Problem Set 1

Graduate Macro II, Spring 2013
The University of Notre Dame
Professor Sims

Instructions: You may consult with other members of the class, but please make sure to turn in your own work. Where applicable, please print out figures and codes. This problem set is in class on Tuesday, January 29.

(1) Moments of ARMA processes: This problem asks you to compute some first and second moments for some popular ARMA processes.

(a) Suppose that a random variable obeys the following stochastic process, assuming that $0 < \rho < 1$:

$$x_t = a + \rho x_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma^2)$$

(i) What is the unconditional mean of x_t, $E(x_t)$?
(ii) What is the unconditional variance of x_t, $var(x_t) = E((x_t - E(x_t)^2)$?
(iii) What is the unconditional covariance between x_t and x_{t-1}, $cov(x_t, x_{t-1}) = E((x_t - E(x_t))(x_{t-1} - E(x_{t-1}))$?

(b) Suppose that a random variable obeys the following stochastic process, assuming that $0 < \theta < 1$:

$$x_t = a + \varepsilon_t + \theta\varepsilon_{t-1} \qquad \varepsilon_t \sim N(0, \sigma^2)$$

(i) What is the unconditional mean of x_t, $E(x_t)$?
(ii) What is the unconditional variance of x_t, $var(x_t) = E((x_t - E(x_t)^2)$?
(iii) What is the unconditional covariance between x_t and x_{t-1}, $cov(x_t, x_{t-1}) = E((x_t - E(x_t))(x_{t-1} - E(x_{t-1}))$?

(2) Using OLS to Estimate a Persistent AR(1) Process: Suppose that some random variable obeys the following stochastic process:

$$x_t = \rho x_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0, 1)$$

This problem asks you to conduct a Monte Carlo exercise to investigate both the "small" and "large" sample properties of the OLS estimator for ρ. Generate a vector of observations of x_t of length T (i.e. $t = 1, ..., T$). Let $T = 200$ for now. Do this N times (N is the number of Monte Carlo "runs"). Let $N = 10000$. In other words, you should have N different observations of x_t, where x_t is a vector with T elements.

(a) Discard the first 100 observations from each data set (this limits the influence of starting values). On each data set, use OLS to come up with an estimate, $\widehat{\rho}$, using the value $\rho = 0.5$ to generate the data.[1] Be sure and include a constant in the regression, even though there is no constant in the DGP (DGP stands for data generating process). Compute the average value of $\widehat{\rho}$ across the Monte Carlo runs, and show a plot of the histogram of $\widehat{\rho}$. Is the estimate biased? In which direction?

[1] In vector notation, if you are estimating: $Y = X\beta + \varepsilon$, where Y is $T \times 1$, X is $T \times q$ (i.e. there are T observations and q regressors, the first of which is a constant (a vector of ones)), and β is $q \times 1$, the OLS estimator of β is $\widehat{\beta} = (X'X)^{-1}X'Y$.

(b) Repeat the exercise in (a) for the following different values of ρ: 0.67, 0.8, 0.9, 0.95, 1. What can you say about the size of the bias in the OLS estimate and the size of the true autoregressive parameter?

(c) Now play around with the size of the sample – i.e. let T take on values of 300, 500, 1000, and 2000 (still throw away the first 100 values). Repeat the exercises from above for the different values of ρ. What happens to the bias of the OLS estimates as T gets larger? What happens to the distribution of OLS estimates as T gets larger? Does the distribution of estimates converge to the true value as T gets large (i.e., loosely speaking, is OLS consistent)? Does the speed of convergence appear to differ at all depending on the true value of ρ?

(3) Estimating the moving average coefficients by simulating the impulse response function: In this problem you will show, via simulation, how to approximate a moving average process by estimating an autoregression and simulating the impulse response function. Suppose that we have the following data generating process:

$$x_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} \qquad \varepsilon_t \sim N(0,1)$$

(a) Re-write this process using lag operator notation.

(b) Write down the generic definition of an impulse response function.

(c) Construct the true impulse response function to a unit impulse (i.e. $\varepsilon_t = 1$) using the following values: $\theta_1 = 0.9$, $\theta_2 = 0.7$, and $\theta_3 = 0.4$. Write out the impulse response function for $h = 10$ periods.

(d) Conduct a Monte Carlo exercise in which you simulate N sets of observations of x_t of length T (let $N = 1000$ and $T = 200$) according to the above MA(3) process using the parameter values given in (c). After discarding the first 100 observations from each data set, estimate an AR(p) on each data set, and use the AR slope estimates to construct an estimate of the impulse response function for $h = 10$ periods. Average the impulse responses across the N Monte Carlo runs. Compare the average estimated IRFs to the true IRF found in part (c). Repeat this for the following different vales of p: 1, 2, 4, and 10. Comment on the relationship between the accuracy of the simulated IRF and the number of lags in the estimated autoregression.

(4) A Helicopter Tour of Unit Roots:[2] Consider the following random walk process:

$$z_t = \rho z_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

Set $\rho = 1$ and $\sigma_\varepsilon = 0.01$. Generate $N = 10{,}000$ different data sets, each with $T = 200$ observations, of $\varepsilon_t,\ t = 1, \ldots, T$, using Matlab's "randn" command. Use these to construct N different data sets with T observations of the z_t process defined above, assuming an initial value of $z_{-1} = 0$ in all N different simulations.

(a) On each of the N different simulated data sets, estimate a regression of z_t on z_{t-1} (and a constant, always use a constant) via OLS. For each of the N different data sets, save your estimate of ρ (the coefficient on z_{t-1}) and the OLS standard error of ρ. Plot a histogram of the estimates of ρ (you can use the "hist" command in Matlab). Report (i) the mean estimate of the $\widehat{\rho}$, (ii) the standard deviation of the N different estimates of $\widehat{\rho}$, and (iii) the mean estimate of the OLS standard error of $\widehat{\rho}$ across the N simulations. How does the mean of the estimates of $\widehat{\rho}$ compare with the true value, 1? How does the mean estimate of the OLS standard errors compare the actual standard deviation across the simulations?

[2]The title of this problem is a play on Sims and Uhlig (1991), "Understanding Unit Rooters: A Helicopter Tour".

(b) Repeat this exercise, but this time generate data with $T = 2000$ observations. Lower N to $N = 1000$ to save on computing time. Collect your estimates of the $\widehat{\rho}$. What is the mean? Plot out the histogram and compare it to the histogram you showed in part (a).

(c) Now go back to the "finite" sample version (i.e. $N = 10,000$ and $T = 200$). Instead of estimating a first order autoregression in levels, estimate a first order autoregression in first differences on the same z_t series you generated. In other words, estimate:

$$\Delta z_t = \gamma \Delta z_{t-1} + u_t$$

Where $\Delta z_t \equiv z_t - z_{t-1}$. What is the "true" value of γ? Estimate this autoregression N different times on the data sets you already generated (note you lose one observation from the first differencing). Report (i) the mean estimate of the $\widehat{\gamma}$, (ii) the standard deviation of the N different estimates of $\widehat{\gamma}$, (iii) the mean estimate of the OLS standard error of $\widehat{\gamma}$ across the N simulations. Compare the mean estimate with the true value of the parameter as well as the mean estimate of the OLS standard error with the actual standard deviation of the estimate across simulations. How do your answers compare with those in (a)?

(d) For each of the N simulations in (c), compute the standard t statistic of the null hypothesis that $\gamma = 0$. What fraction of the time do you get a t statistic greater than 1.965 in absolute value? Does this make sense or not?

Now generate a new series, call it y_t. It follows:

$$y_t = y_{t-1} + \nu_t, \quad \nu \sim N(0, \sigma_u^2)$$

The error terms here are drawn from a separate, independent distribution from the error terms you drew above in constructing the z_t series. Set $\sigma_u = 0.01$, and generate $N = 10,000$ different data sets of y_t with $T = 200$ observations each, using $y_{-1} = 0$ as initial conditions in all simulations.

(e) Consider estimating the following regression:

$$z_t = \alpha + \beta y_t + u_t$$

What is the true value of β? Should y_t be able to explain any of the variation in z_t? On each of the N data sets, estimate this regression. Plot the histogram of the estimates of $\widehat{\beta}$, and report the mean estimated coefficient. For each regression, calculate the t statistic for the null hypothesis that $\widehat{\beta} = 0$. What fraction of the time is the absolute value of the t statistic greater than 1.965?

(f) Consider estimating the following regression:

$$z_t = \alpha + \beta_1 z_{t-1} + \beta_2 y_t + u_t$$

What is the "true" value on β_2? Estimate this regression N times, and save your estimates of β_2. Plot the histogram of estimates and report the mean value. Visually, how does the histogram compare to the one you found in (e)? Also, calculate the t statistic for the null hypothesis that $\widehat{\beta_2} = 0$. What fraction of the time is the absolute value of the t statistic greater than 1.965? How does this compare to (e)?

Now consider a new process, given by:

$$w_t = z_t + e_t$$
$$e_t = 0.5 e_{t-1} + v_t, \;\; v_t \sim N(0, \sigma_v^2)$$

Assume $\sigma_v = 0.01$. Generate $N = 10,000$ different data sets with $T = 200$ observations each for the w_t process, using your existing process for z_t and the newly generated process for e_t.

(g) Estimate the following regression on each simulated data set:

$$w_t = \alpha + \beta z_t + u_t$$

What should the true β be? Collect your estimates of $\widehat{\beta}$ across the N different simulations. Plot a histogram of the $\widehat{\beta}$ and report the mean estimate. Calculate the t statistic for the test of the null hypothesis that $\widehat{\beta}$ equals the true value of β, and report what fraction of the time you reject the null. Compare this rejection rate to the one you reported in (d), (e), and (f).