POLS 707
Paul Johnson
Functions and Plotting Exercises

1. Plotting lines: Get some graph paper!

   (a) Plot this equation in the X-Y plane. Y= 3 + 2.2*X. If X is an independent variable and Y is the dependent variable, describe *in words* the relationship between X and Y.

   (b) Plot this equation, P= a + b*D, in the D-P plane for the following 3 sets of values of the parameters a and b. (Note, the D-P plane is the same as X-Y plane, just with differently labeled axes.)

	a	b
1:	0.0	2
2:	3	0.2
3:	6	-5

   (c) Write one paragraph describing the relationship between two individual-level variables, which we call D and P. You can say what D and P stand for. Suppose you studied the relationship between D and P in three countries (or states or counties or cities or contexts or time periods) and found the relationships that are plotted in your answer to 1b. What would you make of this? Write a paragraph which presents your story about P and D (tell me what the letters P and D refer to and how they are related) and interpret the results illustrated in .

   (d) Most of us can't think clearly in five or six dimensions. Most of us, however, can imagine a relationship between one independent variable and a single dependent variable, and sometimes we can throw in another independent variable and make a 3-D figure out of it.

   Sometimes relationships come along which make it possible to graph the relationship between three or more variables in a more familiar two dimensional graph. Consider the relationship

   $$Y = a + b \cdot X + c \cdot Z$$

   where a, b, and c are parameters (constants given for purposes of discussion), Y and Y are real-valued "continuous" variables, and Z is a variable which takes on only integer values 0, and 1. (Z is often called a "dummy variable" in the social science literature.) For each of the values of Z, you can plot, in the $X-Y$ plane, the relationship between Y and X. Make sure to label which line goes with which value of Z. Oh, by the way, you can choose any values for a, b, and c that you like (as long as $a > 0$, $b \geq 0$, and $c \neq 0$. Make sure you report the values in your answer!).

(e) Here's a slightly more interesting example of the same. Suppose that the equation which describes the relationship between three variables is:

$$Y = a + b \cdot X + c \cdot Z + d \cdot X \cdot Z$$

Now, repeat the exercise in 1d. (Don't forget to choose a value for d!)

(f) Suppose that, in the context of question 1e, the variable Z is a dummy variable representing some kind of "binary" (that means two-valued) variable. For example, Z might be an indicator of ethnicity and $Z = 0$ means a person is caucasian, while $Z = 1$ means a person is not. Suppose that the relationship graphed in 1e describes data about individuals collected in a sample. Give substantively meaningful names for the variables X, Y and Z and create a "story" which explains the observed relationship between X and Y.

Then analyze the figure you created in A5. What is the importance of the parameters "c" and "d"? Which might be called a "slope-shifter" and which might be called an "intercept-shifter"?

(g) Here's a very slightly different example of the problem in 1e. Suppose X represents time. Instead of indexing the relationship by a third variable such as ethnicity, the dummy variable Z indicates the beginning of a "policy intervention." For example, $Z = 0$ for $X < 10$ and $Z = 1$ for $X \geq 10$. The dummy variable Z indicates that there has been a policy change at time 10, then, and this policy change remains in effect after 10 for an indefinitely long time.

Name the variable Y (it should be an indicator of policy impact). Use any nonzero coefficients to graph this equation:

$$Y = a + b \cdot X + c \cdot Z + d \cdot X \cdot Z$$

The graph will look better if you have only one line for the region X $<$ 10 and one in the region X $>$ 10. See what I mean? After you have finished your graph, discuss the meaning of the parameters c and d. What difference is obsereved in Y as X is increased from 9.9 to 10? (Hint: if the equation were written $Y = a + bX + cZ + d \cdot (X - 10) \cdot Z$, how would it's graph differ?).

2. Logs and exponentials are not familiar to many students. You might not have seen these since high school. So try and do these exercises. If I say log(), it is usually a mistake, and I really mean the natural log, the one often denoted ln(). But if you want, you can take the log with any base you want. Just say what base you assume.

   (a) Use a calculator to calculate a few values of the y=log(5*x) and then plot it.

(b) Use a calculator to figure $y = \exp(3 * x)$ for a few values of x and then plot it.

(c) Use the laws of logarithms to figure out these problems

i. log(4*x)

ii. log(4*x*y)

iii. log(4*x*y*z)

(d) Remember that $\exp(x)$ is the same thing as e^x. So figure out how to simplify these:

i. exp(4+x)

ii. exp(4+log(x))

iii. $\exp(4 + x^2 + log(x))$

iv. $\exp([5 + 3 * x + 2] * log(u))$

3. Polynomials.

You need to work yourself back into the high school state of mind so that you don't freak out if somebody says

$$y = a \cdot x^2 + b \cdot x + c \tag{1}$$

Because this equation includes x^2, it is not linear. Rather, it is a quadratic equation, sometimes also called a polynomial. It is not entirely clear to me how you avoid freaking out when you see that, but maybe the best way is to plot it. I can show you how to plot that in R, or I bet you can plot it in many other programs. Pick any domain you want, make it easy, such as [-100,+100]. Pick a few example sets of coefficients for (a, b, c), I don't care what.

After you plot a few of those lines, here are some cool facts about that equation. Try to verify them.

(a) If $a > 0$, the shape of the graph is a "U" (or valley). If $a < 0$, it is a "hill" shape.

(b) The "top" of the hill (or the bottom of the valley) is found when $x = \frac{-b}{2a}$.

(c) The curve "goes through" the horizontal (x) axis at these values:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In mathematical parlance, these values of x are the "roots" of the quadratic equation in 1

(d) As the value of a gets closer and closer to 0, the curve gradually approaches a straight line.