MARTINGALES AND A BASIC APPLICATION

TURNER SMITH

Abstract. This paper will develop the measure-theoretic approach to probability in order to present the definition of martingales. From there we will apply this to the Abracadabra problem, which calculates the expected time necessary for a randomly typed sequence of letters to occur.

Contents

1. Introduction

While probability can be studied without utilizing measure theory, taking the measure-theoretic approach to probability provides significantly more generality. In section 2, we begin the paper by constructing general measure spaces, without considering their uses in probability. In section 3, we provide measure-theoretic definitions for familiar probability topics such as random variables and expectation as well as prove several theorems which are necessary for later proofs. Section 4 will provide the definition of martingales and develop martingales in order to prove Doob's Optional-Stopping Theorem which is instrumental in solving the Abracadabra problem. Section 5 will conclude the paper with a solution of exercise 10.6 in Probability with Martingales. This problem concerns the expected time it takes for a monkey to type the letters ABRACADABRA in that order. This paper closely follows David Williams' Probability with Martingales [1] and many of the proofs presented in this paper can be found in his book.

2. Measure Spaces

Definition 2.1. A collection Σ_0 of subsets of S is called an algebra on S if

(1) $S \in \Sigma_0$
(2) $F \in \Sigma_0 \Rightarrow F^c \in \Sigma_0$
(3) $F, G \in \Sigma_0 \Rightarrow F \cup G \in \Sigma_0$

Date: DEADLINE AUGUST 26, 2011.

Corollary 2.2. *By definition,* $\emptyset = S^c \in \Sigma_0$ *and* $F, G \in \Sigma_0 \Rightarrow F \cap G = (F^c \cup G^c)^c \in \Sigma_0$

These imply that an algebra on S is a family of subsets of S stable under finitely many set operations.

Definition 2.3. A collection Σ of subsets of S is called a σ - algebra on S if Σ is an algebra on S such that whenever $F_n \in \Sigma$ $(\forall n \in \mathbb{N})$, then

$$\bigcup_n F_n \in \Sigma.$$

Additionally, this implies that

$$\bigcap_n F_n = (\bigcup_n F_n^c)^c \in \Sigma.$$

Thus, a σ-algebra on S is a family of subsets of S stable under any countable collection of set operations.

A σ-algebra can be thought of as containing all the known information about a given sample space. This will be important in Section 4 in our discussion of martingales. Not all sets are measurable and the σ-algebra is the collection of sets over which a measure is defined. Perhaps the easiest way to think about σ-algebras is to picture a σ-algebra F as corresponding to collections of yes or no questions with every set in F as an outcome in Ω for which there is an answer to the yes or no question. For example, let Ω be possible arrival times for a bus. Then, one set in F could be all the outcomes where the bus arrives after 10:30am today. If one can ask if the bus arrives after 10:30am then it is obvious that the complement can be asked as well, "Did the bus arrive before or at 10:30am today?" Now we can combine multiple questions with "or's" or "and's", which is the equivalent of taking unions and intersections, to get another yes or no question.

Definition 2.4. Let C be a class of subsets of S. Then $\sigma(C)$, the **σ-algebra generated by** C, is the smallest σ-algebra Σ on S such that $C \subseteq \Sigma$. Alternatively, this is the intersection of all σ-algebras on S which have C as a subclass.

Definition 2.5. The **Borel σ-algebra**, denoted by $\mathcal{B}$, is the σ-algebra generated by the family of open subsets in $\mathbb{R}$.

Because the complement of any open set is closed, a Borel set is thus any that can be written with any countable combination of the set operations union and intersection of closed and open sets. Next we need to define a measure. This will require definitions of countably additive and measure spaces.

Definition 2.6. Let S be a set and let Σ_0 be an algebra on S. Then a non-negative set function $\mu_0 : \Sigma_0 \to [0, \infty]$ is called **countably additive** if $\mu_0(\emptyset) = 0$ and whenever $(F_n : n \in \mathbf{N})$ is a sequence of disjoint sets in Σ_0 with union $F = \bigcup F_n \in \Sigma_0$, then

$$\mu_0(F) = \sum_{\mathbf{n}} \mu_{\mathbf{0}}(F_n).$$

Definition 2.7. A pair (S, Σ), where S is a set and Σ is a σ-algebra on S, is a **measurable space**. An element of Σ is called a Σ-measurable subset of S.

Definition 2.8. Let (S, Σ) be a measurable space, so that Σ is a σ-algebra on S. A map

$$\mu : \Sigma \to [0, \infty]$$

is called a **measure** on (S, Σ) if μ is countably additive. The triple (S, Σ, μ) is then called a **measure space**.

Essentially, a measure assigns a real number to all of the subsets of S. This number is most simply viewed as the "volume" of the set. Measures can be chosen arbitrarily, but they must be countably additive. This prevents a measure made of two sets being combined from being smaller than either of the two individual sets' measures.

Measure spaces translate into probability through the probability triple generally denoted $(\Omega, \mathcal{F}, \mathbb{P})$. Of course, these symbols have some additional intuition and/or meaning when relating to probability.

Definition 2.9. The set Ω is called the **sample space** and consists of elements ω which are called **sample points**. Similarly, the σ-algebra $\mathcal{F}$ is called the family of events, where an event F is an element of $\mathcal{F}$, or, alternatively, an $\mathcal{F}$-measurable subset of Ω.

Definition 2.10. A measure $\mathbb{P}$ is called a **probability measure** if $\mathbb{P}(\Omega) = 1$.

3. Random Variables, Independence, and Expectation

Random variables are a concept which someone who has studied probability at any level would be familiar with. Naturally they occur in the measure-theoretic approach to probability as well. In order to present the definition of a random variable, we first must define Σ-measurable functions.

Definition 3.1. Let (S, Σ) be a measurable space, so that Σ is a σ-algebra on S. Suppose that $h : S \to$. For $A \subseteq \mathbb{R}$, define

$$h^{-1}(A) := \{s \in S : h(s) \in A\}.$$

Then h is called **Σ-measurable** if $h^{-1} : \mathcal{B} \to \Sigma$, that is, $h^{-1}(A) \in \Sigma, \forall A \in \mathcal{B}$.

We write mΣ for the the class of Σ-measurable functions on S, and $(m\Sigma)^+$ for the class of non-negative elements in mΣ. Now we can define what a random variable is.

Definition 3.2. Let $(\Omega, \mathcal{F})$ be our (sample space, family of event). A **random variable** is an element of m$\mathcal{F}$. Thus,

$$X : \omega \to \mathbb{R}, X^{-1} : \mathcal{B} \to \mathcal{F},$$

where X is the random variable.

Definition 3.3. Sub-σ-algebras $\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_n$ are **independent** if, for all events $\mathcal{F}_{i_k} \in \mathcal{F}_i$,

$$\mathbb{P}(\mathcal{F}_1 \cap \mathcal{F}_2 \cap ... \cap \mathcal{F}_n) = \Sigma_{i=1}^{n} \mathbb{P}(\mathcal{F}_i).$$

Definition 3.4. Random variables X_1, X_2,....,X_n are **independent** if their relevant σ-algebras $(\sigma(X_1), \sigma(X_2), ..., \sigma(X_n))$ are independent.

For independence, the familiar notion which does not involve σ-algebras is sufficient for our purposes so it is also presented here.

Definition 3.5. Events E_1, E_2, ... , E_n are independent if and only if whenever $n \in \mathbb{N}$ and i_1, ... , i_n are distinct, then

$$\mathbf{P}(E_{i_1} \cap ... \cap E_{i_n}) = \prod_{k=1}^{n} \mathbf{P}(E_{i_k}).$$

Next, we require a definition for expectation.

Definition 3.6. For a random variable $X \in \mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbf{P})$, we define the **expectation** $E(X)$ of X by

$$E(X) := \int_\Omega X dP = \int_\Omega X(\omega)\mathbf{P}(d\omega).$$

And lastly, we present the definition of conditional expectation.

Definition 3.7. Let (Ω, $\mathcal{F}$, $\mathbf{P}$) be a probability triple, and X a random variable with $E(|X|) < \infty$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Then there exists a random variable Y such that

(1) Y is $\mathcal{G}$ measurable.
(2) $E(|Y|) < \infty$
(3) for every set G in $\mathcal{G}$, we have

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

A random variable Y with properties (1), (2), and (3) is called a **a version of the conditional expectation** $E(X|\mathcal{G})$ of X given $\mathcal{G}$, and we write $Y = E(X|\mathcal{G})$, almost surely.

[1, Theorem 9.2]

Next we will present the Monotone-Convergence Theorem which will begin a chain of proofs required to prove the Dominated-Convergence Theorem and the Bounded Convergence Theorem which are necessary for Section 4.

Theorem 3.8. *If f_n is a sequence of elements of $(m\Sigma)^+$ such that $f_n \to f$, then*

$$\mu(f_n) \to \mu(f) \leq \infty$$

Proof. This is the Monotone Convergence Theorem and the proof can be found in the Appendix of [1]. □

Next we present the Fatou Lemma.

Theorem 3.9. *For a sequence (f_n) in $(m\Sigma)^+$,*

$$\mu(\liminf f_n) \leq \liminf \mu(f_n).$$

Proof. We have

$$\liminf_n f_n = \lim g_k, where g_k := \inf_{n \geq k} f_n. \tag{3.10}$$

For $n \geq k$, we have $f_n \geq g_k$, so that $\mu(f_n) \geq \mu(g_k)$, where

$$\mu(g_k) \leq \inf_{n \geq k} \mu(f_n);$$

and by combining this with an application of MON to (*), we obtain

$$\mu(\liminf_n f_n) = \lim_k \mu(g_k) \leq \lim_k \inf_{n \geq k} \mu(f_n) =: \liminf_n \mu(f_n).$$

□

And now the Reverse Fatou Lemma.

Theorem 3.11. *If (f_n) is a sequence in $(m\Sigma)^+$ such that for some g in $(m\Sigma)^+$, we have $f_n \leq g, \forall n$, and $\mu(g) < \infty$, then*

$$\mu(\limsup f_n) \geq \limsup \mu(f_n).$$

Proof. Apply Fatou Lemma to $(g - f_n)$. □

We now present the two theorems which are used in Section 4 for the proof of Doob's Optional-Stopping Theorem. The first is called the **Dominated-Convergence Theorem**.

Theorem 3.12. *If $|X_n(\omega)| \leq Y(\omega)$ for all n, ω and $X_n \to X$ pointwise almost surely, where $E(Y) < \infty$, then*

$$E(|X_n - X|) \to 0.$$

so that

$$E(X_n) \to E(X)$$

Proof. We have $|f_n - f| \leq 2g$, where $\mu(2g) < \infty$, so by the reverse Fatou Lemma,

$$\limsup \mu(|f_n - f|) \leq \mu(\limsup |f_n - f|) = \mu(0) = 0.$$

Since

$$|\mu(f_n) - \mu(f)| = |\mu(f_n - f)| \leq \mu(|f_n - f|)$$

□

The second theorem is the **Bounded Convergence Theorem**.

Theorem 3.13. *Let $\{X_n\}$ be a sequence of random variables, and let X be a random variable. Suppose that $X_n \to X$ in probability and that for some K in $[0,\infty)$, we have for every n and ω*

$$|X_n(\omega)| \leq K$$

Then

$$E(|X_n - X|) \to 0.$$

Proof. This is a direct consequence of the Dominated-Convergence Theorem and can be obtained by taking $Y(\omega) = K$, for all ω. □

4. Martingales

Now we take $(\Omega, \mathcal{F}, \mathbf{P})$ to be the probability triple which we are referring to. We will now present the definition of martingales after first providing the necessary definitions of filtrations and adapted processes.

Definition 4.1. Now instead of using probability triples as before, we will take a **filtered space**.

$$(\Omega, \mathcal{F}, \mathcal{F}_n, \mathbf{P})$$

$\{\mathcal{F} : n \geq 0\}$ is called a **filtration**. A filtration is an increasing family of sub-σ-algebras of $\mathcal{F}$ such that

$$\mathcal{F}_n \subseteq \mathcal{F}_1 \subseteq ... \subseteq \mathcal{F}_n$$

and $\mathcal{F}_\infty$ is defined as

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

Each filtration can intuitively be thought of as the information available about the events in a sample space after time n. Note that as n increases, more "information" is available and none of the previous information is lost.

Definition 4.2. A process $X = (X_n : n \geq 0)$ is called **adapted** to the filtration $\mathcal{F}_n$ if for each n, X_n is $\mathcal{F}_n$-measurable.

This is essentially stating that the value $X_n(\omega)$ is known at time n. Each of the $X_n(\omega)$ depends only on the information we have up to and including time n, that is they do not depend on information in the future.

Now we provide the definition of a martingale.

Definition 4.3. A process X is called a **martingale** if

(1) X is adapted
(2) $E(|X_n|) < \infty, \forall n$
(3) $E[X_n|\mathcal{F}_{n-1}] = X_{n-1}$, almost surely $(n \geq 1)$.

Definition 4.4. A **supermartingale** is defined the same as a martingale only with (3) replaced by

$$E[X_n|\mathcal{F}_{n-1}] \leq X_{n-1}.$$

Definition 4.5. Similarily, A **submartingale** is defined the same as a martingale only with (3) replaced by

$$E[X_n|\mathcal{F}_{n-1}] \geq X_{n-1}.$$

It is important to note is that a supermartingale decreases on average as the expected value of X_n given all previous information is less than X_{n-1}. Likewise, a submartingale increases on average. Addtionally, if X is a supermartingale, note that $-X$ is a submartingale, and that X is a martingale if and only if it is both a supermartingale and a submartingale.

It is also readily apparent how martingales and supermartingales can be thought of as fair and unfair games, respectively. If you consider $X_n - X_{n-1}$ to be your net winnings per unit stake in game n, with $n \geq 1$. If X is a martingale, you can see that

$$\begin{aligned} E[X_n|\mathcal{F}_{n-1}] = X_{n-1} &\Rightarrow E[X_n|\mathcal{F}_{n-1}] - X_{n-1} = 0 \\ &\Rightarrow E[X_n|\mathcal{F}_{n-1}] - E[X_{n-1}|\mathcal{F}_{n-1}] = 0 \\ &\Rightarrow E[X_n - X_{n-1}|\mathcal{F}_{n-1}] = 0. \end{aligned}$$

The second arrow follows as $E[X_{n-1}|\mathcal{F}_{n-1}] = X_{n-1}$, since X is adapted to $\mathcal{F}_{n-1}$ and thus you know exactly what X_{n-1} is by knowing $\mathcal{F}_{n-1}$. This fits the definition of a fair game as the expected net winnings is zero per game for all games. By using similar method, you can tell that when X is a supermartingale,

$$E[X_n - X_{n-1}|\mathcal{F}_{n-1}] \leq 0.$$

This is an unfair game, as your expected net winnings are negative after each game. The submartingale case is similar, but less interesting as very few people are concerned by a game where they win money on average.

Additionally this provides us with a new way to define martingales as all of the arrows can be proved in reverse.

Next we will define **previsible process**, which is essentially the mathematical expression of a particular gambling strategy.

Definition 4.6. Process $C = (C_n : n \in N)$ is called **previsible** if C_n is $\mathcal{F}_{n-1}$ measurable for $n \geq 1$.

Each C_n represents your particular stake in game n. You are able to use the history of your previous bets up to time $n-1$ to change C_n. Thus, $C_n(X_n - X_{n-1})$ are your winnings on game n and your total winnings up to time n are

$$Y_n = \sum_{1 \leq k \leq n} C_k(X_k - X_{k-1}) =: (C \bullet X)_n.$$

Clearly $(C \bullet X)_0 = 0$ as there are no winnings if no games are played. Additionally, we can obtain the winnings for game n as follows,

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

Now we define **stopping time** as it is important to our definition of Doob's Optional-Stopping Theorem.

Definition 4.7. A map $T : \Omega \to 0, 1, 2, ..., \infty$ is called a **stopping time** if,

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

or

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

For a stopping time T, it is possible to decide whether $\{T \leq n\}$ has occurred based on filtration $\mathcal{F}_n$, meaning that the event $\{T \leq n\}$ is $\mathcal{F}_n$-measurable. For example, if a person gambles until the either play ten games or run out of money, this is a stopping time. However, a person gambling until they win more money than they ever will is not, as that is not knowable with the known information at time n.

Now we must prove a theorem which is integral to our proof of Doob's Optional-Stopping Theorem. Before that, there is a bit of notation which may be new. For $x, y \in \mathbb{R}$,

$$x \wedge y := \min(x, y).$$

Theorem 4.8. *If X is a supermartingale and T is a stopping time, then the* **stopped process** $X^T = (X_{T \wedge n} : n \in \mathbb{Z}^+)$ *is a supermartingale, so that in particular,*

$$E(X_{T \wedge n}) = E(X_0), \forall n.$$

Note that this theorem does not say anything about

$$E(X_T) = E(X_0) \tag{4.9}$$

Now we can present a proof of **Doob's Optional-Stopping Theorem**.

Theorem 4.10. *(a) Let T be a stopping time. Let X be a supermartingale. Then X_T is integrable and*

$$E(X_T) \leq E(X_0)$$

in each of the following situations:

1. *T is bounded (for some N in $\mathbb{N}$, $T(\omega) \leq N, \forall \omega$),*
2. *X is bounded (for some K in $\mathbb{R}^+$, $|X_n(\omega)| \leq K$ for every n and every ω)*
3. *$E(T) < \infty$, and, for some K in $\mathbb{R}^+, |X_n(\omega) - X_{n-1}(\omega)| \leq K, \forall(n, \omega)$*

(b) If any of the conditions 1-3 hold and X is a martingale, then

$$E(X_T) = E(X_0).$$

Proof. First we will prove (a). We know that $X_{T\wedge n}$ is integrable and

$$E(X_{T\wedge n} - X_0) \leq 0 \tag{4.11}$$

because of Theorem 4.8. To prove (1), take n = N. For (2), let $n \to \infty$ in (4.11) using Theorem 3.7. For (3), we have

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

and $E(KT) < \infty$, so that Theorem 3.6 can be used to allow $n \to \infty$ in (4.11) which provides the necessary answer.

Next we prove (b). By applying (a) to X and $(-X)$ you will get two inequalities in opposite directions. This implies equality. □

Corollary 4.12. *Suppose that M is a martingale, the increments $M_n - M_{n-1}$ of which are bounded by some constant K_1. Suppose that C is a previsible process bounded by some constant K_2, and that T is a stopping time such that $E(T) < \infty$. Then,*

$$E(C \bullet M)_T = 0.$$

In essence, this corollary states that, assuming you have no knowledge of future events, you cannot beat a fair game.

Lemma 4.13. *Suppose that T is a stopping time such that for some N in $\mathbb{N}$ and some $\epsilon > 0$, we have, for every n in $\mathbb{N}$:*

$$\mathbb{P}(T \leq n + N|\mathcal{F}_n) > \epsilon, almost surely.$$

Then $E(T) < \infty$.

5. The Abracadabra Problem

To conclude the paper, I will now present a problem which can be solved with the material presented above.

A monkey types letters at random, one per each unit of time, producing an infinite sequence of identically independent random letters. If the monkey is equally likely to type any of the 26 letters, how long on average will it take him to produce the sequence

$$ABRACADABRA$$

Now it may initially appear difficult to solve this problem, but the use of martingale simplifies the problem greatly.

Construct a scenario where before each time $n = 1, 2, ..., m$ another gambler arrives. This gambler bets \$1 that the nth letter will be A. If he loses, then he leaves. If he wins, then he receives \$26 which he then bets all of on the event that the $(n+1)$th letter will be B. If he loses, he leaves. If he wins, then he bets his entire winnings of $\$26^2$ on the event that the (n+2)th letter will be R. This process continues until the entire ABRACADABRA sequence has occurred. Let T be the first time by which the monkey has produced the sequence ABRACADABRA. We are trying to find $E(T)$.

Let C_n^j be the bet of the jth gambler at time n. Define A_n to be the nth letter of the sequence.

$$C_n^j = \begin{cases} 0, & \text{if } n < j \\ 1, & \text{if } n = j \\ 26^k, & \text{if} A_j, \dots, A_{j+k-1} \text{were correct and } n = j + k \\ 0, & \text{otherwise} \end{cases}$$

It is clear that each C_n^j is a previsible process as it is determined only by information up to the $(n-1)$th bet. Now if we define M_n^j to be the martingale of payoffs after n bets for j gamblers, then it will simplify the solution to the problem.

To show that M_n^j is a martingale, we must show that

(1) M_n^j is adapted,
(2) $E(|M_n^j|) < \infty, \forall n$,
(3) $E[M_n^j|\mathcal{F}_{n-1}] = M_{n-1}^j, n \geq 1$

To show (1), simply notice that M_n^j is determined by the event A_n, and whether the letter typed at time n is correct. For (2), note that M_n^j is always positive and is bounded above by 26^n. As $E[|M_n^j|] = E[M_n^j] < 26^n < \infty$, (2) is satisfied. There are two cases for (3). First when the gambler loses before time n, then $E[M_n^j|\mathcal{F}_{n-1}] = 0 = M_{n-1}^j$. If the gambler wins the first $(n-1)$ times, then

$$E[M_n^j|\mathcal{F}_{n-1}] = 26^n \cdot 1/26 + 0 \cdot 25/26 = 26^{n-1} = M_{n-1}^j.$$

We can now apply Doob's Optional-Stopping Theorem because M_n^j is a martingale. All we must do now is show that one of its three conditions are satisfied. For simplicity we will choose (3),

$$E(T) < \infty, \text{and, for some } K \text{ in } \mathbb{R}^+, |X_n(\omega) - X_{n-1}(\omega)| \leq K, \forall (n, \omega). \tag{5.1}$$

First refer to (4.13). Let $N = 11$ and $\epsilon = (1/26)^N$. It is clear that whatever n is, there is always a $(1/26)^{11}$ chance that ABRACADABRA will occur in the next 11 letters. Thus, the condition holds with $N = 11$ and $E(T) < \infty$.

Now we must show the second part of the condition of (5.1). First define

$$X_n := \sum_{j=1}^{\infty} M_n^j = \sum_{j=1}^{n} M_n^j.$$

The second equality holds because after the stopping time, the $n+1$ gamblers have not begun betting so all terms after n are zero.

It is useful to view X_n as the cumulative winnings of every gambler up to and including time n. X_n is a martingale as it is simply the sum of expectations and M_n^j is a martingale. This can be easily shown from the defintions and can be worked as an exercise if a rigourous proof is needed to convince the reader.

Notice that,

$$|X_n - X_{n-1}| \leq 26^{11} + 26^4 + 26.$$

This is because $|X_n - X_{n-1}|$ denotes the maximum payoff at time n. To find the maximum, assume that the monkey has typed everything correctly and find the maximum amount of money that can be won after a single unit of time. Since each gambler wins increase the more correct bets they get in a row, it is easy to see that the first gambler has won \$$26^{11}$ at time 11 if he started winning at the first A. There can be no more winning gamblers until the 4th A because if the 1st gambler wins \$$26^{11}$ then the 2nd, 3rd, ... , 7th gamblers all must lose. The gambler who started winning at the 4th A can win a maximum of \$$26^4$ because there are four more letters that the monkey can type correctly. Lastly, the 11th gambler can also win \$26 because the last letter is an A.

By meeting the requirements for Doob's Optional-Stopping Theorem, we can utilize its conclusion:

$$E(X_T) = E(X_0)$$

$E(X_0) = 0$ because nothing happens at time 0. $E(X_T)$ is the cumulative winnings of all the gamblers after the monkey types ABRACADABRA correctly. As shown earlier, the winnings will equal $26^{11} + 26^4 + 26$. However, in calculating these winnings, we neglected to include the money that the gamblers lose with each subsequent letter. As each gambler arrives, he bets \$1 so after time T, there have been \$$T$ dollars lost because of these \$1 bets. Thus,

$$\begin{aligned} E(X_T) &= E(26^{11} + 26^4 + 26 - T) \\ &= 26^{11} + 26^4 + 26 - E(T) = 0 \\ \iff E(T) &= 26^{11} + 26^4 + 26. \end{aligned}$$

The key reason for the solution being possible this way is the fact that $E(T)$ appears in the calculation for $E(X_T)$. This was possible due to the way we defined the problem, with each gambler betting \$1 upon their arrival. This is a useful trick to remember as this method works for computing $E(T)$ of any pattern in a random sequence of symbols. For example $E(T) = 6^4 + 6^3 + 6^2 + 6^1$ for the event where someone rolls four 6's in a row on a fair six-sided die.

Acknowledgments. It is a pleasure to thank my mentor, Ilya Gehktman, for his knowledgable insights and his flexible scheduling.

References

[1] Michael Williams. Probability with Martingales. Cambridge University Press. 1991.