

[AI2613 Lecture 7] Martingale, Stopping Time, Optional Stopping Theorem

April 30, 2022

1 Martingale

Consider a fair gambling game in which the expected gain in each round is zero. As a result, regardless of how much one bets in each round, the money in expectation remains the same. The balances after each round form a *martingale*.

Definition 1 (Martingale) Let $\{X_n\}_{n \geq 0}$ and $\{Z_n\}_{n \geq 0}$ be two sequences of random variables. Let $Z_n = \sum_{t=0}^n X_t$.¹ We say $\{Z_n\}_{n \geq 0}$ is a martingale w.r.t. $\{X_n\}_{n \geq 0}$ if

$$\mathbf{E}[Z_{n+1} \mid X_0, X_1, \dots, X_n] = Z_n.$$

Sometimes we say a single sequence $\{X_n\}_{n \geq 0}$ is a martingale if it is a martingale w.r.t. itself. Formally, if for every $n \geq 0$, it holds that

$$\mathbf{E}[X_{n+1} \mid X_0, \dots, X_n] = X_n.$$

For convenience, from now on we use $\bar{X}_{i,j} = (X_i, X_{i+1}, \dots, X_j)$ to simplify the notations. The conditional expectation $\mathbf{E}[Z_{n+1} \mid \bar{X}_{0,n}]$ is equivalent to $\mathbf{E}[Z_{n+1} \mid \sigma(\bar{X}_{0,n})]$ where $\sigma(\bar{X}_{0,n})$ is the σ -algebra generated by X_0, \dots, X_n . This motivates us to define martingale in a more general way.

Definition 2 (Martingale (defined by filtration)) Let $\{\mathcal{F}_n\}_{n \geq 0}$ be a sequence of σ -algebras. We call such σ -algebra sequence a *filtration* if it satisfies

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \dots$$

Given a filtration $\{\mathcal{F}_n\}_{n \geq 0}$, let $\{Z_n\}_{n \geq 0}$ be a stochastic process that Z_n is \mathcal{F}_n -measurable for every $n \geq 0$. Then we say $\{Z_n\}_{n \geq 0}$ is a martingale w.r.t. $\{\mathcal{F}_n\}_{n \geq 0}$ if for every $n \geq 0$

$$\mathbf{E}[Z_{n+1} \mid \mathcal{F}_n] = Z_n.$$

Example 1 (1-D Random Walk) Consider a random walk on \mathbb{Z} starting from 0. The probability to the left and the probability to the right are both $\frac{1}{2}$ at each step. Denote the action at the n -th step by a uniform random variable $X_n \in \{-1, +1\}$. Let $S_n = \sum_{k=0}^n X_k$. Then we can verify $\{S_n\}_{n \geq 0}$ is a martingale w.r.t. $\{X_n\}_{n \geq 0}$ (or w.r.t. $\{S_n\}_{n \geq 0}$) by noticing that

$$\mathbf{E}[S_{n+1} \mid \bar{X}_{0,n}] = \mathbf{E}[S_n + X_{n+1} \mid \bar{X}_{0,n}] = S_n + \mathbf{E}[X_{n+1} \mid \bar{X}_{0,n}] = S_n.$$

More generally, if $\mathbf{E}[X_k \mid \bar{X}_{0,n}] = \mu$, we define $Y_k = X_k - \mu$ and $S'_n \triangleq \sum_{k=0}^n Y_k = S_n - (n+1)\mu$. Then S'_n is a martingale w.r.t. $\{Y_n\}_{n \geq 0}$.

¹ Consider the problem of fair gambling where X_n is the gain of n -th round and $Z_n = \sum_{t=1}^n X_t$. $\{Z_n\}_{n \geq 0}$ is not necessarily a Markov chain. The value X_n may depend on information before round $n-1$.

If $\mathbf{E}[Z_{n+1} \mid \mathcal{F}_n] \leq Z_n$ in Definition 2, we call $\{Z_n\}_{n \geq 0}$ a supermartingale w.r.t. $\{\mathcal{F}_n\}_{n \geq 0}$. Similarly, if $\mathbf{E}[Z_{n+1} \mid \mathcal{F}_n] \geq Z_n$, we call it a submartingale.

Example 2 Consider a sequence of random variables $\{X_n\}_{n \geq 0}$ where $E[X_n | \bar{X}_{0,n-1}] = 1$ for all $n \geq 1$. Let $P_n = \prod_{k=0}^n X_k$. Then we can verify $\{P_n\}_{n \geq 0}$ is a martingale w.r.t. $\{X_n\}_{n \geq 0}$ by verifying that

$$E[P_{n+1} | \bar{X}_{0,n}] = E[P_n \cdot X_{n+1} | \bar{X}_{0,n}] = P_n \cdot E[X_{n+1} | \bar{X}_{0,n}] = P_n.$$

Example 3 (Galton-Watson Process) Recall the Galton-Watson process we discussed in the last lecture. Suppose that all the individuals reproduce independently of each other and have the same offspring distribution. Let G_t be the number of individuals of the t -th generation. Each individual of generation t gives birth to a random number of children of generation $t + 1$. Denote by $X_{t,k}$ the number of children of the k -th individual in the t -th generation. Assume that $X_{t,k}$ are i.i.d. and let $\mu \triangleq E[X_{t,k}]$. Then we have $G_{t+1} = \sum_{k=1}^{G_t} X_{t,k}$. Thus,

$$E[G_{t+1} | G_t] = E\left[\sum_{k=1}^{G_t} X_{t,k} \mid G_t\right] = G_t \cdot E[X_{t,1}] = \mu G_t.$$

Define $M_t = \mu^{-t} G_t$. Then

$$E[M_{t+1} | G_t] = \mu^{-t-1} E[G_{t+1} | G_t] = \mu^{-t} G_t = M_t.$$

That is, $\{M_t\}_{t \geq 0}$ is a martingale w.r.t. $\{G_t\}_{t \geq 0}$.

Example 4 (Pólya's urn) Suppose there are some white balls and black balls in an urn. All of these balls are identical except the colors. Consider the following stochastic process: each round we pick a ball uniformly at random and observe its color; then we return the ball, and add an additional ball of the same color into the urn. We repeat the process, and our goal is to study the sequence of colors of the selected balls.

W.l.o.g. assume that we start from only one white ball and one black ball in the urn, and the index of rounds starts from 2. Then after round n , there are exactly n balls in the urn. Let X_n be the number of black balls after round n , and $Z_n = \frac{X_n}{n}$ is the ratio of black balls after round n . Clearly $Z_2 = \frac{1}{2}$. Then we have

$$\begin{aligned} E[Z_{n+1} | \bar{X}_{2,n}] &= \frac{1}{n+1} E[X_{n+1} | \bar{X}_{2,n}] \\ &= \frac{1}{n+1} (Z_n(X_n + 1) + (1 - Z_n)X_n) = \frac{Z_n + X_n}{n+1} = Z_n. \end{aligned}$$

That is, $\{Z_n\}_{n \geq 2}$ is a martingale w.r.t. $\{X_n\}_{n \geq 2}$.

Example 4 shows that X_n does not have to be i.i.d.

2 Optional Stopping Theorem

2.1 Stopping Time

If $\{Z_n\}_{n \geq 0}$ is a martingale w.r.t. $\{\mathcal{F}_n\}_{n \geq 0}$, $E[Z_{n+1} | \mathcal{F}_n] = Z_n$. Take the expectation of both sides, we have

$$E[Z_{n+1}] = E[Z_n] = \dots = E[Z_1] = E[Z_0].$$

That is, for any fixed $t \geq 0$, $\mathbf{E}[Z_t] = \mathbf{E}[Z_0]$. However, when t is a random variable (we denote it by τ), does $\mathbf{E}[Z_\tau] = \mathbf{E}[Z_0]$ still hold? The answer is obviously no in general. Consider the one-dimensional random walk with $Z_0 = 1$ and τ is the first time that $Z_t = 100$. We have $\mathbf{E}[Z_\tau] = 100 \neq \mathbf{E}[Z_0]$. To determine under which condition we could conclude $\mathbf{E}[Z_\tau] = \mathbf{E}[Z_0]$, let's formalize the notion of *stopping time* first.

Definition 3 (Stopping Time) Let $\tau \in \mathbb{N} \cup \{\infty\}$ be a random variable. We say τ is a stopping time defined on a filtration $\{\mathcal{F}_n\}_{n \geq 0}$ if for any $t \geq 0$, $\mathbb{1}[\tau \leq t]$ is \mathcal{F}_t -measurable.

Then we introduce the Optional Stopping Theorem (OST) to show the sufficient conditions that $\mathbf{E}[Z_\tau] = \mathbf{E}[Z_0]$.

For a counter example, in a game, let $\tau' \triangleq$ "the last time I win 5 in streak". Then τ' is not a stopping time.

2.2 Optional Stopping Theorem

Theorem 4 (Optional Stopping Theorem) Suppose that $\{X_t\}_{t \geq 0}$ is a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ and τ is a stopping time with respect to the same filtration. Then $\mathbf{E}[X_\tau] = \mathbf{E}[X_0]$ if at least one of the following holds

1. τ is bounded almost surely, that is, $\exists n \in \mathbb{N}$ such that $\Pr[\tau \leq n] = 1$.
2. $\Pr[\tau < \infty] = 1$ and $\exists M \in \mathbb{N}$ such that $|X_t| \leq M$ for all $t \leq \tau$.
3. $\mathbf{E}[\tau] < \infty$ and $\exists c \in \mathbb{N}$ such that $\mathbf{E}[|X_{t+1} - X_t| | \mathcal{F}_t] \leq c$ for all $t \leq \tau$.

Before proving Theorem 4, we see some applications of OST.

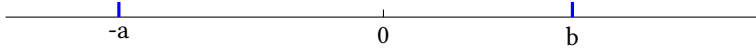
Example 5 (Sex Ratio) Recall the problem of sex ratio we mentioned in the first lecture. Consider the following reproduction strategies

1. Every family keeps having children until they give birth to a boy.
2. Every family keeps having children until their sons are more than their daughters.
3. Every family keeps having children until their sons are more than their daughters or the number of kids is not less than 5.

Assume that the natural birth sex ratio is uniform and every family only gives birth to a child at a time. Fix a family. Let $X_t = \{-1, +1\}$ denote whether the t -th child is a boy, and $Z_n = \sum_{t=1}^n X_t$ denote the number of boys more than girls. This process is a random walk on \mathbb{Z} starting from 0. Let $\tau_1 = \min\{t \geq 0 : X_t = +1\}$, $\tau_2 = \min\{t \geq 0 : Z_t = 1\}$ and $\tau_3 = \min\{\tau_2, 5\}$. Obviously, τ_1 , τ_2 and τ_3 are stopping times corresponding to the 3 strategies.

It is not hard to verify that τ_1 satisfies the third condition in Theorem 4 and τ_3 satisfies all the 3 conditions while τ_2 does not satisfy any conditions in Theorem 4. Thus, we have that the sex ratio is always 1:1 using the first or third strategy. However, if the family adopts strategy 2, since boys are always more than girls when they stop having kids, the sex ratio is unbalanced.

Example 6 (1-D Random Walk with Two Absorbing Barriers) Let $a, b > 0$ be two integers. A man starts the random walk from 0 and stops when he arrives at $-a$ or b . Let τ be the time when the man first reaches $-a$ or b . We want to compute the expected value of τ , that is, the average stopping time of the walk.



Let $X_n \in \{-1, +1\}$ be a uniform random variable, $Z_n = \sum_{k=0}^n X_k$. Then $\tau = \min \{t : Z_t = -a \text{ or } Z_t = b\}$. We know that $\{Z_t\}_{t \geq 0}$ is a martingale w.r.t. $\{X_t\}_{t \geq 0}$ and τ is a stopping time.

First we verify that $\mathbf{E}[Z_\tau] = \mathbf{E}[Z_0]$ by showing $\{Z_t\}$ satisfies the second condition in Theorem 4. Note that $|Z_n|$ is bounded. So, in order to apply OST, we should prove that $\Pr[\tau < \infty] = 1$. Since the probability of ending within the next $a + b$ steps is at least 2^{-a-b} no matter where the current position is, the random walk ends in finite steps with probability 1. Furthermore, if we divide the time into consecutive periods in this manner, in expected finite time, we can meet some period with the event happened, that is, $\mathbf{E}[\tau] < \infty$. So $\{Z_t\}$ also satisfies the third condition. It then follows that $\mathbf{E}[Z_\tau] = \mathbf{E}[Z_0] = 0$. That is

$$-a \cdot \Pr[\text{ending at } -a] + b \cdot (1 - \Pr[\text{ending at } -a]) = 0.$$

This yields $\Pr[\text{ending at } -a] = \frac{b}{a+b}$ and $\Pr[\text{ending at } b] = \frac{a}{a+b}$.

Then we define $\{Y_t\}_{t \geq 0}$ by $Y_t = Z_t^2 - t$. Note that

$$\begin{aligned} \mathbf{E}[Y_{t+1} \mid \bar{X}_{0,t}] &= \mathbf{E}[Z_{t+1}^2 - (t+1) \mid \bar{X}_{0,t}] \\ &= \mathbf{E}[(Z_t + X_{t+1})^2 - (t+1) \mid \bar{X}_{0,t}] \\ &= \mathbf{E}[Z_t^2 + X_{t+1}^2 + 2Z_t \cdot X_{t+1} - (t+1) \mid \bar{X}_{0,t}] \\ &= Z_t^2 + 2Z_t \mathbf{E}[X_{t+1} \mid \bar{X}_{0,t}] + \mathbf{E}[X_{t+1}^2 \mid \bar{X}_{0,t}] - (t+1) \\ &= Z_t^2 + 0 + 1 - (t+1). \end{aligned}$$

That is, $\{Y_t\}_{t \geq 0}$ is a martingale w.r.t. $\{X_t\}_{t \geq 0}$. Note that $\{Y_t\}$ satisfy the third condition of OST. Then $\mathbf{E}[Y_\tau] = \mathbf{E}[Y_0] = 0$. Thus,

$$\mathbf{E}[Z_\tau^2 - \tau] = 0 \Rightarrow \mathbf{E}[\tau] = \mathbf{E}[Z_\tau^2] = a^2 \cdot \frac{b}{a+b} + b^2 \cdot \frac{a}{a+b} = ab.$$

Example 7 (Pattern Matching) Suppose that there is a $\{H, T\}$ -string P of length ℓ (H for “head” and T for “tail”). We flip a coin consecutively until the last ℓ results form exactly the same string as P . How many times do we flip the coin?

Note that if we flip the coin N times and observe the string S consisting of N results. No matter which pattern we choose, by the linearity of expectation, the expected number of occurrence² is

² That means the expected number of substrings exactly the same as P in the resulting string S .

$$E[\text{\# of occurrence of } P \text{ in } S] = \sum_{i=1}^{n-\ell+1} E[\mathbb{1}[S_{i,i+1,\dots,i+\ell-1} = P]] = (n - \ell + 1) \cdot 2^{-\ell}.$$

However, if we would like to compute the first time that pattern P occurs, the pattern itself has an impact on the expected time. Intuitively, let's consider two patterns HT and HH . Assume that the first flipping result is H . Then we consider what happens if the second result fails. Suppose that the desired pattern is HT and H appears. Although we fail, we obtain an H . However, if the desired pattern is HH and the second flipping result is T , then we obtain nothing and the first two flips are a waste. So we should believe that the expected times of the first occurrence of HT is smaller than HH .

We now use the optional stopping theorem to solve this problem. Let $P = p_1 p_2 \dots p_\ell$. For every $n \geq 0$, assume that before $n + 1$ -th flipping there is a new gambler G_{n+1} coming with 1 unit of money to bet that the following ℓ result (i.e., the $n + 1$ -th to $n + \ell$ -th results) are exactly the same as P . At the $n + k$ -th flipping, G_{n+1} will bet that the result is p_k by an all in strategy, that is, if the $n + k$ -th result is p_k then G_{n+1} will have twice as much money as before; otherwise they will lose all. Suppose that the pattern $P = HTHH$ and the flipping results are $HTHHTHTH$. The following table shows the total money of each gambler after flipping.

Gambler	H	T	H	H	T	H	T	H	Money	
1	H	T	H	T					0	1 → 2 → 4 → 8 → 0
2		H							0	1 → 0
3			H	T					0	1 → 2 → 0
4				H	T	H	T	H	32	1 → 2 → 4 → 8 → 16 → 32
5					H				0	1 → 0
2						H	T	H	8	1 → 2 → 4 → 8
5							H		0	1 → 0
5								H	2	1 → 2

Let X_t be the result of t -th flipping, $M_i(t)$ denote the money that G_i has after t -th flipping, and $Z_t \triangleq \sum_{i=1}^t (M_i(t) - 1)$ be the total income of all gamblers after t -th flipping. It is easy to verify that $\{M_i(t)\}_{t \geq 0}$ is a martingale with respect to $\{X_t\}$ since

$$E[M_i(t+1) \mid \bar{X}_{0,t}] = \frac{1}{2} \cdot 2M_i(t) + \frac{1}{2} \cdot 0 = M_i(t).$$

Then by the linearity of expectation we conclude that $\{Z_t\}_{t \geq 0}$ is a martingale with respect to the flipping results $\{X_t\}$ since $E[M_i(t)] = 1$. Let τ be the stopping time defined by the first time that some gambler wins, namely, the first time that P occurs in the flipping results. Applying Condition 2 of OST we obtain that $E[Z_\tau] = E[Z_0] = 0$. Sequentially we have $E[\sum_{i=1}^\tau M_i(\tau) - \tau] = 0$ and $E[\tau] = \sum_{i=1}^\tau E[M_i(\tau)]$.

Note that $M_i(t) = 0$ for $i \leq \tau - \ell$ and $M_i(t) = 2^{\tau-i+1} \chi_{\tau-i+1}$ for $i > \tau - \ell$ where χ_j is defined by

$$\chi_j = \mathbb{1}[p_1 p_2 \dots p_j = p_{\ell-j+1} \dots p_{\ell-1} p_\ell].$$

Hence,

$$\mathbf{E}[\tau] = \sum_{i=\tau-\ell+1}^{\tau} \mathbf{E}[M_i(\tau)] = \sum_{i=1}^{\ell} 2^i \chi_i.$$

Recall the example of HH and HT. If P is HH, $\mathbf{E}[\tau] = 2 + 4 = 6$. If P is HT, $\mathbf{E}[\tau] = 4$. This confirms our hypothesis that $\mathbf{E}[\tau]$ for HH is larger than $\mathbf{E}[\tau]$ for HT.

3 Proof of OST

Theorem 5 (Optional Stopping Theorem (restated)) Suppose that $\{X_t\}_{t \geq 0}$ is a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ and τ is a stopping time with respect to the same filtration. Then $\mathbf{E}[Z_\tau] = \mathbf{E}[Z_0]$ if at least one of the following holds

1. τ is bounded almost surely, that is, $\exists n \in \mathbb{N}$ such that $\Pr[\tau \leq n] = 1$.
2. $\Pr[\tau < \infty] = 1$ and $\exists M \in \mathbb{N}$ such that $|X_t| \leq M$ for all $t \leq \tau$.
3. $\mathbf{E}[\tau] < \infty$ and $\exists c \in \mathbb{N}$ such that $\mathbf{E}[|X_{t+1} - X_t| | \mathcal{F}_t] \leq c$ for all $t \leq \tau$.

Proof. First we show that for every $n \in \mathbb{N}$, $\mathbf{E}[X_{\min\{n, \tau\}}] = \mathbf{E}[X_0]$. Define $Z_n = X_{\min\{n, \tau\}} = X_0 + \sum_{i=0}^{n-1} (X_{i+1} - X_i) \mathbb{1}[\tau > i]$. We verify that $\{Z_n\}_{n \geq 0}$ is a martingale. By definition

$$\mathbf{E}[Z_{n+1} | \mathcal{F}_n] = \mathbf{E}[Z_n + (X_{n+1} - X_n) \mathbb{1}[\tau > n] | \mathcal{F}_n] = Z_n + \mathbb{1}[\tau > n] (\mathbf{E}[X_{n+1} | \mathcal{F}_n] - X_n) = Z_n.$$

So we have $\mathbf{E}[X_{\min\{n, \tau\}}] = \mathbf{E}[Z_n] = \mathbf{E}[Z_0] = \mathbf{E}[X_0]$.

Therefore, this motivates us to decompose X_τ into two terms:

$$\forall n \in \mathbb{N}, X_\tau = X_{\min\{n, \tau\}} + \mathbb{1}[\tau > n] \cdot (X_\tau - X_n).$$

Taking expectation and letting n tend to infinity, we obtain

$$\mathbf{E}[X_\tau] = \mathbf{E}[X_0] + \lim_{n \rightarrow \infty} \mathbf{E}[\mathbb{1}[\tau > n] \cdot (X_\tau - X_n)].$$

Therefore, we only need to verify that each of the three conditions in the statement can guarantee $\lim_{n \rightarrow \infty} \mathbf{E}[\mathbb{1}[\tau > n] \cdot (X_\tau - X_n)] = 0$.

1. If τ is bounded almost surely, then clearly $\mathbf{E}[\mathbb{1}[\tau > n] \cdot (X_\tau - X_n)] = 0$ for sufficiently large n .

2. In this case,

$$\begin{aligned} \mathbf{E}[\mathbb{1}[\tau > n] \cdot (X_\tau - X_n)] &\leq \mathbf{E}[\mathbb{1}[\tau > n] \cdot (|X_\tau| + |X_n|)] \\ &\leq 2M \cdot \mathbf{E}[\mathbb{1}[\tau > n]] \\ &= 2M \cdot \Pr[\tau > n] \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

3. In order to apply our bounds on the gap between consecutive X_t , we write

$$\begin{aligned} \mathbb{1}[\tau > n] \cdot (X_\tau - X_n) &= \sum_{t=n}^{\tau-1} (X_{t+1} - X_t) \\ &\leq \sum_{t=n}^{\tau-1} |X_{t+1} - X_t| \\ &= \sum_{t=n}^{\infty} |X_{t+1} - X_t| \cdot \mathbb{1}[\tau > t]. \end{aligned}$$

Taking expectation and applying the *Fubini's theorem*, we obtain

$$\begin{aligned} \mathbf{E}[\mathbb{1}[\tau > n] \cdot (X_\tau - X_n)] &\leq \sum_{t=n}^{\infty} \mathbf{E}[|X_{t+1} - X_t| \cdot \mathbb{1}[\tau > t]] \\ &= \sum_{t=n}^{\infty} \mathbf{E}[\mathbf{E}[|X_{t+1} - X_t| \cdot \mathbb{1}[\tau > t] \mid \mathcal{F}_t]] \\ &= \sum_{t=n}^{\infty} \mathbf{E}[\mathbf{E}[|X_{t+1} - X_t| \mid \mathcal{F}_t] \cdot \mathbb{1}[\tau > t]] \\ &\leq c \sum_{t=n}^{\infty} \mathbf{E}[\mathbb{1}[\tau > t]], \end{aligned}$$

where the last equality is due to the fact that $\mathbb{1}[\tau > t]$ is \mathcal{F}_t -measurable.

□