

Problem Set 1

Graduate Macro II, Spring 2011
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 due by 5:00 pm on Friday, January 28.

(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 $0 < \rho < 1$:

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

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

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

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

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

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

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

This problem asks you to conduct a Monte Carlo exercise to investigate both the “small” and “large” sample properties of the OLS estimator for $\rho$. 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, $\hat{\rho}$, 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 \(\hat{\rho}\) across the Monte Carlo runs, and show a plot of the histogram of \(\hat{\rho}\). Is the
estimate biased? In which direction?

(b) Repeat the exercise in (a) for the following different values of \(\rho\): 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 \(\rho\). 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 \(\rho\)?

(3) Estimating the moving average coefficients by simulating the impulse re-
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:

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

(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 values 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) Constructing the Dickey Fuller distribution: This problem asks you to construct,
via simulation, the distribution of OLS estimates of an autoregressive parameter in the
presence of a unit root. In particular, suppose that data are generated from the following stochastic process:

\[ x_t = x_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, 1) \]

Suppose, as a researcher, you hypothesize that the series has a unit root, but are not sure and would like to construct a statistical test for the presence of a unit root. Suppose you want to estimate the following regression:

\[ x_t = \rho x_{t-1} + e_t \]

This regression can equivalently be written by first differencing the above:

\[ \Delta x_t = \gamma x_{t-1} + e_t \]

Where \( \Delta x_t = x_t - x_{t-1} \) and \( \gamma = \rho - 1 \). You would like to test the null hypothesis that \( \gamma = 0 \) (equivalently that \( \rho = 1 \)).

Generate \( N \) sets of data with \( T \) observations each (set \( N = 10000 \) and \( T = 200 \)). Use OLS to estimate \( \hat{\gamma} \) for each Monte Carlo run, discarding the first 100 out of the \( T \) observations in each series. Be sure to include a constant in your regression, even though there is no constant in the DGP. For each Monte Carlo run, construct the standard \( t \) statistic for the null hypothesis that \( \gamma = 0 \). The \( t \) statistic can be constructed as follows:

\[ t = \frac{\hat{\gamma}}{se(\hat{\gamma})} \]

In words, the \( t \) statistic is the ratio of the OLS slope estimate to its standard error. The standard linear regression model in matrix notation is: \( Y_t = X_t \beta + e_t \), where the matrix \( X_t \) includes a constant. The variance-covariance matrix of the vector of OLS parameter estimates is given by:

\[
\text{var}(\hat{\beta}) = \frac{e_t'e_t}{q-k} (X'_tX_t)^{-1}
\]

\[ e_t = Y_t - X_t\beta \]

Where \( q \) is the number of observations and \( k \) is the number of estimated parameters. The standard error of \( \hat{\gamma} \) is the square root of the (2,2) element of the variance-covariance matrix formula given above.

What is the mean of the \( t \) statistics across the Monte Carlo runs? Plot out the histogram of \( t \) statistics. What are the 10th, 5th, and 1st percentiles of the distribution of \( t \) statistics? Put differently, if the null hypothesis is \( \gamma = 0 \), what must the standard \( t \) statistic be in order to reject this hypothesis in favor of the alternative that \( \gamma < 0 \) (equivalently that \( \rho < 1 \)) at these different significance levels? Are these critical values different from a “usual” \( t \) test in a cross-sectional study?