Lecture Notes on Probability

MA 111 Spring 2010

Scott Taylor, Colby College

CONTENTS

1. Probability Spaces

Definition 1. An **experiment** is a task which results in a well-defined set of possible outcomes.

The following are examples of experiments (in the probability theory sense):

(a) Flip a coin once and record the result.
(b) Flip a coin twice and record the result in order.
(c) Roll a die and record the result
(d) Randomly choose a Colby student and ask him or her for his or her height.
(e) Give a 16 year old a driver's license and record whether or not they get into an accident within a year.

Definition 2. The **sample space** for an experiment is the set of possible outcomes.

For the experiments above, here are the sample spaces:

(a) $\{H,T\}$
(b) $\{HH,HT,TH,TT\}$
(c) $\{1,2,3,4,5,6\}$
(d) This one is trickier. If we assume that no Colby student is under 3 feet tall or over 8 feet tall, we can take the sample space to be the set of all real numbers between 3 and 8. Other answers, based on different assumptions, are possible.
(e) $\{$ Yes there was an accident, No there wasn't an accident $\}$.

Definition 3. An **event** is a subset of a sample space.

(a) There are four possible events in the sample space $\{H,T\}$. These events are $\{\},\{H\},\{T\},\{H,T\}$.
(b) There are 16 possible events in the sample space $\{HH,HT,TH,TT\}$. What are they?
(c) The phrase "roll an even number" corresponds to the event $\{2,4,6\}$ in the sample space $\{1,2,3,4,5,6\}$.
(d) The phrase "has a height between 4 feet and 5 feet" corresponds to the event $\{h : 4 \leq h \leq 5\}$ in the sample space of the experiment "ask a Colby student their height and record the result."

There is a big mathematical difference between sample spaces with finitely many outcomes and sample spaces with infinitely many outcomes. *From now on, we will only consider finite sample spaces*, unless otherwise indicated.

Definition 4. A **probability space** consists of a sample space for an experiment together with non-negative numbers (called **probabilities**) associated to each outcome, so that the sum of all the probabilities is 1.

The above definition assumes that there are only finitely many outcomes; if there are infinitely many, probabilities need to be attached to events rather than to outcomes. The requirement that the probabilities sum to 1 then needs to be replaced with a different (but analogous) requirement. We won't go into that here.

Example 1. Consider the sample space $\{H,T\}$. We could define probabilities $P(H) = 1/2$ and $P(T) = 1/2$. This sample space, with these probabilities, is then a probability space.

Example 2. Consider the sample space $\{H,T\}$. We could define probabilities $P(H) = .75$ and $P(T) = .25$. This sample space, with these probabilities, is then a probability space.

If we want the probability space to model what happens in the real world, we should assign probabilities to outcomes in such a way that the probability of an outcome is approximately the number of times the outcome occurs divided by the number of times the experiment is performed. The application of probability to real world phenomena is then tricky, since there are some experiments that cannot physically be performed more than once. For example, "At precisely 12 noon EST on April 20, 2010 in Waterville, Maine, measure the temperature".

Definition 5. A probability space is a **uniform probability space** if all outcomes in the sample space have the same probability. We also say that the experiment is **fair** or **unbiased**.

Example 3. Consider the sample space $\{1,2,3,4,5,6\}$ with probabilities $P(n) = 1/6$ for all n. Then this is a uniform probability space.

Example 4. Consider the sample space $\{1,2,3,4,5,6\}$ with probabilities $P(1) = 1/2$ and $P(n) = 1/10$ for all $n \neq 5$. Then this is not a uniform probability space.

We have defined probabilities for outcomes, we also need to define probabilities for events.

Definition 6. Suppose that E is an event in a probability space and that $E = \{x_1, x_2, \ldots, x_n\}$ where each x_i is an outcome. Then the probability $P(E)$ of the event E is defined to be

$$P(E) = P(x_1) + P(x_2) + \ldots + P(x_n).$$

Example 5. Consider the probability space $\{1,2,3,4,5,6\}$ corresponding to rolling a die with

$$\begin{aligned} P(1) &= .1 \\ P(2) &= .2 \\ P(3) &= .3 \\ P(4) &= .4 \\ P(5) &= 0 \\ P(6) &= 0 \end{aligned}$$

(a) The probability of rolling a number less than or equal to 3 is

$$P(1)+P(2)+P(3) = .1+.2+.3 = .6$$

(b) The probability of rolling an even number or a 5 is

$$P(2)+P(4)+P(5)+P(6) = .2+.4+0+0 = .6$$

Notice that if E is an event in a *uniform* probability space then

$$P(E) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in the sample space}}.$$

Example 6. Consider the uniform probability space corresponding to flipping a coin 3 times and recording the results. Since there are three coin flips and two possible outcomes for each coin flip, there are 8 outcomes in the sample space. Since it is specified that it is a uniform probability space, these 8 outcomes are all equally likely.

Calculate the probability of obtaining exactly one H.

Answer: If we mark place the results of the 3 flips in the blanks:

______ ______ ______

we see that there are 3 possibilities for where the H can go. Once the H is placed, the other two spots must be T. Thus, there the event E contains 3 outcomes. Since the probability space is uniform, $P(\text{exactly 1 } H) = 3/8$.

Example 7. Consider the uniform probability space corresponding to flipping rolling a die three times and recording the results. What is the probability of rolling two 6s or two 1s in a row?

Answer:

Since there are 3 rolls and six possible outcomes for each roll, there are $6^3 = 216$ outcomes in the sample space. Since it is specified that it is a uniform probability space, these 216 outcomes are all equally likely.

Consider first the possibility of obtaining two consecutive 6s. If we mark the results of the 3 rolls in the blanks:

$$\_\_\_\_\_\_ \quad \_\_\_\_\_\_ \quad \_\_\_\_\_\_$$

we see that there are 2 possibilities for where the first 6 can go. Once the first 6 is placed, the space immediately after it must also be a 6. The third space can be any of the numbers $\{1,\ldots,6\}$. Making an apparent total of $2\cdot 6 = 12$ possible ways of getting two 6s in a row. *However*, we have double counted the outcome 666 and so there are actually only 11 possibilities .

A similar analysis shows that there are 11 ways of obtaining two 1s in a row. Since there are only 3 rolls, it is impossible to get two 1s in a row and two 6s in a row. Consequently the event contains 22 outcomes. Thus the probability of rolling two 6s or two 1s in a row is $22/216 \approx .1$.

Definition 7. If E and F are sets, $E\cup F$ is the set of all things which are in E **or** F (or both). The set $E\cap F$ is the set of all things which are in **both** E and F.

Example 8. Let $E = \{1,2,a,b,c\}$ and $F = \{a,c,e,g\}$. Then $E\cup F = \{1,2,a,b,c,e,g\}$ and $E\cap F = \{a,c\}$.

In general, if E is an event and if F is an event, then the event phrased as "an outcome in E occurs **or** an outcome in F occurs" is $E\cap F$. The event phrased as "an outcome in E occurs **and** an outcome in F occurs" is $E\cup F$. The following is easy to prove:

$$P(E\cup F) = P(E) + P(F) - P(E\cap F).$$

2. Independence

Definition 8. Suppose that we perform an experiment A and then perform an experiment B and write down the result of A and then the result of B. This is a new experiment which is a combination of A and B. This new experiment has a sample space and the number of outcomes in the sample space is the number of outcomes from A times the number of outcomes from B. If the outcome of B is unaffected by the outcome of A, then we say that the events in the sample space of B are **independent** of the events in the sample space of A. We will also say that B is independent of A.

Definition 9. If the experiment B is independent of A and if E is an event in the sample space of A and if F is an event in the sample space of B then we define

$$P(\text{“}E \text{ then } F'') = P(E)\cdot P(F).$$

Example 9. Let A and B both be the experiment that a fair coin is tossed once and the result recorded. Show that the probability space for the experiment "do A then B" is uniform.

Answer: We could write the results from the experiment in the blanks below.

$$\underset{A \text{ outcome}}{\underline{\qquad\quad}} \qquad \underset{B\text{outcome}}{\underline{\qquad\quad}}$$

Since each outcome from A is equally likely and since each outcome from B is equally likely, in terms of probability we have:

$$\underset{A \text{ outcome}}{\underline{1/2}} \quad \cdot \quad \underset{B\text{outcome}}{\underline{1/2}} \quad = 1/4$$

Since there are 4 possible outcomes, this experiment is unbiased.

Important Note: If two experiments are not independent, you **cannot** multiply probabilities in the way we just did.

Definition 10. If k objects are to be chosen (without replacement) from n objects, the number of ways of doing this is given by the formula:

$$\text{"}n \text{ choose } k'' = \frac{n!}{k!(n-k)!}.$$

Example 10. The number of ways of choosing 3 letters out of a 26 letter alphabet is:

$$\text{"}26 \text{ choose } 3'' = \frac{26!}{3!23!} = 2600.$$

Example 11. Suppose that a fair coin is tossed 8 times and the results recorded. What is the probability that exactly 3 Heads came up?

Solution: The sample space has $2^8 = 256$ outcomes, each of them equally likely (by an argument similar to the preceding example). How many of these outcomes contain exactly 3 heads? There are 8 blanks to fill:

$$\underline{\qquad\quad} \quad \underline{\qquad\quad} \quad \underline{\qquad\qquad\qquad} \quad \underline{\qquad\quad} \quad \underline{\qquad\quad} \quad \underline{\qquad\quad} \quad \underline{\qquad\quad}$$

Exactly 3 of these must contain H. Since 8 choose 3 is 56, the probability of obtaining exactly 3 heads in 8 throws is $56/256 \approx .22$.

2.1. **The Unfinished Game.** Suppose that two players Blaise and Pierre are betting on a game. In this game a coin is to be flipped 5 times and whoever wins the most flips gets the pot of money. For simplicity assume that Blaise always calls H and Pierre always calls T. The goal of the section is to find a method of dividing the pot of money so that if the game is interrupted, the money in the pot can be divided fairly.

Observation 1: The method should depend on how many throws *of the remaining throws* each player needs to win in order to win the game.

Observation 2: If we repeatedly interrupt the game, divide the money using our method, and then resume the game, on average our method should give a method of dividing the money the same (on average) as the method obtained by completing the game.

Example 12. Suppose the game is interrupted after 3 throws and that Blaise is ahead 2 to 1.

Then Blaise needs to win at least one more throw in order to win the game, while Pierre needs to win at least 2 more throws. The possible outcomes corresponding to a win for Blaise are:

$$\text{Blaise wins } = \{HH, HT, TH\}$$

Since two throws remain there are 4 possible ways of finishing the game, so $P(\text{B win}) = 3/4$.

Thus, we should give $3/4$ of the money to Blaise and 1/4 to Pierre. I leave it as an exercise to show that this method of dividing the pot satisfies the requirements.

Example 13. If we reword the problem slightly we get a different game. Suppose that coin is to be flipped until either Blaise or Pierre wins 3 flips. Whoever is the first to win 3 flips, wins the game. If the game is interrupted when Blaise has one 2 flips and Pierre has one 1 flip how should the pot of money be divided?

In this case the possible outcomes of finishing the game are:

$$\{H, TH, TT\}.$$

Of these, the even that Blaise wins is $\{H, TH\}$. The probability of the 4th flip being H is 1/2. The probability of the fourth flip being T and the fifth being H is $1/4$. Thus, the probability that Blaise wins is:

$$P(\text{ Blaise wins }) = P(H) + P(TH) = 1/2 + 1/4 = 3/4.$$

So Blaise should again get 3/4 of the money and Pierre should get 1/4.

Example 14. Suppose that a fair coin is to be flipped 7 times and whoever wins 4 flips wins the game. If the game is interrupted when Blaise has 2 flips and Pierre has 1 flip, how should the money be divided?

Solution: There are 4 flips left to the game, so the sample space has $2^4 = 16$ possible outcomes, all equally likely. Blaise needs at least 2 flips to win. Of the 16 possible outcomes, 4 choose 2 = 6 have exactly two flips. 4 choose 3 = 4 have exactly three flips and 4 choose 4 = 1 have exactly 4 flips. Thus, there are 6 + 4 + 1 = 11 ways for Blaise to win. Blaise's probability of winning is, therefore, 11/16. Blaise should be given 11/16 of the pot.

3. Expected Value

Example 15. Suppose that Troilus and Cressida are tossing a coin. For each toss there is a 2/3 probability that Troilus will win and a 1/3 probability that Cressida will win. If Troilus wins, Cressida pays him \$3. If Cressida wins, Troilus pays her \$6.50. How much, on average, will Cressida win or lose?

Solution: Suppose 3000 games have been played. Troilus has won approximately (2/3)(3000) of the games. And so Cressida has paid him approximately, (\$3)(2/3)(3000) = \$6000. Cressida has won approximately (1/3)(3000) games and has won from Troilus approximately (\$6.50)(1/3)(3000) = \$6500. On average, therefore she wins approximately (\$6500 - \$6000)/3000 = \$0.17.

The calculations in the above example are worth looking at in more detail:

The amount Cressida wins on average is approximately:

$$\frac{\frac{-\$3}{1}\cdot\frac{2}{3}\cdot\frac{3000}{1}+\frac{\$6.50}{1}\cdot\frac{1}{3}\cdot\frac{3000}{1}}{3000}=\frac{(3000)\big(-3)(2/3)+(6.50)(1/3)\big)}{3000}=(-\$3)(2/3)+(\$6.50)(1/3)$$

So we make the following more general definition:

Definition 11. Suppose that a person is playing a game with sample space $\{O_1, O_2, \ldots, O_n\}$. Suppose that if outcome O_i occurs, the player wins $W(O_i)$ dollars. (This number might be negative, which corresponds to a loss for that player.) The **expected value** (or **expectation**) of the game for that player is:

$$E = W(O_1)P(O_1) + \ldots + W(O_n)P(O_n).$$

The expected value is the average amount that the person will win or lose if the game is played many times.

Definition 12. A game is **fair** if the expected value for each person is 0.

Example 16. I propose a betting game to you. We will roll 2 (fair) dice and add the numbers on the top. If the result is an even number other than 2, I give you \$100. If the result is an odd number, you give me \$95. What is your expected value?

Solution: First calculate the sample space the possible ways of getting each number between 2 and 12 by rolling two die and adding the result:

die	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Looking at frequencies we get:

$$\begin{aligned}
\text{P}(2) &= 1/36\\
\text{P}(3) &= 2/36\\
\text{P}(4) &= 3/36\\
\text{P}(5) &= 4/36\\
\text{P}(6) &= 5/36\\
\text{P}(7) &= 6/36\\
\text{P}(8) &= 5/36\\
\text{P}(9) &= 4/36\\
\text{P}(10) &= 3/36\\
\text{P}(11) &= 2/36\\
\text{P}(12) &= 1/36
\end{aligned}$$

Let E be the even "obtain an even number other than 2" and let O be the even "obtain an odd number". Then

$$\begin{aligned}
\text{P(E)} &= 17/36\\
\text{P(F)} &= 18/36
\end{aligned}$$

Thus, your expected value is:

$$100(17/36) - 95(18/36) = 47.2 - 47.5 = \$.3$$

4. Conditional Probability

Definition 13. Suppose that E and F are events in a sample space. The probability of E **given** F is:

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

This is also called the probability of E conditioned on F. It is the right way to take additional knowledge into account.

Example 17. Suppose that a fair die is rolled. What is the probability that a 3 came up, given that an odd number came up?

Solution: Let $E = \{3\}$. Let $F = \{1,3,5\}$. Then $E \cap F = \{3\}$. We have $P(E \cap F) = 1/6$ and $P(F) = 1/2$ so

$$P(E|F) = (1/6)/(1/2) = 1/3.$$

Example 18. Suppose that an unfair coin is tossed 8 times. On any given toss, the probability of getting an H is 3/4 and the probability of getting a T is 1/4. What is the probability that, out of 8 tosses, exactly 3 of them were H, given that the first 3 tosses were tails?

Solution Let F be the event that the first 3 tosses are tails. We have $P(F) = (1/4)^3 = 1/64$. Let E be the event that there are exactly 3 H in 8 tosses. In an outcome in $E \cap F$, there are 3 H and 5 T. Thus the probability of each outcome in $E \cap F$ is $(3/4)^3(1/4)^5 \approx .000412$. There are 5 choose 3 = 10 outcomes in $E \cap F$. Thus $P(E \cap F) = .00412$ Consequently,

$$P(E|F) = .00412/(1/64) = .26368.$$

Example 19. Suppose that an unfair coin is tossed 8 times. On any given toss, the probability of getting an H is 3/4 and the probability of getting a T is 1/4. What is the probability that the first 3 tosses are T, if exactly 3 of the tosses are H?

Solution: Use the same notation as in the previous example. In this case, the probability of $E \cap F$ is still .00412, but now we need to divide by $P(E)$. The probability of getting exactly 3 H in 8 tosses is: (8 choose 3) $*(3/4)^3(1/4)^5$. This is equal to .023. The probability $P(F|E) = .179$.

Example 20. Suppose that you are playing a game show with Monty Hall. There are 3 curtains. Behind two of them are goats and behind the third is a GTO. Monty Hall randomly picks a curtain and opens it, revealing a goat. You randomly pick one of the remaining curtains. What is your probability of winning the car?

Proof. To begin, there are three possibilities for the placement of goats and car; each is equally likely. The probability that Monty Hall picked a goat is 2/3. The probability that you picked the car and that Monty Hall picked a goat is 1/3. Thus, the probability that you picked a car, given that Monty Hall picked a goat is $(1/3)/(2/3) = 1/2$. $\square$

Example 21. Suppose that you are playing a game show with Monty Hall. There are 3 curtains. Behind two of them are goats and behind the third is a GTO. Assume that you want the GTO and not a goat. After you randomly pick a door, Monty Hall opens one of the other two doors revealing a goat. You are then allowed to again choose a door and you are given whatever is behind the door. Is it advantageous to switch?

Solution: Again, there are 3 equally likely possibilities for how the cars and goats are positioned. The probability that you chose correctly the first time is 1/3. The probability that you chose incorrectly the first time is 2/3. Monty's selection of curtain is not random. Thus, the probability that the other door contains the car is 2/3.