Econ 240A - Part II - Fall 2005
Section 1: Convergence Concepts, LLN, CLT and Delta Method

Maria Carolina Caetano

Friday, October 28

1 Review of Concepts:

1.1 Convergence in Probability

Definition 1 *A sequence $\{X_n\}$ of random variables is said to converge in probability to a random variable X, denoted $X_n \longrightarrow_p X$ if*
$P(|X_n - X| > \varepsilon) \longrightarrow 0, \quad \forall \varepsilon > 0$

Definition 2 *A sequence $\{X_n\}$ of random variables is said to converge almost surely to a random variable X, denoted $X_n \longrightarrow_{a.s.} X$ if*
$P(X_n \to X) = 1$

1.1.1 Interesting inequalities:

Theorem 3 *(Markov)* $P(|X| > a) \leq \frac{E(|X|)}{a}$

Theorem 4 *(Chebychev) Suppose* $\mu = E(X)$*, then* $P(|X - \mu| > \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$.

Theorem 5 *(Jensen) If f is a convex function, then* $E(f(X)) \geq f(E(X))$*, provided* $E(|X|), E(|f(X)|) < \infty$.

Theorem 6 *(Minkowski)* $X, Y \in L^p \Rightarrow \|X + Y\|_p \leq \|X\|_p + \|Y\|_p$.

Theorem 7 *(Holder)* $X \in L^p, Y \in L^q, \frac{1}{p} + \frac{1}{q} = 1 \Rightarrow E(|XY|) \leq \|X\|_p \|X\|_q$

Theorem 8 *(Cauchy-Schwarz)* $X, Y \in L^2 \Rightarrow \sqrt{E(X^2)E(Y^2)}.$

Theorem 9 *(Lyapunov)* $X \in L^s; r \leq s \Rightarrow \|X\|_r \leq \|X\|_s$.

1.1.2 Laws of Large Numbers (LLN):

Theorem 10 *(Weak Law of Large Numbers) Let $X_1, X_2, ...$ be a sequence of iid random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 < \infty$. Then*

$$\overline{X}_n \longrightarrow_p \mu,$$

where $\overline{X}_n = \frac{1}{n} \sum\limits_{i=0}^{n} X_i$.

Theorem 11 *(Strong Law of Large Numbers - Kolmogorov) Let $X_1, X_2, ...$ be a sequence of iid random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 < \infty$. Then*

$$\overline{X}_n \longrightarrow_{a.s.} \mu,$$

1.2 Convergence in Distribution:

Definition 12 *A sequence $\{X_n\}$ of random variables is said to converge in distribution to a random variable X, denoted $X_n \longrightarrow_d X$ if*
$P(X_n \leq x) \longrightarrow P(X \leq x), \quad \forall x$ *which is a continuity point of* $P(X \leq \cdot)$.
This definition can also be stated like this:

Definition 13 *A sequence $\{X_n\}$ of random variables, with cdf (cumulative distribution function) F_n is said to converge in distribution to a random variable X, denoted $X_n \longrightarrow_d X$ if*
$F_n(x) \longrightarrow F(x), \quad \forall x$ *where F is continuous.*

Definition 14 *An alternative way of defining convergence in distribution, also known as the* ***Helly-Bray theorem*** *is: $X_n \longrightarrow_d X$ if $E\left(f(X_n)\right) \longrightarrow E(f(X))$, for all function f bounded and continuous.*

1.2.1 Central Limit Theorem (CLT):

Theorem 15 *Let $X_1, X_2, ...$ be a sequence of iid random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 \in (0, \infty)$. Then*

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \longrightarrow_d N(0,1)$$

or, as you can see in other contexts,

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \longrightarrow_d N(0,1)`$$

1.3 Good Results and Techniques to know:

Theorem 16 *Let $X_1, X_2, ...$ be a sequence of random variables, then:*
a) $X_n \longrightarrow_{a.s.} X \Rightarrow X_n \longrightarrow_p X$
b) $X_n \longrightarrow_p X \Rightarrow X_n \longrightarrow_d X$
c) $X_n \longrightarrow_d c \Rightarrow X_n \longrightarrow_p c$*, where c is a constant (or a random variable X with $P(X = c) = 1$)*

Theorem 17 *Let $X_1, X_2, ...$ be a sequence of random variables, then:*
a) $X_n \longrightarrow_p X$ *does not imply* $X_n \longrightarrow_{a.s.} X$
b) $X_n \longrightarrow_d X$*, (X non degenerate random variable) does not imply* $X_n \longrightarrow_p X$

Theorem 18 *(Slutsky) If $X_n \longrightarrow_d X$, $Y_n \longrightarrow_p c$, then*
a) $X_n Y_n \longrightarrow_d cX$
b) $X_n + Y_n \longrightarrow_d X + c$

Theorem 19 *Let X, X_n, Y and Y_n be random variables, then:*
a) $X_n \longrightarrow_{a.s.} X$ *and* $Y_n \longrightarrow_{a.s.} Y \Rightarrow X_n + Y_n \longrightarrow_{a.s.} X + Y$
b) $X_n \longrightarrow_p X$ *and* $Y_n \longrightarrow_p Y \Rightarrow X_n + Y_n \longrightarrow_p X + Y$

Theorem 20 *(Continuous mapping) Let X, X_n, Y and Y_n be random variables, then:*
a) $X_n \longrightarrow_{a.s} X$*, and $g(\cdot)$ is a continuous function* $\Longrightarrow g(X_n) \longrightarrow_{a.s} g(X)$
b) $X_n \longrightarrow_p X$*, and $g(\cdot)$ is a continuous function* $\Longrightarrow g(X_n) \longrightarrow_p g(X)$
c) $X_n \longrightarrow_d X$*, and $g(\cdot)$ is a continuous function* $\Longrightarrow g(X_n) \longrightarrow_d g(X)$

1.4 Characteristic Functions:

Definition 21 *The characteristic function (ch.f.) of a random variable X is defined by:*
$\varphi(t) = E(e^{itX}) = E(\cos(tX)) + iE(\sin(tx))$

1.4.1 Basic properties of characteristic functions:

Theorem 22 *The ch.f. φ is uniformly continuous and symmetric ($\varphi(t) = \varphi(-t)$ for each t.*

Theorem 23 *If x and Y are independent, then $\varphi_{X+Y}(t) = \varphi_x(t)\varphi_Y(t)$*

Theorem 24 *(Uniqueness) If $\varphi_X(t) = \varphi_Y(t)$ for all t, then $X =_d Y$.*

Theorem 25 *(Taylor expansion for ch. f.) If $E(X^2) < \infty$ then*
$\varphi(t) = 1 + itE(X) - \frac{t^2}{2}E(X^2) + o(t^2)$

The most important result concerning ch.f.:

Theorem 26 *(Continuity Theorem) Let $X_1, X_2, ...$ be a sequence of random variables with characteristic functions $\varphi_1(\cdot), \varphi_2(\cdot), ...$*

a) If $X_n \longrightarrow_d X$, then $\varphi_n(t) \longrightarrow \varphi(t)$ for all t.

b) If $\varphi_n(t) \longrightarrow \varphi(t)$ for all t, and $\varphi(t)$ is continuous at 0, then $X_n \longrightarrow_d X$.

1.4.2 The moment generating function:

Definition 27 *The moment generating function (mgf) function of X is the function $\psi(t) = E(e^{tX})$, provided that this expectation exists ($< \infty$) for all t in some neighborhood of the origin.*

Remark 28 *Provided that the expectation exists, the moment generating functions have all the same characteristics of the ch.f., but without the unpleasant complex part.*

Theorem 29 *If $\psi(t)$ exists, then for each k, $E(|X|^k) < \infty$ and $E(X^k) = \psi^{(k)}(0)$.*

Theorem 30 *(Continuity Theorem for mgf) Let $X_1, X_2, ...$ be a sequence of random variables which moment generating functions $\psi_1(\cdot), \psi_2(\cdot), ...$exist in some neighborhood U of 0. Then $X_n \longrightarrow_d X$ if and only if $\psi_n(t) \longrightarrow \psi(t)$ for all t.*

1.5 The Delta Method:

Theorem 31 *(Delta Method) Let Y_n be a sequence of random variables that satisfies $\sqrt{n(\overline{X}_n - \mu)} \longrightarrow_d N(0, \sigma)$. For a given function g and a specific value of θ, suppose that $g'(\theta)$ exists and is not 0. Then*

$\sqrt{n}(g(Y_n) - g(\theta)) \longrightarrow_d N(0, \sigma^2(g'(\theta))^2)$

2 Practicing...

Exercise 32 *We could show that $s^2$is a consistent estimator of σ^2.*

Exercise 33 *We could also show the consistency of s.*

Exercise 34 *Characteristic function of the $U[0, 1]$*

Exercise 35 *Characteristic Function of the Normal distribution*

Exercise 36 *Prove the CLT*

Exercise 37 *Let $X_1, X_2, ...$ be a sequence of iid random variables with $E(X_i) = \mu \neq 0$ and $Var(X_i) = \sigma^2 \in (0, \infty)$. Find the asymptotic variance of $\sqrt{n}\left(\frac{1}{\overline{X}_n} - \frac{1}{\mu}\right)$.*