

Lecture-26: Martingales

1 Martingales

A martingale is a type of stochastic process whose definition formalizes the concept of a fair game.

Definition 1.1. A discrete time stochastic process $X : \Omega \rightarrow \mathbb{R}^N$ is said to be a *martingale* with respect to the filtration $\mathcal{F}_\bullet$ if it satisfies the following three properties for each $n \in \mathbb{N}$,

- (a) *adaptability*: $\sigma(X_n) \subseteq \mathcal{F}_n$,
- (b) *integrability*: $\mathbb{E}|X_n| < \infty$,
- (c) *unbiasedness*: $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n$.

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

Corollary 1.2. 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 1.3 (Simple random walk). Let $\mathcal{F}_\bullet$ be the natural filtration of the random independent sequence $\xi : \Omega \rightarrow \mathbb{R}^N$ with mean $\mathbb{E}\xi_i = 0$ and $\mathbb{E}|\xi_i| < \infty$ for each $i \in \mathbb{N}$. Consider the random walk $X : \Omega \rightarrow \mathbb{R}^N$ 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 $\mathcal{F}_\bullet$. Further, from the linearity of expectation and the finiteness of finitely many individual terms, we have $\mathbb{E}|X_n| \leq \sum_{i=1}^n \mathbb{E}|\xi_i| < \infty$ for each $n \in \mathbb{N}$. In addition, we have for each $n \in \mathbb{N}$,

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

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

Example 1.4 (Product martingale). Let $\mathcal{F}_\bullet$ be the natural filtration of random independent sequence $\xi : \Omega \rightarrow \mathbb{R}^N$ with mean $\mathbb{E}\xi_i = 1$ and $\mathbb{E}|\xi_i| < \infty$ for each $i \in \mathbb{N}$. Consider the random sequence $X : \Omega \rightarrow \mathbb{R}^N$ defined as $X_n \triangleq \prod_{i=1}^n \xi_i$ for each $n \in \mathbb{N}$. Then X is adapted to $\mathcal{F}_\bullet$. Further, from the independence and finiteness of finitely many individual terms, we have $\mathbb{E}|X_n| = \prod_{i=1}^n \mathbb{E}\xi_i < \infty$ for each $n \in \mathbb{N}$. In addition, we have for each $n \in \mathbb{N}$,

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

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

Example 1.5 (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 $P \in \mathcal{M}(\mathbb{Z}_+)$ and finite mean $\mu \triangleq \sum_{j \in \mathbb{N}} jP_j < \infty$. Let X_n denote the size of the n th generation, which is same as the number of offsprings generated by $(n-1)$ th generation. The discrete stochastic process $X : \Omega \rightarrow \mathbb{Z}_+^{\mathbb{Z}_+}$ is called a branching process. Let $X_0 = 1$ and consider the natural filtration $\mathcal{F}_\bullet$ of X . We define $W : \Omega \rightarrow \mathbb{Z}_+^{\mathbb{Z}_+}$ to be a random walk with step-size sequence X such that $W_n \triangleq \sum_{i=0}^n X_i$ is the population until generation n for each $n \in \mathbb{Z}_+$. It follows that W is adapted to $\mathcal{F}_\bullet$ and we can write $X_n = \sum_{i=W_{n-2}+1}^{W_{n-1}} Z_i$ for each $n \in \mathbb{N}$. For any $n \in \mathbb{N}$, conditioning on $\mathcal{F}_{n-1}$ yields

$$\mathbb{E}[X_n | \mathcal{F}_{n-1}] = \mathbb{E}\left[\sum_{i=W_{n-2}+1}^{W_{n-1}} Z_i | \mathcal{F}_{n-1}\right] = \mathbb{E}\left[\sum_{i \in \mathbb{N}} Z_i \mathbb{1}_{(W_{n-2}, W_{n-1}]}(i) | \mathcal{F}_{n-1}\right] = \sum_{i \in \mathbb{N}} \mathbb{E}[Z_i | \mathcal{F}_{n-1}] \mathbb{1}_{(W_{n-2}, W_{n-1}]}(i).$$

From the definition of X , we observe that $\mathcal{F}_{n-1} \subseteq \sigma(X_0, Z_1, \dots, Z_{W_{n-1}})$, and since Z is an independent sequence Z_i is independent of $\mathcal{F}_{n-1}$ for $i > W_{n-1}$. It follows that $\mathbb{E}[Z_i | \mathcal{F}_{n-1}] = \mathbb{E}Z_i = \mu$ for $i > W_{n-1}$, and hence $\mathbb{E}[X_n | \mathcal{F}_{n-1}] = \mu X_{n-1}$. Applying expectation on both sides, and by induction on n , we get

$\mathbb{E}[X_n | \mathcal{F}_0] = X_0 \mu^n$. Consider a positive random sequence $Y : \Omega \rightarrow \mathbb{R}_+^{\mathbb{N}}$ defined for each $n \in \mathbb{N}$ as $Y_n \triangleq \frac{X_n}{\mu^n}$. Then, Y is adapted to $\mathcal{F}_\bullet$. Since X is a non-negative sequence, we have $\mathbb{E}[|Y_n| | \mathcal{F}_0] = \mathbb{E}[Y_n | \mathcal{F}_0] = X_0$. Further, for each $n \in \mathbb{N}$

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

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

Example 1.6 (Doob's Martingale). Consider an arbitrary random sequence $Y : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ with associated natural filtration $\mathcal{F}_\bullet$, and an arbitrary random variable $Z : \Omega \rightarrow \mathbb{R}$ such that $\mathbb{E}|Z| < \infty$. We define a random sequence $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ 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 $\mathbb{E}|X_n| \leq \mathbb{E}[\mathbb{E}[|Z| | \mathcal{F}_n]] = \mathbb{E}|Z| < \infty$. 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 1.7 (Centralized Doob sequence). Let $\mathcal{F}_\bullet$ be the natural filtration for a random sequence $Y : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ with $\mathbb{E}|Y_n| < \infty$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, we can define an associated centralized zero mean random variable $Y_n - \mathbb{E}[Y_n | \mathcal{F}_{n-1}]$. Hence, we can define a centralized zero mean sequence $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ for each $n \in \mathbb{N}$, as $X_n \triangleq \sum_{i=1}^n (Y_i - \mathbb{E}[Y_i | \mathcal{F}_{i-1}])$. 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

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

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

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

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

Lemma 1.8. Consider a martingale $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $\mathbb{E}|f(X_n)| < \infty$ for all $n \in \mathbb{N}$. Then, the random sequence $Y : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ defined for each $n \in \mathbb{N}$ as $Y_n \triangleq f(X_n)$, 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

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

□

Corollary 1.9. Consider a random sequence $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ 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 $Y : \Omega \rightarrow \mathbb{R}_+^{\mathbb{N}}$ and $Z : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ generated by X , such that for each $n \in \mathbb{N}$,

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

- (a) If X is a submartingale with respect to $\mathcal{F}_\bullet$, then so is Y with respect to $\mathcal{F}_\bullet$.
- (b) 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 $\mathbb{E}|f(X_n)| = \mathbb{E}f(X_n) \leq \mathbb{E}|X_n| + |a| < \infty$. We also observe that $\mathbb{E}|g(X_n)| \leq \mathbb{E}|X_n| < \infty$.

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

□

1.1 Stopping Times

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

Definition 1.10. A positive integer valued, possibly infinite, random variable $\tau : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$ 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 1.11. A random sequence $H : \Omega \rightarrow \mathbb{R}^\mathbb{N}$ is *predictable* with respect to 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 1.12. Consider a supermartingale sequence $X : \Omega \rightarrow \mathbb{R}^\mathbb{N}$ and a predictable sequence $H : \Omega \rightarrow \mathbb{R}_+^\mathbb{N}$ with respect to a filtration $\mathcal{F}_\bullet$, where each H_n is non-negative and bounded. Then the random sequence $Y : \Omega \rightarrow \mathbb{R}^\mathbb{N}$ 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

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

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

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

□

1.2 Stopped process

Definition 1.13. Consider a discrete stochastic process $X : \Omega \rightarrow \mathbb{R}^\mathbb{N}$ adapted to a discrete filtration $\mathcal{F}_\bullet$. Let $\tau : \Omega \rightarrow \mathbb{N}$ be a random time for the filtration $\mathcal{F}_\bullet$, then the *stopped process* $X^\tau : \Omega \rightarrow \mathbb{R}^\mathbb{N}$ is defined for each $n \in \mathbb{N}$ as

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

Proposition 1.14. Let $X : \Omega \rightarrow \mathbb{R}^\mathbb{N}$ be a martingale and $\tau : \Omega \rightarrow \mathbb{N}$ a random time both adapted to the same discrete filtration $\mathcal{F}_\bullet$, then the stopped process X^τ is a martingale adapted to $\mathcal{F}_\bullet$.

Proof. Consider a random sequence $H : \Omega \rightarrow \{0, 1\}^\mathbb{N}$ defined by $H_n \triangleq \mathbb{1}_{\{n \leq \tau\}}$ 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

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

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

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

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 1. For any martingale $X : \Omega \rightarrow \mathbb{R}^\mathbb{N}$ and a stopping time $\tau : \Omega \rightarrow \mathbb{N}$ 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 1.15 (Martingale stopping theorem). Let $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ be a martingale and $\tau : \Omega \rightarrow \mathbb{N}$ be a stopping time, both adapted to a common discrete filtration $\mathcal{F}_{\bullet}$. If either of the following conditions holds true.

(a) τ is bounded.

(b) $\mathbb{E}|X_{\tau \wedge n}|$ is uniformly bounded.

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

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.

(a) Let K be the bound on τ then for all $n \geq K$, 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 $n \geq K$.

(b) Dominated convergence theorem implies the result.

(c) We show that this condition implies that $\mathbb{E}|X_{\tau \wedge n}|$ is uniformly bounded, and thus result follows from the previous part. We can write the difference $X_{\tau \wedge n} - X_0 = \sum_{m=1}^{\tau} \mathbb{1}_{\{m \leq n\}} (X_m - X_{m-1})$ using the telescopic sum. From triangle inequality for the absolute function and the fact that $0 \leq \mathbb{1}_{\{m \leq n\}} \leq 1$, we can upper bound the difference $|X_{\tau \wedge n}| - |X_0| \leq |X_{\tau \wedge n} - X_0| \leq \sum_{m=1}^{\tau} |X_m - X_{m-1}|$. We define non-negative predictable sequence $H : \Omega \rightarrow \{0, 1\}^{\mathbb{N}}$ for each $m \in \mathbb{N}$, as $H_m \triangleq \mathbb{1}_{\{m \leq \tau\}}$. From the linearity of expectation, the monotone convergence theorem, non-negativity of H_m , the tower property of conditional expectation, predictability of H , and theorem hypothesis, we can upper bound the mean of this term as

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

Since τ is integrable, we observe that $X_{\tau \wedge n}$ is uniformly bounded by an integrable random variable. $\square$

Corollary 1.16 (Wald's Equation). If τ is a finite mean stopping time for an i.i.d. random sequence $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ such that $\mathbb{E}|X_1| < \infty$, 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 $Z : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ for each $n \in \mathbb{N}$, as $Z_n \triangleq \sum_{i=1}^n (X_i - \mu)$. Then Z is a martingale adapted to the natural filtration of X , and

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

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 $\mathbb{E}[Z_{\tau}] = \mathbb{E} \sum_{i=1}^{\tau} X_i - \mu \mathbb{E}\tau$. $\square$