Coins Flips, Fair Casinos and Martingales

An Introduction to the Generalized ABRACADABRA Theorem

Eytan Chong

May 7, 2025

1 A Coin-Flip Problem

Our journey begins with a classic coin-flip problem:

> *A fair coin is flipped repeatedly until a given sequence of Heads and Tails appears. On average, how many times is the coin flipped?*

Let us unpack this problem and phrase it mathematically. To do so, we introduce the following notation and terminology.

Definition 1. Let $\mathcal{A}$ be an **alphabet**, which is the set of characters from which words are constructed.

In the case of coin-flips, $\mathcal{A} = \{H, T\}$, where H represents Heads and T represents Tails.

Definition 2. A **terminator** is a word that terminates the coin-flipping. The set of all terminators is denoted $\mathcal{T}$.

Definition 3. A word w is said to be **immediately terminated** under $\mathcal{T}$ if

- w ends with a terminator $t \in \mathcal{T}$; and
- w contains no other terminators.

The set of all words immediately terminated under $\mathcal{T}$ is denoted $\mathcal{I}_{\mathcal{T}}$.

Example 4. Let $\mathcal{T} = \{HHT, THH\}$. That is to say, we stop flipping the coin the moment we get HHT or THH. The set of words we might get when playing the game is then

$$\mathcal{I}_{\mathcal{T}} = \{HHT, HHHT, HHHHT, \ldots, THH, HTHH, TTHH, \ldots\}.$$

Note that the word $HTHHT$, despite ending with HHT, is not in $\mathcal{I}_{\mathcal{T}}$. This is because it contains another terminator: $HTHHT$.

To prevent nonsensical scenarios, such as $\mathcal{T} = \{HTT, HT\}$ or $\mathcal{T} = \{HTH, TH\}$, we require that a terminator cannot contain another terminator. Equivalently, $\mathcal{T} \subseteq \mathcal{I}_{\mathcal{T}}$.

We now rephrase our original problem:

> *Let $W_{\mathcal{T}}$ be a word constructed by concatenating letters from $\mathcal{A}$ uniformly at random until $W_{\mathcal{T}} \in \mathcal{I}_{\mathcal{T}}$, and let $L_{\mathcal{T}} = |W_{\mathcal{T}}|$ be its length. What is $\mathbb{E}[L_{\mathcal{T}}]$?*

For now, we will simplify the problem and assume $|\mathcal{T}| = 1$. In the following subsections, we will present two common approaches one might take in answering this (simplified) problem.

1.1 A Naive Approach

Suppose $\mathcal{T} = \{T\}$. Intuitively, because the probability of getting T is $1/2$, one might guess that we will, on average, get one T every two flips, so

$$\mathbb{E}\left[L_{\{T\}}\right] = \frac{1}{\mathbb{P}[T]} = \frac{1}{1/2} = 2.$$

This is indeed the correct answer.

Suppose now that $\mathcal{T} = \{TH\}$. Following a similar line of reasoning, one might conclude that

$$\mathbb{E}\left[L_{\{TH\}}\right] = \frac{1}{\mathbb{P}[TH]} = \frac{1}{1/4} = 4,$$

which is once again the correct answer.

However, this argument quickly breaks down once we consider more complicated terminators. For instance, if $\mathcal{T} = \{THT\}$, the above pattern suggests that

$$\mathbb{E}\left[L_{\{THT\}}\right] = \frac{1}{\mathbb{P}[THT]} = \frac{1}{1/8} = 8.$$

However, empirical evidence suggests that $\mathbb{E}\left[L_{\{THT\}}\right]$ is actually 10, as shown in Figure 1.

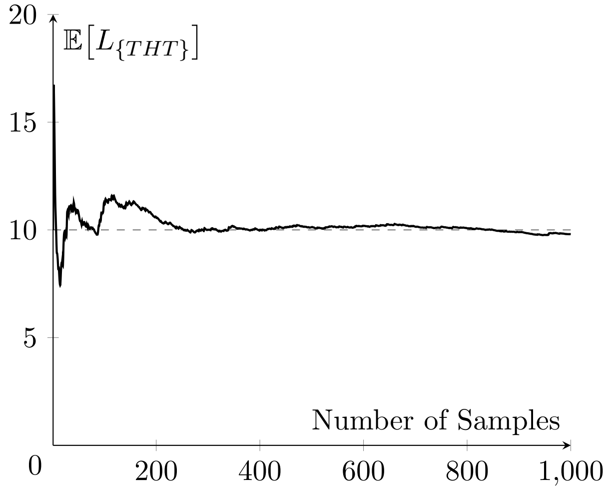

Figure 1: 1 thousand samples of $L_{\{THT\}}$ suggests that $\mathbb{E}\left[L_{\{THT\}}\right]$ is 10, not 8.

1.2 Case-by-Case Analysis

One might observe that depending on which face the coin lands on, the expected number of flips will change accordingly. Thus, by analysing all possible cases, we can form an equation in $\mathbb{E}[L_{\mathcal{T}}]$, which we can then easily solve.

To facilitate further discussion, we first introduce the notion of a left- and right-slice of a word.

Definition 5. The **left-slice** of a word w, denoted $L_n(w)$, refers to the first n characters of w. Analogously, the **right-slice** of w, denoted $R_n(w)$, refers to the last n characters of w.

Example 6. Let $w = HTHH$. The following table gives the left- and right- slices of w for different n.

n	$L_n(w)$	$R_n(w)$
1	H	H
2	HT	HH
3	HTH	THH
4	HTHH	HTHH

To illustrate the method of case-by-case analysis, consider $\mathcal{T} = \{THT\}$.

1. If the first coin is H, we have effectively wasted one flip since starting with H does not contribute to getting THT. Thus,

$$\mathbb{E}[L_\mathcal{T} \mid L_1(W_\mathcal{T}) = H] = \mathbb{E}[L_\mathcal{T}] + 1.$$

2. If the first coin is T, we have two subcases to consider:
   a) If the second coin is T, we have effectively "gone back" to the case where our first coin is T. Thus,

$$\mathbb{E}[L_\mathcal{T} \mid L_2(W_\mathcal{T}) = TT] = \mathbb{E}[L_\mathcal{T}] .$$

   b) If the second coin is H, we have two more subcases to consider:
      i. If the third coin is T, we have reached the terminator. Thus,

$$\mathbb{E}[L_\mathcal{T} \mid L_3(W_\mathcal{T}) = THT] = 3.$$

      ii. If the third coin is H, we have effectively "gone back" to the case where our first coin is H. This means that we wasted 3 flips, so

$$\mathbb{E}[L_\mathcal{T} \mid L_3(W_\mathcal{T}) = THH] = \mathbb{E}[L_\mathcal{T}] + 3.$$

Since all words must start with either H, TT, THT or THH, by the law of total expectation,

$$\begin{aligned}\mathbb{E}[L_\mathcal{T}] = \quad & \mathbb{E}[L_\mathcal{T} \mid L_1(W_\mathcal{T}) = H]\ \mathbb{P}[L_1(W_\mathcal{T}) = H] \\ & + \mathbb{E}[L_\mathcal{T} \mid L_2(W_\mathcal{T}) = TT]\ \mathbb{P}[L_2(W_\mathcal{T}) = TT] \\ & + \mathbb{E}[L_\mathcal{T} \mid L_3(W_\mathcal{T}) = THT]\ \mathbb{P}[L_3(W_\mathcal{T}) = THT] \\ & + \mathbb{E}[L_\mathcal{T} \mid L_3(W_\mathcal{T}) = THH]\ \mathbb{P}[L_3(W_\mathcal{T}) = THH] .\end{aligned}$$

Because the coin is fair, the probability that $L_n(W_\mathcal{T}) = w$ for some arbitrary word w is simply $1/2^n$. Substituting the values we found,

$$\mathbb{E}[L_\mathcal{T}] = \frac{\mathbb{E}[L_\mathcal{T}] + 1}{2^1} + \frac{\mathbb{E}[L_\mathcal{T}]}{2^2} + \frac{3}{2^3} + \frac{\mathbb{E}[L_\mathcal{T}] + 3}{2^3}.$$

After simplification, we get $\mathbb{E}[L_\mathcal{T}] = 10$, which aligns with the results obtained from our simulation.

Of course, this is a perfectly sound solution to the problem, and one can always calculate the correct value of $\mathbb{E}[L_\mathcal{T}]$ using this algorithm. However, it becomes incredibly inefficient and tedious when the terminators become more complicated, rendering it effectively useless.

2 Martingales and the Optional Stopping Theorem

Given the drawbacks of case-by-case analysis, we seek a more simple and elegant way to calculate $\mathbb{E}[L_{\mathcal{T}}]$. For this, we turn to a special mathematical object called a martingale.

2.1 Martingales

Put simply, a martingale is a random process represented by a sequence of random variables $\{X_n\}$, which typically models a gambler's fortune in a fair game.

To motivate our formal definition of a martingale, we investigate what it means for a game to be fair. Consider the following game:

> *Flip a fair coin. If it comes up Heads, we win $1, but if it comes Tails, we lose $1. Repeat this process forever.*

Let X_n be our wealth after the nth coin-flip, and let Y_n represent the outcome of the nth coin-flip. We make two observations regarding Y_n:[1]

- The coin is *fair*: The coin has a 50/50 chance of landing H or T. Mathematically,

$$\mathbb{P}[Y_n = H] = \mathbb{P}[Y_n = T] = \frac{1}{2}.$$

- The coin-flips are *independent*: The outcomes of past flips will not influence the outcomes of future flips. Mathematically,

$$\mathbb{P}[Y_{n+1} \mid Y_1, \ldots, Y_n] = \mathbb{P}[Y_{n+1}]\,.$$

Because of these two properties, we can derive an important fact about our wealth, X_n:

$$\begin{aligned}
&\mathbb{E}[X_{n+1} \mid Y_1, Y_2, \ldots Y_n] \\
&\quad = (X_n + 1)\,\mathbb{P}[Y_{n+1} = H \mid Y_1, \ldots, Y_n] + (X_n - 1)\,\mathbb{P}[Y_{n+1} = T \mid Y_1, \ldots, Y_n] \\
&\quad = (X_n + 1)\,\mathbb{P}[Y_{n+1} = H] + (X_n - 1)\,\mathbb{P}[Y_{n+1} = T] \\
&\quad = \frac{1}{2}\,(X_n + 1) + \frac{1}{2}\,(X_n - 1) = X_n.
\end{aligned}$$

That is to say, ***our expected wealth after the next flip, given that we know all previous outcomes, is exactly our current wealth.*** It is this equation,

$$\mathbb{E}[X_{n+1} \mid Y_1, \ldots, Y_n] = X_n,$$

that is the defining property of a martingale.

Definition 7. A sequence of random variables $\{X_n\}$ is a **martingale** with respect to the sequence $\{Y_n\}$ if

- X_n is a function of $Y_1, \ldots, Y_n$,
- $\mathbb{E}[X_n]$ is finite, and
- $\mathbb{E}[X_{n+1} \mid Y_1, \ldots, Y_n] = X_n$.

[1] Though these properties may seem trivial, it is nevertheless important to highlight them as it may sometimes run against our human intuition. For example, the probability that the next flip is Heads, given that the previous 100 flips were all tails, will still remain at 1/2. This misguided belief that we are more likely to win after a series of losses is commonly known as the Gambler's Fallacy.

Remark. Most of the time, Y_n is the same sequence as X_n. This simply means that

$$\mathbb{E}[X_{n+1} \mid X_1, \ldots, X_n] = X_n.$$

Apart from gambling, another context in which martingales can be applied in is the stock market.

Proposition 8. Let X_n represent the price of a stock on day n. Then $\{X_n\}$ is a martingale with respect to itself.

Proof. Suppose $\mathbb{E}[X_{n+1} \mid X_1, \ldots, X_n] > X_n$. Two things happen:

- Buying the stock today and selling tomorrow yields a profit (in expectation). Demand thus increases.
- Those that own the stock today will not sell today, since its value is expected to increase tomorrow. Supply thus decreases.

This increase in demand and decrease in supply bids up today's stock price.

Now suppose $\mathbb{E}[X_{n+1} \mid X_1, \ldots, X_n] < X_n$. Two things happens:

- Those who want to buy the stock would rather buy it tomorrow, since it will be cheaper. Demand thus decreases.
- Those that own the stock today will want to sell today, since its value is expected to decrease tomorrow. Supply thus increases.

This decrease in demand and increase in supply drives down today's stock price.

From the above two scenarios, it follows that today's stock price will eventually reach an equilibrium, where

$$\mathbb{E}[X_{n+1} \mid X_1, \ldots, X_n] = X_n.$$

This is exactly the condition for $\{X_n\}$ to be a martingale! □

2.2 Stopping Times and Strategies

We now introduce the notion of a stopping time.

Definition 9. A **stopping time** τ with respect to a sequence $\{Y_n\}$ is a random variable taking values in $\mathbb{N} \cup \{\infty\}$ such that for all $n \in \mathbb{N}$, the event $\{\tau = n\}$ depends solely on $Y_1, \ldots, Y_n$. This event is called the gambler's **stopping strategy**.

In layman terms, τ can be thought of as the round at which a gambler quits playing the game. The condition that $\{\tau = n\}$ depends solely on $Y_1, \ldots, Y_n$ means that the gambler quits using only information available to him before round n; he cannot see into the future (view the outcome of $Y_{n+1}, Y_{n+2}, \ldots$) to decide when to stop playing.

Example 10. Suppose a gambler employs a stopping strategy where he quits after playing 10 games. Then his stopping time is simply $\tau = 10$.

Another gambler may employ a different stopping strategy, opting to quit after losing three times in a row. If the gambler plays a fair game with a \$1 stake, the event $\{\tau = n\}$ can be expressed as

$$\Bigg\{ \underbrace{Y_1 = -1,\, Y_2 = -1,\, Y_3 = -1,\, \ldots,\, Y_{n-3} = -1}_{\text{all losses}},\, \underbrace{Y_{n-2} = 1,\, Y_{n-1} = 1,\, Y_n = 1}_{\text{3 wins in a row}} \Bigg\}.$$

2.3 The Optional Stopping Theorem

Is there a stopping strategy that returns a profit (in expectation)? As it turns out, the answer is generally a no. This result is known as Doob's Optional Stopping Theorem, or OST for short.

Theorem 11 (Doob's Optional Stopping Theorem)**.** Let $\{X_n\}$ be a martingale and let τ be a stopping time, both with respect to $\{Y_n\}$. Then $\mathbb{E}[X_\tau] = X_0$ if at least one of the following holds:

1. $|X_n|$ is bounded.
2. τ is bounded.
3. $\mathbb{E}[\tau]$ is finite, and all increments of X are bounded, i.e. there exists a constant C such that for all n,
$$|X_{n+1} - X_n| \leq C.$$

The OST tells us that as long as our stopping strategy is *reasonable enough*, our expected payout, $\mathbb{E}[X_\tau]$, must be equal to the amount we started with, X_0.

To see why all reasonable strategies obey the OST, suppose we somehow came up with a profitable strategy. That is, we managed to force $\mathbb{E}[X_\tau] > X_0$. Then this strategy either

- breaks the validity of our stopping time, or
- breaks all three conditions of the OST.

An invalid stopping time implies that we can somehow look into the future, which is clearly impossible. Furthermore, the three conditions of the OST are hard to break in real life:

- If $|X_n|$ is unbounded, then we either gain or lose an infinite amount of money, which is unrealistic since there is a finite amount of money in the world.
- If $\tau = \infty$, then we must play the game forever. Unfortunately, we have finite lifespans.
- The same problems arise if $\mathbb{E}[\tau] = \infty$, or if $|X_{n+1} - X_n|$ is unbounded.

Thus, a profitable strategy is nigh impossible to come up with, and so for all practical purposes, any strategy we come up with obeys the OST.

2.4 The Gambler's Ruin Problem

A classic application of martingales and the OST is the Gambler's Ruin Problem. Consider the following scenario:

> *We start with \$K. Each round, we flip a fair coin. If it lands H, we gain \$1. If it lands T, we lose \$1. We keep playing until we go bankrupt, or have a total of \$N.*
>
> *What is the probability of going bankrupt?*

Let us formalize this with martingales. Let Y_n be the outcome of the nth flip, and let X_n be our wealth after the nth flip. The stopping time τ is defined as

$$\tau = \min\{n : X_n = 0 \text{ or } X_n = N\}.$$

In our notation, the probability of going bankrupt is $\mathbb{P}[X_\tau = 0]$, while the probability of leaving with \$$N$ is $\mathbb{P}[X_\tau = N]$. Our goal is to find these two probabilities.

Recall back in Section 2.1, we showed that $\{X_n\}$ is a martingale. Further, it is quite clear that X_n is bounded (in particular, $0 \le X_n \le N$). Thus, $\{X_n\}$ and τ satisfy the first scenario of the OST, from which we obtain

$$\mathbb{E}[X_\tau] = X_0 = K.$$

Now observe that we can also write $\mathbb{E}[X_\tau]$ in terms of the desired probabilities:

$$\mathbb{E}[X_\tau] = N\,\mathbb{P}[X_\tau = N] + 0\,\mathbb{P}[X_\tau = 0]\,.$$

It immediately follows from the previous two equalities that

$$\mathbb{P}[X_\tau = N] = \frac{K}{N} \quad \text{and} \quad \mathbb{P}[X_\tau = 0] = 1 - \frac{K}{N}.$$

In context, this result should make sense. The greedier we are, the higher the probability of us going bankrupt. For instance, if we have \$1 thousand and want to gamble until we have \$2 thousand, the probability of us going bankrupt in the process is $1 - 1/2 = 0.5$. However, if we are greedier and change our target to \$10 thousand, the probability of us going bankrupt is now $1 - 1/10 = 0.9$.

3 The ABRACADABRA Theorem

With our knowledge of martingales and the OST, we are now ready to explore the simple and elegant approach to our coin-flip approach.

3.1 The Fair Casino

To set the stage, imagine that you work as a dealer at D'Casino. At D'Casino, there is only one game offered:

> *Each round, you, the dealer, flip a fair-coin. Gamblers go all-in, betting on the outcome of this coin-flip. If they win, they double their money, and they play again. If they lose, the casino takes everything and they go home.*
>
> *This repeats until a terminator (say THT) appears, at which point the casino closes and everybody goes home.*

Let Y_n be the outcome of the nth coin flip, and let the stopping time of the game be τ. Then $\{\tau = n\}$ is the event that the last three coin-flips come up THT. Mathematically,

$$\{\tau = n\} := \{Y_{n-2}Y_{n-1}Y_n = THT\}\,. \tag{3.1}$$

A group of gamblers, obsessed with the sequence THT, frequents D'Casino. Every flip, a new gambler from this group arrives with \$1 and plays the game, betting that the subsequent flips appear T, H, T in that order.

Let R_n and C_n be the total revenue earned and total cost incurred by the gamblers after the nth flip. Define also $X_n = R_n - C_n$ to be the combined wealth of the gamblers after the nth flip.

To illustrate how the game goes, suppose the coins come up $HTTHT$.

n	**Event**	R_n	C_n
1	Gambler #1 bets \$1 that $Y_1 = T$ and loses.	0	1
2	Gambler #2 bets \$1 that $Y_2 = T$ and wins. He bets \$2 that $Y_3 = H$ and loses. We record this as a net loss of \$1.	0	2
3	Gambler #3 bets \$1 that $Y_3 = T$ and wins. He bets \$2 that $Y_4 = H$ and wins. He bets \$4 that $Y_5 = T$ and wins. We record this as a gain of \$8, and a loss of \$1.	8	3
4	Gambler #4 bets \$1 that $Y_4 = T$ and loses.	8	4
5	Gambler #5 bets \$1 that the $Y_5 = T$ and wins. After the fifth coin-flip, the casino closes, so he cannot continue betting. We record this as a gain of \$2 and a loss of \$1.	10	5

We now make three key observations:

Proposition 12. $\mathbb{E}\left[L_{\{THT\}}\right] = \mathbb{E}[C_\tau]$.

Proof. The number of coin-flips, is equal to the number of gamblers. Also, because of the way we recorded losses, each gambler incurs a loss of exactly \$1. Hence, the number of coin-flips made, $L_{\{THT\}}$, is equal to the total cost incurred, C_τ. Taking expectations, $\mathbb{E}\left[L_{\{THT\}}\right] = \mathbb{E}[C_\tau]$. □

Proposition 13. $\mathbb{E}[C_\tau] = \mathbb{E}[R_\tau]$.

We will prove this result later.

Proposition 14. $\mathbb{E}[R_\tau] = R_\tau = 10$.

Proof. By the rules of the game, it is obvious that only the last three gamblers can earn money. Thus, the total revenue R_τ depends solely on the last three coin-flips. However, because the last three coin flips must always be THT (recall (3.1)), it follows that R_τ is a constant. Hence, $\mathbb{E}[R_\tau] = R_\tau = 10$. □

Chaining these three observations together yields

$$\mathbb{E}\left[L_{\{THT\}}\right] = \mathbb{E}[C] = \mathbb{E}[R] = R = 10,$$

which is indeed what we got using case-by-case analysis earlier.

3.2 Proof of Proposition 13

We now present a proof of Proposition 13. We begin by verifying that our wealth $\{X_n\}$ is a martingale with respect to the coin-flips $\{Y_n\}$.

Lemma 15. $\{X_n\}$ is a martingale with respect to $\{Y_n\}$.

Proof. It suffices to show that (1) $\mathbb{E}[X_n]$ is finite, and (2) $\mathbb{E}[X_{n+1} \mid Y_1, \ldots, Y_n] = X_n$.

Notice that X_n attains a maximum when all n gamblers win. Likewise, X_n attains a minimum when all n gamblers lose. Thus, $|X_n| \le n \cdot 2^n$, so $\mathbb{E}[X_n]$ must also be bounded and hence finite.

Let A_n be the total wealth of gamblers that have lost before the nth flip. Correspondingly, let B_n be the total wealth of gamblers that are still betting at the nth flip. Since A_n is constant, we have

$$\mathbb{E}[A_{n+1} \mid Y_1, \ldots, Y_n] = A_{n+1} = A_n.$$

Since the coin is fair and independent, and the gamblers bet double-or-nothing, we have

$$\mathbb{E}[B_{n+1} \mid Y_1, \ldots, Y_n] = \frac{1}{2}(2B_n) + \frac{1}{2}(0) = B_n.$$

Because $X_n = A_n + B_n$, it follows that

$$\mathbb{E}[X_{n+1} \mid Y_1, \ldots, Y_n] = \mathbb{E}[A_{n+1} \mid Y_1, \ldots, Y_n] + \mathbb{E}[B_{n+1} \mid Y_1, \ldots, Y_n] = A_n + B_n = X_n.$$

Thus, $\{X_n\}$ is a martingale with respect to $\{Y_n\}$. □

We now complete our proof of Proposition 13, which requires us to prove that our wealth martingale $\{X_n\}$ and stopping time τ obey the OST.

Proof of Proposition 13. We show that our wealth $\{X_n\}$ and our stopping time τ obeys the OST via the third scenario. That is, $\mathbb{E}[\tau]$ is finite and all increments of X_n are bounded.

It is easy to see that ΔX_n is bounded. A loose upper bound on ΔR_n is $3 \cdot 2^3$, which occurs when the last three gamblers bet on Y_τ and win. Similarly, an upper bound on ΔC_n is 3, which occurs when the last three gamblers bet on Y_τ and lose. It immediately follows that $\Delta X_n \leq \Delta R_n + \Delta C_n$ are bounded.

To see why $\mathbb{E}[\tau]$ is finite, consider the following game:

> *Suppose the terminator has length n. Each round, n coins are flipped. If these n coins matches the terminator (i.e. come up THT), we stop flipping. If not, we continue on with another round.*

Let $\{Y'_n\}$ be the outcome of the nth coin-flip, and let the stopping time for this game be τ'. Its stopping event is given by

$$\{\tau' = 3m\} := \{Y'_{3m-2} = T, Y'_{3m-1} = H, Y'_{3m} = T\},$$

where $m \in \mathbb{N}$. Quite clearly, $\mathbb{E}[\tau] \leq \mathbb{E}[\tau']$.

Let the random variable M be the number of rounds played under this game. Each round, the coin-flips have a $1/8$ chance of matching the terminator. Thus, M follows a geometric distribution with probability of success $p = 1/8$, whence $\mathbb{E}[M] = 1/p = 8$. Since a total of $3M$ coin-flips are made in this game, it follows that $0 < \mathbb{E}[\tau] \leq \mathbb{E}[\tau'] = 3 \cdot 8$, so $\mathbb{E}[\tau]$ is bounded and thus finite.

Hence, $\{X_n\}$ and τ obey the OST, which states $\mathbb{E}[X_\tau] = X_0 = 0$. Since $X_\tau = R_\tau - C_\tau$, we have $\mathbb{E}[R_\tau] = \mathbb{E}[C_\tau]$ as desired. □

3.3 Correlations and the ABRACADABRA Theorem

Since R_τ depends solely on the last few gamblers, we now have an easy way of calculating $\mathbb{E}[L_\mathcal{T}]$.

Example 16. To illustrate, consider yet again the example where $\mathcal{T} = \{THT\}$. Our goal is to calculate R_τ. To do so, we simply imagine that the terminator THT has already been flipped and then work backwards.

- The third-last gambler wins $\$2^3$, since he sees THT.
- The second-last gambler wins \$0, since he sees H and immediately loses.
- The last gambler wins $\$2^1$, since he sees T before the casino closes.

Hence, we have $\mathbb{E}[L_\mathcal{T}] = R_\tau = 2^3 + 2^1 = 10$.

We can abstract this process of calculating R_τ using the *correlation* of two strings.

Definition 17. Let X and Y be two words. The **correlation polynomial** of X and Y, denoted $\rho_z(X, Y)$, is a polynomial in z of maximum degree $|X|$.

The coefficients of $\rho_z(X, Y)$ are determined as follows: place Y under X so that its leftmost character is under the ith character of X (from the right). Then, if all pairs of characters in the overlapping segment are identical, the coefficient of z^i is 1, else it is 0.

Mathematically, using left- and right-slices,

$$\rho_z(X, Y) = \sum_{i=1}^{|X|} z^i \mathbf{1}\{R_i(X) = L_i(Y)\},$$

where the indicator function $\mathbf{1}(P)$ returns 1 if the statement P is true and 0 otherwise.

Example 18. Let $X = HTHTTH$ and $Y = HTTHT$. Then $\rho_z(X, Y) = z^4 + z^1$, as depicted below:

X:	H	T	H	T	T	H					
Y:	H	T	T	H	T						0
		H	T	T	H	T					0
			H	T	T	H	T				t^4
				H	T	T	H	T			0
					H	T	T	H	T		0
						H	T	T	H	T	t^1

Note that in general, $\rho_z(X, Y) \neq \rho_z(Y, X)$. For instance, using the same X and Y as the above example, we have $\rho_z(Y, X) = z^2$.

With this new terminology, one can easily see that $R_\mathcal{T} = \rho_2(t, t)$, where t is the terminator.

Example 19. Once again, suppose $\mathcal{T} = \{THT\}$. Notice that $\rho_z(THT, THT)$ is $z^3 + z^1$, as illustrated below:

X:	T	H	T			
Y:	T	H	T			z^3
		T	H	T		0
			T	H	T	z^1

Thus, $\rho_2(THT, THT) = 2^3 + 2^1 = 10$, which is precisely $R_\mathcal{T}$!

We can summarize this data using a matrix:

	T	H	T
THT	2^3		2^1

Unlike the case-by-case method we explored earlier, this method can easily be applied to terminators of longer lengths, as demonstrated in the following example.

Example 20. Let $\mathcal{T} = \{THHTHHTHH\}$. Computing correlations, we obtain the following matrix:

	T	H	H	T	H	H	T	H	H
$THHTHHTHH$	2^9			2^6			2^3		

Hence, $\mathbb{E}[L_{\mathcal{T}}] = 2^9 + 2^6 + 2^3 = 584$.

If we change the final character to a T, i.e. $\mathcal{T} = \{THHTHHTHT\}$, then our matrix becomes

	T	H	H	T	H	H	T	H	T
$THHTHHTHH$	2^9								2^1

Hence, $\mathbb{E}[L_{\mathcal{T}}] = 2^9 + 2^1 = 514$.

From the above examples, one can see that it is the "self-repetition" of the terminators that determines how long it takes to reach them. For instance, $THHTHHTHH$ self-repeats many times (at the sixth-last and third-last characters), while $THHTHHTHT$ only repeats itself at the last character.

Even if the alphabet $\mathcal{A}$ changes, the core idea remains the same:

Example 21. Consider the following problem:

> *A monkey types one random character on a typewriter every second. On average, how long would it take the monkey to type the word "ABRACADABRA"?*

In this context, our alphabet now contains 26 characters (A, B, C, etc.). To maintain the fairness of the casino, the payout for each win should now be 26 times the bet. Hence, the base of our correlation should be 26.

Comparing the correlation "ABRACADABRA" with itself, we see that our matrix is

	A	B	R	A	C	A	D	A	B	R	A
$ABRACADABRA$	26^{11}							26^4			26^1

The expected time taken is thus $26^{11} + 26^4 + 26$ seconds, or 116.4 million years.

More generally, we can state our result as follows:

Theorem 22 (ABRACADABRA Theorem)**.** Let $\mathcal{T} = \{t\}$ with alphabet $\mathcal{A}$. Then

$$\mathbb{E}[L_{\mathcal{T}}] = \rho_{|\mathcal{A}|}(t, t).$$

This result is known in the literature as the ABRACADABRA Theorem, named after the problem posed in Example 21.

4 The Generalized ABRACADABRA Theorem

We now turn our attention to solving the problem in its most general form. First, let us introduce one more piece of notation:

Definition 23. Let $t \in \mathcal{T}$. We define $[t]$ to be the set of all immediately terminated words w that end with t. Mathematically,

$$[t] = \left\{ w \in \mathcal{I}_{\mathcal{T}} : R_{|t|}(w) = t \right\}.$$

Example 24. If $\mathcal{T} = \{HHT, THH\}$, then

$$[HHT] = \{HHT, HHHT, HHHHT, \ldots\}$$

and

$$[THH] = \{THH, HTHH, TTHH, \ldots\}.$$

4.1 Two Terminators

To build our intuition, we first look at the case where we have two terminators, say $\mathcal{T} = \{THT, HTT\}$.

We once again return to our D'Casino thought experiment. Suppose you quit your job as a dealer at D'Casino. To fill in your absence, D'Casino management has introduced a slight modification to the coin-flipping game:

> *Each round, a fair coin is flipped, and a game is played between two parties, D (the "dealer") and G (the "gambler"):*
>
> *G goes all-in, betting on the outcome of the coin-flip. If G wins, D pays G, and they play again. If G loses, G pays H, and the two stop playing.*
>
> *This repeats until a terminator (either THT or HTT) appears, at which point the casino closes and everybody goes home.*

Let Y_n be the outcome of the nth coin-flip, and let the stopping time of the game be τ.

Suppose we have two groups of gamblers, Group 1 and Group 2, that frequent D'Casino. The gamblers in Group 1 are obsessed with the sequence THT, while those in Group 2 are obsessed with HTT. Every flip, a new gambler from each group arrives with \$1. The two gamblers then play two games simultaneously:

- In the first game, the Group 1 gambler is G and the Group 2 gambler is D. The Group 1 gambler bets that the next few coin-flips will be THT.

- In the second game, the Group 2 gambler is G and the Group 1 gambler is D. The Group 2 gambler bets that the next few coin-flips will be HTT.

Note that the two games share the same coin-flips.

Like before, we consider the revenues gained and costs incurred by each group. Suppose Group i plays as G. We define $R_n(i)$ to be the revenue earned by Group i at the nth flip, and $C_n(i)$ to be the cost incurred by Group i at the nth flip.

To illustrate how the games go, suppose the coin-flips come up $HHTHT$. We focus on Game 1 first. Recall that Group 1 is G, betting on THT, while Group 2 is D. As this is almost identical to what we have seen before, we keep the descriptions brief.

n	**Event**	$R_n(1)$	$C_n(1)$
1	Gambler #1 loses his first bet.	0	1
2	Gambler #2 loses his first bet.	0	2
3	Gambler #3 wins all three bets. He hence earns \$7 overall. For consistency, we record this as a gain of \$8 and a loss of \$1 for Group 1.	8	3
4	Gambler #4 loses his first bet.	8	4
5	Gambler #5 wins his first bet before the casino closes. Like before, we record this as a gain of \$2 and a loss of \$1 for Group 1.	10	5

We now do the same thing for Game 2. Here, Group 2 is G, betting on HTT, while Group 1 is D.

n	**Event**	$R_n(2)$	$C_n(2)$
1	Gambler #1 loses his second bet.	0	1
2	Gambler #2 loses his third bet.	0	2
3	Gambler #3 loses his first bet.	0	3
4	Gambler #4 wins his first and second bet, but the casino closes before his third. We record this as a gain of $4 and a loss of $1 for Group 2.	4	4
5	Gambler #5 loses his first bet.	4	5

Like before, we make three key observations:

- $\mathbb{E}[L_\mathcal{T}] = \mathbb{E}[C_\tau(i)]$ for $i = 1, 2$.
- $\mathbb{E}[C_\tau(i)] = \mathbb{E}[R_\tau(i)]$ for $i = 1, 2$.
- $R_\tau(i)$, where $i = 1, 2$, depends solely on the last three coin-flips.

The proofs of these three observations are almost identical to that of Propositions 12, 13 and 14.

From the first two observations, it is easy to see that

$$\mathbb{E}[L_\mathcal{T}] = \mathbb{E}[R_\tau(1)] = \mathbb{E}[R_\tau(2)]\,. \tag{4.1}$$

The last observation allows us to easily calculate $\mathbb{E}[R_\tau(1)]$ and $\mathbb{E}[R_\tau(2)]$, which we will now do.

- Suppose $W_\mathcal{T} \in [THT]$. Then

$$\begin{aligned}\mathbb{E}[R_\tau(1) \mid W_\mathcal{T} \in [THT]] &= \rho_2(THT, THT) = 2^3 + 2^1 = 10,\\ \mathbb{E}[R_\tau(2) \mid W_\mathcal{T} \in [THT]] &= \rho_2(THT, HTT) = 2^2 = 4.\end{aligned}$$

- Now suppose $W_\mathcal{T} \in [HTT]$. Then

$$\begin{aligned}\mathbb{E}[R_\tau(1) \mid W_\mathcal{T} \in [HTT]] &= \rho_2(HTT, THT) = 2^1 = 2,\\ \mathbb{E}[R_\tau(2) \mid W_\mathcal{T} \in [HTT]] &= \rho_2(HTT, HTT) = 2^3 = 8.\end{aligned}$$

By the law of total expectation, it follows that

$$\begin{aligned}\mathbb{E}[R_\tau(1)] &= 10\,\mathbb{P}[W_\mathcal{T} \in [THT]] + 2\,\mathbb{P}[W_\mathcal{T} \in [HTT]]\,,\\ \mathbb{E}[R_\tau(2)] &= 4\,\mathbb{P}[W_\mathcal{T} \in [THT]] + 8\,\mathbb{P}[W_\mathcal{T} \in [HTT]]\,,\end{aligned}$$

By (4.1), the two are equal, giving us the equation

$$10\,\mathbb{P}[W_\mathcal{T} \in [THT]] + 2\,\mathbb{P}[W_\mathcal{T} \in [HTT]] = 4\,\mathbb{P}[W_\mathcal{T} \in [THT]] + 8\,\mathbb{P}[W_\mathcal{T} \in [HTT]]\,.$$

Further, by the law of total probability,

$$\mathbb{P}[W_\mathcal{T} \in [THT]] + \mathbb{P}[W_\mathcal{T} \in [HTT]] = 1.$$

This gives us a system of two linear equations in two unknowns, which we can easily solve:

$$\mathbb{P}[W_\mathcal{T} \in [THT]] = \mathbb{P}[W_\mathcal{T} \in [HTT]] = \frac{1}{2}.$$

Plugging these values back into (4.1), we finally obtain the expected length:

$$\mathbb{E}[L_\mathcal{T}] = \mathbb{E}[R_\tau(1)] = 10\,\mathbb{P}[W_\mathcal{T} \in [THT]] + 2\,\mathbb{P}[W_\mathcal{T} \in [HTT]] = 10\left(\frac{1}{2}\right) + 2\left(\frac{1}{2}\right) = 6.$$

4.2 Three Terminators

Let us now go one step further and consider the case where we have three terminators, say $\mathcal{T} = \{THT, HTT, HHH\}$.

The set-up is going to remain mostly the same. The only difference is that we now have three groups, each betting on one of the three terminators. Let Group 1 bet on THT, Group 2 on HTT and Group 3 on HHH. These three teams will now play three games:

- In the first game, the Group 1 gambler is G and the Group 2 gambler is D.
- In the second game, the Group 2 gambler is G and the Group 3 gambler is D.
- In the third game, the Group 3 gambler is G and the Group 1 gambler is D.

Observe the cyclical nature of these match-ups. Also, observe that each group plays exactly one game as G, and another as D.

Like before, we define $R_n(i)$ and $C_n(i)$ to be the total revenue earned and total cost incurred by Group i at the nth flip of the game where they play as G.

We make the same three observations:

- $\mathbb{E}[L_{\mathcal{T}}] = \mathbb{E}[C_\tau(i)]$ for $i = 1, 2, 3$.
- $\mathbb{E}[C_\tau(i)] = \mathbb{E}[R_\tau(i)]$ for $i = 1, 2, 3$.
- $R_\tau(i)$, where $i = 1, 2, 3$, depends solely on the last three coin-flips.

From the first two observations, we have

$$\mathbb{E}[L_{\mathcal{T}}] = \mathbb{E}[R_\tau(1)] = \mathbb{E}[R_\tau(2)] = \mathbb{E}[R_\tau(3)]\,. \tag{4.2}$$

Using the last observation, we can easily calculate $\mathbb{E}[R_\tau(i)]$ for $i = 1, 2, 3$:

- Suppose $W_{\mathcal{T}} \in [THT]$. Then

$$\begin{aligned}
\mathbb{E}[R_\tau(1) \mid W_{\mathcal{T}} \in [THT]] &= \rho_2(THT, THT) = 2^3 + 2^1 = 10,\\
\mathbb{E}[R_\tau(2) \mid W_{\mathcal{T}} \in [THT]] &= \rho_2(THT, HTT) = 2^2 = 4,\\
\mathbb{E}[R_\tau(3) \mid W_{\mathcal{T}} \in [THT]] &= \rho_2(THT, HHH) = 0.
\end{aligned}$$

- Suppose $W_{\mathcal{T}} \in [HTT]$. Then

$$\begin{aligned}
\mathbb{E}[R_\tau(1) \mid W_{\mathcal{T}} \in [HTT]] &= \rho_2(HTT, THT) = 2^1 = 2,\\
\mathbb{E}[R_\tau(2) \mid W_{\mathcal{T}} \in [HTT]] &= \rho_2(HTT, HTT) = 2^3 = 8,\\
\mathbb{E}[R_\tau(3) \mid W_{\mathcal{T}} \in [HTT]] &= \rho_2(HTT, HHH) = 0.
\end{aligned}$$

- Lastly, suppose $W_{\mathcal{T}} \in [HHH]$. Then

$$\begin{aligned}
\mathbb{E}[R_\tau(1) \mid W_{\mathcal{T}} \in [HHH]] &= \rho_2(HHH, THT) = 0,\\
\mathbb{E}[R_\tau(2) \mid W_{\mathcal{T}} \in [HHH]] &= \rho_2(HHH, HTT) = 0,\\
\mathbb{E}[R_\tau(3) \mid W_{\mathcal{T}} \in [HHH]] &= \rho_2(HHH, HHH) = 2^3 + 2^2 + 2^1 = 14.
\end{aligned}$$

By the law of total expectation, it follows that

$$\begin{aligned}
\mathbb{E}[R_\tau(1)] &= 10\,\mathbb{P}[W_{\mathcal{T}} \in [THT]] + 2\,\mathbb{P}[W_{\mathcal{T}} \in [HTT]] + 0\,\mathbb{P}[W_{\mathcal{T}} \in [HHH]]\,,\\
\mathbb{E}[R_\tau(2)] &= 4\,\mathbb{P}[W_{\mathcal{T}} \in [THT]] + 8\,\mathbb{P}[W_{\mathcal{T}} \in [HTT]] + 0\,\mathbb{P}[W_{\mathcal{T}} \in [HHH]]\,,\\
\mathbb{E}[R_\tau(3)] &= 0\,\mathbb{P}[W_{\mathcal{T}} \in [THT]] + 0\,\mathbb{P}[W_{\mathcal{T}} \in [HTT]] + 14\,\mathbb{P}[W_{\mathcal{T}} \in [HHH]]\,.
\end{aligned}$$

By (4.2), these three expressions are equal. Further, the probabilities sum up to 1. This gives us a system of linear equations, which we can easily solve. One can verify that the probabilities are

$$\mathbb{P}[W_{\mathcal{T}} \in [THT]] = \frac{11}{28}, \quad \mathbb{P}[W_{\mathcal{T}} \in [HTT]] = \frac{8}{28}, \quad \mathbb{P}[W_{\mathcal{T}} \in [HHH]] = \frac{9}{28}.$$

Substituting this back into (4.2), we obtain

$$\mathbb{E}[L_{\mathcal{T}}] = \mathbb{E}[R_{\mathcal{T}}(3)] = 14\left(\frac{9}{28}\right) = \frac{9}{2}.$$

4.3 Correlation Matrices and Probability Vectors

Before we move on to the most general case, we introduce a more concise way of describing our procedure.

Consider the system of equations we just derived:

$$\begin{aligned}
\mathbb{E}[R_{\mathcal{T}}(1)] &= 10\,\mathbb{P}[W_{\mathcal{T}} \in [THT]] + 2\,\mathbb{P}[W_{\mathcal{T}} \in [HTT]] + \ \ 0\,\mathbb{P}[W_{\mathcal{T}} \in [HHH]]\,,\\
\mathbb{E}[R_{\mathcal{T}}(2)] &= \ \ 4\,\mathbb{P}[W_{\mathcal{T}} \in [THT]] + 8\,\mathbb{P}[W_{\mathcal{T}} \in [HTT]] + \ \ 0\,\mathbb{P}[W_{\mathcal{T}} \in [HHH]]\,,\\
\mathbb{E}[R_{\mathcal{T}}(3)] &= \ \ 0\,\mathbb{P}[W_{\mathcal{T}} \in [THT]] + 0\,\mathbb{P}[W_{\mathcal{T}} \in [HTT]] + 14\,\mathbb{P}[W_{\mathcal{T}} \in [HHH]]\,.
\end{aligned}$$

One might observe that this can be expressed much more simply as a matrix equation! Indeed, we have

$$\begin{pmatrix} \mathbb{E}[R_{\mathcal{T}}(1)] \\ \mathbb{E}[R_{\mathcal{T}}(2)] \\ \mathbb{E}[R_{\mathcal{T}}(3)] \end{pmatrix} = \begin{pmatrix} 10 & 2 & 0 \\ 4 & 8 & 0 \\ 0 & 0 & 14 \end{pmatrix} \begin{pmatrix} \mathbb{P}[W_{\mathcal{T}} \in [THT]] \\ \mathbb{P}[W_{\mathcal{T}} \in [HTT]] \\ \mathbb{P}[W_{\mathcal{T}} \in [HHH]] \end{pmatrix}.$$

We call the matrix on the right the *correlation matrix of $\mathcal{T}$*, since its values are the correlations between pairs of terminators. We also call the vector on the right the *probability vector of $\mathcal{T}$*.

Definition 25. Let $\mathcal{T} = \{t_1, \ldots, t_n\}$ with alphabet $\mathcal{A}$. The **correlation matrix** of $\mathcal{T}$ is an $n \times n$ matrix $\mathbf{M}_{\mathcal{T}} = (m_{ij})$, where

$$m_{ij} = \rho_{|\mathcal{A}|}(t_j, t_i).$$

The **probability vector** of $\mathcal{T}$, denoted $\mathbf{p}_{\mathcal{T}}$, is defined as

$$\mathbf{p}_{\mathcal{T}} = \begin{pmatrix} \mathbb{P}[W_{\mathcal{T}} \in [t_1]] \\ \vdots \\ \mathbb{P}[W_{\mathcal{T}} \in [t_n]] \end{pmatrix}.$$

With these two objects in mind, let us find a general expression for $\mathbb{E}[L_{\mathcal{T}}]$.

4.4 The Generalized ABRACADABRA Theorem

Suppose we now have n terminators, say $\mathcal{T} = \{t_1, \ldots, t_n\}$. Note that these terminators may have different lengths.

We will reuse our set-up again. This time, however, we will have n groups, each betting on one terminator. These groups will play a total of n games:

- Group 1 as G, with Group 2 as D,
- Group 2 as G, with Group 3 as D,

- Group 3 as G, with Group 4 as D,
- $\cdots$,
- Group $n-1$ as G, with Group n as D,
- Group n as G, with Group 1 as D.

We again make three observations:

- $\mathbb{E}[L_{\mathcal{T}}] = \mathbb{E}[C_\tau(i)]$ for $i = 1, \ldots, n$.
- $\mathbb{E}[C_\tau(i)] = \mathbb{E}[R_\tau(i)]$ for $i = 1, \ldots, n$.
- $R_\tau(i)$, where $i = 1, \ldots, n$, depends solely on the last $|t_i|$ coin-flips.

From the last observation, we can express these expected revenues as the product of $\mathbf{M}_{\mathcal{T}}$ and $\mathbf{p}_{\mathcal{T}}$:

$$\begin{pmatrix} \mathbb{E}[R_\tau(1)] \\ \vdots \\ \mathbb{E}[R_\tau(n)] \end{pmatrix} = \mathbf{M}_{\mathcal{T}} \mathbf{p}_{\mathcal{T}}.$$

From the first two observations, we have

$$\mathbb{E}[L_{\mathcal{T}}] = \mathbb{E}[R_\tau(1)] = \mathbb{E}[R_\tau(2)] = \mathbb{E}[R_\tau(3)] = \cdots = \mathbb{E}[R_\tau(n)]\,.$$

Thus, our matrix equation becomes

$$\mathbf{M}_{\mathcal{T}} \mathbf{p}_{\mathcal{T}} = \mathbb{E}[L_{\mathcal{T}}]\,\mathbf{1}_n \implies \mathbf{p}_{\mathcal{T}} = \mathbb{E}[L_{\mathcal{T}}]\,\mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n,$$

where $\mathbf{1}_n$ is the vector consisting of n 1's. Since the probabilities must sum to 1, we must have $\mathbf{1}_n^\intercal \mathbf{p}_{\mathcal{T}} = 1$. Thus,

$$1 = \mathbb{E}[L_{\mathcal{T}}]\,\mathbf{1}_n^\intercal \mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n \implies \mathbb{E}[L_{\mathcal{T}}] = \frac{1}{\mathbf{1}_n^\intercal \mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n}.$$

We also get an expression for the probability vector for free:

$$\mathbf{p}_{\mathcal{T}} = \frac{1}{\mathbf{1}_n^\intercal \mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n}\mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n.$$

This result is the generalized ABRACADABRA Theorem.

Theorem 26 (Generalized ABRACADABRA Theorem)**.** Let $\mathcal{T} = \{t_1, \ldots, t_n\}$. Then

$$\mathbb{E}[L_{\mathcal{T}}] = \frac{1}{\mathbf{1}_n^\intercal \mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n} \quad \text{and} \quad \mathbf{p}_{\mathcal{T}} = \frac{1}{\mathbf{1}_n^\intercal \mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n}\mathbf{M}_{\mathcal{T}}^{-1}\mathbf{1}_n.$$

Remark. $\mathbb{E}[L_{\mathcal{T}}]$ is the sum of the all entries of $\mathbf{M}_{\mathcal{T}}^{-1}$.

With this theorem, we can easily calculate $\mathbb{E}[L_{\mathcal{T}}]$ and the probabilities $\mathbf{p}_{\mathcal{T}}$ using software. Let us now look at some examples of this theorem in action.

Example 27. Let $\mathcal{T} = \{THT, HTT\}$. Our correlation matrix is given by

	T	H	T	H	T	T
THT	2^3		2^1			2^1
HTT		2^2		2^3		

The inverse of $\mathbf{M}_\mathcal{T}$ is

$$\mathbf{M}_\mathcal{T}^{-1} = \frac{1}{36}\begin{pmatrix} 4 & -1 \\ -2 & 5 \end{pmatrix}.$$

Using the generalized ABRACADABRA theorem, we have

$$\mathbb{E}[L_\mathcal{T}] = \frac{1}{\frac{1}{36}(4-1-2+5)} = 6.$$

As mentioned before, the terminators need not be of the same length:

Example 28. Let $\mathcal{T} = \{TT, THT\}$. Our correlation matrix is given by

	T	T	T	H	T
TT	2^2	2^1			2^1
THT		2^1	2^3		2^1

One can calculate the inverse of $\mathbf{M}_\mathcal{T}$ to be

$$\mathbf{M}_\mathcal{T}^{-1} = \frac{1}{28}\begin{pmatrix} 5 & -1 \\ -1 & 3 \end{pmatrix}.$$

Thus, by the generalized ABRACADABRA theorem,

$$\mathbb{E}[L_\mathcal{T}] = \frac{1}{\frac{1}{28}(5-1-1+3)} = \frac{14}{3}.$$

5 Further Questions

In this workshop, we managed to derive closed forms for $\mathbb{E}[L_\mathcal{T}]$ and $\mathbb{P}[W_\mathcal{T} \in [t]]$. There are, however, many more questions we can ask about this game:

- Is there a closed form for $\mathbb{E}[L_\mathcal{T} \mid W_\mathcal{T} \in [t]]$?
- If $W_\mathcal{T} \in [t]$, what is the distribution of $L_\mathcal{T}$? How many words in $[t]$ have length n? Equivalently, given that $L_\mathcal{T} = n$, what is the probability that $W_\mathcal{T} \in [t]$?
- Given n terminators, each at most length k, what is the minimum and maximum value of $\mathbb{E}[L_\mathcal{T}]$?
- What is the significance of $\mathbf{M}_\mathcal{T}^{-1}$? What does it mean to "invert" a correlation matrix?

Slightly modifying our original coin-flip problem also opens up a whole can of worms:

- What if we stopped flipping the coin once we see all terminators?
- What if we allowed up to k appearances of a single terminator?
- If we flip a fair coin n times, what is the probability that a terminator t appears?

References

[1] L.J Guibas and A.M Odlyzko. String overlaps, pattern matching, and nontransitive games. *Journal of Combinatorial Theory, Series A*, 30(2):183–208, 1981.

[2] Vikas Jain. Why is the expected number coin tosses to get HTH is 10? Mathematics Stack Exchange.

[3] Aaron Montgomery. Slick way to solve P(pattern A appears before pattern B) type of problems in a coin toss process. Mathematics Stack Exchange.

[4] Scott M. LaLonde. The Martingale Stopping Theorem. https://math.dartmouth.edu/~pw/math100w13/lalonde.pdf, 2013.

[5] Adam Lelkes. Martingales and the Optional Stopping Theorem. https://doi.org/10.59350/hgcyv-5x530, 2014.

[6] Mihai Nica. A Coin Flip Paradox and the ABRACADABRA Theorem for infinite monkeys: How long does it take? https://www.youtube.com/watch?v=t8xqMxlZz9Y, 2022.

[7] John Sylvester, Nicolás Rivera, Luca Zanetti, and Thomas Sauerwald. Lecture 8: The Optional Stopping Theorem. https://www.cl.cam.ac.uk/teaching/1819/Probablty/materials/Lecture8.pdf, 2019.