Lab Objective

To gain more experience with histograms, box plots, correlations, and line fitting.

Lab Procedures

Long-term interest rates drive much of the economic activity in the U.S. When interest rates are low, people and establishments are more likely to borrow money for purchasing homes or growing their businesses. When interest rates are high, people and establishments are less likely to do so. In this lab, we’ll work with economic data from the U.S. from 1990 to 2010 to explore relationships involving interest rates.

Open the data file USeconstat2.dta. The data are from the Organization for Economic Cooperation and Development.

Data Analysis Caveat: One can look at many relationships with economic (or any) data. It is tempting to assign causal explanations to those relationships. This is risky. Just because there is (or is not) a relationship between two variables, it does not mean there is (or is not) a causal relationship between those variables. There could be many other factors that affect both variables, and these could explain what is seen in the graphs.

Questions:

1. Describe the distribution of long-term interest rates. That is, say where most values are, note any outliers, and say whether the distribution is tightly packed around its mean or is spread out. Also, report the mean and standard deviation.

2. Long-term interest rates right now are around 2.5%. Describe how 2.5% compares to the historical record of long-term interest rates from 1990 - 2010.

3. How did long-term interest rates change annually between 1990 - 2010? Were they (i) generally going up; (ii) generally steady; (iii) generally going down; or (iv) all over the place? Just write a short phrase as your answer.
4. Using the data, describe the relationship between long-term interest rates and household net savings rates. Include in your descriptions a one-number summary of the strength of the association between the two variables.

5. Using the data, describe the relationship between long-term interest rates and unemployment rates. Include in your descriptions a one-number summary of the strength of the association between the two variables.

6. Using the data, describe the relationship between long-term interest rates and gross domestic product. Include in your descriptions a one-number summary of the strength of the association between the two variables.

7. Of the following two variables, which one has the weaker linear association with long-term interest rates: (i) wage rate; or (ii) government net lending rate? Explain your choice in one sentence.

8. What is the correlation between government net borrowing and long-term interest rates? Note that borrowing is the opposite of lending, so that net borrowing equals negative one times net lending.

9. Suppose you had a model that gave reasonable predictions about long-term interest rates in the next year. (This is unrealistic as interest rates are notoriously hard to predict. However, there are many statisticians and economists trying to do so.) Suppose you predict that interest rates next year will be 5.0%. Predict gross domestic product for next year using a regression of interest rates and time period on gdp.

There are multiple ways to use Stata to get an answer for the predicted value:

(a) **matrix and scalar**

This method is just having Stata do the math for you. Stata saves the results from the regression in a matrix called `e(b)` (type `ereturn list` to see a list of all the stored results). To access this, you need to save it into another matrix: `matrix beta = e(b)`, where `beta` is the name of the new matrix. You can see what this matrix looks like with `matrix list beta` and you can access individual elements of `beta` using square brackets and the row/column numbers. So to display the first element, you would type `display beta[1,1]`

Now you can display the predicted value using the necessary formula or you can save it in a scalar.

(b) **margins**

The `margins` command is a very powerful tool in Stata. It allows you to solve almost any estimation command at specific values, whether they be at statistical values like the mean or the median, or at specific values like in our case. The general syntax is: `margins, at( varname1=# [varname2=#])`
where the option `at()` tells Stata what values to solve at, and then you give a list of variables and their values to solve it. If you leave out a variable it will include the mean value by default.

(c) `predict` in current dataset

`predict newvar` will create a `newvar` that contains the regression prediction at each observation.

To `predict` values that are not in the sample, you can create new data after you run the `regression`. Creating new data is difficult to do in the command window, so your best bet is to do what we did in the first problem set and open up the data editor (Data -> Data Editor -> Data Editor (Edit)). Then you can just highlight a cell below all the data and type in the year and interest rate in the cells. With this new data, you can now use `predict` and use the value listed.

(d) `predict` using a new dataset

If you have multiple values you want to predict at, you could actually place all those values in a new dataset with the same variable names. Then when you `use` this new data, you can `predict` and the regression results from the original data will still be there.

10. Create a scatter plot of these variables. Does it suggest any clearly non-linear relationships in the data? Justify your answer in at most two sentences.

11. If interest rates were 1%, could you use the regression equation to predict the corresponding gross domestic product (market prices)? If you think so, write down the predicted value of GDP. If you think not, explain why not in at most one sentence. Only write one answer; writing both answers gets no credit.