AI2613 随机过程	2020–2021 春季学期

Lecture 12 – Martingale (I)

2021 年 5 月 17 日

Lecturer: 张驰豪 | *Scribe*: 杨宽

1 Introduction to Martingale

Today we are talking about a kind of interesting stochastic process. Consider a *fair* gambling game. After each round, the expectation of our money is the same as before. This stochastic process is a *martingale*.

Definition 1 ((Discrete) Martingale (鞅)). A (discrete) stochastic process $\{X_n\}_{n\geq 0}$ is a *martingale* if

$$\forall\, n \geq 0, \qquad \mathbb{E}[X_{n+1} \mid X_0, \ldots, X_n] = X_n.$$

We would like to generalize this definition slightly. For convenience, from now on we will use $\overline{X}_{0,n} = (X_0, \ldots, X_n)$ to simplify our notations.

Definition 2. Suppose $\{X_n\}_{n\geq 0}$ and $\{Z_n\}_{n\geq 0}$ are two stochastic processes. We say $\{Z_n\}_{n\geq 0}$ is a *martingale* with respect to $\{X_n\}_{n\geq 0}$ if

$$\forall\, n \geq 0, \qquad \mathbb{E}\left[Z_{n+1} \mid \overline{X}_{0,n}\right] = Z_n.$$

Remark. Definition 1 is consistent with Definition 2 if we take X_n the same as Z_n in Definition 2. So $\{X_n\}_{n\geq 0}$ is a martingale if $\{X_n\}_{n\geq 0}$ is a martingale with respect to itself.

Definition 3 (Martingale (defined by σ-algebra)). Given a stochastic process $\{X_n\}_{n\geq 0}$, let $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ be the minimum σ-algebra generated by $\overline{X}_{0,n}$. Then $\{\mathcal{F}_n\}_{n\geq 0}$ satisfies

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \cdots.$$

In particular, a sequence of σ-algebra $\{\mathcal{F}_n\}_{n\geq 0}$ satisfying the above condition is called a *filtration*. We say $\{Z_n\}_{n\geq 0}$ is a *martingale* with respect to a filtration $\{\mathcal{F}_n\}_{n\geq 0}$ if

1. $\forall\, n \geq 0$, Z_n is $\mathcal{F}_n$-measurable;
2. $\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] = Z_n$.

Example 4 (One-dimensional Random Walk). Consider a random walk on $\mathbb{Z}$ starting from 0. The probability to left and the probability to right are both 1/2 at each step. Denote by a uniform random variable $X_n \in \{-1, +1\}$ the n-th step. Let $Z_n = \sum_{i=1}^{n} X_n$. Then $\{Z_n\}$ is a martingale with respect to $\{X_n\}$.

It is easy to verify that

1. Z_n is measurable given the σ-algebra $\sigma(\overline{X}_{0,n})$;
2. $\mathbb{E}\left[Z_{n+1} \mid \overline{X}_{0,n}\right] = \mathbb{E}\left[Z_n + X_{n+1} \mid \overline{X}_{0,n}\right] = Z_n + \mathbb{E}\left[X_{n+1} \mid \overline{X}_{0,n}\right] = Z_n$.

Remark. Taking the expectation of the both sides of $\mathbb{E}\left[Z_{n+1} \mid \overline{X}_{0,n}\right] = Z_n$ and applying the law of total expectation we obtain that $\mathbb{E}[Z_{n+1}] = \mathbb{E}[Z_n]$.

Example 5 (Branching Process). Suppose that all the individuals reproduce independently of each other and have the same offspring distribution. Let Z_n be the number of individuals of the n-th generation. Each individual of generation n gives birth to a random number of children of generation $n+1$. Denote by $X_{n,i}$ the number of children of the i-th individual. We further assume that $X_{n,i}$ are i.i.d. random variables with $\mathbb{E}[X_{n,i}] = \mu$.

By assumption we have $\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[Z_{n+1} \mid Z_n]$. Furthermore,

$$\mathbb{E}[Z_{n+1} \mid Z_n = z] = \mathbb{E}\left[\sum_{i=1}^{z} X_{n,i} \mid Z_n = z\right] = \mu \cdot z\,.$$

What we want is to construct a martingale. So we should scale Z_n. Let $M_n = \mu^{-n} \cdot Z_n$. The calculation above justifies that

$$\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = \mu^{-(n+1)} \cdot \mu \cdot Z_n = M_n\,.$$

Remark. In the definition of martingale (such as Definition 3), it is required that $\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] = Z_n$. However, sometimes the stochastic process that we are studying might not satisfy this condition but it still has some important properties and is still widely applied. So we define the following two kinds of stochastic processes as well: If the condition is $\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] \geq Z_n$ for all $n \geq 0$ then it is called a *submartingale* (下鞅); if the condition is $\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] \leq Z_n$ for all $n \geq 0$ then it is called a *supermatingale* (上鞅).

Example 6 (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 3. Then after round n, there are exactly n balls in the urn. Let X_n be the number of white balls after round n, and $Z_n = X_n/n$ is the ratio of white balls after round n. We claim that Z_n is a martingale.

Clearly $Z_2 = 1/2$ and $\mathbb{E}[Z_n] = 1/2$ since white balls and black balls are symmetric. We now compute $\mathbb{E}\left[Z_{n+1} \mid \overline{X}_{2,n}\right]$. Note that at round $n+1$, we pick a white ball with probability Z_n. Thus,

$$\begin{aligned}\mathbb{E}\left[Z_{n+1} \mid \overline{X}_{2,n}\right] &= \frac{1}{n+1} \cdot \mathbb{E}\left[X_{n+1} \mid \overline{X}_{2,n}\right] \\ &= \frac{1}{n+1} \cdot \left(Z_n \cdot (X_n + 1) + (1 - Z_n) \cdot X_n\right) \\ &= \frac{1}{n+1} \cdot (X_n + Z_n) \ = Z_n \,. \end{aligned} \tag{1}$$

Example 7 (Likelihood Ratio). Suppose there is a sequence of numbers $X_1, X_2, \ldots, X_n, \ldots$ independently chosen from an unknown distribution f. We guess that their have distribution g. Then the *likelihood ratio* of our guess is defined by

$$M_n \triangleq \frac{g(x_1) \cdot g(x_2) \cdot \cdots \cdot g(x_n)}{f(x_1) \cdot f(x_2) \cdot \cdots \cdot f(x_n)} \,.$$

We claim that $\{M_n\}$ is a martingale with respect to $\overline{X}_{0,n}$.

In fact, by independence, it is clear to verify that

$$\begin{aligned}\mathbb{E}\left[M_{n+1} \mid \overline{X}_{1,n}\right] &= M_n \cdot \mathbb{E}\left[\frac{g(X_{n+1})}{f(X_{n+1})} \mid \overline{X}_{1,n}\right] \\ &= M_n \cdot \mathbb{E}\left[\frac{g(X_{n+1})}{f(X_{n+1})}\right] \\ &= M_n \cdot \sum_x \mathbf{Pr}[X_{n+1} = x] \cdot \frac{g(x)}{f(x)} \\ &= M_n \cdot \sum_x g(x) \ = M_n \,. \end{aligned}$$

2 Stopping Time

Question. We've already known that $\forall\, n \geq 0$, $\mathbb{E}[Z_n] = \mathbb{E}[Z_0]$. However, if τ is a random variable, could we conclude that $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0]$?

Unfortunately the answer is "no"! For example, consider a one-dimensional random walk starting from 0. Let Z_n be the position after the n-th step, and τ be the first time that $Z_\tau = 1$. It is clear

that $\mathbb{E}[Z_\tau] = 1 \neq Z_0$. Another example is to define τ as $\arg\max_{1\le t\le 100}|Z_t|$, the time to reach the furthest position in the first 100 steps. Obviously we have $\mathbb{E}[Z_\tau] > 0 = Z_0$.
To answer our question, and determine under which condition we could conclude $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0]$, let's formalize *stopping time* first.

Definition 8 (Stopping Time). Let $\tau \in \mathbb{N}$ is a random variables. We say τ is a *stopping time* defined on a filtration $\{\mathcal{F}_n\}_{n\ge 0}$ if

$$\forall\, n \in \mathbb{N}, \qquad \mathbb{1}_{[\tau \le n]} \text{ is } \mathcal{F}_n\text{-measurable}.$$

In other words, for every n the event $\{\tau \le n\}$ is in $\mathcal{F}_n$.

Now consider the filtration generated by a stochastic process $\{X_n\}_{n\ge 0}$. We say τ is a stopping time if the proposition $[\tau \le n]$ is determined by $\overline{X}_{0,n}$ for all n.

Theorem 1 (Optional Stopping Theorem). *Suppose that $\{X_n\}$ is a martingale with respect to a filtration $\{\mathcal{F}_n\}$ and τ is a stopping time with respect to the same filtration. Then $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ if at leat one of the following holds*

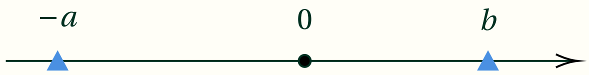
1. *τ is bounded;*
2. *$\mathbf{Pr}[\tau < \infty] = 1$ and $\exists\, M$ such that $|X_n| \le M$ for all $n < \tau$;*
3. *$\mathbb{E}[\tau] < \infty$ and $\exists\, c$ such that $\mathbb{E}[|X_{n+1} - X_n| \mid \mathcal{F}_n] \le c$ for all $n < \tau$.*

The proof of the optional stopping theorem is left as an exercise. We now introduce some examples and applications of this theorem.

Example 9 (Sex Ratio). Suppose that in a villiage, every family keeps having children until they give birth to a boy. If we further assume that the natural birth sex ratio is uniform and every family only gives birth to a child at a time, what is the birth sex ratio in this villiage?
Fix a family. Let $X_n \in \{-1, +1\}$ denote whether the n-th child is a boy, and $Z_n = \sum_{i=1}^n X_i$ denote the number of boys more than girls. Then we define τ by $\tau = \min\{n \colon X_n = 1\}$. Clearly $\{Z_n\}$ is a martingale with respect to $\{X_n\}$ and τ is a stopping time. Note that τ is the time of success in a Bernoulli trial and has a geometric distribution. Thus $\mathbb{E}[\tau] < \infty$. Combining with the fact that $|Z_{n+1} - Z_n| = |X_{n+1}| = 1$, it justifies Condition 3 in Theorem 1. So we conclude that $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0] = 0$ and hence the birth sex ratio in this villiage is still $1:1$.
Suppose that their strategy has changed. Every family keeps giving birth to children until their sons are more than their daughters. Then the optional stopping theorem cannot be applied any longer, since now the stopping time τ has infinite expectation and Z_n is unbounded. To see this, note that τ is the hitting time of 1 in a one-dimensional random walk starting from 0, that is, the hitting time of a *null recurrent* state.

Example 10 (One-Dimensional Random Walk with Two Absorbing Barriers). Consider a one-dimensional random walk (starting from 0) with two absorbing barriers $-a$ and b. There are two natural questions:

1. What is the probability of stopping at $-a$ (or b)?
2. What is the expected number of steps before stopping?



Solution. Let $X_n \in \{-1, +1\}$ be a uniform random variable, $Z_{n+1} = Z_n + X_n$, and $\tau = \min\{n \colon Z_n = -a \vee Z_n = b\}$. Then $\{Z_n\}$ is a martingale w.r.t. $\{X_n\}$, and τ is a stopping time.
Note that $|Z_n|$ is bounded. So, in order to apply the optional stopping theorem, we should prove that $\mathbf{Pr}[\tau < \infty] = 1$. Since the probability of ending within the next $a+b$ steps is at least $2^{-(a+b)}$ no matter where the current position is, we claim that the random walk ends in finite steps with probability 1. It follows that $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0] = 0$. That is

$$-a \cdot \mathbf{Pr}[\text{ending at } -a] + b \cdot \mathbf{Pr}[\text{ending at } b] = 0\,,$$

which yields that the probability of ending at $-a$ and the probability of ending at b are $b/(a+b)$ and $a/(a+b)$, respectively.
We also define $\{Y_n\}_{n \geq 0}$ (which is a common trick) by

$$Y_n \triangleq Z_n^2 - n\,.$$

Claim. We claim that $\{Y_n\}$ is a martingale w.r.t. $\{X_n\}$.

Now we use Condition 3 in Theorem 1. It implies that $\mathbb{E}[Y_\tau] = \mathbb{E}[Y_0] = 0$. By the linearity of expectation, $\mathbb{E}[Y_\tau] = \mathbb{E}\left[Z_\tau^2\right] - \mathbb{E}[\tau]$. It follows that

$$\begin{aligned} \mathbb{E}[\tau] = \mathbb{E}\left[Z_\tau^2\right] &= a^2 \cdot \mathbf{Pr}[\text{ending at } -a] + b^2 \cdot \mathbf{Pr}[\text{ending at } b] \\ &= a^2 \cdot \frac{b}{a+b} + b^2 \cdot \frac{a}{a+b} = ab\,. \end{aligned}$$

Finally, we prove our claim, namely, $\{Y_n\}$ is a martingale with respect to $\{X_n\}$. By definition, we obtain that

$$\begin{aligned} \mathbb{E}\left[Y_{n+1} \mid \overline{X}_{0,n}\right] &= \mathbb{E}\left[Z_{n+1}^2 - (n+1) \mid \overline{X}_{0,n}\right] \\ &= \mathbb{E}\left[(Z_n + X_{n+1})^2 - (n+1) \mid \overline{X}_{0,n}\right] \\ &= Z_n^2 - n - 1 + 2Z_n \cdot \mathbb{E}\left[X_{n+1} \mid \overline{X}_{0,n}\right] + \mathbb{E}\left[X_{n+1}^2 \mid \overline{X}_{0,n}\right] \\ &= Z_n^2 - n = Y_n\,, \end{aligned}$$

which completes our proof.

Example 11 (Pattern Matching). Suppose that there is a $\{\mathsf{H},\mathsf{T}\}$-string P of length ℓ. 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 consisting of N results. No matter which pattern we choose, the expected number of occurrence (i.e., expected number of substrings exactly the same as P of the resulting string) is $(N-\ell+1)/2^{\ell}$ (by the linearity of expectation). 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. Altough 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 these two patterns are different.

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 patter $P = \mathsf{HTHTH}$ 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	$1 \to 2 \to 4 \to 8 \to 0$
2		H							0	$1 \to 0$
3			H	T					0	$1 \to 2 \to 0$
4				H	T	H	T	H	32	$1 \to 2 \to 4 \to 8 \to 16 \to 32$
5					H				0	$1 \to 0$
6						H	T	H	8	$1 \to 2 \to 4 \to 8$
7							H		0	$1 \to 0$
8								H	2	$1 \to 2$

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

$$Z_t = \sum_{i=1}^{t} M_i(t) - 1$$

be the total *net 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 $\{X_n\}$ since

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

Then by the linearity of expectation we conclude that $\{Z_n\}$ is a martingale with respect to the flipping results $\{X_n\}$ since $\mathbb{E}[M_i(t)] = 0$. 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 Theorem 1 we obtain that $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0] = 0$.

We complete our solution by pointing out that G_i lose all for all $i \leq \tau - \ell$ and $M_i(\tau) = 2^{\tau-i+1}\cdot\chi_{\tau-i+1}$ for all $\tau - \ell + 1 \leq i \leq \tau$, where χ_j is defined by

$$\chi_j = \mathbb{1}_{[p_1\cdots p_j = p_{\ell-j+1}\cdots p_\ell]}\,.$$

Hence,

$$0 = \mathbb{E}[Z_\tau] = \sum_{i=1}^{\tau}\mathbb{E}[M_i(\tau)] - \mathbb{E}[\tau] = \sum_{i=\tau-\ell+1}^{\tau} M_i(\tau) - \mathbb{E}[\tau] = \sum_{i=1}^{\ell}\chi_i\cdot 2^i - \mathbb{E}[\tau]\,.$$