Part 1. Statistical analysis. Statistical methods are like those of fine cuisine. You start simple and small and keep adding. Initial steps can be read from “cookbooks,” while later methods require creative adaptation. Also, as with cooking, the rudimentary tools are essential for survival. Statistics as it pertains to economics is called econometrics. Indeed, this is as good a definition of econometrics as I can find. The first step in econometrics is the regression equation

\[ y_t = a_0 + a_1 x_t + \varepsilon_t \]

The variable on the left \((y_t)\) is the dependent variable and the variable on the right hand side \((x_t)\) is the independent variable. Actually, there are two variables on the right, \(x_t\), and the constant 1 (which gets multiplied by \(a_0\)). The final term on the right \((\varepsilon_t)\) is called the error term. An ordinary regression (called OLS for ordinary least squares) finds the best fit line through the data. That is, if \(x_t\) is put on the horizontal axis and \(y_t\) on the vertical and all the points are plotted, a line is fitted through these data points. This line is defined by two coefficients, an estimate of \(a_0\) and an estimate of \(a_1\). (All lines are defined by two numbers.) How we find these two numbers is a question for another day. It involves a page of algebra (but no more than a page).

The interpretation of the coefficients is the heart and soul of econometrics. We ask what is the estimate of \(a_1\) and is it significant? We interpret this coefficient as the effect of \(x_t\) on \(y_t\), or at least we want to interpret it this way. Thus, if \(a_1\) is equal to 3, say, then each one unit increase in \(x_t\) “causes” a 3-unit increase in \(y_t\). The first thing we want to know about the estimate of the coefficient is how accurate the estimate is, so each coefficient has what is known as a “standard error.” The standard error is computed using a simple formula that can be derived (with another page of algebra) from a couple of assumptions about a “normal distribution.” The larger the standard error, the less accurate we believe the estimate to be. The ratio of the coefficient estimate to the standard error is called the t-value. (Actually, this is only true if we are comparing the coefficient to zero. If we are comparing the coefficient estimate to some other number, like 5, which we expect it to be for some reason, then the t-value is the estimate minus this other number, with the difference divided by the standard error.) A t-value that is very large or very low is “significantly” different from zero. When the t-value is significant, we are inclined to say that “there is a significant effect of \(x_t\) on \(y_t\).” Of course, there are a thousand reasons for why we might be wrong, and sorting all this out is the subject of econometrics.

Multiple regression, omitted variables and “corrected \(R^2\). The multiple regression equation with three variables is of the form

\[ y_t = a_0 + a_1 x_{1,t} + a_2 x_{2,t} + a_3 x_{3,t} + \varepsilon_t \]
We might “regress” $y_t$ on $x_{1,t}$ and find that $a_1$ is significant, but then regress $y_t$ on $x_{1,t}$, $x_{2,t}$ and $x_{3,t}$, and find that $a_1$ is no longer significant. This means that the original regression gave us the erroneous conclusion that of significance. The problem is that the original regression suffered from the omitted variables problem. This problem makes us want to keep adding variables. One way to prevent us from adding garbage is to check the corrected $R^2$. The usual $R^2$ always goes up when a variable is added. The corrected $R^2$ goes down when the last variable adds no explanatory power.

1. Use the data set USDATA.prn. Using RATS compute the growth rates of m1, m2 gdp.
   A. Regress the growth rate of gdp on gdp growth lagged one period and lagged two periods. Are these variables significant?
   B. Regress the growth rate of gdp on gdp growth, with one and two-period lags. Also include current and past tbill rates and the money supply growth rates. Does the corrected $R^2$ rise or fall when these variables are added? Is lagged gdp growth still significant?
   C. Create variables that are logs of the original values. Run a bunch of regressions to answer this question: Does current output increase if last period’s investment is higher? (Intuitively, should it?)
   D. Is lagged non-durable consumption or lagged investment a better predictor of current gdp? Give me the regression results.

2. A growth model has the production function $Y_t = F(K_t, N_t, A_t) = A_tK_t^\alpha N_t^{1-\alpha}$, the marginal product of capital $R_t = F'_K(t) = \alpha A_t(K_t / N_t)^{\alpha - 1}$ and wage $w_t = F'_N(t) = (1-\alpha)A_t(K_t / N_t)^{\alpha}$. Investment accumulates according to $K_{t+1} = K_t(1-\delta) + I_t$. Output uses are $Y_t = C_t + I_t + G_t + NX_t$. Labor input is $N_t = 208$. Initial capital is $K_0 = 15,204.89$. The parameters are depreciation: $\delta = 0.08$, productivity, $A_0 = 12$, and capital share $\alpha = 0.25$. Also, the current value for $C_t + G_t + NX_t$ is 5777.86. Compute the values for
   A. Output
   B. Investment
   C. Next period’s capital stock.
   D. The growth rate of capital between periods 0 and 1.
   E. The wage and the rental rates.

3. Read the Volker section of Rukstad.
   A. Was Volker more interested in money growth or credit expansion or can this distinction not be made?
   B. Was Volker’s policy a success? (You must give you criteria for success and then say yes or no.)

4. Does a country need to have backing for its currency? (This should take you a page.)