

Simple Slopes

Following the discovery of a significant interaction, it is advisable to explore the nature of the interaction.¹ It is really not possible to understand the form that an interaction takes without at least plotting the results. For example, if there is a significant interaction between family income (X) and self-esteem (Z) in predicting elementary student aggression, it might be that higher income students are more likely to exhibit aggressive behavior when they have higher self-esteem (a dominance hypothesis). Or, conversely, perhaps it is the combination of lower-self-esteem and lower income that increases the chance of aggression. It is not possible to know which form the interaction takes without going further than the interaction test. The usual recommendation is to center the predictor variables by subtracting the mean of each and then computing the interaction term. I have not used centering in the following equations, however, to make the explanations less complicated.

Simple Slopes

Like simple (or “simple main”) effects tests following a significant interaction with ANOVA, we can investigate (probe) the effects of one independent variable (X) within levels of the other independent variable (Z). With continuous independent variables, probing implies that we examine the effect of X on Y for different values of Z , a simple regression line. When Z is continuous, it is customary to derive the simple regression slopes for few values of Z . Most commonly, arbitrary values of low (one standard deviation below the mean of Z), medium (mean of Z), and high (one standard deviation above the mean of Z) values, but any other theoretical values could be chosen. When Z is binary, simple slopes for values of 0 and 1 are usually used. The slopes can be easily computed from the interaction equation, although the computer usually does the work for us.

$$\hat{Y} = B_1X + B_2Z + B_3XZ + B_0$$

Some basic algebra modifies the equation, so the simple slopes (line equations) can be computed for certain values of Z . The final equation below looks like a simple regression equation with just X as a predictor as long as we know the values of B_1 , B_2 , and B_3 , and we pick a value for Z .

$$\begin{aligned}\hat{Y} &= [B_1X + B_3XZ] + [B_2Z + B_0] \\ &= [B_1 + B_3Z]X + B_2Z + B_0\end{aligned}$$

As a hypothetical example, assume our results for the interaction analysis give the following equation: $\hat{Y} = -2.3X + -.5Z + 1.6XZ + 3.0$ that self-esteem (Z) has a mean of 4.5 and a standard deviation of 1.0. The simple slope for low values of Z could be computed by inserting the value associated with -1 SD below the mean of self-esteem ($4.5 - 1.0 = 3.5$) in this last equation.

$$\begin{aligned}\hat{Y} &= [B_1 + B_3Z]X + [B_2Z + B_0] \\ \hat{Y} &= [-2.3 + (1.6)(3.5)]X + [(-.5)(3.5) + 3.0] \\ \hat{Y} &= 3.3X + 1.25\end{aligned}$$

So, the equation for line when Y is regressed on X for low values of Z (-1 SD below the mean of Z) is $\hat{Y} = 3.3X + 1.25$. I could then plot that line with some values of X from the data set. Using the same equation, plugging in the mean value of Z (4.5) and high values of Z ($4.5 + 1.0 = 5.5$), we get:

¹ There may be instances in which the interaction is marginally significant or for other theoretical reasons, the researcher wishes to compute and test simple slopes, but, typically, follow-ups to the interaction test only occur when the interaction is significant.

For the mean of Z

$$\hat{Y} = [B_1 + B_3Z]X + [B_2Z + B_0]$$

$$\hat{Y} = [-2.3 + (1.6)(4.5)]X + [(-.5)(4.5) + 3.0]$$

$$\hat{Y} = 4.9X + .75$$

For high Z :

$$\hat{Y} = [B_1 + B_3Z]X + [B_2Z + B_0]$$

$$\hat{Y} = [-2.3 + (1.6)(5.5)]X + [(-.5)(5.5) + 3.0]$$

$$\hat{Y} = 6.5X + .25$$

I entered some hypothetical values for X for each of these equations (for my hypothetical study) and found the interaction looks like the figure below.



The plot of the results would support my dominance hypothesis and, because the interaction was significant, we would expect to see non-parallel lines as we do here. The simple slopes just describe the relationship between income and aggression for low, medium, and high values of self-esteem. Each simple slope may or may not differ significantly from zero.

Significance Testing

In addition to computing and plotting the simple slopes, another step in probing the interaction is to test each of the simple slopes for significance. It might be in the example results above, for instance, that the low self-esteem line is not significantly different from zero. The test of each coefficient is a t test formed by the ratio of the coefficient to its standard error, $t_{\text{BatZ}} = B_{\text{atZ}}/SE_{\text{BatZ}}$.

The standard error computation (see Cohen, Cohen, West, & Aiken, 2003, p. 273) requires information from the covariance matrix of regression coefficients, which has the square of the standard errors on the diagonal and the estimates of the covariances among the coefficients on the off diagonal. The estimates of the covariances between any two coefficients represents the estimated association between the coefficient values across the sampling distribution. The covariance matrix is not usually printed by default in statistical packages but can be obtained when requested (`bCOV` in SPSS and `vcov` in R).

Standardized Coefficients

Just as with any regression coefficient, standardized coefficients for the simple slopes can be computed by multiplying by the ratio of standard deviation for X and Y .

$$\beta_{lowZ} = B_{lowZ} \left(\frac{sd_x}{sd_y} \right)$$

Computer Methods

It is far easier to get statistical software to perform the computations, but, unfortunately, most software programs have not created prepackaged procedures for plotting or testing simple slopes. I have created an SPSS macro, which is posted on my webpage that does all of the steps for you for two-way continuous or continuous by binary interactions, <http://web.pdx.edu/~newsomj/macros.htm>. The process macro (for SPSS and R) by Andrew Hayes also tests simple slopes and provides plot points for manual graphing, <http://www.processmacro.org/download.html>. The `pequod` package in R, <https://www.rdocumentation.org/packages/pequod/versions/0.0-5>, which has now been removed from CRAN, is convenient for computing simple slopes, their tests, and plots the interaction. Jeremy Dawson has some very useful Excel sheets for graphing interactions, including three-way interactions, <http://www.jeremydawson.co.uk/slopes.htm>. Finally, the simple slopes can be obtained through recomputation of the predictor variables, capitalizing on the interpretation of the lower-order ("main effect") coefficient (Aiken & West, 1991). The lower-order coefficient for X in the interaction model already represents the simple slope for X at the mean of Z whenever the Z predictor is centered. The intercept in the interaction model is also the intercept for the simple slope for the simple regression equation for the mean of Z if Z has been centered. In the special case when Z is binary and coded 0 and 1, the simple slope for $Z = 0$ group is equal to the slope for the lower-order term for X in the full interaction equation. This equivalence, in general, suggests another method of obtaining simple effects coefficients and their tests—recode the moderator by adding 1 SD to the centered moderator, z (i.e., $z_{low} = z + s_z$), run the regression and look at the lower order coefficient for the X variable (B_1), which is now the simple effect for the low Z value. Similarly, retest the model after centering the Z variable and subtracting 1 SD from the centered moderator, z (i.e., $z_{high} = z - s_z$) and the new coefficient for the X variable (B_1) is the simple slope for the high Z value (Aiken & West, 1991, p. 20).