MARTINGALES AND THE ABRACADABRA PROBLEM

DI AI

Abstract. In this exposition, we will present the definition of martingales using measure theory and an application of it to solve the ABRACADABRA problem, which involves computing the expected time of the first appearance of a pattern in a random sequence of letters.

Contents

1. Introduction

I will present the theory of martingales based on measure theory. Martingales are very useful tools that can be applied to a wide range of interesting mathematical problems. To put it simply, martingales are based on the notion of a fair game, where your expected winnings after a certain number of turns or bets is the same as your expected winnings on the first bet.

The rest of the paper will be structured as follows. Section 2 will be an introduction to measure theory. In section 3, we review concepts from probability theory such as random variables and expectation that are used frequently in martingale theory. Section 4 will focus on developing martingales and the crucial Doob's Optional-Stopping Theorem. I will conclude in section 5 with an application of martingales by solving exercise 10.6 in *Probability with Martingales*. This problem involves finding the expected time it takes for a monkey to type the letters ABRACADABRA correctly, in that order. I have followed *Probability with Martingales* by David Williams [1] closely and all the theorems and definitions presented here can be found in his book.

2. Measure Theory

In order to present martingales rigorously, we must first introduce basic concepts from measure theory such as a probability triple $(\Omega, \mathcal{F}, \mathbf{P})$. Ω is known as the

Date: July 22, 2011.

sample space, which is simply a set containing all of the possible outcomes of an experiment, trial, or some other random process. $\mathcal{F}$ is known as a σ-algebra on Ω.

Definition 2.1. Let S be a set. Then, a collection Σ_0 of subsets of S is called an *algebra on* S (or algebra of subsets of S) if

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

Note that $\emptyset = S^c \in \Sigma_0$ and

$$F, G \in \Sigma_0 \Rightarrow F \cap G = (F^c \cup G^c)^c \in \Sigma_0.$$

Thus, an algebra on S is a family of subsets of S stable under finitely many set operations.

Definition 2.2. A collection Σ of subsets of S is called a *σ-algebra on S* (or σ-algebra of subsets of S) if Σ is an algebra on S such that *whenever* $F_n \in \Sigma$ ($n \in \mathbb{N}$), then

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

In addition, note that

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

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

You can think of a σ-algebra as containing all of the information that we know about a particular sample space. This concept will be important in our discussion of martingales below. It is the collection of sets over which a measure is defined. This is because not all sets are measurable and the σ-algebra contains only sets which are measurable. Another way to think about σ-algebras is to imagine a σ-algebra F corresponding to collections of yes or no questions that one can ask about the outcome of an experiment. Each set in F can be thought of as those outcomes in Ω for which there is an answer to the yes or no question. For example, let Ω be all possible times it will rain today. Then, one set in F could be all the outcomes where it rains after 3:30pm today. Of course, we could also ask the complement, "Did it rain before or at 3:30pm today?" The structure imposed on σ-algebras now becomes more transparent. Note that we can also string together multiple yes or no questions by putting "or's" between them, which is the equivalent of taking unions, to get another yes or no question.

We will now introduce a special type of σ-algebra called a **Borel σ-algebra**. But first, we need to understand what it means for a collection of sets to **generate** a σ-algebra.

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 $C \subseteq \Sigma$. In other words, it is the intersection of all σ-algebras on S which have C as a subclass.

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 set that can be written with any countable combination of the set operations union and intersection of closed and open sets.

Let us define a measure. First we will need the definitions of countably additive measures and measure spaces.

Definition 2.3. 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 $F = \bigcup F_n \in \Sigma_0$ (note that this is an assumption since Σ_0 need not be a σ-algebra), then

$$\mu_0(F) = \sum_n \mu_0(F_n).$$

Definition 2.4. A pair (S, Σ), where S is a set and Σ is a σ-algebra on S, is called a **measurable space**. An element of Σ is called a Σ-measurable subset of S.

Definition 2.5. 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**.

Intuitively, the measure simply assigns a real number to subsets of S. This can be thought of as the "size" or "volume" of a set. While the measure can be chosen arbitrarily, it must be countably additive. This ensures that when we combine two sets, the combined measure is larger than either of the individual measures.

We now define some key vocabulary of the probability triple $(\Omega, \mathcal{F}, \mathbf{P})$.

Definition 2.6. Ω is a set called the **sample space**.

Definition 2.7. A point ω of Ω is called a **sample point**.

Definition 2.8. The σ-algebra $\mathcal{F}$ on Ω is called the family of events, so that an **event** is an element of $\mathcal{F}$, that is, an $\mathcal{F}$-measurable subset of Ω.

3. Random Variables, Independence, and Expectation

Random variables are a key concept that we will work with extensively. Often when solving problems in probability, one will need to define one or more random variables. In order to present the definition of a random variable, we first must study Σ-measurable functions.

Definition 3.1. Let (S, Σ) be a measurable space, so that Σ is a σ-algebra on S. Suppose that $h : S \to \mathbb{R}$. 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 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 events). A **random variable** is an element of m$\mathcal{F}$. Thus,

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

where X is the random variable.

As for the definition of independence, we will use the more familiar notion that does not involve σ-algebras. This definition will be sufficient for our purposes.

Definition 3.3. Events E_1, E_2, ... are independent if and only if whenever $n \in \mathbb{N}$ and i_1,...,i_n are distinct, then

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

Lastly, we need to define expectation.

Definition 3.4. 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

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

And here we will present the definition of conditional expectation.

Definition 3.5. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a 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

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

A random variable Y with properties (1) - (3) is called **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]

We will now present two theorems that will be used later on in the proof of the Doob's Optional-Stopping Theorem. The first one is called the **Dominated-Convergence Theorem**.

Theorem 3.6. *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

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

[1, Theorem 5.9]

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

Theorem 3.7. [1, Theorem 6.2] *If for some finite constant K, $|X_n(\omega)| \leq K$, for all n, ω and $X_n \to X$ pointwise almost surely, then*

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

[1, Theorem 6.2]

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

4. Martingales

From now on, $(\Omega,\ \mathcal{F},\ \mathbf{P})$ will be the probability triple that we will be referring to. In order to define what a martingale is, we first need to define filtrations and adapted processes.

Definition 4.1. Instead of a probability triple, we now take a **filtered space**

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

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

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq ... \subseteq \mathcal{F},$$

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

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

Intuitively, each filtration can be thought of as the information available about the events in a sample space after time n (keep in mind that σ-algebras consist of all the information that we know about events in a sample space).

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.

In other words, the value $X_n(\omega)$ is known to us at time n. Each of the $X_n(\omega)$ depends only on the information we have previously up to and including time n, but does not depend on information in the future, after time n.

Now we present 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 $(n \geq 1)$.

A **supermartingale** is defined similarly, except that the last condition is replaced by

$$E[X_n|\mathcal{F}_{n-1}] \leq X_{n-1},\ \textit{almost surely}\ (n \geq 1),$$

and a **submartingale** is defined with the last condition replaced by

$$E[X_n|\mathcal{F}_{n-1}] \geq X_{n-1},\ \textit{almost surely}\ (n \geq 1).$$

One important thing to notice is that a supermartingale decreases on average, as the expected value of X_n given all previous information is lower than X_{n-1}, and a submartingale increases on average, as the expected value of X_n given all previous information is higher than X_{n-1}. In addition, note that X is a supermartingale if and only if $-X$ is a submartingale, and that X is a martingale if and only if it is both a supermartingale and a submartingale.

We now describe how martingales and supermartingales can be thought of as fair and unfair games, respectively. Take $X_n - X_{n-1}$ and think of it as your net winnings per unit stake in game n, where $n \geq 1$. There is a series of games played at times $n = 1,2,...$ and there is no game at time 0. If X is a martingale, we will see that

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

The second implication follows because $E[X_{n-1}|\mathcal{F}_{n-1}] = X_{n-1}$, since X is adapted to $\mathcal{F}_{n-1}$ and therefore knowing $\mathcal{F}_{n-1}$ will tell you exactly what X_{n-1} is. The last implication follows because the sum (difference) of expectations is the expectation of the sums (differences). This is the definition of a fair game, when your expected net winnings per game is zero for all games. In the supermartingale case, where X is a supermartingale instead of a martingale, we can obtain through similar methods

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

In this case, the game is unfair because your expected net winnings are negative after each game. We ignore the submartingale case because we typically are not concerned with games where you expect to win money after every game.

Lastly, note that this gives us a new way to define a martingale or supermartingale, as all of the above implications can be reversed.

We will now define a **previsible process**, which can be thought of as a mathematical way of expressing a particular gambling strategy.

Definition 4.4. We call a process $C = (C_n : n \in \mathbb{N})$ **previsible** if C_n is $\mathcal{F}_{n-1}$ measurable for $n \geq 1$.

Each C_n represents your particular stake on game n. You can change C_n based upon the history of your previous bets up to and including time $n-1$. Recall that $X_n - X_{n-1}$ is your net winnings per unit stake in game n. It thereby follows that $C_n(X_n - X_{n-1})$ is your winnings on game n and your total winnings up to time n are

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

Since you can't win anything if no games have been played, $(C \bullet X)_0 = 0$. In addition, by taking the difference of two consecutive total winnings, we can recover the winnings on game n.

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

Now we will define what a stopping time is. This will become important later on in our definition of Doob's Optional-Stopping Theorem, which will be the key to solving the ABRACADABRA problem.

Definition 4.5. 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 equivalently,

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

This means that for a stopping time T, it is possible to decide whether $\{T \leq n\}$ has occurred based on the filtration $\mathcal{F}_n$, or in other words the event $\{T \leq n\}$ is $\mathcal{F}_n$-measurable. For example, a gambler who will leave when he runs out of money or has played 100 games or rounds is a stopping time, whereas a gambler who plays until he has won more money then he ever will is not a stopping time because this requires knowledge about both the past and the future, not just the past.

We now present a theorem that will be used in the proof of the Doob's Optional-Stopping Theorem. First, let me introduce some notation. For a, $b \in \mathbb{R}$,

$$a \wedge b := \min(a, b).$$

Theorem 4.6. *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,*

$$E(X_{T\wedge n}) \leq E(X_0), \text{ for all } n.$$

Similarly, if X is a martingale and T is a stopping time, then X^T is a martingale, so that in particular,

$$E(X_{T\wedge n}) = E(X_0), \text{ for all } n.$$

Note that this theorem does not say anything about when

$$E(X_T) = E(X_0) \tag{4.7}$$

for a martingale X. One often desires this to hold and this does indeed hold in general, but not in all cases. Doob's Optional-Stopping Theorem will give us some sufficient conditions when (4.7) holds.

Finally, here is the **Doob's Optional-Stopping Theorem**.

Theorem 4.8.
(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) and T is almost surely finite;*
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 holds 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

$$E(X_{T\wedge n} - X_0) \leq 0 \tag{4.9}$$

because of Theorem 4.6. To prove (1), we can take $n = N$. For (2), let $n \to \infty$ in (4.9) using the Bounded Convergence Theorem. 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 the Dominated-Convergence Theorem applies and justifies letting $n \to \infty$ in (4.9) to obtain the answer that we want.

To prove (b), simply apply (a) to X and to $(-X)$. You will get two inequalities in opposite directions which will imply equality. □

Here is an important corollary of Doob's Optional-Stopping Theorem.

Corollary 4.10. *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,*

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

Roughly speaking, this corollary shows that you cannot beat a fair game no matter what your gambling strategy - provided that you cannot look into the future.

5. ABRACADABRA Problem

Now I will finally present the problem that we will solve using the Martingale theory we have presented above. At first glance, this problem will seem difficult to solve, but will be much easier with the use of martingales.

At each of times 1,2,3,... a monkey types a capital letter at random, the sequence of letters typed forming an independent and identically distributed sequence of random variables each chosen uniformly from amongst the 26 possible capital letters.

Just before each time n = 1,2,..., a new gambler arrives on the scene. He bets \$1 that the n^{th} letter will be A. If he loses, he leaves. If he wins, he receives \$26 all of which he bets on the event that the $(n+1)^{th}$ letter will be B. If he loses, he leaves. If he wins, he bets his whole fortune of \$ 26^2 that the $(n+2)^{th}$ letter will be R and so on through the ABRACADABRA sequence. Let T be the first time by which the monkey has produced the consecutive sequence ABRACADABRA. Show that

$$E(T) = 26^{11} + 26^4 + 26$$

and prove this.

The intuition behind the solution is to note that at each time period, a new gambler comes and bets \$1 before he wins or loses. Thus, after T periods, there will have been a total of T dollars bet. By then taking the expected value of the total winnings of all the gamblers, we are left with this T term, which is what we wanted to solve for. Below is a more formal solution.

Let us first define some variables. Define A_n to be the n^{th} letter of the sequence. Let C_n^j be the bet of the j^{th} gambler betting at time n.

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

We can easily see that each C_n^j is a previsible process because it is determined only by using information from and up to the $(n-1)^{th}$ bet. Now let us define the martingale M_n^j to be the payoff after n bets for the j^{th} gambler. Note that this definition of M_n^j already has C_n^j built into it because we have defined M_n^j to be the total payoff and not the payoff per unit stake. It will be easier to solve the problem by defining M_n^j in this manner. In addition, note that each gambler leaves if he loses any one of his bets. Thus, if the $(n+1)^{th}$ wager is to be made, then $M_n^j = 26^n$ and $M_{n+1}^j = 26^{n+1}$ with probability 1/26 and 0 with probability 25/26. This is because the monkey types each letter randomly and independently of the previous letters, with each letter having an equal probability of being typed at any given moment. 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), we note that M_n^j is determined by the event A_n and whether the letter that is typed at time n is correct. For (2), we note that M_n^j is always positive and that the max value of M_n^j is 26^n, that is when the gambler wins all previous n times. Since $26^n < \infty$ and $E[|M_n^j|] = E[M_n^j] < 26^n$, (2) is satisfied. There are two cases for (3). If the gambler loses anywhere before time n, then given that information $E[M_n^j|\mathcal{F}_{n-1}] = 0 = M_{n-1}^j$. If the gambler wins the first $(n-1)$ times, then

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

Now that we have shown that M_n^j is a martingale, we can apply Doob's Optional-Stopping Theorem. As you know, there are three conditions in the theorem, one of which must be satisfied. I will now show that condition (3) is satisfied, which is reproduced below for convenience.

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

To show that $E(T) < \infty$, we will need the following Lemma.

Lemma 5.2. *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 $\mathbb{N}$:*

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

Then $E(T) < \infty$.

For the ABRACADABRA problem, let $N = 11$ and $\epsilon = \left(\frac{1}{26}\right)^N$. Now, no matter what n is, there is a $\left(\frac{1}{26}\right)^{11}$ chance that ABRACADABRA will be typed in the next 11 letters. In other words, no matter where in the sequence we are, there is a small chance that the next 11 letters will be correct Thus, the condition holds with $N = 11$ and $\epsilon = \left(\frac{1}{26}\right)^{11}$ and $E(T) < \infty$.

Now we need to show the second part of condition (5.1). First, let us 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 haven't even started playing yet and thus all terms after n are zero.

Think of X_n as representing the cumulative winnings of every gambler up to and including time n. Since we have already shown that each M^j is a martingale and X_n is simply a sum and the expectation of a sum is the sum of expectations, this implies that X_n is a martingale as well.

Now, note 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, just simply assume that the monkey has typed everything correctly and find the maximum amount of money that can be won after one 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 4^{th} A because if the 1^{st}

gambler wins \$$26^{11}$ then the 2^{nd}, 3^{rd}, ... , 7^{th} gamblers all must lose. The gambler who started winning at the 4^{th} 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.

Now that we have met the requirements for Doob's Optional-Stopping Theorem, we can apply its conclusion:

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

In this case, $E(X_0) = 0$ because nothing happens at time 0. $E(X_T)$ is simply the cumulative winnings of all the gamblers after the monkey types out ABRACADABRA correctly. As described in the previous paragraph, the winnings will be \$$26^{11} + 26^4 + 26$. There is an interesting case at the 3rd A in ABRACADABRA. While the gambler who bets that the sixth letter will be an A will win \$26, he will lose his money on the next bet because the next letter is a D and not a B. In calculating the total winnings of all the gamblers, we forgot to take into account that each time a new gambler arrives, he bets \$1 regardless of whether he wins or loses. Thus, after time T, there will have been \$$T$ dollars lost because of the initial \$1 bet. Thus,

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

Note that the reason we were able to solve this problem was because $E(T)$ showed up in the calculation for $E(X_T)$. This is because of the unique way this problem was defined, where each gambler bet \$1 at the beginning. Note that this method works for computing $E(T)$ of any pattern in a random sequence of symbols, such as flipping a coin and looking for HHTT. This trick is a useful one to remember, as it makes calculating the expectation of some things much easier.

Acknowledgments. It is a pleasure to thank my mentors, Markus Kliegl and Brent Werness, for their support and encouragement through the REU. They gave me advice on everything from how to structure a math paper to how to use LaTeX. Without them, this would not have been possible.

References

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