

GRAPHING THE INTERACTION BETWEEN TWO QUANTITATIVE VARIABLES

Throughout this document, the following symbols will be used

X_1 : quantitative predictor variable 1

X_2 : quantitative predictor variable 2

C_i : a variable or set of variables used as statistical controls

Y : outcome variable

\hat{Y} : estimate of Y from regression model

The interaction between X_1 and X_2 is tested by estimating the terms of the following regression model:

$$\hat{Y} = a + b_1X_1 + b_2X_2 + b_3(X_1X_2) + \sum b_iC_i$$

where the number of variables being controlled (C) may be as small as 0 (in which case the last term disappears). An interaction between X_1 and X_2 exists if b_3 is statistically different from zero. It is often desirable to graphically depict this interaction by plotting the regression line for X_1 at various levels of X_2 .

A common procedure is to use values of X_1 and/or X_2 that are one standard deviation below the mean, the mean, and one standard deviation above the mean. To construct a graphical depiction of the interaction,

(1) Select at least two but preferably several values of X_1 that will be used in the figure on the X axis. Call the number of values used k . For example, if X_1 is measured on a 1 to 5 scale, you might use the values 1, 2, 3, 4, and 5. In this case, $k = 5$. Alternatively, you could use the mean of X_1 , one standard deviation below the mean of X_1 , and one standard deviation above the mean of X_1 , in which case $k = 3$. Now, create a data file containing $3k$ rows, such that the first k rows contain, for a variable column named $X1$, the values of X_1 selected to be graphed in sequential order. Repeat this for the next 2 sets of k rows. For example, if $k = 5$ using numbers 1 through 5, the first row would have $X1 = 1$, the second would have $X1 = 2$, etc., the fifth row would have $X1 = 5$, the sixth row would have $X1$ set to 1, etc., the tenth row would have $X1$ set to 5, the eleventh row would have $X1$ set to 1, and so forth.

(2) For the first k rows, create a new variable $X2$ and set it equal to the value of X_2 in your data that is one standard deviation below the mean of X_2 . For the next k rows, set $X2$ to the value of X_2 in your data that is equal to the mean of X_2 . For the final k rows, set $X2$ variable equal to the value of X_2 that is one standard deviation above the mean.

(3) Create additional variables that contain the mean of the covariates. For example, if you have two covariates Z and W , create two columns that contain, for every case, the mean of Z in the first column and the mean of W in the second column. (If you have no covariates in the model, then this step is unnecessary).

Suppose, for example, that you were interested in plotting the interaction between X_1 (measured on a 1 to 5 scale) and X_2 (Mean = 5, -1 SD = 3, +1 SD = 7) controlling for Z and W (with means of 6 and 8.5, respectively), your data file might look like this:

X1	X2	Z	W
1	3	6	8.5
2	3	6	8.5
3	3	6	8.5
4	3	6	8.5
5	3	6	8.5
1	5	6	8.5
2	5	6	8.5
3	5	6	8.5
4	5	6	8.5
5	5	6	8.5
1	7	6	8.5
2	7	6	8.5
3	7	6	8.5
4	7	6	8.5
5	7	6	8.5

(4) Using the regression model estimated on the original data (not the data in the table you just created), generate \hat{Y} but *applying that regression model to the data in the table you did just create*. To do this, use a compute statement. For example, suppose the regression model from the analysis of the original data was

$$\hat{Y} = 3.4 - 8.3X_1 - 0.2X_2 + 10.3(X_1X_2) - 0.3W + 0.1Z$$

Then the compute command in SPSS would be

```
compute yhat = 3.4-8.3*X1-0.2*X2+10.3*X1*X2-0.3*W+0.1*Z.
```

where the regression coefficients and intercept in the compute command are from the regression model estimated in your data (again, not the data in the table you just created). This will create a new variable in the data set you created called “yhat.” These are the predicted values of Y for cases with the predictor values used in the data set you created above.

(6) The following command will generate a scatterplot between X_1 and Y , using different markers for the three different values of X_2 you chose.

```
GRAPH
  /SCATTERPLOT(BIVAR)=x1 WITH yhat BY x2.
```

Using the editing features of SPSS, it is possible to edit the graph the command above produces to create a figure that looks like this:

