MARTINGALES AND THEIR APPLICATIONS

MORGAN SHEEHAN

Abstract. The central topic of this expository paper is martingales. We will introduce the notion of a martingale and explore some key martingale results. We will also explore several examples of martingales in discrete and continuous time such as Polya's urn and Brownian motion. Martingales will also be applied to prove several key results in probability and in algebra.

Contents

1. Introduction and Motivation

Imagine you are playing a game involving a flip of a fair coin that costs one dollar to play. If the result of the flip is heads, you are paid two dollars, meaning you profit one dollar. If the flip is tails, you are paid nothing, thus resulting in a loss of the dollar you wagered to play the game. While there is nothing you can do to influence the result of one round of this game to increase your chances of winning, one may wonder whether you can employ certain strategies to increase your chances of winning over the course of several rounds. In particular, what would happen if you changed the amount you wagered each round depending on the result of the previous rounds?

In the 18th century in France, gamblers asked questions like this about games like the one we just described (in fact, the aforementioned game was one example of such a game). The strategies they came up with were called martingales. One example of such a strategy is one for the game we described above where the player would double their bet each time they lost, meaning if the coin came up heads before they ran out of money, they would walk away netting a dollar of profit.

Since then, the term martingale has appeared frequently in literature on the theory behind fair games like the one described in the paragraph above, and the term has come to be the name of a fair game [12]. Throughout this paper, we will use probability to uncover some of the theory behind these fair games. Then we will see that this theory developed can be used to prove some key theorems in probability. Interestingly enough, we will also use martingales to prove a key theorem in algebra, which is pretty amazing considering martingales were originally just a collection of betting strategies—something that seems completely unrelated to algebra.

2. Measure Theory and Probability Background

For the purposes of this paper, we assume that the reader has familiarity with introductory measure theory and probability. The background information necessary to proceed can be found in the second, third, and twenty-first chapters of [8]. Now, we turn our attention to the convergence of random variables. We will present a definition followed by a useful lemma and its partial converse.

Definition 2.1. Let $\{X_n\}$ be a sequence of random variables and X be a random variable defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say X_n converges to X almost surely if

$$\mathbb{P}\left[\lim_{n\to\infty} X_n = X\right] = 1.$$

Lemma 2.2 (Borel-Cantelli Lemma)**.** *Let $A_1, A_2, \ldots$ be an infinite sequence of independent random variables. If* $\sum_{n=1}^{\infty} \mathbb{P}[A_n] < \infty$*, then* $\mathbb{P}[A_n \text{ i.o.}] = 0$.

Proposition 2.3 (Partial Converse to the Borel-Cantelli Lemma)**.** *As before, let $A_1, A_2, \ldots$ be an infinite sequence of independent events. If* $\sum_{n=1}^{\infty} \mathbb{P}[A_n] = \infty$*, then* $\mathbb{P}[A_n \text{ i.o.}] = 1$.

A proof for the Borel-Cantelli Lemma and its partial converse can be found in [8, p. 254]. Now with this lemma and proposition, we will pose a motivating question for this paper involving Bernoulli-p random variables.

Consider an infinite sequence of Bernoulli-p random variables where $p \in (0, 1)$. Note that we restrict p so that $p \in (0, 1)$ because $p = 0$ or $p = 1$ will result in an infinite sequence of 0s or 1s respectively. A natural question to ask about this sequence is what will the longest streak of 1s be. How about the longest streak of 0s? The proposition below provides an answer to that question.

Proposition 2.4. *Given an infinite sequence of Bernoulli-p random variables where $p \in (0, 1)$, arbitrarily long sequences of 1s will occur infinitely many times.*

Although this conclusion can be reached through an application a proof involving the partial converse to the Borel-Cantelli lemma, we will show it is true by applying the Kolmogorov 0-1 law. However, this proof will be saved until Section 5, where the Kolmogorov 0-1 law will be proved with martingales.

Now, we will briefly discuss various results in the last important piece of background information for this paper: conditional probability.

Definition 2.5. An $\mathcal{L}^p$ norm is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where

$$||X||_p = \mathbb{E}[X^p]^{\frac{1}{p}}.$$

We say that a random variable is bounded in $\mathcal{L}^p$ if there is a $K \in \mathbb{R}^+$ such that $\mathbb{E}[|X|^p] \le K$.

Definition 2.6. An $\mathcal{L}^p$ space, typically written as $\mathcal{L}^p\,(\Omega, \mathcal{F}, \mathbb{P})$, is $\{X : ||X||_p < \infty\}$.

Remark 2.7. One can verify that the way we have defined $\mathcal{L}^p$ spaces technically defines an equivalence class since if X and Y are two random variables that agree everywhere except on a set of points with probability 0, then they are in the same $\mathcal{L}^p$ space. However, for the purposes of this paper, we will treat $\mathcal{L}^p$ as a space of functions rather than an equivalence class of functions.

Definition 2.8. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{G}$ be a sub sigma-algebra of $\mathcal{F}$. If $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, the conditional expectation of X given $\mathcal{G}$, denoted $\mathbb{E}[X \mid \mathcal{G}]$, is given by some random variable $Z \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ (that is unique up to a set of probability 0) such that for every bounded $\mathcal{G}$-measurable random variable Y,

$$\mathbb{E}[XY] = \mathbb{E}[ZY].$$

In particular, if $Y = \mathbb{1}_G$ for some $G \in \mathcal{G}$, we have

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

Remark 2.9. The existence of such a $Z \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ is not trivial; however, a proof that there is a $Z \in \mathcal{L}^1(\Omega, \mathcal{G}, \mathbb{P})$ satisfying $\mathbb{E}[XY] = \mathbb{E}[ZY]$ for every $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ will not be given in this paper but can be found in [3].

Now that we have defined conditional expectation, we will introduce a few key properties from [3]. The proofs for these properties will not be presented in this paper.

Proposition 2.10. *For all of these propositions, let $X, Y \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{G}$ and $\mathcal{H}$ be sub sigma-algebras of $\mathcal{F}$.*

(a) $\mathbb{E}[aX + Y \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + \mathbb{E}[Y \mid \mathcal{G}]$.
(b) *If* $Y \leq X$, $\mathbb{E}[Y \mid \mathcal{G}] \leq \mathbb{E}[X \mid \mathcal{G}]$.
(c) *If Y is $\mathcal{G}$-measurable, then* $\mathbb{E}[XY \mid \mathcal{G}] = Y\mathbb{E}[X \mid \mathcal{G}]$.
(d) $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X]$.
(e) *If $\mathcal{H}$ is a sub sigma-algebra of $\mathcal{G}$, then* $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}]$.
(f) *For any scalar, a,* $\mathbb{E}[a \mid \mathcal{G}] = a$.
(g) *If $f : \mathbb{R} \mapsto \mathbb{R}$ is convex and $\mathbb{E}[|X|] < \infty$, then $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$ and $\mathbb{E}[f(X) \mid Y] \geq f(\mathbb{E}[X \mid Y])$ where Y is a $\mathcal{G}$-measurable random variable. Furthermore, $\mathbb{E}[f(X) \mid \mathcal{G}] \geq f(\mathbb{E}[X \mid \mathcal{G}])$.*

Note that although they will not be presented in this paper, conditional probability does have two results that are analogous to the montone convergence theorem and the dominated convergence theorem. These can also be found in [3]. Finally, we introduce a few final definitions that are greatly useful in the study of martingales.

Definition 2.11. Let $\mathcal{F}$ be a sigma-algebra. A filtration $\{\mathcal{F}_n : n \geq 0\}$ is a set of increasing sub sigma-algebras of $\mathcal{F}$. In other words,

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}.$$

Furthermore, we define

$$\mathcal{F}_\infty = \sigma\left(\bigcup_n \mathcal{F}_n\right) \subseteq \mathcal{F}.$$

Definition 2.12. A filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_n, \mathbb{P})$ is a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\{\mathcal{F}_n : n \geq 0\}$.

Definition 2.13. A process $X = \{X_n : n \geq 0\}$ is called adapted to a filtration $\{\mathcal{F}_n\}$ if at each n, X_n is $\mathcal{F}_n$ measurable.

3. Martingales in Discrete Time

Definition 3.1. A process $\{X_n\}$ is called a martingale relative to $(\Omega, \mathcal{F}, \mathcal{F}_n, \mathbb{P})$ if

(i) $\{X_n\}$ is adapted.
(ii) $\mathbb{E}[X_n] < \infty$ for all n.
(iii) $\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] = X_{n-1}$ a.s. $(n \geq 1)$.

Submartingales and Supermartingales are defined in an identical manner except instead of having $\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] = X_{n-1}$ a.s. $(n \geq 1)$, a supermartingale has the condition $\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] \leq X_{n-1}$ a.s. $(n \geq 1)$ and a submartingale has $\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] \geq X_{n-1}$ a.s. $(n \geq 1)$. An intuitive way to think about martingales is by thinking of a fair game. By the same token, submartingales or supermartingales can be thought of as unfair games that are in your favor in the submartingale case or rigged against you in the supermartingale case. We proceed with a couple examples of martingales.

Example 3.2. Let $X_1, X_2, \ldots$ be an infinite sequence of independent random variables with mean 1. Define $M_0 = 1$ and let $M_n = \prod_{k=1}^{n} X_k$.
This is a martingale since

$$\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[M_n X_{n+1} \mid \mathcal{F}_n] = M_n \mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = M_n \mathbb{E}[X_{n+1}] = M_n.$$

Example 3.3. Consider an urn with a white ball and a black ball at time 0. At each time $t = 1, 2, \ldots$, a ball is chosen at random from the urn, and a new ball that is the same color as the ball chosen is placed in the urn. For example, if a black ball is chosen at time $t = 1$, another black ball will be placed in the urn, so the urn would have 2 black balls and 1 white ball. Let B_t be the number of black balls placed in the urn at time t (excluding the one black ball originally in the urn). We will show that a process $\{M_t\}$ defined as $M_t = \frac{B_t+1}{t+2}$ is a martingale with respect to the filtration $\mathcal{F}_t = \sigma(B_1, B_2, \ldots, B_t)$. Let $\mathbb{1}_t$ be the indicator variable of if you put in a black ball at time t. Thus,

$$\mathbb{P}[\mathbb{1}_t = 1 \mid B_{t-1} = k] = \frac{k+1}{t+1} \text{ if and only if } \mathbb{E}[\mathbb{1}_t \mid \mathcal{F}_{t-1}] = \frac{B_{t-1}+1}{t+1}.$$

To show this is a martingale, we will consider $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}]$ and show that it is equal to M_{t-1}.

$$\begin{aligned}
\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] &= \mathbb{E}\left[\frac{B_{t-1} + \mathbb{1}_t + 1}{t+2} \mid \mathcal{F}_{t-1}\right] \\
&= \frac{B_{t-1}+1}{t+2} + \frac{1}{t+2}\mathbb{E}[\mathbb{1}_t \mid \mathcal{F}_{t-1}] \\
&= \frac{B_{t-1}+1}{t+2} + \frac{B_{t-1}+1}{(t+2)(t+1)} \\
&= \frac{B_{t-1}+1}{t+1} = M_{t-1}.
\end{aligned}$$

We therefore see that $\{M_t\}$ is a martingale.

We now proceed with a few Martingale properties before introducing the notions of previsible strategies and stopping times.

Proposition 3.4. *$\{X_n\}$ is a submartingale with respect to $(\Omega, \mathcal{F}, \mathbb{P})$ if and only if $\{-X_n\}$ is a supermartingale with respect to $(\Omega, \mathcal{F}, \mathbb{P})$.*

Proof. $\{X_n\}$ is a submartingale if and only if $\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] \geq X_{n-1}$ if and only if $-\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] \leq -X_{n-1}$ if and only if $\mathbb{E}[-X_n \mid \mathcal{F}_{n-1}] \leq -X_{n-1}$ if and only if $\{-X_n\}$ is a supermartingale. □

Proposition 3.5 (Tower Property for Martingales)**.** *If $\{X_n\}$ is a martingale with respect to $(\Omega, \mathcal{F}, \mathbb{P})$, then*

$$X_m = \mathbb{E}[X_n \mid \mathcal{F}_m] \quad \text{for all} \quad n > m.$$

Proof. We will show this by induction. First, from the definition of a martingale, it follows that

$$X_m = \mathbb{E}[X_{m+1} \mid \mathcal{F}_m].$$

Next, we inductively assume that $X_m = \mathbb{E}[X_{n-1} \mid \mathcal{F}_m]$. By our induction hypothesis, the definition of a martingale, and Proposition 2.10 e, $X_m = \mathbb{E}[X_{n-1} \mid \mathcal{F}_m] = \mathbb{E}[\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] \mid \mathcal{F}_m] = \mathbb{E}[X_n \mid \mathcal{F}_m]$, meaning we can deduce that $\mathbb{E}[X_n \mid \mathcal{F}_m] = X_m$ for all $n > m$. □

Definition 3.6. A process $\mathcal{C} = \{C_n : n \in \mathbb{N}\}$ is previsible if $C_n \in \mathcal{F}_{n-1}$ for each n. Intuitively, one can think of C_n as the bet placed on the nth round of a game. Therefore, your total winnings in the nth round is given by $C_n (X_n - X_{n-1})$.

Definition 3.7. We define the martingale transform of $\{X_n\}$ by C as

$$(C \cdot X)_n = \sum_{k=1}^{n} C_k (X_k - X_{k-1}).$$

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

Proposition 3.8. *Suppose C is a bounded, previsible strategy such that for some $K \in [0, \infty)$, $|C_n(\omega)| \leq K$ for all n, ω. If $\{X_n\}$ is a martingale then so is $(C \cdot X)_n$.*

Proof. As before, we will write $\{Y_n\}$ for $(C \cdot X)_n$. Hence,

$$\mathbb{E}[Y_n - Y_{n-1} \mid \mathcal{F}_{n-1}] = C_n \mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] = 0.$$

Therefore, we see that $\{Y_n\}$ is a martingale. □

Definition 3.9. A stopping time is a random variable $T : \Omega \to \mathbb{Z}^+ \cup \{0\}$ such that

$$\{T \leq n\} = \{\omega : T(\omega) \leq n\} \in \mathcal{F}_n.$$

Equivalently,

$$\{T = n\} = \{\omega : T(\omega) = n\} \in \mathcal{F}_n.$$

Note that T can be ∞ and that constants are stopping times.

To show that these two definitions are equivalent, let T be a stopping time defined in the first way. Therefore,

$$\{T = n\} = \{T \leq n\} \setminus \{T \leq n-1\} \in \mathcal{F}_n.$$

Now, if T is a stopping time by the second definition, note that

$$\{T \leq n\} = \bigcup_{k=1}^{n} \{T = k\} \in \mathcal{F}_n.$$

Notation 3.10. The process $\left\{X_n^T\right\}$ means $\{X_{T\wedge n}\}$, where $T\wedge n$ denotes $\min(T, n)$, and this process is called a martingale stopped at time T. A previsible process for a martingale with a stopping time T can be written as $C_n^{(T)}$. If we bet 1 unit on a game and quit at time T, then your stake for each round is written as $C_n^{(T)} = \mathbb{1}_{\{n\leq T\}}$, and your winnings are given by $\left(C^{(T)} \cdot X\right)_n$.

Lemma 3.11. *If $\{X_n\}$ is a martingale, then $\left\{X_n^T\right\}$ is a martingale. In particular,*

$$\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0].$$

Proof. First, define a process C_n where

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

Note that

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

Hence, $(C \cdot X)_n$ is precisely $\left\{X_n^T\right\}$, and this process is $\mathcal{F}_{n-1}$-measurable and hence previsible since

$$\mathbb{1}_{\{n\leq T\}} = \mathbb{1}_{\{n-1<T\}} = 1 - \mathbb{1}_{\{T\leq n-1\}}.$$

From Proposition 3.8, we know that this means that $(C \cdot X)_n$ is a martingale, meaning $\left\{X_n^T\right\}$ is a martingale.
Since $\mathbb{E}[(C \cdot X)_n] = 0$, we see that $\mathbb{E}[X_n^T - X_0] = 0$. Therefore,

$$\mathbb{E}[X_n^T] = \mathbb{E}[X_0].$$

□

Now that we have introduced some basic Martingale results, we turn our attention to some key theorems.

Theorem 3.12 (Doob's Optional Sampling Theorem)**.** *Let T be a stopping time and $\{X_n\}$ be a martingale. If one of the following is true, then $\mathbb{E}[X_T] = \mathbb{E}[X_0]$*

(i) T is bounded (there exists $N \in \mathbb{Z}^+$ such that $T \leq N$ a.s.).
(ii) $\{X_n\}$ is bounded (there exists $K > 0$ such that $|X_n| \leq K$ for all n) and T is a.s. finite.
(iii) $\mathbb{E}[T] < \infty$ and for some $K \in \mathbb{R}^+$, $|X_n(\omega) - X_{n-1}(\omega)| \leq K$ for all $\omega \in \Omega$ and $n \in \mathbb{N}$.

Proof. From Lemma 3.11, we know that $\mathbb{E}[X_{n\wedge T}] = \mathbb{E}[X_0]$.

If the first condition is true, we know that then we can take $n = N$, and it follows that $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ since $N \wedge T = T$. If the second condition is true, we know that X is dominated by some value K.

Hence, we can apply the dominated convergence theorem and get

$$\mathbb{E}[X_0] = \lim_{n\to\infty} \mathbb{E}[X_{n\wedge T}] = \mathbb{E}\left[\lim_{n\to\infty} X_{n\wedge T}\right] = \mathbb{E}[X_T].$$

Finally, if the third condition is true, then

$$X_{n\wedge T} \leq |X_0| + \sum_{k=0}^{n\wedge T} |X_{k+1} - X_k| \leq |X_0| + KT.$$

Note that T is a.s. finite since $\mathbb{E}[T] < \infty$, meaning $X_{n\wedge T}$ is bounded above, and, again, we can apply the dominated convergence theorem to get

$$\mathbb{E}[X_0] = \lim_{n\to\infty} \mathbb{E}[X_{n\wedge T}] = \mathbb{E}[\lim_{n\to\infty} X_{n\wedge T}] = \mathbb{E}[X_T].$$

□

Remark 3.13. One can also show using a similar proof that given any of the three conditions above, if $\{X_n\}$ is a supermartingale then $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$.

Both Proposition 3.8 and Theorem 3.12 are important because they show that given a fair game, no matter what strategy you use regarding your stake on each round or when you decide to stop playing the game, it will always remain fair.

To conclude this section about martingales, we turn our attention to proving the martingale convergence theorem, which will first require a couple definitions, lemmas, and a corollary.

Definition 3.14. The number of upcrossings $U_N[a,b]\,(\omega)$ of $[a,b]$ made by $n \mapsto X_n\,(\omega)$ is defined to be the largest k such that

$$0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_k < t_k \leq N$$

such that for all $1 \leq j \leq k$, $X_{s_i} < a$ and $X_{t_i} > b$.

Notation 3.15. For a value $x \in \mathbb{R}$, we write $(x)^+$ for $\max(0,x)$ and $(x)^-$ for $\max(0,-x)$.

From the definition of an upcrossing, we have the following inequality which serves as a lower bound for winnings at time N (which are denoted by $Y_N\,(\omega)$ where $\{Y_n\} = (C\cdot X)_n$, where C is previsible):

$$Y_N\,(\omega) \geq (b-a)\,U_N[a,b]\,(\omega) - (X_N\,(\omega) - a)^- \,. \tag{3.16}$$

This equality comes from the fact that for each upcrossing, the minimum payout is $b-a$ since $X_s < a$ and $X_t > b$. The $(X_N\,(\omega) - a)^-$ accounts for the fact that you may lose on the Nth round.

Lemma 3.17. *If $\{X_n\}$ is a supermartingale and $U_N[a,b]$ is the number of upcrossings by time N, then*

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

Proof. As shown in Proposition 3.8, we know that the process $\{Y_n\}$ will be a supermartingale, meaning $\mathbb{E}[Y_N] \leq 0$. Combining this with (3.16), we have

$$0 \geq \mathbb{E}[Y_N\,(\omega)] \geq \mathbb{E}[(b-a)\,U_N[a,b]\,(\omega) - (X_N\,(\omega) - a)^-].$$

By linearity of expectations, we can rearrange the terms from this inequality to get

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

□

Lemma 3.18. *Let $a, b \in \mathbb{R}$ and let $\{X_n\}$ be a supermartingale bounded in $\mathcal{L}^1$ such that*

$$\sup_n \mathbb{E}[|X_n|] < \infty.$$

Then, given $U_N[a,b] \uparrow U_\infty[a,b]$,

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

Proof. From Lemma 3.17, we know

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

By the triangle inequality,

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

Therefore, we see that

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

Taking $N \to \infty$, which can be done by the monotone convergence theorem since $U_N[a,b] \uparrow U_\infty[a,b]$, we get

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

which is the desired result. □

Corollary 3.19. $\mathbb{P}[U_\infty[a,b] = \infty] = 0$.

This corollary comes from the fact that $\mathbb{P}[U_\infty[a,b] = \infty] \neq 0$ would contradict our lemma.

We now turn our attention to the final main result of this section.

Theorem 3.20 (Martingale Convergence Theorem)**.** *Let $\{X_n\}$ be a supermartingale bounded in $\mathcal{L}^1$. Then $\{X_n\}$ converges a.s. to a random variable X_∞, and $\lim_{n\to\infty} X_n$ is a.s. finite.*

Proof. Consider $\{\omega \;:\; X_n(\omega) \text{ does not converge}\}$

$$= \left\{\omega \;:\; \liminf_{n\to\infty} X_n(\omega) < \limsup_{n\to\infty} X_n(\omega)\right\}$$

$$= \bigcup_{\{a,b\in\mathbb{Q}:a<b\}} \left\{\omega \;:\; \liminf_{n\to\infty} X_n(\omega) < a < b < \limsup_{n\to\infty} X_n(\omega)\right\}$$

$$= \bigcup_{\{a,b\in\mathbb{Q}:a<b\}} \{\omega \;: U_\infty[a,b](\omega) = \infty\}$$

Note that for fixed $a, b \in \mathbb{Q}$ where $a < b$, we have shown that

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

Therefore,

$$\mathbb{P}\left[\bigcup_{\{a,b\in\mathbb{Q}:a<b\}} \{\omega \;: U_\infty[a,b](\omega) = \infty\}\right] = 0$$

since it is a countable union of sets each with probability measure 0.

Now, to show X_∞ is finite, we have

$$\begin{aligned}\mathbb{E}[|X_\infty|] &= \mathbb{E}\left[|\lim_{n\to\infty} X_n|\right] \\ &= \mathbb{E}\left[|\liminf_{n\to\infty} X_n|\right] \\ &\le \liminf_{n\to\infty} \mathbb{E}[|X_n|] \\ &\le \sup_n \mathbb{E}[|X_n|] \\ &< \infty.\end{aligned}$$

The first inequality can be justified by Fatou's lemma.

Hence, $\{X_n\}$ converges a.s. to a random variable X_∞, and $\lim_{n\to\infty} X_n$ is a.s. finite. □

4. Random Walks

Definition 4.1. Consider an infinite sequence of independent identically distributed random variables $X_1, X_2, \dots$ such that $\mathbb{P}[X = 1] = \mathbb{P}[X = -1] = \frac{1}{2}$. A random walk on $\mathbb{Z}$ is defined such that $S_0 = 0$ and $S_n = \sum_{k=1}^{n} X_k$.

One can easily verify that this is a martingale because if after n steps you are at a point P, we have

$$\begin{aligned}\mathbb{E}[S_{n+1} \mid S_n = P] &= \mathbb{P}[X_{n+1} = 1]\,(P+1) + \mathbb{P}[X_{n+1} = -1]\,(P-1) \\ &= \frac{1}{2}\,(P+1) + \frac{1}{2}\,(P-1) = P.\end{aligned}$$

Inductively, one can show that the expected position after n steps is 0 for every positive n. A natural question to consider is what is the probability of being at 0 after n steps.

The first observation needed to solve this problem is that it is impossible to be at the origin after an odd number of steps. Next, observe that we are at the origin after $2m$ steps for $m \in \mathbb{Z}$ if exactly m of the steps are 1 and m of the steps are -1. Therefore, we get the expression

$$\mathbb{P}[S_{2m} = 0] = \binom{2m}{m}\left(\frac{1}{2}\right)^{2m} = \frac{(2m)!}{(m!)^2\, 2^{2m}}.$$

Another natural question that arises about a random walk is how many times will a random walk return to 0. Before answering this question, we will introduce some terminology important to this topic.

Definition 4.2. A process is said to be recurrent if it will return to a state infinitely many times. A process is said to be transient if it will return to a state finitely many times.

Returning to our question, by the partial converse to the Borel-Cantelli lemma, if we show that $\sum_{k=1}^{\infty} \mathbb{P}[S_k = 0]$ diverges, then the random walk returns to the origin infinitely often and is recurrent. To solve this, however, we must better understand the behavior of $\frac{(2n)!}{(n!)^2 2^{2n}}$ as $n \to \infty$, which motivates the need for an approximation of $n!$ for large n. This approximation is called Stirling's approximation, and it asserts that $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1 + O\left(n^{-1}\right)\right)$. Below, we will prove a slightly weaker version of Stirling's approximation that will work for our purposes.

Lemma 4.3. *There is an $A \in \mathbb{R}$ such that as $n \to \infty$,*

$$\log(n!) = n\log(n) - n + \frac{1}{2}\log(n) + A + O\left(n^{-1}\right).$$

Equivalently, this implies that as $n \to \infty$,

$$n! = e^A\sqrt{n}\left(\frac{n}{e}\right)^n\left(1 + O\left(n^{-1}\right)\right).$$

Proof. Set $a_n = n\log(n) - n + \frac{1}{2}\log(n)$, and let us evaluate $a_{n+1} - a_n$, We compute

$$\begin{aligned} a_{n+1} - a_n &= (n+1)\log(n+1) - n\log(n) - 1 + \frac{1}{2}\log\left(\frac{n+1}{n}\right) \\ &= n\log\left(1 + \frac{1}{n}\right) + \log(n+1) - 1 + \frac{1}{2}\log\left(1 + \frac{1}{n}\right). \end{aligned}$$

By Taylor's formula, we see that $\log\left(1 + \frac{1}{n}\right) = \frac{1}{n} - \frac{1}{2n^2} + O\left(n^{-3}\right)$. Therefore,

$$\begin{aligned} a_{n+1} - a_n &= 1 - \frac{1}{2n} + \log(n+1) - 1 + \frac{1}{2}\left(\frac{1}{n} - \frac{1}{2n^2}\right) + O\left(n^{-3}\right) \\ &= \log(n+1) + O\left(n^{-2}\right). \end{aligned}$$

Thus, there is some constant C such that

$$|a_n - a_{n+1} - \log(n+1)| \leq \frac{C}{n^2}.$$

Hence the series $\sum\limits_{n=1}^{\infty} a_n - a_{n+1} - \log(n+1)$ converges to some constant A at speed $O\left(n^{-1}\right)$, and we have

$$\log n! - a_n - A = O\left(n^{-1}\right),$$

which is what we set out to prove. □

Remark 4.4. The constant A is known to be $\log\left(\sqrt{2\pi}\right)$, where log denotes the natural logarithm; however, that is not important for the purposes of this paper.

Now that we have established Stirling's Approximation, we get that

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}[S_n = 0] &= \sum_{n=1}^{\infty} \frac{(2n)!}{(n!)^2 2^{2n}} \\ &\approx \sum_{n=1}^{\infty} \frac{e^A\sqrt{2n}\left(\frac{2n}{e}\right)^{2n}}{\left(e^A\sqrt{n}\left(\frac{n}{e}\right)^n\right)^2 2^{2n}} \\ &= \sum_{n=1}^{\infty} \frac{\sqrt{2}}{e^A\sqrt{n}}. \end{aligned}$$

By the p-series test, we can see that this series will diverge, and therefore by the partial converse to the Borel-Cantelli lemma, we can deduce that

$$\mathbb{P}[S_{2n} = 0 \text{ i.o.}] = 1.$$

Corollary 4.5. *A random walk visits each point infinitely many times.*

Proof. Consider some $N \in \mathbb{N}$, and let A be the event that you start at 0 and go in the positive direction N moves in a row. We can easilt verify that

$$\mathbb{P}[A_n] = \binom{N}{N}\left(\frac{1}{2}\right)^N = \left(\frac{1}{2}\right)^N.$$

Now let $n_k = 1, 2, \dots$ be a subsequence of indices for which $S_{n_k} = 0$, and $A_1, A_2, \dots$ be the sequence of events where a random walk goes right N times after each trip to the origin. Note that

$$\sum_{n_k=1}^{\infty} \mathbb{P}[A_{n_k}] = \sum_{n_k=1}^{\infty} \left(\frac{1}{2}\right)^N = \left(\frac{1}{2}\right)^N \sum_{n_k=1}^{\infty} 1.$$

This sum clearly diverges, thus, by the partial converse of the Borel-Cantelli lemma, we can conclude that each point is visited infinitely many times in a random walk with probability one. □

Another thing one could wonder about a random walk in one dimension is about the expected distance from the origin at any given step. To do this, we will consider $\mathbb{E}[S_n^2]$.

$$\begin{aligned}\mathbb{E}[S_n^2] &= \mathbb{E}\left[\left(\sum_{j=1}^{n} X_j\right)\left(\sum_{k=1}^{n} X_k\right)\right] \\ &= \sum_{k=1}^{n} \mathbb{E}\left[X_k \sum_{j=1}^{n} X_j\right].\end{aligned}$$

First, let us consider $\mathbb{E}[X_j X_k]$ for $j \neq k$. By the definition of a random walk, X_j and X_k are independent, meaning if $j \neq k$, then $\mathbb{E}[X_j X_k] = \mathbb{E}[X_j]\mathbb{E}[X_k] = 0$ since the X_i's are mean 0. Thus, we see that

$$\mathbb{E}[S_n^2] = \mathbb{E}\left[\sum_{k=1}^{n} X_k^2\right] = \sum_{k=1}^{n} \mathbb{E}[X_k^2].$$

We can compute that $\mathbb{E}[X_k^2] = 1$, meaning

$$\mathbb{E}[S_n^2] = \sum_{k=1}^{n} 1 = n.$$

Finally, although this does not show that $\mathbb{E}[|S_n|] = O\sqrt{n}$, it does help give intuition behind why that is true. It is known that $\mathbb{E}[|S_n|]$ tends to $\sqrt{\frac{2n}{\pi}}$ [10], but a proof of this fact will not be presented in this paper.

One final aspect about random walks in one dimension that we will touch on is the behavior of a random walk within a certain interval. We will first present a couple of useful lemmas.

Lemma 4.6. *Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables each with mean μ, let $S_n = \sum_{k=1}^{n} X_k$ and $S_0 = 0$ as before, and let T be a stopping time. Then,*

$$\mathbb{E}[S_T] = \mu\mathbb{E}[T].$$

Lemma 4.7. *As before, let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables each with mean 0 and variance σ^2, $S_n = \sum_{k=1}^{n} X_k$, $S_0 = 0$, and T be a stopping time. Then,*

$$\mathbb{E}[S_T^2] = \sigma^2\mathbb{E}[T].$$

Proofs of these lemmas can be found in [7]. Now that these lemmas are out of the way, we proceed with a proposition about how long a random walk will stay in an interval.

Proposition 4.8. *Consider a random walk that starts at 0 and an open interval (a, b) such that $a, b \in \mathbb{Z}$ and $a < 0 < b$. If $T = \inf\{n | S_n \notin (a, b)\}$ is a stopping time, then $\mathbb{E}[T]$ is finite a.s. and, in particular,*

$$\mathbb{E}[T] = -ab.$$

Proof. We first strive to show that $\mathbb{E}[T] < \infty$. To do this, consider the fact that at any time when we are in the interval, if the next $b - a$ moves are to the right, then we will be outside (a, b). The probability of this not happening is $1 - \frac{1}{2^{(b-a)}}$. Note that this is not necessarily the only way out of the interval (suppose $a = -1, b = 1$, clearly it takes 1 step to escape even though $(b - a) = 2$), meaning

$$\mathbb{P}[T > (b - a)] \leq 1 - \frac{1}{2^{(b-a)}}$$

Inductively, assume that

$$\mathbb{P}[T > n\,(b - a)] \leq \left(1 - \frac{1}{2^{(b-a)}}\right)^n.$$

Using the same logic as before, if we are in (a, b) at time $n\,(b - a)$, if the next $b - a$ steps were all to the right we would be out of (a, b). However, this is not necessarily the only way to escape (a, b). Thus,

$$\mathbb{P}[T > (n + 1)\,(b - a)] \leq \left(1 - \frac{1}{2^{(b-a)}}\right)^{n+1}.$$

Therefore, we have inductively shown that $\mathbb{P}[T > n\,(b-a)] \leq \left(1 - \frac{1}{2^{(b-a)}}\right)^n$ for all n. Now, we have the following expression for $\mathbb{E}[T]$:

$$\begin{aligned}
\mathbb{E}[T] &= \sum_{k\geq 0} \mathbb{P}[T > k] \\
&\leq \sum_{k\geq 0} \mathbb{P}\left[T > (b-a)\left\lfloor \frac{k}{(b-a)} \right\rfloor\right] \\
&\leq \sum_{k\geq 0} \left(1 - \frac{1}{2^{(b-a)}}\right)^{\left\lfloor \frac{k}{(b-a)} \right\rfloor} \\
&\leq \sum_{k\geq 0} (b-a)\left(1 - \frac{1}{2^{(b-a)}}\right)^k \\
&= (b-a)\sum_{k\geq 0} \left(\frac{2^{(b-a)}-1}{2^{(b-a)}}\right)^k.
\end{aligned}$$

This is a geometric series whose common ratio is less than one, therefore, we can see that the series converges, and in particular, we have an upper bound for $\mathbb{E}[T]$:

$$\mathbb{E}[T] \leq \frac{b-a}{1 - \frac{2^{(b-a)}-1}{2^{(b-a)}}} < \infty.$$

By the way our stopping time is defined, we know that at the stopping time, $S_T = a$ or $S_T = b$. Hence,

$$\mathbb{E}[S_T] = a\mathbb{P}[S_T = a] + b\mathbb{P}[S_T = b]$$

From Lemma 4.6, we know that $\mathbb{E}[S_T] = \mu\mathbb{E}[T]$, which is just 0 because $\mu = 0$ in a random walk. This information lets us say that

$$0 = a\mathbb{P}[S_T = a] + b\mathbb{P}[S_T = b]$$

equivalently,

$$\mathbb{P}[S_T = a] = \frac{b}{b-a} \text{ and } \mathbb{P}[S_T = b] = \frac{-a}{b-a}. \tag{4.9}$$

Next, consider $\mathbb{E}[S_T^2]$.

$$\mathbb{E}[S_T^2] = a^2\mathbb{P}[S_T = a] + b^2\mathbb{P}[S_T = b].$$

Furthermore, Lemma 4.7 tells us that $\mathbb{E}[S_T^2] = \sigma^2\mathbb{E}[T]$, and in this case $\sigma = 1$, meaning

$$\mathbb{E}[S_T^2] = \mathbb{E}[T].$$

Therefore, $\mathbb{E}[T] = a^2\mathbb{P}[S_T = a] + b^2\mathbb{P}[S_T = b]$. Plugging in our expressions from (4.9) for $\mathbb{P}[S_T = a]$ and $\mathbb{P}[S_T = b]$, we get

$$\mathbb{E}[T] = a^2\frac{b}{b-a} + b^2\frac{a}{b-a} = -ab.$$

Hence, $\mathbb{E}[T] = -ab$, as desired. □

Now that we know the expected amount of time to get out of (a, b), a natural follow up question to consider is which way are you likely to escape. The answer to this problem is analogous to the solution to Gambler's ruin, a famous result in probability. Intuitively, one can easily deduce that whichever endpoint you are closer to is more likely to be hit first; however, we may want to know the exact probability associated with this.

Let $F(x)$ be the function that gives the probability of hitting b before hitting a assuming the random walk starts at x. Clearly, for all $x \leq a, F(x) = 0$ and for all $x \geq b, F(x) = 1$. Given an $x \in (a, b)$, we know that we are equally likely to end up at $x + 1$ and $x - 1$ on the next turn, therefore,

$$F(x) = \frac{1}{2}\left(F(x+1) + F(x-1)\right).$$

This information about F motivates the following theorem about what F will look like on (a, b); however, before solving for F, we need to define the discrete Laplacian and prove the maximum principle.

Definition 4.10. A discrete Laplacian of a function, f, in d-dimensions is defined as

$$\Delta\left(f(x)\right) = \left[\frac{1}{2d}\sum_{|y-x|=1} f(y)\right] - f(x).$$

We say a function is harmonic across a set, S if its Laplacian is 0 on S.
Note that this means that a function is harmonic if and only if

$$f(x) = \frac{1}{2d}\sum_{|y-x|=1} f(y).$$

In one dimension, this means

$$f(x) = \frac{1}{2}\left(f(x+1) + f(x-1)\right).$$

Proposition 4.11 (The Maximum Principle)**.** *A harmonic function attains its maximum value along the boundary of the set on which it is harmonic.*

Proof. If f is constant, then the conclusion immediately follows since f obtains its maximum everywhere.

Now, assume f is non-constant. Take f to be a harmonic function on S and assume for the sake of contradiction that f attains its maximum at some point x on the interior of S, that is, not on $\bar{S}\backslash S$.

Therefore, $f(x) \geq f(y)$ for all $y \in S$. Start at some x where this maximum is attained, and consider its nearest neighbors. If all of its neighbors attain the same value $f(x)$, then find a different point where this maximum is attained. Repeat this process until you find an x' that has a neighbor, z, such that $f(x') > f(z)$. The existence of such an x' is guaranteed by the fact that f is non-constant. Note that this will pose a contradiction since if f is harmonic, $f(x)$ is an average of its nearest neighbors. However, since there is some $z \in S$ such that $f(x) > f(z)$, we know one of xs nearest neighbors will have a function value greater than $f(x)$, and, thus, we get a contradiction. $\square$

Theorem 4.12. *Consider a function* $F : \{a, a+1, \ldots, 0, \ldots, b-1, b\} \mapsto [0, 1]$ *such that* $F(x) = P[S_T = b|S_0 = x]$*, where* T *is a stopping time such that*

$$T = \min\{n \mid S_n = a \text{ or } S_n = b\}.$$

The only solution satisfying such properties is a linear function of the form

$$F(x) = \frac{x-a}{b-a}.$$

Proof. Let S_n be a random walk starting at x, and let $T = \min\{N : S_N = a \text{ or } S_N = b\}$.

By the optional stopping theorem, we know that $\mathbb{E}[S_0] = \mathbb{E}[S_T]$. Therefore, we see that

$$x = \mathbb{E}[S_0] = bF(x) + a(1 - F(x)).$$

Rearranging terms, we get

$$F(x) = \frac{x-a}{b-a}.$$

Now, we want to show uniqueness. Suppose that F_1 and F_2 are two functions with the properties above. Note that it follows that F_1 and F_2 are harmonic since this follows from the conditions above. Therefore, the function $F_1 - F_2$ is harmonic as well. Since $F_1(a) - F_2(a) = 0$ and $F_1(b) - F_2(b) = 1 - 1 = 0$. Therefore, $F_1 - F_2 \leq 0$ on (a, b) due to Proposition 4.11.

Now, consider $F_2 - F_1$. Again, $F_2 - F_1$ is harmonic. Since $F_2(a) - F_1(a) = 0$ and $F_2(b) - F_1(b) = 1 - 1 = 0$, we know that $F_2 - F_1 \leq 0$ on (a, b) too. Since $F_2 - F_1 = -(F_1 - F_2)$ and both $F_2 - F_1$ and $F_1 - F_2$ are bounded above by 0, we know that $F_2 - F_1 = F_1 - F_2 = 0$. Therefore, $F_1 = F_2$, which indicates that F is unique. □

5. Proving the Kolmogorov 0-1 Law and Radon-Nikodym Theorem

Before getting into the proof of the Kolmogorov 0-1 Law, we must first establish a few things about uniform integrability. We begin by explaining some notation and presenting the definition of uniform integrability and a few properties before defining a uniformly integrable martingale.

Notation 5.1. For $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and $F \in \mathcal{F}$, we write $\mathbb{E}[X; F] = \mathbb{E}[X\mathbb{1}_F]$.

Definition 5.2. A class $\mathcal{C}$ of random variables is uniformly integrable if for all $\epsilon > 0$, there is a $K \in \mathbb{R}^+$ such that

$$\mathbb{E}[|X|; |X| > K] < \epsilon \text{ for all } X \in \mathcal{C}.$$

Proposition 5.3. *A class of uniformly integrable random variables is in* $\mathcal{L}^1$*.*

Proof. Let $\mathcal{C}$ be a class of uniform random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and take some $X \in \mathcal{C}$. First note that

$$\mathbb{E}[|X|] = \mathbb{E}[|X|; |X| > K] + \mathbb{E}[|X|; |X| \leq K].$$

Because X is in $\mathcal{C}$, given an $\epsilon > 0$, we can choose K_ϵ such that $\mathbb{E}[|X|; |X| > K] < \epsilon$. Fixing ϵ and its corresponding K_ϵ, we get

$$\mathbb{E}[|X|] = \mathbb{E}[|X|; |X| > K_\epsilon] + \mathbb{E}[|X|; |X| \leq K_\epsilon] \leq \epsilon + K_\epsilon.$$

Therefore, we see that $X \in \mathcal{L}^1$. Since $X \in \mathcal{C}$ was chosen arbitrarily, this holds for all $X \in \mathcal{C}$. □

Lemma 5.4. *If $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, then for any $\epsilon > 0$, there is a $\delta > 0$ such that for any $F \in \mathcal{F}$, $\mathbb{P}[F] < \delta$ implies that $\mathbb{E}[|X|; F] < \epsilon$.*

We will not present a proof of this lemma; however, a proof can be found in [1, p. 126].

Theorem 5.5. *If $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, then*

$$\{\mathbb{E}[X \mid \mathcal{G}] \ : \ \mathcal{G} \text{ is a sub sigma-algebra of } \mathcal{F}\}$$

is a uniformly integrable family.

Proof. First, take some $\epsilon > 0$. We know that we can take $\delta > 0$ such that $\mathbb{P}[F] < \delta$ implies that $\mathbb{E}[|X|; F] < \epsilon$.

Also, since $X \in \mathcal{L}^1$, we know that there is a $K \in \mathbb{R}^+$ such that $\mathbb{E}[|X|] < K\delta$. Now, taking $\mathcal{G}$ to be a sub sigma-algebra of $\mathcal{F}$ and Y to be $\mathbb{E}[X \mid \mathcal{G}]$, we can apply Jensen's inequality (Proposition 2.10 g) to get

$$|Y| \leq \mathbb{E}[|X| \mid \mathcal{G}]$$

a.s. Therefore, we also have $\mathbb{E}[|Y|] \leq \mathbb{E}[|X|]$, and, by Markov's inequality,

$$K\mathbb{P}[|Y| > K] \leq \mathbb{E}[|Y|] \leq \mathbb{E}[|X|] < K\delta.$$

Hence,

$$\mathbb{P}[|Y| > K] < \delta, \text{ and therefore } \mathbb{E}[|Y|; |Y| > K] < \mathbb{E}[|X|; |Y| > K] < \epsilon.$$

□

Definition 5.6. Let $\{M_n\}$ be a martingale relative to $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$. We say $\{M_n\}$ is a uniformly integrable martingale (UI martingale) if $\{M_n\}$ is a uniformly integrable class.

Theorem 5.7. *Let $\{M_n\}$ be a UI Martingale relative to $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, \mathbb{P})$. Then, $M_\infty = \lim_{n\to\infty} M_n$ exists almost surely in the $\mathcal{L}^1$ sense. Furthermore, $M_n = \mathbb{E}[M_\infty \mid \mathcal{F}_n]$ for all n.*

This theorem is the extension of the martingale convergence theorem for UI martingales, and the proof for it is similar to the one presented for Theorem 3.20 and it can be found in [1, pp. 133-134]. With this out of the way, we move on to proving Levy's upward theorem, which will be used in our proof of the Kolmogorov 0-1 law.

Theorem 5.8 (Levy's Upward Theorem)**.** *Let $\xi \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, and define $M_n = \mathbb{E}[\xi \mid \mathcal{F}_n]$ a.s. Define $\eta = \mathbb{E}[\xi \mid \mathcal{F}_\infty]$. Then, $\{M_n\}$ is a UI martingale and $\{M_n\}$ converges almost surely to η in the $\mathcal{L}^1$ sense.*

Proof. First, note that $\mathbb{E}[M_{n+1}|F_n] = M_n$ by Proposition 2.10 e, meaning $\{M_n\}$ is a martingale. Furthermore, since $\xi \in \mathcal{L}^1$, Theorem 5.5 tells us that $\{M_n\}$ is uniformly integrable. By Theorem 5.7, we know that $\lim_{n\to\infty} M_n = M_\infty$ exists in the almost sure sense and in the $\mathcal{L}^1$ sense. It remains to show that $M_\infty = \eta$ a.s.

Without loss of generality, we assume that $\xi \geq 0$. Take two measures on $(\Omega, \mathcal{F}_\infty)$, that are defined such that

$$\mu_1(F) = \mathbb{E}[\eta; F] \text{ and } \mu_2(F) = \mathbb{E}[M_\infty; F] \text{ where } F \in \mathcal{F}_\infty.$$

Now, note that for every $F \in \mathcal{F}_n$, $\mathbb{E}[\eta; F] = \mathbb{E}[\xi; F]$ by the tower property. Thus,

$$\mathbb{E}[\eta; F] = \mathbb{E}[M_n; F] = \mathbb{E}[M_\infty; F],$$

where the second equality comes from [1, pp. 133-134]. Hence we see that μ_1 and μ_2 agree on $\cup_{n=1}^{\infty}\mathcal{F}_n$, meaning they agree on $\mathcal{F}_\infty$.

Note that both η and M_∞ are both $\mathcal{F}_\infty$-measurable. One can see M_∞ is $\mathcal{F}_\infty$-measurable by taking $M_\infty = \limsup_{n\to\infty} M_n$ almost surely. Therefore, we see that $\{\omega \mid \eta > M_\infty\} \in \mathcal{F}_\infty$, and because $\mu_1(F) = \mu_2(F)$, we see that $\mathbb{E}[\eta - M_\infty; \eta > M_\infty] = 0$. Thus, $\mathbb{P}[\eta > M_\infty] = 0$. One can show in an identical manner that $\mathbb{P}[\eta < M_\infty] = 0$.

Therefore, $M_\infty = \eta$ a.s. □

Before, proceeding to the proof of the Kolmogorov 0-1 law, we must define a tail event.

Definition 5.9. Let $\mathcal{F}_1, \mathcal{F}_2, \ldots$ be a sequence of sigma-algebras on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The tail sigma-algebra is defined as

$$\tau = \bigcap_{n=1}^{\infty} \sigma\left(\bigcup_{i=n}^{\infty} \mathcal{F}_i\right).$$

An event, T, is said to be a tail event if $T \in \tau$.

Theorem 5.10 (Kolmogorov 0-1 Law)**.** *Given a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a sequence of independent events* $F_1, F_2, \cdots \in \mathcal{F}$ *and a tail* τ*, for all* $T \in \tau$*,* $\mathbb{P}[T] = 0$ *or* $\mathbb{P}[T] = 1$*.*

Proof. Define $\mathcal{F}_n = \sigma(F_1, F_2, \ldots, F_n)$. Let T be a tail event, and let $\eta = \mathbb{1}_T$. By Levy's upward theorem, we know that

$$\eta = \mathbb{E}[\eta \mid \mathcal{F}_\infty] = \lim_{n\to\infty} \mathbb{E}[\eta \mid \mathcal{F}_n] \text{ a.s.} \tag{5.11}$$

Now, note that for all n, η is independent of $\mathcal{F}_n$, thus, $\mathbb{E}[\eta|\mathcal{F}_n] = \mathbb{E}[\eta]$, which, from (5.11) implies $\eta = \mathbb{E}[\eta] = \mathbb{P}[T]$. Furthermore, we know that since $\eta := \mathbb{1}_T$, we know it takes on values of 0 or 1, meaning $\mathbb{P}[T] = 0$ or $\mathbb{P}[T] = 1$, which is the desired result. □

Now that we have proved the Kolmogorov 0-1 law, we will apply it to answer our question from Section 2 about a sequence of 1s given a sequence of Bernoulli-p random variables. Recall the result.

Proposition 5.12. *Given an infinite sequence of Bernoulli-p random variables where* $p \in (0, 1)$*, arbitrarily long sequences of 1s will occur infinitely many times.*

Proof. Let $\{X_n\}$ represent our sequence of Bernoulli-p random variables and A_m be the event that n consecutive 1s occur starting at the $(nm+1)$th trial. That is, $A_m = \{X_{nm+1} = 1\} \cap \{X_{nm+2} = 1\} \cap \cdots \cap \{X_{nm+n} = 1\}$. Let A be the event that n consecutive 1s occur in $\{X_n\}$. Now consider a sequence $A_1, A_2, \ldots$, each containing n disjoint trials. For each k, let $B_k = \bigcup_{m=k}^{\infty} A_m$ and let $B = \bigcap_{k=1}^{\infty} B_k$. Note that this is just $\lim_{k\to\infty} B_k$ since $\{B_k\}_{k=1}^{\infty}$ is a decreasing sequence of sets.
Note that B is a tail event, meaning Kolmogorov's 0-1 law implies that $\mathbb{P}[B] = 0$

or $\mathbb{P}[B] = 1$. Consider $\mathbb{P}[B]$:

$$\begin{aligned}\mathbb{P}[B] &= \mathbb{P}[\lim_{k\to\infty} B_k] \\ &= \lim_{k\to\infty} \mathbb{P}[B_k] \\ &\geq \lim_{k\to\infty} \mathbb{P}[A_k] \\ &= \lim_{k\to\infty} p^n \\ &= p^n \\ &> 0.\end{aligned}$$

Note that since $\mathbb{P}[B] > 0$ and $\mathbb{P}[B] = 0$ or $\mathbb{P}[B] = 1$, we can deduce that $\mathbb{P}[B] = 1$. Hence, we can conclude that a sequence of n 1s will occur with probability 1. Since n was chosen arbitrarily, this holds for any $n \in \mathbb{Z}^+$. □

Now that we have proved and applied the Kolmogorov 0-1 law, we turn to another application of martingales: a proof of a slightly weaker version of the Radon-Nikodym theorem, but first, we must define an atom.

Definition 5.13. Let $(X, \mathcal{F})$ be a measurable space. $E \in \mathcal{F}$ is said to be an atom if its only proper subset in $\mathcal{F}$ is the emptyset.

In our weaker version of the Radon-Nikodym theorem, we assume that the sigma-algebra in the probability space is separable. This is not always the case. One can, however, generalize the proof we present and prove the Radon-Nikodym theorem without this extra restriction. The proof of this can be found in [1, pp. 148-149].

Theorem 5.14 (Radon-Nikodym Theorem Assuming Separability)**.** *Take* $(\Omega, \mathcal{F}, \mathbb{P})$ *to be a probability space where* $\mathcal{F}$ *is separable in that*

$$\mathcal{F} = \sigma(F_n) \text{ for some } \{F_n : n \in \mathbb{N}\}$$

where F_n *is a subsequence of subsets of* Ω*. Suppose* Q *is a finite measure on* Ω *where*

$$\mathbb{P}[F] = 0 \text{ implies that } Q[F] = 0.$$

Then there is some $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ *such that*

$$Q[F] = \int_F X d\mathbb{P} = \mathbb{E}[X; F] \text{ for all } F \in \mathcal{F}.$$

We can also write

$$\frac{dQ}{d\mathbb{P}} = X \text{ a.s. on } \mathcal{F}.$$

Proof. Recall from earlier that if $X \in \mathcal{L}^1$, then for any $\epsilon > 0$, there is a $\delta > 0$ such that for any $F \in \mathcal{F}$, $\mathbb{P}[F] < \delta$ implies that $\mathbb{E}[X; F] < \epsilon$. From this result, we can show that for all $\epsilon > 0$, there exists δ such that $\mathbb{P}[F] < \delta$ implies that $Q[F] < \epsilon$. Now, define $\mathcal{F}_n = \sigma(F_1, \dots, F_n)$. Now, for each n, let

$$A_{n,1}, \dots, A_{n,k}$$

denote the atoms of $\mathcal{F}_n$.

Now, define a function $X_n : \Omega \to [0, \infty)$ such that if $\omega \in A_{n,j}$,

$$X_n(\omega) = \begin{cases} 0 & \mathbb{P}[A_{n,j}] = 0 \\ \frac{\mathbb{Q}[A_{n,j}]}{\mathbb{P}[A_{n,j}]} & \text{otherwise} \end{cases}.$$

Note that $X_n \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and that

$$\mathbb{E}[X_n; F] = Q[F] \text{ for all } F \in \mathcal{F}_n.$$

From this, we can deduce that $\{X_n\}$ is a martingale relative to the filtration $\mathcal{F}_n$, and since X is non-negative, we know that

$$\lim_{n\to\infty} X_n = X_\infty \text{ exists a.s.}$$

In order to deduce that $X_n \to X_\infty$ a.s. in $\mathcal{L}^1$, however, we must show that $\{X_n\}$ is UI.

Take some $\epsilon > 0$. Now, chose $\delta > 0$ and $K \in \mathbb{R}$ such that

$$K^{-1}Q[\Omega] < \delta.$$

Therefore,

$$\mathbb{P}[X_n > K] \leq K^{-1}\mathbb{E}[X_n] = K^{-1}Q[\Omega] < \delta.$$

Accordingly, this means that

$$\mathbb{E}[X_n; X_n > K] = Q[X_n > K] < \epsilon.$$

Hence, $X_n \to X_\infty$ a.s. in $\mathcal{L}^1$.

We see that this implies that $F \mapsto \mathbb{E}[X; F]$ and $F \mapsto Q[F]$ agree on $\bigcup_n \mathcal{F}_n$, meaning they agree on $\mathcal{F}$.

Now, it remains to show that such an X is unique. Consider f and g to be functions that satisfy such a property. Therefore,

$$Q[F] = \int_F f d\mathbb{P} = \int_F g d\mathbb{P}.$$

Then, $f - g$ is $\mathbb{P}$-integrable and

$$\int_F (f - g)\, d\mathbb{P} = 0.$$

In particular, we have that if $F_1 = \{\omega \in \Omega : f(\omega) > g(\omega)\}$ and $F_2 = \{\omega \in \Omega : f(\omega) < g(\omega)\}$, we have

$$\int_{F_1} (f - g)\, d\mathbb{P} = 0 \text{ and } \int_{F_2} (f - g)\, d\mathbb{P} = 0.$$

This holds if and only if

$$\int_\Omega (f - g)^+ \, d\mathbb{P} = 0 \text{ and } \int_\Omega (f - g)^- \, d\mathbb{P} = 0.$$

Therefore, we have $(f - g)^+ = 0$ a.e. and $(f - g)^- = 0$ a.e. This means $f = g$ a.e., thus completing our proof of uniqueness. $\square$

6. Brownian Motion and Proving the Fundamental Theorem of Algebra

Before defining Brownian motion and proving the fundamental theorem of algebra, we must give some a few definitions and some theorems that are analogous to those proved in Section 3.

Definition 6.1. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\{\mathcal{F}_t\}_{t\in\mathbb{T}}$, where $\mathbb{T}$ is an interval (note that unlike before, our indexing set, $\mathbb{T}$, is continuous), a family of random variables $\{X_t\}_{t\in\mathbb{T}}$ indexed by $\mathbb{T}$ is a stochastic process. A process is called adapted if each random variable X_t is measurable with respect to the corresponding sigma-algebra $\mathcal{F}_t$.

Definition 6.2. Let $\{X_t\}_{t\in\mathbb{T}}$ be an adapted stochastic process with respect to a filtration $\{\mathcal{F}_t\}_{t\in\mathbb{T}}$. The process $\{X_t\}_{t\in\mathbb{T}}$ is said to be progressively measurable if for every $[r, s] \in \mathbb{T}$, the function $\{X_t\}_{t\in[r,s]}$ is (when considered as a function $X(t, \omega)$ on $[r, s] \times \Omega$) measurable with respect to the sigma-algebra $\mathcal{B}([r, s]) \times \mathcal{F}_s$, where $\mathcal{B}([r, s])$ is the Borel sigma-algebra generated by $[r, s]$.

Definition 6.3. Given a probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{T}}, \mathbb{P}\right)$, a process $\{M_t\}_{t\in\mathbb{T}}$ is a martingale if for every $s < t$,

$$M_s = \mathbb{E}[M_t \mid \mathcal{F}_s].$$

Definition 6.4. Let $\{X_t\}_{t\in\mathbb{T}}$ be an progressively measurable stochastic with respect to a filtration $\{\mathcal{F}_t\}_{t\in\mathbb{T}}$ on a probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{T}}, \mathbb{P}\right)$. For a fixed $\omega \in \Omega$, the function $t \mapsto X_t(\omega)$ is called a sample path corresponding to the fixed $\omega \in \Omega$.

Theorem 6.5 (Doob's Optional Stopping Theorem in Continuous Time)**.** *If $\{M_t\}$ is a continuous martingale with respect to a filtration, $\mathcal{F}$, with sample paths that are right-continuous with left limits, then $\{M_t\}$ is progressively measurable and for any finite stopping time, T, the function M_T is measurable. In particular, if T is a bounded stopping time, then M_T is integrable and*

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

Theorem 6.6 (Martingale Convergence Theorem in Continuous Time)**.** *Suppose a continuous martingale $\{M_t\}$ satisfies $\sup_t \mathbb{E}[|M_t|^p] < \infty$ for some $p \geq 1$. Then there is some random variable M_∞ such that*

$$\lim_{t\to\infty} M_t = M_\infty \ a.s.$$

If $p > 1$, then convergence also holds in $\mathcal{L}^p$.

Proofs of these theorems can be found in [5] and [4] respectively. Now, after spending Section 4 discussing random walks, we now turn our attention to Brownian motion (an example of a continuous martingale), the continuous analog of a random walk. We will start by giving a definition for a Brownian motion.

Definition 6.7. A real-valued stochastic process $\{B(t) : t \geq 0\}$ is a standard Brownian Motion if the following conditions are met:

(i) $B(0) = 0$.
(ii) The process has independent increments. This means that for all $t_1, t_2, \ldots, t_n$, the increments $B(t_n) - B(t_{n-1})$, $B(t_{n-1}) - B(t_{n-2})$, $\ldots$, $B(t_2) - B(t_1)$ are pairwise independent random variables.

(iii) Given any $t \geq 0$ and $h > 0$, the increment $B(t+h) - B(t)$ is normally distributed with an expected value of 0 and a variance of h.
(iv) The function $f : t \mapsto B(t)$ is a.s. continuous.

Existence of Brownian Motion is non-trivial, so before trying to prove properties of Brownian motion, one must first verify that it exists. In this paper, we will outline a construction of a function that meets these criteria; however, if the reader would like to see the full construction, they can refer to [2, pp. 23-25].

To begin our outline of a construction, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Define $\mathcal{D}_n = \left\{\frac{k}{2^n}, 0 \leq k \leq 2^n\right\}$, and consider the set of all dyadic points on the interval $[0,1]$. That is,

$$\mathcal{D} = \bigcup_{n=0}^{\infty} \left\{\frac{k}{2^n}, 0 \leq k \leq 2^n\right\}.$$

Let $\{Z_t : t \in \mathcal{D}\}$ be a set of independent, standard normally distributed random variables. Define $B(0) = 0$, $B(1) = Z_1$. We now will define $B(d)$ for every $d \in \mathcal{D}$. Inductively, define $B(d)$ where

$$B(d) = \frac{B(d - 2^{-n}) + B(d + 2^{-n})}{2} + \frac{Z_d}{2^{\frac{n+1}{2}}}.$$

Furthermore, define B to be linear between each pair of adjacent $d \in \mathcal{D}_n$ for all n. One can show that this construction of B satisfies the four properties necessary for it to be a standard Brownian motion, thus proving its existence.

Although there is a lot to prove about Brownian motion in one dimension, such as the fact that Brownian motion is nowhere right differentiable a.s., we will not dive into these facts in this paper. Instead, we will move onto discussing Brownian motion in higher dimensions. We will now present a definition of Brownian motion in d-dimensions.

Definition 6.8. A Brownian motion in $\mathbb{R}^d$ is a d-dimensional vector whose components are independent scalar Brownian motions. Note that for the case $d = 2$ this is called planar Brownian motion.

We will now proceed with a propositions that will be used to prove a lemma. The proof for this proposition will not be presented in this paper; however, it can be found in [9].

Proposition 6.9. *Let $f : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ be continuously differentiable in the first coordinate and twice continuously differentiable in the second coordinate. Furthermore, suppose there is a K such that*

$$|f(t,x)| + |\frac{\partial f}{\partial t}| + \sum_{i=1}^{d} |\frac{\partial f}{\partial x_i}(t,x)| + \sum_{i,j=1}^{d} |\frac{\partial^2 f}{\partial x_i \partial x_j}(t,x)| \leq K e^{K(t+|x|)} \tag{6.10}$$

for all $(t,x) \in [0,\infty) \times \mathbb{R}^d$. If f is harmonic, then $f(t, B_t)$ is a martingale.

Remark 6.11. We say a function g is harmonic on a region A if $\Delta g = 0$ on A, where Δ denotes a continuous version of the laplacian, which is defined as

$$\Delta f(x_1, \ldots, x_k) = \sum_{k=1}^{d} \frac{\partial^2}{\partial x_k^2} f(x_1, \ldots, x_k).$$

Lemma 6.12. *Planar Brownian motion is neighborhood recurrent, but it does not visit a specific point. In particular, if B is a Brownian motion that starts at $z_0 \in \mathbb{C}$, then given some neighborhood of z_0, $N(z_0)$, we have*

$$\{t : B_t \in N(z_0)\}$$

is unbounded a.s., but

$$\{t : B_t = z_0\}$$

is empty.

Proof. Without loss of generality, let $z_0 \neq 0$. We will show that with probability 1, B visits any neighborhood of 0 infinitely often, but it will never hit 0. Let $0 < a < b$ and multiply the function $z \mapsto \log|z|$ by a smooth function that equals 1 for $\{z : |a| \leq |z| \leq b\}$ and is 0 on $\left\{z : \frac{a}{2} \geq |z|, |z| \geq 2b\right\}$. This yields a function, f, that satisfies (6.10), thus, $f(B_t)$ is a martingale by Proposition 6.9. Furthermore,

$$\Delta f = \frac{\partial^2}{\partial x^2} f + \frac{\partial^2}{\partial y^2} f = \frac{y^2 - x^2}{x^2 + y^2} + \frac{x^2 - y^2}{x^2 + y^2} = 0$$

on $\{z : a < |z| < b\}$. Applying the optional stopping theorem to the uniformly integrable martingale $f(B_t)$ with the stopping time $T = \inf\{t : |B_t| \in \{a, b\}\}$ and letting p be the probability B_t hits a before b, we get

$$\log|z_0| = p \log a + (1 - p) \log b.$$

Rearranging terms, we get

$$p = \frac{\log b - \log z_0}{\log b - \log a}.$$

Letting $b \to \infty$ lets us conclude $\{t : B_t \in N(z_0)\}$ is unbounded a.s. Letting $a \to 0$ lets us conclude that $\{t : B_t = z_0\}$ is empty. Therefore, the desired conclusion is reached. □

We are now ready to proceed to the proof of the fundamental theorem of algebra. Recall the statement of the theorem.

Theorem 6.13 (Fundamental Theorem of Algebra)**.** *If p is a non-constant polynomial, then there is some $z \in \mathbb{C}$ for which $p(z) = 0$.*

Proof. Assume for the sake of contradiction that $p(z) \neq 0$ for all $z \in \mathbb{C}$. Then $f = \frac{1}{p}$ is an analytic function on $\mathbb{C}$. Since p is a polynomial, we know that $p \to \infty$ as $z \to \infty$, meaning f is bounded. Let B_t be a Brownian Motion starting at the origin. We know that $\mathrm{Re} f(B_t)$ is a martingale since $\mathrm{Re} f$ is harmonic when f is analytic. Applying the martingale convergence theorem since $\mathrm{Re} f$ is bounded, we get that $\mathrm{Re} f(B_t)$ exists a.s. as $t \to \infty$.

Conversely, $\mathrm{Re} f(\mathbb{C})$ contains more than one element, so we can choose α, β such that

$$\inf \mathrm{Re} f(\mathbb{C}) < \alpha < \beta < \sup \mathrm{Re} f(\mathbb{C}).$$

Now, consider $\{z : \mathrm{Re} f(z) < \alpha\}$ and $\{z : \mathrm{Re} f(z) > \beta\}$, which are non-empty, disjoint open sets. From Lemma 6.12, we know that a Brownian motion visits each neighborhood infinitely many times. Thus,

$$\liminf_{t \to \infty} \mathrm{Re} f(B_t) < \alpha < \beta < \limsup_{t \to \infty} \mathrm{Re} f(B_t).$$

However, this contradicts the convergence of $\mathrm{Re} f(B_t)$. Therefore, if p is a non-constant polynomial, then there is some $z \in \mathbb{C}$ for which $p(z) = 0$. □

Acknowledgments

First and foremost, I thank Peter May for organizing the REU. I also want to thank Daniil Rudenko, Elizaveta Shuvaeva, and Greg Lawler for insightful lectures on many topics ranging from quadratic reciprocity to the Dirichlet problem. Finally, I would also like to extend a big thanks to my mentor, Ethan Naegele. I thoroughly enjoyed our meetings, and this paper would not have been possible without his guidance.

References

[1] David Williams. Probability With Martingales. Cambridge University Press 1991.
[2] Peter Mörters and Yuval Peres. Brownian Motion. https://www.stat.berkeley.edu/ aldous/205B/bmbook.pdf
[3] Steven Lalley. Conditional Expectation. https://galton.uchicago.edu/ lalley/Courses/383/ConditionalExpectation.pdf
[4] Don McLeish. Probability Notes. https://sas.uwaterloo.ca/ dlmcleis/s901/
[5] Steven Lalley. Continuous Martingales I: Fundamentals. https://galton.uchicago.edu/ lalley/Courses/385/ContinuousMG1.pdf
[6] Greg Lawler. Notes on Probability. https://www.math.uchicago.edu/ lawler/probnotes.pdf
[7] Sebastian Roch. Graduate Probability Notes. https://people.math.wisc.edu/ roch/gradprob/gradprob-notes12.pdf
[8] Richard Bass. Real Analysis for Graduate Students. https://bass.math.uconn.edu/v43.pdf
[9] Sam Watson. Brownian Motion, Complex Analysis, and the Dimension of the Brownian Frontier.
[10] Phillip Compeau. Proof of the Random Walk Theorem. https://biologicalmodeling.org/assets/tex/random-walk-theorem.pdf
[11] Rick Durrett. Probability: Theory and Examples. https://services.math.duke.edu/ rtd/PTE/PTE5-011119.pdf
[12] Roger Mansuy The Origins of the Word "Martingale" https://www.jehps.net/juin2009/Mansuy.pdf