THE M&M GAME: FROM MORSELS TO MODERN MATHEMATICS

IVAN BADINSKKI, CHRISTOPHER HUFFAKER, NATHAN MCCUE, CAMERON N. MILLER,
KAYLA S. MILLER, STEVEN J. MILLER, AND MICHAEL STONE

ABSTRACT. To an adult, it's obvious that the day of someone's death is not precisely determined by the day of birth, but it's a very different story for a child. When the third named author was four years old he asked his father, the fifth named author: If two people are born on the same day, do they die on the same day? While this could easily be demonstrated through murder, such a proof would greatly diminish the possibility of teaching additional lessons, and thus a different approach was taken. With the help of the fourth named author they invented what we'll call *the M&M Game*: Given k people, each simultaneously flips a fair coin, with each eating an M&M on a head and not eating on a tail. The process then continues until all M&M'S are consumed, and two people are deemed to die at the same time if they run out of M&M'S together[1]. This led to a great concrete demonstration of randomness appropriate for little kids; it also led to a host of math problems which have been used in probability classes and math competitions. There are many ways to determine the probability of a tie, which allow us in this article to use this problem as a springboard to a lot of great mathematics, including memoryless process, combinatorics, statistical inference, graph theory, and hypergeometric functions.

CONTENTS

1. The Origins of The Game

The M&M Game began as a simple question asked by Steven Miller's curious four year-old son Cam: If two people are born on the same day, do they die on the same day? Of course, needing a way to explain randomness to children (two year old Kayla was there as well), the three Millers took the most logical next step and used M&M'S to give the answer - with a more fun question! This led to what we now call the M&M Game (see Figure 1 for an illustration):

> *You and some friends start with some number of M&M'S. Everyone flips a fair coin at the same time; if you get a head you eat an M&M; if you get a tail you don't. You continue tossing coins together until no one has any M&M'S left, and whoever is the last person with an M&M lives longest and 'wins'.*

FIGURE 1.   The first M&M Game; for young players there is an additional complication in that it matters which colors you have, and the order you place them down.

We can reformulate Cam's question on randomness to: If everyone starts with the same number of M&M'S, what is the chance everyone eats their last M&M at the same time? In the arguments below we'll concentrate on two people playing with c (for Cam) and k (for Kayla) M&M'S, though we encourage you to extend to the case of more people playing, possibly with a biased coin. As we will see in the following analysis, probability games like this one are a great way to see useful but complicated mathematical processes. In the course of our investigations we'll see some nice results in combinatorics and graph theory, and see applications of memoryless processes, statistical

inference and hypergeometric functions. Such consequences are typical of good problems: in addition to being interesting, they serve as an excellent springboard to good concepts.

Recalling that the binomial coefficient $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ denotes the number of ways to choose r objects from n when order doesn't matter, we can compute the probability $P(k, k)$ of a tie when two people start with k M&M'S. If we let $P_n(k, k)$ denote the probability that the game ends in a tie with both people starting with k M&M'S after *exactly* n moves, then

$$P(k, k) = \sum_{n=k}^{\infty} P_n(k, k);$$

note that we are starting the sum at k as it is impossible all the M&M'S are eaten in fewer than k moves (we could start the sum at zero, but since $P_n(k, k) = 0$ for $n < k$ there is no need).

We claim that

$$P_n(k, k) = \binom{n-1}{k-1} \left(\frac{1}{2}\right)^n \binom{n-1}{k-1} \left(\frac{1}{2}\right)^n.$$

This formula follows from the following observation: if the game ends in a tie after n tosses, then each person has *exactly* $k-1$ heads in their first $n-1$ tosses. As we have a fair coin, each string of heads and tails of length n for a player has probability $(1/2)^n$. The number of strings for each person where the first $n-1$ tosses have *exactly* $k-1$ heads, and the nth toss is a head (we need this as otherwise we do not have each person eating their final M&M on the nth move) is $\binom{n-1}{k-1}\binom{1}{1}$. The $\binom{1}{1}$ reflects the fact that the last toss must be a head; as this is just 1 it is common to omit that factor. As there are two players, the probability that each has their k^{th} head after the nth toss is the product, proving the formula.

We have thus shown the following.

Theorem 1.1. *The probability the M&M Game ends in a tie with two people using fair coins and starting with k M&M'S is*

$$P(k, k) = \sum_{n=k}^{\infty} \binom{n-1}{k-1} \left(\frac{1}{2}\right)^n \binom{n-1}{k-1} \left(\frac{1}{2}\right)^n = \sum_{n=k}^{\infty} \binom{n-1}{k-1}^2 \frac{1}{2^{2n}}. \tag{1.1}$$

While the above formula solves the problem, it is unenlightening and difficult to work with. The first difficulty is that it involves an infinite sum over n.[2] Second, it is very hard to sniff out the k-dependence: if we double k, what does that do to the probability of a tie? It is highly desirable to have exact, closed form solutions so we can not only quickly compute the answer for given values of the parameter, but also get a sense of how the answer changes as we vary those inputs. In the sections below we'll look at many different approaches to this problem, most of them trying to convert the infinite sum to a more tractable finite problem.

2. THE BASKETBALL PROBLEM, MEMORYLESS PROCESSES AND THE GEOMETRIC SERIES FORMULA

2.1. **A Basketball Game.** It turns out that we can easily convert the infinite M&M Game sum, equation (1.1), into a finite sum using a powerful observation: we have a *Memoryless Process*.

---

[2]In general we need to be careful and make sure any infinite sum converges; while we are safe here as we are summing probabilities, we can elementarily prove convergence. Note $\binom{n-1}{k-1} \leq n^{k-1}/k!$, and thus the sum is bounded by $k!^{-2} \sum_{n \geq k} n^{2k-2}/2^{2n}$; as the polynomial n^{2k-2} grows significantly slower than the exponential factor 2^{2n}, the sum rapidly converges.

Briefly, what this means is that the behavior of the system only depends on the values of the parameters at a given moment in time, and not on how we got there.

There are many examples where all that matters is the configuration, not the path taken to reach it. For example, imagine a baseball game. If the lead-off hitter singles or walks, the net effect is to have a runner on first and the two results are the same.[3] For another example, consider a game of Tic-Tac-Toe; what matters are where the X's and O's are on the board, not the order they are placed. While chess at first might seem like a perfect example, it fails as many people play that if there is ever a configuration repeated three times in the game then the game is declared a draw; thus in chess we need to know *how* we reached our state, and not just what state we are in.

Before delving into the reduction of the M&M Game into a finite problem, we'll look at a related problem that's a little simpler but illustrates the same point. Moreover, we can easily extract from this problem the famous geometric series formula!

Imagine two of the greatest basketball players of all time, Larry Bird of the Boston Celtics and Magic Johnson of the Los Angeles Lakers[4] are playing a basketball game. Instead of the intense competition which characterized the matches between their teams (see Figure 2) they instead play a one-on-one game of hoops as follows.

> *In this contest, Bird and Magic alternate shooting free throws, with Bird going first. Assume Bird always makes a basket with probability p_L, while Magic always gets a basket with probability p_M. If the probability Bird wins is x_B, what is x_B?*

FIGURE 2. Larry Bird and Magic Johnson, Game 2 of the 1985 NBA Finals (Boston, MA). Image from Steve Lipofsky from Wikipedia Commons (http://www.basketballphoto.com/NBA_Basketball_Photographs.htm).

Note that this is almost a simplified M&M Game: there is only one M&M, but the players take turns flipping their coins. We'll see, however, that it is straightforward to modify the solution.

---

[3] For the baseball purist, there could be a very slight difference as a single breaks up a no-hit attempt, and if the next 26 batters are retired the pitcher might perform differently with a no-hitter on the line!

[4] The players chosen reflect the childhood experiences of the eldest author.

2.2. Solution from the Geometric Series Formula. The standard way to solve this problem uses a geometric series. Similar to the analysis in the introduction, the probability that Bird wins is the sum of the probabilities that Bird wins on his n^{th} shot. We'll see in the analysis below that it's algebraically convenient to define $r := (1 - p_B)(1 - p_M)$, which is the probability they both miss.[5] Let's go through the cases. We assume that p_B and p_M are not both zero; if they were, then neither can hit a basket. Not only would this mean that our ranking of them as two of the all-time greats is wrong, but the game will never end and thus there's no need to do any analysis!

(1) Bird wins on his 1$^{\text{st}}$ shot with probability p_B.
(2) Bird wins on his 2$^{\text{nd}}$ shot with probability $(1 - p_B)(1 - p_B)p_B = rp_B$.
(3) Bird wins on his n^{th} shot with probability $(1 - p_B)(1 - p_M) \cdot (1 - p_B)(1 - p_M) \cdots (1 - p_B)(1 - p_M)p_B = r^{n-1}p_B$.

To see this, if we want Bird to win on shot n then we need to have him and Magic miss their first $n - 1$ shots, which happens with probability $((1 - p_B)(1 - p_M))^{n-1} = r^{n-1}$, and then Bird hits his n^{th} shot, which happens with probability p_B. The important thing to remember here is that we have broken the problem down into all of the possible ways Bird can beat Magic. In doing so, notice how the geometric series is surfacing! This makes sense since we have $n - 1$ trials where Bird and Magic miss, and on the nth shot, Bird makes the basket and wins the game. Thus

$$\text{Prob(Bird wins)} \;=\; x_B \;=\; p_B + rp_B + rp_B^2 + rp_B^3 + \cdots \;=\; p_B \sum_{n=0}^{\infty} r^n.$$

which is a geometric series. As we assumed p_B and p_M are not both zero, $r = (1 - p_B)(1 - p_M)$ satisfies $|r| < 1$ and we can use the geometric series formula to deduce

$$x_B \;=\; \frac{p_B}{1 - r} \;=\; \frac{p_B}{1 - (1 - p_B)(1 - p_M)}.$$

We have made enormous progress. We converted our infinite series into a ***closed-form expression***, and we can easily see how the probability of Bird winning changes as we change p_B and p_M; we display this in Figure 3.

Note the plot supports our intuition. As the probability of Bird making a basket rises to 1, it doesn't matter what Magic's probability is as Bird will almost surely win on his first shot. Further, if the two probabilities are equal then Bird should win more than half of the time, as there is an advantage in going first.

2.3. Solution through Memoryless Process and the Geometric Series Formula. We now give a second solution to the basketball game. Not only does this approach avoid needing to know the geometric series formula, but it gives a proof of it!

Recall the assumptions we made. The probability Bird makes a shot is p_B, the probability Magic hits a basket is p_M, and the probability they both miss is $r := (1 - p_B)(1 - p_M)$. There is a lot hidden in these statements. We are assuming the two never tire; they always make baskets with a fixed probability. We can use this to compute x_B, the probability Bird wins, in another way. Before we wrote x_B as a sum over the probabilities that Bird won in n games. We claim that

$$\text{Prob(Bird wins)} \;=\; x_B \;=\; p_B + rx_B.$$

---

[5]A quick word on notation. We use p to denote probability, and put subscripts B and M so we can easily determine if we're talking about Bird or Magic; we use the letter r for ratio, which will make sense when we see the geometric series with ratio r emerge shortly. There is enormous value in good notation – we can get a better understanding of what is going on simply by glancing down at the formula and quickly parsing the terms.
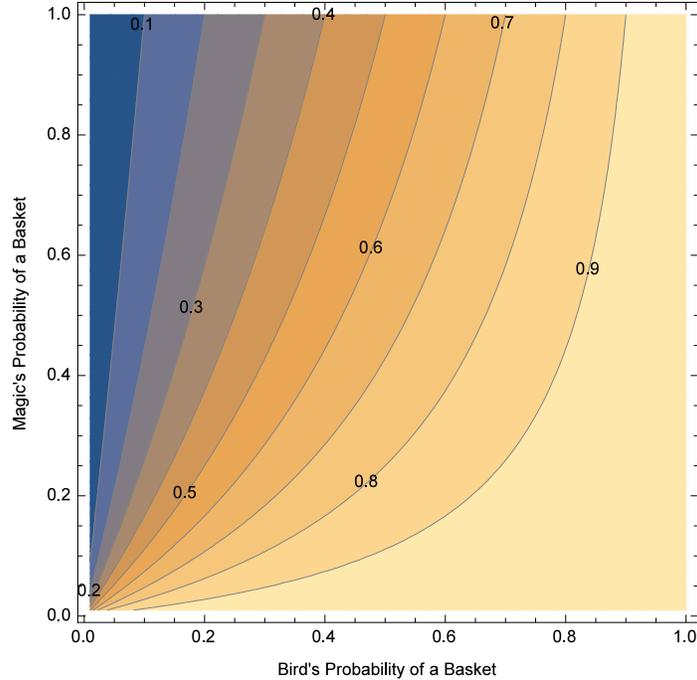
FIGURE 3. Probability Bird, shooting first, gets a basket before Magic.

To see this, note either Bird makes his first basket and wins (which happens with probability p_B) or he misses (with probability $1 - p_B$). If Bird is going to win, then Magic must miss his first shot, and this happens with probability $1 - p_M$. Something interesting happens, however, if both Bird and Magic miss: *we have reset our game to its initial state!* Since both have missed, it's as if we just started playing the game right now. Note this would not be true if we stopped the analysis after Bird misses, as then Magic would have the next shot and the advantage. Since both miss and Bird has the ball again, by definition the probability Bird wins from this configuration is x_B, and thus the probability he wins is $p_B + (1 - p_B)(1 - p_M)x_B$.

Solving for x_B, the probability Bird beats Magic is

$$x_B = \frac{p_B}{1 - r_B}.$$

As this must equal the infinite series expansion from the previous subsection, we deduce the geometric series formula:

$$\frac{p_B}{1 - r} = p_B \sum_{n=0}^{\infty} r^n \quad \text{therefore} \quad \sum_{n=0}^{\infty} r^n = \frac{1}{1 - r}.$$

Remark 2.1. *We have to be a bit careful. It's important to keep track of assumptions. In our analysis $r = (1 - p_B)(1 - p_M)$ with $0 \le p_B, p_M \le 1$ and both p_B and p_M are not zero. Thus we have only proved the geometric series formula if $0 \le r < 1$ (actually, if $p_B = 0$ we cannot divide both sides by p_B, and some care is needed). With a bit more work we can convert this to a proof for all real $|r| < 1$ by splitting the infinite sum into a sum over even and odd powers, and using the formula twice; we encourage you to make this rigorous.*

Let's look closely at what we've done in this subsection. The key observation was to notice that we have a ***memoryless process***. In the infinite sum approach, which led to an infinite geometric series, we cared about each time Bird and Magic miss a free throw. In our new approach we just care about them both missing once. The reason is that if Bird and Magic both miss, the game essentially starts over, and the game has *no memory* of what previously occurred. The advantage to this method is that by reducing the game to the same state we start with, we turn an *infinite* calculation into a *finite* one! In general it is incredibly difficult to come up with a workable expression for an infinite series, and finite expressions are easier to compute. Thus, perhaps there is hope that we can convert the solution to the M&M Game, equation (1.1), into an equivalent finite sum....

2.4. **Lessons.** Before returning to the M&M game, there are a few takeaways worthy of emphasis, all of which will resurface moving forward.

(1) *The Power of Perspective:* In the hoops game, the infinite series may have been daunting. However, after looking at the problem with a different perspective, we saw that we can use a memoryless process to attack an otherwise difficult problem. In fact, the memoryless process is one of the most powerful probability tools we have because it replaces the daunting challenge of infinite calculations with finite ones. Again, any time we can reduce an infinite problem to a finite problem is cause for celebration, as we are making enormous progress! (Technically infinite progress!)

(2) *Circumvent Algebra with Deeper Understanding:* Frequently there is a lot of messy algebra that goes into finding a formula for an infinite sum. The tricks we used to circumnavigate this algebra are great, and we should look for those types of shortcuts as often as possible.

(3) *The Depth of a Problem Is Not Always What You Expect:* Originally, we may have thought we needed the geometric series to solve this problem. It turns out we didn't! This will be a valuable insight for the M&M game in that we should look for ways to simplify problems from infinite sums to finite expressions. That way, we don't have to deal with difficult infinite expressions.

(4) *Math is Fun:* How could anyone think otherwise?

3. MEMORYLESS M&M'S

3.1. **Setup.** Remember (equation (1.1)) that we have an infinite sum for the probability of a tie with both people starting with k M&M'S:

$$P(k,k) = \sum_{n=k}^{\infty} \binom{n-1}{k-1} \left(\frac{1}{2}\right)^{n-1} \frac{1}{2} \cdot \binom{n-1}{k-1} \left(\frac{1}{2}\right)^{n-1} \frac{1}{2}.$$

It's hard to evaluate this series as we have an infinite sum *and* a squared binomial coefficient whose top is changing. Thus instead of evaluating this sum, which is very difficult to do, we want to somehow convert it to something where we have more familiarity. From the hoops game, we should be thinking about how to obtain a *finite* calculation. The trick there was to notice we had a memoryless process, and all that mattered was the game state, not how we reached it. For our problem, we'll have many tosses of the coins, but in the end what matters is where we are, not the string of heads and tails that got us there.

Let's figure out some way to do this by letting $k = 1$. In this case, we can do the same thing we did in the hoops game and boil the problem down into cases. There are four equally likely scenarios each time we toss coins, so the probability of each event occurring is 1/4 or 25%.

(1) Both players eat.
(2) Cam eats an M&M but Kayla does not.
(3) Kayla eats an M&M but Cam does not.
(4) Neither eat.

These four possibilities lead to the infinite series in (1.1), as we calculate the probability the game ends in n tosses. It turns out one of the four events is not needed, and if we remove it we can convert to a finite game.

Let's revisit the lessons of the hoops game. There, we saw that we could create a **memoryless process** by saying if Bird and Magic both missed their free throws, it was as if the game started over. We can do the same thing here: if Cam and Kayla both get tails and therefore don't eat their M&Ms, then it's as if the coin toss never happened. We can therefore ignore the fourth possibility. If you want, another way to look at this is that if we toss two tails then there is no change in the number of M&M'S for either kid, and thus we may pretend such a toss never happened. This allows us to remove all the tosses of double tails, and now after each toss at least one player, possibly both, have fewer M&M'S. As we start with a finite number of M&M'S, the game terminates in a finite number of moves.

Thus instead of viewing our game as having four alternatives each toss, there are only three and they all happen with probability 1/3. To see this, note that if $\Pr(X)$ is the probability that event X happens, we now have a conditional probability problem[6]; for example, what is the probability Cam and Kayla both eat an M&M *given that the outcome is not double tails*? If C denotes the event that Cam gets a head and eats (and C^c the event that he gets a tail), and similarly K for Kayla, then

$$
\begin{aligned}
\Pr(\text{both eat}|\text{at least one eats}) \ &= \ \frac{\Pr(C \cap K)}{\Pr(C \cap K) + \Pr(C \cap K^c) + \Pr(C^c \cap K)} \\
&= \ \frac{1/4}{1/4 + 1/4 + 1/4} \ = \ \frac{1}{3}.
\end{aligned}
$$

We may therefore consider the related game with just three outcomes for each set of tosses, each happening with probability 1/3:

(1) both players eat;
(2) Cam eats an M&M but Kayla does not;
(3) Kayla eats an M&M but Cam does not.

Notice that after each toss the number of M&M'S is decreased by either 1 or 2, so the game ends after at most $2k - 1$ tosses.

3.2. **Solution.** Armed with the reduction from the previous subsection, we can replace the infinite sum of (1.1) with a finite sum.

---

[6]The standard notation is to write $\Pr(A|B)$ for the probability that A happens, given that B happens.

Theorem 3.1. *The probability the M&M Game ends in a tie with two people using fair coins and starting with k M&M'S is*

$$P(k,k) = \sum_{n=0}^{k-1} \binom{2k-n-2}{n} \left(\frac{1}{3}\right)^n \binom{2k-2n-2}{k-n-1} \left(\frac{1}{3}\right)^{k-n-1} \left(\frac{1}{3}\right)^{k-n-1} \frac{1}{3}. \qquad (3.1)$$

Proof. Each of our three possibilities (both eat, just Cam eats, just Kayla eats) happens with probability 1/3. Since the game ends in a tie, we know the final toss must be double heads with both eating, and each must eat exactly $k-1$ M&M'S in the earlier tosses. Let n denote the number of times both eat before the final toss (which again we know must be double heads); clearly $n \in \{0, 1, \ldots, k-1\}$. We thus have $n+1$ double heads, and thus Cam and Kayla must each eat $k-(n+1) = k-n-1$ times when the other doesn't eat.

We see that, in the case where there are $n+1$ double heads (with the last toss being double heads), the total number of tosses is

$$(n+1) + (k-n-1) + (k-n-1) = 2k-n-1.$$

In the first $2k-n-2$ tosses we must choose n to be double heads, then of the remaining $(2k-n-2) - n = 2k-2n-2$ tosses before the final toss we must choose $k-n-1$ to be just heads for Cam, and then the remaining $k-n-1$ tosses before the final toss must all be just heads for Kayla. These choices explain the presence of the two binomial factors. As each toss happens with probability 1/3, this explains those factors; note we could have just written $(1/3)^{2k-n-1}$, but we prefer to highlight the sources. $\qquad \square$

4. Viewing Data

4.1. **Plotting Exact Answer.** Before turning to additional ways to solve the problem, it is worthwhile to pause for a bit and discuss how to view data and use results for small k to predict results for larger ones.

While it is not obvious how we could replace the sum in (3.1) with a nice closed form expression involving k, this finite sum is certainly easier to use than the infinite sum in (1.1). In fact, it's very easy to use the finite sum to compute the exact answer. Below is some simple code to do so in Mathematica (and plot the result).

```
p[k_] := Sum[Binomial[2 k - n - 2, n] Binomial[2 k - 2 n - 2,
    k - n - 1] (1/3)^(2 k - n - 1), {n, 0, k - 1}]
tielist = {};
For[k = 1, k <= 1000, k++, tielist = AppendTo[tielist, {k, p[k]}]]
ListPlot[tielist, AxesLabel -> {"k", "Probability of a tie"}]
```

For example, if $k = 1$ the probability of a tie is 1/3; this is quite reasonable, as there are three equally likely possibilities now and only one of them leads to a tie when both start with one M&M. Some other fun values: if $k = 2$ the probability is $5/27 \approx .185$, if $k = 5$ it is $1921/19683 \approx .098$, if $k = 10$ it falls to almost .066, while for $k = 100$ it's about .020. See Figure 4 for more values.
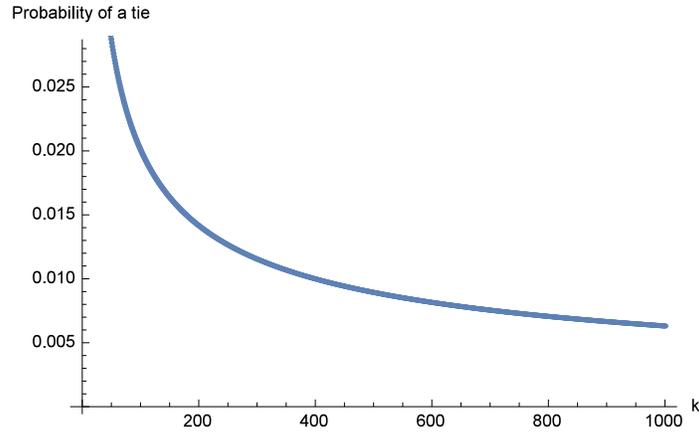
FIGURE 4.   The probability of a tie for $k \leq 1000$.

4.2. **Log-log Plots.** While equation (3.1) gives us a nice formula for finite computations, it is hard to see the k dependence. To try and guess how the answer varies with k we can do a plot, but it's hard to look at the results in Figure 4 and extrapolate to larger values of k. For example, what would you guess for the probability of a tie if there are 200 M&M'S? If there are 2016?

An important skill to learn is how to view data. Frequently rather than plotting the data as given it's better to do a log-log plot. What this means is that instead of plotting the probability of a tie as a function of k, we plot the logarithm of the probability of a tie against the logarithm of k. We do this in Figure 5.
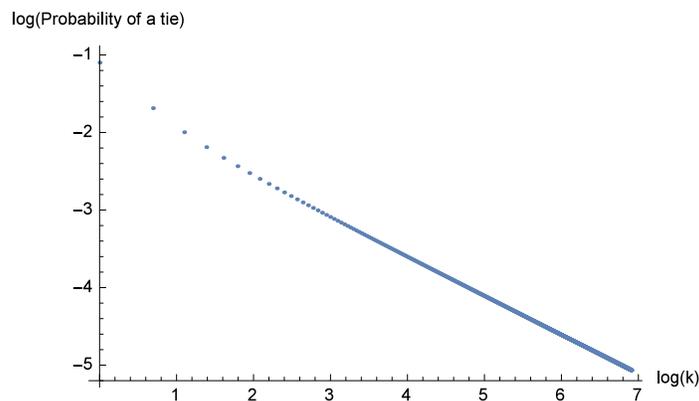

FIGURE 5.   The probability of a tie for $k \leq 1000$.

Notice that the plot here looks *very* linear. Lines are probably the easiest functions to extrapolate, and if this linear relationship holds we should be able to come up with a very good prediction for the logarithm of the probability (and hence by exponentiating obtain the probability). We do this in the next section.

4.3. **Statistical Inference.** Let's try to predict the answer for large values of k from smaller ones. The fifth named author gave a talk on this at the 110th meeting of the Association of Teachers of
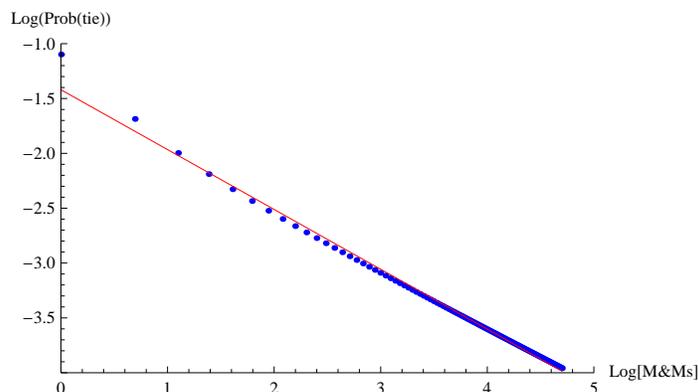
10

FIGURE 6. The probability of a tie for $k \leq 110$. The best fit line is good, but is noticeable non-perfect.

Mathematics in Massachusetts in March 2013, which will explain the prevalence of 110 and 220 below.

Figure 6 gives the log-log plot for $k \leq 110$. Using the Method of Least Squares from Statistics[7] with $P(k)$ the probability of a tie when we start with k M&M'S, we find a predicted best fit line of

$$\log\left(P(k)\right) \approx -1.42022 - 0.545568 \log k,$$

or exponentiating

$$P(k) \approx 0.2412/k^{.5456}.$$

This predicts a probability of a tie when $k = 220$ of about 0.01274, but the answer is approximately 0.0137. While we are close, we are off by a significant amount. (In situations like this it is better to look at not the difference in probabilities, which is small, but the percentage we are off; here we differ by about 10%.)

Why are we so far off? The reason is that small values of k are affecting our prediction more than the should. If we have a main term in the log-log plot which is linear, it will eventually dominate lower order terms *but* those lower order terms could have a sizable effect for low k. Thus, it's a good idea to ignore the smaller values when extrapolating our best fit line.

In Figure 7 we now go from $k = 50$ to 110.

Our new best fit line is

$$\log\left(P(k)\right) \approx -1.58261 - 0.50553 \log k,$$

or exponentiating

$$P(k) \approx 0.205437/k^{.50553};$$

we should compare this to our previous prediction of $0.241662/k^{.5456}$). Using our new formula we predict 0.01344 for $k = 220$, which compares *very* favorably to the true answer of 0.01347.

The point of this section is to give you a brief introduction to the power of statistics and extrapolating, and give you a sense of some of the issues in their use. We are able to get a fairly reasonable prediction with very little work, and if we clean up the data a little we improve to a phenomenal agreement.

---

[7]These formulas can be derived using multivariable calculus and linear algebra. For a derivation, see for example http://web.williams.edu/Mathematics/sjmiller/public_html/105Sp10/handouts/MethodLeastSquares.pdf.
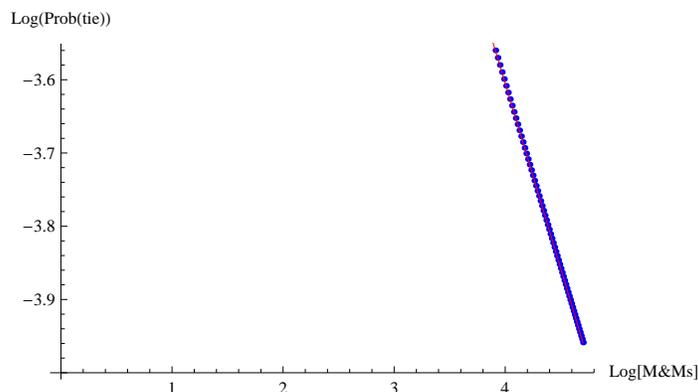
FIGURE 7. The probability of a tie for $50 \le k \le 110$. The best fit line is almost indistinguishable from the data.

5. RECURRENCES

As our goal is not to write a book on this game but rather to show how it leads to many good mathematical items, in the interest of space we will just briefly discuss two final approaches, recurrence relations in this section and hypergeometric functions in the next. See [**?**] for a quick introduction to recurrences.

5.1. Recurrence Review.
If you've seen the Fibonacci numbers

$$\{F_n\}_{n=0}^\infty \;=\; \{0, 1, 1, 2, 3, 5, 8, \dots\},$$

you've seen a terrific example of a linear recurrence with constant coefficients, as they are the unique sequence satisfying

$$F_{n+2} \;=\; F_{n+1} + F_n, \quad F_0 \;=\; 0, \quad F_1 \;=\; 1.$$

Once we know the relation and the first two coefficients, we can find any subsequent value by substituting. Unfortunately, this is costly in practice, as computing the nth term requires us to know all the previous ones.

Fortunately, there are many ways to efficiently solve problems like this, and often these lead to beautiful closed form expressions. To solve the Fibonacci relation we guess $F_n = r^n$. Why is this reasonable? Clearly the Fibonacci sequence is non-decreasing, so $F_{n+2} \le 2F_{n+1}$, which means that every time we increase the index by 1 we at most double our number, so $F_n \le 2^n$. Similarly we find $F_{n+2} \ge 2F_n$; now increasing the index by 2 causes us to at least double, so increasing the index by 1 should yield an increase of at least a factor of $\sqrt{2}$. Thus we expect the Fibonaccis to satisfy a relation such as

$$2^{n/2} \;\le\; F_n \;\le\; 2^n,$$

which is highly suggestive of exponential growth; this is why we try $F_n = r^n$.

Substituting this into the recurrence we obtain the characteristic polynomial for r, which, after dividing both sides by r^n, is

$$r^2 \;=\; r + 1;$$

the solutions to this are

$$r_1 \;=\; \frac{1 + \sqrt{5}}{2}, \quad r_2 \;=\; \frac{1 - \sqrt{5}}{2}.$$

12

A beautiful property of linear recurrences is that an arbitrary linear combination of solutions is a solution, and we find the general solution of the Fibonacci recurrence is

$$F_n = c_1 r_1^n + c_2 r_2^n.$$

As we require $F_0 = 0$ and $F_1 = 1$, after some more algebra we obtain Binet's Formula, the spectacular relation

$$F_n = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n.$$

This formula is amazing: it gives us a simple, closed form expression for the nth Fibonacci number; we can jump to this term in the sequence *without* computing any of the earlier ones![8]

The point of the above is to give a brief glimpse at the rich theory; there is far more that could be said (especially concerning generating function approaches to solve problems such as these), but for our purposes this suffices. The main takeaways is that sometimes we are lucky and able to derive simple closed form expressions, but even if we cannot we are often able to determine the terms by repeated application of the defining relation and initial conditions.

5.2. **The M&M Recurrence.** Even though we have a finite sum for the probability of a tie (equation 3.1), finding that required some knowledge of combinatorics and binomial coefficients. We give an alternate approach which avoids these ideas. It's possible to do it with or without noting that we have a memoryless process. We'll do the memoryless process first as we'll assume we're still clever enough to notice that, and then remark afterwards how we would have found the same formula even if we didn't realize this.

We need to consider a more general problem. We always denote the number of M&M'S Cam has with c, and Kayla with k[9]; we frequently denote this state by (c, k). Then we can rewrite the three equally likely scenarios, each with probability 1/3, as follows:

- $(c, k) \longrightarrow (c - 1, k - 1)$ (double heads and both eat),
- $(c, k) \longrightarrow (c - 1, k)$ (Cam gets a head and Kayla a tail),
- $(c, k) \longrightarrow (c, k - 1)$ (Cam gets a tail and Kayla a head).

If we let $x_{c,k}$ denote the probability the game ends in a tie when we start with Cam having c M&M'S and Kayla having k, we can use the above to set up a recurrence relation. How so? Effectively, on each turn we move from (c, k) in exactly one of the following three ways: either Cam and Kayla both eat an M&M in which case $(c, k) \longrightarrow (c - 1, k - 1)$; only Cam flips heads and eats an M&M in which case $(c, k) \longrightarrow (c - 1, k)$; or Kayla is the only one to eat an M&M in which case $(c, k) \longrightarrow (c, k - 1)$. Now, we can use simpler game states to figure out how the probability of a tie when we start with more M&M, as in each of the three cases we have reduced the total number of M&M'S by at least one. We thus find that the recurrence relation satisfied by $\{x_{c,k}\}$ is

$$x_{c,k} = \frac{1}{3} x_{c-1,k-1} + \frac{1}{3} x_{c-1,k} + \frac{1}{3} x_{c,k-1} = \frac{x_{c-1,k-1} + x_{c-1,k} + x_{c,k-1}}{3}. \tag{5.1}$$

---

[8]Additionally, as the Fibonaccis are all integers Binet's formula must return an integer; at first this might seem unlikely, as our expression involves square-roots and fractions, but fortunately everything that needs to cancel does.

[9]We can see the power of good notation. Using c and k to represent the number of M&M'S Cam and Kayla each have allows us to know exactly what is going on as the math gets more involved.

From our work on the Fibonacci numbers we know that cannot be the full story – we need to specify initial conditions. A little thought says $x_{0,0}$ must be 1 (if they both have no M&M'S then it must be a tie), while $x_{c,0} = 0$ if $c > 0$ and similarly $x_{0,k} = 0$ if $k > 0$ (as in these cases exactly one of them has an M&M, and thus the game cannot end in a tie).

We have made tremendous progress. We use these initial values and the recurrence relation (5.1) to determine $x_{c,k}$. Unfortunately we cannot get a simple closed form expression, but we can easily compute the values by recursion. A good approach is to compute all $x_{c,k}$ where $c + k$ equals sum fixed sum s. We've already done the cases $s = 0$ and $s = 1$, finding $x_{0,0} = 1$, $x_{0,1} = x_{1,0} = 0$.

We now move to $s = 2$. We need only find $x_{1,1}$, as we know $x_{2,0} = x_{0,2} = 0$. Using the recurrence relation we find

$$x_{1,1} = \frac{x_{0,0} + x_{0,1} + x_{1,0}}{3} = \frac{1+0+0}{3} = \frac{1}{3}.$$

Next is the case when the indices sum to 3. Of course, $x_{0,3} = x_{3,0} = 0$, so all we need are $x_{1,2}$ and $x_{2,1}$ (which by symmetry are the same). We find

$$x_{2,1} = x_{1,2} = \frac{x_{1,1} + x_{2,0} + x_{0,2}}{3} = \frac{1/3+0+0}{3} = \frac{1}{9}.$$

We can continue to $s = 4$, and after some algebra easily obtain

$$x_{2,2} = \frac{x_{1,1} + x_{2,1} + x_{1,2}}{3} = \frac{5}{27}.$$

If we continued on with these calculations, we would find that $x_{3,3} = \frac{11}{81}$, $x_{4,4} = \frac{245}{2187}$, $x_{5,5} = \frac{1921}{19863}$, $x_{575,6561} = \frac{11}{81}$, $x_{7,7} = \frac{42635}{531441}$, and $x_{8,8} = \frac{355975}{4782969}$. The beauty of this recursion process is that we have a sure-fire way to figure out the probability of a tie at different states of the M&M game. We leave it as an exercise to the interested reader to compare the computational difficulty of finding $x_{100,100}$ by the recurrence relation versus by the finite sum (3.1).

We end with one final comment on this approach. It's possible to recast this problem as one in counting paths on a graph. In Figure 8 we start with $(c, k) = (4, 4)$, and look at all the possible paths that end in $(0, 0)$. The probability of any path is equal to $(1/3)^t$, where t is the number of terms in the path. It turns out the solution is very similar to the famous Catalan numbers, which count the number of paths from $(0, 0)$ to (n, n) moving in unit horizontal or unit vertical steps and never going above the main diagonal; the difference here is that we now have three possible choices at each turn.

5.3. **Forgetting Memoryless Processes.** In the previous subsection we found a recurrence relation for $x_{c,k}$, but our analysis was based on there only being three options at each step. What if we hadn't noticed there was a memoryless process lurking: would we still have found the same relation? In that case, there would now be four possibilities on each turn, each happening with probability 1/4.

- $(c, k) \longrightarrow (c - 1, k - 1)$ (double heads and both eat),
- $(c, k) \longrightarrow (c - 1, k)$ (Cam gets a head and Kayla a tail),
- $(c, k) \longrightarrow (c, k - 1)$ (Cam gets a tail and Kayla a head),
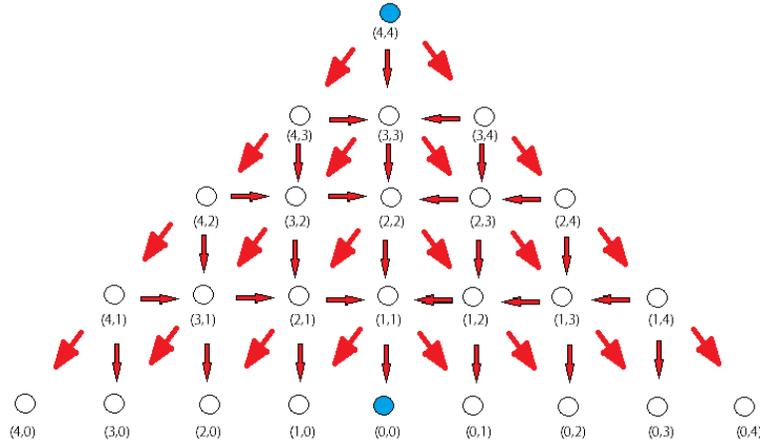- $(c, k) \longrightarrow (c, k)$ (double tails and neither eats).

FIGURE 8. The M&M game when $k = 4$. Count the paths! Answer 1/3 of probability hit (1,1).

We now obtain the following relation:

$$x_{c,k} = \frac{1}{4}x_{c-1,k-1} + \frac{1}{4}x_{c-1,k} + \frac{1}{4}x_{c,k-1} + \frac{1}{4}x_{c,k}.$$

Note that if we bring the $\frac{1}{4}x_{c,k}$ over to the left hand side we relate $\frac{3}{4}x_{c,k}$ to multiples of $x_{c-1,k-1}, x_{c-1,k}$ and $x_{c,k-1}$:

$$\frac{3}{4}x_{c,k} = \frac{1}{4}x_{c-1,k-1} + \frac{1}{4}x_{c-1,k} + \frac{1}{4}x_{c,k-1}.$$

If we then multiply through by $4/3$ we regain our old recurrence, equation (5.1):

$$x_{c,k} = \frac{1}{3}x_{c-1,k-1} + \frac{1}{3}x_{c-1,k} + \frac{1}{3}x_{c,k-1}.$$

This is wonderful: it means if we did not initially notice that there was a memoryless process, doing the algebra suggests there should be one!

5.4. **Revisiting and Generalizing the Hoops Game.** When you learn new concepts in math, it often pays great dividends to revisit earlier problems. Let's go back to the hoops game; not surprisingly, we'll see that, similar to the M&M Game, we can cast it as a double recurrence.

The way the recurrence method worked was we reduced the problem we wished to study to a simpler state; however, that required us to know those answers. Thus it is not enough to deal with just x_B in general, and we need to study $x_{B;b,m}$, which is the probability Bird wins when he needs to make b more baskets to win, Magic needs to make m more to win, *and Bird is currently shooting!*.[10] We find

$$\begin{aligned} x_{B;b,m} = \quad & p_B p_M x_{B;b-1,m-1} + p_B(1-p_M)x_{B;b-1,m} \\ & + (1-p_B)p_M x_{B;b,m-1} + (1-p_B)(1-p_M)x_{B;b,m}. \end{aligned} \tag{5.2}$$

To see why this is true, let's look at the first term. The p_B means Bird got a basket, reducing the number he needs by 1. Note that if $b - 1 = 0$ then Bird wins and the game should stop (we'll deal more with this in a moment). Now it's Magic's turn to shoot. If he gets a basket, which

---

[10]Another option is that we could introduce $x_{M;b,m}$, the corresponding probability where now Magic is shooting.

happens with probability p_M, that reduces his number of baskets needed to $m - 1$, which explains the $p_B p_M x_{B;b-1,m-1}$ term; the other three terms arise from the other possibilities.

We also need the initial conditions. Clearly $x_{B;b,0} = 0$ if $b > 0$ and $x_{B;0,m} = 1$ if $m > 0$, but what should we choose for $x_{B;0,0}$? Well, the way to interpret this is that each needs to make zero baskets and Bird shoots first, so he is the first to reach zero baskets. Thus we set $x_{B;0,0}$ equal to 1. Another justification is that we only reach this situation when Bird makes a basket and then Magic, who shouldn't have been allowed to shoot as Bird just won the game, shoots. These normalizations often are tricky, but can frequently be determined by a good story. This is similar to the convention that $0! = 1$, which we interpret as there is only one way to do nothing (i.e., there is one way to order no elements – there shouldn't be multiple ways to do nothing!).

We have thus found a recurrence for $x_{B;b,m}$. Let's check and make sure it reduces to our previous result when $b = m = 1$. In that case, equation (5.2) becomes

$$x_{B;1,1} = p_B p_M 1 + p_B (1 - p_M) 1 + (1 - p_B) p_M 0 + (1 - p_B)(1 - p_M) x_{B;1,1}.$$

Remembering that we use r for $(1 - p_B)(1 - p_M)$ after some simple algebra we obtain

$$x_{B;1,1} = \frac{p_B}{1 - r},$$

exactly as before! Of course, this now suggests a natural question: what does $x_{B;b,b}$ look like as b grows (let's say under the assumption that $p_B = p_M$)?

6. HYPERGEOMETRIC FUNCTIONS

We end our tour of solution approaches with a method that actually prefers the infinite sum to the finite one, hypergeometric functions (see for example [AS, GR]). These functions arise as the solution of a particular linear second order differential equation:

$$x(1 - x)y''(x) + [c - (1 - a + b)x]y'(x) - aby(x) = 0$$

(this is also called Gauss's differential equation). This equation is useful because every other linear second order differential equation with three singular points (in the case they are at 0, 1, and ∞) can be transformed into it. As this is a second order differential equation there should be two solutions. One is

$$y(x) = 1 + \frac{abx}{c1!} + \frac{a(a+1)b(b+1)x^2}{c(c+1)2!} + \frac{a(a+1)(a+2)b(b+1)(b+2)x^3}{c(c+1)(c+2)3!} + \cdots,$$

so long as c is not a non-positive integer; we denote this solution by $_2F_1(a, b; c; z)$. By choosing appropriate values of a, b and c we recover many special functions. Wikipedia lists three nice examples:

$$\log(1+x) = x \, _2F_1(1, 1; 2; -x), \quad (1-x)^{-a} = \, _2F_1(a, 1; 1; x), \quad \arcsin(x) = x \, _2F_1(1/2, 1/2; 3/2; x^2).$$

By introducing some notation we can write the series expansion more concisely. We define the Pochhammer symbol by

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1) = \frac{(a+n-1)!}{(a-1)!}$$

(where the last equality holds for integer a; for real a we need to interpret the factorial as its completion, the Gamma function). Our solution becomes

$$_2F_1(a, b, c; x) \;=\; \sum_{n=0}^{\infty} \frac{(a)_n (b)_n x^n}{(c)_n n!}.$$

Note the factorials in the above expression suggest that there should be connections between hypergeometric functions and products of binomial coefficients. In this notation, the 2 represents the number of Pochhammer symbols in the numerator, the 1 the number of Pochhammer symbols in the denominator, and the a, b, and c are what we evaluate the symbols at (the first two are the ones in the numerator, the last the denominator). One could of course consider more general functions, such as

$$_sF_t(\{a_i\}, \{b_j\}; x) \;=\; \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_s)_n x^n}{(b_1)_n \cdots (b_t)_n n!}.$$

The solution $_2F_1(a, b, c; x)$ is called a hypergeometric function, and if you look closely at it while recalling the infinite sum solution to the M&M Game you might see the connection. After some algebra where we convert the binomial coefficients in the infinite sum solution (1.1) to the falling factorials that are the Pochhammer symbols, we find the following closed form solution.

Theorem 6.1. *The probability the M&M Game ends in a tie with two people using fair coins and starting with k M&M'S is*

$$P(k, k) \;=\; {}_2F_1(k, k, 1; 1/4) 4^{-k}. \tag{6.1}$$

It is not immediately clear that this is progress; after all, it looks like we've just given a fancy name to our infinite sum. Fortunately, special values of hypergeometric functions are well studied, and a lot is known about their behavior as a function of their parameters. We encourage the interested reader to explore the literature and discover how 'useful' the above is.

7. OEIS

We end with a short bonus section on how to guess formulas. There is an enormous wealth of information available on-line, but often it is hard to figure out what we need and where it resides. A terrific resource is the On-Line Encyclopedia of Integer Sequences (OEIS, http://oeis.org/). This is a wonderful resource with a large number of integer sequences tabulated and stored. You enter some known terms in your sequence, and the site not only tells you what sequences it knows that agree with this, but it provides links, properties and formulas when it can!

For example, if we use our finite series expansion (3.1) or the recurrence relation (5.1) we can easily calculate the probability of a tie for some small k. We give the probabilities for k up to 8 in Table 1. In addition, we also give $3^{2k-1} P(k, k)$. The reason we do this is that looking at the probability of a tie one is struck by the fact that the denominators are all powers of 3; after a little algebra we see that if we multiply by 3^{2k-1} we clear the denominators, and we will obtain a sequence of *integers*. Note that it is very important that we end with integers and not rational numbers if we wish to use the OEIS.

Remark 7.1. *If we didn't notice the right power of 3, we could have reached the same conclusion another way. There are three possibilities each time; from Figure 8 we saw our problem is equivalent to counting how many paths there are from (k, k) to $(0, 0)$. As we end at $(0, 0)$ our last*

k	$P(k,k)$	$3^{2k-1}P(k,k)$
1	1/3	1
2	5/27	5
3	11/81	33
4	245/2187	245
5	1921/19683	1921
6	575/6561	15525
7	42635/531441	127905
8	355975/4782969	1067925

TABLE 1. Probability of a tie as a function of the number of M&M'S the two players have.

step is forced, and the longest path comes when we never get double heads. As we must remove $2k - 2$ M&M'S (remember the last toss of double heads removed 2 of the $2k$), the longest path has $2k - 2 + 1 = 2k - 1$ steps, explaining the presence of this factor as the exponent of 3.

Thus to the M&M Game with two players we can associate the integer sequence 1, 5, 33, 245, 1921, 15525, 127905, 1067925, ... We plug that into the OEIS and find that it knows that sequence! It is sequence A084771 (see http://oeis.org/A084771). The very first comment there on this sequence is that it equals the number of paths in the graph we discussed!

The OEIS is a powerful tool for research. Think back to proofs by induction: if you are told what to prove, it is a lot easier and often the proof writes itself. The OEIS frequently gives you such an advantage.

8. TAKEAWAYS AND FURTHER QUESTIONS

We've seen many different ways of solving the M&M Game, each leading to a different important aspect of mathematics. We end with some quick reflections on some of the valuable lessons this game has to offer.

(1) **Ask Questions:** Great mathematics is everywhere, waiting to be realized and explored. Often some of the deepest mathematics can be extracted from some of the most straightforward problems to state (for an excellent example see Conway's *See and Say* (or *Look and Say*) sequence.

(2) **There are Many Ways to Solve Problems:** Different ways of solving the problem have different advantages. With the recurrence relation, we can calculate the answer for any given number of M&Ms, but it might take a long time. With hypergeometric functions, we get a nice closed-form way of representing our solution. And with a memoryless process we get a finite sum involving well-understood binomial coefficients.

(3) **Experience is a Useful and Great Guide:** We were able to make enormous leaps in the M&M problem because we had the hoops game as a reference. So much of math is interconnected that a lot of times, all it takes for us to solve a difficult problem is remembering what we have done in the past.

(4) **Need to Look at Data the Right Way:** When we looked at the numbers properly, we were able to make progress in solving the M&M problem. This is no different for other math problems: so often, all it takes to figure out a complicated problem is the right lens.

All of these takeaways are great, and they should lead to one last exercise: asking further questions. How long do we expect a game to take? What would happen to the M&M problem if we increased the number of players? What about if all of the players started with different numbers of M&M'S? Maybe the game would yield interesting results if the participants used biased coins.

In one of the first games ever played, Cameron, Kayla and Steven Miller each started with five M&M'S and Kayla tossed five consecutive heads, dying immediately; years later she still talks about that memorable performance. There is a lot known about the longest run of heads or tails in tosses of a fair (or biased) coin (see for example [Sch]). We can ask related questions here. What is the expected longest run of heads or tails by any player in a game? What is the expected longest run of tosses where all players' coins have the same outcome?

We could also revisit the hoops game and consider generalizations there. What if Bird and Magic keep shooting until someone makes k baskets. What's the probability of a tie now? What if you keep shooting until you miss? We could also ask questions about streaks of hits and misses within the game. For another possibility, what if Bird's probability of making a basket and Magic's probability of a basket are independent random variables drawn uniformly[11] on $[0, 1]$: what is the probability that Bird has a greater chance of winning than Magic? If you look at Figure 3 this problem with $k = 1$ is equivalent to finding the area in the unit square above and to the left of the contour line with value .5; the answer turns out to be $\log(2) \approx .693147$. Is there a nice answer for general k? What if instead we ask what is the probability Bird wins if p_B and p_M are independently drawn uniform random variables on $[0, 1]$? If $k = 1$ the answer is $\pi^2/6 - 1 \approx .644934$. It's nice that in both phrasings the answers are interesting numbers, and that the two different interpretations are quite close.

There are plenty of further questions out there, all of which would provide great insights not only into the M&M game and its educational value but also into the study of math in general. We hope you will explore some of these or, even better, ones of your own choosing, and let us know what you find!

REFERENCES

[AS]   M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 1965. Online at http://people.math.sfu.ca/~cbm/aands/.
[GR]   I. Gradshteyn and I. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, New York, 1965.
[M]   S. J. Miller, *The Probability Lifesaver*, Princeton University Press, Princeton, NJ, 2016.
[Sch]   M. Schilling, *The longest run of heads*, College Math. J. **21** (1990), no. 3, 196–207.

E-mail address: ivan.badinski@gmail.com

DEPARTMENT OF ECONOMICS, MIT, CAMBRIDGE, MA 02139

E-mail address: ckh1@williams.edu

---

[11]This means that for any interval $[a, b] \subset [0, 1]$, the probability $p_B \in [a, b]$ is $b - a$, similarly for p_M, and knowledge of p_B gives no information on p_M (or vice-versa).

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

E-mail address: nrm2@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

E-mail address: sjm1@williams.edu,Steven.Miller.MC.96@aya.yale.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

E-mail address: ms14@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267