Notes Section 2

Stationary and nonstationary random variables

Many economic time series are not plausibly characterized as processes with a constant mean; expenditure, income, and price series typically display a tendency to increase over time. Such series are described as nonstationary. To discuss the concept of stationarity, we must consider not only the first two moments of the series—the mean and variance—but also the autocovariance function, or autocorrelation function, of a single series, and the cross–covariance (or cross–correlation) function of a pair of time series. Recall that the covariance of X_t and Y_t is defined as $Cov(X_t, Y_t) = E[(X_t - \mu_x)(Y_t - \mu_y)]$. This definition is only valid for timeseries with constant means; in that context, we may also define $Cov(X_t, Y_{t-\tau})$ as the cross–covariance of those series with a τ–period lag: that is, we are considering the covariance between X at time t and Y τ periods prior (or τ periods hence, since the covariance function is symmetric).

Special cases of these definitions occur when X and Y are the same series; we then consider the autocovariance function γ_k of the series X, whose zero–order element is merely the variance of the series, with other elements referring to the covariance between the series' values at one point in time versus another point in time.

For stationary series—those with a constant mean and variance—both of these concepts may be transformed from covariances into correlations with appropriate scaling. You may recall that the simple correlation coefficient between X and Y is merely the covariance scaled by the product of the standard deviations of the two variables. This implies that the autocorrelation function will have unity as its zero–order term (being the variance of X scaled by the square of the standard deviation of X, that is, the variance), and the autocovariances will be the elements of the γ_k sequence scaled by their variance: the autocorrelations of the series, usually denoted ρ_k. Like the autocovariance function, the autocorrelation function is symmetric, with the k^{th} autocorrelation reflecting the relation between X_t and X_{t-k} or X_{t+k}.

With these building blocks in hand, we may consider the concept of covariance stationarity, or second–order stationarity. For a stochastic process to be covariance stationary, three conditions must be satisfied: the process must have a constant mean μ_x, a constant variance σ_x^2, and its autocovariance function must not be a function of time. That is, $\sigma_{t,s} = \sigma_{t+j,t+j-s}$, so that translating the calculation of the autocovariance function along the time axis does not affect its value: the process measured at two different points in time, e.g. t and s, have an autocovariance depending only on their temporal displacement $k = t - s$. Thus, we may speak of the k^{th} order autocovariance, γ_k, without further reference to time.

We may note that covariance stationarity is itself a weak form of strict stationarity, which would require that the entire distribution of the stochastic process is independent of the measure of time. For a random variable distributed according to the Normal distribution, covariance stationarity implies strict stationarity, since its distribution only depends on first and second moments. In general, strict stationarity is a more restrictive condition, and as it is difficult to test, covariance stationarity will often suffice in applied work.

We may easily determine that a process such as a random walk, or a random walk with drift, cannot be covariance stationary. Such a process might be $X_t = \kappa + X_{t-1} + \epsilon_t$, where ϵ_t is a zero mean random process with a constant variance, σ^2_ϵ, and independent increments (and thus zero autocovariances). We may rewrite the process as $X_t = t\kappa + X_0 + \sum_{s=1}^{t}, \epsilon_s$, where X_0 is the fixed initial condition. The expectation of X_j for $j = 1, ..\tau$ will be $X_0 + \kappa$, $X_0 + 2\kappa$, ..., $X_0 + \tau\kappa$; thus the process has a continuously changing mean (as given by the drift). Likewise, the variance of the process for 1, 2, ... τ observations will be σ^2_ϵ, $2\sigma^2_\epsilon$, $\tau\sigma^2_\epsilon$: that is, the variance increases linearly. This process is clearly nonstationary, as it fails the first two conditions defining a covariance stationary process.

Definitions

A **stochastic process** is a sequence of random variables, $\{X_i\}, i = 1, 2, ...$ If the index is taken as representing time, then the stochastic process is a **time series**. The fundamental problem in time series analysis is that we observe the realization of the stochastic process only once. There are annual data, for example, on the U.S. inflation rate for 1946–1995: 50 real values. But this is only one possible outcome of the underlying stochastic process for the inflation rate over that period. If we could have observed history many times over, we could assemble many samples, each containing a possibly different string of 50 real numbers, and take their average for each year. This would be the **ensemble mean** of the series: the average across the states of nature at any given calendar time. In reality, we can only observe one such history. If the distribution for the inflation rate remains unchanged–essentially, the concept of stationarity–then the particular sequence, or time series, that we observe can be considered as 50 different values from the same distribution. And if the process is not too persistent–if it possesses the property of **ergodicity**–then each element of the sequence will bear some information, and the time average over the elements of the single realization we have will be consistent for the infeasible ensemble mean. A stationary process is ergodic if it is asymptotically independent: that is, if two random variables positioned far apart in the sequence are almost independently distributed. This is true, for instance, for the $AR(1)$ process $z_t = c + \rho z_{t-1} + \epsilon_t$, ϵ a white noise process and $|\rho| < 1$.

Most aggregate time series such as GDP are not stationary because they exhibit time trends. In some cases, financial series are argued to be nonstationary on the grounds of nonconstant variances. Many time series with a trend can be reduced to stationary processes: for instance, a series from which a linear trend has been removed, rendering it stationary, is said to be **trend stationary** (TS). Alternatively, a series may be differenced; if the series is nonstationary, but its difference is stationary, then the process is said to be **difference stationary** (DS). As we shall see, much of the concern over unit roots in economics and finance is related to this distinction between TS and DS processes.

A stochastic process is said to be **white noise** if it satisfies three properties: (a) $E[X_t] = 0$ for all t; (b) the variance of X_t is constant and thus independent, so that the process is said to be homoskedastic; and (c) all autocorrelations γ_k, $|k| \geq 1$, equal zero. This process is covariance stationary and ergodic—but note that not all covariance stationary processes are white noise.

A closely related concept: independently and identically distributed random variables, often labelled iid. The elements of a stochastic process are said to be iid if they possess three properties: (a) $E[X_t] = \mu_x$, a constant not necessarily equal to zero, for all t; (b) the variance of X_t is constant and time–independent: $\sigma_x^2(t) = \sigma_x^2$ for all t; and (c) X_t is distributed independently of X_s for all $t \neq s$. The latter is a stronger condition than the equivalent condition in the definition of white noise, since independence implies zero (auto)correlation but not vice versa. However, if we add the assumption that X is distributed Normal, an assumption of zero autocorrrelations is sufficient to imply independence. An iid sequence is stationary.

A **random walk** process, $X_t = X_{t-1} + \epsilon_t$, combines a white noise error sequence with the level of X. Since $\Delta X_t = \epsilon_t$, the first difference of X is also a white noise process, and stationary; it should be clear that the level process X_t is not stationary. Its mean is time–varying; its variance is infinite, as $T \rightarrow \infty$. and its autocorrelations are nonzero, and die out very slowly. A stochastic process is a **martingale** if it satisfies

$E\left[X_{t+1}|\Omega_t\right] = X_t$, .where Ω_t is the information set, containing at minimum the past history of the process. Since that conditional expectation evaluates to X_t for the random walk process, it is a martingale (given that $E\epsilon_{t+1} = 0$ for the white noise innovation). The difference $\epsilon_{t+1} = X_{t+1} - X_t$ is a **martingale difference sequence**, which generally only requires uncorrelated (rather than the more strictly independent) increments.

An example of a martingale difference process, often used in analysing asset returns, is Engle's autoregressive conditional heteroskedastic (**ARCH**) process. A process is said to be an ARCH process of order 1, or $ARCH(1)$, if it can be written as $g_t = \left(\sqrt{\zeta + \alpha g_{t-1}^2}\right)\epsilon_t$ where ϵ_t is iid with zero mean and unit variance. Since the increments to the process are the iid elements of ϵ, $E[g_t]$ conditioned on its own past history is zero, since that conditional expectation involves conditioning ϵ upon the past history of g. Likewise, $E[g_t^2]$ conditioned on the past history of the g process is merely $\zeta + \alpha g_{t-1}^2$. So the conditional second moment of the process (which is the conditional variance, since the conditional mean is zero) is a function of the history of the process. The g process is strictly stationary and ergodic if $|\alpha| < 1$. If g_t is stationary, then the unconditional second moment may be readily calculated as $E\left(g_t^2\right) = \frac{\zeta}{1-\alpha}$. As we shall see, if $\alpha > 0$, this model captures a characteristic of asset returns: volatility clustering, so that large values (in absolute terms) are followed by large values.