9.8 Uniform integrability of a sequence of r.v.'s

- We prepare this section, especially, later to extend several results for discrete time martingales to continuous time martingales by taking limits (e.g. Doob's optional sampling theorem, Doob-Meyer decomposition).
- Even if a sequence of real-valued r.v.'s $(X_n)_{n=1,2,\dots}$ converges to X in a.s.-sense,

$$X_n \to X \quad \text{in } L^1 \quad \text{or} \quad E[X_n] \to E[X] \quad \text{may not hold.}$$

- [Example] $\Omega = (0,1), P =$Lebesgue meas, $X_n = n1_{(0,\frac{1}{n})}$. Then, $X_n \to X = 0$ a.s., but not in L^1-sense.
- One sufficient condition is given by Lebesgue's convergence theorem:

$$X_n \to X \text{ (a.s.) and } |X_n| \leq^{\exists} Y \in L^1 \Longrightarrow E[X_n] \to E[X].$$

- Uniform integrability provides more general sufficient condition.

[Definition 9.4] (X_n) is uniformly integrable

$$\underset{\text{def}}{\Longleftrightarrow} \lim_{\lambda\to\infty} \sup_n E\left[|X_n|, |X_n| \geq \lambda\right] = 0 \qquad \square$$

• The condition $|X_n| \leq^{\exists} Y \in L^1$ for Lebesgue's convergence theorem implies the uniform integrability of (X_n).

$\because$

$$E\left[|X_n|, |X_n| \geq \lambda\right] \leq E\left[|X_n|, Y \geq \lambda\right] \leq E\left[Y, Y \geq \lambda\right]$$
$$= E[Y 1_{\{Y \geq \lambda\}}] \underset{\lambda\to\infty}{\longrightarrow} 0.$$

1st inequality follows from $\{|X_n| \geq \lambda\} \subset \{Y \geq \lambda\}$, while the last convergence to 0 follows by Lebesgue's convergence theorem. $\square$

- $^{\exists}p > 1$ s.t. $\sup_n E[|X_n|^p] < \infty$
  $\Longrightarrow (X_n)$: uniformly integrable

$\because$

$$E\left[|X_n|, |X_n| \geq \lambda\right] \leq E\left[|X_n|\left(\frac{|X_n|}{\lambda}\right)^{p-1}, |X_n| \geq \lambda\right]$$
$$\leq \frac{1}{\lambda^{p-1}} E[|X_n|^p] \underset{\lambda\to\infty}{\longrightarrow} 0 \quad \text{(uniformly in } n\text{)}. \qquad \square$$

[Lemma 9.12] If (X_n) is uniformly integrable, then we have

(1) $\sup_n E[|X_n|] < \infty$

(2) For $^{\forall}\varepsilon > 0$, $^{\exists}\delta > 0$ such that
$\sup_n E[|X_n|, A] < \varepsilon$ for $^{\forall}A \in \mathcal{F}$ satisfying $P(A) < \delta$. $\square$

[Proof] (1) is easy.

∵
$$E[|X_n|] = E[|X_n|, |X_n| < \lambda] + E[|X_n|, |X_n| \geq \lambda] \\ \leq \lambda + E[|X_n|, |X_n| \geq \lambda]$$

The 2nd term converges to 0 as $\lambda \to \infty$ uniformly in n. Thus, at least for λ large enough, this term is bounded in n. □

We show (2).

∵
$$E[|X_n|, A] = E[|X_n|, A \cap \{|X_n| < \lambda\}] \\ + E[|X_n|, A \cap \{|X_n| \geq \lambda\}] \\ \leq \lambda P(A) + E[|X_n|, |X_n| \geq \lambda]$$

If we take $\lambda > 0$ sufficiently large, the 2nd term is bounded as $\leq \frac{\varepsilon}{2}$ uniformly in n. We take λ in such way and fix it. Then, taking $\delta = \frac{\varepsilon}{2\lambda}$, we have

$$P(A) < \delta \Longrightarrow \lambda P(A) < \frac{\varepsilon}{2},$$

which shows (2). □

[Proposition 9.13] **(Criterion for the uniform integrability)**
(X_n) : uniform integrable
$\Longleftrightarrow$ Two conditions (1), (2) in Lemma 9.12 hold. $\square$

$\because$ [$\Longrightarrow$] follows from Lemma 9.12.
[$\Longleftarrow$] Take $A_n = \{|X_n| \geq \lambda\}$ and set $M = \sup_n E[|X_n|]$.

Then,

$$P(A_n) \underset{\text{Chebyshev}}{\leq} \frac{1}{\lambda} E[|X_n|] \underset{(1)}{\leq} \frac{M}{\lambda}.$$

Therefore, by (2), if $\frac{M}{\lambda} < \delta$ (i.e. if $\lambda > 0$ is large enough), we have $E[|X_n|, A_n] < \varepsilon$. $\square$

[Theorem 9.14] **(Generalization of Lebesgue's convergence thm)**
Assume (X_n) is uniform integrable. Then, we have

(1) $X_n \to X$ (a.s.) $\Longrightarrow X_n \to X$ in L^1. In particular, X is integrable and $\lim_{n\to\infty} E[X_n] = E[X]$ holds.

(2) $X_n \to X$ in prob. $\Longleftrightarrow X_n \to X$ in L^1 □

[Proof] [proof of (1)] • The integrability of X follows by Fatou's lemma:

$$E[|X|] \le \liminf_{n\to\infty} E[|X_n|] < \infty.$$

• Decompose as

$$\begin{aligned} E[|X_n - X|] =& E[|X_n - X|, |X_n - X| < \lambda] \\ &+ E[|X_n - X|, |X_n - X| \ge \lambda]. \end{aligned}$$

• Noting the integrability of X, one can easily show that $(X_n - X)_n$ satisfies two conditions (1), (2) in Lemma 9.12. Thus, it is uniformly integrable so that the 2nd term is bounded as $\le \varepsilon$ uniformly in n, if we take λ large enough.

P: Check that $(X_n - X)_n$ satisfies two conditions (1), (2).

• Take such $\lambda > 0$ and fix it. Then, the 1st term which is rewritten as $E[|X_n - X| 1_{\{|X_n - X| < \lambda\}}]$ converges to 0 as $n \to \infty$ by applying Lebesgue's convergence theorem, since the integrand is bounded in n. Thus, (1) is shown.

[proof of (2)]

[$\Longleftarrow$] holds in general (by Chebyshev's inequality).

[$\Longrightarrow$] Take any subsequence $(X_{n'})$ of (X_n). Then, it also converges in probability to X. Hence, $^\exists$its further subsequence $(X_{n''})$ s.t. $X_{n''} \to X$ (a.s.). $\therefore$ By (1), we have $X_{n''} \to X$ in L^1.

However, since the limit X is determined independently of subsequences $(X_{n'})$ or $(X_{n''})$, this means that X_n itself (without taking any subsequence) converges to X in L^1-sense. □

P: Show the last step. (This can be shown by proof by contradiction (reductio ad absurdum).)

[Theorem 9.15] Assume X_n is integrable and $X_n \to X$ (a.s.) holds. Then, the following 3 conditions are equivalent.

(1) (X_n): uniformly integrable

(2) $X_n \to X$ in L^1

(3) $E[|X|] < \infty$ and $E[|X_n|] \to E[|X|]$

[Proof] (1) $\Longrightarrow$ (2) follows by Theorem 9.14-(1).

(2) $\Longrightarrow$ (3)

$\because$

$$\begin{aligned} \big|E[|X_n|] - E[|X|]\big| &= \big|E[|X_n| - |X|]\big| \\ &\le E\left[\big||X_n| - |X|\big|\right] \le E[|X_n - X|] \longrightarrow 0 \end{aligned} \qquad \square$$

(3) $\Longrightarrow$ (1)

$\because$

$$E[|X_n|, |X_n| \ge \lambda] = E[|X_n|] - E[|X_n|, |X_n| < \lambda]$$

By Assumption (3), 1st term $\to E[|X|]$.

For the 2nd term, if $\lambda > 0$ satisfies $P(|X| = \lambda) = 0$, we have

$$\lim_{n\to\infty} E[|X_n|, |X_n| < \lambda] = E[|X|, |X| < \lambda].$$

Indeed, by $P(|X| = \lambda) = 0$,

$$|X_n| \cdot 1_{\{|X_n|<\lambda\}} \longrightarrow |X| \cdot 1_{\{|X|<\lambda\}} \quad a.s.$$

so that one can apply Lebesgue's convergence theorem.

Thus, if $P(|X| = \lambda) = 0$, we obtain

$$\begin{aligned}\lim_{n\to\infty} E[|X_n|, |X_n| \geq \lambda] &= E[|X|] - E[|X|, |X| < \lambda] \\ &= E[|X|, |X| \geq \lambda].\end{aligned}$$

However, since $E[|X|] < \infty$, (RHS) $< \varepsilon$ if we take $\lambda > 0$ large enough. Moreover, λ for which $P(|X| = \lambda) > 0$ holds are at most countably many ($\because$ the number of λ such that $P(|X| = \lambda) \geq 1/k$ is at most k). Thus, ${}^\exists\lambda > 0$, ${}^\exists n_0$ s.t.

$$E[|X_n|, |X_n| \geq \lambda] < 2\varepsilon$$

holds for $n \geq n_0$. $\therefore$ (X_n) is uniformly integrable. □

Next theorem is known. (We state without proof.)

[Theorem] (Dunford–Pettis's compactness criterion)
(X_n): uniformly integrable
$\Longleftrightarrow$ (X_n) is relatively compact in the space $L^1(\Omega)$
under the weak convergence topology. $\square$

Here, $X_n \to X$ (weakly converge in $L^1(\Omega)$)
$\underset{\text{def}}{\Longleftrightarrow}$ For ${}^\forall Y \in L^\infty(\Omega)$ $(= (L^1)^*)$, $E[X_n Y] \to E[XY]$.

See Dunford-Schwartz "Linear Operators, Part I" or
Dellacherie-Meyer "Probabilities and Potential", Chapter 2.

Summary of discussion on discrete time martingales and submartingales:

$(\Omega, \mathcal{F}, P)$: Probability space
$(\mathcal{F}_n)_{n=1,2,\ldots}$: filtration (or reference family)
$X = (X_n)_{n=1,2,\ldots}$: $(\mathcal{F}_n)$-adapted stochastic process

9.1 Definition of discrete time martingales and submartingales
9.2 Doob decomposition
9.3 Markov time
9.4 Doob's optional sampling theorem
9.5 Doob's inequality
9.6 Submartingale convergence theorem
9.7 Moment inequality

9.8 Uniform integrability of a sequence of r.v.'s

§10 Continuous time martingale

Similar results hold in continuous time case.

10.1 Definition

$(\Omega, \mathcal{F}, P)$: Probability space

$(\mathcal{F}_t) \equiv (\mathcal{F}_t)_{t \geq 0}$: filtration (or reference family)

$\underset{\text{def}}{\Longleftrightarrow}$ · Each $\mathcal{F}_t$ is a sub σ–field of $\mathcal{F}$.

· Increasing, i.e., $0 \leq s < t \Longrightarrow \mathcal{F}_s \subset \mathcal{F}_t$

· Right continuous, i.e., for ${}^{\forall} t \geq 0$, $\mathcal{F}_t = \mathcal{F}_{t+}$ holds,
where $\mathcal{F}_{t+} := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$.

We assume that stochastic process $X = (X_t)_{t \geq 0}$ (i.e. each X_t is a real-valued r.v.) satisfies

- Right continuous having left limits i.e., for ${}^\forall \omega \in \Omega$, $t \in [0, \infty) \mapsto X_t(\omega) \in \mathbb{R}$ is right continuous and has left limit at every t.

We call such X càdlàg (continue à droite limites à gauche).

[Definition 10.1] X is called $(\mathcal{F}_t)$-martingale, if
(1) $(\mathcal{F}_t)$–adapted: X_t is $\mathcal{F}_t$–measurable for ${}^\forall t \geq 0$.
(2) Integrability: $E[|X_t|] < \infty$ for ${}^\forall t \geq 0$.
(3) For $0 \leq^\forall s <^\forall t$, $E[X_t | \mathcal{F}_s] = X_s \quad a.s.$
X is called submartingale if $E[X_t|\mathcal{F}_s] \geq X_s$ $a.s.$ holds in (3) and supermartingale if $E[X_t|\mathcal{F}_s] \leq X_s$ $a.s.$ □

Similar to discrete time case, the following hold also in continuous time: Doob's inequality, Burkholder's inequality, Doob's optional sampling theorem, submartingale convergence theorem, Doob-Meyer decomposition and others.

10.2 Markov time

[Definition 10.2] $\sigma : \Omega \to [0, \infty]$ is $(\mathcal{F}_t)$-Markov time or stopping time $\underset{\text{def}}{\iff}$ $\{\sigma \leq t\} \in \mathcal{F}_t$ for ${}^{\forall} t \geq 0$. □

[Proposition 10.1] The following (1)–(4) are equivalent
(1) σ is a Markov time
(2) $\{\sigma < t\} \in \mathcal{F}_t$ for ${}^{\forall} t \geq 0$
(3) $\{\sigma > t\} \in \mathcal{F}_t$ for ${}^{\forall} t \geq 0$
(4) $\{\sigma \geq t\} \in \mathcal{F}_t$ for ${}^{\forall} t \geq 0$ □

[Proof] (1) $\iff$ (3) and (2) $\iff$ (4) are obvious.
If we assume (1), by $\{\sigma < t\} = \bigcup_{n=1}^{\infty} \{\sigma \leq t - \frac{1}{n}\} \in \mathcal{F}_t$, we obtain (2). Conversely, if we assume (2), since $\{\sigma \leq t\} = \bigcap_{n=N}^{\infty} \{\sigma < t + \frac{1}{n}\} \in \mathcal{F}_{t+\frac{1}{N}}$ for ${}^{\forall} N = 1, 2, \ldots$, the right-continuity of $(\mathcal{F}_t)$ implies (1). □

Markov times are invariant under taking maximum, minimum, sum and limits.

[Proposition 10.2] Let $\sigma, \tau, \sigma_n, n = 1, 2, \ldots$ be Markov times, where the sequence σ_n is assumed to be increasing ($\sigma_1 \leq \sigma_2 \leq \cdots$, more precisely, for ${}^\forall\omega \in \Omega$, $\sigma_1(\omega) \leq \sigma_2(\omega) \leq \cdots$, we write $\sigma_n \nearrow$) or decreasing ($\sigma_1 \geq \sigma_2 \geq \cdots$, we write $\sigma_n \searrow$)

$\Longrightarrow$ The followings are all Markov times:

$$\sigma \vee \tau \equiv \max\{\sigma, \tau\}, \quad \sigma \wedge \tau \equiv \min\{\sigma, \tau\},$$
$$\sigma + \tau, \quad \lim_{n\to\infty} \sigma_n$$

□

[Proof] Same as discrete time case for $\sigma \vee \tau$, $\sigma \wedge \tau$.
For $\sigma + \tau$, note that

$$\{\sigma + \tau < t\} = \bigcup_{r \in \mathbb{Q} \cap (0,t)} \left[\{\sigma < r\} \cap \{\tau < t - r\}\right]$$

and Proposition 10.1.

For the limit, noting

$$\{\lim_{n\to\infty}\sigma_n \le t\} = \bigcap_{n=1}^{\infty}\{\sigma_n \le t\} \quad (\text{when } \sigma_n \nearrow)$$

$$\{\lim_{n\to\infty}\sigma_n < t\} = \bigcup_{n=1}^{\infty}\{\sigma_n < t\} \quad ((\text{when } \sigma_n \searrow),$$

we may use Proposition 10.1. □

[Example] ((first) hitting time) $(X_t)_{t\ge 0}$: $(\mathcal{F}_t)$–adapted $\mathbb{R}^d$–valued continuous stochastic process. For $A\,(\subset \mathbb{R}^d)$, set

$$\sigma_A(\omega) := \inf\{t > 0; X_t(\omega) \in A\},$$

where $\inf\emptyset := \infty$ as before. Then, if A is open or closed set, σ_A is a Markov time. □

$\because$ When A is open, this is shown by

$$\{\sigma_A < t\} = \bigcup_{r\in\mathbb{Q},r<t}\{X_r \in A\} \in \mathcal{F}_t.$$

See the next page for a closed set A. □

P: Show that σ_A is a Markov time for a closed set A.

[Hint] Set $A_n := \{x \in \mathbb{R}^d; \operatorname{dist}(x, A) < \frac{1}{n}\}$. Then, since A_n is open, σ_{A_n} is a Markov time. Obviously, $\sigma_{A_n} \leq \sigma_{A_{n+1}}$ so that the limit $\sigma = \lim_{n\to\infty} \sigma_{A_n}$ exists and it is a Markov time. $\sigma \leq \sigma_A$ is obvious. Show that $\sigma \geq \sigma_A$ by the continuity of (X_t) and the closedness of A. Once this is shown, we see $\sigma_A = \sigma$ so that σ_A is a Markov time. □

[Remark] If the probability space $(\Omega, \mathcal{F}, P)$ is complete and $\mathcal{N} \subset \mathcal{F}_0$ holds for $\mathcal{N} = \{N \in \mathcal{F}; P(N) = 0\}$, then it is known that σ_A is a Markov time for ${}^\forall A \in \mathcal{B}(\mathbb{R}^d)$; see Dellacherie–Meyer Chapter 4, Ikeda–Watanabe Chapter 1. □

For a Markov time σ,

$$\mathcal{F}_\sigma \underset{\text{def}}{=} \{A \in \mathcal{F};\ A \cap \{\sigma \le t\} \in \mathcal{F}_t \text{ for } {}^\forall t \ge 0\}$$

is a σ–field. ($\because$ We may check the axiom of σ–field. Same as the discrete time case.) Equivalently, one can define $\mathcal{F}_\sigma$ as

$$\mathcal{F}_\sigma = \{A \in \mathcal{F};\ A \cap \{\sigma < t\} \in \mathcal{F}_t \text{ for } {}^\forall t \ge 0\}.$$

Proof is similar to that of Proposition 10.1.

In particular, $\sigma(\omega) \equiv s$ (constant) is a Markov time and $\mathcal{F}_\sigma = \mathcal{F}_s$ holds.
($\because$ This is easy from $\{s \le t\} = \Omega$ if $s \le t$, $= \emptyset$ if $s > t$)

[Proposition 10.3] $\sigma, \tau, \sigma_n, n = 1, 2, \ldots$: Markov times. Then,
(1) σ is $\mathcal{F}_\sigma$–measurable
(2) $\tau \leq \sigma$ (i.e. $\tau(\omega) \leq \sigma(\omega), {}^\forall \omega$) $\Longrightarrow \mathcal{F}_\tau \subset \mathcal{F}_\sigma$
(3) (Generalization of right-continuity)

$$\sigma_n \searrow \sigma \ \text{(decreasing, } {}^\forall\omega) \Longrightarrow \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n} = \mathcal{F}_\sigma \qquad \square$$

[Proof] Proof of (1), (2) is the same as the discrete time case, but we give it below.

[Proof of (1)] It's enough to show $\{\sigma \leq s\} \in \mathcal{F}_\sigma$ for ${}^\forall s \geq 0$. For this, we may show $\{\sigma \leq s\} \cap \{\sigma \leq t\} \in \mathcal{F}_t$ for ${}^\forall t \geq 0$. However, this follows by $\{\sigma \leq s\} \cap \{\sigma \leq t\} = \{\sigma \leq s\}$ if $s \leq t$, $= \{\sigma \leq t\}$ if $s \geq t$.

[Proof of (2)] Take ${}^\forall A \in \mathcal{F}_\tau$. $A \in \mathcal{F}_\sigma$ follows, by noting $\tau \leq \sigma$, from $A \cap \{\sigma \leq t\} = A \cap [\{\sigma \leq t\} \cap \{\tau \leq t\}] = [A \cap \{\tau \leq t\}] \cap \{\sigma \leq t\} \in \mathcal{F}_t$ for ${}^\forall t \geq 0$.

[Proof of (3)] $\mathcal{F}_\sigma \subset \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n}$ follows from (2) and $\sigma \le \sigma_n$. To show the converse, take ${}^\forall A \in \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n}$. Then, for ${}^\forall t \ge 0$, $A \cap \{\sigma < t\} = \bigcup_{n=1}^{\infty} [A \cap \{\sigma_n < t\}] \in \mathcal{F}_t$ and this implies $A \in \mathcal{F}_\sigma$. (Note: if we take $\{\sigma \le t\}$, it does not work well.)

□

10.3 Doob's inequality

(X_t): càdlàg is always assumed.

[Theorem 10.4] (1) (X_t): submartingale. Then, for ${}^\forall a > 0$,

$$P\left(\sup_{0\le s\le t} X_s \ge a\right) \le \frac{1}{a} E\left[X_t, \sup_{0\le s\le t} X_s \ge a\right] \le \frac{1}{a} E[X_t^+].$$

(2) (M_t): pth power integrable $(E[|M_t|^p] < \infty, t \ge 0)$ martingale or non-negative submartingale for $p > 1$. Then,

$$E\left[\sup_{0\le s\le t} |M_s|^p\right] \le \left(\tfrac{p}{p-1}\right)^p E[|M_t|^p].$$

□

[Outline of proof] We use approximations from discrete time. Indeed, for $^\forall$division $\pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ of the time interval $[0, t]$, $(X_{t_k})_{k=0,1,2,\dots,n}$ becomes a discrete time $(\mathcal{F}_{t_k})$-submartingale. Therefore, by Doob's inequality in discrete time, we have

$$P\left(\max_{0\le k\le n} X_{t_k} \ge a\right) \le \frac{1}{a} E\left[X_t, \max_{0\le k\le n} X_{t_k} \ge a\right].$$

Then, noting that (X_t) is right-continuous, by making divisions finer and taking the limit, we can complete the proof of (1). More precisely, setting $X^\pi := \max_{0\le k\le n} X_{t_k}$, we can show

- $\pi_1 \subset \pi_2$ (π_2 is finer division) $\Longrightarrow X^{\pi_1} \le X^{\pi_2}$
- $\lim_{|\pi_m|\to 0} X^{\pi_m} = \sup_{0\le s\le t} X_s$

The proof of (2) is the same as that of Theorem 9.10. □

P: Complete the proof of (1). (Hint: First show the conclusion with "$> a$" instead of "$\ge a$". Then show with "$\ge a$".)

[Theorem 10.5] (submartingale convergence theorem)
If a submartingale $(X_t)_{t \geq 0}$ satisfies $\sup_{t \geq 0} E[X_t^+] < \infty$,
$\Longrightarrow {}^{\exists}X := \lim_{t \to \infty} X_t$ (a.s.) and $E[|X|] < \infty$ □

⚇ By discretizing the time as above, we can derive the upcrossing inequality also in continuous time case. For details, see, for example, Revuz-Yor "Continuous martingales and Brownian motion", p.57–p.62. □