Handout for PS 15
Dummy Variables and Interaction Terms in Regressions

Dummy variables are simply variables that have been coded either 0 or 1 to indicate that an observation falls into a certain category. They are also sometimes called indicator variables. We use dummy variables in order to include nominal level variables in a regression analysis.

For instance, assume we are studying political stability in countries, and we want to consider whether a country is a dictatorship versus some other form of government. This variable has a nominal level of measurement – countries are grouped into unordered categories (dictatorship or other).

A dummy variable for dictatorship codes this variable as a 1 for "dictatorship," and 0 for "not dictatorship." Coding the variable in this way means that if we use it as an independent variable, a one-unit increase has meaning. In this case, a one-unit increase in this variable is the difference between dictatorships and other forms of government.

Suppose we included this dummy variable for dictatorship as an independent variable in a regression with deaths due to political violence as the dependent variable, and IMF loans (measures in millions of dollars) as another independent variable. Our results in SPSS are:

Coefficients(a)

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	4.28	1.24		3.449	.000
IMF Loans	3.9	1.7	.056	2.236	.025
Dictatorships	4.9	1.7	.072	2.876	.004

a Dependent Variable: Political Deaths

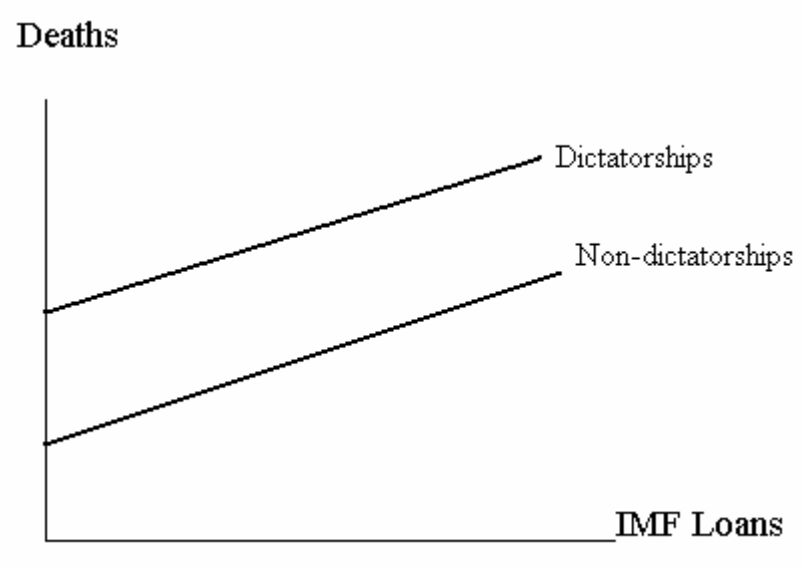
The two most important columns for interpretation are the first column of numbers (the coefficients, or "B" column) and the last column of numbers (the significance level, or "Sig." column). Remember in hypothesis testing we look for "sign and significance" – is the coefficient in the direction you expected, and is the significance less than 0.05? You might also want to look at the R^2, but we'll ignore that for now.

- The constant term tells us that if both of the independent variables were equal to zero, we would expect 4.28 deaths due to political instability in a country.
- The coefficient on IMF loans is positive and statistically significant, telling us we can accept the hypothesis that increasing IMF loans leads to more deaths due to

political instability. Specifically, we expect every extra $1 million in IMF loans to lead to 3.9 more deaths, all else equal.

- The coefficient on dictatorships is positive and statistically significant, telling us we can accept the hypothesis that dictatorships see more deaths due to political instability than non-dictatorships. Specifically, we expect dictatorships to have 4.9 more deaths than non-dictatorships, all else equal.

Graphically, the relationship would look something like this:



Now suppose we thought that IMF loans might have a different effect in dictatorships versus other types of countries. This is a conditional hypothesis, which we can test with an interaction term. We simply multiply the dictatorship dummy variable and the IMF variable to make a new variable called an interaction term. Just as the dummy variable on dictatorships can be regarded as an adjustment to the constant term in the regression for dictatorships, the interaction can be regarded as an adjustment to the slope coefficient on IMF loans for dictatorships.

Coefficients(a)

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	4.355	1.90		2.286	.019
IMF Loans	3.2	1.2	.091	2.605	.009
Dictatorships	0.30	0.26	.063	1.175	.241
Dictatorships X IMF Loans	8.4	2.6	9.9	3.169	.002

a Dependent Variable: Political Deaths

- The constant term tells us that if all three of the independent variables were equal to zero, we would expect 4.355 deaths due to political instability in a country.
- The coefficient on IMF loans is positive and statistically significant, telling us we can accept the hypothesis that increasing IMF loans leads to more deaths due to political instability. Specifically, we expect every extra $1 million in IMF loans to lead to 3.2 more deaths, all else equal.
- The coefficient on dictatorships is positive, but it is not statistically significant. This means we can't rule out the possibility that the coefficient on dictatorships is actually 0. We would reject any hypothesis that said that dictatorships see more deaths due to political instability than non-dictatorships.
- The coefficient on the interaction term is positive and statistically significant, telling us we can accept the hypothesis that IMF loans have a stronger positive relationship to political instability deaths in dictatorships as compared to non-dictatorships. Specifically, for every extra $1 million in IMF loans, we will see 8.4 more deaths in dictatorships over and above any effect we see in non-dictatorships. In this case, every $1 million in IMF loans leads to 8.4 + 3.2 = 11.6 more deaths in dictatorships.

Graphically, the relationship would look something like this: