Stats 300b: Theory of Statistics	Winter 2017

Lecture 14 – February 23

Lecturer: John Duchi — *Scribe: Tri Dao*

Warning: *these notes may contain factual errors*

Reading: None.

Outline: Uniform limits in distribution:

- Stochastic analog of equicontinuity.
- Moduli of continuity of processes

1 Recap: Compactness in functions spaces

Last time we talked about compactness in $\mathcal{C}(T,\mathbb{R})$ (i.e., continuous functions from $T \to \mathbb{R}$).

We defined *modulus of continuity*: $\omega_f(\delta) := \sup_{d(s,t)<\delta}\{|f(t)-f(s)|\}$.

We also stated (but didn't prove) Arzela–Ascoli theorem: $\mathcal{F} \subseteq \mathcal{C}(T,\mathbb{R})$ is compact if and only if:

(i) $\sup_{f\in\mathcal{F}} |f(t_0)| < \infty$ for some t_0.

(ii) $\lim_{\delta\downarrow 0} \sup_{f\in\mathcal{F}} \omega_f(\delta) = 0$ (equicontinuity)

(iii) $\mathcal{F}$ is closed.

How do we get analog of (ii) for stochastic processes, so that sequences $X_n \in L_\infty(T)$ take values in compact sets and have limits?

2 Stochastic analog of equicontinuity

Definition 2.1. *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 1: Let $Z_i \in \mathbb{R}^d$ with $Z_i \stackrel{\text{iid}}{\sim} P$. 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 compact). 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 have partition of T into finite blocks of radius less than δ. 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 1. *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):* $(X_{n,t_1}, \ldots, X_{n,t_k}) \xrightarrow{d}$ *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 $\rho-\text{diam}(T_i^m) \leq \sum_{i=1}^m 0 + \sum_{i=m+1}^\infty 2^i = 2^{-m}$. 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 $\mathcal{C}(T,\mathbb{R})$. By Kolmogorov's extension theorem, there is *some* stochastic process $\{X_t\}_{t\in T_0}$ such that

$$(X_{n,t_1}, \ldots, X_{n,t_k}) \xrightarrow{d} (X_{t_1}, \ldots, X_{t_k}) \quad \text{for all } t_i \in T_0.$$

The standard result on convergence in distribution: for all $S \subset T_0$ with $\text{card}(S) < \infty$, we have

$$2^{-m} \geq \limsup_{n\to\infty} \mathbb{P}\left(\max_i \max_{\substack{s,t\in T_i^m \\ s,t\in S}} |X_{n,t} - X_{n,s}| \geq 2^{-m}\right)$$
$$\geq \mathbb{P}\left(\max_i \max_{\substack{s,t\in T_i^m \\ s,t\in S}} |X_t - X_s| > 2^{-m}\right),$$

where we use the fact that max of a finite number of continuous functions is continuous. But $\mathbb{P}\left(\max_i \max_{\substack{s,t\in T_i^m \\ s,t\in S}} |X_t - X_s| > 2^{-m}\right) \uparrow \mathbb{P}\left(\max_i \sup_{\substack{s,t\in T_i^m \\ s,t\in T_0}} |X_t - X_s| > 2^{-m}\right)$ (monotone convergence). Hence

$$2^{-m} \geq \mathbb{P}\left(\sup_{\substack{\rho(s,t)<2^{-m} \\ s,t\in T_0}} |X_t - X_s| > 2^{-m}\right)$$

because $\rho(s,t) < 2^{-m} \implies \rho_m(s,t) < 1 \Leftrightarrow \rho_m(s,t) = 0$. That is, $s,t \in T_i^m$. By Borel–Cantelli, we have

$$\sum_{m=1}^{\infty} \mathbb{P}\left(\sup_{\substack{\rho(s,t)<2^{-m} \\ s,t\in T_0}} |X_t - X_s| > 2^{-m}\right) \leq \sum_{m=1}^{\infty} 2^{-m} = 1.$$

Thus as $n \to \infty$, $|X_s - X_t| > 2^{-m}$ for some $s,t \in T_0$ such that $\rho(s,t) < 2^{-m}$ happens only *finitely often*. That is, with probability 1, $\sup_{\rho(s,t)<2^{-m}} |X_s - X_t| \leq 2^{-m}$ eventually (for large m).

We can immediately extend $\{X_t\}_{t\in T_0}$ to $\{X_t\}_{t\in T}$ by continuity, and so $\{X_t\}_{t\in T} \in \mathcal{C}_\rho(T, \mathbb{R})$.

Part 3: Show that $X_n \xrightarrow{d} X$ is the space $L_\infty(T)$. This is equivalent to showing that for any Lipschitz function $f\colon L_\infty(T) \to [-1,1]$ (i.e., $|f(x) - f(z)| \leq \sup_{t\in T} |x(t) - z(t)|$ for $x, z \in L_\infty(T)$),

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] \quad \text{as } n \to \infty.$$

Idea: approximate by finer and finer partitions of T.

Define $\pi_m(t) = t_{i,m}$ where $t \in T_i^m$ (our partitions) then $\rho(\pi_m(t), t) \leq 2^{-m}$. We showed that X_t is uniformly continuous in the ρ-metric, so $(X \circ \pi_m)_t := X_{\pi_m(t)} \to X_t$ as $n \to \infty$ (in sup-norm).

We know that $X_n \circ \pi_m$ can only take values $X_{n,\pi_m(t)} = X_{n,t_{i,m}}$ for $i = 1, \ldots, k_m$. We then have that $X_n \circ \pi_m \xrightarrow{d} X \circ \pi_m$ by the FIDI assumption. Thus

$$\lim_{m\to\infty} \lim_{n\to\infty} \mathbb{E}\left[f(X_n \circ \pi_m)\right] = \lim_{m\to\infty} \mathbb{E}[f(X \circ \pi_m)] = \mathbb{E}[f(X)]$$

for any bounded Lipschitz $f\colon L_\infty(T) \to \mathbb{R}$. Hence

$$|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leq |\mathbb{E}[f(X_n \circ \pi_m) - f(X_n)]| + |\mathbb{E}[f(X_n \circ \pi_m) - f(X \circ \pi_m)]| + |\mathbb{E}[f(X \circ \pi_m) - f(X)]|.$$

Notice that the second term is $o(1)$ and the third term goes to 0 as $m \to \infty$. We need to bound the first term:

$$\begin{aligned} |\mathbb{E}f(X_n) - \mathbb{E}f(X_n \circ \pi_m)| &\leq \|f\|_{\text{Lip}}\, \epsilon + \mathbb{P}\left(\sup_{t\in T} \left|X_{n,t} - X_{n,\pi_m(t)}\right| \geq \epsilon\right) \\ &\leq \|f\|_{\text{Lip}}\, \epsilon + \mathbb{P}\left(\sup_{\rho(s,t)<2^{-m}} |X_{n,s} - X_{n,t}| \geq \epsilon\right). \end{aligned}$$

The second term goes to 0 as $m \to \infty$ by stochastic equicontinuity. □