Lecture 8: Probing Moderation

Kyle M. Lang

Institute for Measurement, Methodology, Analysis & Policy
Texas Tech University
Lubbock, TX

April 4, 2016

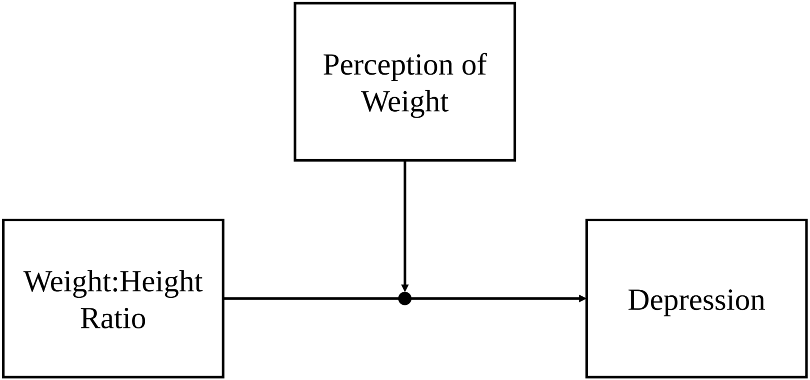
TEXAS TECH
UNIVERSITY

Outline

- Probing moderation via centering
- Alternative probing strategies
- Confidence bands for simple slopes

Starting Point

Last time, we fit the following model:



Interaction Probing

We probed the interaction with the *Pick-a-point* approach.

- Choose interesting values of the moderator Z
- Check the significance of the focal effect $X \rightarrow Y$ at the values we choose for Z.
- Gives us an idea of where in Z's distribution the focal effect is/is not significant.

Previously, we manually calculated the all of the quantities we needed, including a SE for the conditional focal effect.

- There is a simpler way: Centering

Centering

Centering transforms a variable by subtracting a constant (e.g., the variable's mean) from each observation of the variable

- The most familiar form of center is *mean centering*
- We can center on any value
  - When probing interactions, we can center Z on the interesting values we choose during the pick-a-point approach
  - Running the model with Z centered on specific values automatically provides tests of the simple slope conditional on those values of Z

Probing via Centering

Say we want to do a simple slopes analysis to test the conditional effect of X on Y at three levels of $Z = \{Z_1, Z_2, Z_3\}$.

Then, all we need to do is fit the following three models:

$$Y = \alpha + \beta_1 X + \beta_2 (Z - Z_1) + \beta_3 X (Z - Z_1) + e$$

$$Y = \alpha + \beta_1 X + \beta_2 (Z - Z_2) + \beta_3 X (Z - Z_2) + e$$

$$Y = \alpha + \beta_1 X + \beta_2 (Z - Z_3) + \beta_3 X (Z - Z_3) + e$$

The default output for β_1 provides tests of the simple slopes.

Example

```r
## Read in the data:
dataDir <- "../data/"
fileName <- "nlsyData.rds"
dat1 <- readRDS(paste0(dataDir, fileName))
## Moderated Model:
out1 <- lm(depress1 ~ ratio1*perception1, data = dat1)
summary(out1)
```

```
Call:
lm(formula = depress1 ~ ratio1 * perception1, data = dat1)

Residuals:
    Min      1Q  Median      3Q     Max
-2.1218 -0.2830  0.0881  0.3515  1.0835

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         2.41374    0.10375  23.265  < 2e-16 ***
ratio1              0.38126    0.04571   8.341  < 2e-16 ***
perception1         0.11083    0.02941   3.768 0.000165 ***
ratio1:perception1 -0.07702    0.01214  -6.343 2.37e-10 ***
---
```

Example

```
Signif. codes:  0  ***  0.001  **  0.01  *  0.05
      .  0.1     1

Residual standard error: 0.5086 on 8980 degrees of freedom
Multiple R^2:  0.01426,  Adjusted R^2:  0.01393
F-statistic:  43.3 on 3 and 8980 DF,  p-value: < 2.2e-16
```

```r
## Compute critical values of Z:
zMean <- mean(dat1$perception1)
zSD <- sd(dat1$perception1)
dat1$zCen <- dat1$perception1 - zMean
dat1$zHigh <- dat1$perception1 - (zMean + zSD)
dat1$zLow <- dat1$perception1 - (zMean - zSD)
## Test simple slopes:
out2.1 <- lm(depress1 ~ ratio1*zLow, data = dat1)
summary(out2.1)
```

Example

```
Call:
lm(formula = depress1 ~ ratio1 * zLow, data = dat1)

Residuals:
    Min      1Q  Median      3Q     Max
-2.1218 -0.2830  0.0881  0.3515  1.0835

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.68561    0.04066  66.052  < 2e-16 ***
ratio1       0.19234    0.01888  10.190  < 2e-16 ***
zLow         0.11083    0.02941   3.768 0.000165 ***
ratio1:zLow -0.07702    0.01214  -6.343 2.37e-10 ***
---
Signif. codes:  0   ***   0.001   **   0.01   *   0.05
        .   0.1        1

Residual standard error: 0.5086 on 8980 degrees of freedom
Multiple R^2:  0.01426,  Adjusted R^2:  0.01393
F-statistic:  43.3 on 3 and 8980 DF,  p-value: < 2.2e-16
```

Example

```
out2.2 ← lm(depress1 ~ ratio1*zCen, data = dat1)
summary(out2.2)
```

```
Call:
lm(formula = depress1 ~ ratio1 * zCen, data = dat1)

Residuals:
    Min      1Q  Median      3Q     Max
-2.1218 -0.2830  0.0881  0.3515  1.0835

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.76845    0.03076  89.998  < 2e-16 ***
ratio1       0.13477    0.01345  10.021  < 2e-16 ***
zCen         0.11083    0.02941   3.768 0.000165 ***
ratio1:zCen -0.07702    0.01214  -6.343 2.37e-10 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05
        .    0.1          1

Residual standard error: 0.5086 on 8980 degrees of freedom
Multiple R^2:  0.01426,  Adjusted R^2:  0.01393
```

Example

```
F-statistic:  43.3 on 3 and 8980 DF,  p-value: < 2.2e-16
```

```
out2.3 ← lm(depress1 ~ ratio1*zHigh, data = dat1)
summary(out2.3)
```

```
Call:
lm(formula = depress1 ~ ratio1 * zHigh, data = dat1)

Residuals:
    Min      1Q  Median      3Q     Max
-2.1218 -0.2830  0.0881  0.3515  1.0835

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.85128    0.03472  82.115  < 2e-16 ***
ratio1        0.07721    0.01305   5.919 3.36e-09 ***
zHigh         0.11083    0.02941   3.768 0.000165 ***
ratio1:zHigh -0.07702    0.01214  -6.343 2.37e-10 ***
---
Signif. codes:  0    ***    0.001    **    0.01    *    0.05
        .    0.1         1
```

Example

```
Residual standard error: 0.5086 on 8980 degrees of freedom
Multiple R^2:  0.01426,  Adjusted R^2:  0.01393
F-statistic:  43.3 on 3 and 8980 DF,  p-value: < 2.2e-16
```

Compare Approaches

The manual and the centering approaches give identical answers, barring rounding errors with the manual approach:

	Z Low	Z Center	Z High
Manual	0.192335	0.134774	0.077213
Centering	0.192335	0.134774	0.077213

Table: Simple Slopes

	Z Low	Z Center	Z High
Manual	0.018875	0.013449	0.013045
Centering	0.018875	0.013449	0.013045

Table: Standard Errors

A Few Comments on Centering

You will often hear mean centering touted as absolutely necessary or absolutely unnecessary for moderation analysis.

Both sides are partially correct.

Two effects are usually ascribed to mean centering in moderation analysis:

1. Improved interpretation of the conditional effects
2. Reduced multicollinearity between lower-order effects and the interaction term

A Few Comments on Centering

Mean center absolutely *does* have the potential to improve parameter interpretations

- When $X = 0$ or $Z = 0$ are not sensible values, centering is necessary for any plausible interpretation of β_1 or β_2.

Mean centering *can* remove collinearity between lower-order terms and the interaction term

- **BUT**, we don't care

We can get a better sense of what's going on with a synthetic example.

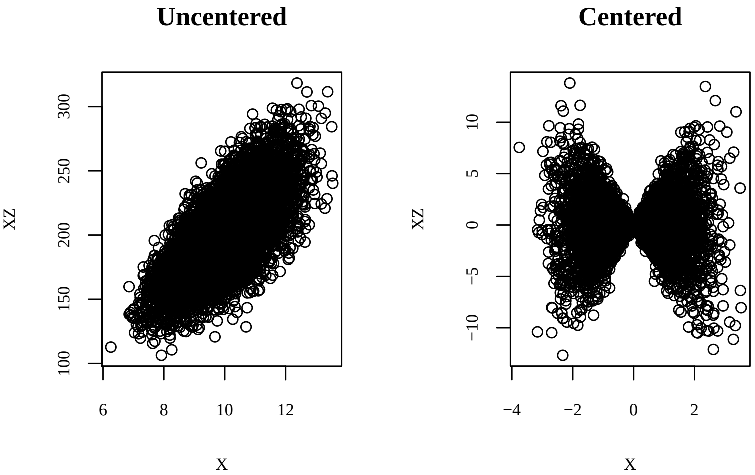
Example

```
n ← 10000
x ← rnorm(n, 10, 1)
z ← rnorm(n, 20, 2)
xz ← x*z
cor(x, xz)
```

```
[1] 0.6994485
```

```
xc ← x - mean(x)
zc ← z - mean(z)
xzc ← xc*zc
cor(xc, xzc)
```

```
[1] -0.02232915
```


Uncentered
Centered
XZ
X
300
250
200
150
100
6
8
10
12
10
5
0
−5
−10
−4
−2
0
2

Example

```
y ← 5*x + 5*z + 2*xz + rnorm(n, 0, 0.5)
out3.1 ← lm(y ~ x*z)
out3.2 ← lm(y ~ xc*z)
out3.3 ← lm(y ~ xc*zc)
summary(out3.1)$coef
```

```
             Estimate  Std. Error     t value  Pr(>|t|)
(Intercept) 0.3110316 0.497848247   0.6247518 0.5321483
x           4.9698297 0.049482362 100.4363872 0.0000000
z           4.9849669 0.024795709 201.0415110 0.0000000
x:z         2.0014878 0.002465063 811.9419413 0.0000000
```

```
summary(out3.2)$coef
```

```
             Estimate  Std. Error    t value Pr(>|t|)
(Intercept) 50.087318 0.049771942  1006.3364        0
xc           4.969830 0.049482362   100.4364        0
z           25.031254 0.002473949 10117.9348        0
xc:z         2.001488 0.002465063   811.9419        0
```

```
summary(out3.3)$coef
```

```
              Estimate  Std. Error     t value Pr(>|t|)
(Intercept) 551.142076 0.004990701 110433.8019        0
xc           45.033944 0.004970146   9060.8897        0
zc           25.031254 0.002473949  10117.9348        0
xc:zc         2.001488 0.002465063    811.9419        0
```

Example

```
sum(out3.1$fitted - out3.3$fitted)
```

```
[1] -2.842171e-13
```

```
summary(out3.1)$r.squared
```

```
[1] 0.9999454
```

```
summary(out3.3)$r.squared
```

```
[1] 0.9999454
```

Example

```
summary(lm(y ~ x*zLow))$coef
```

```
             Estimate  Std. Error   t value Pr(>|t|)
(Intercept) 90.039011 0.070029301 1285.7334        0
x           40.996039 0.006951782 5897.1983        0
zLow         4.984967 0.024795709  201.0415        0
x:zLow       2.001488 0.002465063  811.9419        0
```

```
summary(lm(y ~ xc*zLow))$coef
```

```
              Estimate  Std. Error    t value Pr(>|t|)
(Intercept) 500.642731 0.007057934 70933.3306        0
xc           40.996039 0.006951782  5897.1983        0
zLow         25.031254 0.002473949 10117.9348        0
xc:zLow       2.001488 0.002465063   811.9419        0
```

Example

```
summary(lm(y ~ x*zCen))$coef
```

```
             Estimate  Std. Error   t value Pr(>|t|)
(Intercept) 100.095940 0.050027894 2000.8026        0
x            45.033944 0.004970146 9060.8897        0
zCen          4.984967 0.024795709  201.0415        0
x:zCen        2.001488 0.002465063  811.9419        0
```

```
summary(lm(y ~ xc*zCen))$coef
```

```
             Estimate  Std. Error     t value Pr(>|t|)
(Intercept) 551.142076 0.004990701 110433.8019        0
xc           45.033944 0.004970146   9060.8897        0
zCen         25.031254 0.002473949  10117.9348        0
xc:zCen       2.001488 0.002465063    811.9419        0
```

Example

```
summary(lm(y ~ x*zHigh))$coef
```

```
              Estimate  Std. Error   t value Pr(>|t|)
(Intercept) 110.152870 0.071458433 1541.4957        0
x            49.071849 0.007109273 6902.5127        0
zHigh         4.984967 0.024795709  201.0415        0
x:zHigh       2.001488 0.002465063  811.9419        0
```

```
summary(lm(y ~ xc*zHigh))$coef
```

```
              Estimate  Std. Error    t value Pr(>|t|)
(Intercept) 601.641421 0.007058426 85237.3387        0
xc           49.071849 0.007109273  6902.5127        0
zHigh        25.031254 0.002473949 10117.9348        0
xc:zHigh      2.001488 0.002465063   811.9419        0
```

Back to Work

Question: Okay, so what about our example analysis? Should we center the predictors in this model:

$$Depress = \alpha + \beta_1 Ratio + \beta_2 Perception + \beta_3 Ratio \times Perception + e$$

Back to Work

Question: Okay, so what about our example analysis? Should we center the predictors in this model:

$$Depress = \alpha + \beta_1 Ratio + \beta_2 Perception + \beta_3 Ratio \times Perception + e$$

Answer: Yes.

Back to Work

QUESTION: Okay, so what about our example analysis? Should we center the predictors in this model:

$$Depress = \alpha + \beta_1 Ratio + \beta_2 Perception + \beta_3 Ratio \times Perception + e$$

ANSWER: Yes.

QUESTION MARK II: Why?

Back to Work

QUESTION: Okay, so what about our example analysis? Should we center the predictors in this model:

$$Depress = \alpha + \beta_1 Ratio + \beta_2 Perception + \beta_3 Ratio \times Perception + e$$

ANSWER: Yes.

QUESTION MARK II: Why?

ANSWER THE SECOND: Because a *Weight:Height* ratio of zero is nonsensical and zero is outside the range of *Perception*.

Example

```
dat1$ratioC ← dat1$ratio1 - mean(dat1$ratio1)
dat1$perceptionC ← dat1$perception1 - mean(dat1$perception1)
## Moderated Model:
out1 ← lm(depress1 ~ ratioC*perceptionC, data = dat1)
summary(out1)$coef
```

```
                     Estimate  Std. Error    t value     Pr(>|t|)
(Intercept)        3.07996908 0.005830947 528.210782 0.000000e+00
ratioC             0.13477427 0.013448801  10.021285 1.632952e-23
perceptionC       -0.06719252 0.008311533  -8.084252 7.065839e-16
ratioC:perceptionC -0.07701956 0.012142973  -6.342726 2.366288e-10
```

```
## Compute critical values of Z:
zMean ← mean(dat1$perceptionC)
zSD ← sd(dat1$perceptionC)
dat1$zCen ← dat1$perceptionC - zMean
dat1$zHigh ← dat1$perceptionC - (zMean + zSD)
dat1$zLow ← dat1$perceptionC - (zMean - zSD)
```

Example

```r
## Test simple slopes:
out2.1 <- lm(depress1 ~ ratioC*zLow, data = dat1)
summary(out2.1)$coef
```

```
               Estimate  Std. Error    t value     Pr(>|t|)
(Intercept)  3.13018574 0.008490579 368.665759 0.000000e+00
ratioC       0.19233521 0.018875273  10.189798 2.983313e-24
zLow        -0.06719252 0.008311533  -8.084252 7.065839e-16
ratioC:zLow -0.07701956 0.012142973  -6.342726 2.366288e-10
```

```r
out2.2 <- lm(depress1 ~ ratioC*zCen, data = dat1)
summary(out2.2)$coef
```

```
               Estimate  Std. Error    t value     Pr(>|t|)
(Intercept)  3.07996908 0.005830947 528.210782 0.000000e+00
ratioC       0.13477427 0.013448801  10.021285 1.632952e-23
zCen        -0.06719252 0.008311533  -8.084252 7.065839e-16
ratioC:zCen -0.07701956 0.012142973  -6.342726 2.366288e-10
```

```r
out2.3 <- lm(depress1 ~ ratioC*zHigh, data = dat1)
summary(out2.3)$coef
```

```
                Estimate  Std. Error    t value     Pr(>|t|)
(Intercept)   3.02975242 0.008548655 354.412740 0.000000e+00
ratioC        0.07721333 0.013045300   5.918862 3.360611e-09
zHigh        -0.06719252 0.008311533  -8.084252 7.065839e-16
ratioC:zHigh -0.07701956 0.012142973  -6.342726 2.366288e-10
```

Alternative Probing Strategies

The pick-a-point approach is nice due to its simplicity and ease of interpretation, but the points we choose are totally arbitrary.

- We may be missing important nuances that occur in some of the areas of Z's distribution that we *did not* pick.

Alternative Probing Strategies

The pick-a-point approach is nice due to its simplicity and ease of interpretation, but the points we choose are totally arbitrary.

- We may be missing important nuances that occur in some of the areas of Z's distribution that we *did not* pick.

The *Johnson-Neyman* technique is an alternative approach that removes the arbitrary choices necessary for pick-a-point.

- Johnson-Neyman finds the *region of significance* wherein the conditional effect of X on Y is statistically significant
- Inverts the pick-a-point approach to find what cut-points on the moderator correspond to a critical t value for the conditional β_1.

With pick-a-point, we:

1. Choose conditional values of Z, say Z_1
2. Calculate the simple slope SS_1 and standard error SE_{SS1} associated with Z_1
3. Test SS_1 for significance via a simple Wald-type test:

$$t = \frac{SS_1}{SE_{SS1}} \tag{1}$$

Johnson-Neyman Technique

With Johnson-Neyman, we:

1. Choose an α level for our test and the corresponding critical value of t, say $t_{crit} = 1.96$ to give $\alpha = 0.05$ in large samples.
2. Re-arrange Equation 1 into the following quadratic form:

$$t_{crit}^2 SE_{SS}^2 - SS^2 = 0 \tag{2}$$

3. Solve Equation 2 to find the two values of Z that produce critical t statistics for the conditional focal effect.

The roots produced by the Johnson-Neyman technique delineate the *region of significance*.

- The conditional effect of X on Y is either significant everywhere between these two points or everywhere outside of these two points.
- If only one of the points falls within the observed range of Z, ignore the other point
  - The region of significant is either everywhere above or below the legal root
- If neither of the roots fall within the observed range of Z then, either:
  1. The focal effect is significant across the entire range of Z, or
  2. The focal effect is not significant anywhere within the range of Z

Perspectives on Simple Slopes

Recall the formula for a simple slope:

$$SS = \beta_1 + \beta_3 Z$$

From a graphical perspective, we can think about SS in, at least, two different ways:

1. As a weight for X that we can use to get plots of the conditional effect of X on Y at different levels of Z.
2. We can also consider how SS, itself, smoothly changes as a function of Z.

The latter perspective embodies the spirit of the Johnson-Neyman technique.

A natural quantity to consider is a confidence interval for SS:

$$CI_{SS} = SS \pm t_{crit} \cdot SE_{SS}$$

Last time, we computed a few such intervals for the interesting values of Z we chose for the pick-a-point analysis.

When doing Johnson-Neyman, we can considering the values of CI_{SS} for the entire range of Z.

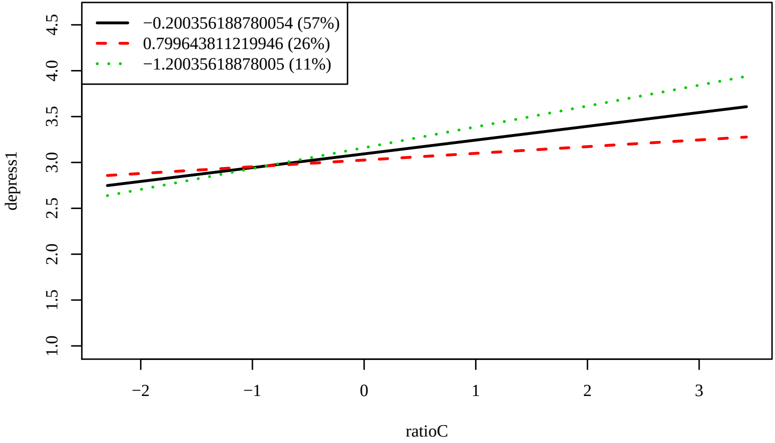
- These CI values define the *confidence bands* of SS and show, for any value of Z, the corresponding CI for SS
  - As a result, we can immediately check any value of Z for a significant simple slope

Implementing the Johnson-Neyman technique by hand is a pain, but we can easily do so by using the **rockchalk** package in R

```
par(family = "serif", cex = 0.75)
library(rockchalk)
## First we need to create a 'plotSlopes' object:
plotOut ← plotSlopes(model = out1,
                     plotx = "ratioC",
                     modx = "perceptionC",
                     plotPoints = FALSE)
## Then we modify 'plotOut' to get the J-N test:
testOut ← testSlopes(plotOut)
```

```
Values of perceptionC OUTSIDE this interval:
      lo        hi
1.327479 2.452667
cause the slope of (b1 + b2*perceptionC)ratioC to be
    statistically significant
```

Example



−0.200356188780054 (57%)
0.799643811219946 (26%)
−1.20035618878005 (11%)
depress1
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
−2
−1
0
1
2
3
ratioC

We can see the significance boundaries by extracting the roots from 'testOut'

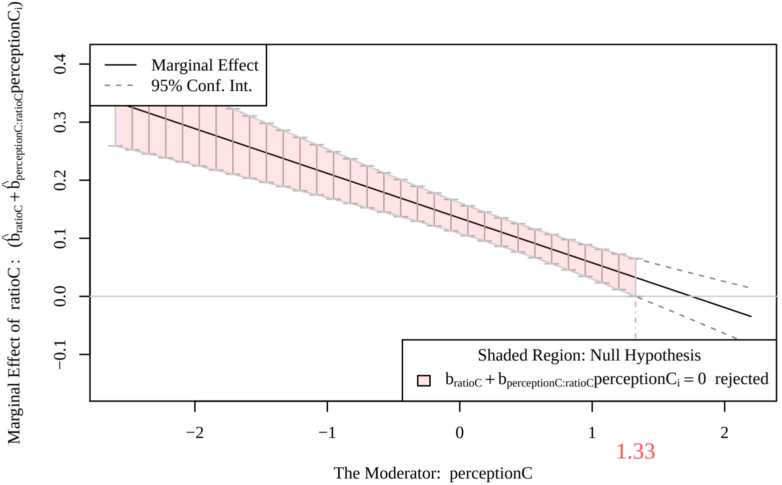
```
testOut$jn$roots
```

```
      lo        hi
1.327479 2.452667
```

We can plot the result:

```
par(cex = 0.75, family = "serif")
plot(testOut)
```

Example



Marginal Effect
95% Conf. Int.
Shaded Region: Null Hypothesis
$b_{ratioC} + b_{perceptionC:ratioC} perceptionC_i = 0$ rejected
Marginal Effect of ratioC : $(\hat{b}_{ratioC} + \hat{b}_{perceptionC:ratioC} perceptionC_i)$
0.4
0.3
0.2
0.1
0.0
−0.1
−2
−1
0
1
2
1.33
The Moderator: perceptionC