Econ 573

Helle Bunzel

ASYMPTOTIC THEORY – Handout #1

Summary

Definition: Converge almost surely

A sequence of random variables $\{X_n\}$ is said to coverage almost surely (a.s.) to a random variable X, denoted by $X_n \xrightarrow{a.s} X$, if

$$\Pr(\lim_{n\to\infty} X_n = X) = 1$$

An equivalent way of defining almost sure convergence is by

$$\lim_{n\to\infty} \Pr(\ |X_m - X\ | < e,\ \text{all } m \geq n) = 1$$

The almost sure convergence is the mode of convergence associated with the strong law of large numbers (SLLN).

Definition: Converge in r'th mean

Let $\{X_n\}$ be a sequence of random variables such that $E(|X_n|^r) < \infty$ for all $n \in N$ and $E(|X|^r) < \infty$ for r > 0, then the sequence converges to X in r'th mean, denoted by $Xn \to X$, if

$$\lim_{n\to\infty} E(|X_n - X|^r) = 0$$

Of particular interest in what follows is the convergence in mean (r = 1) and mean square (r = 2).

Definition: Converge in probability

A sequence of random variables $\{X_n\}$ is said to converge in probability to a random variable X, denoted by $X_n \xrightarrow{P} X$, if

$$\lim_{n\to\infty} \Pr(|X_n - X| < e) = 1$$

Definition: Converge in distribution

A sequence of random variables $\{X_n\}$ with distribution functions $\{F_n(x)\}$ is said to converge in distribution to X, denoted by $X_n \xrightarrow{D} X$ if

$$\lim_{n \to \infty} F_n(x) = F(x)$$

at every continuity point x of F(x).

The mode of convergence related to the central limit theorem (CLT) is that of convergence in distribution.

Another way of looking at the convergence in probability is:

Consider the sequence of stochastic variables

$X_1, X_2, ..., X_n, ...$

and the sequence of their distribution functions

$F_1, F_2, ..., F_n, ...$

X_n converges to X in probability if

$$\lim_{n \to \infty} Pr\ (|X_n - X| > e) = 0 \quad for\ any\ e > 0$$

For a given e and the distribution function F_n ( ) we can find the probability $P(X_n < X - e) + P(X_n > X + e)$.

The sequence of these probabilities is

$a_1, a_2, ..., a_n, ...$

Convergence in probability requires that this sequence has the limit 0.

Convergence in probability is stated as

$$X_n \xrightarrow{P} X \text{ or } \operatorname{plim} X_n = X$$

LEMMA

Let $\{X_n\}$ be a random vector sequence and $g(\cdot)$: $R^K \to R$ a continuous function at X, then

1) $X_n \xrightarrow{a.s.} X \Rightarrow g(X_n) \xrightarrow{a.s.} g(X)$

2) $X_n \xrightarrow{P} X \Rightarrow g(X_n) \xrightarrow{P} g(X)$
3) $X_n \xrightarrow{D} X \Rightarrow g(X_n) \xrightarrow{D} g(X)$

This lemma can be used to prove the following results:

plim(X_n + Y_n) = plim X_n + plim Y_n = μ + ?
plim(X_n • Y_n) = plim X_n • plim Y_n = μ • ?

$$\text{plim}\left(\frac{X_n}{Y_n}\right) = \frac{\text{plim } X_n}{\text{plim } Y_n} = \frac{m}{x}$$

plim g(X_n) = g(plim X_n) = g(μ)

$$\text{plim } X_n^P = (\text{plim } X_n)^P = m^P$$

$$\text{plim } A_T^{-1} = (\text{plim } A_T)^{-1}$$

$$\text{plim } A_T \cdot B_T = \text{plim } A_T \cdot \text{plim } B_T$$

Definition: Consistent estimator

$\hat{q}_n$ is a consistent estimator of ? if

$$\text{plim } \hat{q}_n = q$$

Theorem. Sufficient conditions for convergence in probability

If $\lim_{n \to \infty} E(\hat{q}_n) = q$

and $\lim_{n \to \infty} V(\hat{q}_n) = 0$

then plim $\hat{q} = q$.

Otherwise convergence in probability is often proved using Chesbychef's inequality

$$\Pr\ (|X - m| \geq k \cdot s\ ) \leq \frac{1}{k^2}$$

Convergence in distribution can be proved by finding the limit of the sequence of characteristic functions

f_n = E(exp(i • n • X_n))

$f_n \to f$ for $n \to \infty \Rightarrow F_n \xrightarrow{D} F$

For simple cases this is not necessary.

Example

Let $X_i \sim \text{nid}(\mu, s^2)$

Let $\overline{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n}$

Then the sequence $\{X_n\}$

$\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_n, \ldots$

has the following sequence of distribution functions

$N(\mu, \frac{\sigma}{\sqrt{1}}), N(\mu, \frac{\sigma}{\sqrt{2}}), \ldots, N(\mu, \frac{\sigma}{\sqrt{n}})$

This sequence of distribution functions degenerates to the non-stochastic constant μ because the variance goes to zero.

Instead we can look at $\sqrt{n}\overline{X}$, or usually we look at

$$\sqrt{n}\, Z_n = \sqrt{n}\, \frac{\overline{X} - \mu}{\sigma}$$

$\{\sqrt{n} Z_n\}$ has the following sequence of distribution functions

N(0,1), N(0,1), ..., N(0,1), ...
which obviously has the limit N(0,1).

Otherwise one of the following Lemmas, the Cramer's Theorem, or the Mann-Wald Theorem, can be used to find the limiting distribution function.

Lemma

Let $\{X_n, Y_n\}$ be a sequence of pair of random $k \times 1$vectors.

Then:

(1) If $(X_n - Y_n) \xrightarrow{P} 0$ and $X_n \xrightarrow{D} X \Rightarrow Y_n \xrightarrow{D} Y$:

(2) If $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} 0 \Rightarrow X_n \cdot Y_n \xrightarrow{P} 0$

(3) If $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} C$ (constant) $\Rightarrow (X_n + Y_n) \xrightarrow{D} X + C$

If $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} C$ (constant) $\Rightarrow Y_n X_n \rightarrow CX$

(4) (for Y_n and C kxk non- singular) $Y_n^{-1} X_n \xrightarrow{D} C^{-1} X$

CRAMÉRS THEOREM

Let H_n be kxm matrix, such that

(i) $\underset{n\rightarrow\infty}{plim} H_n = H$

where H is a kxm matrix with rank kxm.

Let ?n be an m element stochastic vector such that

(ii) $\sqrt{n}\,(x_n - x) \xrightarrow{D} N(0, \Sigma)$

where ? is an m element non-stochastic vector, and S is an mxm positive semidefinite matrix.

Define the k element stochastic vector ?n = Hn?n. Then

(1) $\sqrt{n}\,(h_n - H\,x) \xrightarrow{D} N(0, H\,\Sigma\,H')$

MANN-WALD THEOREM

Let z_n be a k element stochastic vector such that

(i) $\text{plim}\,\frac{1}{n}\sum_{t-1}^{n} Z_t Z_t = Q$ or $\text{plim}\,\frac{Z'Z}{n} = Q$

where Q is a kxk positive definite matrix.

Assume

(ii) $u_t \sim \text{iid}(0,s^2)$

(iii) $E(Z_t u_t) = 0$

Then

(1) $\text{plim}\,\frac{1}{n}\sum_{t=1}^{n} Z_t u_t =$ or $\text{plim}\,\frac{Z'u}{n} = 0$

and

(2) $$\sqrt{n}\,\left(\frac{1}{n}\sum_{t=1}^{n} Z_t u_t\right) = \frac{1}{\sqrt{n}}\left(\sum_{t=1}^{n} Z_t\, u_t\right) \xrightarrow{D} N(0, s^2 Q)$$

Using a Taylor series approximation and Cramer's Theorem, the following theorem can be proved.

Theorem. Asymptotic Distribution of a Nonlinear Function

$$\text{If } \sqrt{n}\,(\hat{q} - q) \xrightarrow{D} N(0, \Sigma)$$

and if g(?) is a continuous function and its first and second order derivatives exist, then

$$\sqrt{n}\,(g(\hat{q}) - g(q)) \xrightarrow{D} N(0, G(q)\Sigma G(q)')$$

$$\text{where } G(\theta) = \left\{\frac{\partial g(\theta)}{\partial \theta}\right\}$$

Lemma. plim of sample moment

Let m^k be a sample moment of $\mu_k = E(x^k)$.
Assume that μ_{2k} exists then plim $m_k = \mu_k$.

Lindberg-Feller CLT (Greene Thm 4.14)

Suppose that $x_1, x_2, x_3, ..., x_n$ are a sample of random vectors such that $E[x_i] = m_i, Var[x_i] = Q_i$,

and all mixed third moments of the multivariate distribution are finite. Let

$$\overline{m}_n = \frac{1}{n}\sum_{i=1}^{n} m_i,$$
$$\overline{Q}_n = \frac{1}{n}\sum_{i=1}^{n} Q_i.$$

We assume that

$$\lim_{n\to\infty} \overline{Q}_n = Q,$$

where Q is a finite, positive definite matrix, and that for every i,

$$\lim_{n\to\infty}\left(\overline{nQ}_n\right)^{-1}Q_i = 0.$$

Then

$$\sqrt{n}\left(\overline{x}_n - \overline{m}_n\right) \xrightarrow{d} N(0,Q)$$