Large Sample Theory: Lecture XIV

Charles B. Moss

October 3, 2010

Basic Sample Theory

- The problems set up is that we want to discuss sample theory.
  - First assume that we want to make an inference, either estimation or some test, based on a sample.
  - We are interested in how well parameters or statistics based on that sample represent the parameters or statistics of the whole population.
- The complete statistical term is known as convergence.
  - Specifically, we are interested in whether or not the statistics calculated on the sample converge toward the population estimates.
  - Let $\{X_n\}$ be a sequence of samples. We want to demonstrate that statistics based on $\{X_n\}$ converge toward the population statistics for X.

- Taking a slightly different tack: The classical assumptions for ordinary least squares (OLS) as presented in White, Halbert Asymptotic Theory for Econometricians.
  - **Theorem 1.1**: The following are the assumptions of the classical linear model
    - ★ The model is known to be $y = X\beta + \epsilon$, $\beta < \infty$.
    - ★ X is a nonstochastic and finite $n \times k$ matrix.
    - ★ $X'X$ is nonsingular for all $n \leq k$.
    - ★ $\mathrm{E}(\epsilon) = 0$.
    - ★ $\epsilon \sim N\left(0, \sigma_0^2 I\right)$, $\sigma_0^2 < \infty$.

- Continued
  - Given these assumptions, we can conclude that
    - *Existence* given (i) - (iii) β_n exists for all $n \geq k$ and is unique.
    - *Unibiasedness* given (i) - (v) $\mathrm{E}\left[\beta_n\right] = \beta_0$.
    - *Normality* given (i) - (v) $\beta_n \sim N\left(\beta_0, \sigma^2 \left(X'X\right)^{-1}\right)$.
    - *Efficiency* given (i) - (v) β_n is the maximum likelihood estimator and the best unbiased estimator in the sense that any other unbiased estimator exceeds that of β_n by a positive semi-definite matrix regardless of the value of β_0.

- Continued
  - Existence, unbiasedness normality and efficiency are small sample analogs of asymptotic theory.
    - ★ Unbiased implies that the distribution of β_n is centered around β_0.
    - ★ Normality allows us to construct t-distribution or F-distribution tests for restrictions.
    - ★ Efficiency guarantees that the ordinary least squares estimates have the greatest possible precision.
  - Asymptotic theory involves the behavior of the estimator under the failure of certain assumptions. Specifically, assumptions (ii) or (v).
  - Finally, within the classical linear model the normality of the error term is required to strictly apply t-distributions or F-distributions. However, the central limit theorem can be used if n is large to guarantee that β_n is approximately normal.

Modes of Convergence

- **Definition 6.1.1** A sequence of real numbers $\{\alpha_n\}$, $n = 1, 2, \cdots$ is said to converge to a real number α if for any $\epsilon > 0$ there exists an integer N such that for all $n > N$ we have

$$|\alpha_n - \alpha| < \epsilon \tag{1}$$

  - This convergence is expressed as $\alpha_n \to \alpha$ as $n \to \infty$ or $\lim_{n\to\infty} \alpha_n = \alpha$.
  - This definition must be changed for random variables because we cannot require a random variable to approach a specific value. Instead, we require the probability of the variable to approach a given value. Specifically, we want the probability of the event to equal 1 or zero as n goes to infinity.

- **Definition 6.1.2** (convergence in probability) A sequence of random variables $\{ X_n \}$, $n = 1, 2, \cdots$ is said to converge to a random variable X in probability if for any $\epsilon > 0$ and $\delta > 0$ there exists an integer N such that for all $n > N$ we have $P\left(|X_n - X| < \epsilon\right) > 1 - \delta$. We write

$$X_n \xrightarrow{P} X \tag{2}$$

$n \to \infty$ or $\operatorname{plim}_{n\to\infty} X_n = X$. The last equality reads the probability limit of $X_n$is X. (Alternatively, the if clause may be paraphrased as follows: if $\lim P\left(|X_n - X| < \epsilon\right) = 1$ for any $\epsilon > 0$.
    - Since we are dealing with convergence in the probability, we need to introduce two concepts: convergence in mean square and convergence in distribution.

- **Definition 6.1.3.** (*Convergence in Mean Square*) A sequence $\{\, X_n\}$ is said to *converge* to X in *mean square* if $\lim_{n\to\infty} \mathrm{E}\,(X_n - X)^2 = 0$. We write

$$X_n \overset{M}{\to} X \tag{3}$$

- **Definition 6.1.4.** (*Convergence in Distribution*) A sequence $\{ X_n \}$ is said to *converge* to X *in distribution* if the distribution function F_n of X_n converges to the distribution function F of X at every continuity point of F. We write

$$X_n \xrightarrow{d} X \tag{4}$$

and call F the limit distribution of $\{ X_n \}$. If $\{ X_n \}$ and $\{ Y_n \}$ have the same limit distribution, we write

$$X_n \stackrel{LD}{=} Y_n \tag{5}$$

- **Theorem 6.1.1** (*Chebyshev*)

$$X_n \stackrel{M}{\rightarrow} X \Rightarrow X_n \stackrel{P}{\rightarrow} X \tag{6}$$

- **Theorem 6.1.2**

$$X_n \stackrel{P}{\rightarrow} X \Rightarrow X_n \stackrel{d}{\rightarrow} X \tag{7}$$

- Chebyshev's Inequality

$$P\left[g\left(X_n\right) \geq \epsilon^2\right] \leq \frac{\mathrm{E}\left[g\left(X_n\right)\right]}{\epsilon^2} \tag{8}$$

- **Theorem 6.1.3** Let X_n be a vector of random variables with a fixed finite number of elements. Let g be a function continuous at a constant vector point α. Then

$$X_n \xrightarrow{P} \alpha \Rightarrow g\left(X_n\right) \xrightarrow{P} g\left(\alpha\right) \tag{9}$$

- **Theorem 6.1.4** (*Slutsky*) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} \alpha$, then

$$\begin{aligned} X_n + Y_n &\xrightarrow{d} X + \alpha \\ X_n Y_n &\xrightarrow{d} \alpha X \\ \left(\frac{X_n}{Y_n}\right) &\xrightarrow{d} \frac{X}{\alpha} \text{ if } \alpha \neq 0 \end{aligned} \tag{10}$$