

Exploring Variations of Penney's Game

William Dracopoulos & Derrick Chung
John Abbott College

May 13, 2025

Contents

I	Abstract	4
II	Introduction	4
III	Showing that Penney's Game is unintuitive, and providing some hope that it can be understood	5
IV	Deriving a formula for win probabilities with two players	7
1	The basic variation of Penney's Game	7
2	Allowing for any random outcome generator with any known probability distribution	11
3	Extending to games with sequences of different lengths	12
V	Deriving a formula for win probabilities in games with any number of players	13
VI	Two-player games with equal sequence lengths and even flip outcome likelihoods	14
4	Proving that when sequence S appears, Player S will always win more money than their opponent in the betting game	15
5	Determining how changes of R values affect $\mathbb{P}(A)$	16
6	Proving that the greatest win probability difference occurs when $baa\dots a$ is played against $aaa\dots a$ for $b \neq a$	18
7	Determining the bounds of $\mathbb{P}(A)$ and $\mathbb{P}(B)$	19
8	Using the probability bounds to study extreme games.	19
8.1	Proving that the game tends to be fair with infinite possible flip outcomes but finite sequence length	19
8.2	Determining the bounds for games with infinitely long sequences but finite possible flip outcomes	20
8.3	Proving that the game tends to be fair with infinite possible flip outcomes and infinitely long sequences.	20
VII	Two-player games with different sequence lengths and even flip outcome likelihoods	20

9	Determining when the shorter sequence always has an advantage in games with two players and even flip outcome probabilities.	20
VIII	Verifying the formula with simulations	21
IX	Creating a tool to aid in research of Penney's Game	22
X	Physically modelling unfair coin flips and dice rolls	22
XI	Conclusion	22
XII	References	22

Part I

Abstract

Penney's Game is an example of where our intuitive understanding of probability completely fails; it's a seemingly fair game with a surprisingly complicated strategy. The goal of this research is to explore Penney's Game using new methods and in areas that it has not been explored before. By studying a similar game that is fair, we create a formula for the probability of each player winning Penney's Game depending on the number of players, the sequences chosen by each player, the length of each sequence, the number of possible outcomes of the "flip" (coin flip or dice roll or random number generation), and the probability distribution of the outcomes of this flip. This formula was verified by computer simulations and then packaged into an online calculator. The calculator then revealed patterns which were analyzed using the formula. Among these, it was proven that games tend towards being fair with an infinite number of possible outcomes from the "flip" but not infinite sequence length, and that the shorter sequence always has an advantage, except in some specific and known cases. The work done in this paper is not exhaustive, there is still plenty to explore and discover, and we leave behind tools to help future researchers to do this.

Part II

Introduction

In its basic variation, Penney's Game is simple game invented by Walter Penney involving two players and one coin. Each player predicts which sequence of flips will appear first as the coin is being flipped. For example, Alice predicts that HHT (heads, then heads, then tails) will appear first, while Bob predicts that THT (tails, then heads, then tails) will appear first. The players flip the coin until the last three coin flips match with the sequence predicted by one of the players. That player wins. Note that this does not mean the coin is flipped three times, a match is checked for, and then it's flipped three times again. Instead, after every coin flip, the last three flips are checked to see if they match with a player's prediction.

The game might seem fair, but it really is not. In that example, Alice actually has a 62.5% chance of beating Bob. In fact, for any sequence that one player predicts, there is another that will, beat it on average (Penney, 1969). A formula for finding the exact win probabilities for any sequence in this basic variation of the game was first discovered by the mathematician John Conway, but he never supplied a proof of his "leading number algorithm".

Since Conway, many other researchers have developed other algorithms and ways of thinking about the problem. One way is using Markov chains and martingale arguments, like in Gerber and Li (1981)'s research. However, this paper uses a much simpler way of understanding the game and deriving a formula for win probabilities. The general idea,

which will be explained in depth during the next part, is to think about another game instead, which this time is always fair. This new game, which I'll call the betting game, contains the same probabilities as Penney's Game but allows for a clear and powerful starting point due to its fairness. The idea of the betting game comes simultaneously from two places: a research paper which tackles Penney's game from the perspective of no-arbitrage (Miller, 2019), and Chung (2024)'s talk which mentions describes it in the context of the Optional Stopping Theorem. The latter will be the same approach used in this paper. We extend the betting game analogy further than ever before to allow us to determine probabilities in many variations of Penney's Game.

A formula, (9), will be derived which can calculate the win probabilities in games with:

1. Any number of players
2. Players predicting more or less than the next three flips, and players predicting a different number of flips to each other
3. Any kind of coin, die, or random number generator being "flipped"
4. Unfair coins, loaded dice, and random number generators with any probability distribution

We will also verify this formula with simulations and analyze it to discover new things about Penney's Game in some of the variations covered by the formula.

Part III

Showing that Penney's Game is unintuitive, and providing some hope that it can be understood

The game was introduced as seeming fair but actually being biased. However, sometimes it seems like one player should be advantaged, while in reality the game is completely fair.

Imagine sitting at a table and flipping a coin until there are three consecutive heads. You should find that, on average, HHH appears after the 14th coin flip. Now you do the same and count how many flips it takes to get HHT. This time, you count that HHT appears after only the 8th coin flip on average. We went through this process and collected data for each sequence in the basic variation, which is shown in Figure ??.

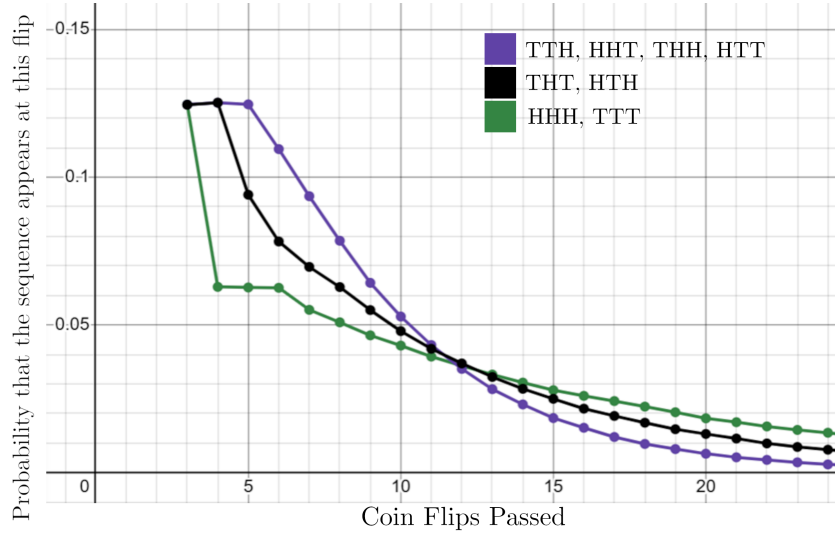


Figure 1: Probability of a sequence appearing for the first time as a function of the number of coin flips. The data was obtained from a simulation of 1 000 000 repetitions. The expected wait time (number of flips until the sequence appears) for the purple, black, and green lines are calculated to be approximately 8 flips, 10 flips, and 14 flips respectively.

fig:probGraph

An exact function for the green line was found by my friend Fenghua Li to be

$$\mathbb{P}_n = \frac{1}{16}A(n-4)$$

$$A(m) = \frac{1}{2^m} \sum_{s=0}^{m-\lfloor \frac{m}{3} \rfloor} \sum_{L=\lceil \frac{s}{2} \rceil}^s \binom{m+1-s}{L} \binom{L}{s-L}$$

This gives the probability of the sequence appearing, \mathbb{P} , for any number of previous coin flips n .

This might lead you to believe that HHT would beat HHH in Penney's Game, since it would appear earlier, but in fact HHT vs. HHH would be a fair game. Figure ?? does not show the whole picture. It does not take into account how the sequences interact with each other. The difference is that the expected wait time comes from one player flipping one coin, but in Penney's Game it's two players checking for their sequence on the same coin. Imagine the last two coin flips were heads. There is now a 50:50 chance that the third flip will land on heads and tails. Therefore it is equally likely that HHT and HHH appears, so the game is fair.

The situation is very different if you switch it around to be THH vs. HHH however. This will actually give the greatest probability difference in the basic variation of the game: the player predicting HHH will only have a 12.5% (or $\frac{1}{8}$) chance of winning. This is because as soon as a single coin flip lands on tails, HHH has no chance of appearing before THH: THH will always get to two heads in a row before HHH gets to three in a row. This leaves only one way for HHH to win: if the first three coin flips land on heads, which is a $\frac{1}{2^3} = \frac{1}{8}$ chance.

We have shown that Penney's Game is unintuitive but not completely impossible to understand. Now we can move on to analyzing it through the betting game to develop a general formula for win probabilities.

Part IV

Deriving a formula for win probabilities with two players

1 The basic variation of Penney's Game

Imagine two players sitting at a table with one coin between them. Each player picks a sequence and they begin to flip the coin until the last flips match with one of the player's sequences. Up until now, the game has been identical to Penney's Game, and that is by design. Now it begins to differ. This is how the betting game works:

Right before each coin flip, both players bet on their chosen sequence appearing in the next flips. Note that this means each player will likely have multiple bets going on at the same time, since a new sequence of bets starts for each flip, and each bet might last multiple flips. The players are betting against each other: if one wins \$2, it's the other player that gives it to them. Although the game ends when one player's sequence appears in the flips, it does not necessarily mean that the player wins. There is no winner nor loser, just an exchange of money.

We will explain using an example: Let's say Alice chooses HHT and Bob chooses THT. We'll start by following the sequence of bets that starts from the first flip. Before the coin is flipped for the first time, Alice bets that the flips will follow her sequence, so she bets that the first will be H, the second H, and the third T. Bob simultaneously bets that the first flip will be T, the second H, and the third T. Both players start by betting \$1 on this. If they win, they will have \$2 and they will use that to bet on the second flip, and so on, doubling each time. If they lose, that \$1 is gone. By random chance, the coin lands on heads. Alice now has \$2 and Bob has lost the bet: his prediction has failed and he lost the \$1 to Alice. He also has to pay \$1 to Alice for her win. So she actually has \$3 and he has -\$1. But Alice continues, she bets the \$2 from her win on the next flip being heads. They flip the coin again and it lands on heads. Alice now has \$5 after winning another \$2 from Bob, and Bob has -\$3. She bets the \$4 from her correct bets so far on tails to finish her sequence of HHT. The third flip lands on heads instead. Alice loses that \$4 to Bob, so they start back where they began: each has \$1.

So this whole thing was useless, and there can be any number of useless sequences of bets like this until the game ends. But remember that each time a coin was flipped, each player starts new bets on their chosen sequence of flips.

This game actually ends one flip later, with Alice's sequence appearing in the coin flips. We will illustrate this game in a table. The top row represents the coin flips that actually

occur. All other rows represent a new sequence of bets that the player makes. Notice that the row labelled 1 shows the sequence of bets that we described above: the sequence that starts at the first coin flip. But there were three more sequences of bets until the game ended.

To simplify, we can think of the bets slightly differently. We can ignore the constant exchange of money between players, and have them only give each other money when the game ends.

Alice:

		H	H	H	T
1		H	H	T	
2			H	H	T
3				H	H
4					H

We colour the successful bets in green and the unsuccessful ones in red. Alice has a single sequence of bets that was successful: row 2. From it, she is at \$8 because her three bets in a row were successful, also meaning that Alice made the game end. From the other sequences, she has nothing. So we can say that Alice has \$8 and therefore Bob owes her \$8. But during this time, what was Bob doing?

Bob:

		H	H	H	T
1		T	H	T	
2			T	H	T
3				T	H
4					T

In all rows except 4, Bob loses the sequence of bets on the first one. So all others after it are not coloured, indicating that they don't have an opportunity to happen at all. He won the first bet in row 4, but his bets were cut off from Alice winning the game, meaning that Bob gets to keep the \$2 he won from it.

In total, Bob owes Alice \$8, but Alice owes Bob \$2. After the exchange, Alice has \$6 and Bob has -\$6.

If either player had received money from another sequence of bets which started at another flip, it would get added to the total.

It is very important to notice that the final share of money is not unique to this game in particular. In any game where Alice's sequence of HHT appears before Bob's sequence of THT, Alice will end up getting \$6 from Bob. This is because all of the flips before the last three do not matter: the prior bets of both Alice and Bob will fail eventually. The only two ways for a player to make money from a sequence of bets are:

1. If the bets go to completion. In other words, if the three bets in a row succeed. This will end the game.
2. If a sequence of bets gets cut off by the game ending.

We can ignore all sequences of bets before the last three because we know they won't give anyone any money. They can't return money in the first way because then the game would be over already. And we cover all bets that might return money in the second way by including the last two sequences of bets.

We can now introduce a better version of the table, which we will use throughout the paper.

	H	H	T	Gain	
$R_A(A):$	H	H	T	$2^3 = 8$	$= 8$
		H	H	0	
			H	0	

It now only features the important rows: the last 3. And it shows that Alice's (denoted by the A within the parentheses) "return" (R) when her own sequence (denoted by the A in the subscript) appears as the coin is flipped will be \$8.

The same table can be made for Bob's return when Alice's sequence appears:

	H	H	T	Gain	
$R_A(B):$	H	H	T	0	$= 2$
		H	H	0	
			T	2	

We have studied sufficiently what happens when Alice's sequence appears. Now let's focus on when Bob's sequence appears by making the tables:

	T	H	T	Gain	
$R_B(A):$	H	H	T	0	$= 0$
		H	H	0	
			H	0	

So Alice makes \$0 from Bob when his sequence appears in the coin flips.

	T	H	T	Gain	
$R_B(B):$	T	H	T	8	$= 10$
		H	H	0	
			T	2	

And Bob makes \$10 from Alice when his sequence appears in the coin flips.

In total, when Bob's sequence appears, Alice gives him \$10.

Note that something interesting is happening here. When Alice's sequence appears, Alice makes \$6 from Bob, but when Bob's sequence appears, he makes \$10 from Alice. This might make it seem like Bob has an advantage in this game, but there's a problem with that.

Here is the point of all this: All Alice and Bob are doing is betting on a fair coin, and either doubling or losing the money they put in. Each individual bet is fair. But because of the linearity of expectation, it doesn't matter if you make a fair bet once or 100 times, on average, you should expect to neither lose nor gain money. Your expected profit should be 0.

But it seems like Bob is doing better than Alice, and that doesn't make sense *assuming* that Bob's sequence and Alice's sequence have an equal chance of appearing. So let's simply assume that they don't. We'll take $\mathbb{P}(A)$ to mean the probability that Alice's sequence appears and ends the game, and $\mathbb{P}(B)$ to mean the probability that Bob's sequence appears and ends the game.

Let's look at it from Alice's perspective:

$$(\$6) \cdot \mathbb{P}(A) + (-\$10) \cdot \mathbb{P}(B) = \$0$$

In words, Alice's net profit when her sequence appears (\$6) weighed by the chance that her sequence appears ($\mathbb{P}(A)$), added to Alice's net profit when Bob's sequence appears (-\$10) weighed by the chance that his sequence appears ($\mathbb{P}(B)$) is her expected profit, which for reasons explained above must be exactly \$0.

Now, of course, either Alice's or Bob's sequence must appear, so

$$\mathbb{P}(A) + \mathbb{P}(B) = 1 \tag{1}$$

We now have two equations and two unknowns, so $\mathbb{P}(A)$ and $\mathbb{P}(B)$ can be found. After some algebra, $\mathbb{P}(A) = 62.5\%$ and $\mathbb{P}(B) = 37.5\%$.

Importantly, this means that if Alice and Bob were playing Penney's Game with the sequences HHT and THT, Alice would have a 62.5% chance of winning and Bob would have a 37.5% chance of winning.

Generalizing beyond this example, we can define the net profit of Player X against Player Y when sequence Z appears (Z being either Player X or Player Y's sequence) as:

$$N_Z(X) = R_Z(X) - R_Z(Y) \tag{2}$$

Next, we can solve for $\mathbb{P}(A)$ and $\mathbb{P}(B)$ generally. From the perspective of Player A:

$$N_A(A) \cdot \mathbb{P}(A) + N_B(A) \cdot \mathbb{P}(B) = 0$$

Solving this with the help of (1), we get

$$\mathbb{P}(A) = \frac{-N_B(A)}{N_A(A) - N_B(A)} = \frac{R_B(B) - R_B(A)}{R_A(A) - R_A(B) + R_B(B) - R_B(A)} \quad (3)$$

And the formula for $\mathbb{P}(B)$ is exactly the same but with all A and B's swapped. This is exactly the same result as found by Miller and it is equivalent to Conway's leading number algorithm (Miller, 2019).

The last and most complicated thing to do is express the return values R mathematically. The return in \$ for sequence Y if the game ends by sequence X appearing can be given by this equation, which simply describes the process done in the tables above:

$$R_X(Y) = \sum_{i=1}^3 \prod_{j=1}^i 2_{[Y_j=X_{3-i+j}]} \quad (4)$$

Where $2_{[Y_j=X_{3-i+j}]}$ is defined as:

$$2_{[Y_j=X_{3-i+j}]} = \begin{cases} 2 & \text{if } Y_j = X_{3-i+j} \\ 0 & \text{if } Y_j \neq X_{3-i+j} \end{cases}$$

And the subscript of a sequence is its index. For example, in X : HTT, $X_1 = H$; $X_2 = T$; $X_3 = T$.

This equation is a specific case of Miller's formula for R .

2 Allowing for any random outcome generator with any known probability distribution

We have described Penney's Game as being played with a fair coin, but this is not necessary. We can extend the formula to account for any random outcome generator with any known probability distribution. For example, a random outcome generator can be a loaded 20-sided die or a coin flipped in a certain way so that it's more likely to land on heads¹. From now on, we will generalize all of these random outcome generators by imagining them as a random number generator that spits out integers $[0, 1, \dots, n-1]$ for n possible outcomes of the "flip". For example, instead of writing "heads" / "H" and "tails" / "T" for a coin, we will write "0" and "1" because $n = 2$. And a 6-sided die will be thought of as a generator outputting numbers 0, 1, 2, 3, 4, 5.

To keep the betting game a valid analogy to Penney's Game, we need to keep the condition that the game must be fair, no matter which outcome generator is being used. This has a simple fix.

With a fair two-sided coin. We said that each bet is a double-or-nothing. This is because for a bet in which you put $\$X$, there is a 50% chance of winning $\$X$ and a 50% chance of losing $\$X$. Mathematically, the expected profit is 0, so

$$W \cdot \mathbb{P}(W) + L \cdot \mathbb{P}(L) = 0$$

¹We provide a way to physically model loaded die and unfair coins in Part X.

Meaning that the net profit from winning, W , weighted by the probability of winning, $\mathbb{P}(W)$, added to the net profit from losing, L , weighted by the probability of losing, $\mathbb{P}(L)$, should be zero.

Since this bet must either be won or lost, $\mathbb{P}(W) + \mathbb{P}(L) = 1$, so $\mathbb{P}(L) = 1 - \mathbb{P}(W)$.

For example, in the case of a fair coin,

$$X \cdot \frac{1}{2} + (-X) \cdot \frac{1}{2} = 0$$

Now we can generalize for any $\mathbb{P}(W)$. We will say that the outcome of any bet will either be losing what you put into it, $L = -X$, or winning some multiple of what you put into it, $W = \lambda X$. We want to see which multiple λ to use as a function of $\mathbb{P}(W)$.

$$(\lambda X) \cdot \mathbb{P}(W) + (-X) \cdot (1 - \mathbb{P}(W)) = 0$$

$$\lambda = \frac{1}{\mathbb{P}(W)} - 1$$

In our table and our formula for R , we don't care about how much is won, but instead how much we're left with after a win. So after winning a bet of X , we will be left with a sum of X' such that

$$\mathbf{X}' = X + \lambda X = X + \frac{X}{\mathbb{P}(W)} - X = \mathbf{X} \frac{\mathbf{1}}{\mathbb{P}(\mathbf{W})}$$

So instead of a double or nothing, it must be a *multiply your bet by $\frac{1}{\mathbb{P}(W)}$* or nothing. We can now include this new rule in our formula for R . Replacing the 2 in (4) by $\frac{1}{\mathbb{P}(W)}$, we get:

$$R_X(Y) = \sum_{i=1}^3 \prod_{j=1}^i \frac{1_{[Y_j = X_{3-i+j}]}}{\mathbb{P}(Y_j)} \quad (5)$$

$\mathbb{P}(Y_j)$ is written instead of $\mathbb{P}(W)$ because, in this formula, the bet is being placed on the j th element of sequence Y appearing, which we define to be $\mathbb{P}(Y_j)$. So $\mathbb{P}(Y_j)$ is the probability of winning. For example, with a fair coin, if sequence Y is HHT and $j = 2$, then $\mathbb{P}(Y_j) = \mathbb{P}(H) = \frac{1}{2} = \mathbb{P}(W)$ and we return to what we had in (4) because $\frac{1}{(\frac{1}{2})} = 2$.

3 Extending to games with sequences of different lengths

As it was previously discussed, a sequence of bets can only return money (and therefore can only matter) if either of these things happen:

1. The bets go to completion. In other words, if all bets in the sequence succeed. This will end the game.
2. A sequence of bets gets cut off by the game ending.

This allowed us to only look at the last 3 sequences of bets in the basic variation, because there will be 3 bets going on when the game ends, because the sequences predict the next three flips. For a game involving sequences of length k , all we have to change is that we should now look at the last k sequences of bets before a player ends the betting. This is again a simple fix in the formula for R , introducing a k in the place of the 3:

$$R_X(Y) = \sum_{i=1}^k \prod_{j=1}^i \frac{1_{[Y_j=X_{k-i+j}]}}{\mathbb{P}(Y_j)} \quad (6)$$

At this point, our formula for R is identical to the one derived by Miller (2019), despite some notation changes.

However, we can make one more small change to allow for games in which the players are predicting a different number of coin flips into the future compared to each other. For a sequence S , we will define its length (the number of flips predicted in the future) to be L_S .

Using the same logic as before, we find that the number of sequences of bets that must be checked when calculating $R_X(Y)$ must be $k = \min\{L_X, L_Y\}$. Additionally, the subscript of X in $1_{[Y_j=X_{k-i+j}]}$ continues to involve the length of X , which has been changed from k to L_X . This completes our work on R :

$$R_X(Y) = \sum_{i=1}^{\min\{L_X, L_Y\}} \prod_{j=1}^i \frac{1_{[Y_j=X_{L_X-i+j}]}}{\mathbb{P}(Y_j)} \quad (7)$$

Part V

Deriving a formula for win probabilities in games with any number of players

The last way that we will generalize our formula is to make it work with any number of players.

We will define n to represent the number of players, and we will store the predicted sequences of all players in a list: $\{S_1, S_2, \dots, S_n\}$.

The first thing we can do is edit (2). This equation says that in the betting game with two players, they pay each other their winnings. We can make this work with more than two players. A way to do this which keeps the game fair for everyone (this is still necessary), is to say that at the end of the game, each player's winnings are paid by all other players equally. For example, in a game with 5 players, if one player wins \$1, all the others will give that player \$0.25 when the game ends. Mathematically, that rule can be expressed like this:

$$N_X(Y) = \frac{1}{n-1} \left[nR_X(Y) - \sum_{i=1}^n R_X(S_i) \right] \quad (8)$$

The problem that is now left is that there are n players, so there are not just two win probabilities (previously simply $\mathbb{P}(A)$ and $\mathbb{P}(B) = 1 - \mathbb{P}(A)$), but n unknown win probabilities.

However, we can generate enough equations to solve for all of them. For each player, the betting game must be fair. Therefore,

$$N_{S_1}(S_1) \cdot \mathbb{P}(S_1) + N_{S_2}(S_1) \cdot \mathbb{P}(S_2) + \dots + N_{S_n}(S_1) \cdot \mathbb{P}(S_n) = 0$$

$$N_{S_1}(S_2) \cdot \mathbb{P}(S_1) + N_{S_2}(S_2) \cdot \mathbb{P}(S_2) + \dots + N_{S_n}(S_2) \cdot \mathbb{P}(S_n) = 0$$

...(one for the net profit of each player)...

$$N_{S_1}(S_n) \cdot \mathbb{P}(S_1) + N_{S_2}(S_n) \cdot \mathbb{P}(S_2) + \dots + N_{S_n}(S_n) \cdot \mathbb{P}(S_n) = 0$$

And finally, someone must win, so

$$\mathbb{P}(S_1) + \mathbb{P}(S_2) + \dots + \mathbb{P}(S_n) = 1$$

Translating this system of linear equations into a matrix equation, we get the general formula to solve for the win probabilities of any number of players:

$$\begin{bmatrix} N_{S_1}(S_1) & N_{S_2}(S_1) & \dots & N_{S_n}(S_1) \\ N_{S_1}(S_2) & N_{S_2}(S_2) & \dots & N_{S_n}(S_2) \\ \vdots & \vdots & \ddots & \vdots \\ N_{S_1}(S_{n-1}) & N_{S_2}(S_{n-1}) & \dots & N_{S_n}(S_{n-1}) \\ 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} \mathbb{P}(S_1) \\ \mathbb{P}(S_2) \\ \vdots \\ \mathbb{P}(S_{n-1}) \\ \mathbb{P}(S_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (9)$$

Note that we were able to remove the equation for the last player and replace it by $\mathbb{P}(S_1) + \mathbb{P}(S_2) + \dots + \mathbb{P}(S_n) = 1$, since only n equations are necessary.

Part VI

Two-player games with equal sequence lengths and even flip outcome likelihoods

Solving (8) and (9) with $n = 2$, we get a formula for the probabilities in a two-player game. This formula is, of course, identical to (3), which was derived earlier on.

$$\mathbb{P}(A) = \frac{R_B(B) - R_B(A)}{R_A(A) - R_A(B) + R_B(B) - R_B(A)} \quad (10)$$

$$\mathbb{P}(B) = \frac{R_A(A) - R_A(B)}{R_B(B) - R_B(A) + R_A(A) - R_A(B)} \quad (11)$$

And since either Player A or Player B must win in a game,

$$\mathbb{P}(A) + \mathbb{P}(B) = 1$$

We will use s to denote the number of possible outcomes of the flip. $s = 2$ for coins, $s = 6$ for 6-sided dice, etc.

4 Proving that when sequence S appears, Player S will always win more money than their opponent in the betting game

In other words, we are making the claim that $R_A(A) > R_A(B)$ in all cases and $R_B(B) > R_B(A)$ in all cases.

We will start by showing that A: 100...0 and B: 000...0 produces the lowest possible $R_A(A)$ and the highest possible $R_A(B)$.

$$R_A(A): \begin{array}{|c|c|c|c|c|c|} \hline 1 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline 1 & 0 & 0 & \dots & 0 & s^k \\ \hline & \ddagger & 0 & \dots & 0 & 0 \\ \hline & & \ddagger & \dots & 0 & 0 \\ \hline & & & \ddots & 0 & 0 \\ \hline & & & & \ddagger & 0 \\ \hline \end{array} = s^k$$

This is the minimum value, since necessarily $A = A$.

$$R_A(B): \begin{array}{|c|c|c|c|c|c|} \hline 1 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline \textcolor{red}{\oplus} & 0 & 0 & \dots & 0 & 0 \\ & \textcolor{green}{0} & \textcolor{green}{0} & \dots & \textcolor{green}{0} & s^{k-1} \\ & & \textcolor{green}{0} & \dots & \textcolor{green}{0} & s^{k-2} \\ & & & \dots & \textcolor{green}{0} & \dots \\ & & & & \textcolor{green}{0} & s \\ \hline \end{array} = \sum_{i=1}^{k-1} s^{k-i}$$

This is the maximum value, since $A \neq B$ or else the game would be invalid.

Therefore, if $s^k > \sum_{i=1}^{k-1} s^{k-i}$, $R_A(A) > R_A(B)$ will always be true. This is because $R_A(A)$ would be greater than $R_A(B)$ even when it's the most disadvantaged.

$\sum_{i=1}^{k-1} s^{k-i}$ can be rewritten as $s^k \sum_{i=1}^{k-1} \left(\frac{1}{s}\right)^i$, and since it's a sum of positive numbers, it must be less than $s^k \sum_{i=1}^{\infty} \left(\frac{1}{s}\right)^i$, a geometric series which is equal to $s^k \frac{s}{s-1} - s^k$.

Now, if $s^k \geq s^k \frac{s}{s-1} - s^k$, then s^k will definitely be larger than $\sum_{i=1}^{k-1} s^{k-i}$ and our claim will be proven.

$$\begin{aligned} s^k &\geq s^k \frac{s}{s-1} - s^k \\ 1 &\geq \frac{s}{s-1} - 1 \\ s &\geq 2 \end{aligned}$$

Which is always the case, since there must be at least two possible outcomes from the flip.

This proves that $R_A(A) > R_A(B)$. If the exact same process is done after swapping the sequences of A and B, we also learn that $R_B(B) > R_B(A)$.

So the player who has chosen a sequence will always get the greatest return when that sequence appears. For example, if Player A's sequence appears in the flips, and the player makes \$10, Player B could not possibly make any more than \$9.

5 Determining how changes of R values affect $\mathbb{P}(A)$

It will be useful to know whether increasing each R value will increase or decrease $\mathbb{P}(A)$. Note that because of (1), a value that increases $\mathbb{P}(A)$ will necessarily decrease $\mathbb{P}(B)$. We focus on $\mathbb{P}(A)$ and the results for $\mathbb{P}(B)$ will simply be the opposite.

Although all R values depend on each other, artificially increasing each value separately gives a good enough idea of how that value affects $\mathbb{P}(A)$.

We start with $R_B(B)$ and increase its value to $R_B(B) + N$ where N is any natural number greater than 0.

Rewriting (10) with this new value, we get:

$$\mathbb{P}_N(A) = \frac{(\mathbf{R}_B(\mathbf{B}) + \mathbf{N}) - R_B(A)}{R_A(A) - R_A(B) + (\mathbf{R}_B(\mathbf{B}) + \mathbf{N}) - R_B(A)}$$

Now, lets check if $\mathbb{P}_N(A) > \mathbb{P}(A)$:

$$\begin{aligned}\mathbb{P}_N(A) &> \mathbb{P}(A) \\ \frac{\mathbb{P}_N(A)}{\mathbb{P}(A)} &> 1\end{aligned}$$

Dividing, simplifying, and factoring, we get that this is true if and only if:

$$\frac{N[R_A(A) - R_A(B)]}{[R_B(B) - R_B(A)][R_A(A) - R_A(B) + N + R_B(B) - R_B(A)]} > 0$$

We know from the previous section that $R_A(A) > R_A(B)$ and $R_B(B) > R_B(A)$, so the above inequality will always hold true, with all terms being positive numbers.

This means that increasing $R_B(B)$ will always increase $\mathbb{P}(A)$, and decreasing $R_B(B)$ will always decrease $\mathbb{P}(A)$.

Next, we will do the same for $R_B(A)$, increasing it to $R_B(A) + N$. This time, we need to note that since $R_B(B)$ must be greater than $R_B(A)$, we have the condition on N that $0 < N < R_B(B) - R_B(A)$.

Rewriting (10), this time with $R_B(A)$ incremented by N , we get:

$$\mathbb{P}_N(A) = \frac{R_B(B) - (\mathbf{R}_B(\mathbf{A}) + \mathbf{N})}{R_A(A) - R_A(B) + R_B(B) - (\mathbf{R}_B(\mathbf{A}) + \mathbf{N})}$$

Now doing the same process as before, we end up with an increase of $R_B(A)$ increasing $\mathbb{P}(A)$ if and only if:

$$\frac{N[R_A(B) - R_A(A)]}{[R_B(B) - R_B(A)][R_A(A) - R_A(B) - N + R_B(B) - R_B(A)]} > 0$$

We know from the previous section that $R_A(B) - R_A(A)$ must be negative and $R_B(B) - R_B(A)$ must be positive. Additionally, the condition that $N < R_B(B) - R_B(A)$ means that $R_B(B) - R_B(A) - N > 0$.

So in the large inequality, the numerator on the left side will always be negative and the denominator will always be positive, so the inequality will be false and increasing $R_B(A)$ will always decrease $\mathbb{P}(A)$ and vice versa.

We can do the exact same process for B and use the idea that increasing $\mathbb{P}(B)$ will necessarily decrease $\mathbb{P}(A)$ to determine that increasing $R_A(A)$ will always decrease $\mathbb{P}(A)$ and increasing $R_A(B)$ will always increase $\mathbb{P}(A)$, and vice versa.

In summary:

$$R_A(A) \uparrow \implies \mathbb{P}(A) \downarrow \tag{12}$$

$$R_A(B) \uparrow \implies \mathbb{P}(A) \uparrow \tag{13}$$

$$R_B(B) \uparrow \implies \mathbb{P}(A) \uparrow \quad (14)$$

$$R_B(A) \uparrow \implies \mathbb{P}(A) \downarrow \quad (15)$$

And vice versa.

Note that swapping the symbols A and B will give you the results of an equivalent analysis of $\mathbb{P}(B)$.

6 Proving that the greatest win probability difference occurs when $baa\dots a$ is played against $aaa\dots a$ for $b \neq a$

As a specific example of the case described in the title, let A: $100\dots 0$ and B: $000\dots 0$ in a game with sequences of length k and with s possible outcomes from the flip.

From the previous section, s^k is the minimum possible value of $R_A(A)$, and this maximizes $\mathbb{P}(A)$ by (12), and $\sum_{i=1}^{k-1} s^{k-i}$ is the maximum possible value of $R_A(B)$, maximizing $\mathbb{P}(A)$ by (13).

Determining the other return values:

$$R_B(B): \begin{array}{|c|c|c|c|c|c|} \hline 0 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline 0 & 0 & 0 & \dots & 0 & s^k \\ & 0 & 0 & \dots & 0 & s^{k-1} \\ & & 0 & \dots & 0 & s^{k-2} \\ & & & \dots & 0 & \dots \\ & & & & 0 & s \\ \hline \end{array} = \sum_{i=0}^{k-1} s^{k-i}$$

Which is the maximum possible value, maximizing $\mathbb{P}(A)$ by (14).

$$R_B(A): \begin{array}{|c|c|c|c|c|c|} \hline 0 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline \textcolor{red}{\perp} & 0 & 0 & \dots & 0 & 0 \\ & \textcolor{red}{\perp} & 0 & \dots & 0 & 0 \\ & & \textcolor{red}{\perp} & \dots & 0 & 0 \\ & & & \dots & 0 & 0 \\ & & & & \textcolor{red}{\perp} & 0 \\ \hline \end{array} = 0$$

Which is the minimum possible value, maximizing $\mathbb{P}(A)$ by (15).

In summary, these two sequences played against each other provide the maximum value of $\mathbb{P}(A)$. This finding will be used in the next section.

7 Determining the bounds of $\mathbb{P}(A)$ and $\mathbb{P}(B)$

Substituting the previous section's values which maximize $\mathbb{P}(A)$ into (10), we get:

$$\mathbb{P}(A) \leq \frac{\sum_{i=0}^{k-1} s^{k-i} - 0}{s^k - \sum_{i=1}^{k-1} s^{k-i} + \sum_{i=0}^{k-1} s^{k-i} - 0}$$

$$\mathbb{P}(A) \leq \frac{1}{2} \sum_{i=0}^{k-1} \frac{1}{s^i} \quad (16)$$

Switching the sequences of A and B, we can also conclude that

$$\mathbb{P}(B) \leq \frac{1}{2} \sum_{i=0}^{k-1} \frac{1}{s^i}$$

From (1), $\mathbb{P}(A) = 1 - \mathbb{P}(B)$. So the previous equation implies that:

$$\mathbb{P}(A) \geq 1 - \frac{1}{2} \sum_{i=0}^{k-1} \frac{1}{s^i}$$

And similarly,

$$\mathbb{P}(B) \geq 1 - \frac{1}{2} \sum_{i=0}^{k-1} \frac{1}{s^i}$$

Now the bounds for the probability that any sequence S wins can be expressed as:

$$1 - \frac{1}{2} \sum_{i=0}^{k-1} \frac{1}{s^i} \leq \mathbb{P}(S) \leq \frac{1}{2} \sum_{i=0}^{k-1} \frac{1}{s^i} \quad (17)$$

8 Using the probability bounds to study extreme games.

8.1 Proving that the game tends to be fair with infinite possible flip outcomes but finite sequence length

$$\lim_{s \rightarrow \infty} \sum_{i=0}^{k-1} \frac{1}{s^i} = \lim_{s \rightarrow \infty} \left[1 + \frac{1}{s} + \frac{1}{s^2} + \cdots + \frac{1}{s^{k-1}} \right] = 1$$

Using this insight in (17),

$$1 - \frac{1}{2} \cdot 1 \leq \lim_{s \rightarrow \infty} \mathbb{P}(S) \leq \frac{1}{2} \cdot 1$$

Therefore $\lim_{s \rightarrow \infty} \mathbb{P}(S) = \frac{1}{2}$ by the Squeeze Theorem, and the game tends toward being fair with more possible flip outcomes.

8.2 Determining the bounds for games with infinitely long sequences but finite possible flip outcomes

$$\lim_{k \rightarrow \infty} \frac{1}{2} \sum_{i=0}^{k-1} \frac{1}{s^i} = \frac{1}{2} \sum_{i=0}^{\infty} \frac{1}{s^i} = \frac{s}{2s-2}$$

Thus taking (17) as $k \rightarrow \infty$, we get:

$$1 - \frac{s}{2s-2} \leq \lim_{k \rightarrow \infty} \mathbb{P}(S) \leq \frac{s}{2s-2} \quad (18)$$

$$\left| \lim_{k \rightarrow \infty} \mathbb{P}(S) - \frac{1}{2} \right| \leq \frac{1}{2s-2}$$

So the limit might not exist, but (and this is supported by simulations) the probability of a sequence winning will tend toward being contained within the range $\frac{1}{2s-2}$ greater than or less than 50% as sequence length approaches infinity.

8.3 Proving that the game tends to be fair with infinite possible flip outcomes and infinitely long sequences.

Using (18) and having $s \rightarrow \infty$,

$$\lim_{s \rightarrow \infty} \left[1 - \frac{s}{2s-2} \right] \leq \lim_{(k,s) \rightarrow (\infty, \infty)} \mathbb{P}(S) \leq \lim_{s \rightarrow \infty} \frac{s}{2s-2}$$

$$\frac{1}{2} \leq \lim_{(k,s) \rightarrow (\infty, \infty)} \mathbb{P}(S) \leq \frac{1}{2}$$

And so by the Squeeze Theorem, $\lim_{(k,s) \rightarrow (\infty, \infty)} \mathbb{P}(S) = \frac{1}{2}$

Part VII

Two-player games with different sequence lengths and even flip outcome likelihoods

9 Determining when the shorter sequence always has an advantage in games with two players and even flip outcome probabilities.

First, we will prove that for all games with A being shorter than B, A: 000...0 and B: 1...100...0 leads to the minimum possible probability for A to win.

$$R_A(A): \begin{array}{|c|c|c|c|c|c|} \hline 0 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline 0 & 0 & 0 & \dots & 0 & s^{L_A} \\ & 0 & 0 & \dots & 0 & s^{L_A-1} \\ & & 0 & \dots & 0 & s^{L_A-2} \\ & & & \dots & 0 & \dots \\ & & & & 0 & s \\ \hline \end{array} = \sum_{i=0}^{L_A-1} s^{L_A-i}$$

Which is the maximum possible value, minimizing $\mathbb{P}(A)$ by (12)

$$R_A(B): \begin{array}{|c|c|c|c|c|c|} \hline 0 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline \cancel{1} & 0 & 0 & \dots & 0 & 0 \\ & \cancel{1} & 0 & \dots & 0 & 0 \\ & & \cancel{1} & \dots & 0 & 0 \\ & & & \dots & 0 & 0 \\ & & & & \cancel{1} & 0 \\ \hline \end{array} = 0$$

Which is the minimum possible value, minimizing $\mathbb{P}(A)$ by (13).

$$R_B(B): \begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & s^{L_A} \\ & 1 & \dots & 1 & \cancel{1} & 0 & \dots & 0 & 0 \\ & & \dots & 1 & \cancel{1} & 1 & \dots & 0 & 0 \\ & & & 1 & \cancel{1} & 1 & \dots & 1 & 0 \\ & & & & \cancel{1} & 1 & \dots & 1 & 0 \\ & & & & & \cancel{1} & \dots & 1 & 0 \\ & & & & & & \dots & 1 & 0 \\ & & & & & & & \cancel{1} & 0 \\ \hline \end{array} = s^{L_A}$$

Which is the minimum possible, value since $B = B$. This minimizes $\mathbb{P}(A)$ by (14).

$$R_B(A): \begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \text{Gain} \\ \hline & & & \cancel{1} & 0 & 0 & \dots & 0 & s^{L_A} \\ & 1 & \dots & 1 & \cancel{1} & 0 & \dots & 0 & 0 \\ & & \dots & 1 & \cancel{1} & 1 & \dots & 0 & 0 \\ & & & 1 & \cancel{1} & 1 & \dots & 1 & 0 \\ & & & & \cancel{1} & 1 & \dots & 1 & 0 \\ & & & & & \cancel{1} & \dots & 1 & 0 \\ & & & & & & \dots & 1 & 0 \\ & & & & & & & 1 & 0 \\ \hline \end{array} = \sum_{i=1}^{L_A-1} s^{L_A-i}$$

NOT DONE

Part VIII

Verifying the formula with simulations

Part IX

Creating a tool to aid in research of Penney's Game

Part X

Physically modelling unfair coin flips and dice rolls

Part XI

Conclusion

Part XII

References

References

- Chung, Derrick (2024). *Fair Game. An exploration of coin flips both real and imaginary.*
- Gerber, Hans U. and Shuo-Yen Robert Li (1981). "The Occurence of Sequence Patterns in Repeated Experiments and Hitting Times in a Markov Chain". In: *Stochastic Processes and their Applications* 11, pp. 101–108.
- Miller, Joshua B. (2019). "Penney's Game Odds from No-Arbitrage". URL: <https://doi.org/10.48550/arXiv.1904.09888>.
- Penney, Walter (1969). "Problem 95. Penney-Ante". In: *Journal of Recreational Mathematics* 2, p. 241.