs

- Distributed Practice
- Longitudinal Data
  - Overview
    - Growth Curve Analysis
      - Main Effect
      - Random Slopes
      - Other Variables
        - Quadratic & Higher Degrees
- Signal Detection Theory
  - Why Do We Need SDT?
  - Sensitivity vs. Response Bias
  - Implementation
  - SDT & Other Independent Variables
  - Logit vs. Probit
Logit and Probit

• How to **link** the binomial response to the continuous model predictors?

• So far, we’ve been using the logit:

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

• **Probit**: Based on the cumulative distribution function of the normal

\[
d' = \text{CDF}(\text{recall}) - \text{CDF}(1-\text{recall})
\]

Area under curve from \(-\infty\) up to this point
Logit and Probit

• Extremely similar, but logit a little less sensitive to extreme values
  • Thus, will probably get qualitatively the same results
• Which to choose?
  • Some literatures (SDT) use $d'$ units -> **Probit** model
  • Otherwise, **logit** has a somewhat easier interpretation
    • Odds / odds ratios
**Probit**

- To use the probit instead of the logit:
  - \[
  \text{model.Probit} \leftarrow \text{glmer(JudgedBPD} \sim 1 + \text{HasBPD} + (1 + \text{HasBPD}|\text{Trainee}), \right.
  \text{data=bpd, family=binomial(link=}'{\text{'probit'}}'){\text{'})}
  \]
- \(\text{(link=}'{\text{'logit'}}'\text{'})\) is the same as the default model

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial (probit)
Formula: JudgedBPD ~ 1 + HasBPD + (1 + HasBPD | Trainee)
Data: bpd