MLE: Lab 5

Shahryar Minhas

January 29, 2015

Interaction Effects

A conditional hypothesis is one in which a relationship between two or more variables depends on the value of one or more other variables. Examples:

- An increase in X is associated with an increase in Y when condition Z is met, but not when condition Z is absent.
- X has a positive effect on Y that gets stronger as Z increases.

Islam & Authoritarianism

Fish (2002) argues that predominantly Muslim societies are distinctly disadvantaged in democratization. To test this he regresses Freedom House scores (`fhrev`) on an Islamic religious tradition variable (`muslim`) that takes the value of one when a country is predominantly Muslim and zero otherwise. He also goes on to add a number of additional controls for level of development (`income`), sociocultural division (`elf`), economic growth (`growth`), British colonial heritage (`britcol`), Communist heritage (`postcom`), and OPEC membership (`opec`).

We will rerun his model and run one of our own that includes an interaction term between Islamic religious tradition and level of economic development. Our hypothesis is that higher levels of economic development should be associated with higher Freedom House scores for non-Muslim countries, but that same relationship should not hold for Muslim countries.

```
## [1] "Fish's Original Model"
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.169      0.676   0.250    0.803
## muslim        -1.245      0.254  -4.892    0.000
## income         1.401      0.178   7.853    0.000
## elf           -0.316      0.389  -0.813    0.418
## growth         0.067      0.045   1.504    0.135
## britcol        0.246      0.267   0.924    0.357
## postcom        0.226      0.314   0.720    0.473
## opec          -1.332      0.455  -2.926    0.004
```

Interpretation

When interpreting the effects of a covariate we need to keep the conditional relationship in mind. In Fish's original model, to interpret the effect of a one unit change in income on Freedom House scores we would just note that β_{income} equaled 1.4.

After incorporating the interaction effect, however, it is no longer as straightforward because:

$$\frac{\partial fhrev}{\partial income} = \hat{\beta}_{income} + \hat{\beta}_{muslim:income} \times muslim$$

We can see that now the effect of income on Freedom House scores is contingent on the value of $\beta_{muslim:income}$ and `muslim`. If `muslim` equals zero, the effect of income on Freedom House scores can be understood through just β_{income}, but if muslim equals one then the effect of income is measured through $\beta_{income} + \beta_{muslim:income}$.

The necessity of taking the interactive relationship into account is also true for how we estimate the effect of a country having an Islamic religous tradition as well.

$$\frac{\partial fhrev}{\partial muslim} = \hat{\beta}_{muslim} + \hat{\beta}_{muslim:income} \times income$$

The best way to understand an interactive effect is through a visualization, so lets do that now. On the y-axis, we will have predicted FH scores, and on the x-axis we will have the range of income values in the Fish dataset. To incorporate the Muslim variable, we will use two different colored lines to show the conditional effect of a country having an Islamic religious tradition.

Are these Marginal Effects Significant?

In the figure above, we have done a pretty decent job in showing how our interactive effect works in predicting FH scores across the different values of economic development and religious tradition. However, before we can fully understand the substantive meaning of this interaction effect we also need to take into account the standard errors around the interactive effect. This will help us to answer the question of whether marginal effect of our covariate is significant.

From above we know that:

$$\frac{\partial fhrev}{\partial income} = \hat{\beta}_{income} + \hat{\beta}_{muslim:income} \times muslim$$

Lets denote the `muslim` variable by M and `income` by I, then we can express the variance of $\frac{\partial fhrev}{\partial income}$ as:

$$\begin{aligned} Var(\frac{\partial fhrev}{\partial I}) &= Var(\hat{\beta}_I + \beta_{M:I} \times M) \\ &= Var(\hat{\beta}_I) + Var(\hat{\beta}_{M:I} \times M) + 2Cov(\hat{\beta}_I, \beta_{M:I} \times M) \\ &= Var(\hat{\beta}_I) + M^2\, Var(\hat{\beta}_{M:I}) + 2\, M\, Cov(\hat{\beta}_I, \beta_{M:I}) \end{aligned}$$

We can summarize what we have learned so far in a neat 2 × 2 (remind yourself that the standard error is just the square root of the estimated variance):

	M = 0	M = 1
Marginal Effect of Income	$\hat{\beta}_I$	$\hat{\beta}_I + \hat{\beta}_{M:I}$
Standard Error	$\sqrt{var(\hat{\beta}_I)}$	$\sqrt{var(\hat{\beta}_I) + M^2 var(\hat{\beta}_{M:I}) + 2M cov(\hat{\beta}_I, \hat{\beta}_{M:I})}$

Now lets do a simple numerical example for how to get a confidence interval around the marginal effect. Before we can start we need to get the necessary parameter values, specifically, we need the coefficient estimates, variances, and covariances. We can find each of these by using the variance-covariance martix from our regression results, we can access this matrix by calling the `vcov` function on our model:

The diagnoals of this matrix are filled with the variance estimates of our regressions and the off-diagnoals their covariances. From this matrix we can pull out the necessary information we need to calculate the confidence interval around the marginal effect of income on FH scores.

Marginal effect of Income on FH scores when muslim = 0: $\hat{\beta}_I + \hat{\beta}_{M:I} \times 0 = ?$

Corresponding standard error: $\sqrt{var(\hat{\beta}_I)} = ?$

Marginal effect of Income on FH scores when muslim = 1: $\hat{\beta}_I + \hat{\beta}_{M:I} \times 1 = ?$

Corresponding standard error:

$\sqrt{var(\hat{\beta}_I) + M^2 var(\hat{\beta}_{M:I}) + 2M cov(\hat{\beta}_I, \hat{\beta}_{M:I})} = ?$

Given this information we can fill in the table that we presented before:

	Muslim = 0	Muslim = 1
Marginal Effect of Income	?	?
Standard Error	?	?

From this information we can now answer the question that we started with this section. To calculate the 95% confidence interval, we follow the same procedure as always:

Upper 95% CI for marginal effect of income when muslim=0: ?

Lower 95% CI for marginal effect of income when muslim=0: ?

Upper 95% CI for marginal effect of income when muslim=1: ?

Lower 95% CI for marginal effect of income when muslim=1: ?

Tables are boring lets make a plot to illustrate this result, specifically what is known as a marginal effects plot.

Now lets do this same analysis to assess the signifiance of the marginal effect of muslim on Freedom House scores at the varying levels of income in the Fish dataset.

I find that marginal effect plots are useful to an extent but I prefer thinking of the interactive relationship in a holistic sense that allows us to directly relate the scenarios we are modeling to the dependent variable.

To do this we're going to bring back the plot that we started with and add some confidence intervals to it.