Reading Group: Probability With Martingales Ch13

LEUNG Man Fung, Heman

Summer 2020

Uniform integrability

Motivation

- Convergence in probability is easy to establish, e.g.
  - WLLN for independent RVs
  - Ergodic theorem for dependent RVs (discussed last semester in recursive TAVC)
  - Dominated convergence theorem
- Convergence in $\mathcal{L}^p$-norm is harder to establish on the other hand
- Uniform integrability is a necessary and sufficient condition to link them

An “absolute continuity” property

- Lemma 13.1.1
  - Suppose that $X \in \mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$
  - Then, given $\epsilon > 0$, $\exists \delta > 0$ s.t. for $F \in \mathcal{F}$, $P(F) < \delta \implies E(|X|; F) < \epsilon$
- Proof
  - If the conclusion is false, then, for some $\epsilon_0 > 0$, we can find $\{F_n\}$ consists of elements of $\mathcal{F}$ s.t.

$$P(F_n) < 2^{-n}, E(|X|; F_n) \geq \epsilon_0$$

    - Construction of “contracting” events
  - Let $H := \limsup F_n$. Then BC1 shows that $P(H) = 0$
  - Yet reverse Fatou lemma shows that $E(|X|; H) \geq \limsup_{n\to\infty} E(|X|; F_n) = \epsilon_0$
  - Contradiction arises since $P(H) = 0 \implies E(|X|; H) = 0$

An “absolute continuity” property

- Corollary 13.1.2
  - Suppose that $X \in \mathcal{L}^1$ and that $\epsilon > 0$
  - Then $\exists K \in [0, \infty)$ such that $E(|X|; |X| > K) < \epsilon$
- Proof
  - Let δ be as in lemma 13.1.1
  - Since $KP(|X| > K) \leq E(|X|)$, we can choose K such that $P(|X| > K) \leq \delta$
  - Application of lemma 13.1.1 yields the result

UI family

- A class $\mathcal{C}$ of RVs is called uniformly integrable (UI) if given $\epsilon > 0$,

$$\exists K \in [0, \infty) \text{ s.t. } E(|X|; |X| > K) < \epsilon, \forall X \in \mathcal{C}$$

- For such a class $\mathcal{C}$, we have (with K_1 relating to $\epsilon = 1$) for every $X \in \mathcal{C}$,

$$\begin{aligned} E(|X|) &= E(|X|; |X| > K_1) + E(|X|; |X| \leq K_1) \\ &\leq 1 + K_1 \end{aligned}$$

  - The first term comes from choice of K_1 and corollary 13.1.2
  - The second term comes from idea of Markov's inequality
- This means that a UI family is bounded in $\mathcal{L}^1$ but the converse is not true
  - Counterexample: Take $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1], \mathcal{B}[0,1], \text{Leb})$
  - Let $E_n = \left(0, \frac{1}{n}\right)$ and $X_n = nI_{E_n}$
  - Then $E(|X_n|) = 1, \forall n$ so that $\{X_n\}$ is bounded in $\mathcal{L}^1$
  - However, for any $K > 0$, we have for $n > K$, $E(|X_n|; |X_n| > K) = nP(E_n) = 1$
  - This means $\{X_n\}$ is not UI. Here, $X_n \to 0$ but $E(X_n) \nrightarrow 0$

Two sufficient conditions for the UI property

- First condition: boundedness in $\mathcal{L}^p$ where $p > 1$
  - Suppose that $\mathcal{C}$ is a class of RVs bounded in $\mathcal{L}^p$ for some $p > 1$
  - Thus, for some $A \in [0, \infty)$, $E(|X|^p) < A, \forall X \in \mathcal{C}$
  - Then $\mathcal{C}$ is UI
- Proof
  - If $v \geq K > 0$, then $v^{1-p} \leq K^{1-p} \implies v \leq K^{1-p} v^p$
  - Hence, for $K > 0$ and $X \in \mathcal{C}$, we have

$$E(|X|; |X| > K) \leq K^{1-p} E(|X|^p; |X| > K) \leq K^{1-p} A$$

  - The result follows from the fact that we can choose K based on the value of $\epsilon := K^{1-p} A$
- Idea
  - Boundedness in $\mathcal{L}^p$ for some $p > 1$ implies boundedness in $\mathcal{L}^1$
    - Which is a property of UI family
    - While $\mathcal{L}^p$ provides a “faster” convergence

Two sufficient conditions for the UI property

- Second condition: dominated by an integrable non-negative variable
  - Suppose that $\mathcal{C}$ is a class of RVs which is dominated by an integrable non-negative variable Y:

$$|X(\omega)| \leq Y(\omega), \forall X \in \mathcal{C} \text{ and } E(Y) < \infty$$

  - Then $\mathcal{C}$ is UI
- Proof
  - For $K > 0$ and $X \in \mathcal{C}$, we have

$$E(|X|; |X| > K) \leq E(Y; Y > K) < \epsilon$$

  - where the last inequality comes from corollary 13.1.2
- Remark
  - It is precisely this which makes dominated convergence theorem works for our $(\Omega, \mathcal{F}, \mathbb{P})$
  - An extension of dominated convergence theorem to the whole class $\mathcal{C}$

UI property of conditional expectation

- Theorem 13.4.1
  - Let $X \in \mathcal{L}^1$. Then the class $\{E(X|\mathcal{G}) : \mathcal{G} \text{ a sub-}\sigma\text{-algebra of } \mathcal{F}\}$ is uniformly integrable
  - Formally, the definition of the class $\mathcal{C}$ is $Y \in \mathcal{C}$ if and only if Y is a version of $E(X|\mathcal{G})$ for some sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$
- Proof
  - Let $\epsilon > 0$ be given
  - By lemma 13.1.1, we can choose $\delta > 0$ such that, for $F \in \mathcal{F}$, $P(F) < \delta \implies E(|X|; F) < \epsilon$
  - Choose K so that $K^{-1}E(|X|) < \delta$
  - Now let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let Y be any version of $E(X|\mathcal{G})$
  - By Jensen's inequality, $|Y| \leq E(|X||\mathcal{G})$ a.s. (absolute function is convex)
  - Hence $E(|Y|) \leq E(|X|)$ by tower property and $KP(|Y| > K) \leq E(|Y|) \leq E(|X|)$
  - By the choice of K, we now have $P(|Y| > K) < \delta$ from last inequality
  - But $\{|Y| > K\} \in \mathcal{G}$, so that $E(|Y|; |Y| \geq K) \leq E(|X|; |Y| \geq K) < \epsilon$ completes the proof
    - By $|Y| \leq E(|X||\mathcal{G})$, property of conditional expectation and lemma 13.1.1

Convergence of random variables

Convergence in probability

- Definition
  - Let $\{X_n\}$ be a sequence of RVs and X be a RV
  - We say that $X_n \overset{p}{\to} X$ if for every $\epsilon > 0$

$$\lim_{n\to\infty} P(|X_n - X| > \epsilon) \to 0$$

- Lemma 13.5.1: almost sure convergence implies convergence in probability
  - $X_n \overset{a.s.}{\to} X \implies X_n \overset{p}{\to} X$
- Proof
  - Suppose that $X_n \overset{a.s.}{\to} X$ and that $\epsilon > 0$
  - Then by reverse Fatou lemma for sets,

$$\begin{aligned} 0 &= P(|X_n - X| > \epsilon,\ \text{i.o.}) = P\left(\limsup\{|X_n - X| > \epsilon\}\right) \\ &\geq \limsup P(|X_n - X| > \epsilon) \end{aligned}$$

  - The result is proved by non-negativity of probability and sandwich theorem

Bounded convergence theorem

- Let $\{X_n\}$ be a sequence of RVs and X be a RV
- Suppose that $X_n \xrightarrow{p} X$ and that for some $K \in [0, \infty)$, we have $|X_n(\omega)| \leq K, \forall n, \forall \omega$
- Then $E(|X_n - X|) \to 0$

- Proof
  - Let's check that $P(|X| \leq K) = 1$. By assumption, for $k \in \mathbb{N}$,

$$P(|X| > K + k^{-1}) \leq P(|X - X_n| > k^{-1}), \forall n$$

  - $X_n \xrightarrow{p} X$ implies $P(|X| > K + k^{-1}) = 0$
  - Hence $P(|X| > K) = P\left(\cup_k \{|X| > K + k^{-1}\}\right) = 0$
  - Now let $\epsilon > 0$ be given
  - Choose n_0 such that $P\left(|X_n - X| > \frac{1}{3}\epsilon\right) < \frac{\epsilon}{3K}$ when $n \geq n_0$
  - Then, for $n \geq n_0$,

$$E(|X_n - X|) = E\left(|X_n - X|; |X_n - X| > \frac{1}{3}\epsilon\right) + E\left(|X_n - X|; |X_n - X| \leq \frac{1}{3}\epsilon\right)$$
$$\leq 2KP\left(|X_n - X| > \frac{1}{3}\epsilon\right) + \frac{1}{3}\epsilon \leq \epsilon$$

- Remark
  - This proof shows that convergence in probability is a natural concept (how?)

A necessary and sufficient condition for $\mathcal{L}^1$ convergence

- Theorem 13.7.1
  - Let $\{X_n\}$ be a sequence in $\mathcal{L}^1$ and let $X \in \mathcal{L}^1$
  - Then $X_n \xrightarrow{\mathcal{L}^1} X$, equivalently $E(|X_n - X|) \to 0$, if and only if $X_n \xrightarrow{p} X$ and $\{X_n\}$ is UI
- Remarks
  - The "if" part is more useful since it improves dominated convergence theorem
    - This can be seen from 13.3 the second sufficient condition of UI
  - The "only if" part is less surprising
    - Convergence in $\mathcal{L}^p, p \geq 1$ implies convergence in probability

- Proof of “if” part
  - Suppose that $X_n \xrightarrow{p} X$ and $\{X_n\}$ is UI. For $K \in [0, \infty)$, define $\varphi_K : \mathbb{R} \to [-K, K]$ by

$$\varphi_K(x) := \begin{cases} K & , x > K \\ x & , |x| \le K \\ -K & , x < -K \end{cases}$$

  - Let $\epsilon > 0$ be given. By the UI property of $\{X_n\}$ and corollary 13.1.2, choose K so that

$$E\big[|\varphi_K(X_n) - X_n|\big] < \frac{\epsilon}{3}, \forall n; E\big[|\varphi_K(X) - X|\big] < \frac{\epsilon}{3}$$

  - Note that $|\varphi_K(x) - \varphi_K(y)| \le |x - y| \implies \varphi_K(x) \xrightarrow{p} \varphi_K(y)$ by taking probability
  - Applying bounded convergence theorem, we can choose n_0 such that, for $n \ge n_0$,

$$E\big[|\varphi_K(X_n) - \varphi_K(X)|\big] < \frac{\epsilon}{3}$$

  - Minkowski inequality shows that, for $n \ge n_0$,

$$E\big(|X_n - X|\big) = E\big[|X_n - \varphi_K(X_n) + \varphi_K(X) - X + \varphi_K(X_n) - \varphi_K(X)|\big] < \epsilon$$

- Proof of “only if” part
  - Suppose that $X_n \to X$ in $\mathcal{L}^1$. Let $\epsilon > 0$ be given
  - Choose N such that $n \geq N \implies E(|X_n - X|) < \frac{\epsilon}{2}$
  - By lemma 13.1.1, we can choose $\delta > 0$ such that whenever $P(F) < \delta$, we have

$$E(|X_n|; F) < \epsilon, 1 \leq n \leq N; \quad E(|X|; F) < \frac{\epsilon}{2}$$

    - The second inequality probably comes from choice of N instead of lemma 13.1.1
  - Since $\{X_n\}$ is bounded in $\mathcal{L}^1$, we can choose K such that $K^{-1} \sup_r E(|X_r|) < \delta$
  - Then for $n \geq N$, we have $P(|X_n| > K) < \delta$ (by idea in Markov inequality) and

$$E(|X_n|; |X_n| > K) \leq E(|X|; |X_n| > K) + E(|X - X_n|) < \epsilon$$

    - By lemma 13.1.1 and choice of N
  - For $n \leq N$, we have $P(|X_n| > K) < \delta$ and $E(|X_n|; |X_n| > K) < \epsilon$ by choice of δ
  - Hence $\{X_n\}$ is a UI family
  - Since $\epsilon P(|X_n - X| > \epsilon) \leq E(|X_n - X|) \to 0$, we have $X_n \xrightarrow{p} X$

Concluding remarks

Comments

- UI allows us to establish stronger $\mathcal{L}^1$ convergence from weaker convergence in probability
    - This is appealing as there are more standard devices for convergence in probability
- UI appears naturally in conditional expectation, which is central to martingale property
    - Thus UI martingale is studied in next chapter