HANDOUT on Convergence, WLLN and CLT

1 Various Modes of Convergence

Definition 1 (Convergence in probability) *A sequence of r. var. $\{h_n\}$ converges in probability to a constant (non-random) scalar h, if, for any $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(|h_n - h| > \epsilon) = 0.$$

In this case, h is known as the probability limit or plim h_n. This definition is equivalent to $h_n \xrightarrow{p} h$.

The definition of convergence in probability implies that for each $\epsilon > 0$ there exists a $n_\epsilon > 0$ such that the probability of $|h_n - h| < \epsilon$ is arbitrarily close to one for all $n \geq n_\epsilon$. Convergence in probability can be recognized as the natural extension of the concept of convergence for deterministic sequences. Next, we define the concepts of orders in magnitude for sequences of random variables.

Definition 2 (Small order in probability) *The sequence of r.var. $\{h_n\}$ is said to be of small order in probability n^a, a being a real number, if:*

$$plim \frac{h_n}{n^a} = 0.$$

This is denoted by $h_n = o_p(n^a)$; equivalently, for any $\epsilon > 0$, $\lim_{n\to\infty} P(\frac{h_n}{n^a} > \epsilon) = 0$, or $\frac{h_n}{n^a} \xrightarrow{p} 0$.

We usually use this definition for $a = 0$, in which case $h_n = o_p(1)$.

Definition 3 (Large order in probability) *The sequence of random variables $\{h_n\}$ is said to be of large order in probability n^a, a being a real number, if for every $\epsilon > 0$, there exists a positive real number m_ϵ such that*

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{|h_n|}{n^a} > m_\epsilon\right] = 0.$$

This is denoted by $h_n = O_p(n^a)$.

The notion of large order in probability can be seen here to imply that h_n/n^a converges in probability to a random variable or a constant, that is, it does not diverge to infinity but it

can converge to zero. Both types of order in probability are very useful in asymptotic analysis because the can be linked to consistency and convergence in distribution, as discussed below. Before that, we need to define the notion of convergence in probability to vectors or matrices.
A vector (or matrix) h_n is said to converge in probability to h if the i^{th} (or $(i,j)^{th}$) element of h_n converges in probability to the i^{th} (or $(i,j)^{th}$) element of h. We write $h_n = O_p(n^a)$ or $h_n = o_p(n^a)$ to indicate that all elements of the vector or matrix individually satisfy the stated order in probability.

Definition 4 (Consistency) *Let $\hat{\delta}_n$ be a sequence of estimators of some unknown parameter vector of constants δ. Then $\hat{\delta}_n$ is said to be a consistent estimator of δ if and only if plim $\hat{\delta}_n = \delta$, or, equivalently, $\hat{\delta}_n \xrightarrow{p} \delta$.*

If plim $\hat{\delta}_n \neq \delta$, then the estimator is said to be inconsistent. Note that consistency implies that $\hat{\delta}_n - \delta = o_p(1)$. It is a weak property because it only says that as $n \to \infty$, the estimator converges in probability to the true value. This convergence property is a sense in which $\hat{\delta}_n$ becomes closer to δ as the sample size increases, and it is reasonable to be concerned if such a weak property is not satisfied.
In deriving asymptotic distributions, it will be convenient to appeal to the weaker notion of convergence in distribution. For this definition, we revert to the earlier notation involving h_n rather than $\hat{\delta}_n$ because we will apply this notion not only to estimators, but also to other quantities of interest.

Definition 5 (Convergence in distribution) *The sequence of random vectors h_n with corresponding distribution functions $\{F_T(\cdot)\}$ converges in distribution to the random vector h with distribution function $\{F(\cdot)\}$ if and only if*

$$\lim_{n\to\infty} |F_n(c) - F(c)| = 0$$

at every point of continuity $\{c\}$. This is denoted by $h_n \xrightarrow{d} h$.

The distribution of h is known as the limiting (or asymptotic) distribution of h_n. If h_n converges in distribution, then $h_n = O_p(1)$. In practice, our focus is not just on establishing that h_n converges in distribution, but also on characterizing the exact nature of its limiting distribution.

2 The Weak Law of Large Numbers and the Central Limit Theorem

WLLN 1 *Let $\{h_i, i = 1, \ldots, n\}$ be an $p \times 1$ sequence of independent random variables with means $\mu_i = E[h_i]$ that exist and are finite. If $\sup_t E||h_i||^{1+\delta} < M < \infty$, for some $M, \delta > 0$ and $||\cdot||$ the Euclidean norm, then:*

$$\frac{1}{n}\sum_{i=1}^{n} h_i \xrightarrow{p} \mu,$$

where $\mu = \lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)$ exists and is finite.

CLT 1 *Let $\{h_i, i = 1, \ldots, n\}$ be an $p \times 1$ sequence of independent random variables with means $\mu_i = E[h_i]$ that exist and are finite. If $Var(h_i) = \Sigma$, a positive definite (pd) matrix of constants, and $\sup_t E||h_i||^{2+\delta} < M < \infty$, for some $M, \delta > 0$, then:*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(h_i - \mu_i) \xrightarrow{d} \mathcal{N}(0, \Sigma),$$