Notes on Weak Convergence Theory

Chang Feng

April 22, 2025

1 Measures on Metric Spaces

First a notation: Let P be probability measure on $(S, \mathcal{S})$, f be a function on S, write

$$Pf = \int_S f(x) P(dx).$$

Definition 1.1 (Weak Convergence of Probability Measures)**.** We say P_n weakly converge to P ($P_n \Rightarrow P$) if $P_n f \to Pf$ for every bounded continuous real function f on S.

Definition 1.2 (Tightness)**.** A probability measure P on $(S, \mathcal{S})$ is tight if

$$\forall \epsilon > 0, \exists K \text{ compact s.t. } P(K) > 1 - \epsilon$$

Definition 1.3 (Separating Class)**.** A class $\mathcal{A} \subset \mathcal{S}$ is called a separating class if for any probability measures P and Q,

$$P(A) = Q(A), \forall A \in \mathcal{A} \implies P(M) = Q(M), \forall M \in \mathcal{S}.$$

That is if P and Q agree on $\mathcal{A}$, then they agree on $\mathcal{S}$.

Recall a π-system means closed under finite intersection.

Proposition 1.4. *If $\mathcal{A}$ is a π-system generating $\mathcal{S}$, then it is a separating class for $\mathcal{S}$.*

Theorem 1.5. *If S is separable and complete, then every probability measure on $(S, \mathcal{S})$ is tight.*

Proof. Since S is separable (have a countable dense subset), there is for each k, a sequence $A_{k1}, A_{k2}, \ldots$ of open $1/k$-balls covering S. Choose n_k large enough that

$$P\left(\bigcup_{i \le n_k} A_{ki}\right) > 1 - \frac{\epsilon}{2^k}$$

Considr the set $\bigcap_{k \ge 1} \bigcup_{i \le n_k} A_{ki}$. This is totally bounded because inside the intersection, it is a bunch of finite union of $1/k$-balls. By completeness, this totally bounded set has compact closure K.

Clearly $P(K^c) \le \sum_{i=1}^{\infty} \epsilon/2^k = \epsilon$, so $P(K) > 1 - \epsilon$. $\square$

2 Prohorov Metric

Let (S, d) be metric space. $\mathcal{B}(S)$ be the borel σ-algebra and $\mathcal{P}(S)$ be the family of all borel-probability measures on S. Turns out we can make $\mathcal{P}(S)$ into a metric space with the Prohorov metric.

Definition 2.1 (Prohorov Metric)**.** Let $\mathcal{C}$ be collection of all closed subsets of S.

$$\rho(P, Q) = \inf\{\epsilon > 0 : P(F) \leq Q(F^\epsilon) + \epsilon \text{ for all } F \in \mathcal{C}\}.$$

where

$$F^\epsilon = \{x \in S : \inf_{y \in F} d(x, y) < \epsilon\}$$

To see this is indeed a metric, see [EK Ch3.1 p96].

Proposition 2.2 (Probabilistic Interpretation)**.**

$$\rho(P, Q) = \inf_{\mu \in \mathcal{M}(P,Q)} \left\{ \inf \left\{ \epsilon > 0 : \mu(d(x, y) \geq \epsilon) \leq \epsilon \right\} \right\}$$

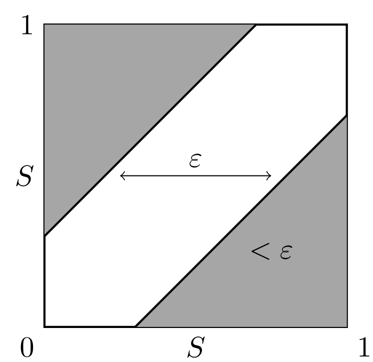

Figure 1: visualization for $S = [0, 1]$

Theorem 2.3.

(i) If S separable, then $\mathcal{P}(S)$ is separable.

(ii) If (S, d) separable and complete, then $(\mathcal{P}(S), \rho)$ is separable and complete.

One of the main result from all this construction is the following.

Theorem 2.4 (Skorokhod Representation Theorem)**.** *Let (S, d) be separable. Suppose $(P_n)_{n\in\mathbb{N}} \subset \mathcal{P}(S)$, $P \in \mathcal{P}(S)$, such that $\rho(P_n, P) \to 0$.*
Then there exist some probability space $(\Omega, \mathcal{F}, v)$ on which S-valued R.V.'s $(X_n)_{n\in\mathbb{N}}$ and X lives, with distributions $(P_n)_{n\in\mathbb{N}}$ and P respectively, such that

$$X_n \to X \quad a.s.$$

Theorem 2.5 (Portmanteau Theorem)**.** *These five conditions are equivalent:*

(i) $P_n \Rightarrow P$.

(ii) $P_n f \to P f$ for all bounded, uniformly continuous f.

(iii) $\limsup_n P_n F \leq PF$ for all closed F.

(iv) $\liminf_n P_n G \geq PG$ for all open G.

(v) $P_n A \to PA$ for all P-continuity sets A.

2.1 Prohorov's theorem

Definition 2.6 (Relative Compactness)**.** Let Π be a family of probability measures on $(S, \mathcal{S})$. Say Π is relatively compact if every sequence $\{P_n\} \subset \Pi$ contains a weakly converging subsequence.

Example 2.7. Consider space $(C, \mathcal{C})$. Suppose finite dimensional projections of P_n weakly converge to that of P, i.e. $P_n \pi^{-1}_{t_1,\ldots,t_k} \Rightarrow P\pi^{-1}_{t_1,\ldots,t_k}$. It's not necessarity true that $P_n \Rightarrow P$. We also need relative compactness of $\{P_n\}$!

Proof sketch: relative compactness implies subsequence $P_{n_i} \Rightarrow Q$ for some probability measure Q. By continuous mapping theorem, $P_{n_i}\pi^{-1}_{t_1,\ldots,t_k} \Rightarrow Q\pi^{-1}_{t_1,\ldots,t_k}$. Therefore, the finite dimensional projections of P and Q must agree. Since finite dimensional projections is a separating class, this implies $P = Q$.

Now suppose $\{P_n\}$ is relatively compact and we only know that the finite dimensional projections of P_n converge weakly to some measure $\mu_{t_1,\ldots,t_k}$ on $(\mathbb{R}^k, \mathcal{R}^k)$. By similar arguments, we can conclude that there exist some probability measure P such that $P_n \Rightarrow P$. ❖

Now, how do you show relatively compactness?

Theorem 2.8. *If Π is tight, then it is relatively compact.*

Proposition 2.9. *If $(S, \mathcal{S})$ is complete and separable, then tightness $\iff$ relative compactness*

3 Space C

Space $C = C[0,1]$ is the space of continous function on interval $[0,1]$. We equip it with the uniform topology, induced by the metric

$$\rho(x,y) = \|x-y\| = \sup_{t\in[0,1]} |x(t)-y(t)|$$

3.1 Weak Convergence and Tightness in C

Theorem 3.1. *Let P_n, P be probability measures on $(C,\mathcal{C})$. If the finite-dimensional distributions of P_n converge weakly to those of P, and if $\{P_n\}$ is tight, then $P_n \Rightarrow P$.*

Definition 3.2 (Modulus of Continuity)**.**

$$w_x(\delta) = w(x,\delta) = \sup_{|s-t|\le\delta} |x(s)-x(t)|, \quad 0<\delta\le 1.$$

A function $x(\cdot)$ is uniformly continuous if and only if $\lim_{\delta\to 0} w_x(\delta) = 0$.

3.2 Maximal Inequalitites

Let $\xi_1,\dots,\xi_n$ be random variables (don't need to be stationary nor independent), let $S_k = \xi_1 + \dots + \xi_k$ ($S_0 = 0$), and put

$$M_n = \max_{k\le n} |S_k|.$$

Let $m_{ijk} = |S_j - S_i| \wedge |S_k - S_j|$ and denote

$$L_n = \max_{0\le i\le j\le k\le n} m_{ijk}.$$

We derive upper bounds for $P[M_n \ge \lambda]$ using the follwing two inequalities:

$$M_n \le L_n + |S_n|. \tag{3.1}$$

$$M_n \le 3L_n + \max_{k\le n} |\xi_k|. \tag{3.2}$$

From these we can bound M_n by having a bound on L_n and S_n or $\max_{k\le n}|\xi_k|$.

Derivations of the inequalities:

From $|S_k| \le |S_n - S_k| + |S_n|$ and $|S_k| \le |S_k| + |S_n|$ follows $|S_k| \le \min\{|S_k|, |S_n - S_k|\} + |S_n| = m_{0kn} + |S_n|$, which gives the first inequality

$$M_n \le L_n + |S_n|.$$

A useful claim:

$$|S_n| \le 2L_n + \max_{k\le n} |\xi_k|.$$

proof of claim:

- Case when $|S_n| = 0$: Trivially true .

- Case when $|S_n| > 0$: Observe $|S_0| = 0 < |S_n - S_0| = |S_n|$, but $|S_n| \geq |S_n - S_n| = 0$. Therefore there exist some k, $1 \leq k \leq n$ such that

$$|S_k| \geq |S_n - S_k| \quad \text{but} \quad |S_{k-1}| < |S_n - S_{k-1}|$$

For this k,

$$|S_n - S_k| = m_{0kn} \leq L_n \quad \text{and} \quad |S_{k-1}| = m_{0,k-1,n} \leq L_n$$

Therefore

$$|S_n| \leq |S_{k-1}| + |\xi_k| + |S_n - S_k| \leq 2L_n + |\xi_k|$$

Then we have the 2nd inequality:

$$M_n \leq 3L_n + \max_{k \leq n} |\xi_k|.$$

Theorem 3.3. *Suppose that $\alpha > \frac{1}{2}$ and $\beta \geq 0$ and that $u_1, \ldots, u_n$ are nonnegative numbers such that*

$$P[m_{ijk} \geq \lambda] \leq \frac{1}{\lambda^{4\beta}} \left(\sum_{i<l\leq k} u_l \right)^{2\alpha}, \quad 0 \leq i \leq j \leq k \leq n,$$

for $\lambda > 0$. Then

$$P[L_n \geq \lambda] \leq \frac{K}{\lambda^{4\beta}} \left(\sum_{0<l\leq n} u_l \right)^{2\alpha}$$

for $\lambda > 0$, where $K = K_{\alpha,\beta}$ depends only on α and β.

Theorem 3.4. *Suppose that $\alpha > \frac{1}{2}$ and $\beta \geq 0$ and that $u_1, \ldots, u_n$ are nonnegative numbers such that*

$$\mathbb{P}\left[|S_j - S_i| \geq \lambda\right] \leq \frac{1}{\lambda^{4\beta}} \left(\sum_{i<l\leq j} u_l \right)^{2\alpha}, \quad 0 \leq i \leq j \leq n,$$

for $\lambda > 0$. Then

$$\mathbb{P}\left[M_n \geq \lambda\right] \leq \frac{K'}{\lambda^{4\beta}} \left(\sum_{0<l\leq n} u_l \right)^{2\alpha}$$

for $\lambda > 0$, where $K' = K'_{\alpha,\beta}$ depends only on α and β.