Lecture-25: Martingales

1 Martingales

Definition 1.1. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A **filtration** is an increasing sequence of σ-fields denoted by $\mathcal{F}_\bullet = (\mathcal{F}_n \subseteq \mathcal{F} : n \in \mathbb{N})$, with nth σ-field denoted by $\mathcal{F}_n$.

Definition 1.2. For a discrete stochastic process $X : \Omega \to \mathbb{R}^{\mathbb{N}}$, its **natural filtration** is defined as $\mathcal{F}_n \triangleq \sigma(X_1, \dots, X_n)$.

Definition 1.3. A random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ of random variables is said to be **adapted** to the filtration $\mathcal{F}_\bullet$ if $\sigma(X_n) \subseteq \mathcal{F}_n$ for all $n \in \mathbb{N}$.

Remark 1. For any random sequence X adapted to a filtration $\mathcal{F}_\bullet$, we also have $\sigma(X_1, \dots, X_n) \subseteq \mathcal{F}_n$ for each $n \in \mathbb{N}$.

Definition 1.4. A discrete stochastic process $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ is said to be a **martingale** with respect to the filtration $\mathcal{F}_\bullet$ if it satisfies the following three properties for each $n \in \mathbb{N}$.

i_ **Integrability.** $\mathbb{E}|X_n| < \infty$.

ii_ **Adaptability.** $\sigma(X_n) \subseteq \mathcal{F}_n$.

iii_ **Unbiasedness.** $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$.

If the equality in third condition is replaced by $\leqslant$ or $\geqslant$, then the process is called **supermartingale** or **submartingale**, respectively.

Corollary 1.5. *For a martingale X adapted to a filtration $\mathcal{F}_\bullet$, we have $\mathbb{E}X_n = \mathbb{E}X_1$ for each $n \in \mathbb{N}$.*

Example 1.6 (Simple random walk). Let $\xi : \Omega \to \mathbb{R}^{\mathbb{N}}$ be an independent random sequence with mean $\mathbb{E}\xi_i = 0$ and $\mathbb{E}|\xi_i| < \infty$ for each $i \in \mathbb{N}$. Let $\mathcal{F}_\bullet$ be the natural filtration of the random sequence ξ. Consider the random walk $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ with step-size sequence ξ such that $X_n \triangleq \sum_{i=1}^n \xi_i$ for each $n \in \mathbb{N}$, then X is adapted to $\mathcal{F}_\bullet$. From the linearity of expectation and the finiteness of finitely many individual terms, we have $\mathbb{E}|X_n| \leqslant \sum_{i=1}^n \mathbb{E}|\xi_i| < \infty$. Further, we have

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] = \mathbb{E}[X_n + \xi_{n+1} \mid \mathcal{F}_n] = X_n.$$

Thus, the random walk X is a martingale with respect to filtration $\mathcal{F}_\bullet$.

Example 1.7 (Product martingale). Let $\xi : \Omega \to \mathbb{R}^{\mathbb{N}}$ be an independent random sequence with mean $\mathbb{E}\xi_i = 1$ and $\mathbb{E}|\xi_i| < \infty$ for each $i \in \mathbb{N}$. Let $\mathcal{F}_\bullet$ be the natural filtration of random sequence ξ. Consider the random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined as $X_n \triangleq \prod_{i=1}^n \xi_i$ for each $n \in \mathbb{N}$, then X is adapted to $\mathcal{F}_\bullet$. From the independence and finiteness of finitely many individual terms, we have $\mathbb{E}|X_n| = \prod_{i=1}^n \mathbb{E}\xi_i < \infty$. Further, we have

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] = \mathbb{E}[X_n \xi_{n+1} | \mathcal{F}_n] = X_n.$$

Thus, the random sequence X is a martingale with respect to filtration $\mathcal{F}_\bullet$.

Example 1.8 (Branching process). Consider a population where each individual i can produce an independent random number of offsprings Z_i in its lifetime, with a common distribution $P : \mathbb{Z}_+ \to [0, 1]$ and finite mean $\mu \triangleq \sum_{j \in \mathbb{N}} j P_j < \infty$. Let X_n denote the size of the nth generation, which is same as the number of offsprings generated by $(n-1)$th generation. The discrete stochastic process $X : \Omega \to \mathbb{Z}_+^{\mathbb{N}}$ is called a branching process. Let $X_0 = 1$ and consider the natural filtration $\mathcal{F}_\bullet$ of X. We can write $X_n = \sum_{i=1}^{X_{n-1}} Z_i$. Conditioning on $\mathcal{F}_{n-1}$ yields,

$$\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] = \mathbb{E}[\sum_{i=1}^{X_{n-1}} Z_i \mid \mathcal{F}_{n-1}] = \mathbb{E}[\sum_{i\in\mathbb{N}} Z_i \mathbb{1}_{\{i \leqslant X_{n-1}\}} | \mathcal{F}_{n-1}] = \sum_{i\in\mathbb{N}} \mathbb{E}[Z_i \mid \mathcal{F}_{n-1}] \mathbb{1}_{\{i\leqslant X_{n-1}\}} = \sum_{i=1}^{X_{n-1}} \mu = \mu X_{n-1}.$$

Applying expectation on both sides, and by induction on n, we get $\mathbb{E}[X_n] = \mu^n$. Consider a positive random sequence $Y : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ defined by $Y_n \triangleq \frac{X_n}{\mu^n}$ for each $n \in \mathbb{N}$, adapted to $\mathcal{F}_\bullet$. Since X is a non-negative sequence, we have $\mathbb{E}|Y_n| = \mathbb{E}Y_n = 1$. Further,

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \frac{1}{\mu^{n+1}}\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = \frac{X_n}{\mu^n} = Y_n.$$

It follows that Y is a martingale with respect to filtration $\mathcal{F}_\bullet$.

Example 1.9 (Doob's Martingale). Consider an arbitrary random sequence $Y : \Omega \to \mathbb{R}^{\mathbb{N}}$ with associated natural filtration $\mathcal{F}_\bullet$, and an arbitrary random variable $Z : \Omega \to \mathbb{R}$ such that $\mathbb{E}|Z| < \infty$. We define a random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ as $X_n \triangleq \mathbb{E}[Z|\mathcal{F}_n]$ for each $n \in \mathbb{N}$. From the definition of conditional expectation, X is adapted to $\mathcal{F}_\bullet$. Further, from the Jensen's inequality for conditional expectation applied to the convex absolute function, we get $\mathbb{E}|X_n| \leqslant \mathbb{E}[\mathbb{E}[|Z| \mid \mathcal{F}_n]] = \mathbb{E}|Z| < \infty$. Further, from the tower property of conditional expectation

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_{n+1}]|\mathcal{F}_n] = \mathbb{E}[Z|\mathcal{F}_n] = X_n.$$

Thus, X is a martingale with respect to $\mathcal{F}_\bullet$, and called a *Doob-type* martingale.

Example 1.10 (Centralized Doob sequence). For any sequence of random variables $Y : \Omega \to \mathbb{R}^{\mathbb{N}}$ with $\mathbb{E}|Y_n| < \infty$ for all $n \in \mathbb{N}$ and its natural filtration $\mathcal{F}_\bullet$, the centralized random variable $Y_i - \mathbb{E}[Y_i \mid \mathcal{F}_{i-1}]$ has a zero mean for each $i \in \mathbb{N}$. Consider, the centralized zero mean sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined by $X_n \triangleq \sum_{i=1}^n (Y_i - \mathbb{E}[Y_i \mid \mathcal{F}_{i-1}])$ for each $n \in \mathbb{N}$. By the definition of condition expectation and filtration, the random sequence X is adapted to the filtration $\mathcal{F}_\bullet$. From the triangle inequality and the conditional Jensen's inequality applied to convex absolute function, we get

$$\mathbb{E}|X_n| \leqslant \sum_{i=1}^n \mathbb{E}|Y_i - \mathbb{E}[Y_i|\mathcal{F}_{i-1}]| \leqslant \sum_{i=1}^n \Big(\mathbb{E}|Y_i| + \mathbb{E}|\mathbb{E}[Y_i \mid \mathcal{F}_{i-1}]|\Big) \leqslant 2\sum_{i=1}^n \mathbb{E}|Y_i| < \infty.$$

Further, from the linearity and the tower property of conditional expectation, we have

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_n + Y_{n+1} - \mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] \mid \mathcal{F}_n] = X_n.$$

Thus, X is a martingale with respect to this filtration $\mathcal{F}_\bullet$, and called *centralized Doob martingale*.

Lemma 1.11. *Consider a martingale $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ adapted to a filtration $\mathcal{F}_\bullet = (\mathcal{F}_n \subseteq \mathcal{F} : n \in \mathbb{N})$ defined on the probability space $(\Omega, \mathcal{F}, P)$, and a convex function $f : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}|f(X_n)| < \infty$ for all $n \in \mathbb{N}$. Then, the random sequence $Y : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined by $Y_n \triangleq f(X_n)$ for each $n \in \mathbb{N}$, is a submartingale with respect to the filtration $\mathcal{F}_\bullet$.*

Proof. We observe that Y is adapted to the filtration $\mathcal{F}_\bullet$ and integrable by hypothesis. From the conditional Jensen's inequality applied to convex function f, we get

$$\mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n] \geqslant f(\mathbb{E}[X_{n+1} \mid \mathcal{F}_n]) = f(X_n).$$

□

Corollary 1.12. *Consider a random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined on the probability space $(\Omega, \mathcal{F}, P)$, with its natural filtration $\mathcal{F}_\bullet$. Let $a \in \mathbb{R}$ be a constant, and consider two random sequences $Y : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ and $Z : \Omega \to \mathbb{R}^{\mathbb{N}}$ generated by X, such that for each $n \in \mathbb{N}$,*

$$Y_n \triangleq (X_n - a)_+ = (X_n \vee a) - a, \qquad\qquad Z_n \triangleq X_n \wedge a.$$

i. *If X is a submartingale with respect to $\mathcal{F}_\bullet$, then so is Y with respect to $\mathcal{F}_\bullet$.*

ii. *If X is a supermartingale with respect to $\mathcal{F}_\bullet$, then so is Z with respect to $\mathcal{F}_\bullet$.*

Proof. Clearly, both sequences Y and Z are adapted to $\mathcal{F}_\bullet$. Defining $x \mapsto f(x) \triangleq (x-a)_+$ and $x \mapsto g(x) \triangleq x \wedge a$ for all $x \in \mathbb{R}$, we observe that f is convex and non-decreasing and g is concave and non-decreasing. The function f is positive, and hence $\mathbb{E}|f(X_n)| = \mathbb{E}f(X_n) \leqslant \mathbb{E}|X_n| + |a| < \infty$. We also observe that $\mathbb{E}|g(X_n)| \leqslant E|X_n| < \infty$.

i. From the conditional Jensen's inequality applied to the convex non-decreasing function f and the fact that $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \geqslant X_n$, we get $\mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n] \geqslant f(\mathbb{E}[X_{n+1} \mid \mathcal{F}_n]) \geqslant f(X_n)$.

ii. From the conditional Jensen's inequality applied to the concave non-decreasing function f and the fact that $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] \leqslant X_n$, we get $\mathbb{E}[g(X_{n+1}) \mid \mathcal{F}_n] \leqslant g(\mathbb{E}[X_{n+1} \mid \mathcal{F}_n]) \leqslant g(X_n)$.

□

1.1 Stopping Times

Consider a discrete filtration $\mathcal{F}_\bullet = (\mathcal{F}_n \subseteq \mathcal{F} : n \in \mathbb{Z}_+)$.

Definition 1.13. A positive integer valued, possibly infinite, random variable $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ is said to be a **random time** with respect to the filtration $\mathcal{F}_\bullet$, if the event $\{\tau = n\} \in \mathcal{F}_n$ for each $n \in \mathbb{N}$. If $P\{\tau < \infty\} = 1$, then the random time τ is said to be a **stopping time**.

Definition 1.14. A random sequence $H : \Omega \to \mathbb{R}^{\mathbb{N}}$ is **predictable** with respect to the the filtration $\mathcal{F}_\bullet$, if $\sigma(H_n) \subseteq \mathcal{F}_{n-1}$ for each $n \in \mathbb{N}$. For a process X adapted to $\mathcal{F}_\bullet$, we define

$$(H \cdot X)_n \triangleq \sum_{m=1}^{n} H_m (X_m - X_{m-1}).$$

Theorem 1.15. *Consider a supermartingale sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ and a predictable sequence $H : \Omega \to \mathbb{R}_+^{\mathbb{N}}$ with respect to a filtration $\mathcal{F}_\bullet$, where each H_n is non-negative and bounded. Then the random sequence $Y : \Omega \to \mathbb{R}^N$ defined by $Y_n \triangleq (H \cdot X)_n$ for each $n \in \mathbb{N}$ is a supermartingale with respect to $\mathcal{F}_\bullet$.*

Proof. From the definition of Y, it follows that Y is adapted to $\mathcal{F}_\bullet$. From the tower property of conditional expectation, and predictability, non-negativity, and boundedness of H, we obtain

$$\mathbb{E}|Y_n| \leqslant \sum_{m=1}^{n} \mathbb{E}[H_m \mathbb{E}[|X_m - X_{m-1}| \mid \mathcal{F}_{m-1}]] \leqslant \sup_{m \leqslant n} H_m \sum_{m=1}^{n} (\mathbb{E}|X_m| + \mathbb{E}|X_{m-1}|) < \infty.$$

Further, from the definition of Y, the predictability of H, and the supermartingale property of X,

$$\mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[H_{n+1}(X_{n+1} - X_n) + Y_n \mid \mathcal{F}_n] = H_{n+1}(\mathbb{E}[X_{n+1}|\mathcal{F}_n] - X_n) + Y_n \leqslant Y_n.$$

□

1.2 Stopped process

Definition 1.16. Consider a discrete stochastic process $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ adapted to a discrete filtration $\mathcal{F}_\bullet$. Let $\tau : \Omega \to \mathbb{N}$ be a random time for the filtration $\mathcal{F}_\bullet$, then the **stopped process** $X^\tau : \Omega \to \mathbb{R}^{\mathbb{N}}$ is defined for each $n \in \mathbb{N}$ as

$$X_n^\tau \triangleq X_{\tau \wedge n} = X_n \mathbb{1}_{\{n \leqslant \tau\}} + X_\tau \mathbb{1}_{\{n > \tau\}}.$$

Proposition 1.17. *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a martingale with a discrete filtration $\mathcal{F}_\bullet$. If $\tau : \Omega \to \mathbb{N}$ is an integer random time for the filtration $\mathcal{F}_\bullet$, then the stopped process X^τ is a martingale.*

Proof. Consider a random sequence $H : \Omega \to \{0,1\}^{\mathbb{N}}$ defined by $H_n \triangleq \mathbb{1}_{\{n \leqslant \tau\}}$ for each $n \in \mathbb{N}$. Then H is a non-negative and bounded sequence. Further H is predictable with respect to $\mathcal{F}_\bullet$, since the event

$$\{n \leqslant \tau\} = \{\tau > n-1\} = \{\tau \leqslant n-1\}^c = (\cup_{i=0}^{n-1} \{\tau = i\})^c = \cap_{i=0}^{n-1} \{\tau \neq i\} \in \mathcal{F}_{n-1}.$$

In terms of the non-negative, predictable, and bounded sequence H, we can write the stopped process

$$X_{\tau \wedge n} = X_0 + \sum_{m=1}^{\tau \wedge n} (X_m - X_{m-1}) = X_0 + \sum_{m=1}^{n} \mathbb{1}_{\{m \leqslant \tau\}} (X_m - X_{m-1}) = X_0 + (H \cdot X)_n.$$

From the previous theorem, it follows that X^τ is a martingale, and we have $\mathbb{E}X_{\tau \wedge n} = \mathbb{E}X_{\tau \wedge 1} = \mathbb{E}X_1$. □

Remark 2. For any martingale $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ and a stopping time $\tau : \Omega \to \mathbb{N}$ adapted to $\mathcal{F}_\bullet$, we have $\mathbb{E}X_{\tau\wedge n} = \mathbb{E}X_1$, for all $n \in \mathbb{N}$. Since τ is finite almost surely, it follows that the stopped process X^τ converges almost surely to X_τ, i.e. $P\{\lim_{n\in\mathbb{N}} X_{\tau\wedge n} = X_\tau\} = 1$.

We are interested in knowing under what conditions will we have convergence in mean.

Theorem 1.18 (Martingale stopping theorem). *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a martingale and $\tau : \Omega \to \mathbb{N}$ be a stopping time, both adapted to a common discrete filtration $\mathcal{F}_\bullet$. If either of the following conditions holds true.*

(i) *τ is bounded,*

(ii) *$X_{\tau\wedge n}$ is uniformly bounded,*

(iii) *$\mathbb{E}\tau < \infty$, and for some real positive K, we have $\sup_{n\in\mathbb{N}} \mathbb{E}[|X_n - X_{n-1}| \mid \mathcal{F}_{n-1}] < K$.*

Then X_τ is integrable and the stopped process X^τ converges in mean to X_τ, i.e. $\lim_{n\in\mathbb{N}} \mathbb{E}X_{\tau\wedge n} = \mathbb{E}X_\tau = \mathbb{E}X_1$,

Proof. We show this is true for all three cases.

(i) Let K be the bound on τ then for all $n \geqslant K$, we have $X_{\tau\wedge n} = X_\tau$, and hence it follows that $\mathbb{E}X_1 = \mathbb{E}X_{\tau\wedge n} = \mathbb{E}X_\tau$ for all $n \geqslant K$.

(ii) Dominated convergence theorem implies the result.

(iii) We can write the difference $X_{\tau\wedge n} - X_0 = \sum_{m=1}^{\tau} \mathbb{1}_{\{m\leqslant n\}}(X_m - X_{m-1})$ using the telescopic sum. From triangle inequality for the absolute function and the fact that $0 \leqslant \mathbb{1}_{\{m\leqslant n\}} \leqslant 1$, we can upper bound the difference $|X_{\tau\wedge n}| - |X_0| \leqslant |X_{\tau\wedge n} - X_0| \leqslant \sum_{m=1}^{\tau} |X_m - X_{m-1}|$. We define non-negative predictable sequence $H : \Omega \to \{0,1\}^{\mathbb{N}}$ as $H_n \triangleq \mathbb{1}_{\{n\leqslant\tau\}}$ for $n \in \mathbb{N}$. From the linearity of expectation, the monotone convergence theorem, the tower property of conditional expectation, predictability of H, and theorem hypothesis, we can upper bound the mean of this term as

$$\mathbb{E}\sum_{m=1}^{\tau} |X_m - X_{m-1}| = \sum_{m\in\mathbb{N}} \mathbb{E}[H_m \mathbb{E}[|X_m - X_{m-1}| \mid \mathcal{F}_{m-1}]] \leqslant K\mathbb{E}\sum_{m\in\mathbb{N}} H_m = K\mathbb{E}\tau.$$

Since τ is integrable, we observe that $X_{\tau\wedge n}$ is uniformly bounded by an integrable random variable. The result follows from dominated convergence theorem.

□

Corollary 1.19 (Wald's Equation). *If τ is a stopping time for the discrete* i.i.d. *random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ with $\mathbb{E}|X| < \infty$ and $\mathbb{E}\tau < \infty$, then*

$$\mathbb{E}\sum_{i=1}^{\tau} X_i = \mathbb{E}\tau\mathbb{E}X_1.$$

Proof. Let $\mu = \mathbb{E}X$ and define a random sequence $Z : \Omega \to \mathbb{R}^{\mathbb{N}}$ such that $Z_n \triangleq \sum_{i=1}^{n}(X_i - \mu)$ for each $n \in \mathbb{N}$, Then Z is a martingale adapted to natural filtration of X, and

$$\mathbb{E}[|Z_n - Z_{n-1}| \mid \mathcal{F}_{n-1}] = \mathbb{E}[|X_n - \mu|] \leqslant \mu + \mathbb{E}|X_1|.$$

Thus, $\sup_{n\in\mathbb{N}} \mathbb{E}[|Z_n - Z_{n-1}| \mid \mathcal{F}_{n-1}] < \infty$, and from the Martingale stopping theorem, we have $\mathbb{E}Z_\tau = \mathbb{E}Z_1 = 0$. The result follows from the observation that $\mathbb{E}[Z_\tau] = \mathbb{E}\sum_{i=1}^{\tau} X_i - \mu\mathbb{E}\tau$. □