University of Waterloo
Faculty of Mathematics

Centre for Education in
Mathematics and Computing

Senior Math Circles
November 19, 2008
Probability I

Probability

Probability is the study of uncertain events or outcomes. Games of chance that involve rolling dice or dealing cards are one obvious area of application. We can apply probability models to study many other phenomena such as the number of vehicles passing through an intersection in a ten minute period, or the length of time it takes to search the internet on a topic.

In each of these situations, the result is uncertain. We will not actually know the result until after we collect the data. In all of these situations, we can use probability models to describe the patterns of variation we are likely to see.

Experiments and Sample Spaces

Definitions:
An *experiment* is any process that gives rise to data.
The data is often *numerical* such as the number on the top face of a die, or the blood pressure of a patient. However, data can also be *categorical* such as what comes up when a coin is tossed, or the colour that litmus paper turns when placed in an unknown solution.
We call the result of the experiment an *outcome*.
The *sample space* is the set of all possible outcomes of an experiment.

The main features of these experiments is that there is more than one possible outcome, and the experiments are repeatable. Also, when we repeat the experiment we will not necessarily get the same outcome.

Examples:
Our experiment consists of tossing a fair coin and observing the top face, then the sample space $S = \{H, T\}$

Our experiment consists of tossing a fair coin twice and observing the top face on each trial, then $S = \{HH, HT, TH, TT\}$

Our experiment consists of tossing a fair coin twice and counting the number of Heads, then $S = \{0, 1, 2\}$.

Our experiment consists of counting the number of cars passing through the intersection of King St. and University Ave. between 4 and 5 pm on a Wednesday evening. Then $S = \{0, 1, 2, 3, ...\}$. Note that some of these outcomes will have an extremely low chance of occurring. Some might be physically impossible. However, this is a convenient way to think of this sample space.

Our experiment involves measuring the weight of a randomly chosen student at the University of Waterloo. Here S is a subset of the real numbers.

Definitions:
A sample space is *discrete* if it contains a finite or a countably infinite number of outcomes. The sample spaces in the first four examples are discrete sample spaces.

A sample space is *continuous* if it contains a non-countably infinite number of possible outcomes. The last example above would be considered a continuous sample space. (In reality our measuring instruments only measure to a certain number of decimal points, so all sample spaces could be considered discrete. However, it is very convenient to allow for continuous sample spaces since we can then use continuous probability models.)

Events

Definition:
An *event* is a subset of a sample space. Another way to think of this is that an event is a collection of outcomes that share some property.

Examples:
A die is rolled. Define the event A to be that we observe and odd number. Then $S = \{1, 2, 3, 4, 5, 6\}$, and $A = \{1, 3, 5\}$.

A die is rolled. Define the event B to be that we observe a number bigger than 3. Then $B = \{4, 5, 6\}$.

Definitions:
The *union* of two events is the collection of outcomes that are in one event or the other event or both. We write $A \cup B$ and say "A or B".
The *intersection* of two events is the collection of outcomes that are in both events. We write $A \cap B$, or, more simply, AB, and say "A and B".

Example: Using the events A and B when rolling a die defined above, $A \cup B = \{1, 3, 4, 5, 6\}$ (odd *or* bigger than 3 *or both*), and $AB = \{5\}$ (odd *and* bigger than 3).

Probability Distributions for Discrete Sample Spaces

We often call the individual possible outcomes of an experiment *points*. With a discrete sample space, we can arrange the outcomes (points) so that one of them is the first, another is the second, etc. So let the points in the sample space be labeled $1, 2, 3, \ldots$. We assign to each point i of S a real number p_i which is called the *probability* of the outcome labeled i. These probabilities must be non-negative and sum to one.
So the p_i's must satisfy:
$p_i \geq 0$, and
$p_1 + p_2 + p_3 + \ldots = 1$

Definition:
Any set of p_i's that satisfies these results is a *probability distribution* defined on the discrete sample space.
Of course, to be useful in practice, we want a probability distribution that reflects the experimental situation.

Examples:
A *fair* die is rolled and we note the number on the top face. Then $S = \{1, 2, 3, 4, 5, 6\}$. Since the die is fair, the probability distribution $p_1 = \frac{1}{6}, p_2 = \frac{1}{6}, \ldots, p_6 = \frac{1}{6}$ will describe the situation.

Suppose the die was weighted so that odd numbers occurred twice as often as even numbers. Then a reasonable probability distribution would be $p_1 = \frac{2}{9}, p_2 = \frac{1}{9}, p_3 = \frac{2}{9}, p_4 = \frac{1}{9}, p_5 = \frac{2}{9}, p_6 = \frac{1}{9}$. Note that these sum to one and are each greater than or equal to 0.

J. G. Kalbfleisch in his book *Probability and Statistical Inference, Volume 1: Probability* gives an example of a brick-shaped die. Suppose a die is brick-shaped, with the numbers 1 and 6 on the square faces and the numbers 2,3,4,5 on the rectangular faces. The square faces are $2cm \times 2cm$, and the rectangular faces are $2cm \times (2 + k)cm$, where $-2 < k < \infty$. Then we have $p_1 = p_6$, and $p_2 = p_3 = p_4 = p_5$. If $p_1 = p_6 = p$ and $p_2 = p_3 = p_4 = p_5 = q$, then $2p + 4q = 1$, so $p = \frac{1}{2} - 2q$.
Now these probabilities will be a function of k. If $k = 0$, then $p = q = \frac{1}{6}$. As k approaches -2, q approaches 0. As k gets bigger, p approaches 0.
We could rewrite these probabilities as $q = \frac{1}{6} + \theta$ and $p = \frac{1}{6} - 2\theta$, where θ is a function of k. By creating several dice with different values of k, we could experiment to try to determine the relationship between θ and k. Alternatively, we could try to use the laws of physics and develop a mathematical model to relate θ and k.

Probabilities of Events

Once we have defined a probability distribution for the individual outcomes, we define the probability of any event defined on the sample space as the sum of the probabilities assigned to the outcomes in that event.

Examples:
If A is the event that we roll an odd number when a die is rolled, the probability of event A is $P(A) = p_1 + p_3 + p_5$. From the three previous examples, this would be $\frac{1}{2}$ if the die were fair, $\frac{2}{3}$ if it were weighted towards the odd numbers as described, and $\frac{1}{2}$ if it were brick shaped, regardless of k.

We can find $P(A \cup B)$ in a couple of ways. First, we can add the probabilities of the outcomes in A or B or both. So if A is the event that we get an odd number, and B is the event we get a number greater than 3 when we roll a die, $P(A \cup B) = p_1 + p_3 + p_4 + p_5 + p_6$. For the probability distributions above, this will be $\frac{5}{6}$ or $\frac{8}{9}$ or $\frac{5}{6} - \theta$.
Note though, that if we take the outcomes in A and add to them the outcomes in B, we have all the outcomes in $A \cup B$, except the outcomes in $A \cap B$

are there twice. This suggests that $P(A \cup B) = P(A) + P(B) - P(AB)$. Here $P(A \cup B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{6} = \frac{5}{6}$, for the case of the fair die. We can verify the other cases.

Conditional Probability

It is very common in many applications to want to know the probability of an event *conditional on* the occurrence of some other event. For example, hypertension (high blood pressure) is very common the population. Being overweight is also common. If we sample someone from the population and see he is overweight, what is the probability that he has high blood pressure as well.

Here let A be the event that an individual has high blood pressure and B be the event that an individual is overweight. Then the probability that someone has high blood pressure *given that* he is overweight is written $P(A|B)$.

To calculate this probability, we note that we only want to deal with the overweight people. So our sample space is confined only to those people in the set B. Of those overweight people, we only want the hypertensive ones. So they must be overweight *and* hypertensive (so in the set AB).

So we have $P(A|B) = \dfrac{P(AB)}{P(B)}$. This also gives us a useful way to calculate $P(AB)$, since $P(AB) = P(A|B)P(B)$
Note that we could just as well compute $P(AB) = P(B|A)P(A)$. If we put these together, we get a very useful theorem named after Thomas Bayes ($\sim$ 1702 - 1761).

Bayes' Theorem:
Let A and B be events defined on the same sample space.
Then $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$.

Example:
For the events A (we roll an odd number) and B (we roll a number bigger than 3), we have the probability that we roll an odd number given that we roll a number bigger than 3 is $P(A|B) = \dfrac{P(AB)}{P(B)} = \dfrac{1}{3}$, for the equal

probability case. This makes sense since there are three numbers bigger than 3 and only one of them is odd. Note that if we use the case where the odds are weighted more heavily, $P(A|B) = (\frac{2}{9}) \div (\frac{4}{9}) = \frac{1}{2}$.

Independent Events

Definition:
We define events A and B to be *independent* if and only if $P(AB) = P(A)P(B)$.

Note that if events are independent,
$P(A|B) = \dfrac{P(AB)}{P(B)} = \dfrac{P(A)P(B)}{P(B)} = P(A)$. This makes intuitive sense since it says that if the events are independent, knowing that B has occurred does not change the probability that A also occurred.

Example:
Are the events A and B for the die independent? Here for the equal outcome case, $P(AB) = \frac{1}{6} \neq P(A)P(B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$. So A and B are not independent. Knowing you have a number greater than 3 decreases your chances of getting an odd number.

Reference:
Kalbflesich, J.G., *Probability and Statistical Inference, Volume 1: Probability, Second Edition*, Springer-Verlag, New York, 1985.

Problem Set

1. Suppose a die is weighted so that when it is rolled, the probability of seeing any number on the top face is proportional to the number on the face. Give the probability distribution that would apply.

2. Suppose we draw one card from a well-shuffled deck. Let A be the event that we get a spade, and B be the event we get an ace. Are these events independent? Suppose we add an extra ace of spades to the deck. Are the events A and B independent now?

3. Two teams play a best-of-seven series. All games must end with one team winning and play stops as soon as one of the teams has won four games. Describe a sample space for this experiment. If the teams are evenly matched, what probabilities should be assigned? What is the probability that the series will go the full seven games?

4. A bag contains some blue and green hats. On each turn, Julia removes one hat without looking, with each hat in the bag being equally likely to be chosen. If it is green, she adds a blue hat into the bag from her supply of extra hats, and if it is blue, she adds a green hat to the bag. The bag initially contains 4 blue hats and 2 green hats. What is the probability that the bag again contains 4 blue hats and 2 green hats after two turns?

5. Billy and Crystal each have a bag of 9 balls.The balls in each bag are numbered 1 to 9. Billy and crystal each remove one ball from their own bag. Let b be the sum of the numbers on the balls remaining in Billy's bag. Let c be the sum of the numbers on the balls remaining in Crystal's bag. Determine the probability that b and c differ by a multiple of 4.

6. In a 4×4 grid, three coins are randomly placed in different squares. Determine the probability that no two coins lie in the same row or column.

7. You want to find someone whose birthday matches yours, so you approach people on the street asking each one for their birthday. Ignoring leap year babies, if we assume that each person you approach has probability $\dfrac{1}{365}$ of sharing your birthday, how many people would you have to approach to have a 50% chance of finding someone with the same birthday as you?

8. If there are n people in a room, what is the probability that there will be at least two people with the same birthday? What is the smallest value of n so that this probability is greater than 50%?

9. Efron's Dice: We are going to play a game with a set of dice. It is a simple game. You get a die and I get a die and we roll them. Whoever gets the highest number wins. The only rules are that since it is my game we must play with my dice, but you will always know all about the dice, and, even more importantly, you will always get first choice of the die you want to roll. I will then have to choose another one. Here are the dice:
   Die A: Has 4's on four faces and 0's on two faces.
   Die B: Has 3's on all six faces.
   Die C: Has 2's on four faces and 6's on two faces.
   Die D: Has 1's on three faces and 5's on three faces.
   What is the probability that Die A's number will be greater than Die B's? Die B's greater than Die C's? Die C's greater than Die D's? Which die do you want?

10. Pairwise Best - Worst Paradox: You are to play a game in which there are three urns. In each urn there are some balls and each ball has a number on it. You choose an urn and your opponent chooses an urn from the remaining ones. You each choose one ball at random from your urn. Who ever draws the ball with the higher number wins. In urn A, there is one ball numbered 3. In urn B, there are 56 balls numbered 2, 22 balls numbered 4, and 22 balls numbered 6. In urn C, there are 51 balls numbered 1 and 49 balls numbered 5. Which urn would you select? Which urn would be your worst choice? Now, a third player comes along. Each of you will get an urn and each will draw one ball. The player with the highest number will win. Now which urn do you choose?

11. In the game *Risk*, there are times when one player tosses three dice and another player tosses two dice. The person who tosses the largest number "wins". What is the probability distribution for the largest number tossed by the player who tosses three dice? What is the probability that the player who tosses the three dice wins?

12. (From Kalbfleisch, J. G. *Probability and Statistical Inference, Volume 1: Probability*)There is a box that contains two coins. One coin is a regular coin with Heads on one side and Tails on the other. The other coin has Head on both sides. You select a coin and toss it, and it comes up Heads. What is the probability that you tossed the two headed coin?

13. (From Kalbfleisch, J. G. *Probability and Statistical Inference, Volume 1: Probability*)The probability that a student knows the correct answer to a question on a multiple choice exam is p. If he doesn't know the correct answer, he chooses one of the k possible answers at random. If the student correctly answers the question,what is the probability he knew the answer?

14. (From Kalbfleisch, J. G. *Probability and Statistical Inference, Volume 1: Probability*) There are two diagnostic tests for a disease. Among those who have the disease, 10% give negative results on the first test, and independently of this, 5% give negative results on the second test. Among those who do not have the disease, 80% give negative results on the first test, and, independently, 70% give negative results on the second test. Twenty percent of those tested actually have the disease.

    (a) If both tests are negative, what is the probability that the person tested has the disease?

    (b) If both tests are positive, what is the probability that the person tested has the disease?

    (c) If the first test gives a positive result, what is the probability that the second test will also be negative?

15. The Monte Hall Problem: The famous "Monte Hall Problem" goes something like this. You are a contestant on a game show where Monte Hall is the host. You have gotten to the point where you are to choose a prize for your efforts on the show. You are shown three doors, and told that there is an outstanding prize behind one of three doors, and there are prizes of no real value behind the other two doors. (The good prize could be a car, a fabulous vacation, etc., while the other "prizes" could be a goat and a year's supply of carrots, or a leaky row boat and a fishing line). You select a door (say door number 2). Before you can see what is behind the door you selected, Monte says "Let me show you what is behind door number 3". He then reveals one of the lousy "prizes". After much laughter from the audience, Monte says, "You have selected door number 2. Do you now want to switch your choice and select door number 1?" The audience will shout out their advice, and after a short time of frantic indecision, you

choose to either stay with your original pick (door number 2), or switch to door number 1 and accept the prize behind that door.

So should you switch? Some would argue that since there are now two doors remaining and the prize is equally likely to be behind either door, there is no benefit to switching. Is this argument correct?

16. There are n seats on an airplane and n passengers have bought tickets. Unfortunately, the first passenger to enter the plane has lost his ticket and, so, he just chooses a seat at random and sits in it. Thereafter, each of the remaining passengers enters one at a time and either sits in their assigned seat if it is empty, or, if someone is sitting in their seat, chooses a seat at random from those that are empty. What is the probability that the last passenger to enter will end up sitting in her assigned seat?