STA447/2006 Lecture 7

Midterm 2: next week.

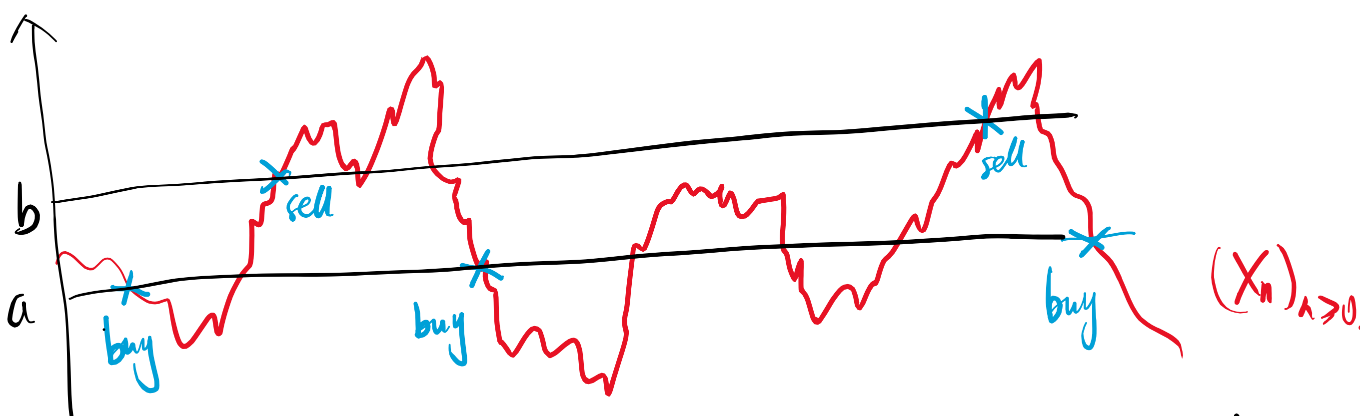
- Same format / timing / location / etc.
- Cover Lec 5 – 7.

Thm (MG convergence)

$(X_n)_{n\geq 0}$ martingale, $\sup_{n\geq 0} \mathbb{E}[|X_n|] < +\infty$ $(\forall n)$

then $\exists X_\infty$ s.t. $\mathbb{P}(X_n \to X_\infty) = 1$.

Proof idea: "up-crossing".



Intuition: you may only trade finitely many times using this strategy.

$$\mathbb{P}(\exists\, a,b \in \mathbb{R}, \overset{a<b}{\text{s.t.}}\ \#\text{up-crossings for } [a,b] = +\infty)$$

$$= \mathbb{P}(\exists\, a,b \in \mathbb{Q}, \overset{a<b}{\text{s.t.}}\ \#\text{up-crossing for } [a,b] = +\infty).$$

$$\leq \sum_{\substack{a,b\in\mathbb{Q}\\ a<b}} \mathbb{P}(\#\text{up-crossing for } [a,b] = +\infty).$$

To show convergence, it suffices to prove

$$\boxed{\forall a<b, \quad \mathbb{P}(\#\text{ up-crossing } [a,b] = +\infty) = 0.} \quad (*)$$

Bounding up-crossings:

$W_n :=$ amount of money made using trading strategy.

$$= \sum_{j=1}^{n} B_j (X_j - X_{j-1})$$

where $B_j =$ "holding stock at j-th step". $\in \{0,1\}$.

$$B_j = \begin{cases} 1 & \text{when } B_{j-1}=0,\ X_{j-1} \leq a \quad \text{"buy"} \\ 0 & \text{when } B_{j-1}=1,\ X_{j-1} \geq b \quad \text{"sell"} \\ B_{j-1} & \text{otherwise.} \quad \text{"keep".} \end{cases}$$

Key observation: $(W_n)_{n\geq 0}$ is MG.

$$\mathbb{E}[W_{n+1} | \mathcal{F}_n] = W_n + \boxed{\mathbb{E}[\underbrace{B_{n+1}}_{\text{determined by } (X_1 \cdots X_n).} \cdot (X_{n+1} - X_n) | \mathcal{F}_n]} = W_n.$$

$B_{n+1} \cdot \mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] = 0$

Let $U_n := \#$ upcrossing $[a,b]$ up to time n.

we have

$$W_n \geq (b-a)U_n - |X_n - a|.$$

$$\mathbb{E}[W_n] \geq (b-a)\mathbb{E}[U_n] - \left(\mathbb{E}[|X_n|] + a\right).$$

$$\overset{\shortparallel}{0 = \mathbb{E}[W_0]}$$

So we have $\mathbb{E}[U_n] \leq \frac{\mathbb{E}[|X_n|]+a}{b-a} \leq \frac{C+a}{b-a}$

The bound is indp of n.

So $\forall a,b$, $\mathbb{E}[\#\text{upcrossing } [a,b] \text{ over } [0,+\infty)] < +\infty$.

$\mathbb{P}(\#\text{upcrossing } [a,b] = +\infty) = 0$, which proves $(*)$.

Next question: $\mathbb{E}[X_\infty] \overset{?}{=} \mathbb{E}[X_0]$.

Counter-example: $\varepsilon_n = \begin{cases} +1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2}. \end{cases}$

$X_0 = 1$, $X_n = X_{n-1} + \varepsilon_n$

$Y_n = X_{\min\{n,T\}}$ $T =$ hitting time of 0.

$(Y_n)_{n\geq 1}$ is MG.

$$\mathbb{E}[|Y_n|] = \mathbb{E}[Y_n] = \mathbb{E}[Y_0] = 1 < +\infty.$$

By MG convergence, $Y_n \to Y_\infty$. (a.s.)

Indeed, since $(X_n)_{n\geq 0}$ is irreducible and recurrent MG.

$\mathbb{P}(T < +\infty) = 1$, So $Y_n \to 0$ (a.s.).

$$0 = \mathbb{E}[Y_\infty] \neq \mathbb{E}[Y_0] = 1.$$

Recall uniform integrability.

$\forall \varepsilon > 0, \exists K$, s.t. $\sup_{n\geq 0} \mathbb{E}\left[|X_n| \cdot \mathbb{1}_{|X_n|\geq K}\right] \leq \varepsilon$.

Thm. $(X_n)_{n\geq 0}$ u.i., $X_n \xrightarrow{a.s.} X_\infty$

then we have $\mathbb{E}|X_n - X_\infty| \longrightarrow 0$

and in MG convergence context, $\mathbb{E}[X_\infty] = \mathbb{E}[X_0]$.

- General result for sequence of r.v.'s
- u.i. implies $\sup_{n\geq 0} \mathbb{E}[|X_n|] \leq C < +\infty$

We only need to verify UI to establish $\begin{cases} \text{a.s.} \\ L^1 \end{cases}$

Proof of thm:

$$\mathbb{E}|X_n - X_\infty|$$

$$\leq \mathbb{E}\left[\left|X_n \cdot \mathbb{1}_{|X_n|\leq K} - X_\infty \mathbb{1}_{|X_\infty|\leq K}\right|\right]$$

$$+ \mathbb{E}\left[|X_n| \cdot \mathbb{1}_{|X_n|>K}\right] + \mathbb{E}\left[|X_\infty| \mathbb{1}_{|X_\infty|>K}\right]$$

Bounded by K + ptwise convergence

$\Downarrow$

$\mathbb{E}[|\cdots|] \to 0$

by UI, taking $K = K_\varepsilon$

$$\mathbb{E}\left[|X_n| \mathbb{1}_{|X_n|>K}\right] \leq \varepsilon$$

By Fatou's Lemma

$X_n \to X_\infty$.

$$\mathbb{E}\left[|X_\infty|\right] \leq \liminf_{n\to+\infty} \mathbb{E}[|X_n|] < +\infty$$

$\forall \varepsilon > 0$, $\exists K'_\varepsilon$

s.t. $\mathbb{E}\left[|X_\infty| \cdot \mathbb{1}_{|X_\infty|>K'_\varepsilon}\right] < \varepsilon$

Taking $K = \max\{K_\varepsilon, K'_\varepsilon\}$.

we conclude

$$\limsup_{n\to+\infty} \mathbb{E}|X_n - X_\infty| \leq 2\varepsilon \qquad \forall \varepsilon > 0.$$

So we have $\mathbb{E}|X_n - X_\infty| \to 0$.

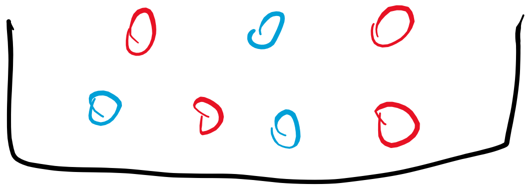
---

e.g. Random harmonic series

$$M_n = \sum_{j=1}^{n} \frac{1}{j} X_j \quad \text{where } X_j = \begin{cases} +1 & \text{w.p. } \tfrac{1}{2} \\ -1 & \text{w.p. } \tfrac{1}{2}. \end{cases}$$

$$\text{UI MG} \implies M_n \xrightarrow[L^1]{a.s.} M_\infty.$$

eg. Polya's Urn.



Box

Initially $\begin{cases} 1 & \text{red} \\ 1 & \text{blue} \end{cases}$

We let $M_n :=$ Proportion of red.

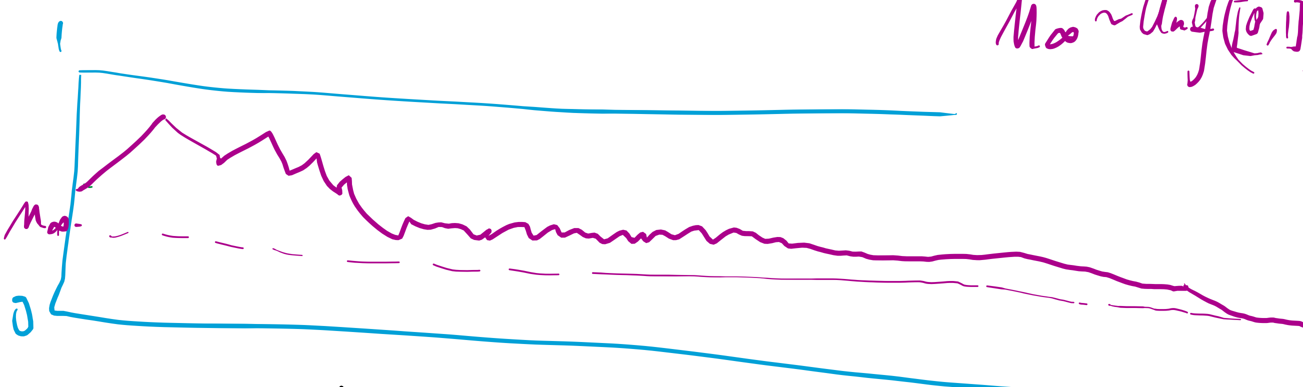
For each time, add a new ball $\begin{cases} \text{red} & \text{w.p } M_n \\ \text{blue} & \text{w.p. } 1-M_n. \end{cases}$

$(M_n)_{n\geq 0}$ is an MG.

$M_n \in [0,1]$ a.s. $\forall n$ So U.I.

$$M_n \xrightarrow[L^1]{a.s.} M_\infty.$$

Actually, we can show that $M_\infty \sim \text{Unif}([0,1])$.



In general, $M_\infty \sim$ Beta distr depending on initial configuration.

Connection between OST and MG convergence

$(X_n)_{n\geq 1}$ UI martingale

T is a stopping time.

$\boxed{T \text{ does not need to be finite w.p. } 1.}$

$Y_n := X_{n\wedge T}$ for $n \geq 0$

$(Y_n)_{n\geq 0}$ is also an MG.

Claim: $(Y_n)_{n\geq 0}$ is uniformly integrable

So $Y_n \xrightarrow[L^1]{a.s.} Y_\infty = X_T$.

(If $T < +\infty$, $Y_n = X_{n\wedge T} \xrightarrow{a.s.} X_T$

If $T = +\infty$, $Y_n = X_n \xrightarrow{a.s.} X_\infty$).

By L^1 convergence, we have.

$$\mathbb{E}[X_T] = \mathbb{E}[Y_n] = \mathbb{E}[Y_0] = \mathbb{E}[X_0].$$

(This arguments only requires $(X_{n\wedge T})_{n\geq 0}$ to be UI) e.g. $X_0 = 1$, $X_{n+1} = \begin{cases} \frac{3}{2} X_n & \text{w.p. } 1/2 \\ \frac{1}{2} X_n & \text{w.p. } 1/2. \end{cases}$

$(X_n)_{n \geq 0}$ is MG.

$$\log X_n = \sum_{i=1}^{n} \log(Z_i) \quad \text{with } Z_i = \begin{cases} \log(\frac{3}{2}) & \text{w.p. } 1/2 \\ \log(\frac{1}{2}) & \text{w.p. } 1/2. \end{cases}$$

$$\mathbb{E}[Z_i] < 0.$$

By SLLN, $\frac{1}{n} \sum_{i=1}^{n} \log(Z_i) \xrightarrow{a.s.} \mathbb{E}[Z_1] < 0$.

So $X_n \xrightarrow{a.s.} 0$

$$T := \inf\{t : X_t \geq 10\}.$$

$(X_{n \wedge T})_{n \geq 0}$ is bdd (and therefore UI) martingale.

Applying the generalized version of OST.

$$X_{n \wedge T} \xrightarrow[L^1]{a.s.} X_T.$$

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] = 1.$$

$$1 = \mathbb{E}[X_T]$$

$$= \mathbb{P}(T=+\infty)\cdot 0 + \mathbb{P}(T<+\infty)\cdot \mathbb{E}[X_T \mid T<+\infty]$$

$$\geq 10 \cdot \mathbb{P}(T<+\infty).$$

So $\mathbb{P}(T=+\infty) \geq 0.9.$

How about L^p convergence? $(p>1)$

Technical tool: Doob's L^p maximal inequality.

$(X_n)_{n\geq 0}$ martingale.

$$\mathbb{P}\left(\max_{0\leq t\leq n} |X_t| \geq a\right) \leq \frac{\mathbb{E}[|X_n|^p]}{a^p}$$

(Proof similar to L^1 version).

Integrating the tail bound,

$$\mathbb{E}\left[\max_{0\leq t\leq n} |X_t|^p\right] \leq \frac{p}{p-1}\,\mathbb{E}\left[|X_n|^p\right]. \quad \leq C<+\infty$$

Application to L^p convergence.

$$\mathbb{E}\left[\boxed{|X_n - X_\infty|^p}\right] \overset{?}{\longrightarrow} 0.$$

$\longrightarrow 0$ a.s.

Need a dominance function and apply DCT.

$$|X_n - X_\infty|^p \le 2^{p-1}\left(|X_n|^p + |X_\infty|^p\right).$$

$$\le \underbrace{2^{p-1}\left(\sup_{n\ge 0} |X_n|^p + |X_\infty|^p\right)}_{\text{Dominance function}}.$$

Assumption: $\sup_{n\ge 0} \mathbb{E}\left[|X_n|^p\right] \le C < +\infty.$

By Fatou's lemma, $\mathbb{E}\left[|X_\infty|^p\right] < +\infty.$

Using Doob's L^p ineq, we can show

$$\mathbb{E}\left[\sup_{n\ge 0} |X_n|^p\right] \le \frac{p}{p-1} C < +\infty.$$

By DCT, we have $X_n \xrightarrow[L^p]{a.s.} X_\infty.$

$$\mathbb{E}\left[|X_n - X_\infty|^p\right] \longrightarrow 0.$$

Doob's MG.

Let X be a random variable

$$\mathbb{E}[|X|] < +\infty.$$

"filtration" $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \cdots$

flow of information.

e.g. In stats. X is the underlying signal,

$\mathcal{F}_n$ contains first n observations.

$$M_n = \mathbb{E}[X \mid \mathcal{F}_n].$$

"best prediction" based on info up to time n.

$(M_n)_{n \geq 0}$ is MG.

$$\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n].$$

$$\text{(tower law)} = \mathbb{E}[X \mid \mathcal{F}_n] = M_n$$

Fact: $(M_n)_{n \geq 0}$ is UI.

Proof: $\mathbb{E}\left[|M_n| \cdot \mathbb{1}_{|M_n| \geq K}\right]$

$$= \mathbb{E}\left[\left|\mathbb{E}[X|\mathcal{F}_n]\right| \cdot \mathbb{1}_{|\mathbb{E}[X|\mathcal{F}_n]| \geq K}\right]$$

$$\leq \mathbb{E}\left[|X| \cdot \mathbb{1}_{\max_{n\geq 0} |M_n| \geq K}\right].$$

By Doob's max ind ineq,

$$\mathbb{P}\left(\max_{0\leq i\leq n} |M_i| \geq K\right) \leq \frac{\mathbb{E}|M_n|}{K} \leq \frac{\mathbb{E}|X|}{K}$$

So, $\mathbb{P}\left(\max_{n\geq 0} |M_n| \geq K\right) \leq \frac{\mathbb{E}|X|}{K}. \xrightarrow{K\to+\infty} 0$

So $|X| \cdot \mathbb{1}_{\max_{n\geq 0} |M_n| \geq K} \xrightarrow{K\to+\infty} 0$, dominated by $|X|$.

So by DCT. $\lim_{K\to+\infty} \mathbb{E}\left[|X| \cdot \mathbb{1}_{\max_{n\geq 0} |M_n| \geq K}\right] = 0.$

This implies $(M_n)_{n\geq 0}$ UI.

So we have

$$M_n \xrightarrow[L^1]{a.s.} M_\infty.$$

Remark: we can show

$$M_\infty = \mathbb{E}[X \mid \mathcal{F}_\infty]$$

where $\mathcal{F}_\infty = \bigcup_{n=1}^{+\infty} \mathcal{F}_n$. all the info in the infinite seq.

Application in statistics.

$$\theta \sim \pi \quad (\text{prior distribution})$$

$$X_1, X_2, \cdots, X_n, \cdots \mid \theta \overset{iid}{\sim} P_\theta.$$

$$\pi(\theta \mid X_1 \cdots X_n) = \frac{\pi(\theta) \cdot P_\theta(X_1) \cdots P_\theta(X_n)}{\int \pi(\theta') P_{\theta'}(X_1) \cdots P_{\theta'}(X_n) \, d\theta'}.$$

Goal: to estimate $g(\theta) \in \mathbb{R}$.

$$\hat{g}_n := \mathbb{E}[g(\theta) \mid X_1, \cdots, X_n].$$ is Doob's MG.

$X = g(\theta)$, $\mathcal{F}_n = (X_1, X_2, \cdots, X_n)$.

$$\hat{g}_n \xrightarrow[L^1]{a.s.} g_\infty = \mathbb{E}[g(\theta) \mid X_1, X_2, \cdots X_n, \cdots]$$

Doob (1953) shows that as long as

the model is identifiable,

we have $g_\infty = g(\theta)$.

Indeed, Polya's Urn model, is equivalent

~~to~~ estimating θ in $Ber(\theta)$

with iid data, where

$\theta \sim \pi = Beta(\alpha, \beta)$