MATH 171 LECTURE NOTE 5: MARTINGALES

HANBAEK LYU

1. Conditional expectation

Let X, Y be discrete RVs. Recall that the expectation $\mathbb{E}(X)$ is the 'best guess' on the value of X when we do not have any prior knowledge on X. But suppose we have observed that some possibly related RV Y takes value y. What should be our best guess on X, leveraging this added information? This is called the *conditional expectation of X given $Y = y$*, which is defined by

$$\mathbb{E}[X|Y=y] = \sum_x x\mathbb{P}(X=x|Y=y). \tag{1}$$

This best guess on X given $Y = y$, of course, depends on y. So it is a function in y. Now if we do not know what value Y might take, then we omit y and $\mathbb{E}[X|Y]$ becomes a RV, which is called the *conditional expectation of X given Y*.

Exercise 1.1. Let X, Y be discrete RVs. Show that for any function $g:\mathbb{R}\to\mathbb{R}$,

$$\mathbb{E}_X[Xg(Y)\,|\,Y] = g(Y)E_X[X\,|\,Y]. \tag{2}$$

Exercise 1.2 (Iterated expectation)**.** Let X, Y be discrete RVs. Use Fubini's theorem to show that

$$\mathbb{E}[X] = \mathbb{E}_Y[\mathbb{E}_X[X|Y]]. \tag{3}$$

In case when X or Y are continuous RVs, we simply replace the sum by integral and PMF by PDF. For instance, if both X and Y are continuous with PDFs f_X, f_Y and joint PDF $f_{X,Y}$, then

$$\mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y=y}(x)\,dx, \tag{4}$$

where $f_{X|Y=y}$ is the conditional PDF of X given $Y = y$ defined by

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}. \tag{5}$$

To summarize how we compute the iterated expectations when we condition on discrete and continuous RV:

$$\mathbb{E}[\mathbb{E}[X\,|\,Y]] = \begin{cases} \sum_y \mathbb{E}[X\,|\,Y=y]\mathbb{P}(Y=y) & \text{if } Y \text{ is discrete} \\ \int_{-\infty}^{\infty} \mathbb{E}[X\,|\,Y=y] f_Y(y)\,dy & \text{if } Y \text{ is continuous.} \end{cases} \tag{6}$$

An important use of iterated expectation is that we can compute probabilities using conditioning, since probability of an event is simply the expectation of the corresponding indicator variable.

Exercise 1.3 (Iterated expectation for probability)**.** Let X, Y be RVs.

(i) For any $x\in\mathbb{R}$, show that $\mathbb{P}(X\le x) = \mathbb{E}[\mathbf{1}(X\le x)]$.

(ii) By using iterated expectation, show that

$$\mathbb{P}(X\le x) = \mathbb{E}_Y[\mathbb{P}(X\le x\,|\,Y)]. \tag{7}$$

In order to properly develop our discussion on martingales in the following sections, we need to generalize the notion of conditional expectation $\mathbb{E}[X\,|\,Y]$ of a RV X given another RV Y. Recall that this was the a collection of 'best guesses' of X given $Y = y$ for all y. But what if we only know, say, $Y \geq 1$? Can we condition on this event as well?

More concretely, suppose Y takes values from $\{1,2,3\}$. Regarding Y, the following outcomes are possible:

$$\mathscr{E}_Y := \{\{Y=1\},\{Y=2\},\{Y=3\},\{Y=1,2\},\{Y=2,3\},\{Y=1,3\},\{Y=1,2,3\}\}. \tag{8}$$

For instance, the information $\{Y = 1,2\}$ could yield some nontrival implication on the value of X, so our best guess in this scenario should be

$$\mathbb{E}[X\,|\,\{Y=1,2\}] = \sum_x x\mathbb{P}(X = x|\{Y=1,2\}). \tag{9}$$

More generally, for each $A \in \mathscr{E}_Y$, the best guess of X given $A \in \mathbb{E}_Y$ is the following conditional expectation

$$\mathbb{E}[X\,|\,A] = \sum_x x\mathbb{P}(X = x|A). \tag{10}$$

Now, what if we don't know which event in the collection $\mathscr{E}_Y$ to occur? As we did before to define $\mathbb{E}[X\,|\,Y]$ from $\mathbb{E}[X\,|\,Y = y]$ by simply not specifying what value y that Y takes, we simply do not specify which event $A \in \mathscr{E}_A$ to occur. Namely,

$$\mathbb{E}[X\,|\,\mathscr{E}_Y] = \text{best guess on } X \text{ given the information in } \mathscr{E}_Y. \tag{11}$$

In general, this could be defined for any collection of events $\mathscr{E}$ in place of $\mathscr{E}_Y$. Mathematically, we understand $\mathbb{E}[X\,|\,\mathscr{E}]$ as[1]

$$\mathbb{E}[X\,|\,\mathscr{E}] = \text{the collection of } \mathbb{E}[X\,|\,A] \text{ for all } A \in \mathscr{E}. \tag{12}$$

2. DEFINITION AND EXAMPLES OF MARTINGALES

Let $(X_t)_{t\geq 0}$ be the sequence of observations of the price of a particular stock over time. Suppose that an investor has a strategy to adjust his portfolio $(M_t)_{t\geq 0}$ according to the observation $(X_t)_{t\geq 0}$. Namely,

$$M_t = \text{Net value of portfolio after observing } (X_k)_{0\leq k\leq t}. \tag{13}$$

We are interested in the long-term behavior of the 'portfolio process' $(M_t)_{t\geq 0}$. Martingales provide a very nice framework for this purpose.

Martingale is a class of stochastic processes, whose expected increment conditioned on the past is always zero. Recall that the simple symmetric random walk has this property, since each increment is i.i.d. and has mean zero. Martingales do not assume any kind of independence between increments, but it turns out that we can proceed quite far with just the unbiased conditional increment property.

In order to define martingales properly, we need to introduce the notion of 'information up to time t'. Imagine we are observing the stock market starting from time t. We define

$$\mathscr{E}_t := \text{collection of all possible events we can observe at time } t \tag{14}$$

[1] For more details, see [Dur10].

$$\mathscr{F}_t := \bigcup_{k=1}^{t} \mathscr{E}_k = \text{collection of all possible events we can observe up to time } t. \tag{15}$$

In words, $\mathscr{E}_t$ is the information available at time t and $\mathscr{F}_t$ contains all possible information that we can obtain by observing the market up to time t. We call $\mathscr{F}_t$ the *information* up to time t. As a collection of events, $\mathscr{F}_t$ needs to satisfy the following properties[2]:

(i) (*closed under complementation*) $A \in \mathscr{F}_t \Longrightarrow A^c \in \mathscr{F}_t$;
(ii) (*closed under countable union*) $A_1, A_2, A_3, \cdots \in \mathscr{F}_t \Longrightarrow \bigcup_{k=1}^{\infty} A_k \in \mathscr{F}_t$.

Note that as we gain more and more information, we have

$$\mathscr{F}_s \subseteq \mathscr{F}_t \qquad \forall\, t \geq s \geq 0. \tag{16}$$

In other words, $(\mathscr{F}_t)_{t\geq 0}$ is an increasing set of information, which we call a *filtration*. The roll of a filtration is to specify what kind of information is observable or not, as time passes by.

Example 2.1. Suppose $(\mathscr{F}_t)_{t\geq 0}$ is a filtration generated by observing the stock price $(X_t)_{t\geq 0}$ of company A in New York. Namely, $\mathscr{E}_t$ consists of the information on the values of the stock price X_t at day t. Given $\mathscr{F}_{10}$, we know the actual values of $X_0, X_1, \cdots, X_{10}$. For instance, X_8 is not random given $\mathscr{F}_{10}$, but X_{11} could still be random. On the other hand, if $(Y_t)_{t\geq 0}$ is the stock price of company B in Hong Kong, then we may have only partial information for $Y_0, \cdots, Y_{10}$ given $\mathscr{F}_t$. ▲

Now we define martingales.

Definition 2.2. Let $(\mathscr{F}_t)_{t\geq 0}$ be a filtration and $(M_t)_{t\geq 0}$ be discrete-time stochastic processes. We call $(M_t)_{t\geq 0}$ a *martingale* with respect to $(\mathscr{F}_t)_{t\geq 0}$ if the following conditions are satisfied: For all $t \geq 0$,

(i) (*finite expectation*) $\mathbb{E}[|M_t|] < \infty$.
(ii) (*measurability*[3]) $\{M_t = m\} \in \mathscr{F}_t$ for all $m \in \mathbb{R}$.
(iii) (*conditional increments*) $\mathbb{E}[M_{t+1} - M_t \,|\, A] = 0$ for all $A \in \mathscr{F}_t$.

If (iii) holds with "=" replaced by "≤" (resp., "≥"), then $(M_t)_{t\geq 0}$ is called a *supermartingale* (resp., *submartingale*) with respect to $(X_t)_{t\geq 0}$, respectively.

When appropriate, we will abbreviate the condition (iii) as

$$\mathbb{E}[M_{t+1} - M_t \,|\, \mathscr{F}_t] = 0. \tag{17}$$

If martingale is a fair gambling strategy, then one can think of supermartingale and submartingale as unfavorable and favorable gambling strategies, resepctively. For instance, expected winning in gambling on an unfavorable game should be non-increasing in time. This is an immediate consequence of the definition and iterated expectation.

Proposition 2.3. *Let $(M_t)_{t\geq 0}$ be a stochastic process and let $(\mathscr{F}_t)_{t\geq 0}$ be a filtration.*

(i) *If $(M_t)_{t\geq 0}$ is a supermartingale w.r.t. filtration $(\mathscr{F}_t)_{t\geq 0}$, then $\mathbb{E}[M_n] \leq \mathbb{E}[M_m]$ for all $n \geq m \geq 0$.*
(ii) *If $(M_t)_{t\geq 0}$ is a submartingale w.r.t. filtration $(\mathscr{F}_t)_{t\geq 0}$, then $\mathbb{E}[M_n] \geq \mathbb{E}[M_m]$ for all $n \geq m \geq 0$.*
(iii) *If $(M_t)_{t\geq 0}$ is a martingale w.r.t. filtration $(\mathscr{F}_t)_{t\geq 0}$, then $\mathbb{E}[M_n] = \mathbb{E}[M_m]$ for all $n \geq m \geq 0$.*

---

[2]We are requiring $\mathscr{F}_t$ to be a σ-*algebra*, but we avoid using this terminology.

[3]In this case, we say "M_t is measurable w.r.t. $\mathscr{F}_t$", but we avoid using this terminology.

Proof. (ii) and (iii) follows directly from (i), so we only show (i). Let $(M_t)_{t\geq 0}$ be a supermartingale w.r.t. $(X_t)_{\geq 0}$. Recall that for each $m \in \mathbb{R}$, $\{M_t = m\} \in \mathscr{F}_t$. Hence

$$\mathbb{E}[M_{t+1} - M_t \,|\, M_t = m] \leq 0 \tag{18}$$

since M_t is a supermartingale. Hence by conditioning on the values of M_t,

$$\mathbb{E}[M_{t+1} - M_t] = \mathbb{E}[\mathbb{E}[M_{t+1} - M_t \,|\, M_t]] \leq 0. \tag{19}$$

Then the assertion follows by an induction. □

In order to get familiar with martingales, it is helpful to envision them as a kind of simple symmetric random walk. In general, one can subtract off the mean of a given random walk to make it a martingale.

Example 2.4 (Random walks)**.** Let $(X_t)_{t\geq 1}$ be a sequence of i.i.d. increments with $\mathbb{E}[X_i] = \mu < \infty$. Let $S_t = S_0 + X_1 + \cdots + X_t$. Then $(S_t)_{t\geq 0}$ is called a *random walk*. Define a stochastic process $(M_t)_{t\geq 0}$ by

$$M_t = S_t - \mu t. \tag{20}$$

For each $t \geq 0$, let $\mathscr{F}_t$ be the information obtained by observing $S_0, S_1, \cdots, S_t$. Then $(M_t)_{t\geq 0}$ is a martingale with respect to the filtration $(\mathscr{F}_t)_{t\geq 0}$. Indeed, we have

$$\mathbb{E}[|M_t|] = \mathbb{E}[|S_t - \mu t|] \leq \mathbb{E}[|S_t| + |\mu t|] = \mathbb{E}[|S_t|] + \mu t < \infty, \tag{21}$$

and for any $m \in \mathbb{R}$,

$$\{M_t = m\} = \{S_t - \mu t = m\} = \{S_t = m + \mu t\} \in \mathscr{F}_t. \tag{22}$$

Furthermore, Since X_{t+1} is independent from $S_0, \cdots, S_t$, it is also independent from any $A \in \mathscr{F}_t$. Hence

$$\mathbb{E}[M_{t+1} - M_t \,|\, A] = \mathbb{E}[X_{t+1} - \mu \,|\, A] \tag{23}$$

$$= \mathbb{E}[X_{t+1} - \mu] = \mathbb{E}[X_{t+1}] - \mu = 0. \tag{24}$$

▲

Example 2.5 (Products of indep. RVs)**.** Let $(X_t)_{t\geq 0}$ be a sequence of independent RVs such that $X_t \geq 0$ and $\mathbb{E}[X_t] = 1$ for all $t \geq 0$. For each $t \geq 0$, let $\mathscr{F}_t$ be the information obtained by observing $M_0, X_0, \cdots, X_t$. Define

$$M_t = M_0 X_1 X_2 \cdots X_t, \tag{25}$$

where M_0 is a constant. Then $(M_t)_{t\geq 0}$ is a martingale with respect to $(\mathscr{F}_t)_{t\geq 0}$. Indeed, the assumption implies $\mathbb{E}[|M_t|] < \infty$ and that $\{M_t = m\} \in \mathscr{F}_t$ for all $m \in \mathbb{R}$ since M_t is determined by $M_0, X_1 \cdots, X_t$. Furthermore, since X_{t+1} is independent from $X_1, \cdots, X_t$, for each $A \in \mathscr{F}_t$,

$$\mathbb{E}[M_{t+1} - M_t \,|\, A] = \mathbb{E}[M_t X_{t+1} - M_t \,|\, A] \tag{26}$$

$$= \mathbb{E}[(X_{t+1} - 1)(M_0 X_1 \cdots X_t) \,|\, A] \tag{27}$$

$$= \mathbb{E}[X_{t+1} - 1 \,|\, A]\, \mathbb{E}[(M_0 X_1 \cdots X_t) \,|\, A] \tag{28}$$

$$= \mathbb{E}[X_{t+1} - 1]\, \mathbb{E}[(M_0 X_1 \cdots X_t) \,|\, A] = 0. \tag{29}$$

This multiplicative model a reasonable for the stock market since the changes in stock prices are believed to be proportional to the current stock price. Moreover, it also guarantees that the price will stay positive, in comparison to additive models. ▲

Exercise 2.6 (Exponential martingale)**.** Let $(X_t)_{t\geq 0}$ be a sequence of i.i.d. RVs such that their moment generating function exists, namely, there exists $\theta > 0$ for which

$$\varphi(\theta) := \mathbb{E}[\exp(\theta X_k)] < \infty \qquad \forall k \geq 0. \tag{30}$$

Let $S_t = S_0 + X_1 + \cdots + X_t$. Define

$$M_t = \exp(\theta S_n)/\varphi(\theta)^t. \tag{31}$$

Show that $(M_t)_{t\geq 0}$ is a martingale with respect to filtration $(\mathscr{F}_t)_{t\geq 0}$, where $\mathscr{F}_t$ is the information obtained by observing $S_0, \cdots, S_t$.

The following lemma allows us to provide many examples of martingales from Markov chains.

Lemma 2.7. *Let $(X_t)_{t\geq 0}$ be a Markov chain on state space Ω with transition matrix P. For each $t \geq 0$, let f_t be a function $\Omega \to \mathbb{R}$ such that*

$$f_t(x) = \sum_{y\in\Omega} P(x,y) f_{t+1}(y) \qquad \forall x \in \Omega. \tag{32}$$

Then $M_t = f_t(X_t)$ defines a martingale with respect to filtration $(\mathscr{F}_t)_{t\geq 0}$, where $\mathscr{F}_t$ is the information obtained by observing $X_0, \cdots, X_t$.

Proof. First note that for any $x \in \Omega$,

$$\mathbb{E}[M_{t+1} - M_t \,|\, X_t = x] = \mathbb{E}\left[\left(\sum_{y\in\Omega} P(x,y) f_{t+1}(y)\right) - f_t(x)\right] = 0. \tag{33}$$

Now fix $A \in \mathscr{F}_t$. By conditioning on the value of X_t and using Markov property,

$$\mathbb{E}[M_{t+1} - M_t \,|\, A] = \mathbb{E}[f_{t+1}(X_{t+1}) - f_t(X_t) \,|\, A] \tag{34}$$

$$= \mathbb{E}\left[\mathbb{E}\left[f_{t+1}(X_{t+1}) - f_t(X_t) \,|\, A, X_t\right] \,|\, A\right] \tag{35}$$

$$= \mathbb{E}\left[\mathbb{E}\left[f_{t+1}(X_{t+1}) - f_t(X_t) \,|\, X_t\right] \,|\, A\right] = 0. \tag{36}$$

This shows the assertion. □

Example 2.8 (Simple random walk)**.** Let $(X_t)_{t\geq 1}$ be a sequence of i.i.d. RVs with

$$\mathbb{P}(X_k = 1) = p, \qquad \mathbb{P}(X_k = -1) = 1 - p. \tag{37}$$

Let $S_t = S_0 + X_1 + \cdots + X_t$. Note that $(S_t)_{t\geq 0}$ is a Markov chain on $\mathbb{Z}$. Define

$$M_t = \left(\frac{1-p}{p}\right)^{S_n}. \tag{38}$$

Then $(M_t)_{t\geq 0}$ is a martingale with respect to filtration $(\mathscr{F}_t)_{t\geq 0}$, where $\mathscr{F}_t$ is the information obtained by observing $S_0, \cdots, S_t$.

In order to see this, define a function $h_t(x) = ((1-p)/p)^x$. According to Lemma 2.7, it suffice to show that h is a harmonic function with respect to the Gambler's chain. Namely,

$$\sum_{y\in\mathbb{Z}} P(x,y) h_t(y) = p h(x+1) + (1-p) h(x-1) \tag{39}$$

$$= p\left(\frac{1-p}{p}\right)^{x+1} + (1-p)\left(\frac{1-p}{p}\right)^{x-1} \tag{40}$$
$$= (1-p)\left(\frac{1-p}{p}\right)^{x} + p\left(\frac{1-p}{p}\right)^{x} = \left(\frac{1-p}{p}\right)^{x} = h_t(x). \tag{41}$$

Hence by Lemma 2.7, $(M_t)_{t\geq 0}$ is a martingale with respect to the filtration $(\mathscr{F}_t)_{t\geq 0}$. ▲

Example 2.9 (Simple symmetric random walk)**.** Let $(X_t)_{t\geq 1}$ be a sequence of i.i.d. RVs with

$$\mathbb{P}(X_k = 1) = \mathbb{P}(X_k = -1) = 1/2. \tag{42}$$

Let $S_t = S_0 + X_1 + \cdots + X_t$. Note that $(S_t)_{t\geq 0}$ is a Markov chain on $\mathbb{Z}$. Define

$$M_t = S_t^2 - t. \tag{43}$$

Then $(M_t)_{t\geq 0}$ is a martingale with respect to $(S_t)_{t\geq 0}$.

For each $t \geq 0$, define a function $f_t : \mathbb{Z} \to \mathbb{R}$ by $f_t(x) = x^2 - t$. By Lemma 2.7, it suffices to check if $f_t(x)$ is the average of $f_{t+1}(y)$ with respect to the transition matrix of S_t. Namely,

$$\sum_{y\in\mathbb{Z}} P(x,y) f_{t+1}(y) = \frac{1}{2} f_{t+1}(x+1) + \frac{1}{2} f_{t+1}(x-1) \tag{44}$$
$$= \frac{(x+1)^2 - t}{2} + \frac{(x-1-t)^2 - t}{2} = x^2 - t = f_t(x). \tag{45}$$

Hence by Lemma 2.7, $(M_t)_{t\geq 0}$ is a martingale with respect to filtration $(\mathscr{F}_t)_{t\geq 0}$, where $\mathscr{F}_t$ is the information obtained by observing $S_0, \cdots, S_t$. ▲

3. Basic properties of martingales

In this section, we study some of the basic properties of martingales. We begin with the relation between martingale increments and convex functions. Recall that a function $\varphi : \mathbb{R} \to \mathbb{R}$ is *convex* if

$$\varphi(\lambda x + (1-\lambda)y) \leq \lambda\varphi(x) + (1-\lambda)(\varphi(y)), \qquad \forall\ \lambda \in [0,1] \text{ and } x, y \in \mathbb{R}. \tag{46}$$

For instance, $\varphi(x) = x^2$ and $\exp(x)$ are convex functions.

Exercise 3.1 (Jensen's inequality)**.** Let X be any RV with $\mathbb{E}[X] < \infty$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be any convex function, that is,

$$\varphi(\lambda x + (1-\lambda)y) \leq \lambda\varphi(x) + (1-\lambda)(\varphi(y)), \qquad \forall\ \lambda \in [0,1] \text{ and } x, y \in \mathbb{R}. \tag{47}$$

Jensen's inequality states that

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]. \tag{48}$$

(i) Let $c := \mathbb{E}[X] < \infty$. Show that there exists a line $f(x) = ax + b$ such that $f(c) = \varphi(c)$ and $\varphi(x) \geq f(x)$ for all $x \in \mathbb{R}$.

(ii) Verify the following and prove Jensen's inequality:

$$\mathbb{E}[\varphi(X)] \geq \mathbb{E}[f(X)] = a\mathbb{E}[X] + b = f(c) = \varphi(c) = \varphi(\mathbb{E}[X]). \tag{49}$$

(iii) Let X be RV, A an event, φ be the convex function as before. Show the Jensen's inequality for the conditional expectation:

$$\varphi(\mathbb{E}[X|A]) \leq \mathbb{E}[\varphi(X)\,|\,A]. \tag{50}$$

Proposition 3.2. *Let $(M_t)_{t\geq 0}$ be a submartingale with respect to a filtration $(\mathscr{F}_t)_{t\geq 0}$. Let $\varphi:\mathbb{R}\to\mathbb{R}$ be a convex function.*

(i) *If $(M_t)_{t\geq 0}$ is a martingale, then $(\varphi(M_t))_{t\geq 0}$ is a submartingale w.r.t. $(\mathscr{F}_t)_{t\geq 0}$.*
(ii) *If $\varphi:\mathbb{R}\to\mathbb{R}$ is a non-decreasing function, then $(\varphi(M_t))_{t\geq 0}$ is a submartingale w.r.t. $(\mathscr{F}_t)_{t\geq 0}$.*

Proof. We first show (i). Fix $A\in\mathscr{F}_t$. Since $(M_t)_{t\geq 0}$ is a martingale, for each $A'\in\mathscr{F}_t$,

$$\mathbb{E}[M_{t+1}\,|\,A'] - \mathbb{E}[M_t\,|\,A'] = \mathbb{E}[M_{t+1}-M_t\,|\,A'] = 0. \tag{51}$$

Note that for each $m\in\mathbb{R}$, since $\{M_t=m\}\in\mathscr{F}_t$, by using Jensen's inequality,

$$\mathbb{E}[\varphi(M_{t+1})-\varphi(M_t)\,|\,M_t=m] = \mathbb{E}[\varphi(M_{t+1})\,|\,M_t=m]-\varphi(m) \tag{52}$$
$$\geq \varphi(\mathbb{E}[M_{t+1}\,|\,M_t=m])-\varphi(m) \tag{53}$$
$$= \varphi(\mathbb{E}[M_t\,|\,M_t=m])-\varphi(m) = \varphi(m)-\varphi(m)=0. \tag{54}$$

Then by iterated expectation and using the fact that $A\cap\{M_t=m\}\in\mathscr{F}_t$,

$$\mathbb{E}[\varphi(M_{t+1})-\varphi(M_t)\,|\,A] = \mathbb{E}\left[\mathbb{E}[\varphi(M_{t+1})-\varphi(M_t)\,|\,A,M_t]\right]=0. \tag{55}$$

So $(\phi(M_t))_{t\geq 0}$ is a submartingale. This shows (i).

To show (ii), note that since $(M_t)_{t\geq 0}$ is a submartingale, for any $A'\in\mathscr{F}_t$,

$$\mathbb{E}[M_{t+1}\,|\,A'] - \mathbb{E}[M_t\,|\,A'] = \mathbb{E}[M_{t+1}-M_t\,|\,A'] \geq 0. \tag{56}$$

By Jensen's inequality and since φ is non-decreasing,

$$\mathbb{E}[\varphi(M_{t+1})\,|\,A'] \geq \varphi(\mathbb{E}[M_{t+1}\,|\,A']) \geq \varphi(\mathbb{E}[M_t\,|\,A']). \tag{57}$$

Then following the similar argument as before shows that $(\phi(M_t))_{t\geq 0}$ is a submartingale. □

Example 3.3. Let $(M_t)_{t\geq 0}$ be a martingale. Since $\varphi(x)=x^2$ is a convex function, $(M_t^2)_{t\geq 0}$ is a submartingale. ▲

The following observation is a martingale analogue of the formula $\mathbb{E}[X^2]-\mathbb{E}[X]^2=\mathrm{Var}(X)$.

Proposition 3.4. *If $(M_t)_{t\geq 0}$ is a martingale with respect to filtration $(\mathscr{F}_t)_{t\geq 0}$, then we have*

$$\mathbb{E}[M_{t+1}^2-M_t^2\,|\,\mathscr{F}_t] = \mathbb{E}[(M_{t+1}-M_t)^2\,|\,\mathscr{F}_t]\geq 0. \tag{58}$$

Proof. By the martingale condition, we have

$$\mathbb{E}[M_{t+1}-M_t\,|\,\mathscr{F}_t] = \tag{59}$$

Hence by expanding the square, we get

$$\mathbb{E}[(M_{t+1}-M_t)^2\,|\,\mathscr{F}_t] = \mathbb{E}[M_{t+1}^2\,|\,\mathscr{F}_t]-2M_t\mathbb{E}[M_{t+1}\,|\,\mathscr{F}_t]+\mathbb{E}[M_t^2\,|\,\mathscr{F}_t] \tag{60}$$
$$= \mathbb{E}[M_{t+1}^2\,|\,\mathscr{F}_t]-\mathbb{E}[M_t^2\,|\,\mathscr{F}_t]. \tag{61}$$

□

Exercise 3.5 (Long range martingale condition)**.** Let $(M_t)_{t\geq 0}$ be a martingale with respect to a filtration $(\mathscr{F}_t)_{t\geq 0}$. For any $0\leq k<n$, we will show that

$$\mathbb{E}[(M_n-M_k)\,|\,\mathscr{F}_k]=0. \tag{62}$$

(i) Suppose (62) holds for fixed $0 \le k < n$. For each $A \in \mathcal{F}_k$, to show that

$$\mathbb{E}[M_{n+1} - M_k \mid A] = \mathbb{E}[M_{n+1} - M_n \mid A] + \mathbb{E}[M_n - M_k \mid A] \tag{63}$$
$$= \mathbb{E}[M_{n+1} - M_n \mid A] = 0. \tag{64}$$

(ii) Conclude (62) for all $0 \le k < n$ by induction.

Proposition 3.6 (Orthogonality of martingale increments)**.** *Let $(M_t)_{t\ge0}$ be a martingale with respect to filtration $(\mathcal{F}_t)_{t\ge0}$.*

(i) *For each $0 \le j \le k < n$, we have*

$$\mathbb{E}[(M_n - M_k)M_j] = 0. \tag{65}$$

(ii) *For each $0 \le j \le k < n$, we have*

$$\mathbb{E}[(M_n - M_k)(M_j - M_i)] = 0. \tag{66}$$

Proof. To show (i), we use conditioning on the values of M_j. By noting that $\{M_j = m\} \in \mathcal{F}_j \subseteq \mathcal{F}_k$ for all $m \in \mathbb{R}$, we have

$$\mathbb{E}[(M_n - M_k)M_j \mid M_j = m] = m\mathbb{E}[M_n - M_k \mid M_j = m] = 0, \tag{67}$$

where we have used Exercise 3.5 for the last equality. Hence

$$\mathbb{E}[(M_n - M_k)M_j] = \mathbb{E}[\mathbb{E}[(M_n - M_k)M_j \mid M_j]] = 0. \tag{68}$$

Moreover, (ii) follows from (i) immediately since

$$\mathbb{E}[(M_n - M_k)(M_j - M_i)] = \mathbb{E}[(M_n - M_k)M_j] - \mathbb{E}[(M_n - M_k)M_i]. \tag{69}$$

This shows the assertion. □

Exercise 3.7 (Pythagorian theorem for martingales)**.** Let $(M_t)_{t\ge0}$ be a martingale with respect to a filtration $(\mathcal{F}_t)_{t\ge0}$. Use Proposition 3.6 to show that

$$\mathbb{E}[(M_t - M_0)^2] = \sum_{k=1}^{t} \mathbb{E}[(M_k - M_{k-1})^2]. \tag{70}$$

4. Gambling strategies and stopping times

In this section, we derive some properties of martingales in relation to stopping times viewed as gambling strategies. The following observation is natural if we think of supermartingale as a random walk with negative drift; or as betting on an unfavorable game.

Next, we prove one of the most famous result in martingale theory, which says

$$\text{“You can't beat an unfavorable game.”} \tag{71}$$

To formulate this, let $(M_t)_{t\ge0}$ be a supermartingale w.r.t. $(X_t)_{t\ge0}$. Think of $(X_t)_{t\ge0}$ as the progression of stock market, and $(M_t)_{t\ge0}$ is the price of the stock of company A. Say we have an investment strategy so that we can determine

$$H_{t+1} = \text{Amount of stock of company A that we hold between time } t \text{ and } t+1 \tag{72}$$

based on the observation $M_0, X_0, \cdots, X_t$. Then the stock price will change from M_t to M_{t+1},

$$H_{t+1}(M_{t+1} - M_t) = \text{Net gain occured between time } t \text{ and } t+1. \tag{73}$$

Hence if we let W_0 denote the initial fortune, then

$$W_t := W_0 + \sum_{k=1}^{t} H_k(M_k - M_{k-1}) = \text{Total fortune at time } t. \tag{74}$$

The following theorem tells us that, if $(M_t)_{t\geq 0}$ is declining stock on average, then no matter what strategy $(H_t)_{t\geq 0}$ we use, we will always lose our fortune.

Theorem 4.1. *Let $(M_t)_{t\geq 0}$ be a supermartingale w.r.t. a filtration $(\mathscr{F}_t)_{t\geq 0}$. Suppose $(H_t)_{t\geq 0}$ is such that*

(i) *(Predictability) $\{H_{t+1} = h\} \in \mathscr{F}_t$ for all $h \in \mathbb{R}$.*
(ii) *(Boundedness) $0 \leq H_t \leq c_t$ for some constant $c_t \geq 0$ for all $t \geq 0$.*

Let $(W_t)_{t\geq 0}$ be as defined at (74). Then $(W_t)_{t\geq 0}$ is a supermartingale w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$.

Proof. By (ii) and a triangle inequality,

$$|W_t| \leq |W_0| + \sum_{k=1}^{t} c_t(|M_k| + |M_{t-1}|). \tag{75}$$

Taking expectation shows $\mathbb{E}[|W_t|] < \infty$. Moreover, for each $w \in \mathbb{R}$,

$$\{W_t = w\} = \left\{ W_0 + \sum_{k=1}^{t} H_k(M_k - M_{k-1}) = w \right\} \in \mathscr{F}_t \tag{76}$$

since the event on the right hand side can be written as the union of suitable events involving the values of $H_0, \cdots, H_t$ and $M_0, \cdot, M_t$, which are all events belonging to $\mathscr{F}_t$.

Lastly, fix $A \in \mathscr{F}_t$. For each $h \in [0, c_t]$, since $\{H_{t+1} = h\} \in \mathscr{F}_t$, we also have $A \cap \{H_{t+1} = h\} \in \mathscr{F}_t$. Since M_t is a supermartingale, we have

$$\mathbb{E}[W_{t+1} - W_t \,|\, A \cap \{H_{t+1} = h\}] = \mathbb{E}[H_{t+1}(M_{t+1} - M_t) \,|\, A \cap \{H_{t+1} = h\}] \tag{77}$$

$$= h\mathbb{E}[M_{t+1} - M_t \,|\, A] \leq 0. \tag{78}$$

Hence, by conditioning on the values of H_{t+1}, we get

$$\mathbb{E}[W_{t+1} - W_t \,|\, A] = \mathbb{E}\left[\mathbb{E}[W_{t+1} - W_t \,|\, A, H_{t+1}] \,|\, A\right] \leq 0. \tag{79}$$

This shows the assertion. □

Next, we define stopping times with respect to a given filtration. This generalizes the version of stopping times we introduced for Poisson processes (See Def. 3.3 in Lecture note 3). The definition is as you would expect:

Definition 4.2. A RV $T \geq 0$ is a *stopping time* w.r.t. a filtration $(\mathscr{F}_t)_{t\geq 0}$ if

$$\{T = s\}, \{T \neq s\} \in \mathscr{F}_s \qquad \forall s \geq 0. \tag{80}$$

Example 4.3. Let $(N_1(t))_{t\geq 0}$ and $(N_2(t))_{t\geq 0}$ be the counting processes of PP(λ_1) and PP(λ_2) (not necessarily indep.). For each $t \geq 0$, let $\mathscr{F}_t$ be the information obtained by observing $(N_i(t))_{0\leq s\leq t}$ for $i = 1, 2$. Let $T_k^{(i)}$ denote the kth arrival time for the ith process.

Clearly, for each $i \in \{1, 2\}$ and $k \geq 0$, $T_k^{(i)}$ is a stopping time w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$. Namely, for each $t \geq 0$,

$$\{T_k^i = t\} = \{N^{(i)}(t) = k,\, N^{(i)}(t^-) = k - 1\} \in \mathscr{F}_t. \tag{81}$$

Also, let

$$\tilde{T}_k := k\text{the arrival time for the merged process } N_1 + N_2. \tag{82}$$

This is also a stopping time w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$. Namely, for each $t \geq 0$,

$$\{\tilde{T}_k = t\} = \{N^{(1)}(t) + N^{(2)}(t) = k,\, N^{(1)}(t^-) + N^{(2)}(t^-) = k-1\} \in \mathscr{F}_t. \tag{83}$$

On the other hand, let $(N^{(3)}(t))_{t\geq 0}$ be the counting process of another Poisson process of rate λ_3, which is independent from the other processes. Let $T_k^{(3)}$ denote the kth arrival time of this process. Is this a stopping time w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$? No, since for each $t \geq 0$, the event

$$\{T_k^{(3)} = t\} = \{N^{(3)}(t) = k,\, N^{(3)}(t^-) = k-1\} \tag{84}$$

cannot be determined from observing the other two processes $N^{(1)}$ and $N^{(2)}$ up to time t. ▲

Example 4.4 (Constant betting up to a stopping time)**.** Let $(M_t)_{t\geq 0}$ be a supermartingale w.r.t. $(X_t)_{t\geq 0}$. Let T be a stopping time for $(X_t)_{t\geq 0}$. Consider the gambling strategy of betting \$1 up to time T. In this case, the wealth at time t is

$$W_t = W_0 + \sum_{k=1}^{t} H_k(M_k - M_{k-1}) = M_{T\wedge t}, \tag{85}$$

where $T \wedge t = \min(T, t)$. The above relation can be easily verified by considering whether $t > T$ or $t \leq T$. ▲

Theorem 4.5. *Let $(M_t)_{t\geq 0}$ be a supermartingale w.r.t. a filtration $(\mathscr{F}_t)_{t\geq 0}$. Let T be a stopping time for $(\mathscr{F}_t)_{t\geq 0}$. Then $M_{T\wedge t}$ is a supermartingale w.r.t. $(\mathscr{F}_t)_{t\geq 0}$. Furthermore,*

$$\mathbb{E}[M_{T\wedge t}] \leq \mathbb{E}[M_0] \qquad \forall t \geq 0. \tag{86}$$

In particular, if $(M_t)_{t\geq 0}$ is a martingale, then

$$\mathbb{E}[M_{T\wedge t}] = \mathbb{E}[M_0] \qquad \forall t \geq 0. \tag{87}$$

Proof. Let W_t be as defined as in (4.4). Since $W_t = M_{T\wedge t}$, it suffices to show that W_t is a supermartingale. But this is the content of Theorem 4.1. Moreover, by Proposition (2.3), we have

$$\mathbb{E}[M_{T\wedge t}] \leq \mathbb{E}[W_t] \leq \mathbb{E}[W_0] = \mathbb{E}[M_0]. \tag{88}$$

The same argument holds for the submartingale case with the reversed inequalities. Then the martingale case follows. □

Exercise 4.6. Let $(M_t)_{t\geq 0}$ be a supermartingale w.r.t. a filtration $(\mathscr{F}_t)_{t\geq 0}$. We will directly show that $(M_{T\wedge t})_{t\geq 0}$ is a supermartingale w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$.

(i) For each fixed $A \in \mathscr{F}_t$, show that

$$A \cap \{T \leq t\} = A \cap \left(\bigcup_{k=1}^{t} \{T = k\} \right) \in \mathscr{F}_t, \tag{89}$$

$$A \cap \{T \geq t+1\} = A \cap \left(\bigcup_{k=1}^{t} \{T \neq k\} \right) \in \mathscr{F}_t. \tag{90}$$

(ii) Use conditioning on whether $T \leq t$ or not to show

$$\mathbb{E}[M_{T\wedge(t+1)} - M_{T\wedge t} \,|\, A] = \mathbb{E}[M_T - M_T \,|\, A \cap \{T \leq t\}]\mathbb{P}(T \leq t) \tag{91}$$

$$+ \mathbb{E}[M_{t+1} - M_t \,|\, A \cap \{T \geq t+1\}]\mathbb{P}(T \geq t+1) \tag{92}$$

$$\leq 0. \tag{93}$$

Conclude that $(M_{T\wedge t})_{t\geq 0}$ is a supermartingale w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$.

5. Applications of martingales at stopping times

If a martingale is bounded up to a stopping time, then the expected value of the martingale at stopping time equals the initial expectation. This observation will be useful in some applications.

Proposition 5.1. *Let $(Y_t)_{\geq 0}$ be a stochastic process such that $\mathbb{E}[|Y_t|] < \infty$ for all $t \geq 0$. Let $T \geq 0$ be a RV. Suppose*

$$\mathbb{P}(T < \infty) = 1, \quad |\mathbb{E}[Y_T \,|\, T > t]| < \infty, \quad |\mathbb{E}[Y_{T\wedge t} \,|\, T > t]| < \infty. \tag{94}$$

Then

$$\mathbb{E}[Y_T] = \lim_{t\to\infty} \mathbb{E}[Y_{T\wedge t}]. \tag{95}$$

Proof. We first condition on whether $T \leq t$ or $T > t$ to write

$$\mathbb{E}[Y_T] = \mathbb{E}[Y_T \,|\, T \leq t]\mathbb{P}(T \leq t) + \mathbb{E}[Y_T \,|\, T > t]\mathbb{P}(T > t) \tag{96}$$

$$= \mathbb{E}[Y_{T\wedge t} \,|\, T \leq t]\mathbb{P}(T \leq t) + \mathbb{E}[Y_T \,|\, T > t]\mathbb{P}(T > t). \tag{97}$$

Similarly, we also write

$$\mathbb{E}[Y_{T\wedge t}] = \mathbb{E}[Y_{T\wedge t} \,|\, T \leq t]\mathbb{P}(T \leq t) + \mathbb{E}[Y_{T\wedge t} \,|\, T > t]\mathbb{P}(T > t). \tag{98}$$

Subtracting these two equations, we get

$$|\mathbb{E}[Y_T] - \mathbb{E}[Y_{T\wedge t}]| \leq |\mathbb{E}[Y_T \,|\, T > t]|\,\mathbb{P}(T > t) + |\mathbb{E}[Y_{T\wedge t} \,|\, T > t]|\,\mathbb{P}(T > t). \tag{99}$$

Note that since $\mathbb{P}(T \leq t) \to \mathbb{P}(T < \infty) = 1$ as $t \to \infty$, $\mathbb{P}(T > t) \to 0$ as $t \to \infty$. Then by the hypothesis, the right hand side converges to zero as $t \to \infty$. This shows the assertion. □

An immediate consequence for martingales is the following:

Proposition 5.2. *Suppose $(M_t)_{\geq 0}$ is a martingale and T is a stopping time.*

$$\mathbb{P}(T < \infty) = 1, \quad |\mathbb{E}[M_T \,|\, T > t]| < \infty, \quad |\mathbb{E}[M_{T\wedge t} \,|\, T > t]| < \infty. \tag{100}$$

Then

$$\mathbb{E}[M_T] = \mathbb{E}[M_0]. \tag{101}$$

Proof. By Theorem 4.5 and Proposition 5.1,

$$\mathbb{E}[M_0] = \lim_{t\to\infty} \mathbb{E}[M_0] = \lim_{t\to\infty} \mathbb{E}[M_{T\wedge t}] = \mathbb{E}[M_T]. \tag{102}$$

□

Example 5.3 (Gambler's ruin)**.** Let $(X_t)_{t\geq 1}$ be i.i.d. RVs with

$$\mathbb{P}(X=1)=p, \qquad \mathbb{P}(X=-1)=1-p. \tag{103}$$

Let $S_t = S_0+X_1+\cdots+X_t$ with $S_0 = i$. Let $\mathscr{F}_t$ be the information obtained by observing $S_0,\cdots,S_t$. Let $h(x) = ((1-p)/p)^x$ as in Exercise 2.8. We have seen $M_t := h(S_t)$ is a martingale w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$.

Fix integers $a < S_0 < b$, and define

$$T = \min\{t\geq 0 \,|\, S_t = a \text{ or } S_t = b\}. \tag{104}$$

So T is the first time that the random walk hits a or b. Note that T is a stopping time, and since S_t is an irreducible Markov chain on $\mathbb{Z}$, its hitting time to any state is almost surely finite, so $\mathbb{P}(T<\infty)=1$. Moreover, $S_T, S_{T\wedge t}\in[a,b]$. Hence by Proposition 5.2, we have

$$\left(\frac{1-p}{p}\right)^i = \mathbb{E}[h(S_T)] = \left(\frac{1-p}{p}\right)^a \mathbb{P}(S_T=a) + \left(\frac{1-p}{p}\right)^b (1-\mathbb{P}(S_T=a)). \tag{105}$$

Assuming $p\neq 1/2$, solving for $\mathbb{P}(S_T=a)$, this gives

$$\mathbb{P}(S_T=a) = \frac{\left(\frac{1-p}{p}\right)^b - \left(\frac{1-p}{p}\right)^x}{\left(\frac{1-p}{p}\right)^b - \left(\frac{1-p}{p}\right)^a}. \tag{106}$$

▲

Example 5.4 (Duration of fair games)**.** Let $(X_t)_{t\geq 1}$ be i.i.d. RVs with

$$\mathbb{P}(X=1)=\mathbb{P}(X=-1)=1/2. \tag{107}$$

Let $S_t = S_0+X_1+\cdots+X_t$ with $S_0 = 0$. Let $\mathscr{F}_t$ be the information obtained by observing $S_0,\cdots,S_t$. Note that $(S_t)_{t\geq 0}$ itself is a martingale w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$. Also, we have seen that $M_t := S_t^2 - t$ is a martingale w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$ in Exercise 2.9.

Fix integers $a<0<b$, and define

$$T = \min\{t\geq 0 \,|\, S_t = a \text{ or } S_t = b\}. \tag{108}$$

Again T is a stopping time w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$. As before, we also have $\mathbb{P}(T<\infty)$ and $S_T, S_{T\wedge t}\in[a,b]$. Hence by Proposition 5.2, we get

$$0 = \mathbb{E}[S_0] = \mathbb{E}[S_T] = a\mathbb{P}(S_T=a) + b\mathbb{P}(S_T=b). \tag{109}$$

Since $\mathbb{P}(S_T=b) = 1-\mathbb{P}(S_T=a)$, the first equation yield

$$\mathbb{P}(S_T=a) = \frac{b}{b-a}, \quad \mathbb{P}(S_T=b) = \frac{-a}{b-a}. \tag{110}$$

For the other martingale, we need to stop it at time $T\wedge t$ since $M_{T\wedge t} = S_{T\wedge t}^2 - (T\wedge t)$ may depend on t. Proposition 5.2 yields

$$0 = \mathbb{E}[M_0] = \mathbb{E}[M_{T\wedge t}] = \mathbb{E}[S_{T\wedge t}^2 - (T\wedge t)] = \mathbb{E}[S_{T\wedge t}^2] - \mathbb{E}[T\wedge t]. \tag{111}$$

Notice that, as $t\to\infty$,

$$\mathbb{E}[T\wedge t] = \sum_{k=0}^{t} \mathbb{P}(T\geq k) \nearrow \sum_{k=0}^{\infty} \mathbb{P}(T\geq k) = \mathbb{E}[T]. \tag{112}$$

Also, since $S^2_{T\wedge t} \le \max(a^2, b^2)$, we can use Proposition 5.1 to get

$$\mathbb{E}[S_T^2] = \lim_{t\to\infty} \mathbb{E}[S^2_{T\wedge t}]. \tag{113}$$

Hence (111) yields $\mathbb{E}[T] = \mathbb{E}[S_T^2]$. Thus

$$\mathbb{E}[T] = a^2\mathbb{P}(S_T = a) + b^2\mathbb{P}(S_T = b) \tag{114}$$

$$= \frac{a^2 b}{b-a} - \frac{b^2 a}{b-a} = \frac{ab(a-b)}{b-a} = -ab. \tag{115}$$

This shows the expected duration of a fair game is $\mathbb{E}[T] = |ab|$.

Observe that for each fixed $a < 0$, $\mathbb{E}[T] = |ab| \to \infty$ as $b \to \infty$. In other words, if one starts gambling on fair coin flips, each time winning or losing \$1, then the expected time $\mathbb{E}[T]$ that he reaches $-\$1$, or ruins, is infinity. In terms of random walk, this shows that the first return time of a simple symmetric random walk has infinite expected value. In terms of $M/M/1$ queue, this shows that if the arrival and service rate is the same, then there is no stationary distribution for the queue size. ▲

Next, we recall the following computation of normal MGF.

Exercise 5.5 (MGF of normal RVs)**.** Let $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0,1)$. Using the fact that $\mathbb{E}[e^{tZ}] = e^{t^2/2}$, show that

$$\mathbb{E}[e^{tY}] = \exp(\sigma^2 t^2/2 + t\mu). \tag{116}$$

The following example illustrates how martingales can be used in a risk management situation.

Example 5.6 (Cramér's estimate of ruin)**.** Let S_n denote the total assets of an insurance company at the end of year n. During year n, premiums totaling $c > 0$ dollars are received, while claims totaling Y_n dolors are paid. Hence

$$S_n = S_{n-1} + c - Y_n. \tag{117}$$

Let $X_n = c - Y_n$ be the net profit during year n. We may assume that X_n's are i.i.d. with distribution $N(\mu, \sigma^2)$. We are interested in estimating the probability of bankruptcy. Namely, let

$$B = \{S_n < 0 \text{ for some } n \ge 1\} = \{\text{Bankruptcy}\}. \tag{118}$$

We will show that

$$\mathbb{P}(B) \le \exp\left(-\frac{2\mu S_0}{\sigma^2}\right). \tag{119}$$

Hence, the company will avoid bankruptcy by maximizing the mean profit per year μ and the initial asset S_0, while minimizing the uncertainty σ of profit per year.

To begin, notice that $S_n = S_0 + X_1 + \cdots + X_n$ is a random walk. Let $\mathscr{F}_t$ be the information obtained by observing $S_0, S_1, \cdots, S_t$. According to Exercise 5.5, we have

$$\varphi(\theta) := \mathbb{E}[e^{\theta X_k}] = \exp(\sigma^2\theta^2/2 + \theta\mu) < \infty. \tag{120}$$

Hence we can use the following exponential martingale (see Exercise 2.6)

$$M_t(\theta) := \frac{\exp(\theta S_n)}{\varphi(\theta)} \tag{121}$$

with respect to the filtration $(\mathscr{F}_t)_{t\geq 0}$. Moreover, note that

$$\varphi(-2\mu/\sigma^2) = \exp(2\mu^2/\sigma^2 - 2\mu^2/\sigma^2) = \exp(0) = 1. \tag{122}$$

Hence

$$M_t := M_t(-2\mu/\sigma^2) = \exp\left(-\frac{2\mu S_n}{\sigma^2}\right) \tag{123}$$

is a martingale w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$.

Let $T = \min\{k \geq 1 \,|\, S_k < 0\}$ be the first time of hitting negative asset, which is a stopping time w.r.t. the filtration $(\mathscr{F}_t)_{t\geq 0}$. Also note that

$$M_T = \exp\left(-\frac{2\mu S_T}{\sigma^2}\right) > 1 \quad \text{a.s.}, \tag{124}$$

since $S_T < 0$. Since S_t is a supercritical random walk ($\mu > 0$), $\mathbb{P}(T = \infty) > 0$. Hence we need to stop the martingale at truncated stopping time $T \wedge t$, using Proposition 5.2. Noting that $M_t \geq 0$ for all $t \geq 0$, this yields

$$\exp\left(-\frac{2\mu S_0}{\sigma^2}\right) = \mathbb{E}[M_0] = \mathbb{E}[M_{T\wedge t}] = \mathbb{E}[M_T \,|\, T \leq t]\,\mathbb{P}(T \leq t) + \mathbb{E}[M_t \,|\, T > t]\mathbb{P}(T > t) \tag{125}$$

$$\geq \mathbb{E}[M_T \,|\, T \leq t]\,\mathbb{P}(T \leq t) \geq \mathbb{P}(T \leq t). \tag{126}$$

Note that the last equality follows since S_T given $T \leq t$ is zero almost surely and $M_t \geq 0$ for all $t \geq 0$. To finish, note that $\mathbb{P}(T \leq t)$ is the probability of having bankruptcy by time t. Hence

$$\mathbb{P}(B) = \mathbb{P}(T < \infty) = \lim_{n\to\infty} \mathbb{P}(T \leq n) \leq \exp\left(-\frac{2\mu S_0}{\sigma^2}\right). \tag{127}$$

▲

References

[Dur10] Rick Durrett, *Probability: theory and examples*, Cambridge university press, 2010.