EE 178/278A Probabilistic Systems Analysis
Spring 2014 Tse/Hussami Lecture 2

Introduction to Probability (cont.)[1]

Reading: Chapter 1 from Bertsekas and Tsitsiklis.

We first review the example of rolling two fair dice.

Roll two fair dice. Then the same space $\Omega = \{(i,j) : 1 \le i,j \le 6\}$. Each of the 36 outcomes has equal probability, $\frac{1}{36}$. Let's look at the event A that the sum of the dice is at least 10 (which holds for the pairs in the set $\{(5,5),(6,6),(6,4),(4,6),(5,6),(6,5)\}$), and the event B that there is at least one 6. In this example, our probability assignment is **uniform**, i.e., all the outcomes have the *same* probability (which must be $\frac{1}{|\Omega|}$, where $|\Omega|$ denotes the size of Ω). In such circumstances, the probability of any event A is clearly just

$$\mathbf{P}(A) = \frac{\text{\# of outcomes in } A}{\text{\# of outcomes in } \Omega} = \frac{|A|}{|\Omega|}.$$

So for uniform assignments, computing probabilities reduces to *counting* outcomes! Using this observation, it is now easy to compute the probabilities of the two events A and B above: $\mathbf{P}(A) = \frac{6}{36} = \frac{1}{6}$, and $\mathbf{P}(B) = \frac{11}{36}$.

We now introduce some set-theoretic notation:

1. Union of sets: The union of sets A and B, denoted by $A \cup B$, is the set of elements that belong to either A or B or both:
$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

2. Intersection of sets: The intersection of A and B, denoted by $A \cap B$, is the set of elements that belong to both A and B:
$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

3. Complement: The complement of A, denoted by A^c, is the set of all elements that are not in A:
$$A^c = \{x : x \notin A\}.$$

4. Subset: Set A is a subset of a set B, denoted $A \subset B$, if $x \in A \Rightarrow x \in B$.

Two events are said to be disjoint when they cannot occur simultaneously, meaning they do not have any common outcomes.

We now state some useful properties:

[1] Part of this note is adapted from the notes of EECS 70 at Berkeley.

1. The probability of the union of two sets satisfies the following relation:

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B).$$

   Therefore,

$$\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$$

   holds in general. When events A and B are disjoint, $\mathbf{P}(A \cap B) = 0$. Therefore, in this case

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B).$$

2. The probability of a set and its complement are related by the following equation: $\mathbf{P}(A) = 1 - \mathbf{P}(A^c)$.

3. $A \subset B \Rightarrow \mathbf{P}(A) \leq \mathbf{P}(B)$.

In the dice example, the set B can be written as the union of two sets. Let $C = \{(i,j) : i = 6\}$ and $D = \{(i,j) : j = 6\}$. Then we have $B = C \cup D$. We can compute the probability of the set B as $\mathbf{P}(B) = \mathbf{P}(C) + \mathbf{P}(D) - \mathbf{P}(C \cap D) = 1/6 + 1/6 - 1/36$, where the intersection of the sets C and D only contains the element $(6,6)$.

Birthday Paradox

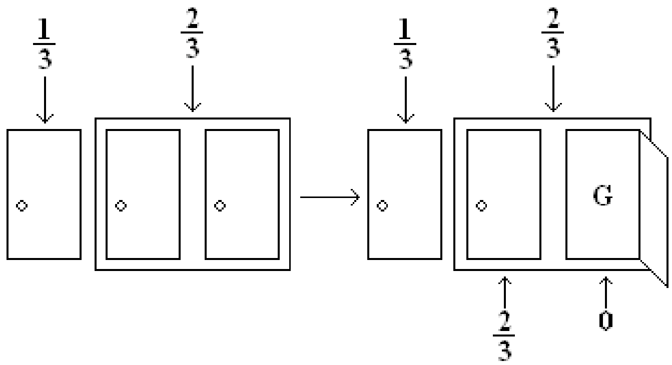
The "birthday paradox" is a remarkable phenomenon that examines the chances that two people in a group have the same birthday. It is a "paradox" not because of a logical contradiction, but because it goes against intuition. For ease of calculation, we take the number of days in a year to be 365. If we consider the case where there are n people in a room, then $|\Omega| = 365^n$. Let A = "At least two people have the same birthday," and let A^c = "No two people have the same birthday." It is clear that $\mathbf{P}(A) = 1 - \mathbf{P}(A^c)$. We will calculate $\mathbf{P}(A^c)$, since it is easier, and then find out $\mathbf{P}(A)$. How many ways are there for no two people to have the same birthday? Well, there are 365 choices for the first person, 364 for the second, $\ldots$, $365 - n + 1$ choices for the n^{th} person, for a total of $365 \times 364 \times \cdots \times (365 - n + 1)$. (Note that this is just sampling without replacement with 365 bins and n balls; as we saw in the previous Note, the number of outcomes is $\frac{365!}{(365-n)!}$, which is what we just got.) Thus we have $\mathbf{P}(A^c) = \frac{|A^c|}{|\Omega|} = \frac{365 \times 364 \times \cdots \times (365-n+1)}{365^n}$. Then $\mathbf{P}(A) = 1 - \frac{365 \times 364 \times \cdots \times (365-n+1)}{365^n}$. This allows us to compute $\mathbf{P}(A)$ as a function of the number of people, n. Of course, as n increases $\mathbf{P}(A)$ increases. In fact, with $n = 23$ people you should be willing to bet that at least two people do have the same birthday, since then $\mathbf{P}(A)$ is larger than 50%! For $n = 60$ people, $\mathbf{P}(A)$ is over 99%.

The Monty Hall Problem

In an (in)famous 1970s game show hosted by one Monty Hall, a contestant was shown three doors; behind one of the doors was a prize, and behind the other two were goats. The contestant picks a door (but doesn't open it). Then Hall's assistant (Carol), opens one of the other two doors, revealing a goat (since Carol knows where the prize is, she can always do this). The contestant is then given the option of sticking with his current door, or switching to the other unopened one. He wins the prize if and only if his chosen door is the correct one. The question, of course, is: Does the contestant have a better chance of winning if he switches doors?

Intuitively, it seems obvious that since there are only two remaining doors after the host opens one, they must have equal probability. So you may be tempted to jump to the conclusion that it should not matter whether or not the contestant stays or switches. We will see that actually, the contestant has a better chance of picking the car if he or she uses the switching strategy. We will first give an intuitive pictorial argument, and then take a more rigorous probability approach to the problem.

To see why it is in the contestant's best interests to switch, consider the following. Initially when the contestant chooses the door, he or she has a $\frac{1}{3}$ chance of picking the car. This must mean that the other doors combined have a $\frac{2}{3}$ chance of winning. But after Carol opens a door with a goat behind it, how do the probabilities change? Well, the door the contestant originally chose still has a $\frac{1}{3}$ chance of winning, and the door that Carol opened has no chance of winning. What about the last door? It must have a $\frac{2}{3}$ chance of containing the car, and so the contestant has a higher chance of winning if he or she switches doors. This argument can be summed up nicely in the following picture:



What is the sample space here? Up to the point where the contestant makes his final decision, there are three random choices made: the game show host's choice of where to put the car, the contestant's initial choice of door, and Carol's choice of which door to open. We can therefore describe the outcome of the game using a triple of the form (i, j, k), where $i, j, k \in \{1, 2, 3\}$. The values i, j respectively specify the location of the prize and the initial door chosen by the contestant. The value k represents the door chosen by Carol (when the contestant chooses the wrong door initially, Carol does not have a choice and can only open one of the remaining doors). Note that the sample space has exactly 12 outcomes, 6 of which have $i = j$, and the other 6 $i \neq j$.

Now the assignment of probabilities. A reasonable probability model is to think of all the outcomes where $i \neq j$ to have probability $\frac{1}{9} = \frac{1}{3} \cdot \frac{1}{3}$ because both the car is equally likely to be behind any of the 3 doors, and Bob is equally likely to choose any of the 3 doors. Carol in this case does not have a choice. For the outcomes where $i = j$, each of them occurs with probability $\frac{1}{18} = \frac{1}{9} \cdot \frac{1}{2}$, because now we take into account the fact that Carol can choose one out of 2 doors after the car position and Bob's initial choice are determined.

Let's return to the Monty Hall problem. Recall that we want to investigate the relative merits of the "sticking" strategy and the "switching" strategy. Let's suppose the contestant decides to switch doors. The event A we are interested in is the event that the contestant wins. Which outcomes (i, j, k) are in A? Well, since the contestant is switching doors, the event A does not include any outcomes where the contestant's initial choice j is equal to the prize door i. And all outcomes where $i \neq j$ correspond to a win for the contestant, because Carol must open the second non-prize door, leaving the contestant to switch to the prize door. So A consists of all outcomes (i, j, k) in which $i \neq j$. How many of these outcomes are there? Well, there are 6 pairs of (i, j) in which $i \neq j$. So $\mathbf{P}(A) = \frac{6}{9} = \frac{2}{3}$. That is, using the switching strategy, the contestant wins with probability $\frac{2}{3}$! It should be intuitively clear (and easy to check formally — try it!) that under the sticking strategy his probability of winning is $\frac{1}{3}$. (In this case, he is really just picking a single random door.) So by switching, the contestant actually improves his odds by a huge amount!

This is one of many examples that illustrate the importance of doing probability calculations systematically, rather than "intuitively." Recall the key steps in all our calculations:

- What is the **sample space** (i.e., the experiment and its set of possible outcomes)?
- What is the **probability** of each outcome (outcome)?
- What is the **event** we are interested in (i.e., which subset of the sample space)?
- Finally, compute the probability of the event by adding up the probabilities of the outcomes inside it.

Whenever you meet a probability problem, you should always go back to these basics to avoid potential pitfalls. Even experienced researchers make mistakes when they forget to do this — witness many erroneous "proofs", submitted by mathematicians to newspapers at the time, of their (erroneous) claim that the switching strategy in the Monty Hall problem does not improve the odds.