DISCRETE TIME MARTINGALES AND APPLICATIONS TO RANDOM WALKS

ADITYA RAMAN

Abstract. This paper discusses conditional expectations, discrete time martingales, and important theorems involving discrete time martingales. We prove Doob's Optional Stopping Theorem as well as the Martingale Convergence Theorem. Next, we discuss some applications of discrete time martingales to random walks. In specific, we solve the Gambler's Ruin problem by applying martingale theory to asymmetric simple random walks on the real line. This paper presupposes an elementary understanding of measure theory and measure theoretic probability.

Contents

Definitions and Notation

Definition 0.1. Given a random experiment, a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ consists of:

(i) A *sample space* Ω which is the set of all outcomes of the experiment. An individual outcome of the experiment, i.e. an element of Ω, is denoted by ω.
(ii) An *event space* $\mathcal{F}$ which denotes the σ-algebra of all measurable events, where an event is a collection of outcomes, i.e. a subset of Ω.
(iii) A *probability measure* $\mathbb{P}$ which is a set function from $\mathcal{F}$ to $[0, 1]$, assigning probabilities to measurable events.

Definition 0.2. Given a set A, the characteristic function of A, $\mathbb{1}_A$, is defined as follows:

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise} \end{cases}$$

Date: September 2022.

Definition 0.3. The *Borel σ-algebra* on the real numbers, $\mathcal{B}(\mathbb{R})$, is the smallest σ-algebra on $\mathbb{R}$ which contains all open sets in $\mathbb{R}$. In this paper, we denote $\mathcal{B} := \mathcal{B}(\mathbb{R})$.

Definition 0.4. Given a class $\mathcal{C}$ of subsets of Ω, the *σ-algebra generated by $\mathcal{C}$*, notated $\sigma(\mathcal{C})$, is the smallest σ-algebra containing $\mathcal{C}$. $\sigma(\mathcal{C})$ is the intersection of all σ-algebras containing $\mathcal{C}$.

Given a random variable X, $\sigma(X)$ is the smallest σ-algebra with respect to which X is $\mathcal{F}$-measurable, and consists of the sets $\{X \in B\}, B \in \mathcal{B}$.

Definition 0.5. Two sub σ-algebras $\mathcal{G}$ and $\mathcal{H}$ are *independent* if, for all $G \in \mathcal{G}$ and $H \in \mathcal{H}$, we have

$$\mathbb{P}(G \cap H) = \mathbb{P}(G)\mathbb{P}(H)$$

A random variable X is independent of a sub σ-algebra $\mathcal{G}$ if $\sigma(X)$ and $\mathcal{G}$ are independent.

Definition 0.6. A *stochastic process* $\{X_t \mid t \in T\}$ is a collection of random variables on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, indexed by some set T. In this paper, we consider discrete stochastic processes, indexed by the non-negative integers $\mathbb{N}_0$ or the natural numbers $\mathbb{N}$.

1. Conditional Expectation

We first introduce some motivation for the notion of conditional expectation by considering the case of two discrete random variables on a probability space. Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and X and Z are two discrete random variables, such that

$$X(\Omega) = \{x_1, x_2, \ldots, x_m\} \text{ and } Z(\Omega) = \{z_1, z_2, \ldots, z_n\}.$$

The expectation of X conditioned on the event $\{Z = z_j\}$ is defined as

$$\mathbb{E}(X \mid Z = z_j) = \sum_{i=1}^{m} x_i \mathbb{P}(X = x_i \mid Z = z_j).$$

This enables us to define a new random variable Y as follows:

$$Y(\omega) = \mathbb{E}(X \mid Z = z_j) := y_j \text{ on } \{\omega \in \Omega \mid Z(\omega) = z_j\}.$$

The above random variable Y is notated $\mathbb{E}(X|Z)$ and is called the conditional expectation of X given Z.

The disjoint sets $\{Z = z_j\}, j = 1, 2, \ldots, n$ partition the sample space Ω into n components on which the random variable Z is constant. The σ-algebra $\mathcal{G} = \sigma(Z)$ generated by Z consists of the sets $\{Z \in B\}, B \in \mathcal{B}$, and therefore consists of the 2^n possible disjoint unions of the sets $\{Z = z_j\}, j = 1, 2, \ldots, n$. Since $Y^{-1}(y_j) = \{Z = z_j\} \in \sigma(Z)$, we have that Y is $\mathcal{G}$-measurable. Moreover, since Y is constant

on the set $\{Z = z_j\}$, we have

$$\begin{aligned}\int_{\{Z=z_j\}} Y d\mathbb{P} = y_j \mathbb{P}(Z = z_j) &= \mathbb{E}(X \mid Z = z_j)\mathbb{P}(Z = z_j) \\ &= \sum_{i=1}^{m} x_i \mathbb{P}(X = x_i \mid Z = z_j)\mathbb{P}(Z = z_j) \\ &= \sum_{i=1}^{m} x_i \mathbb{P}(X = x_i \cap Z = z_j) \\ &= \int_{\{Z=z_j\}} X d\mathbb{P}.\end{aligned}$$

In other words, $\mathbb{E}(Y\mathbb{1}_{\{Z=z_j\}}) = \mathbb{E}(X\mathbb{1}_{\{Z=z_j\}})$. Since, for every G in $\mathcal{G}$, G is a disjoint union of sets of the form $\{Z = z_j\}$, we have that $\mathbb{1}_G$ is a sum of $\mathbb{1}_{\{Z=z_j\}}$'s. By linearity of expectation, it follows that $\mathbb{E}(Y\mathbb{1}_G) = \mathbb{E}(X\mathbb{1}_G)$. In other words,

$$\int_G Y d\mathbb{P} = \int_G X d\mathbb{P} \ \forall G \in \mathcal{G}.$$

The above example extends to a more general definition of conditional expectation with respect to a sub σ-algebra.

Theorem 1.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability measure space, and X be a random variable with finite expectation, i.e. $\mathbb{E}(|X|) < \infty$. Let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$. Then there exists a random variable Y such that:*

(i) Y is $\mathcal{G}$-measurable,
(ii) $\mathbb{E}(|Y|) < \infty$,
(iii) For every set $G \in \mathcal{G}$,

$$\int_G Y d\mathbb{P} = \int_G X d\mathbb{P}.$$

This random variable is almost surely unique, in the sense that, if there exists another random variable Y' satisfying the same properties, then $Y = Y'$, a.s. or $\mathbb{P}(Y = Y') = 1$.

Definition 1.2. A random variable satisfying the above properties is called a *version* of the conditional expectation of X given $\mathcal{G}$, and we write $Y = \mathbb{E}(X|\mathcal{G})$, a.s.

We do not prove the existence or almost sure uniqueness of the conditional expectation of X given $\mathcal{G}$ in this paper, but interested readers may refer to [1], pp. 85-87 for a proof.

Intuitively, we may think about the conditional expectation of X given $\mathcal{G}$ as being the best estimator of X given the partial information reflected in the coarser sub σ-algebra $\mathcal{G}$. In other words, given an outcome ω of a random experiment, the value $\mathbb{E}(X|\mathcal{G})(\omega)$ is the expected value of $X(\omega)$ given the set of values $Z(\omega)$ for every $\mathcal{G}$-measurable random variable Z.

Conditional expectation has several nice properties, similar to the expectation of a random variable. For example, conditional expectation is linear in the sense that

$$\mathbb{E}(aX + bY|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G}), \text{ a.s.}$$

Moreover, analogues of the Monotone Convergence Theorem (MCT), Fatou's Lemma, as well as the Dominated Convergence Theorem (DCT) hold true in the case of conditional expectations. For example, if

$$0 \leq X_n \nearrow X, \text{ then } \mathbb{E}(X_n|\mathcal{G}) \nearrow \mathbb{E}(X|\mathcal{G}), \text{ a.s.}$$

Some important properties which we shall refer to in the future are listed below.

Proposition 1.3.

(a) If X is $\mathcal{G}$-measurable, then $\mathbb{E}(X|\mathcal{G}) = X$, a.s.

(b) (Tower Property) If $\mathcal{H}$ is a sub σ-algebra of $\mathcal{G}$, then

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})|H) = \mathbb{E}(X|\mathcal{H}), \text{ a.s.}$$

(c) (Pulling out known factors) If Z is $\mathcal{G}$-measurable and bounded, then

$$\mathbb{E}(ZX|\mathcal{G}) = Z\mathbb{E}(X|\mathcal{G}), \text{ a.s.}$$

(d) (Independence) If X is independent of $\mathcal{G}$, then

$$\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X), \text{ a.s.}$$

Proof. Refer to [1], pp. 89-90. □

2. Martingales

Now that we have defined conditional expectation and listed some basic properties, we are ready to introduce the notion of martingales. Before that, we now consider a filtered space.

Definition 2.1. A *filtered space* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, P)$ consists of the usual probability space, as well as a *filtration* $\{\mathcal{F}_n\}_{n\in\mathbb{N}_0}$ which is a family of sub σ-algebras of $\mathcal{F}$ increasing in the sense that

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}.$$

Furthermore, we define

$$\mathcal{F}_\infty := \sigma(\bigcup_n \mathcal{F}_n) \subset \mathcal{F}$$

which is the smallest σ-algebra containing the entire family of sub σ-algebras. Intuitively, we may think of the filtration as the partial information available to us at time n, which consists of the values $Z(\omega)$ for every $\mathcal{F}_n$-measurable random variable Z. The increasing nature of the filtration reflects the idea that the partial information available to us increases as time elapses.

Definition 2.2. A stochastic process $X = \{X_n \mid n \in \mathbb{N}_0\}$ is called *adapted* to the filtration $\{\mathcal{F}_n\}$ if for each n, X_n is $\mathcal{F}_n$-measurable.

Definition 2.3. Given a stochastic process X, the *natural filtration* $\{\mathcal{F}_n\}_{n\in\mathbb{N}_0}$ is defined as

$$\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n).$$

In other words, the σ-algebra $\mathcal{F}_n$ is the smallest σ-algebra with respect to which $X_0, X_1, \ldots, X_n$ are measurable.

Note that, if $\{\mathcal{F}_n\}$ is the natural filtration, then the information we have about ω at time n consists of the values

$$X_0(\omega), X_1(\omega), \ldots, X_n(\omega).$$

Now we are ready to introduce the notion of martingales.

Definition 2.4. A stochastic process X is called a *martingale* relative to $(\{\mathcal{F}_n\}, P)$ if

(i) X is adapted,
(ii) $\mathbb{E}(|X_n|) < \infty \ \forall n$,
(iii) $\mathbb{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1}$, a.s.

A *supermartingale* is defined similarly, except that equality in property (iii) is replaced by $\mathbb{E}(X_n|\mathcal{F}_{n-1}) \leq X_{n-1}$. For *submartingales*, property (iii) is replaced by $\mathbb{E}(X_n|\mathcal{F}_{n-1}) \geq X_{n-1}$. We note that X is a submartingale if and only if $-X$ is a supermartingale. Moreover, a martingale is both a supermartingale and a submartingale. Furthermore, we note that if X is a martingale, then

$$\mathbb{E}(X_n - X_0|\mathcal{F}_{n-1}) = \mathbb{E}(X_n|\mathcal{F}_{n-1}) - \mathbb{E}(X_0|\mathcal{F}_{n-1}) = X_{n-1} - X_0$$

by Proposition 1.3(a), since X_0 is $\mathcal{F}_{n-1}$-measurable. Therefore, $X - X_0$ is a martingale, and hence we may restrict attention to martingales which are null at zero, i.e. $X_0 = 0$.

Before viewing martingales in terms of gambling and fair games, we give a simple example of a martingale.

Example 2.5 (Sums of independent zero mean random variables)**.** Let $X_1, X_2, ...$ be a sequence of independent random variables with finite expectation. Let $\mathbb{E}(X_k) = 0$ for all k. Define $S_0 := 0$ and

$$S_n = X_1 + X_2 + \ldots + X_n.$$

Further define $\mathcal{F}_0 := \{\varnothing, \Omega\}$ and

$$\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n).$$

Then, for $n \geq 1$, we have

$$\mathbb{E}(S_n|\mathcal{F}_{n-1}) = \mathbb{E}(S_{n-1}|\mathcal{F}_{n-1}) + \mathbb{E}(X_n|\mathcal{F}_{n-1}) = S_{n-1} + \mathbb{E}(X_n) = S_{n-1}.$$

by Proposition 1.3(a) and Proposition 1.3(d). Therefore, $S = \{S_n \mid n \in \mathbb{N}_0\}$ is a martingale.

We may consider martingales as describing fair games. For example, given a martingale $X = \{X_n \mid n \in \mathbb{N}_0\}$, we may think of the difference $X_n - X_{n-1}$ as the winnings per unit stake at time n, in a series of games played at times $n = 1, 2, \ldots$. By the definition of a martingale, we have

$$\mathbb{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1} = \mathbb{E}(X_{n-1}|\mathcal{F}_{n-1})$$

which implies

$$\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) = 0$$

by linearity of conditional expectation. Intuitively, this means that, given the history of the game up and until time $n-1$, the expected winnings per unit stake at time n is zero, or in other words, the game series played is fair. On the other hand, we have

$$\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) \leq 0$$

for supermartingales, or that the game series is unfavourable to the player.

Definition 2.6. A stochastic process $C = \{C_n \mid n \in \mathbb{N}\}$ is called *previsible* if C_n is $\mathcal{F}_{n-1}$-measurable.

When viewing martingales in terms of fair games, we may consider C_n as the stake or bet being played on game n. The condition that C_n must be $\mathcal{F}_{n-1}$-measurable simply means that the player is not clairvoyant, and that the value of the stake being played on game n is entirely determined by the history of the game series up and until time $n-1$.

Definition 2.7. Let X be a martingale and C be a previsible process. The *martingale transform of X by C* denoted by $Y := C \bullet X$ is defined as $Y_0 := 0$ and

$$Y_n = \sum_{k=1}^{n} C_k(X_k - X_{k-1}).$$

Once again, if we consider the difference $X_n - X_{n-1}$ as the winnings per unit stake at time n, and C_n as the stake being played on game n, then Y_n reflects the player's cumulative winnings at time n. Therefore, we may think of the stochastic process Y as being the *winnings process* of the game series.

Theorem 2.8. *If C is a bounded previsible process and X is a martingale (resp. supermartingale), then $C \bullet X$ is a martingale (resp. supermartingale) null at zero.*

Proof. Let $Y := C \bullet X$. Then we have

$$Y_n - Y_{n-1} = C_n(X_n - X_{n-1})$$

and therefore

$$\mathbb{E}(Y_n - Y_{n-1}|\mathcal{F}_{n-1}) = \mathbb{E}(C_n(X_n - X_{n-1})|\mathcal{F}_{n-1}) = C_n\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}).$$

by Proposition 1.3(c) since C_n is bounded and $\mathcal{F}_{n-1}$-measurable. By definition, the term on the right hand side of the equation is equal to zero if X is a martingale and less than or equal to zero if X is a supermartingale. □

3. Doob's Optional Stopping Theorem

In this section, we state and prove Doob's Optional Stopping Theorem. In order to do so, we must first define and explain what a stopping time is.

Definition 3.1. Let $\{\mathcal{F}_n\}$ be a filtration. A map $T : \Omega \to \mathbb{N}_0$ is called a *stopping time* if

$$\{T \le n\} = \{\omega \in \Omega \mid T(\omega) \le n\} \in \mathcal{F}_n \ \forall n$$

or equivalently

$$\{T = n\} = \{\omega \in \Omega \mid T(\omega) = n\} \in \mathcal{F}_n \ \forall n.$$

Intuitively, a stopping time defines an index or time at which the player decides to stop playing the game. Importantly, the decision to stop playing immediately after game n is entirely dependent on the history up and until time n, since $\{T = n\} \in \mathcal{F}_n$. We now give a simple example of a stopping time.

Example 3.2. Let A be an adapted process (with respect to a filtration $\{\mathcal{F}_n\}$) and let B be a Borel set. Define

$$T = \inf\{n \ge 0 \mid A_n \in B\}.$$

T defines the first instance at which the stochastic process hits the set B, and is also called a *hitting time*. We have

$$\{T \le n\} = \bigcup_{k=1}^{n} \{A_k \in B\} \in \mathcal{F}_n$$

since A is an adapted process. Hence, T defines a stopping time.

Definition 3.3. Let X be a martingale and T be a stopping time. Define a new stochastic process X^T, *the process X stopped at T*, as follows:

$$X_n^T(\omega) = X_{\min\{T(\omega),n\}}(\omega).$$

Before we prove Doob's Optional Stopping Theorem, we require the following lemma.

Lemma 3.4.

(a) If X is a martingale and T is a stopping time, then X^T is a martingale. In particular, $\mathbb{E}(X_n^T) = \mathbb{E}(X_{\min\{T,n\}}) = \mathbb{E}(X_0)\ \forall n$.

(b) If X is a supermartingale and T is a stopping time, then X^T is a supermartingale. In particular, $\mathbb{E}(X_n^T) \leq \mathbb{E}(X_0)\ \forall n$.

Proof. Let us first consider the case where X is a martingale. Once again, we think of the difference $X_n - X_{n-1}$ as the winnings per unit stake at time n. Consider a stake process in which the player always bets a single unit and stops betting immediately after time T. This enables us to define a stake process C^T, where

$$C_n^T = \mathbb{1}_{\{n \leq T\}}.$$

Clearly, C^T is bounded by 1. Additionally, C_n^T can only take on values 0 or 1, and

$$\{C_n^T = 0\} = \{T \leq n-1\} \in \mathcal{F}_{n-1}$$

since T is a stopping time. The above equation simply expresses that the bet played at time n is zero if and only if the stopping time is less than or equal to $n-1$. Furthermore,

$$\{C_n^T = 1\} = \{C_n^T = 0\}^c \in \mathcal{F}_{n-1}$$

since σ-algebras are closed under complementation.

It follows that C^T is a bounded previsible process and therefore $C^T \bullet X$ is a martingale null at zero by Theorem 2.8.

Next, the associated winnings process is $C^T \bullet X$, where

$$(C^T \bullet X)_n = \sum_{k=1}^{n} \mathbb{1}_{\{k \leq T\}}(X_k - X_{k-1}) = X_{\min\{T,n\}} - X_0.$$

Therefore, we have

$$C^T \bullet X = X^T - X_0.$$

It follows that $X^T - X_0$ is a martingale null at zero, and hence X^T is a martingale. By definition,

$$\mathbb{E}(X_n^T | \mathcal{F}_{n-1}) = X_{n-1}^T. \tag{3.5}$$

Taking expectations on both sides, we have

$$\mathbb{E}(X_n^T) = \mathbb{E}(\mathbb{E}(X_n^T | \mathcal{F}_{n-1})) = \mathbb{E}(X_{n-1}^T).$$

Recursively, it follows that $\mathbb{E}(X_n^T) = \mathbb{E}(X_0^T) = \mathbb{E}(X_0)$.

The proof of (b) follows by the exact same argument, with equality in (3.5) replaced by the supermartingale inequality $\mathbb{E}(X_n^T | \mathcal{F}_{n-1}) \leq X_{n-1}^T$. □

Theorem 3.6 (Doob's Optional Stopping Theorem)**.**

(a) Let X be a supermartingale and T be a stopping time. Then X_T is integrable and $\mathbb{E}(X_T) \leq \mathbb{E}(X_0)$ if one of the following conditions holds:
  (i) T is bounded;
  (ii) X is bounded and T is a.s. finite;
  (iii) $\mathbb{E}(T) < \infty$ and there exists $K > 0$ such that

$$|X_n - X_{n-1}| \leq K \ \forall n.$$

(b) If any of the above conditions holds and X is a submartingale, then X_T is integrable and $\mathbb{E}(X_T) \geq \mathbb{E}(X_0)$.
(c) If any of the above conditions holds and X is a martingale, then X_T is integrable and $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

Proof. Consider first the case where X is a supermartingale. By Lemma 3.4, X^T is a supermartingale such that $\mathbb{E}(X_n^T) \leq \mathbb{E}(X_0)$ for all n. In other words, $X_{\min\{n,T\}}$ is integrable and

$$\mathbb{E}(X_{\min\{n,T\}} - X_0) \leq 0 \ \forall n \tag{3.7}$$

If condition (i) holds, there exists $N > 0$ such that $T(\omega) \leq N$ for every $\omega \in \Omega$. Substituting $n = N$ into (3.7) yields $\mathbb{E}(X_T) \leq \mathbb{E}(X_0)$.

If condition (ii) holds, there exists $K > 0$ such that $|X_n| \leq K$ for every n. Since T is almost surely finite, we have pointwise convergence a.s.

$$\lim_n \left(X_{\min\{n,T\}} - X_0\right) = X_T - X_0$$

and

$$|X_{\min\{n,T\}} - X_0| \leq 2K \ \forall n.$$

By the Bounded Convergence Theorem,

$$\mathbb{E}(X_T - X_0) = \lim_n \mathbb{E}(X_{\min\{n,T\}} - X_0) \leq 0.$$

Finally, if condition (iii) holds, we have

$$|X_{\min\{n,T\}} - X_0| = \left| \sum_{k=1}^{\min\{n,T\}} (X_k - X_{k-1}) \right| \leq \sum_{k=1}^{\min\{n,T\}} |X_k - X_{k-1}| \leq KT \ \forall n.$$

Since $\mathbb{E}(KT) = K\mathbb{E}(T) < \infty$, then by the Dominated Convergence Theorem,

$$\mathbb{E}(X_T - X_0) = \lim_n \mathbb{E}(X_{\min\{n,T\}} - X_0) \leq 0.$$

For a proof of (b), we note that, since X is a submartingale, $-X$ is a supermartingale. Applying (a) to $-X$ yields that $-X_T$ is integrable and

$$\mathbb{E}(-X_T) \leq \mathbb{E}(-X_0)$$

which implies that $\mathbb{E}(X_T) \geq \mathbb{E}(X_0)$. Moreover, X_T is integrable since $\mathbb{E}(|-X_T|) = \mathbb{E}(|X_T|) < \infty$.

For a proof of (c), we have that X is a supermartingale as well as a submartingale. Applying (a) and (b) to X shows that X_T is integrable and $\mathbb{E}(X_T) = \mathbb{E}(X_0)$. □

4. The Martingale Convergence Theorem

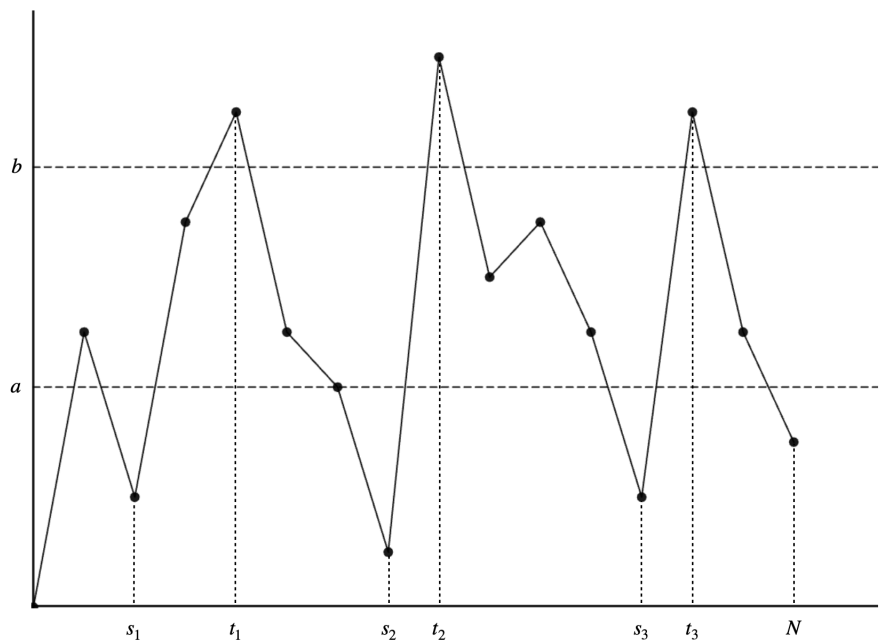
In this section, we discuss the convergence of supermartingales. Given a supermartingale X, it is a natural question to ask whether $\lim_n X_n$ exists. As we shall see, a necessary and sufficient condition for the limit to exist is for the supermartingale to be bounded in $\mathcal{L}^1$. In order to prove the Martingale Convergence Theorem, we first prove Doob's Upcrossing Lemma.

Definition 4.1. Given a stochastic process X and an outcome $\omega \in \Omega$, consider the sample path $n \mapsto X_n(\omega)$. Given $a, b \in \mathbb{R}$, the number $U_N[a,b](\omega)$ of *upcrossings* of $[a,b]$ made by the sample path by time N is defined as

$$\sup\{k \in \mathbb{N}_0 \mid 0 \leq s_1 < t_1 < s_2 < t_2 < \ldots < s_k < t_k \leq N \text{ and } X_{s_i}(\omega) < a, X_{t_i}(\omega) > b\}.$$

Intuitively, given a sample path, the number of upcrossings of an interval $[a,b]$ is the number of times the stochastic process passes upwards through the interval by time N. Figure 1 below illustrates the number of upcrossings for a given sample path. In this case, $U_N[a,b](\omega) = 3$.

Figure 1. Number of upcrossings of a sample path $n \mapsto X_n(\omega)$



Now consider a stochastic process X where $X_n - X_{n-1}$ defines the winnings per unit stake on game n. We define a betting strategy C as follows:

Fix $a, b \in \mathbb{R}$ with $a < b$.
WHILE TRUE
    IF $X < a$
        DO
            Play a unit stake
        WHILE $X \leq b$

We formally define $C_1 = \mathbb{1}_{\{X_0<a\}}$ and, for $n \geq 2$,

$$C_n = \mathbb{1}_{\{C_{n-1}=1\}}\mathbb{1}_{\{X_{n-1}\leq b\}} + \mathbb{1}_{\{C_{n-1}=0\}}\mathbb{1}_{\{X_{n-1}<a\}}.$$

The above equation simply states that the bet on game n, C_n, equals 1 if and only if the previous bet C_{n-1} equals 1, and the process X is not above b, i.e. $X_{n-1} \leq b$, or if the previous bet C_{n-1} equals 0 and the process X falls below a, i.e. $X_{n-1} < a$.

Proposition 4.2. *C is a previsible process with respect to the natural filtration $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$.*

Proof. In order to show that C is a previsible process, we proceed by induction on n.

Clearly C_1 is $\mathcal{F}_0$-measurable, since C_1 can only take on values 0 or 1, and

$$C_1^{-1}(1) = \{X_0 < a\} \in \mathcal{F}_0 = \sigma(X_0).$$

Therefore, the base case $n = 1$ holds. Now assume that C_n is $\mathcal{F}_{n-1}$-measurable. We have

$$C_{n+1} = \mathbb{1}_{\{C_n=1\}}\mathbb{1}_{\{X_n\leq b\}} + \mathbb{1}_{\{C_n=0\}}\mathbb{1}_{\{X_n<a\}}.$$

and

$$C_{n+1}^{-1}(1) = (\{C_n = 1\} \cap \{X_n \leq b\}) \cup (\{C_n = 0\} \cap \{X_n < a\}).$$

Since we assume that C_n is $\mathcal{F}_{n-1}$-measurable, then $\{C_n = 1\}, \{C_n = 0\} \in \mathcal{F}_{n-1} \subset \mathcal{F}_n$, and therefore it follows that $C_{n+1}^{-1}(1) \in \mathcal{F}_n$ since X is adapted to the natural filtration $\mathcal{F}_n$. By induction on n, we have that C is a previsible process. □

Proposition 4.3. *Given a stochastic process X and $Y := C \bullet X$ with C as defined above,*

$$Y_N(\omega) \geq (b-a)U_N[a,b](\omega) - (X_N(\omega) - a)^-.$$

Proof. Let s_i and t_i be defined as in Definition 4.1. For every upcrossing of $[a,b]$, we play unit bets in the times $s_{i+1}, s_{i+2}, \ldots, t_i$ and therefore $Y_N(\omega)$ is incremented by

$$\sum_{k=s_{i+1}}^{t_i} (X_k - X_{k-1}) = X_{t_i} - X_{s_i} > b - a.$$

In other words, each upcrossing of $[a,b]$ increments our total winnings by at least $b - a$.

On the other hand, the term $(X_N(\omega) - a)^-$ accounts for the game series following the last upcrossing by bounding the potential loss if X_N drops below a. Suppose that, after the final upcrossing, the process X drops below a at time r. Then we have $X_r < a$ and we play unit stakes until time N, since our process cannot exceed b in this time (otherwise there would be another upcrossing). Hence, the change to our winnings process is

$$\sum_{k=r+1}^{N} (X_k - X_{k-1}) = X_N - X_r > X_N - a.$$

Since X_N may be below a, the potential loss after the final upcrossing cannot exceed $(X_N - a)^-$. It follows that Proposition 4.3 holds true. □

With this inequality, we now state and prove Doob's Upcrossing Lemma.

Lemma 4.4 (Doob's Upcrossing Lemma)**.** *Suppose X is a supermartingale and $U_N[a,b]$ is the number of upcrossings of $[a,b]$ by X by time N. Then*

$$(b-a)\mathbb{E}(U_N[a,b]) \leq \mathbb{E}((X_N - a)^-).$$

Proof. The process C is previsible by Proposition 4.2 and is clearly bounded above by 1. Since $Y := C \bullet X$, then by Theorem 2.7, Y is a supermartingale null at zero. Therefore $\mathbb{E}(Y_N) \leq 0$. Since $Y_N \geq (b-a)U_N[a,b] - (X_N - a)^-$ by Proposition 4.3, then it follows that

$$\mathbb{E}((b-a)U_N[a,b] - (X_N - a)^-) \leq 0$$

which implies that

$$(b-a)\mathbb{E}(U_N[a,b]) \leq \mathbb{E}((X_N - a)^-).$$

□

Corollary 4.5. *Let X be a supermartingale bounded in $\mathcal{L}^1$ such that $\sup_n \mathbb{E}(|X_n|) < \infty$. Fix $a, b \in \mathbb{R}$ with $a < b$. Define $U_\infty[a,b] = \lim_N U_N[a,b]$. Then*

$$(b-a)\mathbb{E}(U_\infty[a,b]) \leq |a| + \sup_n \mathbb{E}(|X_n|) < \infty.$$

As a consequence,

$$\mathbb{P}(U_\infty[a,b] = \infty) = 0.$$

This Corollary states that, given a supermartingale bounded in $\mathcal{L}^1$, the number of upcrossings of any arbitrary interval $[a,b]$ is almost surely finite. We shall use this property to prove the Martingale Convergence Theorem.

Proof. By Lemma 4.4, we have

$$(b-a)\mathbb{E}(U_N[a,b]) \leq \mathbb{E}((X_N - a)^-) \leq \mathbb{E}(|X_N - a|) \leq \mathbb{E}(|X_N|) + |a|.$$

Therefore,

$$(b-a)\mathbb{E}(U_N[a,b]) \leq |a| + \sup_n \mathbb{E}(|X_n|).$$

Since $U_N[a,b]$ is monotonically non-decreasing in N (since the number of upcrossings of the interval $[a,b]$ as N increases can only increase, never decrease), then $U_N[a,b] \nearrow U_\infty[a,b]$. By the Monotone Convergence Theorem,

$$\mathbb{E}(U_\infty[a,b]) = \lim_N \mathbb{E}(U_N[a,b]) \leq |a| + \sup_n \mathbb{E}(|X_n|)$$

since the right hand side of the inequality is independent of N. This completes the proof. □

Theorem 4.6 (The Martingale Convergence Theorem)**.** *Let X be a supermartingale bounded in $\mathcal{L}^1$. Then, the sequence converges to a random variable X with finite expectation, i.e. $\mathbb{E}(|X|) < \infty$. As a result, X is almost surely finite.*

Proof. Define

$$\begin{aligned} S &:= \{\omega \in \Omega \mid X_n(\omega) \text{ does not converge to a limit in } [-\infty, \infty]\} \\ &= \{\omega \mid \liminf X_n(\omega) < \limsup X_n(\omega)\} \\ &= \bigcup_{\{a,b \in \mathbb{Q} \mid a<b\}} \{\omega \mid \liminf X_n(\omega) < a < b < \limsup X_n(\omega)\}. \end{aligned}$$

Let

$$S_{a,b} := \{\omega \mid \liminf X_n(\omega) < a < b < \limsup X_n(\omega)\}.$$

Then,

$$S = \bigcup_{\{a,b\in\mathbb{Q}\}} S_{a,b}.$$

However, if ω lies in $S_{a,b}$, then $X_{s_i}(\omega) < a < b < X_{t_i}(\omega)$ for infinitely many values of i, meaning that the process X crosses $[a,b]$ upwards infinitely often. Therefore, $\omega \in \{U_\infty[a,b] = \infty\}$. Hence,

$$S_{a,b} \subset \{U_\infty[a,b] = \infty\}$$

so that $\mathbb{P}(S_{a,b}) = 0$ by Corollary 4.5 and the monotonicity of measure. Since S is a countable union of sets $S_{a,b}$, it follows that $\mathbb{P}(S) = 0$. As a result,

$$X = \lim_n X_n \text{ exists a.s. in } [-\infty, \infty].$$

Moreover, by Fatou's Lemma, we have

$$\mathbb{E}(|X|) = \mathbb{E}(\liminf |X_n|) \leq \liminf \mathbb{E}(|X_n|) \leq \sup \mathbb{E}(|X_n|) < \infty$$

so that X has finite expectation. Hence, X is finite almost surely. □

5. Applications to Random Walks

In this section, we discuss applications of martingale theory to random walks. In particular, we solve a famous problem known as The Gambler's Ruin Problem by considering asymmetric simple random walks. We first describe the problem below:

Problem 5.1 (Gambler's Ruin Problem)**.** *Consider an unfair coin-flipping game with two players, where Player 1 has probability of winning p and Player 2 has probability of winning $q = 1-p$, where $p, q \in (0,1)$ with $p+q = 1$ and $p \neq q$. Player 1 starts with a dollars whereas Player 2 starts with b dollars. After each coin flip, the loser transfers 1 dollar to the winner. The game ends when one player has all the money. What are the respective probabilities of Player 1 and Player 2 winning the game?*

In order to arrive at an answer, we model the problem in terms of an asymmetric random walk, which we first define below.

Definition 5.2. Consider a sequence of independent, identically distributed (i.i.d.) random variables $X_1, X_2, \ldots$ with

$$\mathbb{P}(X_i = 1) = p \text{ and } \mathbb{P}(X_i = -1) = q$$

where

$$p, q \in (0,1),\ p + q = 1,\ \text{and } p \neq q.$$

Define $S_0 := 0$ and

$$S_n = \sum_{i=1}^{n} X_i.$$

Then, the stochastic process $S = \{S_n \mid n \in \mathbb{N}_0\}$ is known as an *asymmetric random walk.*

In the case of the above problem, let us define the random variable X_i as follows:

$$X_i = \begin{cases} 1, & \text{if Player 1 wins the } i\text{th game} \\ -1, & \text{if Player 2 wins the } i\text{th game} \end{cases}$$

Then, $\mathbb{P}(X_i = 1) = p$ and $\mathbb{P}(X_i = -1) = q$. Further define $S_0 := 0$ and

$$S_n = \sum_{i=1}^{n} X_i$$

such that the process S is an asymmetric simple random walk.

As defined above, S_n measures the cumulative gain/loss of Player 1 at time n since the start of the game, since the process S is incremented by 1 each time Player 1 wins, and reduced by 1 each time Player 1 loses.

Next, we define

$$\mathcal{F}_n := \sigma(X_1, X_2, \ldots, X_n)$$

and a stochastic process $M = \{M_n \mid n \in \mathbb{N}_0)$ as follows:

$$M_n := \left(\frac{q}{p}\right)^{S_n}.$$

Since the process X is clearly adapted to the natural filtration $\{\mathcal{F}_n\}$, we have that X_i is $\mathcal{F}_i$-measurable and therefore $\mathcal{F}_n$-measurable for $1 \le i \le n$. It follows that S_n is $\mathcal{F}_n$-measurable, and hence M_n is $\mathcal{F}_n$-measurable.

Additionally, since $|X_i| \le 1$, we have

$$|M_n| \le \max\left\{\left(\frac{q}{p}\right)^{n}, \left(\frac{q}{p}\right)^{-n}\right\}$$

such that each M_n is bounded and has finite expectation.

Furthermore,

$$\begin{aligned}
\mathbb{E}(M_n|\mathcal{F}_{n-1}) &= \mathbb{E}\left(M_{n-1}\left(\frac{q}{p}\right)^{X_n} \mid \mathcal{F}_{n-1}\right) \\
&= M_{n-1}\mathbb{E}\left(\left(\frac{q}{p}\right)^{X_n} \mid \mathcal{F}_{n-1}\right) \\
&= M_{n-1}\mathbb{E}\left(\left(\frac{q}{p}\right)^{X_n}\right) \\
&= M_{n-1}\left(p\left(\frac{q}{p}\right) + q\left(\frac{q}{p}\right)^{-1}\right) \\
&= M_{n-1}.
\end{aligned}$$

Here, we used Proposition 1.3(c) since M_{n-1} is bounded and $\mathcal{F}_{n-1}$-measurable, as well as Proposition 1.3(d) since X_n is independent of $\mathcal{F}_{n-1}$. It follows that M is a martingale.

Now, let us define

$$T_a := \inf\{n \ge 0 \mid S_n = -a\} \text{ and } T_b := \inf\{n \ge 0 \mid S_n = b\}.$$

Clearly, T_a is a stopping time since

$$\{T_a \le n\} = \bigcup_{i=0}^{n}\{S_i = -a\} \in \mathcal{F}_n.$$

By the same argument, T_b is a stopping time as well. We further define

$$T := \min\{T_a, T_b\}.$$

We have that T is a stopping time as well, since

$$\{T \le n\} = \{T_a \le n\} \cup \{T_b \le n\} \in \mathcal{F}_n.$$

Further note that the event $\{T = T_a\}$ is the set of outcomes for which the random walk hits the boundary $-a$ before hitting the boundary b. In other words, $\{T = T_a\} = \{\text{Event Player 1 loses}\}$. Similarly, $\{T = T_b\} = \{\text{Event Player 1 wins}\}$. We are therefore interested in calculating $\mathbb{P}(T = T_a)$ and $\mathbb{P}(T = T_b)$.

Proposition 5.3. $\mathbb{E}(T) < \infty$. *As a consequence, T is a.s. finite.*

Proof. Let $m = b + a$. We divide up the random variables (X_i) into sets $A_1, A_2, \ldots$ as follows:

$$\begin{aligned} A_1 &= \{X_1, X_2, \ldots, X_m\} \\ A_2 &= \{X_{m+1}, X_{m+2}, \ldots, X_{2m}\} \\ A_3 &= \{X_{2m+1}, X_{2m+2}, \ldots, X_{3m}\} \\ &\vdots \end{aligned}$$

such that each set A_k contains exactly m of the X_i's. Define the event

$$E_k := \{\omega \in \Omega \mid X(\omega) = 1 \text{ for all } X \in A_k\}.$$

If the event E_k occurs, then the random walk moves to the right by m steps during the time $(k-1)m+1, (k-1)m+2, \ldots, km$, and necessarily exits the interval $[a, b]$ by crossing above b.

If $T(\omega) > (k-1)m$, then we have $-a < S_{(k-1)m} < b$. If the event E_k then occurs, we must have $S_{km} > b$ and therefore

$$T \leq km. \tag{5.4}$$

We may therefore consider $E_1, E_2, \ldots$ as a sequence of opportunities for the random walk to exit the interval $[-a, b]$ by crossing above b.

By the independence of the X_i's, we have $\mathbb{P}(E_k) = p^m$ for every k. We define a new random variable K as follows:

$$K(\omega) = \inf\{k \in \mathbb{N} \mid \omega \in E_k\}.$$

We note that K is a geometric random variable with probability of success parameter p^m. Therefore, $\mathbb{E}(K) = \frac{1}{p^m} < \infty$ which implies that K is a.s. finite. Moreover, by the definition of K, ω almost surely lies in $E_{K(\omega)}$, which implies that $T \leq Km$ by (5.4). It follows that

$$\mathbb{E}(T) \leq \mathbb{E}(Km) = m\mathbb{E}(K) < \infty.$$

Hence, T is finite almost surely. □

Next, we consider M^T, the process M stopped at T, which is a martingale by Lemma 3.4. Since $M_n^T = M_{\min\{n,T\}}$, we have

$$M_{\min\{n,T\}} \leq \max\left\{\left(\frac{q}{p}\right)^{-a}, \left(\frac{q}{p}\right)^{b}\right\}$$

for all n. It follows that M^T is a bounded martingale.

Next, since M^T is a bounded martingale and T is a stopping time that is a.s. finite, then by Doob's Optional Stopping Theorem, we have

$$\mathbb{E}(M_T^T) = \mathbb{E}(M_0^T) = \mathbb{E}(M_0) = 1.$$

However, $M_T^T = M_{\min\{T,T\}} = M_T$. Therefore, we have

$$\mathbb{E}(M_T) = 1. \tag{5.5}$$

Next, we wish to calculate the probabilities $\mathbb{P}(T = T_a)$ and $\mathbb{P}(T = T_b)$ in order to know the respective probabilities of the random walk hitting each boundary at time T. Since $\mathbb{P}(T < \infty) = 1$, then $T = T_a$ or $T = T_b$ almost surely. In other words,

$$\mathbb{P}(T = T_a) + \mathbb{P}(T = T_b) = 1. \tag{5.6}$$

Moreover, by partitioning the expectation $\mathbb{E}(M_T)$ on whether the event $\{T = T_a\}$ occurs or not, we have

$$\begin{aligned}\mathbb{E}(M_T) &= \mathbb{E}(M_T \mathbb{1}_{\{T=T_a\}}) + \mathbb{E}(M_T \mathbb{1}_{\{T=T_b\}}) \\ &= \mathbb{E}\left(\left(\frac{q}{p}\right)^{-a} \mathbb{1}_{\{T=T_a\}}\right) + \mathbb{E}\left(\left(\frac{q}{p}\right)^{b} \mathbb{1}_{\{T=T_b\}}\right) \\ &= \mathbb{P}(T = T_a)\left(\frac{q}{p}\right)^{-a} + \mathbb{P}(T = T_b)\left(\frac{q}{p}\right)^{b}.\end{aligned}$$

From (5.5), we have that $\mathbb{E}(M_T) = 1$. It follows that

$$\mathbb{P}(T = T_a)\left(\frac{q}{p}\right)^{-a} + \mathbb{P}(T = T_b)\left(\frac{q}{p}\right)^{b} = 1. \tag{5.7}$$

Finally, solving the system of linear equations (5.6) and (5.7) gives

$$\mathbb{P}(T = T_a) = \frac{\left(\frac{q}{p}\right)^{a} - \left(\frac{q}{p}\right)^{a+b}}{1 - \left(\frac{q}{p}\right)^{a+b}} \tag{5.8}$$

and

$$\mathbb{P}(T = T_b) = 1 - \mathbb{P}(T = T_a) = \frac{1 - \left(\frac{q}{p}\right)^{a}}{1 - \left(\frac{q}{p}\right)^{a+b}}. \tag{5.9}$$

The above equations (5.8) and (5.9) give the probabilities of Player 2 winning and Player 1 winning respectively, and so we are done.

Acknowledgements

I would like to thank my mentor, Chris Wilson, for all the invaluable feedback and advice provided during the planning as well as writing of this paper. I would also like to thank Professor Peter May for organising this wonderful REU program, which helped me gain invaluable insight into what research in mathematics looks like.

References

[1] David Williams. Probability with Martingales. Cambridge University Press, 1991.
[2] Patrick Billingsley. Probability and Measure. `https://www.colorado.edu/amath/sites/default/files/attached-files/billingsley.pdf`.
[3] Nic Freeman. Martingale Theory and Applications. `https://people.maths.bris.ac.uk/~mabat/MARTINGALE_THEORY_2016/NIC_FREEMAN_notes.pdf`.
[4] Sebastien Roch. Lecture Notes on Measure-theoretic Probability Theory. `https://people.math.wisc.edu/~roch/grad-prob/index.html`.
[5] Desmos. Graphing Calculator. `https://www.desmos.com/calculator`.