

Advanced Probability

-Martingales

=====

(15th October 2018, Monday)

Chapter 2. Martingales in Discrete Time

2.1. Definitions.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

- A **Filtration** for $(\Omega, \mathcal{F}, \mathbb{P})$ is a sequence $(\mathcal{F}_n)_{n \geq 0}$ of σ -algebras s.t. for all $n \geq 0$, we have

$$\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$$

Set $\mathcal{F}_\infty = \sigma(\mathcal{F}_n : n \geq 0)$ then $\mathcal{F}_\infty \subset \mathcal{F}$. We allow $\mathcal{F}_\infty \neq \mathcal{F}$. We interpret n as times and \mathcal{F}_n as the extent of knowledge at time n .

- A **Random process(in discrete time)** is a sequence of random variables $(X_n)_{n \geq 0}$. It has a natural filtration $(\mathcal{F}_n^X)_{n \geq 0}$ given by

$$\mathcal{F}_n^X = \sigma(X_0, \dots, X_n)$$

That is, the knowledge obtained from X_n by time n . We say $(X_n)_{n \geq 0}$ is **adapted to** $(\mathcal{F}_n)_{n \geq 0}$ if X_n is \mathcal{F}_n -measurable for all $n \geq 0$. This is equivalent to having $\mathcal{F}_n^X \subset \mathcal{F}_n$, for all $n \geq 0$. (Here, X_n are real-valued)

- We would say $(X_n)_{n \geq 0}$ is **integrable** if X_n is integrable for all $n \geq 0$.
- A **martingale** is an *adapted, integrable random process* $(X_n)_{n \geq 0}$ s.t. for all $n \geq 0$,

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n \quad \text{a.s.}$$

In the case $\mathbb{E}[X_{n+1} | \mathcal{F}_n] \leq X_n$ a.s., $(X_n)_n$ is called a **super-martingale** and in the case $\mathbb{E}[X_{n+1} | \mathcal{F}_n] \geq X_n$ a.s., $(X_n)_n$ is called a **sub-martingale**.

Optional Stopping

- A random variable $T : \Omega \rightarrow \{0, 1, 2, \dots\} \cup \{\infty\}$ is a **stopping time** if $\{T \leq n\} \in \mathcal{F}_n$ for all $n \geq 0$.
- For a stopping time T , we set $\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for all } n \geq 0\}$. It is easy to check \mathcal{F}_T is indeed a σ -algebra and that if $T(\omega) = n$ for all $\omega \in \Omega$, then T is a stopping time and $\mathcal{F}_T = \mathcal{F}_n$.
- Given X , define $X_T(\omega) = X_{T(\omega)}(\omega)$ whenever $T(\omega) < \infty$ and define the **stopped process** X^T by

$$X_n^T(\omega) = X_{T(\omega) \wedge n}(\omega) \quad \text{for } n \geq 0$$

Proposition 2.2.1.) Let X be an adapted process. Let S, T be stopping times for X . Then

- (a) $S \wedge T$ is a stopping time for X .
- (b) \mathcal{F}_T is a σ -algebra.

- (c) If $S \leq T$ then $\mathcal{F}_S \subset \mathcal{F}_T$.
- (d) $X_T 1_{T < \infty}$ is an \mathcal{F}_T -measurable random variable.
- (e) X^T is adapted.
- (f) If X is integrable, then X^T is also integrable.

proof)

- (a) $\{S \wedge T \leq n\} = \{S \leq n\} \cup \{T \leq n\} \in \mathcal{F}_n$ for all $n \geq 0$, so $S \wedge T$ is a stopping times
- (b) Directly from the definition, we see that $\phi \mathcal{F}_T$. Also, given $A \in \mathcal{F}_T$ and a sequence $(A_m)_m \subset \mathcal{F}_T$, we have

$$\begin{aligned} A^c \cap \{T \leq n\} &= \{T \leq n\} - A \cap \{T \leq n\} \in \mathcal{F}_n \Rightarrow A^c \in \mathcal{F}_T \\ (\cup_m A_m) \cap \{T \leq n\} &= \cup_m (A_m \cap \{T \leq n\}) \in \mathcal{F}_n \Rightarrow \cup_m A_m \in \mathcal{F}_T \end{aligned}$$

hence \mathcal{F}_T is a σ -algebra.

- (c) Let $A \in \mathcal{F}_S$. Then $A \cap \{T \leq n\} = A \cap \{S \leq n\} \cap \{T \leq n\} \in \mathcal{F}_n$, hence $A \in \mathcal{F}_T$.
- (d) For each $t \in \mathbb{R}$, we have $\{X_T 1_T > t\} = \cup_m \{X_m > t, T = n\}$ so for any $n \geq 0$,

$$\{X_T 1_T > t\} \cap \{T \leq n\} = \cup_{m=1}^n \{X_m > t, T = n\} \in \mathcal{F}_n$$

and so $X_T 1_T$ is \mathcal{F}_T -measurable.

- (e) By definition of being a stopping time, for any $t \in \mathbb{R}$,

$$\{(X^T)_n > t\} = \{T > n, X_n > t\} \cup \left(\cup_{m=0}^n \{T = m, X_m > t\} \right) \in \mathcal{F}_n$$

so X^T is adapted.

- (f) First consider the case where X is non-negative integrable. Then

$$\mathbb{E}(X_n^T) = \mathbb{E}(\mathbb{E}(X_n^T | T)) = \sum_{m \geq n} \mathbb{P}(T = m) \mathbb{E}(X_m) + \mathbb{P}(T > n) \mathbb{E}(X_n) < \infty$$

for any n , so we have the result for non-negative X .

For the general case, divide X into a non-negative and a negative part.

(End of proof) \square

Theorem 2.2.2) (Optional stopping theorem) Let X be a super-martingale and let S, T be bounded stopping times with $S \leq T$ a.s. Then

$$\mathbb{E}[X_T] \leq \mathbb{E}[X_S]$$

proof) Fix $n \geq 0$ such that $T \leq n$ a.s. Then

$$\begin{aligned} X_T &= X_S + \sum_{S \leq k < T} X_{k+1} - X_k \\ &= X_S + \sum_{k=0}^n (X_{k+1} - X_k) 1_{S \leq k < T} \end{aligned}$$

Now $\{S \leq k\}$ is in \mathcal{F}_k and $\{T > k\}$ is in \mathcal{F}_k , so

$$\begin{aligned} \mathbb{E}[(X_{k+1} - X_k) 1_{S \leq k < T}] &= \mathbb{E}[\mathbb{E}[(X_{k+1} - X_k) 1_{S \leq k < T} | \mathcal{F}_k]] \\ &= \mathbb{E}[\mathbb{E}[X_{k+1} - X_k | \mathcal{F}_k] 1_{S \leq k < T}] \end{aligned}$$

but since (X_n) was a super-martingale, $\mathbb{E}[X_{k+1} - X_k | \mathcal{F}_k] \leq 0$ a.s. and therefore $\mathbb{E}[(X_{k+1} - X_k) 1_{S \leq k < T}] \leq 0$ a.s. Hence $\mathbb{E}(X_T) \leq \mathbb{E}(X_S)$.

(End of proof) \square

•Note that X is a sub-martingale *if and only if* $(-X)$ is a super-martingale, and that X is a martingale *if and only if* X and $(-X)$ are super-martingales. Hence, we obtain sub-martingale and martingale versions of the theorem :

$$\begin{aligned} \text{If } (X_n) \text{ is a sub-martingale, } \mathbb{E}[X_T] &\geq \mathbb{E}[X_S] \\ \text{If } (X_n) \text{ is a martingale, } \mathbb{E}[X_T] &= \mathbb{E}[X_S] \end{aligned}$$

Theorem 2.2.3.) Let X be an adapted integrable process. Then the followings are equivalent.

- (a) X is a super-martingale.
- (b) for all bounded stopping times T and stopping time S ,

$$\mathbb{E}(X_T|\mathcal{F}_S) \leq X_{S \wedge T} \quad \text{a.s.},$$

- (c) for all stopping times T , the stopped process X^T is a super-martingale,
- (d) for all bounded stopping times T and all stopping times S with $S \leq T$ a.s.,

$$\mathbb{E}(X_T) \leq \mathbb{E}(X_S)$$

★ The theorem gives an inverse statement of the optional stopping theorem.

proof)

(a) \Rightarrow (b) Suppose X is a super-martingale and S, T are stopping times. Let $T \leq n$, for some $n < \infty$. Then

$$X_T = X_{S \wedge T} + \sum_{k=0}^T (X_{k+1} - X_k) 1_{S \leq k < T} \dots \dots (*)$$

Let $A \in \mathcal{F}_S$. Then $A \cap \{S \leq k\} \in \mathcal{F}_k$ and $\{T > k\} \in \mathcal{F}_k$ so

$$\mathbb{E}[(X_{k+1} - X_k) 1_{S \leq k < T} 1_A] = \mathbb{E}[\mathbb{E}[(X_{k+1} - X_k) 1_{S \leq k < T} 1_A | \mathcal{F}_k]] \leq 0$$

and

$$\begin{aligned} \mathbb{E}[(X_T - X_{S \wedge T}) 1_A] &= \mathbb{E}\left[\sum_{n=0}^T (X_{k+1} - X_k) 1_{S \leq k < T} 1_A\right] \leq 0 \\ \Rightarrow \mathbb{E}[X_T 1_A] &\leq \mathbb{E}[X_{S \wedge T} 1_A] \end{aligned}$$

But since this inequality is true for any $A \in \mathcal{F}_S$ and noting that $X_{S \wedge T} \in \mathcal{F}_S$, we see

$$\mathbb{E}[X_T | \mathcal{F}_S] \leq X_{S \wedge T} \quad \text{a.s.}$$

The implications (b) \Rightarrow (c) and (c) \Rightarrow (d) are obvious.

(d) \Rightarrow (a) Let $m \leq n$ and $A \in \mathcal{F}_n$. Set $T = m 1_A + n 1_{A^c}$. Then T is a stopping with $T \leq n$. Then

$$\mathbb{E}(X_n 1_A - X_m 1_A) = \mathbb{E}(X_n) - \mathbb{E}(X_T) \leq 0$$

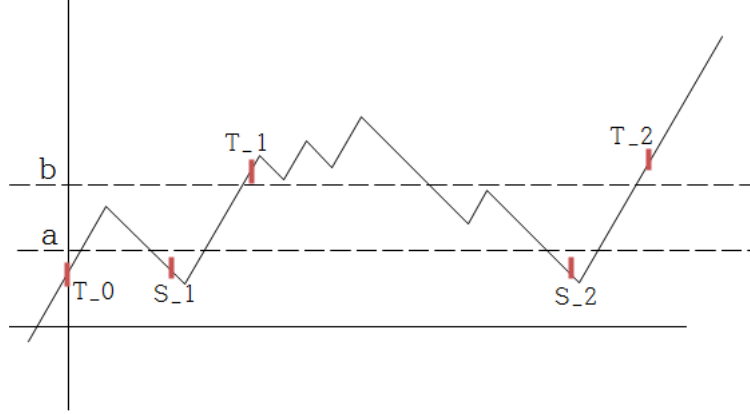
(note, if $\omega \in A$ then $(X_n 1_A - X_m 1_A)(\omega) = X_n(\omega) - X_m(\omega)$ and 0 otherwise) so

$$\mathbb{E}[X_n | \mathcal{F}_m] \leq X_m$$

(End of proof) \square

2.3. Doob's upcrossing inequality

- Let X be a random process and let $a, b \in \mathbb{R}$ s.t. $a < b$. Fix $\omega \in \Omega$. By an **upcrossing** of $[a, b]$ by $X(\omega)$, we mean an interval of times $\{j, j+1, \dots, k\}$ s.t. $X_j(\omega) < a$, $X_k(\omega) > b$.
- Write $U_n[a, b](\omega)$ for the number of disjoint upcrossings contained in $\{0, 1, \dots, n\}$, and $U_n[a, b] \nearrow U[a, b]$ as $n \rightarrow \infty$.



Theorem 2.3.1.) (Doob's upcrossing inequality) Let X be a *super-martingale*. Then

$$(b - a)\mathbb{E}[U[a, b]] \leq \sup_{n \geq 0} \mathbb{E}[(X_n - a)^-]$$

(Recall, $x^- = (-x) \vee 0$)

In fact, in this theorem, we prove $(b - a)\mathbb{E}[U_n[a, b]] \leq \mathbb{E}[(X_n - a)^-]$.

proof) Set $T_0 = 0$ and define recursively for $k \geq 0$,

$$S_{k+1} = \inf\{m \geq T_k : X_m < a\}, \quad T_{k+1} = \sup\{m \geq S_{k+1} : X_m > b\}$$

Note that if $T_k < \infty$, then $\{S_k, S_k + 1, \dots, T_k\}$ is an upcrossing of $[a, b]$ by X , and T_k is the time of completion of the k -th upcrossing. Also note that $U_n[a, b] \leq n$. For $m \leq n$, we have

$$\{U_n[a, b] = m\} = \{T_m \leq n < T_{m+1}\}$$

On this event,

$$X_{T_k \wedge n} - X_{S_k \wedge n} = \begin{cases} X_{T_k} - X_{S_k} \geq b - a & \text{if } k \leq m \\ X_n - X_{S_k} \geq X_n - a & \text{if } k \geq m + 1, S_{m+1} \leq n \\ 0 & \text{otherwise} \end{cases}$$

Hence

$$\begin{aligned} \sum_{k=1}^n (X_{T_k \wedge n} - X_{S_k \wedge n}) &\geq (b - a)U_n[a, b] + X_n - a \\ &\geq (b - a)U_n[a, b] - (X_n - a)^- \end{aligned}$$

Since X is a super-martingale and $T_k \wedge n$ and $S_k \wedge n$ are *bounded stopping times* with $S_k \leq T_k$, by optional stopping theorem, we have

$$\mathbb{E}(X_{T_k \wedge n}) \leq \mathbb{E}(X_{S_k \wedge n})$$

By $\mathbb{E}(\sum_{k=1}^n (X_{T_k \wedge n} - X_{S_k \wedge n})) \leq 0$ we get

$$(b - a)\mathbb{E}(U_n[a, b]) \leq \mathbb{E}[(X_n - a)^-]$$

Apply monotone convergence, with $n \rightarrow \infty$, then we are done.

(End of proof) \square

This theorem does not seem to have any significance at the moment, but it will turn out to be important later on.

2.4. Doob's maximal inequalities.

Define $X_n^* = \sum_{k \geq n} |X_k|$

In the next two theorems, we see that the martingale (or sub-martingale) property allows us to obtain estimates on this X_n^* in terms of expectations for X_n .

Theorem 2.4.1) (Doob's maximal inequality) Let X be a *martingale* or a *non-negative sub-martingale*. Then for all $\lambda \geq 0$,

$$\lambda \mathbb{P}(X_n^* \geq \lambda) \leq \mathbb{E}(|X_n| 1_{\{X_n^* \geq \lambda\}}) \leq \mathbb{E}(|X_n|)$$

proof) If X is a martingale, then $|X|$ is a non-negative sub-martingale. It suffices to consider the case where X is a non-negative sub-martingale.

Set $T = \inf\{k \geq 0 : X_k \geq \lambda\} \wedge n$. Then T is a stopping time and $T \leq n$, so by optional stopping, has

$$\begin{aligned} \mathbb{E}(X_n) &\geq \mathbb{E}(X_T) = \mathbb{E}(X_T 1_{X_n^* \geq \lambda}) + \mathbb{E}(X_T 1_{X_n^* < \lambda}) \\ &= \mathbb{E}(\lambda 1_{X_n^* \geq \lambda}) + \mathbb{E}(X_n 1_{X_n^* < \lambda}) \end{aligned}$$

and

$$\mathbb{E}(X_n 1_{X_n^* \geq \lambda}) \geq \lambda \mathbb{P}(X_n^* \geq \lambda)$$

(End of proof) \square

Theorem 2.4.2) (Doob's L^p -inequality) Let X be a *martingale* or a *non-negative sub-martingale*. Then, for all $p > 1$ and $q = p/(p-1)$, we have

$$\|X_n^*\|_p \leq q \|X_n\|_q$$

proof) Again, it suffices to consider when X is a non-negative sub-martingale. Fix $k < \infty$. Then

$$\begin{aligned} \mathbb{E}[(X_n^* \wedge k)^p] &= \mathbb{E} \int_0^k p \lambda^{p-1} 1_{\{X_n^* \geq \lambda\}} d\lambda \quad (\text{integration by parts}) \\ &= \int_0^k p \lambda^{p-1} \mathbb{P}(X_n^* \geq \lambda) d\lambda \quad (\text{Fubini}) \\ &\leq \int_0^k p \lambda^{p-2} \mathbb{E}(X_n 1_{X_n^* \geq \lambda}) d\lambda \quad (\text{Doob's maximal inequality}) \\ &= \frac{p}{p-1} \mathbb{E}(X_n (X_n^* \wedge k)^{p-1}) \\ &\leq q \|X_n\|_p \|X_n^* \wedge k\|_p^{p-1} \quad (\text{H\"older's inequality}) \end{aligned}$$

Hence, $\|X_n^* \wedge k\|_p \leq q \|X_n\|_p$. Apply monotone convergence theorem with $k \rightarrow \infty$, then we have the desired result.

(End of proof) \square

Doob's maximal and L^p inequalities have different versions which apply under the same hypothesis to

$$X^* = \sum_{n \geq 0} |X_n|$$

since $X_n^* \nearrow X^*$. Letting $n \rightarrow \infty$ in Doob's maximal inequality gives

$$\lambda \mathbb{P}(X^* \geq \lambda) \leq \lim_{n \rightarrow \infty} \lambda \mathbb{P}(X_n^* \geq \lambda) \leq \sup_{n \geq 0} \mathbb{E}(|X_n|)$$

We can then replace $\lambda \mathbb{P}(X^* > \lambda)$ by $\lambda \mathbb{P}(X^* \geq \lambda)$ by taking limits from the right in λ .

Similarly, for $p \in (1, \infty)$ by monotone convergence,

$$\|X^*\|_p \leq q \sup_{n \geq 0} \|X_n\|_p$$

=====
(19th October, Friday)

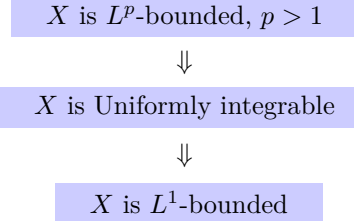
2.5. Doob's martingale convergence theorems

We are going to study three different martingale convergence theorems. They are all important.

- We say that a random process X is **L^p -bounded** if $\sum_{n \geq 0} \|X_n\|_p < \infty$.
- We say that X is **uniformly integrable** if

$$\sup_{n \geq 0} \mathbb{E}(|X_n| 1_{|X_n| > \lambda}) \rightarrow 0 \quad \text{as } \lambda \rightarrow \infty$$

- If X is L^p bounded for some $p > 1$, then this implies that X is uniformly integrable. This again implies that X is L^1 bounded. The first implication follows from Hölder inequality. The second implication is true because $\mathbb{E}(|X_n|) = \mathbb{E}(|X_n| 1_{|X_n| \leq \lambda}) + \mathbb{E}(|X_n| 1_{|X_n| > \lambda}) \leq \lambda + \mathbb{E}(|X_n| 1_{|X_n| > \lambda})$.



Theorem 2.5.1 (*Almost sure martingale convergence theorem*) Let X be an L^1 -bounded super-martingale. Then there exists an integrable and \mathcal{F}_∞ -measurable random variable X_∞ such that

$$X_n \rightarrow X \quad \text{a.s. as } n \rightarrow \infty$$

proof) For a sequence of real numbers $(x_n)_{n \geq 0}$, as $n \rightarrow \infty$, $(x_n)_n$ either converges or $|x_n| \rightarrow \infty$, or $\liminf_n x_n < \limsup_n x_n$. In the last case, since the rationals are dense in \mathbb{R} , there exist $a, b \in \mathbb{Q}$ such that $\liminf_n x_n < a < b < \limsup_n x_n$.

Set $\Omega_0 = \Omega_\infty \cap (\bigcap_{a, b \in \mathbb{Q}, a < b} \Omega_{a, b})$ where $\Omega_\infty = \{\liminf |X_n| < \infty\}$, $\Omega_{a, b} = \{U[a, b] < \infty\}$ (Recall that $U[a, b]$ is the number of upcrossings). Then $X_n(\omega)$ converges for all $\omega \in \Omega_0$. By Fatous' lemma,

$$\mathbb{E}(\liminf |X_n|) \leq \liminf \mathbb{E}|X_n| < \infty$$

so this implies $\mathbb{P}(\Omega_\infty) = 1$. By Doob's inequality, for $a < b$, has

$$(b - a)\mathbb{E}(U[a, b]) \leq |a| + \sup_{n \geq 0} \mathbb{E}|X_n| < \infty$$

and therefore $\mathbb{P}(\Omega_{a, b}) = 1$. Putting this together, we deduce that $\mathbb{P}(\Omega_0) = 1$, and we can find a random variable X_∞ defined by

$$X_\infty = \lim_{n \rightarrow \infty} X_n 1_{\Omega_0}$$

Then $X_n \rightarrow X_\infty$ a.s. Also X_∞ is \mathcal{F}_∞ -measurable and $|X_\infty| \leq \liminf |X_n|$ so $\mathbb{E}(|X_\infty|) < \infty$. Hence X_∞ is integrable.

(End of proof) \square

Remark : Every non-negative integrable super-martingale is L^1 -bounded, hence it converges a.s.

Theorem 2.5.2 (L^1 martingale convergence theorem) Let $(X_n)_{n \geq 0}$ be a uniformly integrable martingale. Then there exists a random variable $X_\infty \in L^1(\mathcal{F}_\infty)$ such that

$$X_n \xrightarrow{n \rightarrow \infty} X_\infty \quad \text{a.s. and in } L^1$$

Moreover, $X_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ a.s. for all $n \geq 0$.

Conversely, for all $Y \in L^1(\mathcal{F}_\infty)$, on choosing version X_n of $\mathbb{E}(Y | \mathcal{F}_n)$ for all n , we obtain a uniformly integrable martingale $(X_n)_{n \geq 0}$ such that

$$X_n \xrightarrow{n \rightarrow \infty} Y \quad \text{a.s. and in } L^1$$

We can think of this theorem as establishing the bijection

$$\text{unif. integrable martingale/a.s.} \leftrightarrow L^1(\mathcal{F}_\infty)$$

proof) Let $(X_n)_{n \geq 0}$ be a uniformly integrable martingale. By the almost sure martingale convergence theorem, there exists $X_\infty \in L^1(\mathcal{F}_\infty)$ s.t. $X_n \rightarrow X_\infty$ a.s. Since X is uniformly integrable, it also follows that $X_n \rightarrow X_\infty$ in L^1 . (see PM, Thm 2.5.1. and 6.2.3.)

Next, for $m \geq n$,

$$\begin{aligned} \|X_n - \mathbb{E}(X_\infty | \mathcal{F}_n)\|_1 &= \|\mathbb{E}(X_m - X_\infty | \mathcal{F}_n)\|_1 \\ &= \|X_m - X_\infty\|_1 \rightarrow 0 \quad \text{as } m \rightarrow \infty \end{aligned}$$

Hence $X_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ a.s.

For the converse statement, suppose $Y \in L^1(\mathcal{F}_\infty)$ and let X_n be a version of $\mathbb{E}(Y | \mathcal{F}_n)$ for all n . Then $(X_n)_{n \geq 0}$ is a martingale by the tower property, and is uniformly integrable by **Lemma 1.5.1**. Hence there exists $X_\infty \in L^1(\mathcal{F}_\infty)$ such that $X_n \rightarrow X_\infty$ a.s. and in L^1 . For all $n \geq 0$ and all $A \in \mathcal{F}_n$, we have

$$\mathbb{E}(X_\infty 1_A) = \lim_{m \rightarrow \infty} \mathbb{E}(X_m 1_A) = \lim_{n \leq m \rightarrow \infty} \mathbb{E}(\mathbb{E}(Y 1_A | \mathcal{F}_m)) = \mathbb{E}(Y 1_A)$$

where the second equality follows because $\mathbb{E}(X_m | \mathcal{F}_n) = \mathbb{E}(Y | \mathcal{F}_n)$. Now $X_\infty, Y \in L^1(\mathcal{F}_\infty)$ and $\cup_n \mathcal{F}_n$ is a π -system generating \mathcal{F}_∞ . Hence, by Dynkin's lemma,

$$X_\infty = Y \quad \text{a.s.}$$

(End of proof) \square

Theorem 2.5.3) (L^p -martingale convergence theorem) Let $p \in (1, \infty)$. Let $(X_n)_{n \geq 0}$ be an L^p -bounded martingale. Then there exists a random variable $X_\infty \in L^p(\mathcal{F}_\infty)$ s.t.

$$X_n \rightarrow X_\infty \quad \text{a.s. and in } L^p$$

Moreover, $X_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ a.s. for all $n \geq 0$.

Conversely, for all $Y \in L^p(\mathcal{F}_\infty)$, on choosing a version X_n of $\mathbb{E}(Y | \mathcal{F}_n)$ for all n , we obtain an L^p -bounded martingale such that $X_n \rightarrow Y$ a.s. and in L^p .

This is very similar to the statement of L^1 -martingale convergence theorem. Indeed, the proof is also very similar.

proof) Let (X_n) be an L^p -bounded martingale. By *a.s. martingale convergence theorem*, there exists $X_\infty \in L^1(\mathcal{F}_\infty)$, $X_n \rightarrow X_\infty$ a.s.

By *Doob's L^p -inequality*, $\|X^*\|_p \leq q \sup_{n \geq 0} \|X_n\|_p < \infty$, where $X^* = \sup_{n \geq 0} |X_n|$. Also, since $|X_n - X_\infty|^p \leq (2X^*)^p$ for all n , we may apply dominated convergence theorem to deduce that $X_n \rightarrow X_\infty$ in L^p . Then $X_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ a.s. for all n , as in the L^1 -convergence.

For the converse statement, suppose $Y \in L^p(\mathcal{F}_\infty)$ and let X_n be a version of $\mathbb{E}(Y | \mathcal{F}_n)$. Then $(X_n)_{n \geq 0}$ is a martingale by the tower property and by Jensen inequality,

$$\|X_n\|_p = \|\mathbb{E}(Y | \mathcal{F}_n)\|_p \leq \|Y\|_p$$

Let $X_n \rightarrow X_\infty$ a.s. and in L^p for $X_\infty \in L^p(\mathcal{F}_\infty)$, using the previous part. Then proceed as in the proof of L^1 -convergence to prove that in fact $Y = X_\infty$ a.s.

(End of proof) \square

=====
(22nd October, Monday)

Recall that, for a stopping time T and a random process X , X_T has been defined only on $\{T < \infty\}$. Given an almost sure limit X_∞ for X , we define $X_T = X_\infty$ on $\{T = \infty\}$. Then the optional stopping theorem extends to all stopping times for uniformly integrable martingales.

Theorem 2.5.5.) Let X be a uniformly integrable martingale and let T be any stopping time. Then $\mathbb{E}(X_T) = \mathbb{E}(X_0)$. Moreover, for all stopping time S and T , we have

$$\mathbb{E}(X_T|\mathcal{F}_S) = X_{S \wedge T} \quad \text{a.s.}$$

This theorem is an extension of Optional stopping theorem, **Theorem 2.2.2** and **Theorem 2.2.3**.

proof) By the optional stopping time theorem and **2.2.3**, when applied to the bounded stopping time $T \wedge n$, we have

$$\begin{aligned} \mathbb{E}(X_{T \wedge n}) &= \mathbb{E}(X_0) \\ \mathbb{E}(X_{T \wedge n}|\mathcal{F}_S) &= X_{S \wedge T \wedge n} \end{aligned}$$

In order to get the claim by letting $n \rightarrow \infty$, we need to prove $X_{T \wedge n} \rightarrow X_T$ a.s. and in L^1 . This will imply that

$$\mathbb{E}(X_{T \wedge n}|\mathcal{F}_S) \rightarrow \mathbb{E}(X_T|\mathcal{F}_S) \quad \text{in } L^1$$

Claim : $X_{T \wedge n} \rightarrow X_T$ a.s. and in L^1

proof) By the L^1 martingale convergence theorem, there exists $X_\infty \in L^1(\mathcal{F}_\infty)$ s.t. $X_n \rightarrow X_\infty$ a.s. and in L^1 and $X_n = \mathbb{E}(X_\infty|\mathcal{F}_n)$. This implies $X_{T \wedge n} \rightarrow X_T$ a.s. as $n \rightarrow \infty$. (if $T < \infty$, the convergence trivial, and in the case $T = \infty$, the convergence justified the previous statement). Since $F_{T \wedge n} \subset F_n$, by **Theorem 2.2.3**. and the tower property we have

$$X_{T \wedge n} = \mathbb{E}(X_n|\mathcal{F}_{T \wedge n}) = \mathbb{E}(X_\infty|\mathcal{F}_{T \wedge n})$$

By **Lemma 1.5.1**, $(X_{T \wedge n})_{n \geq 0}$ is uniformly integrable. Hence

$$X_{T \wedge n} \rightarrow X_T \quad \text{in } L^1$$

(End of proof) \square

Backward martingale

- A **backward filtration** $(\hat{\mathcal{F}}_n)_{n \geq 0}$ is a sequence of σ -algebras such that $\mathcal{F} \supset \hat{\mathcal{F}}_n \supset \hat{\mathcal{F}}_{n+1}$.
- This also defines $\hat{\mathcal{F}}_\infty = \bigcap_{n \geq 0} \hat{\mathcal{F}}_n$

Theorem 2.5.4.) (*Backward martingale convergence theorem*) For all $Y \in L^1(\mathcal{F})$, we have

$$\mathbb{E}(Y|\hat{\mathcal{F}}_n) \rightarrow \mathbb{E}(Y|\hat{\mathcal{F}}_\infty) \quad \text{a.s. and in } L^1 \quad \text{as } n \rightarrow \infty$$

Note that we do not need a uniformly integrability condition, because our assumption of backward filtration already implies uniform convergences.

proof) Write $X_n = \mathbb{E}(Y|\hat{\mathcal{F}}_n)$ for all $n \geq 0$. Fix $n \geq 0$, by the Tower property, $(X_{n-k})_{0 \leq k \leq n}$ is a martingale for the filtration $(\hat{\mathcal{F}}_{n-k})_{0 \leq k \leq n}$. For $a < b$, the number $U_n[0, \infty]$ of upcrossings of $[a, b]$ by $(X_k)_{0 \leq k \leq n}$ equals the number of upcrossings of $[-b, -a]$ by the process $(-X_{n-k})_{0 \leq k \leq n}$. Hence by (the note on) **Theorem 2.3.1**,

$$(b - a)\mathbb{E}(U_n[a, b]) \leq \mathbb{E}((X_0 - b)^+)$$

and so by monotone convergence,

$$(b - a)\mathbb{E}(U[a, b]) \leq \mathbb{E}((X_0 - b)^+) \leq \mathbb{E}(|X|) + |b| \leq \mathbb{E}(|Y|) + |b| < \infty$$

where the third inequality follows because of Jensen's inequality. Also,

$$\mathbb{E}(\liminf |X_n|) \leq \liminf \mathbb{E}|X_n| \leq \mathbb{E}|Y| < \infty$$

With these properties in hand, we can apply the same proof used to prove almost sure martingale convergence theorem to show that $\mathbb{P}(\hat{\Omega}_0) = 1$, where $\hat{\Omega}_0 = \{X_n \text{ converges as } n \rightarrow \infty\}$ - observe that $\hat{\Omega}_0 = \{\liminf_n |X_n| < \infty\} \cap (\bigcap_{a,b \in \mathbb{Q}, a < b} \{U[a, b] < \infty\})$ and we see that each set in the intersection has measure 1, and therefore $\mathbb{P}(\hat{\Omega}_0) = 1$.

Set $X_\infty = 1_{\hat{\Omega}_0} \lim_{n \rightarrow \infty} X_n$. Then $X_\infty \in L^1(\hat{\mathcal{F}}_\infty)$ and $X_n \rightarrow X_\infty$ a.s. Now $(X_n)_{n \geq 0}$ is uniformly integrable (by **Lemma 1.5.1**), so $X_n \xrightarrow{L^1} X_\infty$. Finally, for all $A \in \hat{F}_\infty$, we have

$$\mathbb{E}((X_\infty - \mathbb{E}(Y|\hat{\mathcal{F}}_\infty))1_A) = \lim_{n \rightarrow \infty} \mathbb{E}((X_n - Y)1_A) = 0$$

This implies $X_\infty = \mathbb{E}(Y|\hat{\mathcal{F}}_\infty)$ a.s.

(End of proof) \square

3. Applications of martingale theory

Sums of independent random variables

Let $S_n = X_1 + \dots + X_n$, where $(X_n)_{n \geq 0}$ is a sequence of independent random variables.

Theorem 3.1.1) (*Strong Law of Large Numbers*) Let $(X_n)_{n \geq 0}$ be a sequence of independent identically distributed (*i.i.d*) integrable random variables. Set $\mu = \mathbb{E}(X_1)$. Then

$$S_n/n \rightarrow \mu \quad \text{a.s. and in } L^1$$

proof) Define $\hat{\mathcal{F}}_n = \sigma(S_m : m \geq n)$, $\mathcal{T}_n = \sigma(X_m : m \geq n+1)$ and $\mathcal{T} = \bigcap_{n \geq 1} \mathcal{T}_n$. Then $\hat{\mathcal{F}}_n = \sigma(S_n, \mathcal{T}_n)$ and $(\hat{\mathcal{F}}_n)_{n \geq 1}$ is a backward filtration. Since $\sigma(X_1, S_n)$ is independent of \mathcal{T}_n , we have

$$\mathbb{E}(X_1|\hat{\mathcal{F}}_n) = \mathbb{E}(X_1|S_n) \quad \text{a.s.}$$

For $k \leq n$ and all Borel sets B , we have

$$\mathbb{E}(X_k 1_{\{S_n \in B\}}) = \mathbb{E}(X_1 1_{\{S_n \in B\}})$$

by symmetry $(X_k, S_n) \stackrel{d}{=} (X_1, S_n)$ in distribution, so $\mathbb{E}(X_k|S_n) = \mathbb{E}(X_1|S_n)$ a.s. But

$$\mathbb{E}(X_1|S_n) + \dots + \mathbb{E}(X_n|S_n) = \mathbb{E}(S_n|S_n) = S_n \quad \text{a.s.}$$

so $\mathbb{E}(X_1|\hat{\mathcal{F}}_n) = S_n/n$ almost surely. Then by backward martingale convergence theorem, has $S_n/n \rightarrow Y$ a.s. and in L^1 for some random variable Y . Then $Y \in \mathcal{T}$. By Kolmogorov's 0-1 law [PM **Theorem 2.6.1**], Y is almost surely a constant. Hence

$$Y = \mathbb{E}(Y) = \lim \mathbb{E}(S_n/n) = \mu \quad \text{a.s.}$$

where the second equality follows from L^1 convergence $S_n/n \rightarrow Y$.

(End of proof) \square

Since a.s. convergence implies convergence in probability, we have the following corollary.

Corollary 3.1.2) (*Weak law of large numbers*) Let $(X_n)_{n \geq 1}$ be a sequence of i.i.d. integrable r.v.. Set $\mu = \mathbb{E}(X_1)$. Then

$$\mathbb{P}(|\frac{S_n}{n} - \mu| > \epsilon) \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad \forall \epsilon > 0$$

=====
(24th October, Wednesday)

3.2. Non-negative martingale and change of measure

- Given a random variable X , \mathcal{F} -measurable with $X \geq 0$ and $\mathbb{E}(X) = 1$, we can define a new probability measure for $\tilde{\mathbb{P}}$ on \mathcal{F} by

$$\tilde{\mathbb{P}}(A) = \mathbb{E}(X1_A) \quad \forall A \in \mathcal{F}$$

Moreover, by [PM, Prop 3.1.4], given $\tilde{\mathbb{P}}$, this equation determines X uniquely, up to a.s. modification. We say $\tilde{\mathbb{P}}$ **has a density w.r.t.** \mathbb{P} and X is a version of the density.

- Let $(\mathcal{F}_n)_{n \geq 0}$ be a filtration in \mathcal{F} and assume $\mathcal{F} = \mathcal{F}_\infty$. Let $(X_n)_{n \geq 0}$ be an adapted random process, with $X_n \geq 0$ and $\mathbb{E}(X_n) = 1$ for all n . We can define, for each n , a new probability measure $\tilde{\mathbb{P}}_n$ on \mathcal{F}_n by

$$\tilde{\mathbb{P}}_n(A) = \mathbb{E}(X_n 1_A) \quad \forall A \in \mathcal{F}_n$$

Since we require each X_n to be \mathcal{F}_n -measurable, this equation determines X_n uniquely up to a.s. modification.

Proposition 3.2.1.) The measures $\tilde{\mathbb{P}}_n$ are consistent. That is

$$\tilde{\mathbb{P}}_{n+1}|_{\mathcal{F}_n} = \tilde{\mathbb{P}}_n \quad \forall n \quad \text{iff} \quad (X_n)_{n \geq 0} \quad \text{is a martingale}$$

Moreover, there is a measure $\tilde{\mathbb{P}}$ on \mathcal{F} , which has a density w.r.t \mathbb{P} such that

$$\tilde{\mathbb{P}}|_{\mathcal{F}_n} = \tilde{\mathbb{P}}_n \quad \forall n \quad \text{iff} \quad (X_n)_n \quad \text{is a uniformly integrable martingale}$$

proof) (The proof was an exercise.) For the first point,

$$\begin{aligned} \tilde{\mathbb{P}}_n(A) &= \tilde{\mathbb{P}}_{n+1}(A|\mathcal{F}_n) = \mathbb{E}(X_{n+1}1_A|\mathcal{F}_n) = \mathbb{E}(X_{n+1}|\mathcal{F}_n)1_A \quad \forall A \in \mathcal{F}_n \\ &\Leftrightarrow \mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n \quad \text{a.s.} \quad \Leftrightarrow (X_n) \text{ is a martingale} \end{aligned}$$

For the second point, suppose $\tilde{\mathbb{P}}|_{\mathcal{F}_n} = \tilde{\mathbb{P}}_n \quad \forall n$. Then $\tilde{\mathbb{P}}_n|_{\mathcal{F}_m} = (\tilde{\mathbb{P}}|_{\mathcal{F}_n})|_{\mathcal{F}_m} = \tilde{\mathbb{P}}_m$ whenever $n \geq m$, so we find that $(X_n)_n$ is a martingale. Since we assumed that $\mathbb{E}(X_n) = 1$ for all n , by almost everywhere martingale convergence theorem, we find a random variable X such that $X_n \rightarrow X$ a.s. Now for any $A \in \mathcal{F}$, we may find $N \geq 0$ such that $A \in \mathcal{F}_k$ for all $k \geq N$, so

$$\tilde{\mathbb{P}}(A) = \tilde{\mathbb{P}}(A|\mathcal{F}_k) = \tilde{\mathbb{P}}_k(A) = \mathbb{E}(X_k 1_A) \xrightarrow{k \rightarrow \infty} \mathbb{E}(X 1_A)$$

and therefore $\tilde{\mathbb{P}}(A) = \mathbb{E}(X 1_A)$. Hence for all $n \geq 0$, we have

$$\mathbb{E}(X 1_A|\mathcal{F}_n) = \tilde{\mathbb{P}}(A|\mathcal{F}_n) = \mathbb{E}(X_n 1_A) \quad \forall A \in \mathcal{F}_n$$

and therefore $X_n = \mathbb{E}(X|\mathcal{F}_n)$. This shows that $(X_n)_n$ is uniformly integrable martingale.

For the converse direction, assume that $(X_n)_n$ is uniformly integrable. Then by L^1 -martingale convergence theorem, we may find $X \in \mathcal{F}_\infty$ such that $X_n \rightarrow X$ in L^1 and a.s. Define $\tilde{\mathbb{P}}(A) = \mathbb{E}(X 1_A)$. Then $\tilde{\mathbb{P}}(A|\mathcal{F}_n) = \mathbb{E}(\mathbb{E}(X 1_A|\mathcal{F}_n)) = \mathbb{E}(X_n 1_A)$ for any $n \geq 0$ and $A \in \mathcal{F}_n$, and therefore $\tilde{\mathbb{P}}|_{\mathcal{F}_n} = \tilde{\mathbb{P}}_n$.

(End of proof) \square

Theorem 3.2.3) (*Radon-Nikodym theorem*) Let μ and ν be σ -finite measures on a measurable space (E, \mathcal{E}) . Then the followings are equivalent :

- $\nu(A) = 0$ for all $A \in \mathcal{E}$ such that $\mu(A) = 0$, i.e. ν is **absolutely continuous** with respect to μ .
- There exists a measurable function f on E such that $f \geq 0$ and $\nu(A) = \mu(f 1_A)$ for all $A \in \mathcal{E}$.

The function f which is unique up to modification μ -a.e. is called (a version of) the **Radon-Nikodym derivative** of ν with respect to μ . We write $f = d\nu/d\mu$ almost surely.

We will give a proof for the case *where \mathcal{E} is countably generated*. We assume there is a sequence $(G_n : n \in \mathbb{N})$ of subsets of E which generates \mathcal{E} . This holds, for example, whenever \mathcal{E} is the Borel σ -algebra of a topology with countable basis. A further martingale argument is required to prove the general case, but we omit it.

proof) The direction (b) \Rightarrow (a) is obvious. So we aim to prove (a) \Rightarrow (b)

By assumption, there is a countable partition of E by measurable sets on which both μ and ν are finite. (since μ, ν are σ -finite.) It suffices to show (b) holds on each of these sets, so we can reduce to the case where μ, ν are finite.

The case $\nu(E) = 0$ is clear, as we can just take $f \equiv 0$. So assume $\nu(E) > 0$. Then $\mu(E) > 0$ by (a). Write $\Omega = E$ and $\mathcal{F} = \mathcal{E}$ and consider the probability measures

$$\mathbb{P} = \mu/\mu(E) \quad \text{and} \quad \tilde{\mathbb{P}} = \nu/\nu(E) \quad \text{on } (\Omega, \mathcal{F})$$

It will suffice to show that there is a random variable $X \geq 0$ such that $\tilde{\mathbb{P}}(A) = \mathbb{E}(X1_A)$ for all $A \in \mathcal{F}$.

Set $\mathcal{F}_n = \sigma(G_k : k \leq n)$. There exists $m \in \mathbb{N}$ and a partition of Ω by events A_1, \dots, A_m such that $\mathcal{F}_n = \sigma(A_1, \dots, A_m)$ (e.g. choose $A_1 = G_1$, $A_2 = G_2 \setminus G_1$, $A_3 = G_3 \setminus (G_1 \cup G_2)$ and so on). Set $X_n = \sum_{j=1}^m a_j 1_{A_j}$ where $a_j = \tilde{\mathbb{P}}(A_j)/\mathbb{P}(A_j)$ if $\mathbb{P}(A_j) > 0$ and $a_j = 0$ otherwise. Then $X_n \geq 0$, $X_n \in \mathcal{F}_n$.

Observe that $(\mathcal{F}_n)_{n \geq 0}$ is a filtration and $(X_n)_{n \geq 0}$ is a non-negative martingale adapted to $(\mathcal{F}_n)_{n \geq 0}$. We will show that $(X_n)_{n \geq 0}$ is *uniformly integrable*. Once shown this, by the L^1 -martingale convergence theorem, there exists $X \geq 0$ such that $\mathbb{E}(X1_A) = \mathbb{E}(X_n1_A)$ for all $A \in \mathcal{F}_n$. Define a probability measure \mathbb{Q} on \mathcal{F} by

$$\mathbb{Q}(A) = \mathbb{E}(X1_A) \quad \forall A \in \mathcal{F}$$

Then $\mathbb{Q} = \tilde{\mathbb{P}}$ on $\cup_n \mathcal{F}_n$ which is a π -system generating \mathcal{F} . Hence $\mathbb{Q} = \tilde{\mathbb{P}}$ on \mathcal{F} , by uniqueness of extension.[PM, Thm 1.7.1], which implies (b). (In fact we do not have to use the extension theorem... we just have $\mathcal{F} = \cup_n \mathcal{F}_n$)

It remains to show that $(X_n)_n$ is uniformly integrable. Given $\epsilon > 0$, we can find $\delta > 0$ such that $\tilde{\mathbb{P}}(B) < \epsilon$ for all $B \in \mathcal{F}$ with $\mathbb{P}(B) < \delta$. (If not, there would be a sequence of sets $(B_n)_n \subset \mathcal{F}$ with $\mathbb{P}(B_n) < 2^{-n}$ and $\tilde{\mathbb{P}}(B_n) \leq \epsilon$ for all n . Then by Borel-Cantielli lemma, $\mathbb{P}(\limsup B_n) = 0$, but $\tilde{\mathbb{P}}(\limsup B_n) > \epsilon$, which contradicts (a)). Set $\lambda = 1/\delta$. Then by Markov inequality,

$$\mathbb{P}(X_n > \lambda) \leq \frac{\mathbb{E}(X_n)}{\lambda} = \frac{1}{\lambda} = \delta \quad \forall n$$

so $\mathbb{E}(X_n 1_{X_n > \lambda}) = \tilde{\mathbb{P}}(X_n > \lambda) < \epsilon$ for all n . Hence $(X_n)_n$ is uniformly integrable by its definition

(End of proof) \square