MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070J Fall 2013
Lecture 10 10/4/2013

Martingales and stopping times

Content.

1. Martingales and properties.
2. Stopping times and Optional Stopping Theorem.

1 Martingales

We continue with studying examples of martingales.

- **Brownian motion.** A standard Brownian motion $B(t)$ is a martingale on $C[0, \infty)$, equipped with the Wiener measure, with respect to the filtration $\mathcal{B}_t, t \in \mathbb{R}_+$, defined as follows. Let $\Im_t$ be the the Borel σ-field on $C[0, t]$ generated by open and closed sets with respect to the sup norm: $\|f\| = \max_{z\in[0,t]} |f(z)|$. For every set $A \subset C[0, t]$ consider an extension $\bar{A} \subset C[0, \infty)$ as a set of all $f \in C[0, \infty)$ such that the restriction of f onto $[0, t]$ belongs to A. Then define $\mathcal{B}_t$ to be the set of all extensions $\bar{A}$ of sets $A \in \Im_t$. Informally, $\mathcal{B}_t$ is the field representing the information about the paths f available by observing it during the time interval $[0, t]$. Naturally $\mathcal{B}_s \subset \mathcal{B}_t \subset \mathcal{B}$ when $s \le t$, where $\mathcal{B}$ is the Borel σ-field of the entire space $C[0, \infty)$.

  We now check that B_t is indeed a martingale. Fix $0 \le s < t$. We have

$$\begin{aligned}\mathbb{E}[B(t)|\mathcal{B}_s] &= \mathbb{E}[B(t) - B(s) + B(s)|\mathcal{B}_s] \\ &= \mathbb{E}[B(t) - B(s)|\mathcal{B}_s] + \mathbb{E}[B(s)|\mathcal{B}_s] \\ &= 0 + B(s) \\ &= B(s)\end{aligned}$$

  where we use the fact that, due to the independent increments property, $B(t) - B(s)$ is independent from $\mathcal{B}_s$, implying $\mathbb{E}[B(t) - B(s)|\mathcal{B}_s] = \mathbb{E}[B(t) - B(s)] = 0$; and $B(s) \in \mathcal{B}_s$, implying $\mathbb{E}[B(s)|\mathcal{B}_s] = B(s)$.

- **Brownian motion with drift**. Now consider a Brownian motion with drift μ and standard deviation σ. That is consider $B_\mu(t) = \mu t + \sigma B(t)$, where B is the standard Brownian motion. It is straightforward to show that $B_\mu(t) - \mu t$ is a martingale. Also it is simple to see that $(B_\mu(t) - \mu t)^2 - \sigma^2 t$ is also a martingale.

- **Wald's martingale**. Suppose $X_n, n \in \mathbb{N}$ is an i.i.d. sequence with $\mathbb{E}[X_1] = \mu$, such that the exponential moment generating function $M(\theta) \triangleq \mathbb{E}[e^{\theta X_1}] < \infty$ for some $\theta > 0$. Let $S_n = \sum_{1 \le k \le n} X_k$. Then $Z_n = \frac{\exp(\theta S_n)}{M^n(\theta)}$ is a martingale. It is straightforward to check this.

2 Properties of martingales

Note that if X_n is submartingale, then $-X_n$ is supermartingale and vice verse.

Proposition 1. *Suppose X_n is a martingale and ϕ is a convex function such that $\mathbb{E}[|\phi(X_n)|] < \infty$. Then $\phi(X_n)$ is a submartingale.*

Proof. We apply conditional Jensen's inequality:

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

□

As a corollary we obtain that if X_n is a martingale and $\mathbb{E}[|X_n|^p] < \infty$ for all n for some $p \ge 1$, then $|X_n|^p$ is submartingale. Note that if X_n was submartingale and ϕ was non-decreasing, then the same result applies.

Definition 1. *A sequence of random variables H_n is defined to be* predictable *if $H_n \in \mathcal{F}_{n-1}$.*

Here is sort of an artificial example. Let X_n be adapted to filtration $\mathcal{F}_n$. Then $H_n = X_{n-1}, H_0 = H_1 = X_0$, is predictable. We simply "played" with indices. Instead now consider $H_n = \mathbb{E}[X_n|\mathcal{F}_{n-1}]$. By the definition of conditional expectations, $H_n \in \mathcal{F}_{n-1}$, so it is predictable.

Theorem 1 (**Doob's decomposition**). *Every submartingale $X_n, n \ge 0$ adapted to filtration $\mathcal{F}_n$, can be written in a unique way as $X_n = M_n + A_n$ where M_n is a martingale and A_n is an a.s. non-decreasing sequence, predictable with respect to $\mathcal{F}_n$.*

Proof. Set $A_0 = 0$ and define A_n recursively by $A_n = A_{n-1} + \mathbb{E}[X_n|\mathcal{F}_{n-1}] - X_{n-1} \geq A_{n-1}$. Let $M_n = X_n - A_n$. By induction, we have that $A_n \in \mathcal{F}_{n-1}$ since $A_{n-1} \in \mathcal{F}_{n-2} \subset \mathcal{F}_{n-1}$ and $\mathbb{E}[X_n|\mathcal{F}_{n-1}], X_{n-1} \in \mathcal{F}_{n-1}$. Therefore A_n is predictable. We now need to show that M_n is a martingale. To check that $\mathbb{E}[|M_n|] < \infty$, it suffices to shot the same for A_n, since by assumption X_n is a sub-martingale and therefore $\mathbb{E}[|X_n] < \infty$. Now, we establish finiteness of $\mathbb{E}[|A_n|]$ by induction, for which it suffices to have finiteness of $\mathbb{E}[|\mathbb{E}[X_n|\mathcal{F}_{n-1}]|]$ which follows by conditional Jensen's inequality which gives $|\mathbb{E}[X_n|\mathcal{F}_{n-1}]| \leq \mathbb{E}[|X_n|\,|\mathcal{F}_{n-1}]$ and the tower property which gives $\mathbb{E}[\mathbb{E}[|X_n|\,|\mathcal{F}_{n-1}]] = \mathbb{E}[|X_n|] < \infty$. We now establish that $\mathbb{E}[M_n|\mathcal{F}_{n-1}] = M_{n-1}$. We have

$$\begin{aligned}
\mathbb{E}[M_n|\mathcal{F}_{n-1}] &= \mathbb{E}[X_n - A_n|\mathcal{F}_{n-1}] \\
&= \mathbb{E}[X_n|\mathcal{F}_{n-1}] - A_n \\
&= X_{n-1} - A_{n-1} \\
&= M_{n-1}.
\end{aligned}$$

Thus M_n is indeed a martingale. This completes the proof of the existence part.

To prove uniqueness, we assume that $X_n = M_n' + A_n'$ is any such decomposition. Then

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] = \mathbb{E}[M_n' + A_n'|\mathcal{F}_{n-1}] = M_{n-1}' + A_n' = X_{n-1} - A_{n-1}' + A_n'$$

since, by assumption, M_n' is a martingale and A_n' is predictable. Then we see that A_n' satisfies the same recursion as A_n, implying $A_n' = A_n$. It then immediately follows that $M_n' = M_n$. $\square$

The concept of predictable sequences is useful when we want to model a gambling scheme, or in general any controlled stochastic process, where we observe the state X_{n-1} of the process at time $t = n-1$, take some decision at time $t = n$ and observe the realization X_n at time $t = n$. This realization might actually depend on our decision. In the special case of gambling, suppose X_n is your wealth at time n if in each round you bet one dollar. In particular, $X_n - X_{n-1} = \pm 1$. If you instead bet an amount H_n each time, then your net at time n is $\sum_{1 \leq m \leq n} H_m(X_m - X_{m-1})$. Modelling X_n as an i.i.d. coin toss sequence with the usual product σ-field on $\{0,1\}^\infty$, we have X_n is adapted to the σ-field on $\{0,1\}^n$ and H_n is predictable.

Is there a scheme which allows us to beat the system? In fact, yes. This is the famous double the bet strategy. We make $H_1 = 1$ and $H_n = 2H_{n-1}$ if $X_{n-1} - X_{n-2} = -1$ and stop once $X_{n-1} - X_{n-2} = 1$, then assuming we won the first time in round k, our net worth is $-1 - 2 - \cdots - 2^{k-1} + 2^k = 1$.

If $\mathbb{P}(X_{n-1} - X_{n-2} = 1) > 0$, and the games have independently distributed outcomes, then we win a.s. a dollar at some random time k.

As we will see this was possible because we did not put a bound on how long we are allowed to play and how much we are allowed to lose. We now establish that placing either of this conditions makes beating the gambling system impossible.

For now we establish the following result.

Theorem 2. *Suppose X_n is a supermartingale and $H_n \geq 0$ is predictable. Then $Z_n = \sum_{1\leq m\leq n} H_m(X_m - X_{m-1})$ is also a supermartingale.*

Proof. We have

$$\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = \mathbb{E}[H_{n+1}(X_{n+1} - X_n)|\mathcal{F}_n] + \mathbb{E}[\sum_{1\leq m\leq n} H_m(X_m - X_{m-1})|\mathcal{F}_n].$$

Since H is predictable, then $H_{n+1} \in \mathcal{F}_n$, implying that the first summand is equal to

$$H_{n+1}\mathbb{E}[(X_{n+1} - X_n)|\mathcal{F}_n] \leq 0,$$

where inequality follows since X_n is supermartingale. On the other hand,

$$\sum_{1\leq m\leq n} H_m(X_m - X_{m-1}) = Z_n \in \mathcal{F}_n,$$

implying that its expectation is $\mathbb{E}[Z_n|\mathcal{F}_n] = Z_n$. This completes the proof. □

3 Stopping times

In the example above, showing how to create a winning gambling system, notice that part of the strategy was to stop the first time when we win. Thus the game is interrupted at a random time, which depends on the observed conditions. We now formalize this using the concept of *stopping times*. The end goal of this section is establishing the following result: if the gambling strategy involves a stopping time which is bounded, then our expected gain is non-positive.

Definition 2. *Given a filtration $\{\mathcal{F}_t\}_{t\in T}$ on a sample space Ω, a random variable $\tau : \Omega \to T$ is called a stopping time, if the event $\{\tau \leq t\} = \{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$, for every t.*

Here we consider again the two cases $T = \mathbb{N}$ or $T = \mathbb{R}_+$. Note, that we do not need to specify a probability measure here. The concept of stopping times is framed only in terms of measurability with respect to σ-fields.

- Consider the filtration corresponding to a sequence of random variables $X_1, \dots, X_n, \dots$. Namely $\Omega = \mathbb{R}^\infty$ and $\mathcal{F}_n$ is Borel σ-field of $\mathbb{R}^n$. Fix some real value x. Given any sample $\omega = (\omega_1, \dots, \omega_n, \dots) \in \mathbb{R}^\infty$ define

$$\tau(\omega) = \inf\{n : \sum_{1 \le k \le n} \omega_k \ge x\}.$$

Namely, $\tau(\omega)$ is the smallest index at which the sum of the components is at least x. Then τ is a stopping time. Indeed, the event $\tau \le n$ is completely specified by the portion $\omega_1, \dots, \omega_n$ of the sample. In particular

$$\{\omega : \tau(\omega) \le n\} = \cup_{1 \le k \le n}\{\omega : \sum_{1 \le i \le k} \omega_i \ge x\}.$$

Each of the events in the union on the right-hand side is measurable with respect to $\mathcal{F}_n$. This example is the familiar Wald's stopping time: τ is the

smallest n for which a random walk $X_1 + \cdots + X_n \ge x$, except for, we do not say that the random sequence should be i.i.d. and, in fact, we do not say anything about the probability law of the sequence at all.

- Consider the standard Wiener measure on $C([0, \infty))$, that is consider the standard Brownian motion. Then, given $a > 0$, $T_a = \inf\{t : B(t) = a\}$ is a stopping time with respect to the filtration $\mathcal{B}_t, t \in \mathbb{R}_+$, where $\mathcal{B}_t$ is the filtration described in the beginning of the lecture. This is the familiar hitting time of the Brownian motion.

The main result in terms of stopping times that we wish to establish is as follows.

Theorem 3. *Suppose X_n is a supermartingale and τ is a stopping time, which is a.s. bounded: $\tau \le M$ a.s. for some M. Then $\mathbb{E}[X_\tau] \le \mathbb{E}[X_0]$. In other words, if there exists a bound on the number of rounds for betting, then the expected net gain is non-positive, provided that in each round the expected gain is non-positive.*

This theorem will be established as a result of several short lemmas.

Lemma 1. *Suppose τ is a stopping time corresponding to the filtration $\mathcal{F}_n$. Then the sequence of random variables $H_n = 1\{\tau \ge n\}$ is predictable.*

Proof. H_n is a random variable which takes values 0 and 1. Note that the event $\{H_n = 0\} = \{\tau < n\} = \{\tau \le n - 1\}$. Since τ is a stopping time, then the event $\{\tau \le n - 1\} \in \mathcal{F}_{n-1}$. Thus H_n is predictable. $\square$

Given two real values a, b we use $a \wedge b$ to denote $\min(a, b)$.

Corollary 1. *Suppose X_n is supermartingale and τ is a stopping time. Then $Y_n = X_{n\wedge\tau}$ is also a supermartingale.*

Proof. Define $H_n = 1\{\tau \geq n\}$. Observe that

$$\sum_{1\leq m\leq n} H_m(X_m - X_{m-1}) = -H_0X_0 + \sum_{0\leq m\leq n-1} X_m(H_m - H_{m+1}) + H_nX_n. \tag{1}$$

Note, $H_0 = 1\{\tau \geq 0\} = 1$. $H_m - H_{m+1} = 1\{\tau \geq m\} - 1\{\tau \geq m+1\} = 1\{\tau = m\}$. Therefore, the expression on the right-hand side of (1) is equal to $X_{n\wedge\tau} - X_0$. By Theorem 2, the left-hand side of (1) is a supermartingale. We conclude that $Y_n = X_{n\wedge\tau}$ is a supermartingale. □

Now we are ready to obtain our end result.

Proof of Theorem 3. The process $Y_n = X_{n\wedge\tau}$ is a supermartingale by Corollary 1. Therefore

$$\mathbb{E}[Y_M] \leq \mathbb{E}[Y_0]$$

But $Y_M = X_{M\wedge\tau} = X_\tau$ and $Y_0 = X_{0\wedge\tau} = X_0$. We conclude $\mathbb{E}[X_\tau] \leq \mathbb{E}[X_0]$. This concludes the proof of Theorem 3. □

4 Additional reading materials

- Durrett [1] Chapter 4.

References

[1] R. Durrett, *Probability: theory and examples*, Duxbury Press, second edition, 1996.

15.070J / 6.265J Advanced Stochastic Processes
Fall 2013