



University of Waterloo  
Faculty of Mathematics



Centre for Education in  
Mathematics and Computing

## Senior Math Circles November 19, 2008 Probability II

### Probability

#### Counting

There are many situations where each of the outcomes in a sample space is equally likely. In these situations, the probability of an event  $A$  can be calculated by dividing the number of outcomes in  $A$  by the total number of outcomes in  $S$ .

If we have 3 objects that are all different, we can arrange them in a row in 6 ways. For example, if the objects are denoted by the letters  $X, Y, Z$ , then there 6 ways to arrange them:

$XYZ, XZY, YXZ, YZX, ZXY, ZYX$ . We have 3 choices for the first letter, for each of those we have 2 choices for the second letter, and then there is only one choice left for the third letter. So there are  $3 \times 2 \times 1 = 6$  different ways they can be arranged.

If there are  $n$  distinct objects, we can arrange them in

$n \times (n - 1) \times (n - 2) \times \dots \times 2 \times 1$  ways.

We define  $n!$  as  $n! = n \times (n - 1) \times (n - 2) \times \dots \times 2 \times 1$ , to represent this number.

In many applications, we need to know the number of ways to select a subset of outcomes from the entire set. For example, we might want to select 2 items from 4 where we do not really care about the order we select them (for example,  $XY$  and  $YX$  are treated the same way).

If we have letters  $W, X, Y, Z$ , there are 6 subsets of size 2:  $WX, WY, WZ, XY, XZ, YZ$ . To count these, we could say that there are 4 choices for the first letter, and three for the second. But that counts  $XY$  and  $YX$  differently. There

are  $2!$  ways to arrange the two letters we chose, so each selected pair is counted twice, so, overall, there are  $\frac{4 \times 3}{2!} = 6$  ways to choose 2 letters from 4 if the order does not matter.

We define a special notation for this, and say “4 choose 2”:  $\binom{4}{2} = \frac{4!}{2!2!}$  is the number of ways to choose 2 items from 4 if order does not matter. In general, the number of ways to choose  $x$  items from  $n$  if order does not matter is:  $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ .

Examples:

In Lotto 649, you choose 6 numbers from 49. A random set of 6 numbers is chosen on the night of the draw. To calculate the probability that your numbers will be drawn, we first note that the sample space consists of  $\binom{49}{6}$  possible ways to choose 6 numbers from 49. Because the numbers are chosen at random, each of the  $\binom{49}{6} = 13,983,816$  possible combinations is equally likely. Hence the probability that your numbers will be drawn is:  $P(\text{win}) = \frac{1}{\binom{49}{6}} = \frac{1}{13,983,816} = 0.00000007151124$ .

Problem:

The various other prizes from Lotto 649 are taken from the Lotto 649 website and given below. Can you verify them? (Some are quite tricky).

Match	Prize	Probability
6 of 6	Share of 80.50% of Pools Fund	$\frac{1}{13,983,816}$
5 of 6 + Bonus	Share of 5.75% of Pools Fund	$\frac{1}{2,330,636}$
5 of 6 (no bonus)	Share of 4.75% of Pools Fund	$\frac{1}{55,491}$
4 of 6	Share of 9.00% of Pools Fund	$\frac{1}{1,032}$
3 of 6	\$10	$\frac{1}{57}$
2 of 6 + Bonus	\$5	$\frac{1}{81}$

## Independent Experiments

If we conduct two experiments so that the outcome of the second experiment is not influenced by the outcome of the first experiment, we say that the experiments are *independent*.

Example:

If experiment 1 consists of tossing a fair coin, and experiment 2 consists of rolling a fair dice independently of the results of experiment 1. Then the

sample space for the combined experiment (experiment 1 followed by experiment 2) is:  $S = \{(H, 1), (H, 2), (H, 3), (H, 4), (H, 5), (H, 6), (T, 1), (T, 2), (T, 3), (T, 4), (T, 5), (T, 6)\}$ .

Since the coin and die are fair, then the probability of a Head on the coin is  $\frac{1}{2}$ , and the probability of a 1 on the die is  $\frac{1}{6}$ . All 12 events in this sample space should be equally likely since the result of the second experiment did not depend on the result of the first. Hence, the probability of the combined outcomes (H,1) is  $\frac{1}{12} = \frac{1}{2} \times \frac{1}{6}$ .

In general, if  $p_i$  is the probability of outcome  $i$  in experiment 1, and  $q_j$  is the probability of outcome  $j$  in experiment 2, and the experiments are independent, then  $p_i \times q_j$  is the probability observing of outcome  $i$  in experiment 1 followed by outcome  $j$  in experiment 2.

## Independent Repetitions

A special case of independent experiments occurs when experiment 2 is an *independent repetition* of experiment 1. Then, unless things have changed somehow, the probability distribution for the outcomes will be the same for each repetition. So the probability of outcome  $i$  on repetition 1 and outcome  $j$  on repetition 2 is  $p_i \times p_j$ .

Examples:

A fair coin is tossed 4 times independently. Since there are 2 possibilities for each toss (H or T), there will be  $2^4$  possible outcomes. Since the coin is fair, each of these will be equally likely, so the probability distribution will assign probabilities  $\frac{1}{16}$  to each point in the combined sample space.

A fair coin is tossed, independently, until a head appears. Then the sample space is  $S = \{H, (T, H), (T, T, H), (T, T, T, H), \dots\}$ .

Since the tosses are independent, the probabilities assigned to the points of  $S$  are:  $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots\}$ .

This is our first example of a finite sample space with an infinite number of outcomes. But can a sum of an infinite number of real numbers equal 1?

Our sum is  $T = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots$ . This is an example of an *infinite geometric series*, where each term after the first in the sum is  $\frac{1}{2}$  of the term before it. There is a formula for finding the sums of such series. If  $a$  is the

first term in the sum, and  $r$  is the common ratio between all terms in the sum, then  $T = \frac{a}{1-r}$ , providing  $-1 < r < 1$ . This works here since  $a = \frac{1}{2}$ , and  $r = \frac{1}{2}$ , so  $T = 1$ .

## Bernoulli Trials - The Binomial Distribution

A common application occurs when we have *Bernoulli Trials* - independent repetitions of an experiment in which there are only two possible outcomes. Label the outcomes as a “Success” or a “Failure”. If we have  $n$  independent repetitions of an experiment, and  $P(\text{Success}) = p$  on each repetition, then:

$$P(x \text{ Successes in } n \text{ repetitions}) = \binom{n}{x} p^x (1-p)^{(n-x)}.$$

Example:

A multiple choice exam consists of 25 questions with 5 responses to each question. If a student guesses randomly for each question, what is the probability he will answer 13 questions correctly? at least 13 questions correctly?

Solution:

This is a Bernoulli Trials situation with  $P(\text{Success}) = \frac{1}{5}$ .

$$\text{Hence, } P(13 \text{ correct}) = \binom{25}{13} \left(\frac{1}{5}\right)^{13} \left(\frac{4}{5}\right)^{12} = 0.00029.$$

The probability of at least 13 correct, is:

$$P(13 \text{ correct}) + P(14 \text{ correct}) + \dots + P(25 \text{ correct}) = 0.00037.$$

## Negative Binomial Distribution

In a sequence of Bernoulli Trials, suppose we want to know the probability of having to wait  $x$  trials before seeing our  $k$ th success. If the  $k$ th success occurs on the  $x$ th trial, we must have  $k-1$  successes in the first  $x-1$  trials and then a success on the  $x$ th trial. So the probability is:

$$P(k\text{th Success on } x\text{th trial}) = \binom{x-1}{k-1} p^{k-1} (1-p)^{x-k} * p = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

Note that there is no upper limit on  $x$ , so our sample space here is

$$S = \{k, k+1, k+2, \dots\}.$$

## Expected Values

In the above examples, we are counting the number of successes in  $n$  trials or the number of trials until some event occurs. It is natural to want to know, *on average* how many successes will we see, or *on average* how many trials until some event occurs.

Definition:

“*On average*” is meant here to be a theoretical concept referring to repeating the process over and over again and averaging the numbers from an *infinite* number of trials. We refer to this theoretical average as the *Expected Value* for the number of successes or the number of trials.

If  $p_i$  represents the probability of  $i$  successes, or  $i$  trials until some event, then the *Expected Number* of successes or trials is

$0 \times p_0 + 1 \times p_1 + 2 \times p_2 + \dots + n \times p_n$  where we are counting the number of successes. If we are counting the number of trials, we get a sum that goes to infinity.

Example:

If you buy one ticket on Lotto 649 per week, how many weeks *on average* will you have to wait until you win the big prize?

Solution:

From the earlier example, the probability of winning the big prize on any draw is:  $p = \frac{1}{\binom{49}{6}} = \frac{1}{13,983,816}$ .

So the probability that you have to wait  $x$  weeks to win the big prize is:

$$P(\text{wait } x \text{ weeks}) = (1 - p)^{x-1}p.$$

Thus, the expected number of weeks to wait is

$$E = 1(p) + 2(1 - p)(p) + 3(1 - p)^2(p) + \dots$$

This is a combined arithmetic-geometric sequence. There are several ways to sum this. If you consider  $E - (1 - p)E$ , you can verify that  $E = 1/p$ . So you need to wait 13,983,816 weeks on average to win the big prize.

Example (Montmort Letter Problem):

Suppose  $n$  letters are distributed at random into  $n$  envelopes so that one letter goes into each envelope. On average, how many letters will end up in their correct envelopes? Does this depend on  $n$ ?

There are two ways to do this question - the hard way and the easy way. Figuring out the probability that  $x$  letters end up in the correct envelopes is not that easy. So let's try this another way.

Define  $X_i$  to be 0 if letter  $i$  does not end up in the correct envelope and define  $X_i$  to be 1 if letter  $i$  does end up in the correct envelope. Then  $T = \sum X_i$  is the number of letters in the correct envelopes. Then the expected value of  $T$  is the sum of the expected values of the  $X_i$ 's. Now  $P(X_i = 1) = \frac{1}{n}$ , since there are  $n$  choices for letter  $i$ . So the expected value of  $X_i$  is also  $\frac{1}{n}$ , and so the expected value of  $T$  is  $n(\frac{1}{n}) = 1$ .

So, on average, regardless of  $n$ , we expect to find only one letter in the correct envelope.

## Problem Set

1. The Paradox of Chevalier de Mere: This famous problem was presented by a French gambler (Chevalier de Mere) to Pascal who, in turn, discussed it with Fermat. Pascal and Fermat began the formal study of probability. The problem was “Is it more likely to roll at least one 1 in four rolls of a fair die, or to roll at least one pair of 1’s in 24 rolls of two dice”? de Mere argued that the probability of a 1 when a fair die is rolled is  $\frac{1}{6}$ . The probability of rolling two 1’s when two dice are rolled is  $\frac{1}{36}$ . So he claimed that the pair of dice must be rolled 6 times more often to have the same probability of rolling at least one pair of 1’s, as rolling at least one 1 in four rolls of a fair die. Was he right?
2. For Lotto 649, suppose you buy 5 different tickets for a draw. What is the probability that you win the grand prize? If you use this strategy every week for a year, what is the probability that you will win the grand prize at least once? How many tickets would you need to buy every week for one year so that the probability you win the grand prize at least once is 10%?
3. 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  $\frac{1}{365}$  of sharing your birthday, *on average* how many people would you have to approach to find someone with the same birthday as you?
4. Eight red blocks and four blue blocks are arranged at random in a row. What is the probability that no two blue blocks are side by side?
5. You will be presented with 10 offers to buy your car. You, quite naturally, want to choose the best offer. However, you see the offers one at a time, and if you refuse any offer, you do not get another chance at that offer. You know if you choose an offer at random, there is probability  $\frac{1}{10}$  of choosing the best offer. So you try the following strategy: Examine the first four offers but refuse each one, noting which of these is the best. Then for each of the other offers, refuse it if it is lower than the best one you saw in the first four, and accept it if it is better than the best one you saw in the first four. If you have not selected an offer by the time the last one comes in, you take the last one (by now you are desperate). What is the probability that by using this strategy you will select the best offer?
6. In a sequence of  $p$  zeros and  $q$  ones, the  $i$ th term,  $t_i$ , is called a *change point* if  $t_i \neq t_{i-1}$ , for  $i = 2, 3, 4, \dots, p + q$ . For example, the sequence 0, 1, 1, 0, 0, 1, 0, 1 has  $p = q = 4$ , and five change points  $t_2, t_4, t_6, t_7, t_8$ . For all possible sequences of  $p$  zeros and  $q$  ones, with  $p \leq q$ , determine:
  - (a) the minimum and maximum number of change points.
  - (b) the average number of change points.
7. A group of  $n$  married couples arrives at a dinner party and is seated around a circular table. The distance between the members of a couple is defined to be the number of people sitting between them measured either clockwise or counter-clockwise, whichever gives the smaller result.
  - (a) Considering all possible seating arrangements of the  $2n$  people, what is the average distance between a particular couple  $A$ ?
  - (b) Considering all possible seating arrangements for the  $2n$  people, what is the average number of couples, per arrangement, where both members of the couple are seated side-by-side?

- (c) Repeat the question if the  $2n$  people are attending the theatre and are seated in a row. (The distance between members of a couple is defined here as the number of people sitting between them).
8. A sample of size  $n$  is chosen at random *with replacement* from the integers  $1, 2, \dots, N$ . (In sampling *with replacement*, an item is drawn, we note which item it was, and it is then replaced. Note that not all integers need to be different under sampling with replacement). Find the average number of *different* integers in the sample. What happens if  $n \ll N$ ?
9. A random graph with  $n$  vertices is generated by connecting pairs of vertices at random. Each of the possible  $\binom{n}{2}$  possible edges is inserted with probability  $p$ , independently of other edges. Find the average number of triangles in the graph.
10. A group of  $n$  people each toss their hats in the air and the hats are picked up randomly. Each person who gets his or her own hat leaves and the remaining people toss the hats again. The process continues until every person has received his or her own hat again. Find the average number of rounds of tosses required.