Puzzling Problems in Probability

April 16-17, 2011

The most important questions of life are indeed, for the most part, really only problems of probability – Pierre Simon Laplace

1 Introduction

There are many ways to approach probability for the first time, as the subject is relevant to a variety of scientific disciplines and has a rich theory in and of itself. One could teach a probability class with either an emphasis on mathematical rigor, engineering applications, biological problems, or operations analysis. Indeed, five departments at Stanford offer an introductory probability class.

We will approach probability from a puzzle solving perspective. Our goal is to motivate the fundamental ideas of probability using problems. Thus, rather than subjecting the reader to a list of definitions and theorems and only subsequently providing examples, we will inspect a series of interesting questions and use these questions to develop some basic probabilistic concepts.

2 The Birthday Paradox

Problem. What is the chance that anyone in this room (of n people) has a birthday matching any one else's?

As is the true in many probability problems, it is easier to investigate the opposite question: What is the chance that no one in this room has a birthday probability any one else's?

Why is this the logical thing to do? Here are two intuitive ideas that have been formalized by probability theory:

- The sum of probabilities taken over all possible outcomes is 1. For example, if a coin is tossed, a head turns up with probability 1/2 and a tails turns up with probability 1/2, so the probability that either event occurs is 1. In our problem, the sum of all possible outcomes (some two people share a birthday vs. no two people share a birthday) is also set equal to 1.
- If two events A and B are mutually exclusive but together include all possible events, then the probability that A occurs is 1 minus the probability that B occurs. More formally, $\mathbb{P}(A) = 1 - \mathbb{P}(B)$. In our case, this is simply $\mathbb{P}(\text{Some two people share a birthday}) = 1 - \mathbb{P}(\text{No two people share a birthday})$

We can now proceed to calculate the probability that no two people in this room share the same birthday. First of all, notice that if there were 366 people in this room, then there necessarily have to be two people in this room who share the same birthday (it's impossible to have 366 distinct birthdays).

How would we actually go about solving this problem? We would probably write a list with everyone's name and birthday on it. We could go down the list and check whether the birthday we are currently analyzing has appeared previously on the list.

Let R_j =the event that we have to stop on the j^{th} item on the list, because we have arrived at a duplicate birthday.

Let T_j =the event that the once we get to the j^{th} birthday on the list, we are able to continue reading down the list because we have still not arrived at a duplicate.

By the bullet points above, $\mathbb{P}(R_j) = 1 - \mathbb{P}(T_j)$.

Notice that, if we can go through the entire list without stopping, then we know that each of the birthdays is unique. That is, D_n is true if and only if each of the $R_1, \ldots, R_n$ does not occur. Therefore, the probabilities of these events are related by the following equation:

$$\mathbb{P}(D_n) = 1 - (\mathbb{P}(R_1) + \mathbb{P}(R_2) + \cdots + \mathbb{P}(R_n))$$

Notice that $\mathbb{P}(R_1) = 0$ because we could never stop on the first entry of the list. To calculate $\mathbb{P}(R_2)$, notice that we will only stop if the day in position 2 is the same as that in position 1. If we suppose that all birthdays are distributed uniformly throughout the year, this event occurs with a probability $\frac{1}{365}$. Therefore, $\mathbb{P}(D_2) = \frac{364}{365} = \left(1 - \frac{1}{365}\right)$. Let's continue. What is the probability $\mathbb{P}(R_3)$ that we stop once we reach the third name? For this event to occur, we need the third birthday to be the same as either the first or second birthday; this happens with probability $\frac{2}{365}$. By the formula above, $\mathbb{P}(D_3) = 1 - (\mathbb{P}(R_1) + \mathbb{P}(R_2) + \mathbb{P}(R_3)) = 1 - \left(\frac{1}{365} + \frac{2}{365}\right) =$

Notice that we will never stop on the first birthday on the list, because it is always going to be unique. Therefore, $\mathbb{P}(R_1) = 0$ and $\mathbb{P}(T_1) = 1$. The second birthday on our list is going to be different from the previous birthday with a probability $\frac{364}{365}$, and it is going to be the same with probability $\frac{1}{365}$. Therefore, $\mathbb{P}(R_2) = \frac{1}{365}$ and $\mathbb{P}(T_2) = 1 - \frac{364}{365}$. Let's suppose that we were able to get to the third name in the list and want to find the probability that we continue past the third point. In this case, $\mathbb{P}(R_3) = \frac{2}{365}$ and $\mathbb{P}(T_3) = 1 - \frac{2}{365}$ because we will only stop if the third birthday is the same as one of the previous two. The birthdays are only unique if we can read through the entire list without stopping. This is simply $\mathbb{P}(T_1 \text{ and } T_2\text{and }, \ldots, \text{ and } T_n)$ which is typically notated $\mathbb{P}(\cap_{i=1}^n T_i)$. For the same reason that the probabilities of two coin flips turning up HH is the (probability that a single coin flip turns up H)^2, we have

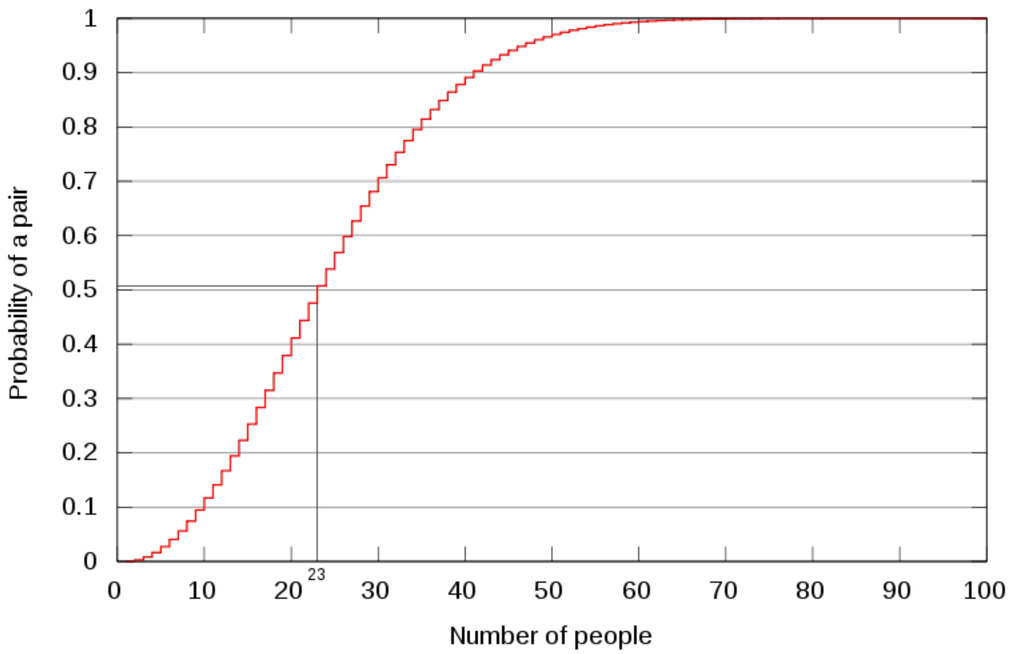
$$\mathbb{P}(\text{no two people have the same birthday}) = \mathbb{P}(\cap_{i=1}^n T_i) = \mathbb{P}(T_1)\,\mathbb{P}(T_2)\ldots\mathbb{P}(T_n)$$

$$= (1)\left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)\cdots\left(1 - \frac{n-1}{365}\right)$$

Therefore, the probability that some two people do have the same birthday is

$$= 1 - \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)\cdots\left(1 - \frac{n-1}{365}\right)$$

The surprising thing about this result (and the reason it is called a paradox) is that this probability increases very rapidly. In fact, in a group of 23 people, the probability that at least two of them share a birthday is greater than $\frac{1}{2}$! Below is a plot of how the probability that some two people share the same birthday grows as a function of the number of people in the group.



3 Catching a Cautious Counterfeiter

Problem. The king's minter boxes his coins 100 to a box. In each box, he puts 1 false coin. The king suspects the minter and from each of 100 boxes draws a random coin and has it tested. What is the chance the minter's corruption goes undetected? What happens if we replace both 100's by n's?

In the first case, the probability that not a single false coin is selected is:

$$\mathbb{P}\left(0 \text{ false coins chosen in any of the 100 boxes}\right) = \mathbb{P}\left(0 \text{ false coins in a single box }\right)^{100}$$

$$= \left(1 - \frac{1}{100}\right)^{100}$$

For general n, this becomes $\left(1 - \frac{1}{n}\right)^n$. Notice that as $n \to \infty$, this tends to $\frac{1}{e} \approx 0.367879$. A rigorous derivation of this fact requires calculus, but we can see this limiting behavior by simply writing out some terms.

n	$\mathbb{P}\left(0 \text{ false coins}\right)$
1	0
2	0.250
3	0.296
4	0.316
10	0.349
100	0.366
1000	0.3677

4 Catching A Greedy Counterfeiter

Problem. The king's minter boxes his coins n to a box. Each box contains m false coins. The king suspects the minter and randomly draws 1 coin from each of n boxes and has these tested. What is the chance that the sample of n coins contains exactly r false coins?

The probability that any single coin that the king draws is a false coin is simply the fraction of false coins in any box: $\frac{m}{n}$. The probability that exactly r of these false coins are selected is a binomial distribution: $\mathbb{P}(r \text{ false coins}) = \binom{n}{r}\left(\frac{m}{n}\right)^r\left(1-\frac{m}{n}\right)^{n-r}$. This distribution arises because we want r false coins drawn, which each occur with probability $\frac{m}{n}$, and $n-r$ true coins drawn, which each occur with probability $1-\frac{m}{n}$. Further, there are $\binom{n}{r}$ ways to order these true and false coins.

5 Playing Matchmaker

Problem. Eight eligible bachelors and seven beautiful models happen randomly to have purchased single seats in the same 15-seat row of a theater. On the average, how many pairs of adjacent seats are ticketed for marriageable couples?

Consider an example outcome: $BBMMBBMBMBMMBBMM$

In this case, there are 9 BM and MB pairs. We want the average number of adjacent pairs that are either BM or MB. If a pair is BB or MM, then we score that pair a 0, if they are either MB or BM, they we score that pair a 1. The chance of a marriageable couple in the first two seats is $\frac{8}{15}\times\frac{7}{14}+\frac{7}{14}\times\frac{8}{14}=\frac{8}{15}$. This is in fact the expected number of marriageable couples in the first two seats–we have a $\frac{8}{15}$ probability of a 1 and a $\frac{7}{15}$ probability of a 0, so the expected number of couples in the first two seats is $1\times\frac{8}{15}+0\times\frac{7}{15}=\frac{8}{15}$. This calculation applies over any adjacent pair, and since there are 14 adjacent pairs, the total expected number of marriageable pairs is $7\frac{7}{15}$.

This problem helps motivate the idea of an expected value. The expected value of a random variable is the weighted average of possible outcomes. More precisely, the expected value of a random variable X that takes on values x_i is defined as $\mathbb{E}(X)=\sum_{i=1}^{\infty}x_i\mathbb{P}(X=x_i)$. To see why this is a reasonable definition, notice that for any set of numbers, the average can be written in a form similar to the expected value. For example, $(1+0+8+6+6+1+6)/7 = 0\times\frac{1}{7}+1\times\frac{2}{7}+6\times\frac{3}{7}+8\times\frac{1}{7}=4$.

The reason we were able to equate an expected value with a probability is due to the fact that the expected value of an indicator variable on an event (a variable that takes a value 0 when the event does not occur and a1 when the event occurs) is equal to the probability of that event happening. Let $\mathbb{I}(X)$ be the indicator of the event X. Then, $\mathbb{E}(\mathbb{I}(X)) = 0\times\mathbb{P}(\mathbb{I}(X)=0)+1\times\mathbb{P}(\mathbb{I}(X)=1) = 1\times\mathbb{P}(X \text{ occurs}) = \mathbb{P}(X \text{ occurs})$.

The other main idea we used was that the expected value of a sum of events (the expected number of couples over all possible adjacent seats) was equal to the sum of the expected values of each event (the expected number of couples for a single pair of adjacent seats). In general, if X and Y are random variables, $\mathbb{E}(aX+bY)=a\mathbb{E}(X)+b\mathbb{E}(Y)$. This property is known as the linearity of expectation, and it comes up in a variety of applications. However, don't let the notation get in the way of understanding–the reasoning in the first part of this problem solution provides the intuition necessary to solve probability questions.

6 Collecting Coupons

Problem. Coupons in cereal boxes are numbered 1 to 5, and a set of one of each is required for a prize. With one coupon per box, how many boxes on the average are required to make a complete set?

In approaching this problem, it is easier to think of the number of boxes on average required to get the first coupon, the number required to get the second, etc.

Let's step back now, and consider a more general setting. Suppose you have a coin where you get successes with probability p and failures with probability $1-p$. What is the expected number of times you have to flip this coin before you get a heads?

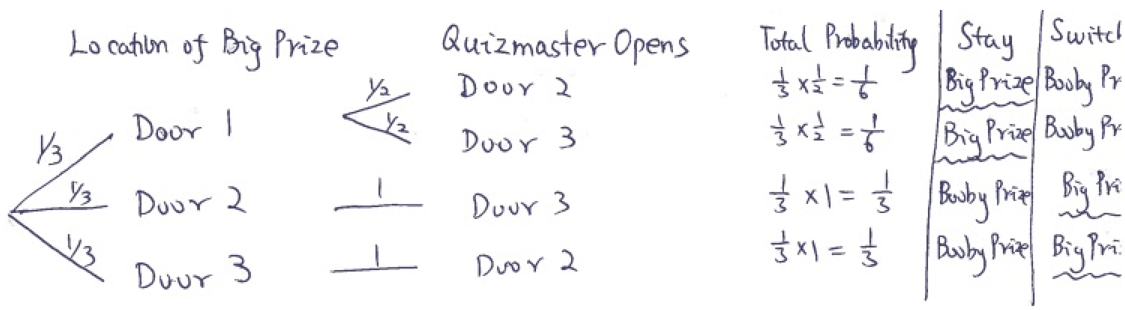
Intuitively, this number should be inversely proportional to the probability of the event–the larger the probability of a success, the less time we would expect it to take before we achieved one. It turns out that we can use the definition of expected value to find that, if X is the number of flips required before a success is achieved, $\mathbb{E}(X)=\frac{1}{p}$.

We can apply this result to our current problem. Clearly, our first coupon comes from our first box. After that, we can treat the number of boxes it takes before we get a second coupon like we treat the number of coins we have to flip before we get a success. For each box, we have a probability of $\frac{4}{5}$ of getting a new coupon, so the expected number of boxes required to get two coupons is $1 + \frac{5}{4} = \frac{9}{4}$. Likewise, after getting the first two coupons, we have a probability $\frac{3}{5}$ of getting a third coupon, so the expected time to go from having two to three coupons is $\frac{5}{3}$. In the end, we have that the expected number of boxes we have to go through before we get all of the coupons is $5\left(\frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{2} + 1\right) \approx 11.42$.

7 Monty Hall Problem

Problem. Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No.1, and the host who knows what's behind the doors, opens another door, say No.3, which has a goat. He then says to you, "Do you want to pick door No.2?" Is it to your advantage to switch your choice?

The following table shows the probability of every possible outcome if the player initially picks door No.1.



Therefore, the player should switch–doing so doubles the probability of winning from $\frac{1}{3}$ to $\frac{2}{3}$.

8 Conclusion

I hope these notes gave you a fun introduction to the richness and surprises that characterize probability theory. I know that I glossed over most of the formal topics traditionally covered in probability classes, but I wanted to emphasize the intuition that lies behind approaching questions about chance. Good luck with your studies, and I hope you choose to learn more probability in the future!

References

[1] Many of the problems discussed here were taken from Frederick Mosteller's "Fifty Challenging Problems in Probability"

[2] For a comprehensive introduction to probability theory, Sheldon Ross' "A First Course in Probability" and William Feller's "An Introduction to Probability Theory and its Applications" are considered classics.