CS 2800: Discrete Math	Oct. 17, 2011

Lecture

Lecturer: John Hopcroft | *Scribe: Dan & June*

Review

1 Monte Hall Problem

The Monte Hall Problem is an interesting excercise in conditional probability. In the scenario, you are on a game show competing for a grand prize. The game consists of you choosing among three doors and you are awarded whatever prize hides behind the door you choose. You know that one of the doors is hiding a sports car and the other two are hiding goats. Obviously, it is your objective to choose the door hiding the car. However, you do not know which door is hiding what prize.

When the game begins you select a door (let's say door A), but it is not opened yet. Because there are two goats, at least one of the two doors you did not choose must be hiding a goat. So, as an added twist, the host of this show - Monte Hall - will open this door (let's say door B) and reveal a goat. He then poses this question: " Do you wish to change your selection to the other remaining door (door C)? "

What do you do? Many people will first assume that switching doors will neither help or hurt your chances. Their reasoning is that when you initially choose, each door has a 1/3 chance of hiding a car and Monte revealing one door doesn't change this uniform probability. This turns out to be *false*. In reality, you double your chances of winning the car if you choose to switch. For a mathematical explanation, you'll need to turn to conditional probability and Bayes Theorem (covered in the next lecture). However, we can develop some intuition to convince you of this startling effect.

Let's consider the only two cases. Case 1 is when you initially choose a door with a goat behind it. Then, Monte will reveal the only remaining goat and, clearly, you should switch to the final door (which must be hiding the car). In Case 2, you initially choose the door with the car behind it. Then, Monte will reveal either of two remaining goats and ask you if you wish to switch. Obviously, you will lose the car if you switch and it is better to keep your first choice.

Hopefully you now see why switching is twice likely to get you the car: it is the better choice in Case 1 and Case 1 is *twice as likely to occur*! If you repeated the game many times, switching your choice will give you a car twice as often as it will give you a goat. So, if anyone ever presents you with an option like this, you should *always* switch!

2 Probability Intro

Probablility is formally covered in the next lecture. This quick introduction will serve to first build intuition on the subject.

A common way to define probability is with counting. For example, we can continuously flip a fair coin and count how many times heads comes up. Over time, we'll surely notice that about half of the outcomes are heads. So, we call the probablility of a single flip bing heads 1/2. A direct result of this definition is that probabilities must always be between 0 and 1 inclusive. Yes, somethings can have probability 0: the probability that weather in Ithaca will be exactly 120 degrees tomorrow is virtually 0.

Another product of this definition is that the sum the probabilities of all possible events must be 1. In the example above, this is simply 1/2 for heads + 1/2 for tails = 1. Again, a formal treatment of this property is covered in the next lecture. A good way of remembering this fact is by understanding that *something must happen*. When you toss a perfect fair coin, either it lands on heads or tails. There are no other options. Since a probability of 1 leaves no room for other events to occur, it makes sense that the sum of the probabilities of the two outcomes equals 1.

Let's say somebody now told you that they had devised a system of choosing an integer greater than 1 such that the probability of a number being picked is equal to its reciprocal. For example, the probability that you choose $5 = 1/5$ and 100 has a 1/100 chance of being picked. Is this possible? The answer is no. To check schemes like these, the only thing you have to ensure is that the sum of all probabilities equals 1. If you sum the infinite series $1/2 + 1/3 + 1/4$ ... (a.k.a the harmonic series) you will notice that it diverges, $\sum_{i>1} \frac{1}{i} = \infty$. Notice, if we changed the scheme so that the probability is proportional to the *square* of the reciprocal, then something different occurs. The infinite sum $1/4 + 1/9 + 1/16$ ... actually converges to $\pi^2/6$. While this also doesn't sum to 1, it is a more manageable situation. We can now simply normalize the probability of each event so that the sum does equal one, but their relative probabilities do not change. For the reciprocal squared example, we would multiply each probability by $6/\pi^2$.

3 Poker

Poker involves being dealt a 5 card hand from a 52 card deck. The 52 cards are organized by suit and rank. There are 4 suits, and 13 ranks. The suits are spade, diamond, heart, and club. The ranks are $2, 3, \ldots, 10, J, Q, K, A$. Certain combinations of the 5 cards are called, flushes, royal flushes, full house, etc. The rarer the named combination your hand is the more likely you are to win. In Poker the order of your cards does not matter. (If you are unfamiliar with the game, check it out on wikipedia - gambling with money is unadvised.) The number of possible 5 card hands is $\binom{52}{5}$

Example A *flush* is a 5 card hand, where all the cards are of the same suit, eg all cards are diamonds, or all cards are clubs, etc. Here we calculate the probability of being dealt a flush.

First we find the number of hands that are a flush. There are 4 possible suits, spades, diamonds, hearts, and clubs. For each suit there are 13 cards of that suit. Now, there are $\binom{13}{5}$ ways to get a flush in a given suit. So, there are $4\binom{13}{5}$ hands that are flushes. The probability of getting a flush is then

$$4\frac{\binom{13}{5}}{\binom{52}{5}} \approx 0.00198$$

Example A hand with a pair is a 5 card hand where exactly two cards have the same rank. Here we calculate the probability of being dealt a pair.

First, for a given rank there are 4 cards so there are $\binom{4}{2}$ ways to select a pair. There are 13 ranks, so there are $13 * \binom{4}{2}$ ways to just select a pair. Now we have to combine this with the number of ways the remaining 3 cards can be picked. Well, for the hand to be a pair, none of the remaining three can be of the same rank as the pair (otherwise we'll have three-of-a-kind or four-of-a-kind). So, there are 12 remaining ranks and 3 cards to choose. We can't forget, though, that each of the 12 remaining ranks has 4 cards per suit. Thus, the number of ways to choose the remaining 3 cards is $\binom{3}{12} * 4^3$. Multiplying our two counts together and dividing by the total number of possible hands, we find: $13 * \binom{4}{2} * \binom{12}{3} * 4^3 / \binom{52}{5}$ which is approximately .4225.

4 Birthday Paradox

How many people must be in a room, such that with probability at least a half, two people have the same birthday?

We assume no twins, that everyone's birthday is independent of everyone else's.

Let's look at this with the probability that no one has the same birthday. The first person has probability 1. The second person has probability $\frac{364}{365}$ of having a different birthday. So given that no one has had the same birthday before, the probability the k^{th} person has a different birthday is $\frac{365-k}{365}$. So the accumulated probability of making it to person k without a birthday is $\frac{365}{365}\frac{364}{365} \cdots \frac{365-k}{365}$.

Now we can solve $\frac{365}{365}\frac{364}{365} \cdots \frac{365-k}{365} \geq \frac{1}{2}$. It works out to be roughly $\sqrt{365} \approx 22$

Now, if you look at the probabilty mapped as, P x number of people, you'll notice a sharp threshold. This is known as a phase transition and plays many important parts in research.

5 Hashing

We dealt with hashing before, and had to worry about collisions. How many elements do I have to hash, before the probability of a there is a collison.

If you use the sqrt argument, even if your array was 10^6, then you could only hash 1000 elements before a collision occured.

6 Infinite Sample Spaces

Example Suppose we are flipping a coin. What is the probability, that first heads occurs at the i^{th} flip? This means for all flips prior to i, the flip yielded a tails. We get the series $\frac{1}{2}, \ldots, \left(\frac{1}{2}\right)^{i-1} \frac{1}{2}$. Note, if we add up the sequence, we'll get 1.

It is a geometric series. ... insert

$$
\begin{aligned}
S &= 1 + a^1 + a^2 + \ldots \\
aS &= a + a^2 + a^3 + \ldots \\
S - aS &= a \\
S &= \frac{a}{1-a}
\end{aligned}
$$