Probability

80% of the time it works every time.

Math Circle

February 4th, 2018

Today we are going to do two things. The first, is that we are going to review (or learn for the first time) some of the basic rules of probability. Second, we are going to learn about Bayes rule which, I'm not kidding, might just change your life. Or at least change the way that you think about life.

1. First let's warm up with some soft balls. Compute the probability of each event, assuming that the coins, die, etc. are all totally fair random.

   (a) If I roll a die, what are the chances of rolling a 6?

   (b) What are the chances of rolling two of the same numbers in the row? Three times in a row?

(c) If I roll two six sided dice, what are the chances that the combined roll will be a 7?

(d) Prove that if I have two dice with n sides. The sides on each dice are labeled $1, 2, \ldots, n-1, n$. Prove that the chances of rolling an $n+1$ is always exactly $\frac{1}{n}$. *Hint, there is a harder, and an easy way to prove this. If you are trying to take a very big sum, you are probably barking up the wrong tree.

(e) If I flip a coin and roll a die, what are the chances that I get a heads, and the number on the die is even? What is the probability that neither happens?

(f) **People often talk about something having a 'good chance' of something happening. Like, "there is a good chance that it will rain tomorrow," or "there is a good chance that eating this entire pie by myself is a mistake." What probability does this translate to? Well, if you roll six 6 sided dice, there is a 'good change' that you'll roll a 1 at least once. And if you roll 20 20 sided dice, then there is a 'good chance' that you will roll a 1 at least once. If you consider the problem of rolling n n sided dice, then 'good chance' converges to a fixed number as you increase n. What is that number?**

2. In order to talk about probability in a more abstract way, we need to make two things more precise. A set of events $\mathscr{A} = \{a_1, a_2, \dots\}$, and a probability function P. P is a sort of unusual function, because it doesn't take a number as input, but rather it takes as input a subset of $\mathscr{A}$. As an example, if $\mathscr{A} = \{$we will finish this entire handout, we won't finish this entire handout$\}$, then I could say that $P(\{$we will finish this entire handout$\}) = 0$. You can pick just about anything for $\mathscr{A}$, but P has to obey a few rules

   - If x is a set of events (could be just one event, could be more) then $P(x) \geq 0$.
   - If a_1, $a_2 \dots a_n$ are events, then $P(\{a_1, a_2, \dots, a_n\}) = P(\{a_1\}) + P(\{a_2\}) + \cdots + P(\{a_n\})$.
   - $P(\mathscr{A}) = 1$

   (a) If I'm talking about a problem where you flip a fair coin, what is $\mathscr{A}$? What is P for every possible subset of $\mathscr{A}$?

(b) If I'm talking about the problem of rolling a fair dice, what is $\mathscr{A}$? How can I express the English statement "an even number was rolled" using the above notation? What about "A prime number was rolled?"

(c) If you roll a fair die what are the chances that you get either an even number, or a prime? What about a prime and an even number?

(d) Let's say that $x, y \subset \mathscr{A}$. The set $x \cup y$ is called the union of x and y. It's defined as the set of things that are in either x or y. Similarly we call $x \cap y$ as the intersection of x and y, and it's just the set of things that are in both x and y. Using a convincing English argument, justify the identity $P(x \cup y) = P(x) + P(y) - P(x \cap y)$.

(e) Now try to prove that $P(x \cup y) = P(x) + P(y) - P(x \cap y)$ using only the three properties above.

(f) Use this rule to redo problem 2.c

3. Before we talk about Bayes theorem, there is one more thing that we have to touch on, conditional probability. When I write that $P(x|y)$ you should read that as "The probability that x happens, if y happened." In other words, on the condition that y happens, what is the condition that x also happens?

   (a) What is the $P(x|y)$ if x is rolling an even number on a fair die, and y is rolling a prime?

(b) We define $P(x|y) = \frac{P(x \cap y)}{P(y)}$ mathematically. Compare your answers from 2.c and 3.a. Do they agree with the definition?

(c) Finally we can talk about Bayes theorem. Bayes theorem says that for any x, y then

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}. \tag{1}$$

Can you prove Bayes theorem?

4. Now that we finally have Bayes theorem, we can use it to resolve some interesting paradoxes. problems b - d are taken from the website http://www.daviddeley.com and problem e came from real life data.

(a) You are at a fancy dinner. At the dinner, you overhear two guests talking about their families. They both say that they have two children. One of them says that their eldest is a girl, and the other says that they have at least one girl. Before you could hear any more of their conversation, the both start choking on their Hors d'oeuvre, and fall over dead. You can't help but wonder "what is the probability that each of the guests had two girls?"

(b) This is a very famous problem called the Monty hall problem. It is as follows. Monty Hall hosted a game show called Let's Make a Deal. He presented his game show contestant with three doors numbered 1, 2, and 3. Behind one of the doors is the grand prize. The contestant chooses a door. Then Monty, who knows what's behind each door, opens up one of the two remaining doors which doesn't have the grand prize behind it. Monty then asks the contestant, "Do you want to stay with your original choice, or would you like to switch to the other remaining door?"

Should the contestant stay with her original choice, should she change to the other door, or does it not make any difference?

(c) Three cards are in a hat. One card is white on both sides; the second is white on one side and red on the other; the third is red on both sides. The dealer shuffles the cards, takes one out and places it flat on the table. The side showing is red. The dealer now says, "Obviously this is not the white-white card. It must be either the red-white card or the red-red card. I will bet even money that the other side is red." Is this a fair bet?

(d) A doctor has two drugs, A and B, which he can prescribe to patients with a certain illness. The drugs have been rated in terms of their effectiveness on a scale of 1 to 6, with 1 being the least effective and 6 being the most effective. Studies show that drug A is uniformly effective at a value of 3. Drug B varies in its effectiveness. 54% of the time it scores a value of 1, and 46% of the time it scores a value of 5.

The doctor, wishing to provide his patients with the best possible care, asks his statistician friend which drug has the highest probability of being the most effective. The statistician says, "It is clear that drug A is the most effective drug 54% of the time. Thus drug A is your best bet."

Later a new drug C becomes available. Studies show that on the scale of 1 to 6, 22% of the time this drug scores a 6, 22% of the time it scores a 4, and 56% of the time it scores a 2.

The doctor, again wishing to provide his patients with the best possible care, goes back to his statistician friend and asks him which drug has the highest probability of being the most effective. The statistician says, "Well, seeing as there's this new drug C on the market, your best bet is now drug B, and drug A is your worst bet."

Show that the statistician is right.

(e) The table below shows the success rates and numbers of treatments for treatments involving both small and large kidney stones, where Treatment A includes all open surgical procedures and Treatment B is percutaneous nephrolithotomy (which involves only a small puncture). The numbers in parentheses indicate the number of success cases over the total size of the group.

Small Stones	93% (81/87)	87% (234/270)
Large Stones	73% (192/263)	69% (55/80)
Both	78% (273/350)	83% (289/350)

It would appear that treatment A is more effective in either case, but is less effective if you don't know what case you are in. How is this possible?

(f) There is a **TERRIBLE** disease out there called maidupitus which afflicts about 500,000 out of the around 300 million Americans. You test positive for maidupitus. The test that you took isn't 100% accurate, 1% of the time it will say that you have the disease even if you don't and 1% of the time it will say that you don't have the disease, even if you do. What are you chances of having maidupitus?

Finally, I wanted to show you a YouTube video that explains Bayes rule in a visual way, and is full of real life examples, which I like a lot. It's called "A visual guide to Bayesian thinking" by the channel "Julia Galef". The URL is:

`https://www.youtube.com/watch?v=BrK7X_XlGB8`