Stat210B: Theoretical Statistics **Lecture Date: February 20, 2007**

Lecture 10

Lecturer: Michael I. Jordan *Scribe: Alex Shyr*

In Empirical Process theory, the notion of a sequence of stochastic processes converging to another process is important. The scalar analogy of this convergence is the CLT. This lecture is an introduction to Donsker's Theorem, one of the fundamental theorems of Empirical Process theory.

1 Weak Convergence (aka Conv. in Law, Conv. in Distribution)

Given the usual sample space $(\Omega, \mathcal{F}, P)$, random element $X : \Omega \to \mathcal{X}$. Let $\mathcal{A}$ be a σ-field of $\mathcal{X}$.

Define $C(\mathcal{X}, \mathcal{A})$ to be the space of continuous, bounded function class on $\mathcal{X}$, which is measurable on $\mathcal{A}$.

A sequence of probability measures Q_n converges weakly to Q if $Q_n f \to Qf, \forall f \in C(\mathcal{X}, \mathcal{A})$. Note that $\mathcal{A}$ must be smaller than the Borel σ-field $\mathcal{B}(\mathcal{X})$. An alternative field that works is the projection σ-field generated by the coordinate projection maps.

2 Continuous Mapping Theorem (van der Vaart, 1998, Cha.18)

Since weak convergence does not hold for all probability measures, we need conditions on the set $\mathcal{C}$ on which the limiting random element concentrates.

Definition 1. A set $\mathcal{C}$ is *separable* if it has a countable, dense subset.

A point X in $\mathcal{X}$ is *regular* if

$$\forall \text{ neighborhood } V \text{ of } X, \exists \text{ a uniformly continuous } g \text{ with } g(X) = 1 \text{ and } g \leq V.$$

Theorem 2. *Let H be an $\mathcal{A}/\mathcal{A}'$ measurable map from $\mathcal{X}$ into another metric space $\mathcal{X}'$. If H is continuous at each point of some separable, $\mathcal{A}$-measurable set $\mathcal{C}$ of regular points, then*

$$X_n \xrightarrow{\mathcal{L}} X \text{ and } P(X \in \mathcal{C}) = 1 \qquad \Rightarrow \qquad HX_n \xrightarrow{\mathcal{L}} HX$$

Some useful notes:

- a common function space $\mathcal{X}$ is $D[0,1]$, which is the set of all $\mathbf{R}$-valued, *cadlag* functions
- $d(x,y) = \sup_{0 \leq t \leq 1} |x(t) - y(t)|$ defines a metric, and closed balls for d generate the projection σ-field
- every point of $D[0,1]$ is regular, but $D[0,1]$ is not separable ...
- BUT the limit processes we will talk about concentrate on $C[0,1]$, which is separable.

Theorem 3 ("Stochastic equicontinuity" or "Asymptotic tightness"). *Let $X_1, ..., X_n$ be random elements of $D[0,1]$. Suppose that $P(X \in \mathcal{C}) = 1$ for some separable $\mathcal{C}$. Then $X_n \xrightarrow{\mathcal{L}} X$ iff*

(i) *Fidi convergence of X_n to X (ie. $\Pi_S X_n \xrightarrow{\mathcal{L}} \Pi_S X \quad \forall$ finite $S \subseteq [0,1]$)*

(ii) *$\forall \epsilon > 0, \delta > 0, \exists$ a grid $0 = t_0 < t_1 < ... < t_n = 1$ s.t. $\limsup_n P\{\max_i \sup_{J_i} |X_n(t) - X_n(t_i)| > \delta\} < \epsilon$, where $J_i = [t_i, t_{i+1})$*

3 Donsker's Theorem (for standard empirical process)

The first version of Donsker's theorem deals with the convergence of the empirical process U_n of random variables drawn uniformly from the unit interval, where

$$U_n t = \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} 1_{\{\xi_i \leq t\}} - t), \qquad \text{and } \xi_i \overset{iid}{\sim} U[0,1]$$

Definition 4 (Brownian Bridge). U is a Brownian Bridge iff

(i) $\forall$ finite subset $S \in [0,1]$, $\Pi_S U$ is Gaussian with zero mean,

(ii) covariances $E[U(s)U(t)] = s(1-t), \quad \forall 0 \leq s \leq t \leq 1$, and

(iii) U only has continuous sample paths.

Theorem 5. $U_n \xrightarrow{\mathcal{L}} U$, *where U is a Brownian Bridge.*

Proof. First check (i) of Theorem 3.

$$\begin{aligned} E[U_n s U_n t] &= \frac{1}{n}\sum_i E[(1_{\{\xi_i \leq t\}} - t)(1_{\{\xi_i \leq s\}} - s)] \\ &= \frac{1}{n}\sum_i \{P(\xi_i \leq s) - tP(\xi_i \leq s) - sP(\xi_i \leq t) + st\} \\ &= s(1-t) \end{aligned}$$

□

References

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.