WELCOME TO PSY 653 LAB!

MODULE 02:

ORTHOGONAL CONTRASTS, POLYNOMIAL CONTRASTS AND MODERATED REGRESSION

*Thanks to Gemma Wallace for her help with these slides

OBJECTIVES

- Part 1: Orthogonal contrasts
- Part 2: Polynomial contrasts
- Part 3: Moderated regression

PART 1: ORTHOGONAL CONTRASTS

A Quick Review of Planned Contrasts

- Contrasts test a *specific* hypothesis related to group means based upon some prior information about the groups.
- Contrasts assign a weight to each of the groups in your predictor variable
    - Weights should add up to zero
    - Assigning a weight of zero means that group will not be included in the contrast
    - Order of the weights corresponds to order of groups in predictor variable

AN EXAMPLE OF A PLANNED CONTRAST SET-UP:

	Control	Drug A	Drug B
Contrast 1	-1	1/2	1/2
Contrast 2	0	-1	1

- Contrast 1 tests whether or not the control group differs from the groups that received a drug treatment

- Contrast 2 tests whether or not the two drugs differ in their effect.

ORTHOGONAL CONTRASTS

A set of contrasts is orthogonal if:

- The number of contrasts = df (number of groups -1)
- You have at least 3 groups to compare
- Two contrasts are *orthogonal* if the pairwise products of the corresponding coefficients for each term sum to zero:

	Control	Drug A	Drug B
Contrast 1	-1	1/2	1/2
Contrast 2	0	-1	1

These contrasts are orthogonal because: (-1 *0) + (1/2 * -1) + (1/2 * 1) = 0

RELEVANT FORMULAS

We can use orthogonal contrasts to get a lot of information about a dataset, even if we don't have a full ANOVA table!

<u>Sums of Squares for each contrast:</u>

n = n/cell

cj is a set of contrast weights

Ŷj is the mean of the DV in group j

$$SS_{contrast} = \frac{n(\Sigma c\bar{Y}_j)^2}{\Sigma c^2_j}$$

<u>Sum of Squares Treatments:</u>

$$\mathbf{SS_{treatments} = SS_{contrast1} + SS_{contrast2} + \ldots SS_{contrastk}}$$

<u>Sum of Squares Total:</u>

$$\mathbf{SS_{total} = SS_{contrast1} + SS_{contrast2} + \ldots SS_{contrastk} + SS_{error}}$$

	Mean	Contrast 1	Contrast 2
Group 1	35	-1	1
Group 2	40	-1	-1
Group 3	45	2	0

Note: We could use different contrast weights, depending on our specific research question and a priori information.

N for this study was 90 (i.e., 30 subjects/cell in this design)

$SS_{total} = 5000$

	Mean	Contrast 1	Contrast 2
Group 1	35	-1	1
Group 2	40	-1	-1
Group 3	45	2	0

Contrast 1 tests whether the mean of Y for treatment group 3 was significantly different than the means of the other two groups.

Contrast 2 tests the hypothesis that the mean of group 1 was significantly different than the mean of group 2.

	Mean	Contrast 1	Contrast 2
Group 1	35	-1	1
Group 2	40	-1	-1
Group 3	45	2	0

In this demo we walk through how to test the hypothesis that the mean of group 3 was significantly different from the mean of participants in groups 1 and 2 (i.e., testing the significance of Contrast 1)

You will calculate contrast 2 by yourself on the practice activity ☺

1) Calculate Sums of Squares for the contrast

	Mean	Contrast 1
Group 1	35	-1
Group 2	40	-1
Group 3	45	2

$$SS_{contrast} = \frac{n(\Sigma c\bar{Y}_j)^2}{\Sigma c^2_j}$$

N/cell = 30

$SS_{contrast1}$ = (30 * ((35 * -1) + (40 * -1) + (45 * 2))^2) / ((-1)^2 + (-1)^2 + (2)^2)

$SS_{contrast1}$ = (30 * (15)^2)/6

$SS_{contrast1}$ = (30 * 225)/6

$SS_{contrast1}$ = 6750/6 = **1125**

2) CALCULATE ETA-SQUARED FOR CONTRAST

$\mathbf{Eta^2_{contrast} = SS_{contrast}/SS_{total}}$

Note: if you are not given SS_{total} you can calculate this from the SD and mean of Y!

$\mathbf{SS_{total} = SS_{contrast1} + SS_{contrast2} + SS_{error} = 5000}$ (this was given to us)

$\mathbf{Eta^2_{contrast1}} = 1125/5000 = \mathbf{0.225}$

22.5% variance in Y explained by this contrast

2) CALCULATE ETA-SQUARED FOR CONTRAST

$$Eta^2_{contrast} = SS_{contrast}/SS_{total}$$

$$SS_{total} = SS_{contrast1} + SS_{contrast2} + SS_{error}$$

1125 is the $SS_{contrast}$ from the previous slide

5000 is the SS_{total} that was given to us beforehand

$$Eta^2_{contrast1} = \boxed{1125}/\boxed{5000} = \mathbf{0.225}$$

22.5% variance in Y explained by this contrast

RESOURCES ON ORTHOGONAL CONTRASTS

- http://www.jds-online.com/files/JDS-563.pdf

PART 2: POLYNOMIAL CONTRASTS

WHAT ARE ORTHOGONAL POLYNOMIAL CONTRASTS (AKA TREND CONTRASTS) AND WHY USE THEM?

Allow us to evaluate non-linear relations between a categorical predictor and an outcome

- With continuous predictors we can model these by squaring X to test a quadratic effect, cubing X to test a cubic effect, etc.
- With categorical predictors, we can use specific polynomial contrasts to test different effects. *These contrast levels can be found online, or in the table on slide 23!*

This type of coding system should be used only with an ordinal variable in which the levels are equally spaced

EFFECTS WE CAN EVALUATE WITH MORE THAN 3 LEVELS OF A CATEGORICAL PREDICTOR:

Linear: If we increase the dose level the Y values will increase, and we can select the best level based on the highest dose. $\mathbf{Y = a + bX}$

Quadratic: If we increase the dose level the Y values will be increased until certain dose after that the level of dosage will have a negative effect. $\mathbf{Y = a + b_1X + b_2X^2}$

Cubic: The dosage would increase Y values after certain dosage and then decrease and if we increase more the dosage level, the Y values will increase. $\mathbf{Y = a + b_1X + b_2X^2 + b_3X^3}$

CREATE A NEW R-PROJECT AND R-NOTEBOOK!

Download the “cogtest.csv” file from Canvas and save it into your R-project file

Load libraries

````
## Load libraries
```{r}
library(psych)
library(tidyverse)
library(olsrr)
```
````

NEW DATASET DESCRIPTION

Researchers were interested in the effect of time spent in practice on the performance of a visual discrimination task. Subjects were randomly assigned to different levels of practice, following which a test of visual discrimination is administered, and the number of correct responses is recorded for each subject. 40 subjects were randomly assigned to practice 0 minutes, 2 minutes, 4 minutes, 6 minutes, 8 minutes, 10 minutes, 12 minutes, or 14 minutes.

There are two variables:

practice = minutes spent practicing, this was assigned by the experimenter

score = the number of correct answers on the test

````
```{r}
cog <- read_csv("cogtest.csv")
```

Parsed with column specification:
cols(
  subject = col_double(),
  practice = col_double(),
  score = col_double()
)
````

FILTER PRACTICE TO ONLY 4 CONDITIONS

```
# Filter prectice
```{r}
cog <- filter(cog, practice == 0 | practice == 4 | practice == 8 | practice == 12)
```
```

We selected four evenly spaced conditions to test non-linear effects between practice and score

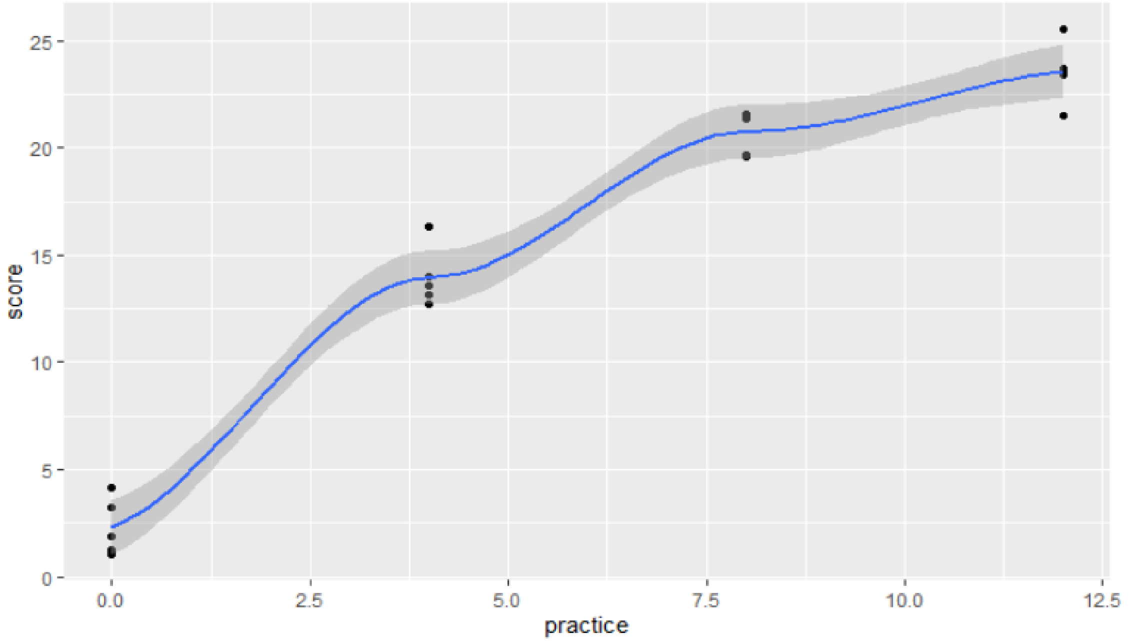
We could include more than four levels of practice if we wanted to, but we use four levels in this example for a simpler interpretation

VISUALIZE THE RELATION BETWEEN X AND Y

````
```{r}
ggplot(cog, aes(x = practice, y = score)) +
  geom_point() +
  geom_smooth(method = "loess")
```
````

`geom_smooth()` using formula 'y ~ x'

pseudoinverse used at -0.06
neighborhood radius 8.06
reciprocal condition number 4.896e-017
There are other near singularities as well. 64.964
pseudoinverse used at -0.06
neighborhood radius 8.06
reciprocal condition number 4.896e-017
There are other near singularities as well. 64.964



The slope appears to become less steep as practice time increases. Since there may be a curve to the regression line, we should test for more than a linear effect. Since we have four levels of our categorical variable, we can test both a quadratic and cubic effect.

The coefficients used for calculating sums of squares are:

Number of treatment	Degree of polynomial	T1	T2	T3	T4	T5	T6	Divisor $k=\sum c_i^2$
		Treatment totals						
2	1	-1	+1					2
3	1	-1	0	+1				2
	2	+1	-2	+1				6
4	1	-3	-1	+1	+3			20
	2	+1	-1	-1	+1			4
	3	-1	+3	-3	+1			20
5	1	-2	-1	0	+1	+2		10
	2	+2	-1	-2	-1	+2		14
	3	-1	+2	0	-2	+1		10
	4	+1	-4	+6	-4	+1		70
6	1	-5	-3	-1	+1	+3	+5	70
	2	+5	-1	-4	-4	-1	+5	84
	3	-5	+7	+4	-4	-7	+5	180
	4	+1	-3	+2	+2	-3	+1	28
	5	-1	+5	-10	+10	-5	+1	252

https://www.ndsu.edu/faculty/horsley/Polycnst.pdf

- We have 4 treatment levels

The coefficients used for calculating sums of squares are:

Number of treatment	Degree of polynomial	Treatment totals T1	T2	T3	T4	T5	T6	Divisor $k=\sum c_i^2$
2	1	-1	+1					2
3	1	-1	0	+1				2
	2	+1	-2	+1				6
4	1	-3	-1	+1	+3			20
	2	+1	-1	-1	+1			4
	3	-1	+3	-3	+1			20
5	1	-2	-1	0	+1	+2		10
	2	+2	-1	-2	-1	+2		14
	3	-1	+2	0	-2	+1		10
	4	+1	-4	+6	-4	+1		70
6	1	-5	-3	-1	+1	+3	+5	70
	2	+5	-1	-4	-4	-1	+5	84
	3	-5	+7	+4	-4	-7	+5	180
	4	+1	-3	+2	+2	-3	+1	28
	5	-1	+5	-10	+10	-5	+1	252

Linear Contrasts

Quadratic contrasts

Cubic contrasts

https://www.ndsu.edu/faculty/horsley/Polycnst.pdf

SPECIFY THE POLYNOMIAL CONTRASTS

Syntax of ifelse() function

```
ifelse(test_expression, x, y)
```

TRUE FALSE

````
```{r}
cog <- mutate(cog,

              linear = ifelse(practice == 0, -3, NA),
              linear = ifelse(practice == 4, -1, linear),
              linear = ifelse(practice == 8,  1, linear),
              linear = ifelse(practice == 12, 3, linear),

              quadratic = ifelse(practice == 0,  1, NA),
              quadratic = ifelse(practice == 4, -1, quadratic),
              quadratic = ifelse(practice == 8, -1, quadratic),
              quadratic = ifelse(practice == 12, 1, quadratic),

              cubic = ifelse(practice == 0, -1, NA),
              cubic = ifelse(practice == 4,  3, cubic),
              cubic = ifelse(practice == 8, -3, cubic),
              cubic = ifelse(practice == 12, 1, cubic)
              )
```
````

All of these contrasts come from the table on the previous slide!

We created three new variables:
linear specifies contrasts for testing a linear effect
quadratic specifies contrasts for testing a quadratic effect
cubic specifies contrasts for testing a cubic effect

STEP 1: REGRESS SCORE ON LINEAR EFFECT

```{r}
m3 <- lm(score ~ linear, data = cog)
ols_regress(m3)
```

```
                        Model Summary
--------------------------------------------------------------
R                       0.953       RMSE                2.635
R-Squared               0.909       Coef. Var          17.400
Adj. R-Squared          0.903       MSE                 6.943
Pred R-Squared          0.886       MAE                 2.207
--------------------------------------------------------------
 RMSE: Root Mean Square Error
 MSE: Mean Square Error
 MAE: Mean Absolute Error

                               ANOVA
---------------------------------------------------------------------
                Sum of
               Squares        DF    Mean Square       F         Sig.
---------------------------------------------------------------------
Regression    1240.975         1       1240.975    178.727    0.0000
Residual       124.981        18          6.943
Total         1365.957        19
---------------------------------------------------------------------

                                  Parameter Estimates
----------------------------------------------------------------------------------------
      model      Beta    Std. Error    Std. Beta      t        Sig      lower     upper
----------------------------------------------------------------------------------------
(Intercept)    15.144         0.589                 25.702    0.000    13.906    16.382
     linear     3.523         0.264        0.953    13.369    0.000     2.969     4.076
----------------------------------------------------------------------------------------
```

Evidence of linear effect

The model testing the linear effect between practice and score explained 90.9% of the variance in score, and the linear trend was statistically significant at p<0.001.

This model fits the data pretty well, but since we observed a potential curved relationship when we plotted the data, there could be a better way to examine this relationship.

STEP 2: REGRESS SCORE ON LINEAR & QUADRATIC EFFECT

```
```{r}
m4 <- lm(score ~ linear + quadratic, data = cog)
ols_regress(m4)
```
```

```
                        Model Summary
-------------------------------------------------------------
R                       0.990       RMSE                1.274
R-Squared               0.980       Coef. Var           8.411
Adj. R-Squared          0.977       MSE                 1.623
Pred R-Squared          0.972       MAE                 0.944
-------------------------------------------------------------
 RMSE: Root Mean Square Error
 MSE: Mean Square Error
 MAE: Mean Absolute Error

                              ANOVA
--------------------------------------------------------------------
               Sum of
              Squares        DF    Mean Square       F         Sig.
--------------------------------------------------------------------
Regression   1338.373         2        669.187    412.425     0.0000
Residual       27.584        17          1.623
Total        1365.957        19
--------------------------------------------------------------------

                                  Parameter Estimates
------------------------------------------------------------------------------------------
      model       Beta    Std. Error    Std. Beta       t        Sig      lower     upper
------------------------------------------------------------------------------------------
(Intercept)     15.144         0.285                 53.168    0.000     14.543    15.745
     linear      3.523         0.127        0.953    27.655    0.000      3.254     3.791
  quadratic     -2.207         0.285       -0.267    -7.748    0.000     -2.808    -1.606
------------------------------------------------------------------------------------------
```

Evidence of quadratic effect

The model testing the linear and quadratic effects between practice and score explained 98.0% of the variance in score, which is 7.1% higher than the model that only tested the linear relation.

The quadratic term is statistically significant, indicating that there is a substantial curve to the relation between practice and score (i.e., it's not linear). We need to maintain the quadratic term in the model.

STEP 3: REGRESS SCORE ON LINEAR, QUADRATIC & CUBIC EFFECT

```{r}
m5 <- lm(score ~ linear + quadratic + cubic, data = cog)
ols_regress(m5)
```

```
                         Model Summary
---------------------------------------------------------------
R                       0.990       RMSE                 1.309
R-Squared               0.980       Coef. Var            8.641
Adj. R-Squared          0.976       MSE                  1.712
Pred R-Squared          0.969       MAE                  0.954
---------------------------------------------------------------
 RMSE: Root Mean Square Error
 MSE: Mean Square Error
 MAE: Mean Absolute Error

                                 ANOVA
-----------------------------------------------------------------------
                Sum of
               Squares        DF    Mean Square       F          Sig.
-----------------------------------------------------------------------
Regression    1338.559         3        446.186    260.572    0.0000
Residual        27.397        16          1.712
Total         1365.957        19
-----------------------------------------------------------------------

                                  Parameter Estimates
-----------------------------------------------------------------------------------------
      model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
-----------------------------------------------------------------------------------------
(Intercept)    15.144         0.293                 51.755    0.000     14.524    15.764
     linear     3.523         0.131        0.953    26.921    0.000      3.245     3.800
  quadratic    -2.207         0.293       -0.267    -7.542    0.000     -2.827    -1.586
      cubic     0.043         0.131        0.012     0.330    0.746     -0.234     0.321
-----------------------------------------------------------------------------------------
```

No evidence of cubic effect

We tested the cubic term to determine if there is a second bend to the relationship between practice and score. Since we have at least four levels of our categorical predictor, we can test the cubic effect.

This model explains the same amount of variance in score as the previous model that only included the linear and quadratic effects (i.e., adding the cubic effect does not increase the explanatory power of the model).

The cubic term is not significant, indicating that there is not a second bend to the relationship. **Therefore, the quadratic model is the best fit for these data.**

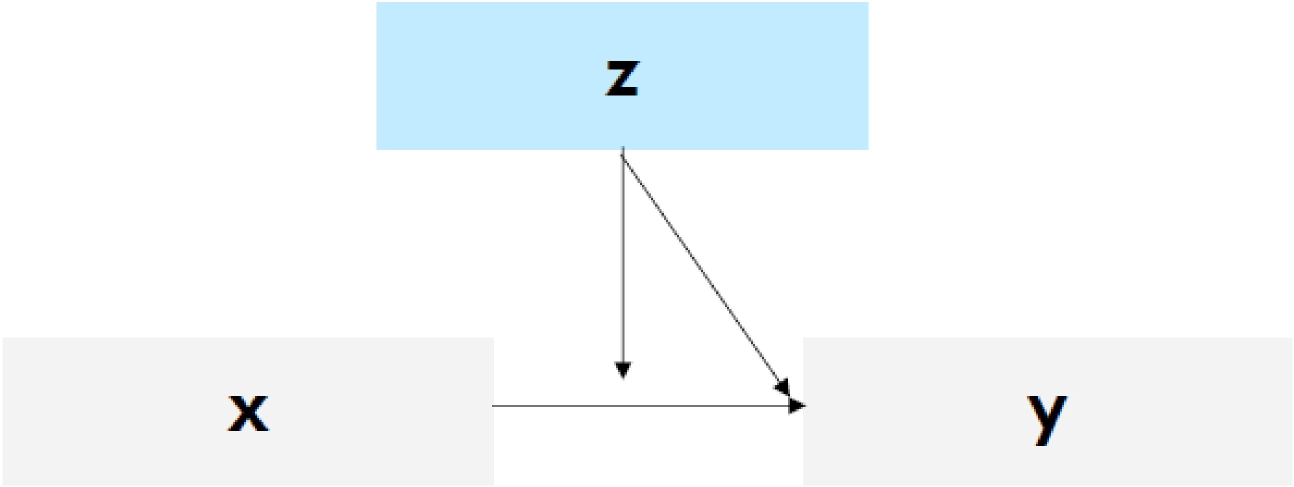
ADDITIONAL RESOURCES ON CATEGORICAL VARIABLE CODING SYSTEMS:

- https://www.researchgate.net/post/Linear_quadratic_and_cubic_polynomial_contrasts
- https://www.ndsu.edu/faculty/horsley/Polycnst.pdf

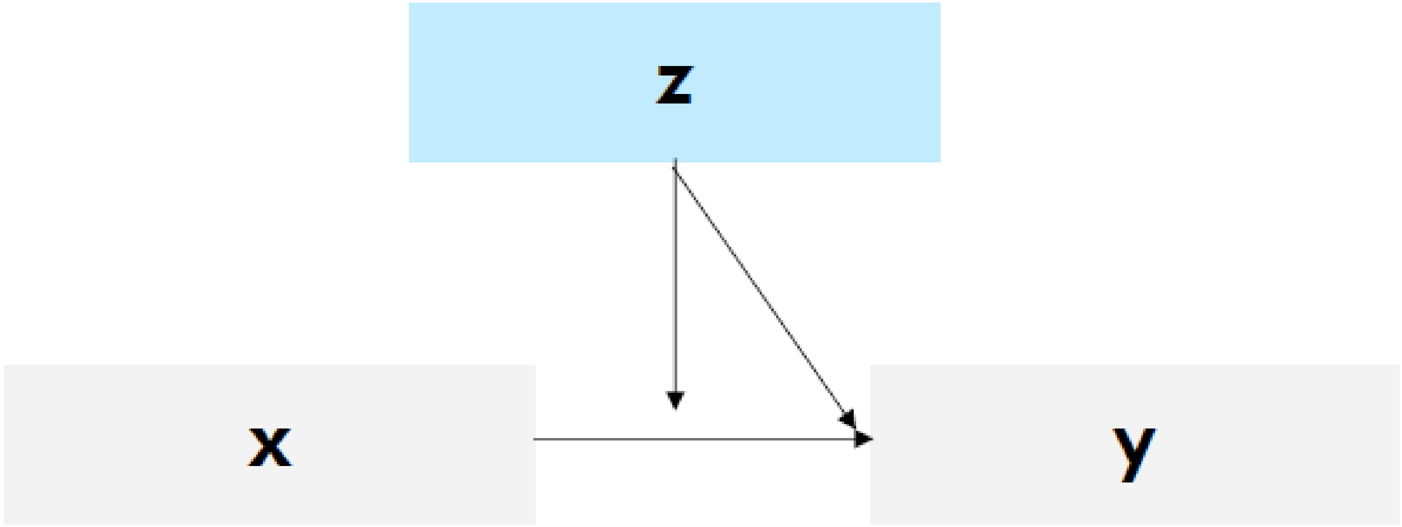
PART 3: MODERATED REGRESSION

What is Moderation?

- A moderator is a variable that changes (i.e., moderates) the relationship between two (or more) other variables
- Moderation models are used to determine if the magnitude and/or direction of a certain regression slope varies as a function of some third variable



What is moderation?

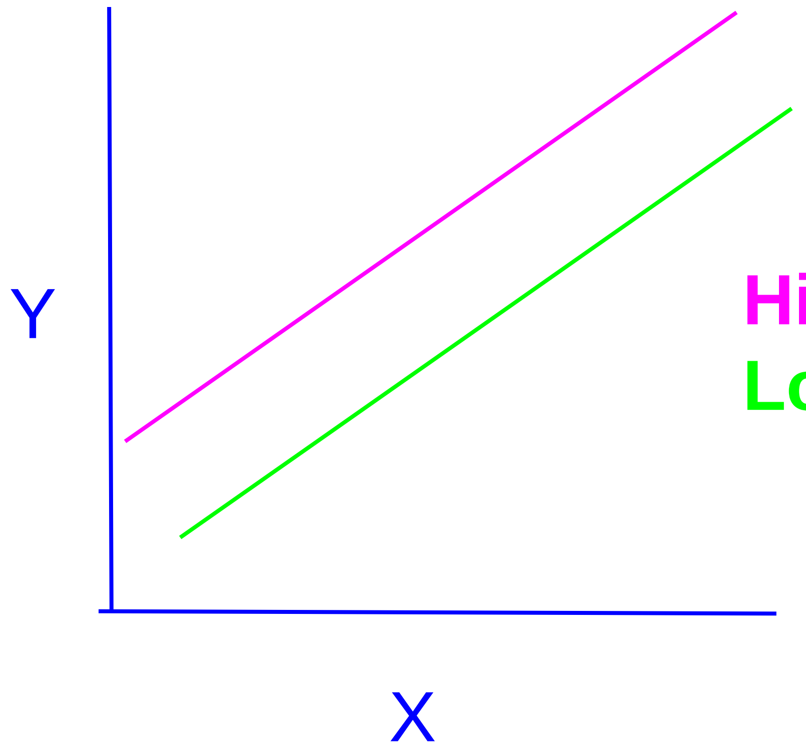


An interaction term = the predicted difference in the effect of x on y for a one-unit increase in z

(i.e., how the simple slope of x on y changes as z increases)

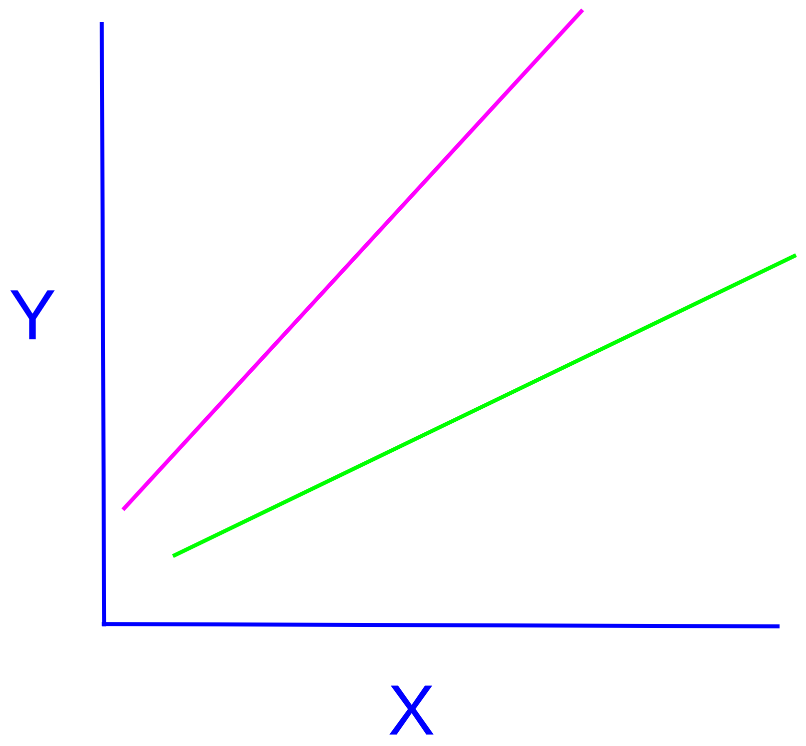
What does an interaction (aka moderation) effect look like?

No interaction observed:



High level of Z
Low level of Z

Interaction observed:



Note: crossed lines would also indicate an interaction

CREATE A NEW R-NOTEBOOK!

Download the "moderation_demo.csv" file from Canvas and save it into your R-project file

Load libraries

```
## Load libraries
```{r}
library(psych)
library(tidyverse)
library(olsrr)
```
```

Read in Data

```
## Read in data
```{r}
dat <- read_csv("moderation_demo.csv")
```
```

```
Parsed with column specification:
cols(
  att1 = col_double(),
  att2 = col_double(),
  att3 = col_double(),
  att4 = col_double(),
  att5 = col_double(),
  group1 = col_double(),
  group2 = col_double(),
  out1 = col_double(),
  out2 = col_double(),
  out3 = col_double(),
  out4 = col_double()
)
```

Calculate descriptives

This is a simulated dataset (i.e., the variables don't have specific meaning)

```
## Get descriptives
```{r, rows.print = 11}
describe(dat)
```
```

	vars	n	mean	sd	median	trimmed	mad	min	max
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
att1	1	692	1.14	1.52	0	0.88	0.00	0	9
att2	2	692	1.94	0.88	2	1.88	1.48	1	9
att3	3	692	1.48	1.44	1	1.30	1.48	0	9
att4	4	692	1.17	1.39	1	0.98	1.48	0	9
att5	5	692	1.26	1.51	1	1.06	1.48	0	9
group1	6	692	2.58	1.17	2	2.50	1.48	1	9
group2	7	692	1.59	0.49	2	1.62	0.00	1	2
out1	8	692	1.39	1.41	2	1.32	0.00	0	9
out2	9	692	1.21	1.55	1	0.97	1.48	0	9
out3	10	692	1.38	1.43	2	1.30	0.00	0	9
out4	11	692	1.39	1.48	2	1.29	0.00	0	9

1-11 of 11 rows | 1-10 of 13 columns

Our variables of interest

```
## Get descriptives
```{r, rows.print = 11}
describe(dat)
```
```

	vars	n	mean	sd	median	trimmed	mad	min	max
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
att1	1	692	1.14	1.52	0	0.88	0.00	0	9
att3	3	692	1.48	1.44	1	1.30	1.48	0	9
out4	11	692	1.39	1.48	2	1.29	0.00	0	9

Predictors

Outcome

1-11 of 11 rows | 1-10 of 13 columns

Create the cross-product of the two predictors

```
## Create the cross product of att1 & att3
```{r}
dat <- mutate(dat, att1att3 = att1*att3)
```
```

The cross-product is the interaction term. We'll include this as an additional predictor in the regression model.

Create the cross product

```
## Create the cross product of att1 & att3
```{r}
dat <- mutate(dat, att1att3 = att1*att3)
```
```

	att1	att3	att1att3
1	3	1	3
2	0	0	0
3	3	2	6
4	0	1	0
5	0	0	0
6	4	0	0
7	0	1	0
8	0	1	0
9	1	2	2
10	0	2	0
11	2	1	2
12	0	2	0
13	2	0	0
14	0	2	0
15	2	3	6
16	2	2	4

Run the main effects model

```
### Main Effects model
```{r}
modME <- lm(out4  ~ att1 + att3, data = dat)
ols_regress(modME)
```
```

This is just a regular multiple linear regression.
We first want to examine the main effects between each predictor and the outcome before adding the interaction term.

Main effects model results

Model Summary

R	0.453	RMSE	1.317
R-Squared	0.206	Coef. Var	94.839
Adj. R-Squared	0.203	MSE	1.735
Pred R-Squared	0.179	MAE	0.905

RMSE: Root Mean Square Error
MSE: Mean Square Error
MAE: Mean Absolute Error

ANOVA

	Sum of Squares	DF	Mean Square	F	Sig.
Regression	309.277	2	154.638	89.148	0.0000
Residual	1195.156	689	1.735		
Total	1504.432	691			

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	0.655	0.080		8.185	0.000	0.498	0.812
att1	0.049	0.033	0.050	1.480	0.139	-0.016	0.114
att3	0.458	0.035	0.448	13.190	0.000	0.390	0.527

Run the same model with the interaction term
(aka a Moderated Regression)

```
### Interaction model
```{r}
modINT <- lm(out4 ~ att1 + att3 + att1att3 , data = dat)
ols_regress(modINT)
```
```

att1att3 is the cross-product variable we made earlier
It represents att1*att3

Model Summary

R	0.462	RMSE	1.311
R-Squared	0.213	Coef. Var	94.435
Adj. R-Squared	0.210	MSE	1.720
Pred R-Squared	0.176	MAE	0.907

RMSE: Root Mean Square Error
MSE: Mean Square Error
MAE: Mean Absolute Error

ANOVA

	Sum of Squares	DF	Mean Square	F	Sig.
Regression	321.154	3	107.051	62.243	0.0000
Residual	1183.278	688	1.720		
Total	1504.432	691			

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	0.757	0.089		8.539	0.000	0.583	0.931
att1	-0.032	0.045	-0.033	-0.716	0.474	-0.121	0.056
att3	0.393	0.043	0.384	9.210	0.000	0.309	0.477
att1att3	0.049	0.019	0.140	2.628	0.009	0.012	0.085

Model results with interaction term

Simple Slopes

Interaction

Interpreting the moderated regression

Model Summary

R	0.462	RMSE	1.311
R-Squared	0.213	Coef. Var	94.435
Adj. R-Squared	0.210	MSE	1.720
Pred R-Squared	0.176	MAE	0.907

RMSE: Root Mean Square Error
MSE: Mean Square Error
MAE: Mean Absolute Error

ANOVA

	Sum of Squares	DF	Mean Square	F	Sig.
Regression	321.154	3	107.051	62.243	0.0000
Residual	1183.278	688	1.720		
Total	1504.432	691			

Parameter Estimates

model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
(Intercept)	0.757	0.089		8.539	0.000	0.583	0.931
att1	-0.032	0.045	-0.033	-0.716	0.474	-0.121	0.056
att3	0.393	0.043	0.384	9.210	0.000	0.309	0.477
att1att3	0.049	0.019	0.140	2.628	0.009	0.012	0.085

att1: This is the predicted change in our outcome for every 1 unit increase in att1 ***when att3 is 0.***

att3: This is the predicted change in our outcome for every 1 unit increase in att3 ***when att1 is 0.***

att1att3: This is the predicted **change in the effect of att1** on our outcome **for every one unit increase in att3.**

Compare the two models via hierarchical regression

```
# Compare the two models
```{r}
anova(modME,modINT)
```
```

```
Analysis of Variance Table

Model 1: out4 ~ att1 + att3
Model 2: out4 ~ att1 + att3 + att1att3
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1    689 1195.2
2    688 1183.3  1    11.877 6.9058 0.008783 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Does including the interaction term significantly improve the amount of variance our model explains in out4?

Compare the two models via hierarchical regression

```
# Compare the two models
```{r}
anova(modME,modINT)
```
```

```
Analysis of Variance Table

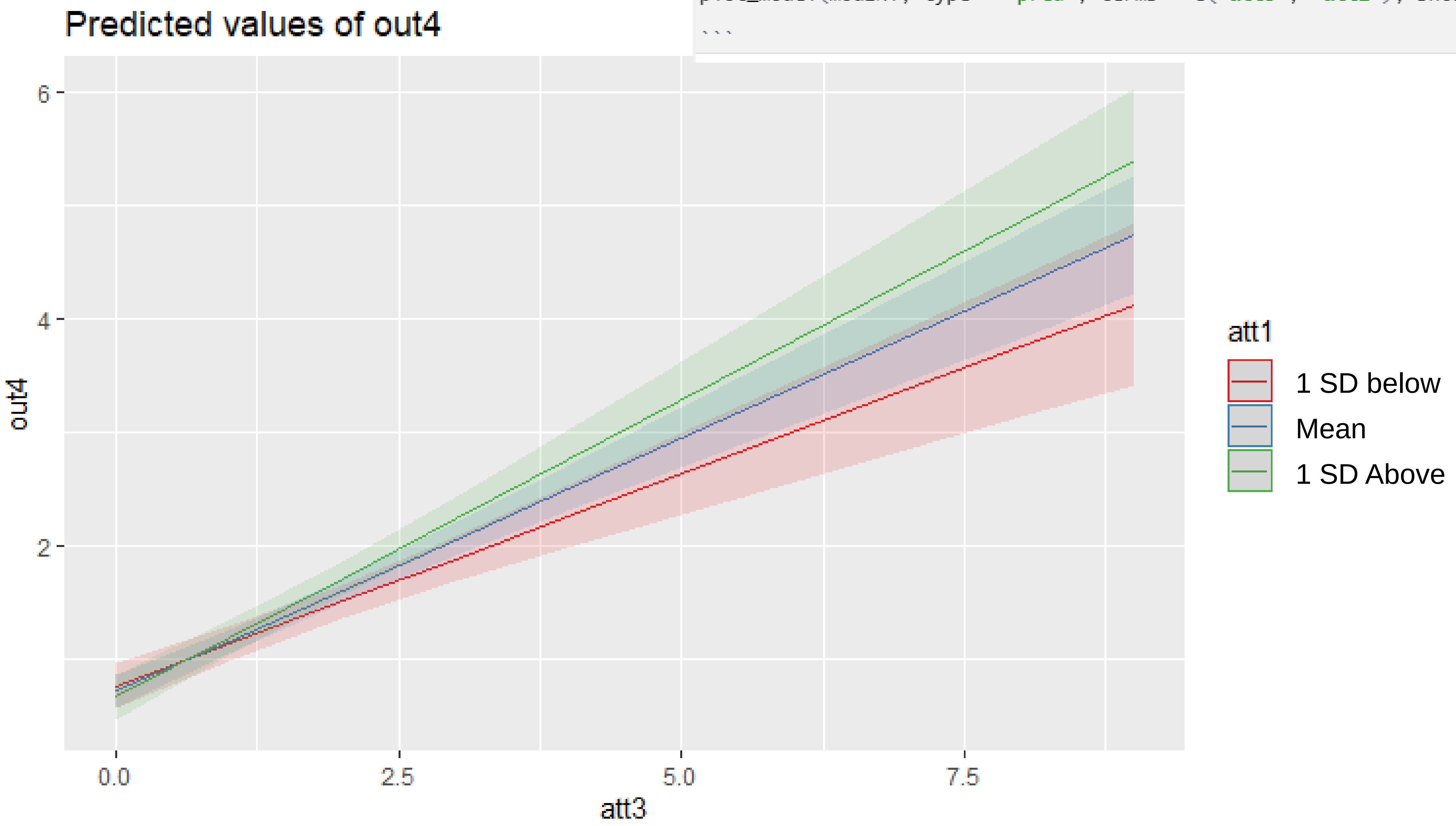
Model 1: out4 ~ att1 + att3
Model 2: out4 ~ att1 + att3 + att1att3
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1    689 1195.2
2    688 1183.3  1    11.877 6.9058 0.008783 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

At $p<0.05$, the moderated regression model explained significantly more variance in out4 than the main effects model

Our interaction - Visualized

```{r}
library(sjPlot)
library(sjmisc)

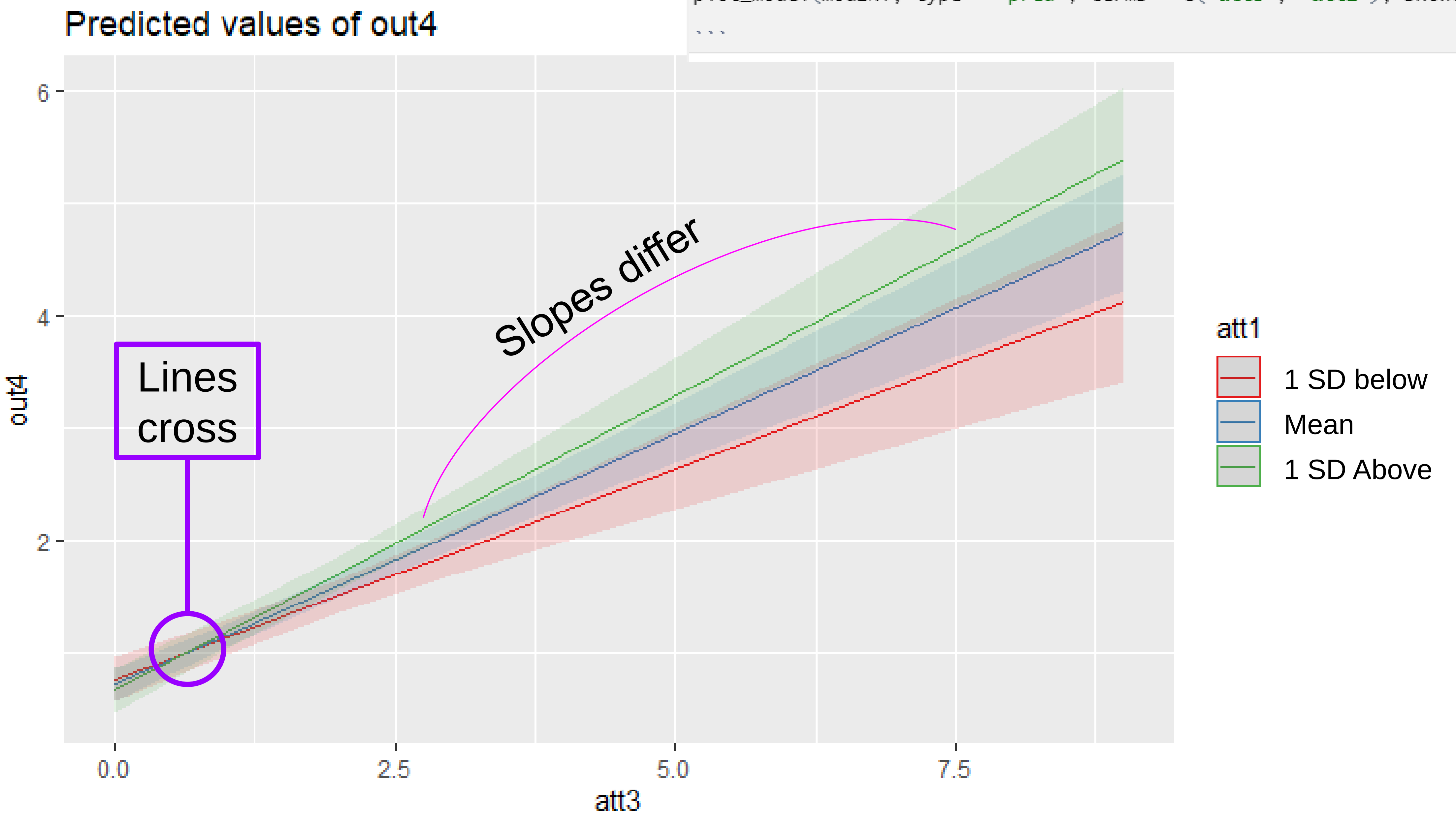
plot_model(modINT, type = "pred", terms = c("att3", "att1"), show.intercept = TRUE)
```



Our interaction - Visualized

```{r}
library(sjPlot)
library(sjmisc)

plot_model(modINT, type = "pred", terms = c("att3", "att1"), show.intercept = TRUE)
```



Alternative way to specify interactions: Specify the interaction in the regression equation using a “*”

```
### Interaction model
```{r}
modINT <- lm(out4 ~ att1 + att3 + att1*att3 , data = dat)
ols_regress(modINT)
```
```

```
                        Model Summary
--------------------------------------------------------------
R                       0.462       RMSE                1.311
R-Squared               0.213       Coef. Var          94.435
Adj. R-Squared          0.210       MSE                 1.720
Pred R-Squared          0.176       MAE                 0.907
--------------------------------------------------------------
 RMSE: Root Mean Square Error
 MSE: Mean Square Error
 MAE: Mean Absolute Error

                              ANOVA
---------------------------------------------------------------------
                Sum of
               Squares        DF    Mean Square      F         Sig.
---------------------------------------------------------------------
Regression     321.154         3        107.051    62.243    0.0000
Residual      1183.278       688          1.720
Total         1504.432       691
---------------------------------------------------------------------

                                  Parameter Estimates
---------------------------------------------------------------------------------------
      model      Beta    Std. Error    Std. Beta      t        Sig      lower    upper
---------------------------------------------------------------------------------------
(Intercept)     0.757         0.089                 8.539    0.000     0.583    0.931
       att1    -0.032         0.045        0.041   -0.716    0.474    -0.121    0.056
       att3     0.393         0.043        0.439    9.210    0.000     0.309    0.477
  att1:att3     0.049         0.019        0.090    2.628    0.009     0.012    0.085
---------------------------------------------------------------------------------------
```

Example APA write up

The differential effect of att1 on out4, using att3 as a moderator, was examined among 692 participants. A moderation model was estimated, out4 was regressed on att1, att3, and the interaction between the two. The simple slope of att1 was not statistically significant (b = -0.03, 95%CI -0.12, 0.06) and the simple slope of att3 (b = 0.39, 95%CI 0.31, 0.48) was statistically significant. The interaction term is statistically significant (b = 0.05, 95%CI 0.01, 0.09), indicating that the effect of att1 on out4 is larger as att3 increases.

Additional considerations for moderation

- Power is important!
  - N needed to detect interaction effect can be up to 9x larger than for detecting main effects (e.g., Wahlsten, 1991)
  - For every interaction term you add, N needed to detect effect increases

- You can examine interactions between more than two variables
  - E.g., 3-way interactions

- Interpretation is easier with categorical predictors
  - E.g. You can turn continuous variables into categorical by using cut-off scores