Stats 300b: Theory of Statistics **Winter 2017**

Lecture 14 – February 22

Lecturer: John Duchi *Scribes: Emily Diana and Zihan Lin*

Warning: *these notes may contain factual errors*

Reading: VdV 18-19

Outline:

- Moduli in distribution
- Moduli of continuity
- Convergence in distribution
- Prohorov's Theorem
- Stochastic equi-continuity

Recap Last week, we defined the modulus of continuity of a process

$$\Delta_n(\theta) := R_n(\theta) - R(\theta) - (R_n(\theta_0) - R(\theta_0))$$

$$\omega_n(\delta) := \sup_{d(\theta,\theta_0)\leq\delta} |\Delta_n(\theta)|$$

Theorem 1. *If $R(\theta) \geq R(\theta_0) + \lambda d(\theta,\theta_0)^\alpha$ and $\mathbb{E}[\omega_n(s)] \leq c\frac{\delta^\beta}{\sqrt{(n)}}$, then if $\hat{\theta}_n = \arg\min_\Theta R_n(\theta)$,*

$$(\hat{\theta}_n - \theta_0) = o_p(n^{-\frac{1}{2(\alpha-\beta)}})$$

Example 1: *Suppose $X \sim P$, $\min P(X \geq \theta_0), P(X \leq \theta_0) \geq \frac{1}{2}$. $P(X = \theta_0) = p_0 > 0$ so med(X)=θ_0. Set $l(\theta, X) := |\theta - X|, R(\theta) = \mathbb{E}[|x-\theta|]$. Then $\theta_0 = \arg\min_\theta R(\theta)$ =med(x), and $R(\theta) \geq R(\theta_0) + p_0|\theta-\theta_0|$. Then, $\theta_n = \arg\hat{\min}_\theta R_n(\theta) = \frac{1}{n}\sum_{i=1}^n |X_i - \theta|$ satisfies $\hat{\theta_n} - \theta_0 = O_p(n^{-k})$ for any $k < \infty$. (i.e., we want to see how much the absolute value wiggles around a neighborhood of size δ)*

New Material *Convergence in distribution in metric spaces and uniform versions of convergence in distribution. Recall that $X_N \xrightarrow{d} X$ if, and only if,*

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$$

for all bounded continuous f. (This is the definition whether X_n and X are real-valued or valued in any other space.)

Definition 0.1. *Let $\mathbb{D}$ be a metric space. A random variable $X : \Omega \to \mathbb{D}$ is **tight** if, $\forall \varepsilon > 0$, $\exists K \subset \mathbb{D}$ compact such that*

$$\mathbb{P}[X \in K] \geq 1 - \varepsilon. \tag{1}$$

Definition 0.2. *A sequence of $\mathbb{D}$-valued random variable, $\{X_n\}_{n\in\mathbb{N}}$ is **asympototically tight** if, forall $\varepsilon > 0$, $\exists K \subset \mathbb{D}$ compact such that*

$$\limsup_{n\to\infty} \mathbb{P}[X_n \notin K^\delta] < \varepsilon$$

for all $\delta > 0$ where $K^\delta := \{y \in \mathbb{D} : \operatorname{dist}(y, K) < \delta\}$.

Theorem 2 (Pohorov)**.** *Let $X_N : \Omega_n \to \mathbb{D}$.*

- *If $X_N \xrightarrow{d} X$ where X is tight, then X_n is asympototically tight.*
- *If X_N is asymptotically tight, then there exists a subsequence $\{X_n\}$ and tight X such that $X_{n_k} \xrightarrow{d} X$ as $k \to \infty$.*

Example 2: *Continuous functions on compact sets. Let (T, d) be a compact metric space. Let*

$$L^\infty(T) := \left\{ f : T \to \mathbb{R} \,:\, f \text{ measureable}, \sup_T f(x) < \infty \right\}.$$

Take $\mathbb{D} = L^\infty(T)$. Let $\ell : T \times X \to \mathbb{R}$ be continuous and $X_i \overset{iid}{\sim} \mathbb{P}$, $X_i \in X$.

Define $Z_n : T \to \mathbb{R}$ by

$$Z_n(\cdot) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \ell(\cdot, X_i) - \mathbb{P}\ell(\cdot, X_i) = \sqrt{n}\,(\mathbb{P}_n - \mathbb{P})]\ell(\cdot, X).$$

Then Z_n is a continuous function of its argument $t \in T$, and so $Z_n \in L^\infty(T)$. (Note that if $t \subset T$ is countable and dense, then Z_n is determined completely by T_0.) Let $t_1, \cdots, t_k \in T$. Then

$$\Big(Z_n(t_1), \cdots Z_n(t_k)\Big) \xrightarrow{d} N(0, cov(\ell(t_i, x), \ell(t_j, x))).$$

Big question: is there a random version of this, i.e. some random function $Z \in L^\infty(T)$ s.t. $\{Z_n(t)\}_{t\in T} \xrightarrow{d} Z$? ♣ **Compactness in functional space:**

Evidently, if $X_n \in L_\infty(T)$, we must understand compactness in $L_\infty(T)$. For us, limits will be in $\mathcal{C}(T, R)$. $\mathcal{C}(T, R)$ =continuous function. $f : T \to \mathbb{R}$, $||f - g||_\infty$ is metric.

Standard compactness theorem: Arzela-Ascoli theorem

Definition 0.3. *For a function $f : T \to \mathbb{R}$, modulus of continuity is*

$$w_f(\delta) = \sup_{d(s,t)<\delta} \{|f(t) - f(s)|\}$$

Definition 0.4. *A collection $\mathcal{F}$ is uniformly equicontinuous if*

$$\lim_{s\downarrow 0}\sup_{f\in\mathcal{F}} w_f(\delta) = 0$$

Equivalently, $\forall\epsilon > 0$, $\exists\delta > 0$ s.t. if $d(s,t) < \delta$, then $|f(s) - f(t)| \leq \delta$ for all $f \in \mathcal{F}$.

Theorem 3 (Arzela-Ascoli)**.** *Let (T, d) be a compact metric space. Then:*

1. *$\mathcal{F} \subset \mathcal{C}(T, \mathbb{R})$ is relatively compact in the sup norm. That is, $Cl(\mathcal{F})$ (Closure of $\mathcal{F}$) is compact*
2. *$\mathcal{F}$ is uniformly equicontinuous and $\exists t_0 \in T$, s.t. $\sup_{f\in\mathcal{F}} |f(t_0)| < \infty$*

are equivalent.

In stochastic case, we will look into the stochastic analogue of equicontinuity to show asymptotic rightness.

Definition 0.5. *Let $X_n \in L_\infty(T)$ be random variables. We say that X_n are* asymptotically equicontinuous *if for all $\eta, \epsilon > 0$, there is a finite partition $T_1, \ldots, T_k$ of T such that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\max_i \sup_{s,t\in T_i} |X_{n,s} - X_{n,t}| \geq \epsilon\right) \leq \eta.$$

Recall that we have a sequence Ω_n of sample spaces, with $X_n \colon \Omega_n \to L_\infty(T)$, and $X_{n,t}(\omega)$ is the value of $X_n(\omega)$ at t.

This definition is saying that for any fix amount of separation, we can divide T into little blocks so that the process is stochastically continuous within each block.

Example 3: *Let $Z_i \in \mathbb{R}^d$ with $Z_i \overset{\text{iid}}{\sim} P$. Assume that $\mathbb{E}[\|Z_i\|^2] < \infty$ and $\mathbb{E}[Z_i] = 0$. Define $X_{n,t} := \frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i^T t$ for $t \in T \subseteq \mathbb{R}^d$ (with $T = \{t \in R^d : \|t\|_2 \leq M\}$). Hence for any $\delta > 0$,*

$$\mathbb{P}\left(\sup_{\|t-s\|\leq\delta} |X_{n,t} - X_{n,s}| \geq \epsilon\right) = \mathbb{P}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_i\right\| \delta \geq \epsilon\right) \leq \frac{\mathbb{E}\|Z_i\|^2 \delta^2}{\epsilon^2} \quad \text{by Chebyshev.}$$

Hence if we choose δ small enough, we get $\frac{\delta^2\mathbb{E}\|Z_i\|^2}{\epsilon^2} \leq \eta$ for any desired $\eta > 0$.

As T is compact, we can use a finite number of δ balls covering T. Hence X_n is a sequence of asymptotically equicontinuous random variables. ♣

With stochastic equicontinuity, we put down finite balls centered at t^i in our set T. Somehow if $X_{n,t} \approx X_{n,t^i}$ for some t^i, then if X_{n,t^i} converges in distribution to something, $X_{n,t}$ should too.

Theorem 4 (Weak convergence in $L^\infty(T)$)**.** *The following are equivalent:*

(i) *$X_i \in L_\infty(T)$ and $X_n \xrightarrow{d} X \in L_\infty(T)$ where X is tight.*

(ii) *(a) Finite dimensional convergence (FIDI): $(X_{n,t_1}, \ldots, X_{n,t_k}) \xrightarrow{d} Z(t_1, \ldots t_k)$ (something) for any $t_1, \ldots, t_k \in T$, and $k < \infty$.*

(b) X_n are (asymptotically) stochastically equicontinuous.

Proof *We only care about (ii) $\implies$ (i).*

Part 1: Construct $T_0 \subset T$ a countable dense subset and work there. *Let $m \in \mathbb{N}$, construct partitions $T_1^m, \ldots, T_{k_m}^m$ of T such that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\max_i \sup_{s,t\in T_i^m} |X_{n,s} - X_{n,t}| \geq 2^{-m} \right) \leq 2^{-m}.$$

Assume without loss of generality that partitions are nested.

For each m, define

$$\rho_m(s,t) = \begin{cases} 0 & \text{if } s,t \in T_i^m \text{ for some } i \\ 1 & \text{otherwise.} \end{cases}$$

Define $\rho(s,t) := \sum_{m=1}^{\infty} 2^{-m}\rho_m(s,t)$. So then $\rho-\mathrm{diam}(T_i^m) \leq \sum_{i=1}^{m} 0 + \sum_{i=m+1}^{\infty} 2^i = 2^{-m}$. Let $t_{i,m} \in T_i^m$ (arbitrary) and define

$$T_0 := \bigcup_{m=1}^{\infty} \bigcup_{i=1}^{k_m} \{t_{i,m}\}$$

then T_0 is countable and dense in T for ρ-metric. Moreover, T is totally bounded in the ρ-metric.

Part 2: Use metric ρ to extend process to $\mathcal{C}(T, \mathbb{R})$.

Part 3: Show that $X_n \xrightarrow{d} X$ is the space $L_\infty(T)$. *Use equicontinuity to approximate X_n, X with finite partitions* □