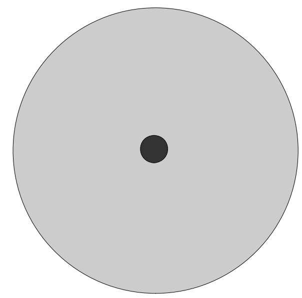
Darts

The purpose of this section is to uncover a fundamental counting result. This is a result that comes up again and again in my encounters with high school students who have a simple problem they're trying to solve and in my work with graduate students modeling random processes. Some think it's obvious and others can see why it's true. I call it "the dartboard theorem" but that's just <u>my</u> name. The point is that you're trying to do something again and again and each time you have the same probability of success. How many times do you have to try?

I'm throwing darts at a dart board. Suppose that, on each throw, I have probability 1/20 of hitting the bulls-eye. The problem is: how many darts do I have to throw to get the first bulls-eye?



Of course, on different occasions, I'll need different numbers of throws. I might be real lucky and get it on the first or second throw. Or I might be unlucky and need something like 100 throws. So the question to ask is this—What is the *average* number of throws required?

We start with some empirical work. Actually I'm going to "stand a little closer to the dart board"—instead of a probability of 1/20, I'll suppose it's 1/6. Then we can use dice instead of darts, and a bullseye will be signaled by the rolling of a **6**.

<u># Rolls to get first 6</u>	
<u>#Rolls</u>	*<u>#Trials</u>*
1	14
2	11
3	9
4	6
5	5
6	9
7	5
8	2
9	3
10	4
11	1
12	2
13	3
15	1
18	2
22	1
23	1
28	1
Total	80

There are forty students and I get them each to perform the experiment twice—roll the die as many times as needed to get the first six, and record the number of rolls required, and then repeat. I tabulate the results—of the 80 trials, 14 got the six on the very first roll, 11 got their first six on the second roll, and so forth. The "record" went to one student who rolled 28 times to get his first six.

The sum of the second column is the total number of experiments which is 80. Note that some of the possible outcomes (e.g. 14 and 16) didn't occur at all, but on another occasion likely would.

Now the average # of rolls per experiment is found by adding up all the outcomes we obtained and dividing by the total number of experiments. That is, we have to add 14 1's, and 11 2's, and 9 3's, etc. and divide by 80:

$$\text{Avge \# rolls } = \frac{14\times 1+11\times 21+9\times 3+...+1\times 28}{80} = \frac{482}{80} \approx 6.025 \,.$$

Note that the numerator really is the sum of 80 numbers, but I've grouped the 14 1's together, etc. I was impressed (and surprised) that the answer was so close to 6.

But is 6 really the answer? Time perhaps for some analysis. I ask for ideas, but the class is strangely quiet. They're not entirely sure what is wanted here. I let them talk about it a while in their groups, and I find that the most common activity is to take the above list of possible numbers of rolls, and try to assign a probability to each. This develops into an important approach.

One of the interesting things about this problem is the large number of different approaches or modes of reasoning that can be used—some complex, involving a good technical workout, and others unbelievably simple.

1. Calculating the probability of each outcome.
To ease their way into the probability idea, I rewrite the above empirical expression as:

$$\text{Average \# rolls } = \frac{14}{80}\times 1+\frac{11}{80}\times 2+\frac{9}{80}\times 3+...+\frac{1}{80}\times 28=\frac{482}{80}$$

In this form, we regard it as a *weighted sum*—each of the candidates, 1, 2, 3, etc. is weighted by the *proportion* of times it occurred, and then everything is added up.

What makes this an average is that the sum of the weights is 1:

$$\frac{14}{80}+\frac{11}{80}+\frac{9}{80}+...+\frac{1}{80}=1.$$

Now in the above expression, the weights used are the *proportions* that came from the 80 trials. The idea is that the theoretically exact average ought to have the same form—a weighted average of the possible outcomes—but with the experimental *proportions* replaced by the theoretical *probabilities* of the different outcomes. That is, the theoretical average should be obtained as

$$\text{theoretical average } = \quad p_1\times 1+p_2\times 2+p_3\times 3+...$$

where p_i is the probability of getting the first six on the *i*th roll.

The sum goes on forever. Although any really large number of rolls is very unlikely, it's certainly theoretically possible.

By this approach, in order to calculate the average, we have to first find a formula for the p_i and then perform the summation.

We find the p_i as follows. To get the first six on the *i*th roll we need a non-six on every one of the first *i*–1 rolls (probability 5/6 for each) and a six on the *i*th roll (probability 1/6). The probability of all these things happening is:

$$p_i=\left(\frac{5}{6}\right)^{i-1}\cdot\frac{1}{6}.$$

Calculating the p_i
For example, the probability p_1 of getting six on the very first roll is clearly 1/6:

$$p_1=\frac{1}{6}$$

What is the probability p_2 of getting the first six on the second roll? For this we need a non-six on the first roll and a six on the second. Now the probability of a non-six on the first roll is 5/6 (5 of the 6 possibilities), and the probability of a six on the second roll is 1/6, and the answer will be the product:

$$p_2=\frac{5}{6}\cdot\frac{1}{6}.$$

Armed with this the theoretical average number of rolls is:

$$\text{Theoretical Avge } = \left[p_1\times 1\right]+\left[p_2\times 2\right]+\left[p_3\times 3\right]+...$$

$$=\left[\frac{1}{6}\times 1\right]+\left[\frac{1}{6}\left(\frac{5}{6}\right)\times 2\right]+\left[\frac{1}{6}\left(\frac{5}{6}\right)^2\times 3\right]+\left[\frac{1}{6}\left(\frac{5}{6}\right)^3\times 4\right]+...$$

$$=\frac{1}{6}\left[1+2r+3r^2+4r^3+...\right]$$

Here I have reorganized the sum and replaced 5/6 by *r*, so we can more clearly see the form of the sum. This is not a geometric series, but there is an elementary argument (see problem 12) to show that the sum is:

$$\frac{1}{6}\left[\frac{1}{(1-r)^2}\right]=\frac{1}{6}\left[\frac{1}{(1/6)^2}\right]=6$$

and that's the answer we were expecting.

This "analytic" approach is definitely not what I wanted to begin with. Technically it is the most demanding and least intuitive of all. But perhaps the other approaches are more elusive, more inventive and harder to find. Certainly they are more fun. Read on!

2. *A naive argument.*
Now an interesting thing happens. A few students have been puzzling over a simple idea that they had at the beginning. Now having seen (endured?) the technical intricacies of approach #1, they naturally assume that their "two-line" approach is wrong, so they come and ask me to show them why.

Essentially, their argument goes like this: since on average 1/6 of all my rolls are **6**'s, there are on average six rolls for every **6** thrown. So on average I have to throw six times to get a **6**.

This argument is so simple it just has to be wrong.

But it isn't.

So where's the flaw?

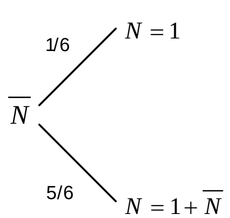
And my answer to them is that there's no flaw at all. The approach is perfectly sound and just needs a bit of shoring up. For example, it can be made a bit more rigorous as follows. Suppose I throw the die 6000 times. I'd expect to get 1000 **6**'s. Now imagine the 6000 outcomes all in a line. The set of outcomes from right after each **6** to the next **6** can be regarded as a typical sequence of throws to get a six. What's the average number of throws in such a sequence?—Well it's the total number of throws over the number of sequences, which is 6000/1000 which is 6.

This idea, of considering a large number of throws, can be made quite rigorous using standard probabilistic results. It's certainly wonderfully intuitive, and in that sense it's the most "convincing."

3. *The recursive approach*, *or "how does it feel?"*.
Let N denote the number of rolls to the first six. Then N will take different values on different occasions, and we are going to look for a recursive formula for its average which we'll denote by $\overline{N}$.
Think about the first roll. With probability 1/6 we'll get a six right away in which case N=1. But with probability 5/6 we won't and in this case our expectation for the number of rolls we will now require is $\overline{N}$ again (are not we right back where we started?) for a total of $1+\overline{N}$, that "1" being the one we've just had.

So there are two possibilities for the first roll, one gives us N=1 and the other gives us an average N of $1+\overline{N}$. We display this in a diagram:

$$\overline{N} = \frac{1}{6}\cdot 1 + \frac{5}{6}\cdot(1+\overline{N})$$



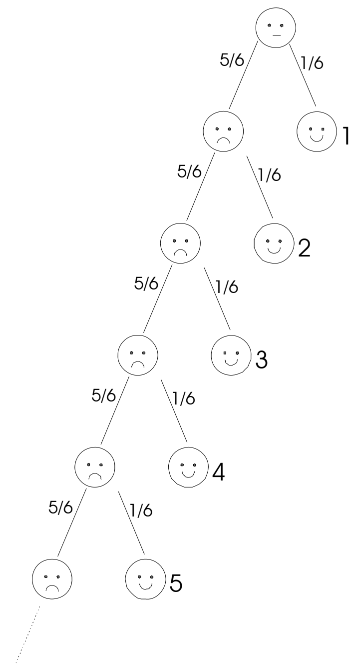
This gives us a "recursive" formula for $\overline{N}$, which can easily be solved to give $\overline{N}$ =6.

In many ways the recursive approach is the most sophisticated. It was not explicitly suggested by any student, but I found that it was already lurking in a number of minds. Recall your experiments, I said. You roll the die for the first time, and maybe you get a six right away, but maybe you don't. Maybe you get a **2**. In that case you pick up the die and get ready to roll again, *and hold it—hold it right there***.** Ask yourself:

How does it feel?
(Bob Dylan)

Well, it feels as if you're right back where you started, that the first roll was wasted, that whatever is the average number of rolls required to get the six, *that's the number you're looking at right now***.**

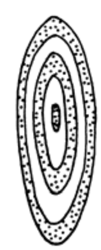
This is a very powerful style of mathematical argument and it's well worth playing with a bit. Some nice insights can be gained by studying its relationship to the first approach. Indeed, the first approach can be illustrated by drawing what might be called the full tree:



That first approach calculated the average number of rolls as an infinite sum, each term of which corresponds to a route from the top of the tree to one of the happy faces. The number of rolls at each happy face is the number of branches in the route (recorded at the right) and the probability of that path is the product of the probabilities along the branches. The first approach was really an accounting of all these routes.

Now the essential insight of this third approach is that there's a self-similarity in that tree structure!—that the stuff from the first sad face down is essentially a copy of the whole tree.

Let's return to the dart board and state our result. What we've done for 1/6 will work for any bulls-eye probability. For example, if at each throw you get a bulls-eye with probability 1/20, you'd expect to need 20 throws on average to get each one.



The dart-board theorem.
If on average a proportion p of all my throws are bulls-eyes then on average the number of throws required to get a bulls-eye is 1/p.

Problems

1. With a standard die, what is the average number of rolls required to get an even number (any one will do).

2. A dart board has a red bulls-eye at the centre of a blue circle. Suppose 80% of all my throws actually hit the board, and *of those*, 50% get inside the blue circle and *of those*, 10% hit the bulls-eye. How many throws are required, on average, to hit the bulls-eye?

3. On average, how many tosses of a fair coin are needed to get the first repetition, that is, two heads in a row or two tails in a row? Use the first approach—that is, catalogue all the different possible routes and assign a probability and a number of rolls to each.

4. On average, how many tosses of a fair coin are needed to get two heads in a row? Use the recursive approach—that is, find a recursive formula for the average, and then solve it. Illustrate your formula with a tree diagram.

5. Problem 3 can also be done using the recursive approach, but the argument is just a bit trickier than that for problem 4. Give it a shot.

6. What is the average number of times I have to toss a fair coin to get heads and tails each at least twice?

7. What is the average number of times I have to roll the die to get every even outcome at least once?

8. With a standard die, what is the average number of rolls required to get:

(a) two **6**'s

(b) a **6** and a **5**

(c) two **6**'s and a **5**

(d) two **6**'s and two **5**'s.

The recursive approach works very nicely for these problems if you do them in the order given.

9. (a) Suppose I have a biased coin which comes up Heads with probability 2/3 and Tails with probability 1/3. How many flips does it take, on average, to get both outcomes at least once?

10. I deal cards one at a time face up from a standard shuffled deck, and stop when I get the ace of spades. On average, how many cards do I have to deal?

11. In the first rather technical approach, $p_i = \left(\frac{5}{6}\right)^{i-1} \frac{1}{6}$ was the probability of getting the first six on the ith roll. Verify that the p_i have sum 1 (as expected). For this you will need the formula for the sum of an infinite geometric series:

$$a + ar + ar^2 + ar^3 + \ldots \quad = \quad \frac{a}{1-r}.$$

12. The infinite sum $1 + 2r + 3r^2 + 4r^3 + \ldots$ can be evaluated (for $-1 < r < 1$) using the same trick we used to find the sum of the infinite geometric series. Let the sum be S. Calculate rS and then look at $S - rS$.