# 241B Lecture
## Ergodic Stationarity

The key concept in extending analysis to time series, is a stochastic process. A stochastic process is simply a formal name given to a sequence of random variables. If the index denotes time, then the stochastic process is called a time series. For the time series $\{Z_t\}$, a realization $\{z_t\}$ is called a sample path.

### Need for Ergodic Stationarity

The fundamental problem in time-series analysis is that we can observe the realization of the process only once. That is, we have only one sequence of observations corresponding to annual US GDP growth for the post-war period. The sequence is just one of many that could have arisen from the stochastic process, if history had taken a different course we would have a different realization. If we could observe history many times over, then we could assemble a collection of strings for post-war GDP growth. Suppose that 2000 corresponds to the 55th element of the series. Then to estimate the average growth rate for 2000, we would simply average the 55th element across the realizations to generate the ensemble mean. In the language of general equilibrium economics, the ensemble mean is the average across all possible states of nature at any point in calendar time.

Of course, it is not possible to observe many different alternative histories. But if the distribution of the growth rate remains unchanged over time (a property referred to as **stationarity**) the particular string of 55 numbers that we do observe can be viewed as 55 observations from the same distribution. Further, if the process is not too persistent (what is called **ergodicity** has this property) each element from the string will contain some information not available from the other elements and, as show below, the time average over the elements of the single string will be consistent for the ensemble mean.

### Various Classes of Stochastic Processes

#### Stationary Processes

A stochastic process $\{Z_t\}$ $(t = 1, 2, \ldots)$ is strictly stationary if, for any given finite integer $r$ and for any set of subscripts $t_1, t_2, \ldots, t_r$, the joint distribution of $(Z_t, Z_{t_1}, Z_{t_2}, \ldots, Z_{t_r})$ depends only on $t_1 - t, t_2 - t, \ldots, t_r - t$ but not on $t$. For example, the distribution of $(Z_1, Z_3)$ is the same as the distribution of $(Z_5, Z_7)$. As the distribution depends only on the relative positions within the sequence, rather than the absolute position given by $t$, the mean, variance and all other higher moments (if they exist) are the same across $t$. The definition also implies that any transformation of a stationary process is itself stationary, that is if $\{Z_t\}$ is stationary then $\{f(Z_t)\}$ is stationary. (The function $f$ needs to be measurable so that $f(Z_t)$ is a well-defined random variable. Any continuous function is measurable. We assume henceforth that $f$ is measurable without use of the word measurable, when $f(Z_t)$ is understood to be a random variable.) For example $\{Z_t Z_t'\}$ is stationary if $\{Z_t\}$ is stationary.

**Example (i.i.d. sequences):** A sequence of independent and identically distributed random variables is a stationary process that exhibits no serial dependence.

**Example (constant sequence):** Draw $Z_1$ from some distribution and then set $Z_t = Z_1$ $(t = 2, 3, \ldots)$. The value of the process is exactly determined at the initial observation. This process is a stationary process that exhibits maximal serial dependence.

Evidently, if a vector process $\{Z_t\}$ is stationary, then each element of the vector forms a univariate stationary process. The converse, however, is not true.

**Example (element-wise vs. joint stationarity):** Let $\{U_t\}$ $(t = 1, 2, \ldots)$ be a scalar i.i.d. process. Create a two-dimensional process $\{Z_t\}$ as

$$Z_{t1} = U_t \quad \text{and} \quad Z_{t2} = U_1.$$

Each scalar process is stationary (they are merely the first two examples). The vector process $\{Z_t\}$ is not jointly stationary, however, because the (joint) distribution of $Z_1 = (U_1, U_1)'$ differs from that of $Z_2 = (U_2, U_1)'$.

Most aggregate time series, such as GDP, are not stationary because they exhibit time trends. A less obvious example of nonstationarity occurs with many asset prices, which are alleged to have increasing variance. Many time series with trends can be rendered stationary. A process is called trend stationary if it is stationary after subtracting from it a (usually linear) function of time. A process is called difference stationary if the first difference $Z_t - Z_{t-1}$ is stationary.

### Covariance Stationary Processes

A stochastic process $\{Z_t\}$ is covariance (weakly) stationary if:

(i) $E(Z_t)$ does not depend on $t$, and

(ii) $Cov(Z_t, Z_{t-j})$ exists, is finite and depends only on $j$ but not on $t$.

The relative, not absolute, position in a sequence is all that matters for a covariance stationary process. Evidently, if a stochastic process is strictly stationary *and* if the covariances are finite, then the sequence is weakly stationary (hence the term strict). An example of a covariance stationary but not strictly stationary process is the ARCH process given below.

The $j$-th order autocovariance for a covariance stationary process, denoted $\Gamma_j$, does not depend on $t$. The quantity is defined as

$$\Gamma_j \equiv Cov(Z_t, Z_{t-j}) \quad j = 0, 1, 2, \ldots$$

Covariance stationarity also implies

$$\Gamma_j = \Gamma'_{-j}.$$

For a scalar series, $\gamma_j = \gamma_{-j}$.

For a scalar series, take a string of $n$ consecutive values $(Z_t, Z_{t+1}, \ldots, Z_{t+n-1})$. By covariance stationarity, the covariance matrix of this string is the same as the covariance matrix for $(Z_1, \ldots, Z_n)$ and is a band spectrum matrix

$$Var\left(Z_{t+1}, \ldots, Z_{t+n}\right) = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \gamma_{n-2} & \cdots & \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_{n-1} & \cdots & \gamma_2 & \gamma_1 & \gamma_0 \end{bmatrix}.$$

This is called the autocovariance matrix of the process. The $j$-th order autocorrelation coefficient is defined as

$$\rho_j \equiv \frac{\gamma_j}{\gamma_0} = \frac{Cov\left(Z_t, Z_{t-j}\right)}{Var\left(Z_t\right)} \quad (j = 1, 2, \ldots).$$

For $j = 0$, $\rho_j = 1$. The plot of $\left\{\rho_j\right\}$ against $j = 0, 1, 2, \ldots$ is called the correlogram.

White Noise Processes

An important class of covariance stationary processes are white noise processes, which have zero mean and are serially uncorrelated.

**Definition (white noise processes):** A covariance stationary process $\{Z_t\}$ is white noise if

$$EZ_t = 0 \text{ and } Cov\left(Z_t, Z_{t-j}\right) = 0 \text{ for all } j \neq 0.$$

Clearly an independent and identically distributed sequence with mean 0 and finite variance is a special case of a white noise process. For this reason, it is called an independent white noise process.

**Example (white noise process that is not strictly stationary):** Let $w$ be a random variable uniformly distributed in the interval $(0, 2\pi)$. Define

$$Z_t = \cos(tw) \quad (t = 1, 2, \ldots).$$

It can be shown that $EZ_t = 0$, $Var(Z_t) = \frac{1}{2}$ and $Cov(Z_t, Z_s) = 0$ for $t \neq s$. So $\{Z_t\}$ is white noise. However, the values are clearly dependent, so the sequence is not independent white noise. Further, the process is not strictly stationary. (Suppose $tw = 2\pi$ vs. $tw = \pi/4$. The joint distribution of two observations will depend on their relative placement in the sequence.)

Ergodicity

A stationary process is said to be ergodic if, for any two bounded functions $f : \mathbb{R}^k \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}$,

$$\lim_{n\to\infty} \left[ \left| E\left[ f(Z_t, \ldots, Z_{t+k})\, g(Z_{t+n}, \ldots, Z_{t+n+m}) \right] \right| - \left| E\left[ f(Z_t, \ldots, Z_{t+k}) \right] \right| \left| E\left[ g(Z_{t+n}, \ldots, Z_{t+n+m}) \right] \right| \right] = 0.$$

Heuristically, a stationary process is ergodic if it is asymptotically independent, that is, if any two random variables positioned far apart in the sequence are almost independently distributed. (Note: $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$ if $X$ and $Y$ are uncorrelated, but we are looking at arbitrary functions of $X$ and $Y$ and so independence is the term to be used here.) A stationary process that is ergodic is called ergodic stationary. Ergodic stationarity is integral in developing large-sample theory because of the following property.

**Ergodic Theorem:** Let $\{Z_t\}$ be an ergodic stationary process with $EZ_t = \mu$. Then

$$\bar{Z}_n \equiv \frac{1}{n} \sum_{t=1}^{n} Z_t \overset{as}{\to} \mu.$$

By assuming $EZ_t = \mu$, we assume the mean exists and is finite. The Ergodic Theorem is a substantial generalization of the Kolmogorov LLN as serial dependence is allowed for. Indeed, all we need is for the dependence to vanish asymptotically. The theorem is even more useful when combined with the fact that for any measurable function $f(\cdot)$:

$$\{Z_t\} \text{ ergodic stationary } \Rightarrow \{f(Z_t)\} \text{ ergodic stationary.}$$

Hence any moment of an ergodic stationary process, provided it exists and is finite, is consistently estimated by the corresponding sample moment. For example, if

$\{Z_t\}$ is ergodic stationary and $E(Z_t Z_t')$ exists and is finite, then $\frac{1}{n}\sum Z_t Z_t'$ is a consistent estimator of $E(Z_t Z_t')$.

**Examples (ergodic stationary processes):**

(i) The AR(1) process satisfying

$$Y_t = \alpha + \rho Y_{t-1} + U_t \quad \text{with} \quad |\rho| < 1.$$

(ii) Independent white noise. (Note, white noise processes with independence weakened to no serial correlation are not necessarily ergodic stationary as the previous example makes clear.)