Chapter 23

Difference Equations, Markov Processes and Probability

Loesje: `Domino effect: Once you drop a good idea, the rest will follow.`

You might not have known this when you purchased this book, but as an added bonus I'm going to share a wonderful strategy to win at Roulette. You can make millions with no risk. In fact, as soon as I finish writing this chapter (as I'm so altruistic I want to share this secret with you), I'll be flying back to Vegas to win some more....

Sadly, a lot of people fall for scams like the above. In this chapter we'll talk about what looks like a sure, safe bet, and show why it isn't. What I like about this problem is that it's connected to a lot of great math, and can be understood without a huge amount of mathematical (or gambling) pre-requisites. Specifically, we'll see how some real world problems can be modeled by **recurrence relations**. We'll quickly develop just enough of the theory to solve a few interesting problems, and end the chapter with a short primer on the subject for those who want more. Often these topics are covered in courses on discrete mathematics or differential equations; the reason they fit in this book is they can be used to compute interesting probabilities.

23.1 From the Fibonacci Numbers to Roulette

The goal of this section is to understand a very popular strategy for Roulette, and connect it to some mathematics you hopefully have seen before, the Fibonacci numbers.

23.1.1 The Double-plus-one strategy

To simplify our discussion, we'll talk about an easier version of Roulette (see Figure 23.1). We'll assume that every time the wheel spins the ball either lands on a red or a

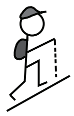
Figure 23.1: A roulette wheel (image from Toni Lozano).

black number, and each outcome happens 50% of the time. The actual game is a bit more complicated, but the strategy we describe below would work in that case too. (The real game usually has 18 red, 18 black and 2 green places, so that red and black each occur about 47.37% of the time.) To make life easy, we're only allowing bets on red or black, and we're eliminating the two green numbers (the greens provide a huge advantage to the casino). Say we bet \$1 on red (if we bet on black the result is similar). If red comes up we win \$1; this means we get back our original dollar plus an additional one. If, however, black comes up then we lose our dollar.

Obviously, our goal is to make money. Here's a famous strategy, called **Double-plus-one**. Bet \$1 on red. If it comes up red, great, we're up a dollar. If not, we're down a dollar and now bet \$2. If we win, we're now up a grand total of one dollar. What if we lose? If we lose, we're now down \$3. In this case, we bet \$4. If we win, we're now up a dollar (we lost \$3 previously and just one a dollar), while if we lose we're down \$7, and now we bet \$8.

Hopefully the pattern is clear. We keep doubling our bet until we win. When we win, we recoup all our losses and an extra dollar. As *eventually* a red should turn up, *eventually* we should be up a dollar. We then just keep repeating until we've made whatever amount we desire.

What's wrong with this? There are two problems; one requires just some common sense, while the other requires knowing a bit how Vegas works (and why they listen to mathematicians!). The first issue, of course, is that at some point we may need to bet \$1,267,650,600,228,229,401,496,703,205,376 (or 2^{100} dollars), and we 'may' not have that much money! In order not to worry about such 'trivialities', we assume the existence of a rich, but very eccentric, aunt or uncle. This kind family member has unlimited financial reserves, and will advance us whatever amount of



money we need to cover our bets, but won't just give us a dollar directly. Why won't they just give us a dollar? That's beyond the scope of this book – we just focus on the mathematics here! The purpose of assuming a rich, eccentric aunt or uncle is to remove the difficulty of needing a large bankroll for the problem, though after we analyze the problem I urge you to modify the argument in the case when you have a fixed, finite amount of money.

What's the other problem? This one turns out to be far more serious. We haven't talked too much about how the bets can be done. It turns out that each casino sets both lower and *upper* bounds on how much you can wager on a given spin. For example, the lower limit may be \$1 and the upper limit might be \$30. If this is the case, if the first five spins are black we're in trouble. If that happens, we've lost $1 + 2 + 4 + 8 + 16 = 31$ dollars. Our method tells us to bet \$32, but we can only bet \$30, and our system breaks down. We're in even more trouble if we get another black. The problem is that when we win, we win small, but when we lose, we lose *big*.

This should suggest the following natural, and very important, problem: *If we play n times, what's the chance we get 5 or more consecutive blacks?* Interestingly, the same mathematics that we can use to study the Fibonacci numbers can be applied to solve this problem, too. We'll therefore pause and quickly review the Fibonacci numbers, and then return to Roulette.

23.1.2 A quick review of the Fibonacci numbers

Let's briefly recall the Fibonacci numbers, though at first there doesn't seem to be any connection. The Fibonacci numbers are the sequence $F_0 = 0$, $F_1 = 1$, $F_2 = 1$, $F_3 = 3$, $F_4 = 5$, $F_5 = 8$, and in general $F_{n+2} = F_{n+1} + F_n$. This is an example of a **linear recurrence relation** (also called a **difference equation**). It's linear as the unknown term depends linearly on previous terms; note we don't have terms multiplying each other, or exponentials of terms. There are many ways to solve this. A great approach is through generating functions (see §19.2 for such a proof), but in the interest of time and to make the exposition self-contained we'll now give the proof by **Divine Inspiration**. Essentially, the way this works is you guess the answer, and see that you're right! Obviously the trouble is that, in general, it's hard to just guess the answer to a difficult math problem! What saves the method is that there's actually a large class of problems where we can just look and rightly guess. For those who want to see more of the general theory, just read on to §23.2.

Let's try $F_n = r^n$ for some r. This is a reasonable guess. It means each term is r times the previous. If we had the simpler relation $G_{n+1} = 2G_n$ then the solution is $G_n = 2^n$, as each term is 2 times the previous. We'll expand on this idea later. If we substitute our guess into the recurrence $F_{n+2} = F_{n+1} + F_n$, we get $r^{n+2} = r^{n+1} + r^n$. This simplifies to $r^2 = r + 1$, or $r^2 - r - 1 = 0$, which by the quadratic formula has two roots: $r_1 = (1 + \sqrt{5})/2 \approx 1.618$ and $r_2 = (1 - \sqrt{5})/2 \approx -.618$. The polynomial $r^2 - r - 1$ is called the **characteristic polynomial of the recurrence relation**.

It turns out that for a linear recurrence relation, any linear combination of solutions of the characteristic polynomial is a solution to the recurrence. In other words, if you plug in $F_{n+2} = c_1 r_1^n + c_2 r_2^n$ for *any* choice of c_1 and c_2, you'll find it solves the recurrence relation because r_1 and r_2 solve the characteristic polynomial; it's a good idea to check this to get a feel for how linearity helps. While we can use any choice of c_1 and c_2, we want our sequence to start off with a 0 when $n = 0$ and a 1 when $n = 1$. In other words, $c_1 + c_2 = 0$ and $c_1 r_1 + c_2 r_2 = 1$. Solving for c_1 and c_2 we find $c_1 = -c_2 = 1/\sqrt{5}$, and

Binet's formula. Let $F_{n+2} = F_{n+1} + F_n$, with $F_0 = F_1 = 1$. Then

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

Don't worry if this is a bit incomprehensible right now. We'll talk about recurrence relations in general, and the Fibonacci numbers in particular, in more detail and more leisurely below. Right now, all that matters is you leave this problem knowing that there exists a method to solve linear recurrence relations. Binet's formula is very efficient. It allows us to jump forward and calculate F_{100} without going through all the intermediate terms. While it's of course nice to avoid tedious algebra, if we didn't know the advanced theory we could compute F_{100}, assuming we're very patient. We just keep using the recurrence relation $F_{n+2} = F_{n+1} + F_n$ to find more and more terms, eventually getting $F_{100} = 354, 224, 848, 179, 261, 915, 075$.

It's worth commenting a bit on the Divine Inspiration; what made us think that $a_n = r^n$ would be a good guess? Here's one argument that suggests this is a good thing to try. The Fibonacci series is strictly increasing, so $F_{n-2} < F_{n-1} < F_n$. As $F_n = F_{n-1} + F_{n-2}$, we have

$$2F_{n-2} \;<\; F_n \;<\; 2F_{n-1}.$$

After some algebra, we see $F_n < 2^n$. The lower bound is a bit harder. From $2F_{n-2} < F_n$, we see that every time the index increases by 2, our Fibonacci number at least doubles. Continuing this line backwards, we get

$$F_n \;>\; 2F_{n-2} \;>\; 2^2 F_{n-4} \;>\; 2^3 F_{n-6} \;>\; \cdots \;>\; 2^{n/2} F_0$$

(at least if n is even). In other words, $F_n > 2^{n/2} = (\sqrt{2})^n$. We've sandwiched the n^{th} Fibonacci number between two exponential bounds; it grows at least as fast as $(\sqrt{2})^n$, and at most as fast as 2^n. It's thus reasonable to guess it grows like r^n for some r. For large n Binet's formula says F_{n+1} is approximately $\frac{1+\sqrt{5}}{2}$ larger than F_n; note this constant is about 1.61803, sandwiched beautifully between our lower bound of $\sqrt{2} \approx 1.414$ and our upper bound of 2.

23.1.3 Recurrence Relations and Probability

Why are recurrence relations helpful for our roulette problem? Let's try to compute the probability that, in n spins of the wheel, we have at least 5 consecutive blacks. We'll call this probability a_n. It turns out to be easier to compute b_n, the probability that in n spins we *do not* have at least 5 consecutive blacks. Note that a_n is just $1 - b_n$, so if we can find one we can surely find the other. This is a powerful principle in probability, namely that complementary events have probabilities summing to 1.

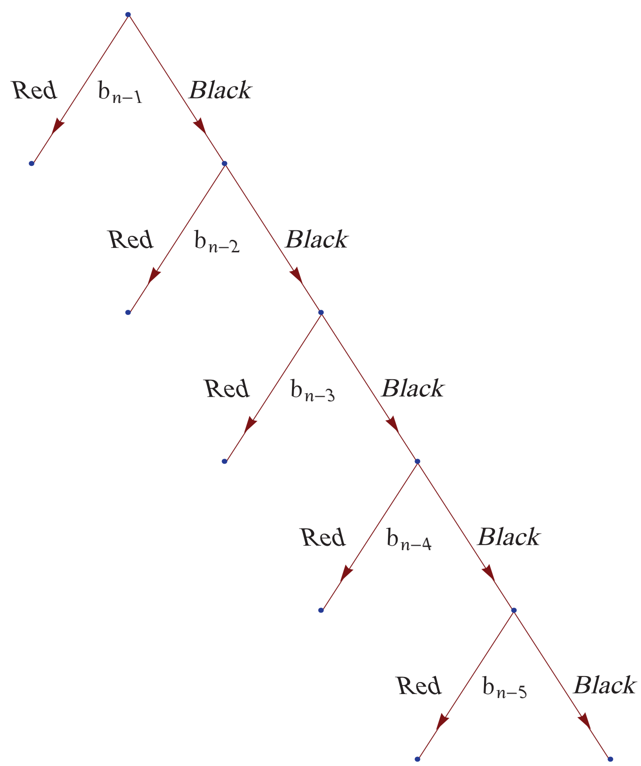


Figure 23.2: Developing the recurrence relation for not having 5 consecutive blacks in n spins of our roulette wheel.

There are many names for this, including **Law of Total Probability**. Let's use this to get a recurrence relation for b_n. We sketch what happens as we spin in Figure 23.2.

What is b_n? Well, there are two possibilities for the first spin, and each happens with probability 1/2. Half the time we get a red, half the time we get a black. What is the probability we do not have 5 consecutive black spins in n spins, *given that the first spin is a red?* The answer to this question is just b_{n-1}; since the first spin is a red, it can't contribute to 5 consecutive blacks. We now analyze the branch coming from a first spin of black. There are two possibilities for the second spin, again each happening half the time: a red spin, a black spin. If we start off black then red, which happens $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ of the time, then the probability that we don't have 5 consecutive blacks is just b_{n-2}.

Continuing along these lines, we find

$$b_n \;=\; \frac{1}{2}b_{n-1} + \frac{1}{4}b_{n-2} + \frac{1}{8}b_{n-3} + \frac{1}{16}b_{n-4} + \frac{1}{32}b_{n-5}.$$

Why do we stop here, why aren't there more terms? Well, if we start off with five consecutive black spins, then there's no chance that we won't have 5 consecutive black spins! It's precisely for this reason that we're trying to find b_n and not a_n. We now have the recurrence relation. All that remains is to find the initial conditions. This isn't too bad; it's just

$$b_0 \;=\; b_1 \;=\; b_2 \;=\; b_3 \;=\; b_4 \;=\; 1.$$

Why are each of these 1? If we have fewer than 5 spins, we can't have at least 5 consecutive blacks! We can either modify the advanced theory or just use the recurrence relation to find the b_n's or the a_n's. After some algebra, we find the a_n's are

$$0, 0, 0, 0, 0, \frac{1}{32}, \frac{3}{64}, \frac{1}{16}, \frac{5}{64}, \frac{3}{32}, \frac{7}{64}, \frac{255}{2048}, \frac{571}{4096}, \dots,$$

or in decimal form,

$$0, 0, 0, 0, 0, 0.03125, 0.046875, 0.0625, 0.078125, 0.09375, 0.109375, 0.124512, \dots.$$

By the time we get to $n = 100$, there is an 81.01% chance that we'll have at least 5 consecutive blacks. At $n = 200$ the probability climbs to 96.59%, while at $n = 400$ it's 99.89%.

23.1.4 Discussion and Generalizations

Our roulette problem has a lot of beautiful features. We can extract a nice mathematical formulation from it which we can solve. Without too much trouble, we can write a simple program to use the recurrence relation and initial conditions to find the probabilities. This illustrates just a small subset of the different types of math that can arise in a probability problem. It also shows the importance of looking at the right object; the recurrence relation is a bit cleaner if we go for the probability of not having 5 consecutive blacks, rather than what we desire (namely the probability of having at least 5 consecutive blacks).

We end with one final feature about this problem. Say we desire the probability of not getting 5 consecutive blacks in 100 spins. We saw we could set up the recurrence

ERROR: typecheck
OFFENDING COMMAND: image

STACK:

-dictionary-
-mark-
-savelevel-