Regression-Based Tests for Moderation

Brian K. Miller, Ph.D.

1

Presentation Objectives

1. Differentiate between mediation & moderation
2. Differentiate between hierarchical and stepwise regression
3. Run and interpreting hierarchical regression in SPSS
4. Compute interaction terms
5. Mean center variables
6. Graphing interactions

2

1. Differentiate between mediation and moderation

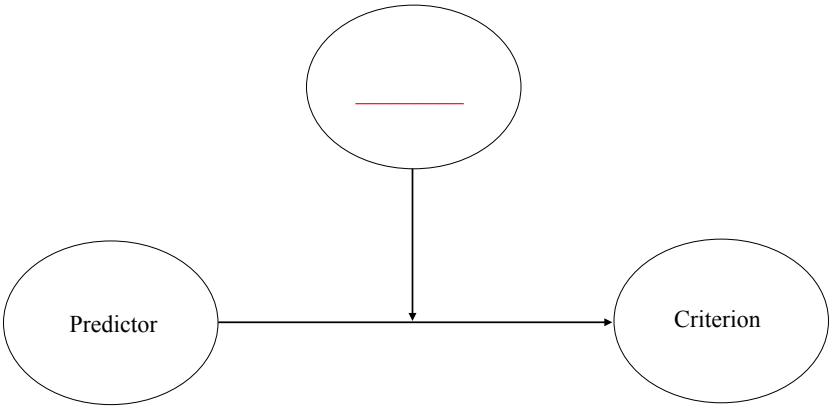
3

Moderation vs. Mediation

- _____: third variable that affects strength of relationship between two other variables
  - Ex: relationship between performance and salary _____ _____ gender
  - Ex: relationship between X_1 and Y is strong, especially if X_2 is also strong
- _____: third variable that acts as _____ _____ between two other variables
  - Ex: performance mediates relationship between job knowledge and salary
  - Ex: X_1 "causes" X_2 which "causes" Y

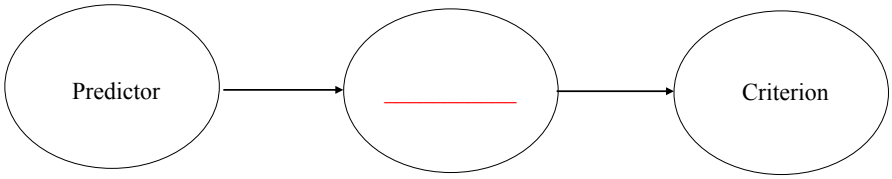
4

_____ Diagram



5

_____ Diagram



6

2. Differentiate between hierarchical regression & stepwise regression

7

Hierarchical Regression

- Variables entered in equation at various stages
- Important statistics and tests
  - Change in R^2 (i.e. _____)
  - Change in *F*-score (i.e. ΔF)
  - Overall equation _____
- Uses:
  - _____ variables
  - Interaction terms
- NOT same as _____ regression

8

Stepwise Regression

- All variables entered at one stage
- Software enters every _____ of variables
  - Seeks to _____ R^2
  - Capitalizes on chance characteristics of THIS sample ☹
  - Like throwing spaghetti against the wall to see what sticks ☹
- Not recommended _____ by some journals

9

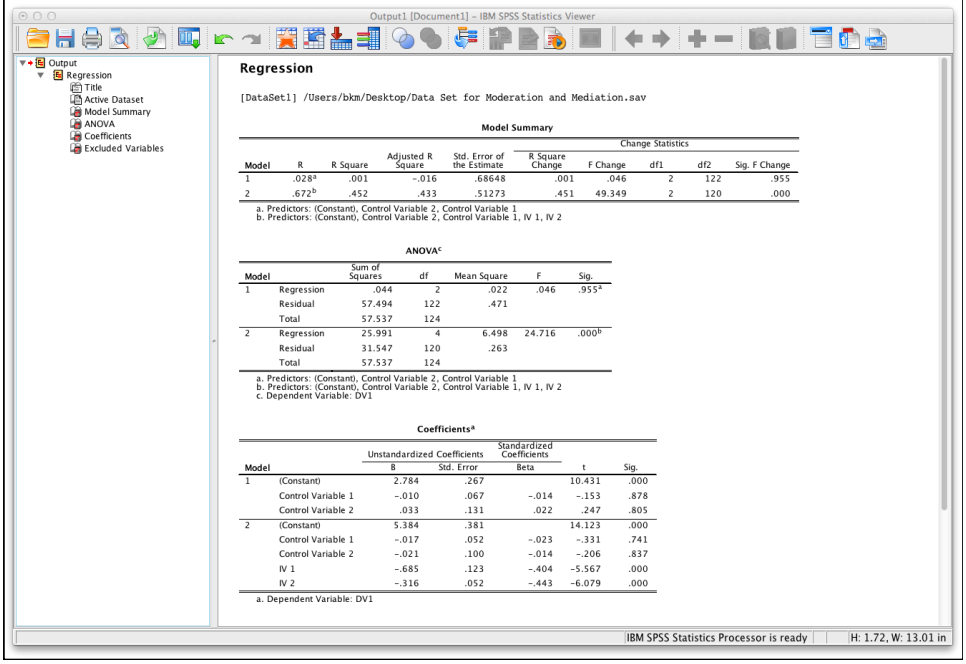
3. Run hierarchical regression in SPSS

10

SPSS Example #1

- Click "Analyze" / "Regression" / "Linear"
- Choose "DV1" as Dependent Variable
- Choose "Control1", "Control2" as IVs
  - Is there theoretical rationale for these control variables?
  - However, no need for stated hypotheses
- Click "Next"
- Choose "IV1", "IV2" as Independent Variables
- Click "Statistics"
- Put check in box for "R squared change"
- Click "Continue", click "OK"

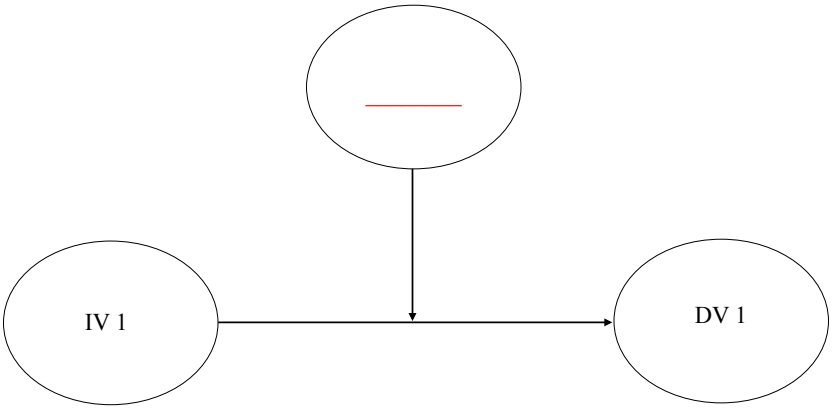
11



4. Compute Interaction Terms

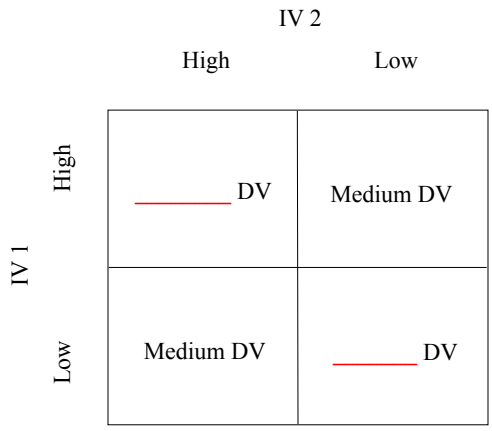
13

Review Diagram



14

Reconceptualizing Diagram as 2 x 2



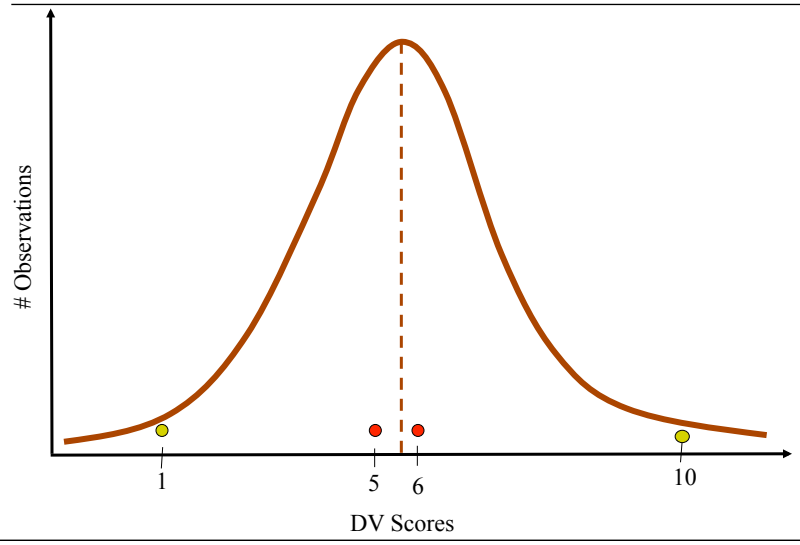
15

Caution!

- Never _____ continuous variables (e.g. cutting in half at the median)
- Results in loss of useful information
- Treats scores close to each other as if far from each other
- In _____, no need to dichotomize since all IVs are **categorical**
- _____ subsumes ANOVA
- But, different method of creating interaction terms needed

16

Problem:
_____ Continuous Variable



17

Moderation Effects

- Strength of relationship between IV and DV _____ _____ Moderator
  - If one is low on the moderator (IV2) the correlation between IV1 and DV is different from same correlation for those high on moderator (IV2)
  - In regression
    - correlations manifest themselves as _____, or
    - _____ of lines
  - So…slope of lines is different based upon gender
- Caution: Moderating variables is ALSO **an IV**

18

Moderation Effects (cont'd)

- MUST include bot_____(i.e. both IVs) in regression before including _____(cross product of IV1 and IV2)
  - Hierarchical regression model 1 has both main effects but **NO** interaction term
  - Hierarchical regression model 2 has main effects **PLUS** interaction term

19

Example Hypothesis

1. There is an interaction effect between IV1 and IV2 in the prediction of DV, such that high levels of IV1 combined with high levels of IV2 will lead to higher levels of DV than will low levels of either or both of IV1 and IV2.

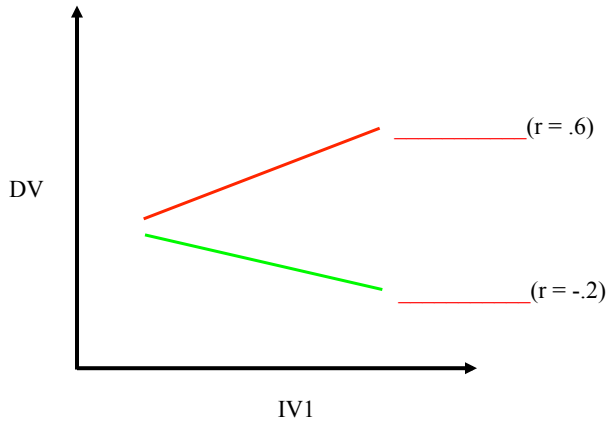
20

Alternatively Written Hypotheses

2. The strength of the relationship between IV1 and DV depends upon IV2, such that IV1 is strongest when IV2 is high and weakest when IV2 is low.
3. There is a positive relationship between IV1 and DV especially if IV2 is also high.

21

Plotting Interactions



22

Creating Multiplicative Terms

- Most (but not all!) IVs in regression are _____
- For 2 IVs, create third term that serves as interaction
- Simply _____ both IVs to create new term
- *Heads up:* sometimes new term is _____ with component terms

23

Collinearity Issues

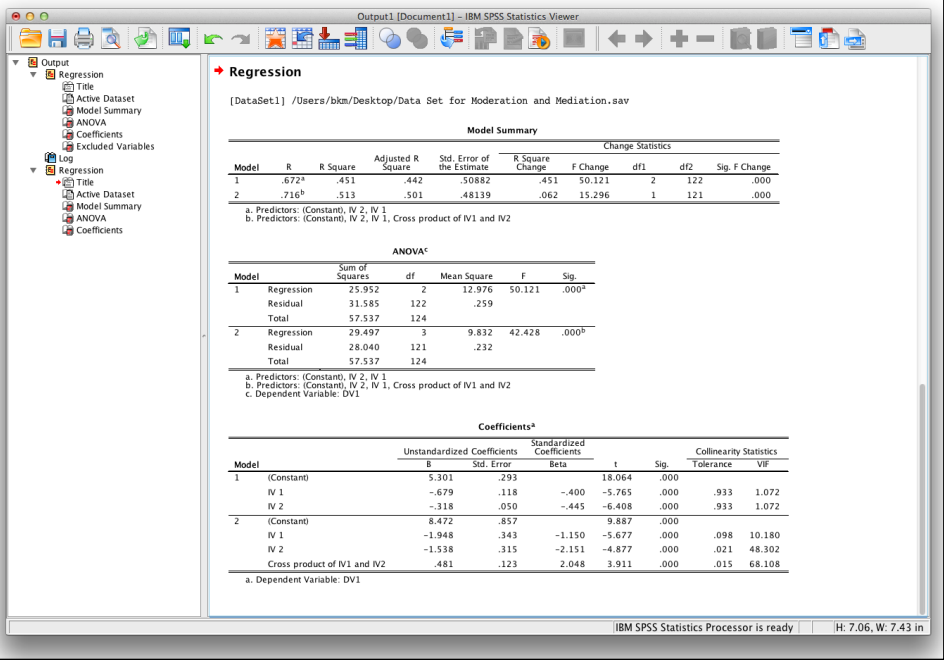
- Product of two variables is almost always collinear with its _____
- When two variables are so strongly correlated with each other that they affect interpretation of regression
  - Can make _____ exceed **± 1.0** ☹
- Tolerance / _____ (VIF)
  - Tolerance is reciprocal of VIF
  - Collinearity indicated if:
    - Tolerance < .10, or…
    - …VIF > 10

24

SPSS Example #2: Moderated Multiple Regression

- Click Analyze/Regression/Linear or Dialog Recall button
- Click "Reset" to start with all new variables (i.e. remove the control variables previously used)
- Choose "DV1" as DV
- Choose "IV1" and "IV2" as IVs
- Click "Next"
- Choose new term "IV1xIV2" (interaction term already created from product of two IVs above) as only IV in second step of hierarchical regression model
- Click "Statistics"
- Put check in boxes for "R square change" and "Collinearity Diagnostics"
- Click "Continue", click "OK"

25



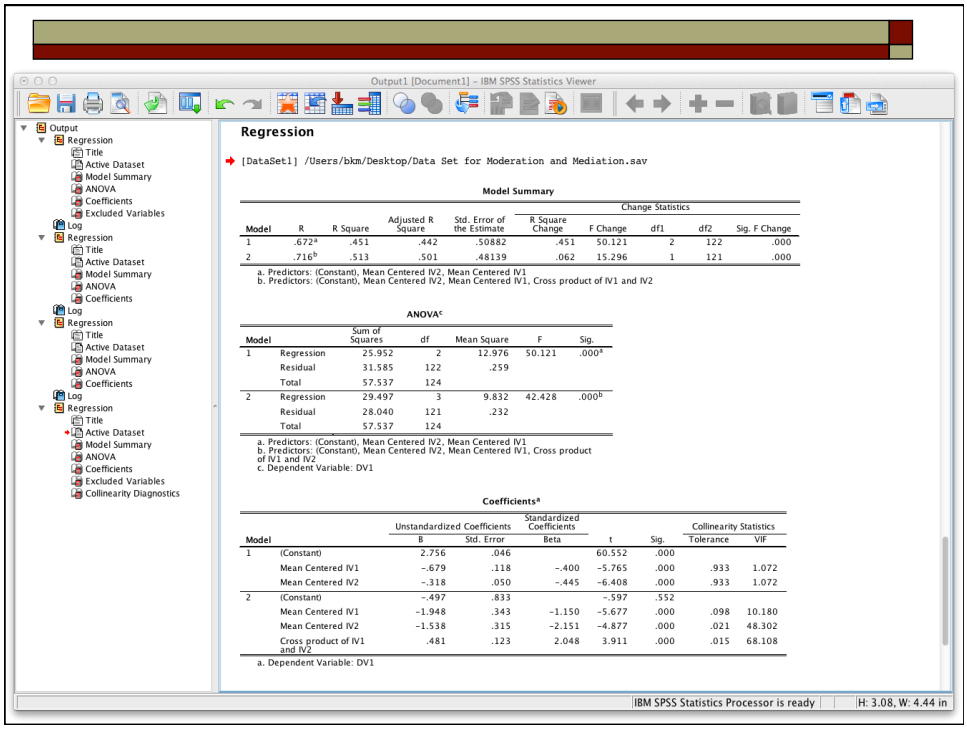
5. Mean Centering Variables

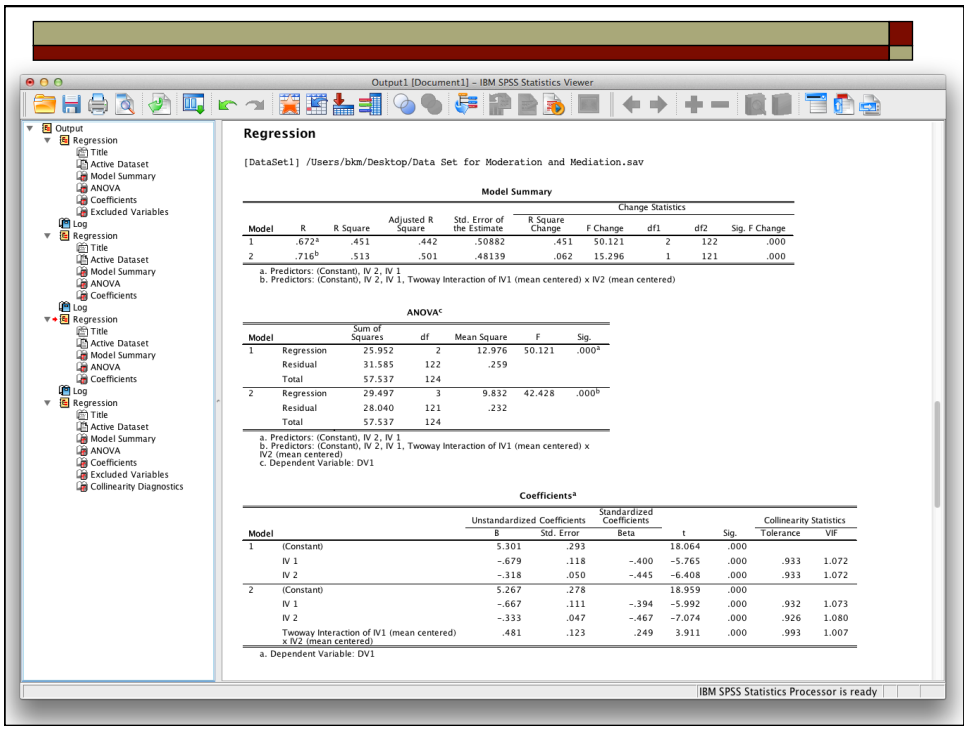
27

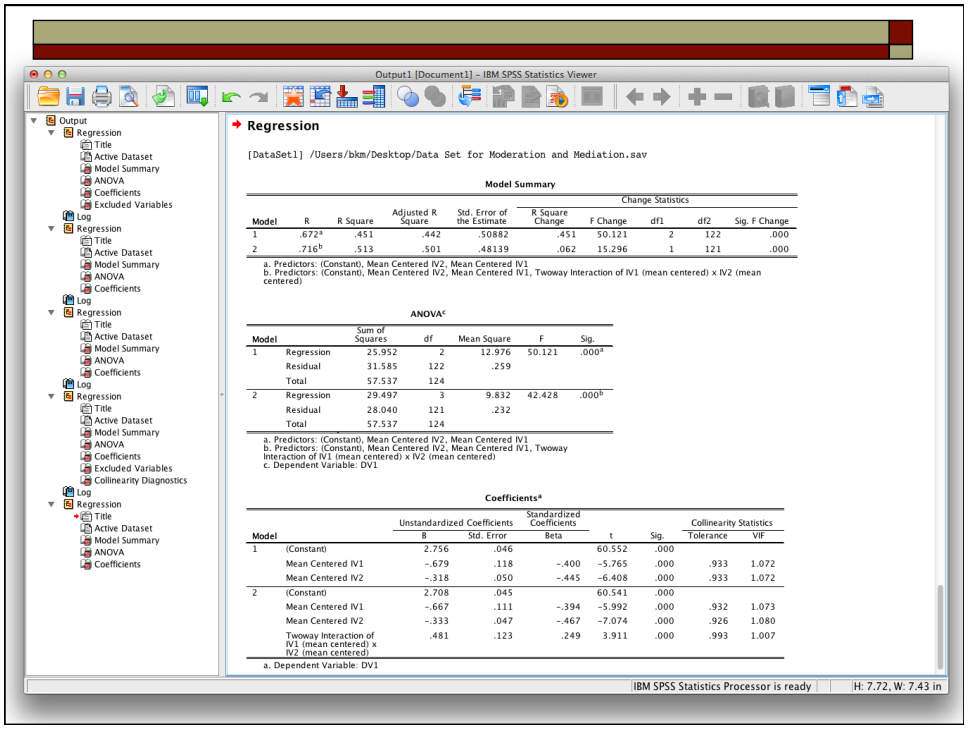
Mean Centering Variables

- Mean centering of variables reduces impact of _____
- This is NOT same as _____ variables!
- Allows for better interpretation of regression weights
- Requires:
  - Calculation of _____ of the variable
  - Creation of new variable that…
  - …is _____ between measured variable and mean of variable

28


Output1 [Document1] – IBM SPSS Statistics Viewer
Output
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Excluded Variables
Log
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Log
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Log
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Excluded Variables
Collinearity Diagnostics
Regression
[DataSet1] /Users/bkm/Desktop/Data Set for Moderation and Mediation.sav
Model Summary
Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics: R Square Change | F Change | df1 | df2 | Sig. F Change
1 | .672[a] | .451 | .442 | .50882 | .451 | 50.121 | 2 | 122 | .000
2 | .716[b] | .513 | .501 | .48139 | .062 | 15.296 | 1 | 121 | .000
a. Predictors: (Constant), Mean Centered IV2, Mean Centered IV1
b. Predictors: (Constant), Mean Centered IV2, Mean Centered IV1, Cross product of IV1 and IV2
ANOVA[c]
Model | | Sum of Squares | df | Mean Square | F | Sig.
1 | Regression | 25.952 | 2 | 12.976 | 50.121 | .000[a]
Residual | 31.585 | 122 | .259
Total | 57.537 | 124
2 | Regression | 29.497 | 3 | 9.832 | 42.428 | .000[b]
Residual | 28.040 | 121 | .232
Total | 57.537 | 124
a. Predictors: (Constant), Mean Centered IV2, Mean Centered IV1
b. Predictors: (Constant), Mean Centered IV2, Mean Centered IV1, Cross product of IV1 and IV2
c. Dependent Variable: DV1
Coefficients[a]
Model | | Unstandardized Coefficients: B | Unstandardized Coefficients: Std. Error | Standardized Coefficients: Beta | t | Sig. | Collinearity Statistics: Tolerance | Collinearity Statistics: VIF
1 | (Constant) | 2.756 | .046 | | 60.552 | .000 | |
Mean Centered IV1 | -.679 | .118 | -.400 | -5.765 | .000 | .933 | 1.072
Mean Centered IV2 | -.318 | .050 | -.445 | -6.408 | .000 | .933 | 1.072
2 | (Constant) | -.497 | .833 | | -.597 | .552 | |
Mean Centered IV1 | -1.948 | .343 | -1.150 | -5.677 | .000 | .098 | 10.180
Mean Centered IV2 | -1.538 | .315 | -2.151 | -4.877 | .000 | .021 | 48.302
Cross product of IV1 and IV2 | .481 | .123 | 2.048 | 3.911 | .000 | .015 | 68.108
a. Dependent Variable: DV1
IBM SPSS Statistics Processor is ready
H: 3.08, W: 4.44 in


Output1 [Document1] – IBM SPSS Statistics Viewer
Output
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Excluded Variables
Log
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Log
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Log
Regression
Title
Active Dataset
Model Summary
ANOVA
Coefficients
Excluded Variables
Collinearity Diagnostics
Regression
[DataSet1] /Users/bkm/Desktop/Data Set for Moderation and Mediation.sav
Model Summary
Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics: R Square Change | F Change | df1 | df2 | Sig. F Change
1 | .672[a] | .451 | .442 | .50882 | .451 | 50.121 | 2 | 122 | .000
2 | .716[b] | .513 | .501 | .48139 | .062 | 15.296 | 1 | 121 | .000
a. Predictors: (Constant), IV 2, IV 1
b. Predictors: (Constant), IV 2, IV 1, Twoway Interaction of IV1 (mean centered) x IV2 (mean centered)
ANOVA[c]
Model | | Sum of Squares | df | Mean Square | F | Sig.
1 | Regression | 25.952 | 2 | 12.976 | 50.121 | .000[a]
Residual | 31.585 | 122 | .259
Total | 57.537 | 124
2 | Regression | 29.497 | 3 | 9.832 | 42.428 | .000[b]
Residual | 28.040 | 121 | .232
Total | 57.537 | 124
a. Predictors: (Constant), IV 2, IV 1
b. Predictors: (Constant), IV 2, IV 1, Twoway Interaction of IV1 (mean centered) x IV2 (mean centered)
c. Dependent Variable: DV1
Coefficients[a]
Model | | Unstandardized Coefficients: B | Unstandardized Coefficients: Std. Error | Standardized Coefficients: Beta | t | Sig. | Collinearity Statistics: Tolerance | Collinearity Statistics: VIF
1 | (Constant) | 5.301 | .293 | | 18.064 | .000 | |
IV 1 | -.679 | .118 | -.400 | -5.765 | .000 | .933 | 1.072
IV 2 | -.318 | .050 | -.445 | -6.408 | .000 | .933 | 1.072
2 | (Constant) | 5.267 | .278 | | 18.959 | .000 | |
IV 1 | -.667 | .111 | -.394 | -5.992 | .000 | .932 | 1.073
IV 2 | -.333 | .047 | -.467 | -7.074 | .000 | .926 | 1.080
Twoway Interaction of IV1 (mean centered) x IV2 (mean centered) | .481 | .123 | .249 | 3.911 | .000 | .993 | 1.007
a. Dependent Variable: DV1
IBM SPSS Statistics Processor is ready



6. Graphing Interactions

32

Graphing Interactions

- DV on _____(Y) axis
- IV on _____(X) axis
- Calculate values of Moderator (other IV) 1 sd above and below mean
- Calculate values of IV 1 sd above and below mean
- Insert value of Moderator at lower sd in _____ _____
- Insert value of Moderator at upper sd in regression equation

33

Using Regression Formula to Plot Interactions

Use the regression equation:

$$Y = 2.708 - .667IV_1 - .333IV_2 + .481IV_1IV_2$$

Find 4 different values (points) for Y

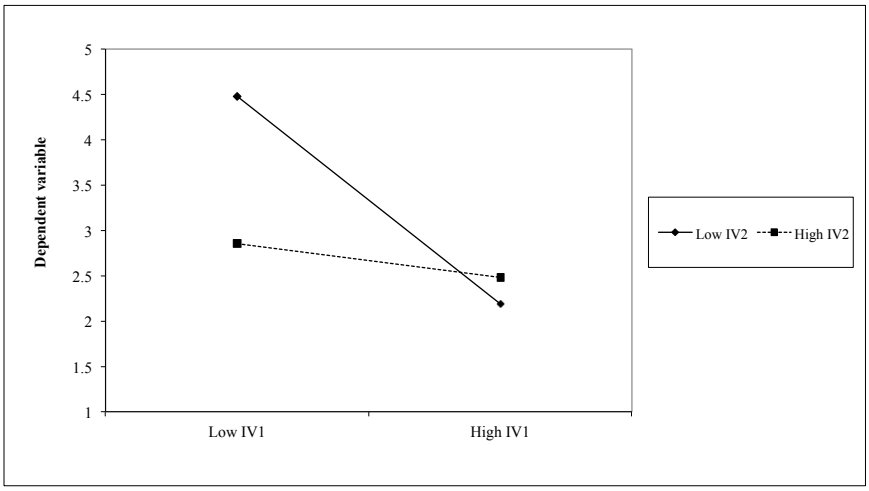
34

Using Regression Formula (cont'd)

- Calculate one equation for:
  - Insert value of IV2 at 1sd above mean
  - Insert value of IV1 at 1sd above mean
- Calculate another equation for:
  - Insert value of IV2 at 1sd below mean
  - Insert value of IV1 at 1sd below mean
- Calculate one equation for:
  - Insert value of IV2 at 1sd above mean
  - Insert value of IV1 at 1sd below mean
- Calculate another equation for:
  - Insert value of IV2 at 1sd below mean
  - Insert value of IV1 at 1sd above mean
- Connect the dots (i.e. draw the line segments)

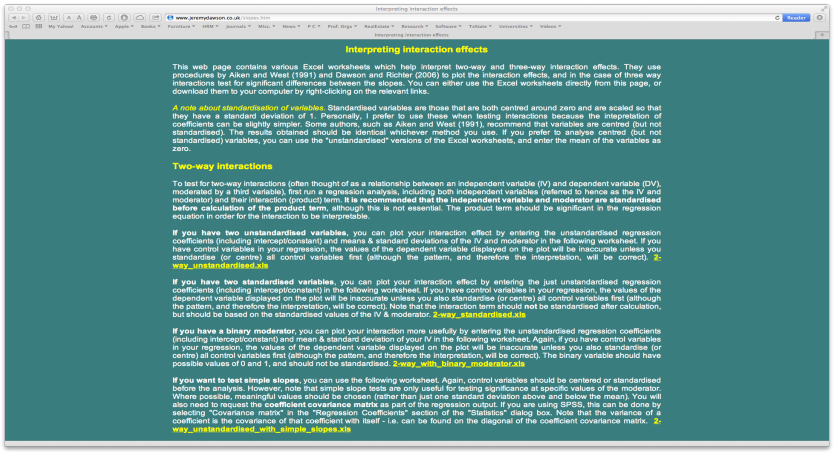
35

Sample Graph



Using Online Calculators Instead

- See http://www.jeremydawson.co.uk/slopes.htm



Interpreting interaction effects

This web page contains various Excel worksheets which help interpret two-way and three-way interaction effects. They use procedures by Aiken and West (1991) and Dawson and Richter (2006) to plot the interaction effects, and in the case of three way interactions test for significant differences between the slopes. You can either use the Excel worksheets directly from this page, or download them to your computer by right-clicking on the relevant links.

A note about standardisation of variables. Standardised variables are those that are both centred around zero and are scaled so that they have a standard deviation of 1. Personally, I prefer to use these when testing interactions because the interpretation of coefficients can be slightly simpler. Some authors, such as Aiken and West (1991), recommend that variables are centred (but not standardised). The results obtained should be identical whichever method you use. If you prefer to analyse centred (but not standardised) variables, you can use the "unstandardised" versions of the Excel worksheets, and enter the mean of the variables as zero.

Two-way interactions

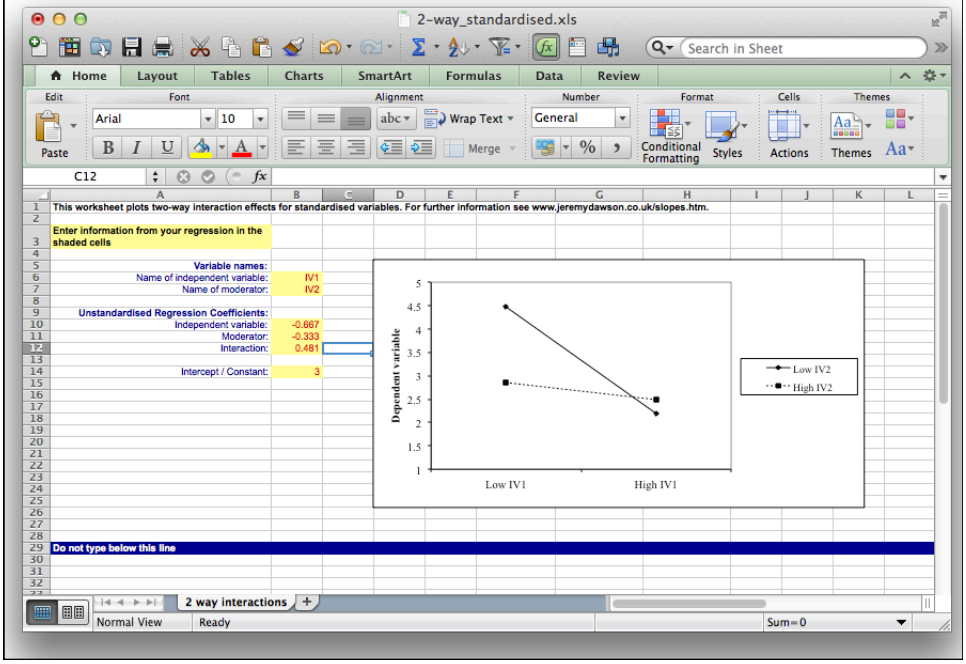
To test for two-way interactions (often thought of as a relationship between an independent variable (IV) and dependent variable (DV), moderated by a third variable), first run a regression analysis, including both independent variables (referred to hence as the IV and moderator) and their interaction (product) term. **It is recommended that the independent variable and moderator are standardised before calculation of the product term**, although this is not essential. The product term should be significant in the regression equation in order for the interaction to be interpretable.

If you have two unstandardised variables, you can plot your interaction effect by entering the unstandardised regression coefficients (including intercept/constant) and means & standard deviations of the IV and moderator in the following worksheet. If you have control variables in your regression, the values of the dependent variable displayed on the plot will be inaccurate unless you standardise (or centre) all control variables first (although the pattern, and therefore the interpretation, will be correct). 2-way_unstandardised.xls

If you have two standardised variables, you can plot your interaction effect by entering the just unstandardised regression coefficients (including intercept/constant) in the following worksheet. If you have control variables in your regression, the values of the dependent variable displayed on the plot will be inaccurate unless you also standardise (or centre) all control variables first (although the pattern, and therefore the interpretation, will be correct). Note that the interaction term should **not** be standardised after calculation, but should be based on the standardised values of the IV & moderator. 2-way_standardised.xls

If you have a binary moderator, you can plot your interaction more usefully by entering the unstandardised regression coefficients (including intercept/constant) and mean & standard deviation of your IV in the following worksheet. Again, if you have control variables in your regression, the values of the dependent variable displayed on the plot will be inaccurate unless you also standardise (or centre) all control variables first (although the pattern, and therefore the interpretation, will be correct). The binary variable should have possible values of 0 and 1, and should not be standardised. 2-way_with_binary_moderator.xls

If you want to test simple slopes, you can use the following worksheet. Again, control variables should be centered or standardised before the analysis. However, note that simple slope tests are only useful for testing significance at specific values of the moderator. Where possible, meaningful values should be chosen (rather than just one standard deviation above and below the mean). You will also need to request the **coefficient covariance matrix** as part of the regression output. If you are using SPSS, this can be done by selecting "Covariance matrix" in the "Regression Coefficients" section of the "Statistics" dialog box. Note that the variance of a coefficient is the covariance of that coefficient with itself - i.e. can be found on the diagonal of the coefficient covariance matrix. 2-way_unstandardised_with_simple_slopes.xls



Things We Learned

1. Mediation and moderation are very different
2. Hierarchical regression is NOT stepwise regression
3. Interpreting hierarchical regression output is easy
4. Computing interaction terms is just simple multiplication
5. Mean centered variables offset collinearity
6. Graphing interactions is easier with Excel

39

That's all folks!!!

BrianMillerPhD@gmail.com

40