Stochastic Processes[1]

These notes provide some brief theoretical background that will be useful throughout the semester. We are not yet concerned with econometrically estimating time series. Instead we want to think about the theoretical data generating process that we will later try to estimate. For now, we will look only at univariate processes. Because our focus is on random process,we will refer to these as *stochastic processes.*

Autoregressive 1 Processes

We now begin to describe several common types of time series. We begin with an AR(1) (for "autoregressive") process. An AR(1) time series takes the following form:

$$x_t = \delta + \alpha x_{t-1} + u_t \tag{1}$$

This is called an AR(1) process because the variable, x_t, depends on exactly one lag of itself. δ is a constant. u_t is a random error term. We assume that it is a mean-zero *white noise* process. This entails that, on average, it equals zero and that u_t and its lags are uncorrelated. Knowing u_{t-1} tells us nothing about u_t.

Examining (1), we see that $x_t$has a fixed point which it would stay at if it started there and there is no randomness. We can solve for this by setting all us equal to their mean (zero) and dropping the time subscript:

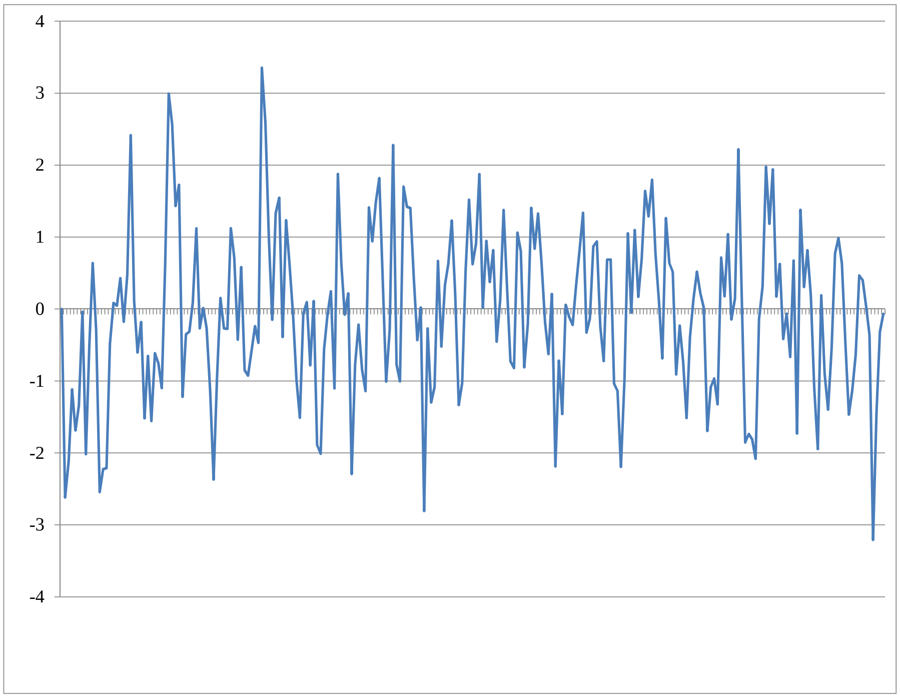
$$x = \frac{\delta}{1-\alpha} \tag{2}$$

This fixed point may or may not be the mean. We will examine this point in greater detail soon. It is better to think of it as akin to the concept of a steady state which you have seen in your macro courses.

The value of α tells us a great deal about this process. We consider five cases:

#1 $\alpha \in (0, 1)$. In this case, x_t is positively autocorrelated in that a high value of x_t likely leads to a high value of x_{t+1}. Below, I plot two examples for $\alpha = 0.5$. The first has no constant ($\delta = 0$). The second sets $\delta = 0.2$. In both cases, I assume that u_t is normally distributed with a mean equal to zero and a standard deviation (and variance) equal to one.

Table 1: $\alpha = 0.5$, $\delta = 0$



In both cases, x_t always has a tendency to revert back to its mean. This will be important when we discuss the concept of stationarity.

Because the process is mean reverting, (2) is the true mean. This does not mean that a given sample mean is likely to equal the process's true mean. For this case, the time series is also ergodic. Thus as $T \rightarrow \infty$, the sample mean will converge to $\frac{\delta}{1-\alpha}$.

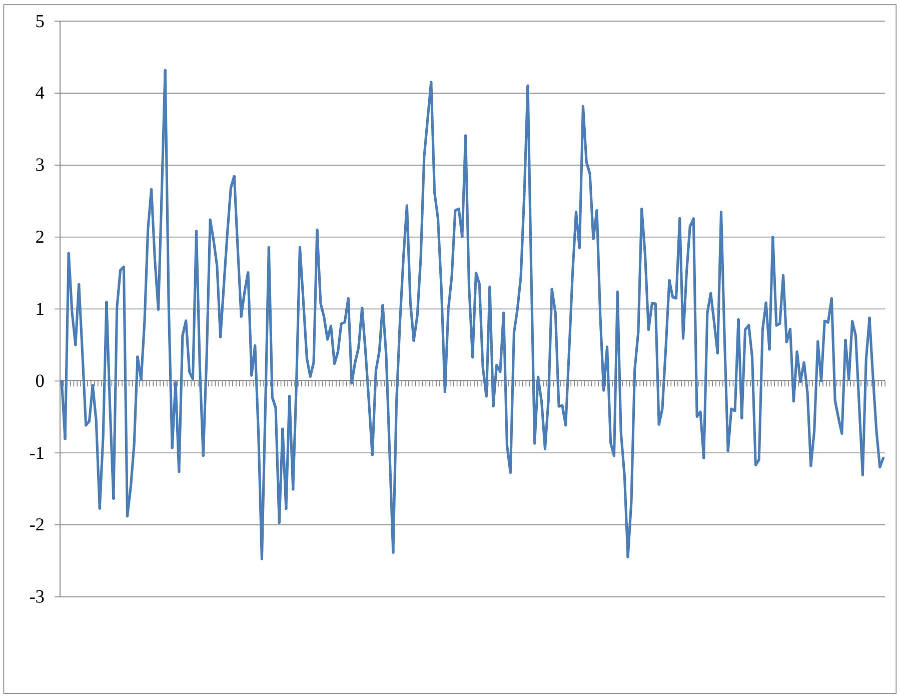
Increasing δ has the effect of increasing the time series's true mean. This is evident when we compare the prior two graphs.

For an AR(1) process, α is a measure of the series's persistence. Larger values imply that shocks have longer lasting effects. Consider the case where $\alpha = 0.9$.

Note that the time series now experiences extended periods where it is either above or below its mean (zero). This is because the value in period t now has a major effect on the value on period $t + 1$.

[1]These are undergraduate lecture notes. They do not represent academic work. Expect typos, sloppy formatting, and occasional (possibly stupefying) errors.

Table 2: $\alpha = 0.5$, $\delta = 0.2$



#2 $\alpha \in (1, \infty)$. In this case, x_t is positively autocorrelated as in #1. But it now behaves explosively, there is no tendency to revert back to a mean or trend. Figure 5 plots this case for $\alpha = 1.2$ and $\delta = 0$.

Because the process is explosive, (2) is not generally the true mean.

#3 $\alpha \in (-1, 0)$. Setting $\delta = 0$, and $\alpha = -0.5$ shows that this process is again mean-reverting. But now, a positive value of x_t implies that a negative value of x_{t+1} is likely and vice-versa. The system is *oscillatory*.

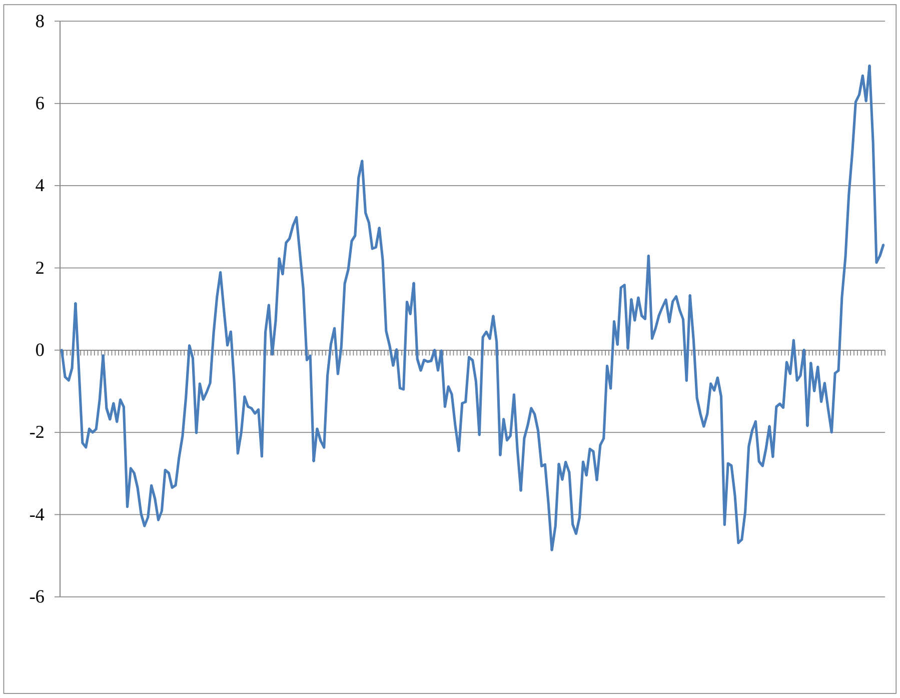
#4 $\alpha \in (-\infty, -1)$. Here, the system is both explosive (as in #1) and oscillatory (as in #4).

#5 $\alpha = 1, -1$. This is known as a *unit root*. We will discuss these later in the course.

We now consider an alternate way of wiring an AR(1) process. If (1) is true in period t, then we can re-date this equation to make it hold for period $t - 1$:

$$x_{t-1} = \delta + \alpha x_{t-2} + u_{t-1} \tag{3}$$

Table 3: $\alpha = 0.9$, $\delta = 0$



This is a common trick. As long as the definition of t is arbitrary, then we can always re-date a time series either backward or forward as long as we re-date all of the time subscripts by the same number of periods. We can then insert (3) into (1):

$$x_t = \alpha^2 x_{t-2} + u_t + \alpha u_{t-1} + (1 + \alpha)\delta \tag{4}$$

We can then re-date (3) backwards one more period:

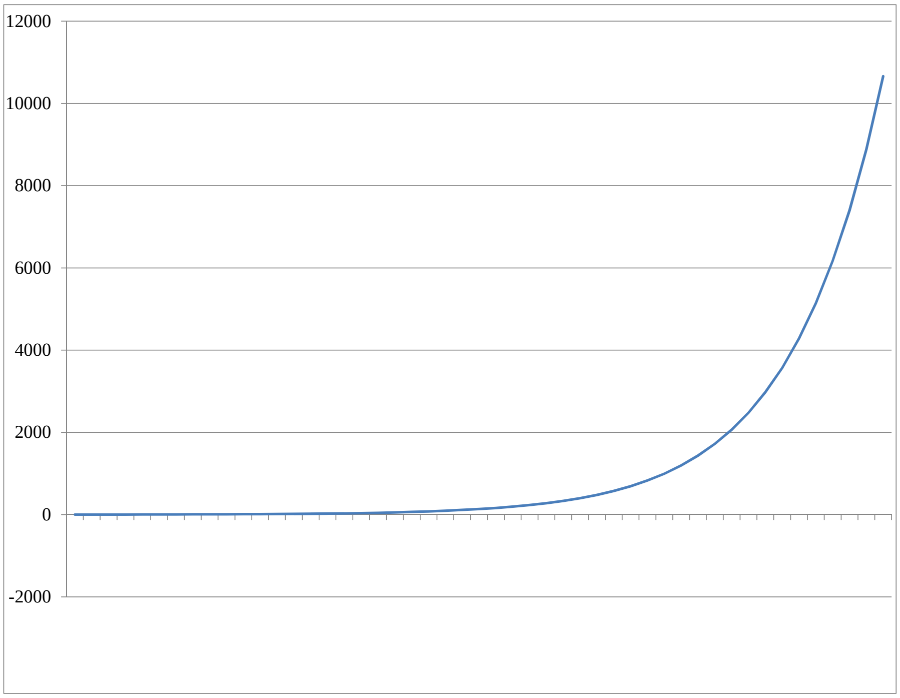
$$x_{t-2} = \delta + \alpha x_{t-3} + u_{t-2} + \delta \tag{5}$$

And then we can insert (5) into (4) to eliminate x_{t-2}:

$$x_t = \alpha^3 x_{t-3} + u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + (1 + \alpha + \alpha^2)\delta \tag{6}$$

The goal is to continue until a pattern emerges. Note that each time we iterate backward, we move the x term back one period and add one more error term. If we do this an infinite number of times and if $|\alpha| < 1$:

Table 4: $\alpha = 0.9$, $\delta = 0$



$$x_t = u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \alpha^3 u_{t-3} + \alpha^4 u_{t-4}... + (1 + \alpha + \alpha^2 + \alpha^3...)\delta \tag{7}$$

This only works because as $n \to \infty$, $\alpha^n \to 0$. That would not be the case if $|\alpha| \geq 1$.

We can now calculate the time series's moments. To again calculate the mean a slightly different way, we just insert the mean of u_t, which is zero by assumption, into (7):

$$E[x_t] = \frac{\delta}{1 - \alpha} \tag{8}$$

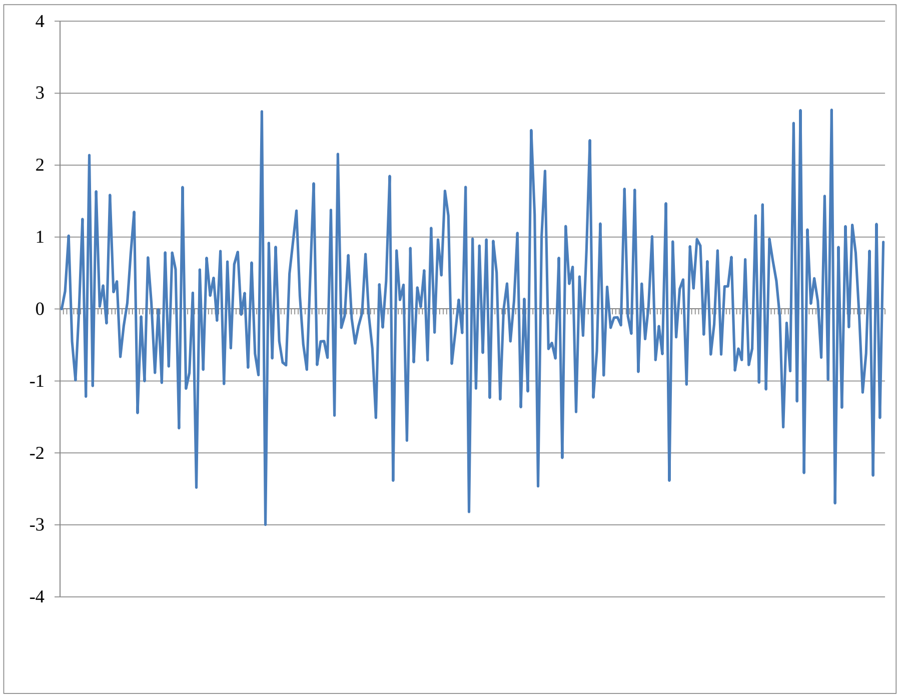
To derive (8), I am using the formula for an infinite geometric series to eliminate $(1 + \alpha + \alpha^2 + \alpha^3...)\delta$. Crucially, the mean does not depend on t. It is thus mean-stationary.

By assumption, $E[u_t, u_s] = 0$, that is the error terms are uncorrelated. We can thus use the statistical result that the variance of the sum of independent variables is simply the sum of the variances: Thus:

$$v[x_t] = (1 + \alpha^2 + \alpha^4...)\sigma_u^2 = \frac{\sigma_u^2}{1 - \alpha^2} \tag{9}$$

where I again use the formula for an infinite geometric series. Once again, this moment is

Table 5: $\alpha = -0.5$, $\delta = 0$



independent of t. The process is thus at least weakly stationary.[2] It is therefore possibly useful to include in an econometric specification.

Equation (7) is known as a $MA(\infty)$ process. "MA" stands for moving-average. An MA process consists only of independent random error terms. For example, an MA(2) process is:

$$x_t = \delta + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} \tag{10}$$

Equation (10) is MA(2) because it depends on two lags. AN MA(0) process is as *white noise.*

An AR(1) process can always be written as a $MA(\infty)$ process. In general a $MA(\infty)$ process cannot be written as an AR(1) process. If it can, the $MA(\infty)$ process is said to be *invertable.*

Any finite-order MA process is weakly stationary. To see this, note that the mean is simply δ, independent of time. The variance will also be a function only of σ_u^2 and the θs.

We now consider the AR(1) process where $|\alpha| \geq 1$. Note that (7) no longer applies. So

[2] It can be shown that it is also stationary for all moments.

instead iterate back t periods:

$$x_t = u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \alpha^3 u_{t-3} + ...\alpha^t u_o + (1 + \alpha + \alpha^2 + ...\alpha^t)\delta + \alpha^{t-1} x_0 \tag{11}$$

The mean of x_t now depends on the term $\alpha^{t-1} x_0$. It is time dependent and the process is non-stationary. We cannot ordinarily use this time series in our econometric specifications. Doing so would renders the results uninteresting.

ARMA processes:

An ARMA(p,q) process combines AR and MA components. For example, an ARMA (1,2) process is:

$$x_t = \delta + \alpha x_{t-1} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} \tag{12}$$

An ARMA(3,0) process includes no MA components. We would thus typically refer to this as an AR(3) process:

$$x_t = \delta + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \alpha_3 x_{t-3} + u_t \tag{13}$$

One mistake I have seen students make is to assume, for example, that an AR(3) includes only the third lag. This is not generally true (it is a special case, however). The p in an AR(p) process refers only to the oldest lag term in the process.

There is no simple formula for the stationarity of an ARMA(p,q) process.

Theoretical vs. Applied Time Series)

The focus of this class is almost entirely on applied time series: how we can use the methods of the class to better answer interesting economic questions. But it is worth briefly discussing the distinction between this and theoretical time series, which includes finding new estimators.

Suppose we have an AR(1) process. We could thus decide to regress x_t on its lag, x_{t-1}. Although we lose one observation (the earliest one which has no lag), there is no obvious problem. We may reasonably decide to use OLS.

Now suppose that I propose an alternate estimator, Fabulous Most Circles. Which is better? We may be able to analytically (using just math without simulations) show that OLS is better in the ways we usually care about: consistency, efficiency, etc. But suppose that we cannot.

Another approach is to choose values of δ, α, and the variance of u_t from (1) and simulate 1 million draws using one of many available software packages. We can then apply each estimator to the simulated data and see which provides better estimates. We can repeat this process 1 million times. Maybe one estimator always does better. Or maybe it depends on the value-of α or some other parameter.