Interactions among Continuous Predictors

- Today's Class:
  - Simple main effects within two-way interactions
  - Conquering TEST/ESTIMATE/LINCOM statements
  - Regions of significance
  - Three-way interactions (and beyond…)

Representing the Effects of Predictors

- From now on, we will think carefully about exactly **<u>how</u>** the **predictor variables** are entered into the **model for the means** (i.e., by which a predicted outcome is created for each person)

- Why don't people always care? Because the scale of predictors:
  - Does NOT affect the amount of outcome variance accounted for (R^2)
  - Does NOT affect the outcomes values predicted by the model for the means (so long as the same predictor fixed effects are included)

- Why should this matter to us?
  - ***Because the Intercept = expected outcome value when X = 0***
  - Can end up with nonsense values for intercept if X = 0 isn't in the data, so we need to change the scale of the predictors to include 0
  - Scaling becomes more important **once interactions are included** or once random intercepts are included (i.e., variability around fixed intercept)

Adjusting the Scale of Predictors

- For **continuous** (quantitative) predictors, **we** will make the intercept interpretable by **centering**:
  - **Centering** = subtract a constant from each variable value so that **the 0 value** falls within the range of the new centered predictor variable
  - Typical → Center around predictor's mean: $\text{Centered } X_1 = X_1 - \overline{X_1}$
    - Intercept is then expected outcome for "average X_1 person"
  - Better → Center around meaningful constant *C*: $\text{Centered } X_1 = X_1 - C$
    - Intercept is then expected outcome for person with that constant (even 0 may be ok)

- For **binary** predictors, it can be more convenient to treat them as "continuous" than as "categorical" by doing manual coding:
  - Accomplished via "dummy coding" (aka, "reference group coding")
    → Two-group example using *Gender*: 0 = Men, 1 = Women
    (or 0 = Women, 1 = Men)

Interactions: $y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i + e_i$

- **Interaction = Moderation:** the effect of a predictor depends on the value of the interacting predictor
  - Either predictor can be "the moderator" (interpretive distinction only)
- <u>Interactions can always be evaluated</u> for any combination of categorical and continuous predictors, although traditionally…
  - <u>In "ANOVA":</u> By default, all possible interactions are estimated
    - Software does this for you; oddly enough, nonsignificant interactions usually still are kept in the model (even if only significant interactions are interpreted)
  - <u>In "ANCOVA":</u> Continuous predictors ("covariates") do not get to be part of interaction terms → "homogeneity of regression assumption"
    - There is no reason to assume this – it is a testable hypothesis!
  - <u>In "Regression":</u> No default – the effects of predictors are as you specify
    - Requires most thought, but gets annoying because in regression programs you usually have to manually create the interaction as an observed variable:
    - e.g., XZinteraction = centeredX * centeredZ

Interaction variables are created on the fly in MIXED instead! ☺

Main Effects of Predictors within Interactions

- Main effects of predictors within interactions should remain in the model regardless of whether or not they are significant
  - An interaction is an over-additive (enhancing) or under-additive (dampening) effect, so *what it is additive to* must be included

- **The role of a two-way interaction is to adjust its main effects…**
- However, the idea of a "main effect" no longer applies… each main effect is ***conditional*** on the interacting predictor = 0

- e.g., Model of Y = W, X, Z, X*Z:
  - The effect of W is still a "main effect" because it is not part of an interaction
  - The effect of X is now the conditional main effect of X *specifically when Z=0*
  - The effect of Z is now the conditional main effect of Z *specifically when X=0*

- Note that this is a different type of conditionality than just "holding the other predictors constant" (which means constant at any value)
  - Constant at the 0 value of the interacting predictor(s), specifically

Interactions: Why 0 Matters

- Y = Student achievement (GPA as percentage out of 100)
  X = Parent attitudes about education (measured on 1-5 scale)
  Z = Father's education level (measured in years of education)
- $\mathbf{GPA_i = \beta_0 + (\beta_1 * Att_i) + (\beta_2 * Ed_i) + (\beta_3 * Att_i * Ed_i) + e_i}$
  $\mathbf{GPA_i = 30 + (1 * Att_i) + (2 * Ed_i) + (0.5 * Att_i * Ed_i) + e_i}$
- **Interpret β_0:**
- **Interpret β_1:**
- **Interpret β_2:**
- **Interpret β_3: Attitude as Moderator:**

  Education as Moderator:

- **Predicted GPA** for **attitude of 3** and **Ed of 12**?

  $\mathbf{75 = 30 + 1*(3) + 2*(12) + 0.5*(3)*(12)}$

Model-Implied Simple Main Effects

- **Original:** $GPA_i = \beta_0 + (\beta_1 * Att_i) + (\beta_2 * Ed_i) + (\beta_3 * Att_i * Ed_i) + e_i$
  $GPA_i = 30 + (1 * Att_i) + (2 * Ed_i) + (0.5 * Att_i * Ed_i) + e_i$
- Given any values of the predictor variables, the model equation provides predictions for:
  - Value of outcome (model-implied intercept for non-zero predictor values)
  - Any conditional (simple) main effects implied by an interaction term
  - **Simple (Conditional) Main Effect = what it is + what *modifies* it**

- Step 1: **Identify** all terms in model involving the predictor of interest
  - e.g., Effect of Attitudes comes from: $\beta_1 * Att_i + \beta_3 * Att_i * Ed_i$
- Step 2: **Factor out** common predictor variable
  - Start with $[\beta_1 * Att_i + \beta_3 * Att_i * Ed_i] \rightarrow [Att_i (\beta_1 + \beta_3 * Ed_i)] \rightarrow Att_i$ **(new** β_1**)**
  - Value given by ( ) is then the model-implied coefficient for the predictor
- Step 3: **Calculate** model-implied simple effect and SE
  - Let's try it for **a new reference point of attitude = 3 and education = 12**

Model-Implied Simple Main Effects

- Old Equation using uncentered predictors:
  $\mathbf{GPA_i = \beta_0 + (\beta_1 * Att_i) + (\beta_2 * Ed_i) + (\beta_3 * Att_i * Ed_i) + e_i}$
  $\mathbf{GPA_i = 30 + (1 * Att_i) + (2 * Ed_i) + (0.5 * Att_i * Ed_i) + e_i}$
- New equation using centered predictors ($Att_i - 3$ and $Ed_i - 12$):
  $\mathbf{GPA_i = \_\_ + \_\_\_ * (Att_i - 3) + \_\_\_ * (Ed_i - 12) + \_\_\_ * (Att_i - 3) * (Ed_i - 12) + e_i}$

- **Intercept: expected value of GPA when $Att_i = 3$ and $Ed_i = 12$**
  $\mathbf{\beta_0 = 75}$
- **Simple main effect of Att if $Ed_i = 12$**
  $\mathbf{\beta_1 * Att_i + \beta_3 * Att_i * Ed_i \rightarrow Att_i(\beta_1 + \beta_3 * Ed_i) \rightarrow Att_i(1 + 0.5 * 12)}$
- **Simple main effect of Ed if $Att_i = 3$**
  $\mathbf{\beta_2 * Ed_i + \beta_3 * Att_i * Ed_i \rightarrow Ed_i(\beta_2 + \beta_3 * Att_i) \rightarrow Ed_i(2 + 0.5 * 3)}$
- **Two-way interaction of Att and Ed:**
  $\mathbf{(0.5 * Att_i * Ed_i)}$

Significance of Model-Implied Fixed Effects

- We now know how to calculate simple (conditional) main effects:
  Effect of interest = what it is + what *modifies* it
  e.g., Effect of Attitudes = β_1 + β_3*Ed

- But if we want to test whether that new effect is ≠ 0, we also need its **standard error** (**SE** to get Wald test *t- or z*-value → *p*-value)

- Even if the simple (conditional) main effect is not *directly* given by the model, its estimate and SE are still *implied* by the model

- **3 options** to get the new simple (conditional) main effect estimate and SE (in order of least to most annoying):

1. **Ask the software to give it to you** using your original model (e.g., ESTIMATE in SAS, TEST in SPSS, LINCOM in STATA, NEW in Mplus… most programs will do this if you know how to ask)

Significance of Model-Implied Fixed Effects

2. **Re-center your predictors** to the interacting value of interest (e.g., make attitudes=3 the new 0 for attitudes) and **re-estimate** your model; repeat as needed for each value of interest

3. **Hand calculations** (what the program does for you in option #1)

For example: **Effect of Attitudes =** $\boldsymbol{\beta_1 + \beta_3 * Ed}$

- SE^2 = sampling variance of estimate → e.g., $Var(\beta_1) = SE_{\beta 1}^2$
- $\boldsymbol{SE_{\beta 1}^2 = Var(\beta_1) + Var(\beta_3) * Ed + 2Cov(\beta_1, \beta_3) * Ed}$
    - Values come from "asymptotic (sampling) covariance matrix" (COVB)
    - Variance of a sum of terms always includes 2*covariance among them
    - Here, this is because what each main effect estimate could be is related to what the other main effect estimates could be
    - Note that if a main effect is unconditional, its $SE^2 = Var(\beta)$ only

1. Model-Implied Simple Main Effects

- Old Equation using uncentered predictors:
  $GPA_i = \beta_0 + (\beta_1 * Att_i) + (\beta_2 * Ed_i) + (\beta_3 * Att_i * Ed_i) + e_i$
  $GPA_i = 30 + (1 * Att_i) + (2 * Ed_i) + (0.5 * Att_i * Ed_i) + e_i$
- **Intercept: predicted GPA if $Att_i = 3$ and $Ed_i = 12$?**
- **Simple main effect of Att if $Ed_i = 12$? $Att_i(\beta_1 + \beta_3 * Ed_i)$**
- **Simple main effect of Ed if $Att_i = 3$? $Ed_i(\beta_2 + \beta_3 * Att_i)$**

```
ECHO 'Requesting Model-Implied Fixed Effects From Previous Slide'.
MIXED y WITH att ed  /METHOD = REML  /PRINT = SOLUTION TESTCOV
/FIXED = att ed att*ed
/TEST  = "Pred GPA if Att=3 Ed=12" intercept 1 att 3 ed 12 att*ed 36
/TEST  = "Effect of Att if Ed=12"  att 1 att*ed 12
/TEST  = "Effect of Ed if Att=3"   ed 1 att*ed 3.
```

In TEST statements, the variables refer to their betas; the numbers refer to the operations of their betas.

These estimates would be given directly by the fixed effects instead if you re-centered the predictors as: Att-3, Ed-12.

Requesting Model-Implied Fixed Effects

- **To request <u>predicted outcomes</u> (= intercepts):**
  - Need to start with **`"intercept 1"`** (for β_0)
  - ALL model effects must be included or else are held = 0
  - Note: predictors on CLASS/BY statements must be given a value (more on this next week)

- **For example: regression after centering both predictors**

$$\mathbf{GPA_i = \beta_0 + (\beta_1 * Att_i - 3) + (\beta_2 * Ed_i - 12) + (\beta_3 * Att_i - 3 * Ed_i - 12) + e_i}$$

```
"GPA if Att=5 Ed=16" intercept 1 att __ ed __ att*ed __

"GPA if Att=1 Ed=12" intercept 1 att __ ed __ att*ed __

"GPA if Att=3 Ed=20" intercept 1 att __ ed __ att*ed __
```

Requesting Model-Implied Fixed Effects

- **To request <u>predicted slopes</u> (= simple main effects):**
  - **DO NOT** start with **`"intercept 1"`** (β_0 does **not** contribute to slopes)
  - **NOT ALL** model effects must be included (**only** what modifies the slope)
  - Note: predictors on CLASS/BY statements must be given a value if they modify the slope in an interaction (more on this next week)

- **For example: regression after centering both predictors**

$$\mathbf{GPA_i = \beta_0 + (\beta_1 * Att_i - 3) + (\beta_2 * Ed_i - 12) + (\beta_3 * Att_i - 3 * Ed_i - 12) + e_i}$$

```
"Att Slope if Ed=10" intercept 0 att __ ed __ att*ed __
"Att Slope if Ed=18" intercept 0 att __ ed __ att*ed __

"Ed Slope if Att=2"  intercept 0 att __ ed __ att*ed __
"Ed Slope if Att=5"  intercept 0 att __ ed __ att*ed __
```

Regions of Significance for Main Effects

- For continuous predictors, there may not be specific values of the moderator at which you want to know significance…

- For example, age*woman (in which 0=man, 1=woman):

$$y_i = \beta_0 + (\beta_1 * Age_i - 85) + (\beta_2 * Woman_i) + (\beta_3 * Age_i - 85 * Woman_i) + e_i$$

- Age slopes are only possible for two specific values of woman:

```
"Age Slope for Men"    age85 __ woman __ age85*woman __
"Age Slope for Women"  age85 __ woman __ age85*woman __
```

- But there are many ages to request gender differences for…

```
"Gender Diff at Age=80" age85 __ woman __ age85*woman __
"Gender Diff at Age=90" age85 __ woman __ age85*woman __
```

Regions of Significance for Main Effects

- An alternative approach for continuous moderators is known as **regions of significance** (see Hoffman 2014 ch 2. for refs)
- Rather than asking if the simple main effect of gender is still significant at a particular age, we can find the **boundary ages** at which the gender effect becomes non-significant
- We know that: EST / SE = *t*-value → if |*t*| > |1.96|, then *p*<.05
- So we work backwards to find the EST and SE such that:

$$\pm \text{t} = \pm 1.96 = \frac{\text{Slope Estimate}}{\sqrt{\text{Variance of Slope Estimate}}}, \text{ where:}$$

$$\text{Gender Slope (Gender Difference) Estimate} = \beta_2 + \beta_3 \left(\text{Age} - 85\right)$$

$$\text{Variance of Slope Estimate} = \text{Var}\left(\beta_2\right) + \boxed{2\text{Cov}\left(\beta_2\beta_3\right)\left(\text{Age} - 85\right)} + \text{Var}\left(\beta_3\right)\left(\text{Age} - 85\right)^2$$

- Need to request "asymptotic covariance matrix" (COVB)
  - Covariance matrix of fixed effect estimates (SE^2 on diagonal)

Regions of Significance for Main Effects

$$\pm t = \pm 1.96 = \frac{\text{Slope Estimate}}{\sqrt{\text{Variance of Slope Estimate}}}, \text{ where:}$$

$$\text{Gender Slope (Gender Difference) Estimate} = \beta_2 + \beta_3(\text{Age} - 85)$$

$$\text{Variance of Slope Estimate} = \text{Var}(\beta_2) + 2\text{Cov}(\beta_2\beta_3)(\text{Age} - 85) + \text{Var}(\beta_3)(\text{Age} - 85)^2$$

- For example, age*woman (0=man, 1=woman), age = moderator:
  $y_i = \beta_0 + (\beta_1 * \text{Age}_i - 85) + (\beta_2 * \text{Woman}_i) + (\beta_3 * \text{Age}_i - 85 * \text{Woman}_i) + e_i$

- **β_2 = −0.5306* at age=85, Var(β_2) → SE^2 for β_2 = 0.06008**
- **β_3 = −0.1104* unconditional, Var(β_3) → SE^2 for β_3 = 0.00178**
- **Covariance of β_2 SE and β_3 SE = 0.00111**

- **Regions of Significance for Moderator of Age = 60.16 to 79.52**
  - The gender effect β_2 is predicted to be significantly negative above age 79.52, non-significant from ages 79.52 to 60.16, and significantly positive below age 60.16 (because non-parallel lines will cross eventually).

More Generally…

- Can decompose a **2-way interaction** by testing the simple effect of X at different levels of Z (and vice-versa)
  - Use TESTs to request simple effects at any point of the interacting predictor
  - Regions of significance are useful for continuous interacting predictors

- More general rules of interpretation, given a **3-way interaction**:
  - *Simple (main) effects move the intercept*
    - 1 possible interpretation for each simple main effect
    - Each <u>simple main effect</u> is conditional on other two variables = 0
  - *The 2-way interactions (3 of them in a 3-way model) move the simple effects*
    - 2 possible interpretations for each 2-way interaction
    - Each <u>simple 2-way interaction</u> is conditional on third variable = 0
  - *The 3-way interaction moves each of the 2-way interactions*
    - 3 possible interpretations of the 3-way interaction
    - Is highest-order term in model, so is unconditional (applies always)

Practice with 3-Way Interactions

- Intercept = 5, Effect of X = 1.0, Effect of Z = 0.50, Effect of W = 0.20
- **X*Z = .10 (applies specifically when W is 0)**
  - #1: for every 1-unit ΔX,
  - #2: for every 1-unit ΔZ,
- **X*W = .01 (applies specifically when Z is 0)**
  - #1: for every 1-unit ΔX,
  - #2: for every 1-unit ΔW,
- **Z*W = .05 (applies specifically when X is 0)**
  - #1: for every 1-unit ΔZ,
  - #2: for every 1-unit ΔW,
- **X*Z*W = .001 (unconditional because is highest order)**
  - #1: for every 1-unit ΔX,
  - #2: for every 1-unit ΔZ,
  - #3: for every 1-unit ΔW,

Practice with 3-Way Interactions

- Model: $y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 W_i + \beta_4 X_i W_i$
  $+\beta_5 X_i Z_i + \beta_6 Z_i W_i + \beta_7 X_i Z_i W_i + e_i$
- Formula to get simple main effects:
  - Simple effect of X =
  - Simple effect of Z =
  - Simple effect of W =
- Formula to get simple 2-way interactions:
  - Simple X*Z =
  - Simple X*W =
  - Simple Z*W =

Interpreting Interactions: Summary

- Interactions represent "moderation" – the idea that the effect of one variable depends upon the level of other(s)

- <u>The main effects WILL CHANGE in once an interaction with them is added, because they now mean different things:</u>
  - Main effect → Simple effect specifically when interacting predictor = 0
  - Need to have 0 as a meaningful value for each predictor for that reason

- <u>Rules for interpreting conditional (simple) fixed effects:</u>
  - Intercepts are conditional on (i.e., get adjusted by) main effects
  - Main effects are conditional on two-ways (become 'simple main effects')
  - Two-ways are conditional on three-ways… And so forth
  - Highest-order term is unconditional – same regardless of centering

Creating Predicted Outcomes

- Figures of predicted outcomes will be essential in describing interaction terms (especially in talks and posters)

- Three ways to get them (in order of most to least painful):

1. In excel: input fixed effects, input variable values, write equation to create predicted outcomes for each row
   - Good for pedagogy, but gets old quickly (and error-prone)

2. Via programming statements:
   - Per prediction: Use SAS ESTIMATE or SPSS TEST
   - For a range of predictor values: Use STATA MARGINS

3. Via "fake people" (most useful in SPSS and SAS)
   - Add cases to your data with desired predictor values (no outcomes)
   - Ask program to save predicted outcomes for all cases (FIXPRED in SPSS)
   - Fake cases won't contribute to model, but will get predicted outcomes