http://www.unt.edu/rss/class/Jon/SPSS\_SC/Module9/Cars\_SimpleSlopesAnalysis.sps

#

###############################################################################################################

#

#

############ Simple Slopes Analysis during Testing of Moderation with Regression ############

#

# Using the "Cars.sav" dataset which is available here:

# http://www.unt.edu/rss/class/Jon/SPSS\_SC/Module3/Cars.sav

# Once the data set has been loaded, save the dataset as a different file name (e.g. Cars2.sav).

DATASET ACTIVATE DataSet1.

# We will be using only the following variables: mpg weight accel

# Delete the variables we will not be using (engine, horse, year, origin, cylinder, filter$).

# Delete the 8 cases with missing values on MPG.

# Should now have 398 cases of complete data for the three variables of interest (mpg, weight, accel).

### Our general model in this example is mpg (outcome) predicted by weight (predictor) with accel (moderator).

# First, we need to center both of our predictor variable (weight) and moderator variable (accel).

# Start by getting the mean of each.

DATASET ACTIVATE DataSet1.

DESCRIPTIVES VARIABLES=weight accel

/STATISTICS=MEAN SUM STDDEV VARIANCE RANGE MIN MAX SEMEAN KURTOSIS SKEWNESS.

# Next, we subtract the mean of each variable from itself to create the centered variables.

COMPUTE CenteredWeight=weight - 2960.37.

EXECUTE.

COMPUTE CenteredAccel=accel - 15.54.

EXECUTE.

# Next, we need to create the interaction term, also known as the product term because; it

# is the product of the centered predictor and centered moderator.

COMPUTE InteractionWeAc=CenteredWeight \* CenteredAccel.

EXECUTE.

# Now, we are ready to conduct the multiple regression to determine if we have a significant interaction effect.

# Remember: (1) enter the centered preditor and centered moderator as one block and then enter the centered

# predictor, centered moderator, and interaction term in a second block -- this will allow us to see if the R-square-

# change is significant which, with a significant coefficient for the interaction, would mean a significant interaction

# is present. \*This is not the same as conducting a stepwise regression.

# Please NOTE: You should also remember to specify in the 'Statistics...' that you want the R-squared-change,

# the Covariance matrix (of the coefficients), Descriptives, and Part and partial correlations. You may also want

# to specify the usual plots/graphs; such as the standardized residual vs. predicted plot, histogram, and normality plots.

REGRESSION

/DESCRIPTIVES MEAN STDDEV CORR SIG N

/MISSING LISTWISE

/STATISTICS COEFF OUTS BCOV R ANOVA CHANGE ZPP

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT mpg

/METHOD=ENTER CenteredWeight CenteredAccel

/METHOD=ENTER CenteredWeight CenteredAccel InteractionWeAc

/RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).

# First, notice that the correlations between the interaction term and its component (centered) parts are trivial

# in magnitude; with (r = -.356) between InteractionWeAc and CenteredWeight, and with (r = .173) between

# InteractionWeAc and CenteredAccel. This indicates that centering was successful in decreasing multicollinearity.

# The multiple correlation for our interaction model was .826 and the regression equation was:

# y = 23.091 + (-.00738\*CenteredWeight) + (.38335\*CenteredAccel) + (-.00045\*CenteredWeight\*CenteredAccel)

# Examining the multiple correlation (and multiple correlation squared) for the 'main effects model' (containing

# only the centered predictor and centered moderator) in comparison to the 'interaction effects model' (containing

# all three terms; centered predictor, centered moderator, & interaction term), we find a significant F-value (F = 21.586)

# which indicates a significant improvement. In other words, the change in R-square from .666 to .683 = .017 represents

# a significant improvement in the amount of variance explained by our model. However, 1.7% really is not a great

# improvement. We also see that the B-weight (un-standardized regression coefficient) is significant (t = -4.646, p < .000).

# As a reminder, that t-value is simply the B-weight (-.00045) divided by its standard error (.00010). This significant t-value

# also reinforces the idea that we have a significant interaction; even though the coefficient value associated with it is

# quite small.

# Bottom line interpretation, thus far, is: we have a significant interaction.

# So, we need to explore the nature of that interaction using simple slopes analysis.

# First, calculate the simple slope for the 'low group' where low corresponds to 1 standard deviation below the mean of

# CenteredAccel (-1 SD = -2.77640).

# You may be confused because we centered the predictor and moderator; but check the output. Centering only changes the

# location of the mean, it does not change the standard deviation (see the Descriptive Statistics table at the top/beginning of

# the output file and the Descriptive Statistics table at the top of the Regression output).

# So, we use the regression equation from above and substitute -2.77640 (one standard deviation below the mean) in place

# of 'CenteredAccel', which results in the following equation:

# y = 23.091 + (-.00738\*CenteredWeight) + (.38335\*-2.77640) + (-.00045\*CenteredWeight\*-2.77640)

# which can be reduced down to give us the simple slope we wanted:

# y = 23.091 + (.38335\*-2.77640) + (-.00738\*CenteredWeight) + (-.00045\*CenteredWeight\*-2.77640)

# y = 23.091 + -1.064333 + [-.00738 + (-.00045\*-2.77640)]\*CenteredWeight

# y = 22.02667 + [-.00738 + .00124938]\*CenteredWeight

# y = 22.02667 + -.00613062\*CenteredWeight

# So, the simple slope for the 'low group' is -.00613062.

# Next, calculate the simple slope for the 'high group' where high corresponds to 1 standard deviation above the mean of

# CenteredAccel (+1 SD = 2.77640).

# y = 23.091 + (-.00738\*CenteredWeight) + (.38335\*2.77640) + (-.00045\*CenteredWeight\*2.77640)

# which can be reduced down to give us the simple slope we wanted:

# y = 23.091 + (.38335\*2.77640) + (-.00738\*CenteredWeight) + (-.00045\*CenteredWeight\*2.77640)

# y = 24.15533 + [-.00738 + (-.00045 \* 2.77640)]\*CenteredWeight

# y = 24.15533 + -0.00862938\*CenteredWeight

# So, the simple slope for the 'high group' is -.00862938.

# Second, we have to calculate the standard error for each simple slope/group.

# To do this; we need to refer to the 'Coefficient Correlations' table we specified in the output.

# We will need 3 things from the lowest row (Model 2) of this table.

# (1) the variance of the coefficient for CenteredWeight = .000000090 (which also corresponds to the squared coefficient

# for CenteredWeight [.00030 \* .00030]).

# (2) the variance of the coefficient for InteractionWeAc = .000000009 (which corresponds to the squared coefficient for

# CenteredAcceel [.00010 \* .00010] minus rounding error).

# (3) the COVARIANCE of the coefficients for CenteredWeight and InteractionWeAc = .000000009

# Now we can calculate the Standard Error of each group specific simple slope using the formula:

# std.err. = sqrt of [(1) + S\*S\*(2) + 2\*S\*(3)]

# which can also be written as:

# std.err. = sqrt of [.000000090 + S\*S\*.000000009 + 2\*S\*.000000009]

# where S corresponds to the group specific value for standard deviation; for the low group: S = -2.77640 and

# for the high group: S = 2.77640.

# For the low group: S = -2.77640.

# std.err. low = sqrt of [.000000090 + -2.77640\*-2.77640\*.000000009 + 2\*-2.77640\*.000000009]

# std.err. low = sqrt of [.0000001094]

# std.err. low = .0003307567

# For the high group: S = 2.77640.

# std.err. high = sqrt of [.000000090 + 2.77640\*2.77640\*.000000009 + 2\*2.77640\*.000000009]

# std.err. high = sqrt of [.0000002094]

# std.err. high = .0004576024

# Finally, we can conduct the t-tests to determine if each group specific simple slope is significantly different from zero.

# We calculate our t-value by simply dividing the slope by the standard error (for each group).

# BUT, remember we are doing multiple t-tests here, so we risk inflation of type 1 error rates. To overcome this, we

# must apply a simple Bonferroni correction to our alpha level (.05 / 2 comparisons = .025).

# For the low group:

# -.00613062 / .0003307567 = -18.53513

# For the high group:

# -.00862938 / .0004576024 = -18.85781

# To find our critical t-value, we use degrees of freedom (df) = N - k - 1 where N is the number of cases,

# and k is the number of predictors (including the predictor, moderator, & interaction term).

# df = 398 - 3 - 1 = 394

# This df yeilds a two-tailed critical t-value of 3.347468181818182 with alpha = .001.

# We can see that both simple slopes are significantly different from zero.

#####################################################################################################

#

############ GRAPHING ############

## Keep in mind; this is not typically done with continuous variables (regression setting) the way it is typically done with

## categorical variables (ANOVA setting). Also note that SPSS is a limiting factor here where as R would

## be much more suited for this type of graphing.

# Showing how cases with low Acceleration times differ from individuals with high Acceleration

# times when predicting MPG using Weight.

# REMEMBER ALSO: we are not using the centered variables, nor are we using the interaction term. We are

# merely creating a graph which 'should' show how individuals with low Acceleration times differ from individuals

# with high Acceleration times when predicting MPG using Weight.

# Using the mean of accel (non centered, original variable) and the 'Recode into Different Variables' function from the 'Transform' menu of the taskbar

# we can create a new variable which breaks Acceleration (accel variable) into two groups (below & above mean = 15.54).

RECODE accel (Lowest thru 15.54=1) (15.55 thru Highest=2) INTO AccelGroups.

EXECUTE.

# Now go into the data editor window and then the variable view tab (at bottom) to change the values and

# measurement scale of the AccelGroups variable. Set the values to: 1 = Low and 2 = High; then set the

# measurement scale to Nominal.

# Method 1:

# Go to Graphs, Scatter, Simple, and run a standard scatterplot graph function but with separate panels for

# each acceleration group.

GRAPH

/SCATTERPLOT(BIVAR)=weight WITH mpg

/PANEL ROWVAR=AccelGroups ROWOP=CROSS

/MISSING=LISTWISE.

# Then we can double click on the graph(s) to enter the chart editor. Once in the chart editor, right click on the

# points in one (or the other) scatter plot and select 'Add Fit Line at Total'.

# Notice both lines have a slope that is significantly different from zero (horizontal).

# Method 2: Showing the same thing; but with a single scatterplot.

# Go to graphs, choose Chart Builder..., then from the 'Gallery' tab, 'Choose from'

# and select 'Scatter/Dot'. Click and drag to the preview panel the 'Grouped Scatter' choice (multi-colored

# dots with no lines). Then drag and place the MPG variable on the Y-axis box, the Weight variable on

# the X-axis box and AccelGroups onto the 'Set color' box.

# Now click OK.

\* Chart Builder.

GGRAPH

/GRAPHDATASET NAME="graphdataset" VARIABLES=weight mpg AccelGroups MISSING=LISTWISE

REPORTMISSING=NO

/GRAPHSPEC SOURCE=INLINE.

BEGIN GPL

SOURCE: s=userSource(id("graphdataset"))

DATA: weight=col(source(s), name("weight"))

DATA: mpg=col(source(s), name("mpg"))

DATA: AccelGroups=col(source(s), name("AccelGroups"), unit.category())

GUIDE: axis(dim(1), label("Vehicle Weight (lbs.)"))

GUIDE: axis(dim(2), label("Miles per Gallon"))

GUIDE: legend(aesthetic(aesthetic.color.exterior), label("AccelGroups"))

SCALE: cat(aesthetic(aesthetic.color.exterior), include("1.00", "2.00"))

ELEMENT: point(position(weight\*mpg), color.exterior(AccelGroups))

END GPL.

# Again; we can double click on the graph to enter the chart editor. Once in the chart editor, right click on the

# points in the scatter plot and select 'Add Fit Line at Subgroups'.

# Notice both lines have a slope that is significantly different from zero (horizontal).

# Method 3: Showing the same thing; but with a single 3 dimensional scatterplot.

# Go to graphs, choose Chart Builder..., then from the 'Gallery' tab, 'Choose from'

# and select 'Scatter/Dot'. Click and drag to the preview panel the 'Grouped 3-D Scatter' choice (multi-colored

# dots in a grey box). Then drag and place the MPG variable on the Y-axis box, the Weight variable on

# the X-axis box, MPG on the Z-axis, and AccelGroups onto the 'Set color' box.

# Now click OK.

\* Chart Builder.

GGRAPH

/GRAPHDATASET NAME="graphdataset" VARIABLES=weight mpg accel AccelGroups MISSING=LISTWISE

REPORTMISSING=NO

/GRAPHSPEC SOURCE=INLINE.

BEGIN GPL

SOURCE: s=userSource(id("graphdataset"))

DATA: weight=col(source(s), name("weight"))

DATA: mpg=col(source(s), name("mpg"))

DATA: accel=col(source(s), name("accel"))

DATA: AccelGroups=col(source(s), name("AccelGroups"), unit.category())

COORD: rect(dim(1,2,3))

GUIDE: axis(dim(1), label("Time to Accelerate from 0 to 60 mph (sec)"))

GUIDE: axis(dim(2), label("Vehicle Weight (lbs.)"))

GUIDE: axis(dim(3), label("Miles per Gallon"))

GUIDE: legend(aesthetic(aesthetic.color.exterior), label("AccelGroups"))

SCALE: cat(aesthetic(aesthetic.color.exterior), include("1.00", "2.00"))

ELEMENT: point(position(accel\*weight\*mpg), color.exterior(AccelGroups))

END GPL.

# Notice with all these graphs, the slopes are similar, but each is significantly differnt from zero.

# The b-weight (non-standardized coefficient) for the interaction term in the regression was significant

# which tells us there is a significant interaction; the graph (which ever one is chosen) simply shows

# us what we discovered in the regression and to a lesser extent, simple slopes analysis.

#######################################################################################

# Some resources:

# Jaccard, J., Turrisi, R., & Wan, C. (1990). Interaction Effects in Multiple Regression. Sage University Paper

series on Quantitative Applications in the Social Sciences, 07-072. Newbury Park, CA: Sage.

# Wu & Zumbo (2008) pdf link available here:

# http://www.springerlink.com/content/2m6k0747k1q1w446/