Chapter 4 Notes and elaborations for STAT 141-Introduction to Statistics

Assignment

Chapter 4 is fairly well written, for the most part. You will want to read section 4.1, 4.2 and 4.3 very carefully. Note that "subjective probability" as mentioned in section 4.1 isn't mathematical probability at all and as such I won't be testing it.

Do the following exercises:

4.1: 1-10 all, 13, 15, 17, 21, 23, 27, 31, 33.
4.2: 1, 2, 3, 5, 9, 15, 17, 23.
4.3: 1, 5, 11, 15, 21, 25, 29, 31, 33, 35.

This chapter introduces basic finite discrete probability theory. Most of the following material is covered in the book, but please read through this as well. As usual, let's start with some definitions.

4.0 Definitions and some basic set theory

In this section, we give some basic definitions and set theory necessary for the discussion of probability.

Definition 4.0.0 The **outcome** of an experiment is the result of the experiment.

When you perform an experiment, you get a result. This result is an outcome. For example, if your experiment is to roll a standard die, you will roll one of 1, 2, 3, 4, 5, or 6. Whatever you roll is the outcome of the experiment.

Definition 4.0.1 The **sample space** is the set of all possible outcomes of an experiment and is denoted as S. The sample space is written as a set.

Examples 4.0.0

a. Your experiment is rolling a die. Then S = {1,2,3,4,5,6}.

b. Your experiment is planting ten seeds to count the number that germinate Then S={0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10}, where the numbers in the sample space represent the number of seeds that germinated.

c. Your experiment is to toss a coin three times: S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}.

By the way, there is no order in sets. For example, {1, 4, 5} = {4, 5, 1} = {5, 1, 4}. However, the notion for the outcomes DO have order here. "HTH" means heads, then tails, and finally a heads were flipped, and in that order. It is important to be able to list the (basic) outcomes clearly, for if we know all possible outcomes, determining probabilities of specific events is easy.

Definition 4.0.2 An **event** is a subset of the sample space.

Event are also written as sets. Before we give some examples of events, we need the concept of the empty set.

Definition 4.0.3 The **empty set** is a set that contains nothing and is denoted as {} or Ø (it depends on the author's preference). It might seem strange, but we consider this as a possible event. Note that S is can also be consider an event. They're useful—like the number zero.

Example 4.0.1

The experiment is rolling a die. Let A be the event that the roll of a die is even. Let B be the event that the roll is greater than 3. Let C be the event "roll greater than 7". Let D be the event "roll a 1, or a 6, or a number between 1 and 6". Find A, B, C, and D.

$$A=\{2,4,6\},\ B = \{4,5,6\},\ C = \emptyset,\ D = S$$

Also note that an event must be well-defined. For example, the event "roll a nice number" isn't well-defined because nice means different things for different people.

Definition 4.0.4 The **complement** of an event E is the event where E doesn't happen. We denote the complement of an event with a bar over the name of the event (not to be confused with notation for the mean of a sample!). So $\overline{E}$ is the complement of E.

Example 4.0.2

Suppose E={all red cars in the US}. The complement of E is $\overline{E}$ = {all cars in the US that are not red}. From Example 4.0.1, the complement of A is $\overline{A}$ = {1,3,5}, and the complement of B is $\overline{B}$ = {1,2,3}.

Note that the complement of Ø is S and the complement of S is Ø. Some books use the notation E^C instead of a bar above.

We can build new events from old events using the **and** and **or** operators from set theory.

The event “E and F” means that both E and F happen. It’s what the two events have in common, their intersection.

The event “E or F” means that either E or F or both happen. Mathematically speaking, it’s their union.

Example 4.0.3

Using A and B from above, we see that the event (A and B) = {4, 6}, and (A or B) = {2, 4, 5, 6}.

Suppose that we are drawing a card from a standard deck of playing cards. Let’s name a few events:

A = “Card is a heart.”
B = “Card is red.”
C = “Card is a face card.”
D = “Card is a king.”
E = “Card is black.”

Consider the new events:

• A and D. This would mean the card is both a heart, and a king. There is one card that is both, the king of hearts.

• A and B. Since hearts are red, this would be all the hearts. You might write A and B = A.

• A or B. Either red or hearts. . .is just all the red cards. A or B = B.

• C or E. This is the whole deck. Every card is either red or black. C or E = S.

• A and E. There are no black hearts. At least not in a deck of cards! A and E = $\emptyset$.

• B and C. Red cards that are face cards. There are 6. The jack, queen, king of hearts are three of them. The jack, queen, king of diamonds are the other three.

4.1 Basic Probability

Associated with each basic outcome is a number called a **probability**. We denote probability with a capital P followed by the event in parentheses, e.g., P(E) is the probability of E happening. Here are some facts about probability that you are expected to know:

• The probability of each outcome is a number between 0 and 1. These can be written as percentages. For example, a probability of 0.04 is the same as a 4% chance (recall that we move the decimal place two places to the right to convert to percentages and vice versa). These can also be expressed as fractions, e.g., a probability of 0.25 is the same as a probability of 1/4.

• The probability of the entire sample space is 1, i.e., P(S) =1, and the probability of the empty event is 0, that is, P($\emptyset$) = 0. Note that other events might have probability zero or one as well.

• The sum of the probabilities of all the basic outcomes in the sample space must be 1. For example, if your experiment is rolling a fair die, the probability of rolling any given number is 1/6. There are six numbers in the sample space so the sum of the probabilities of all of them is 1:

• The probability of a basic outcome is the percentage of time the outcome happens if the experiment is performed an infinite number of times. If you're trying to roll a 4, theoretically you should roll a 4 about once in every six rolls. However, it may take 7 or 8 or more rolls to roll a 4. But, if you were to roll the die an infinite number of times (boring way to spend eternity), you would roll a 4, on average, once every six rolls.

• The probability of an event is calculated by adding up the probabilities of all the outcomes comprising the event. For example, suppose your experiment is rolling a die. You probably know intuitively that the chance of rolling an even number is 50% because half the numbers are even. This is given symbolically by:

$$P(\{2,4,6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2.$$

When the experiment is rolling a die, all the outcomes have equal probabilities. These are special types of experiments.

Definition 4.1.1 Experiments where all outcomes have the same probability are called **equally likely experiments**.

Rolling a (fair) die is an equally likely experiment. So is drawing a single card from a deck, randomly choosing a candy bar when there are three of each kind in a box, having a child (you have the same probability of having a boy or a girl), etc.

The probability of an event when the experiment is equally likely is always the number of outcomes in E divided by the number of outcomes in S.

$$P(E) = \frac{\text{Number of outcomes in } E}{\text{Number of outcomes in } S}$$

So, for E={roll an even number}, there are three possible outcomes in E. There are a total of six possible outcomes in S. Thus we have P({roll an even number}) = 1/2, as we found above.

Example 4.1.1

Suppose you have a box with 5 Snickers, 3 Milky Ways, and two 3 Musketeers. (Six Musketeers?) Suppose also that you can not tell the type of candy bar by touch alone. The experiment is as follows: someone blindfolds you and you randomly choose a candy bar from the box.

Let A={draw a snickers}, B={draw a milky way}, and C={draw a 3 musketeers}. Find the following probabilities.

P(A), P(B), P(C), P(A or B), P(A or C), P(B and C), and the P(diabetes from stress eating candy bars over probability problems). :-)

This is an equally likely experiment. Which means we need to know how many outcomes there are in each event. We'll take than number and divide by the total number of outcomes. There are a total of 10 candy bars in the box.

- P(A). There are 5 snickers in the box. So 5 divided by a total of 10 candy bars in the box is 5/10 = 0.5. I frequently use percentages, so call it 50%.

- P(B). There are 3 Milky Ways in the box. So 3/10 = 30%.

- P(C). There are two 3 Musketeers in the box. 20%.

- P(A or B). 5 Snickers or 3 Milky Ways is 8 candy bars. Thus, 8/10 = 80%.

- P(B or C). None of the candy bars are both Milky Ways and 3 Musketeers. So 0/10 = 0%.

Not all experiments are equally likely so we can't always use the equally likely formula! Here is an example of such a case.

Example 4.1.2

Suppose your experiment is planting 5 seeds and checking two weeks after planting to see how many seeds germinated. Our sample space is S = {0, 1, 2, 3, 4, 5}, where the numbers in the sample space represents the number of seeds that germinated. Suppose also that we know the following probabilities:

S = s	P(s)
0	.35
1	.05
2	.05
3	.25
4	.25
5	.05

So the table above is giving the probability that 0 seeds germinate is .35, the probability that exactly 1 seed germinates is .05, the probability that exactly 4 seeds germinate is .25, and so on.

Use the table above to find the following probabilities. Remember that we can find the probability of an event by adding all the probabilities of each basic outcome that is contained in that event.

(a) P({no more than 2 seeds germinated}).

No more than two means either 0, 1, or 2 seeds germinated. This can be expressed as P({0,1,2}) and is calculated by adding the respective probabilities. So,

$$P(\{\text{no more than 2 seeds germinated}\}) = P(\{0,1,2\}) = 0.35 + 0.05 + 0.05 = 0.45.$$

(b) P({at least 4 seeds germinated})

At least four means either 4 or five seeds germinated. This can be expressed as

$$P(\{4,5\}) = 0.25+0.05 = 0.30$$

Definition 4.1.1 Two events are said to be **mutually exclusive** if they cannot both occur together. (Some may call these two events disjoint events. I'll try to stick with mutually exclusive.)

Consider, again, rolling a die. You cannot roll a 2 and a 4 in a single roll; the events are mutually exclusive. If A and B are mutually exclusive, then their intersection is empty—there are no basic outcomes in common. It follows that P(A and B) = 0 if and only if A and B are mutually exclusive. If and only if means that if P(A and B) = 0, then A and B are mutually exclusive, and if A and B are mutually exclusive, then P(A and B) = 0. (This is for finite discrete probability only.)

Also, note that complimentary events are always mutually exclusive. For every event, every basic outcome in the sample space is either in that event or not in that event. That fact leads us to the following useful rule: For any event E, $P(E) + P(\bar{E}) = 1$. Do a little basic algebra and we have: **the probability of a complimentary event.**

Equation 4.1.1 $P(\bar{E}) = 1 - P(E)$ for any event E.

Sometimes, finding the probability of the complement of an event is less computationally tedious than calculating the probability of the event itself.

Example 4.1.3

Consider again the experiment in Example 4.1.2. Find the probability that at least one seed germinated.

"At least one" means that 1, 2, 3, 4, or 5 seeds germinated. So, we add all the respective probabilities. However, it is much simpler to recognize that 'at least one' is the complement of 'none'. So, the probability we seek is given by

$$P(\{1,2,3,4,5\}) = 1-P(\{0\}) = 0.65.$$

Clearly, using the complement here is much easier than adding all the other probabilities.

Now we are ready to find probabilities of combinations of events! We'll start with the addition rule. Note that the addition rule does not mean we are adding events themselves. It means we are finding the probability that the union of the events happens.

The addition rule for mutually exclusive events is

Equation 4.1.2 P(A or B) = P(A) + P(B).

This works only because A and B have no common basic outcomes.

Consider once more the experiment in Example 4.1.2. Let A be the event that exactly 1 seed germinates, and B be the event that exactly 5 seeds germinate. These events are mutually exclusive, so P(A or B) = 0.05 +0.05 = 0.10.

If events are not mutually exclusive, addition rule gets a little messier. Consider that when we take the union of two events whose intersection is not empty, we are essentially counting the common outcomes twice. For example, if we are rolling a die and A={1,2,3} and B={2,3,4}, then A or B = {1,2,3,4}. When calculating probabilities of union of events, we account for this 'double counting' by subtracting the probability of the intersection. So, the general **addition rule** becomes

Equation 4.1.3 P(A or B) = P(A) + P(B) - P(A and B).

Example 4.1.4

Suppose our experiment is to draw one card from a standard deck of 52 cards. Find the probability of drawing a diamond or an even non-face card. (Even non-face card would be a 2, 4, 6, 8, or 10.) I'll just call them even cards henceforth.

I'm going to use the addition rule, but with words in their places. We seek: P(diamond or even card).

P(diamond or an even card) = P(diamond) + P(even card) - P(diamond and an even card)

We know that there are 13 diamonds. So P(diamond) = 13/52.

There are 5 even cards in 4 suits. That's 20 total even cards. P(even card) = 20/52.

There are 5 even cards that are diamonds. P(diamond and an even card) = 5/52.

Put it all together and: P(diamond or an even card) = 13/52 + 20/52 - 5/52 = 28/52.

Example 4.1.5

Instead of rolling a single die, let's shake things up a bit and roll two dice at the same time. Suppose we are considering the sums of the rolls, e.g., if you roll a 2 and a 4, the sum is 6. Our sample space is the shaded area of the table to the right. Each shaded cell has a probability of as each outcome is equally likely. Find the following probabilities:

	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

(a) P(roll a sum of 8 or a sum of 5)

The events "roll a sum of 8" and "roll a sum of 5" are mutually exclusive, so we have

P({8,5}) = P({8}) + P({5}) = 5/36 + 4/36 = 9/36
= 1/4.

(b) P(sum is even or sum is a multiple of 3)

The events "sum is even" and "sum is a multiple of 3" are not mutually exclusive, so we have

P({even, multiple of 3}) = P({even}) + P({multiple of 3}) - P({even and multiple of 3})

(You may need to really count carefully here!)

= 18/36 + 12/36 - 6/36
= 2/3.

Now we move on to somewhat trickier stuff.

4.2 Conditional Probability

What's the probability of rolling a 3 on a fair die? Well, that's 1/6. What if I roll the die and hide it from you and tell you "the roll is odd." Now what are the chances of rolling a 3? Well, I just reduced the sample space to the numbers 1, 3 and 5. So the probability of rolling a 3 knowing that the roll is odd is 1/3.

What's the probability of rolling smaller than or equal to 3 on a fair die? Well, that's 3/6 = 1/2. What if I roll the die and hide it from you and tell you "the roll is odd." Now what are the chances of

Conditional probability is denoted P(A|B) and it is the probability of A given that B has already happened or will happen. Sometimes common sense or intuition will easily provide the answer to these types of problems. The trick to seeing conditional probability is that, if we know that we're conditioning on B, then we simply reduce the sample space to only the outcomes in B (because B has already happened, or will happen before A).

Equation 4.2.0 $$P(A|B)=\frac{P(A \text{ and } B)}{P(B)}$$ This is the definition of condition probability.

i.e., the probability of A given B is the probability of both A and B happening divided by the probability of B. As stated previously, many times intuition will help us find conditional probabilities.

Do note, that you can't condition on an event with zero probability.

Example 4.2.0

Consider the experiment of Example 4.1.5. Let A be the event that we roll an even sum, and let B be the event that we roll a sum of 6. Find the following conditional probabilities:

(a) P(A|B)

This is the probability of A given that B happened. Our intuition will work just fine here since we're considering the probability of rolling an even sum given that we've already rolled a 6. We know this probability is 1 because 6 is an even number. However, let's employ the formula for kicks! So we need to find P(A and B) and P(B). We have

$$P(A \text{ and } B) = P(\text{sum is even and sum is 6})$$

Now, if the roll adds to 6 then we know the sum is even. So:

$$P(\text{sum is even and sum is 6}) = P(\text{sum is 6}) = P(B).$$

Thus, we have $P(A|B) = P(A \text{ and } B) / P(B) = P(B) / P(B) = 1$.

(b) $P(B|A)$

This is different. We're looking for the probability of B given that A has happened. So, we need to find the probability we rolled a sum of 6 given that we've rolled an even sum. This is not intuitively clear. So, we must use the formula. Notice that we are conditioning on A, so P(A) will be our denominator. (You might want to look at the table from example 4.1.5 to find these probabilities. Note that 18 outcomes are even numbers, and of those, 5 are the number 6.)

$$P(A \text{ and } B) = P(\text{sum is even and sum is } 6) = P(\text{sum is } 6) = P(B) = 5/36.$$

We also have that

$$P(A) = 18/36.$$

So, employing the formula, we get

$$P(B|A) = P(A \text{ and } B) / P(A) = (5/36) / (18/36) = 5/18.$$

We already have the addition rule in our toolbox. Now we need the multiplication rule. Note that the multiplication rule does not mean we are multiplying events themselves. It means we are finding the probability that the intersection of the events happens.

The **multiplication rule** states that, for any two events A and B, we have

Equation 4.2.1 $P(A \text{ and } B) = P(A)\, P(B \mid A)$ or equivalently, $P(A \text{ and } B) = P(B)\, P(A \mid B)$.

Note that this formula is derived from the definition of condition probability. This will be needed in the next examples.

Example 4.2.1

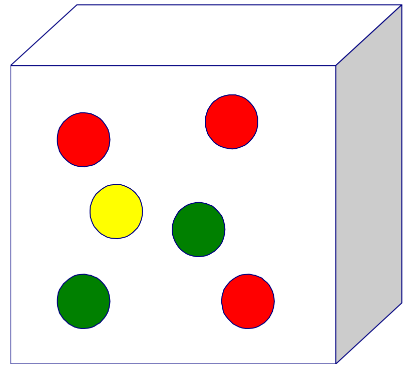
Suppose 9% of working-age people with disabilities are unemployed. Assume a 5% unemployment rate among working-age people. 11% of working-age people have disabilities. What are the chances that a randomly selected working-age person is both unemployed and has a disability?

Let D be "working-age person with a disability" and E equal to "working-age person is unemployed".

From the problem we have $P(E \mid D) = 9\%$, $P(E) = 5\%$, and $P(D) = 11\%$. We're looking for P(E and D). From the multiplication rule,

$$P(E \text{ and } D) = P(D)\, P(E \mid D) = 11\% * 9\% = 0.99\%$$

Example 4.2.2



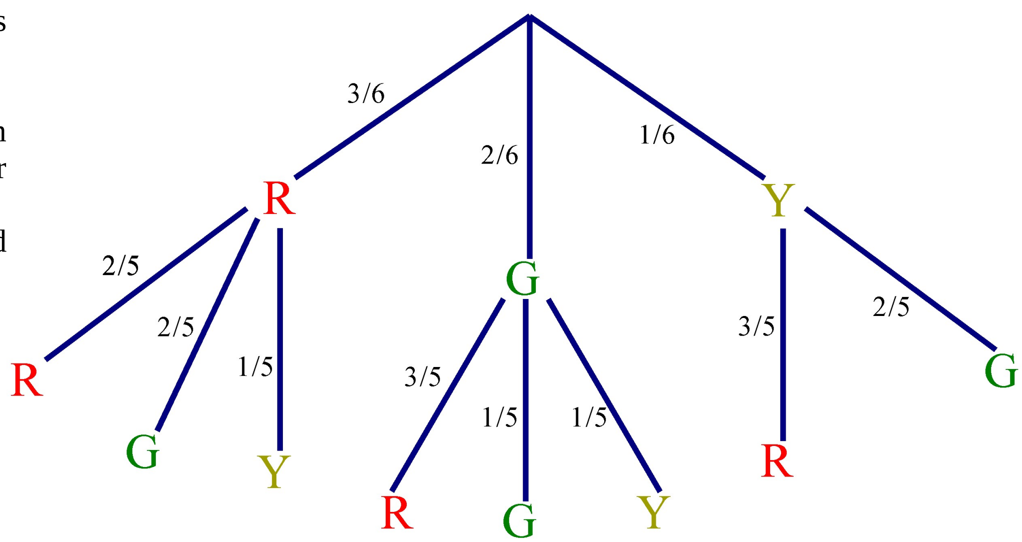
Suppose we have six balls in a box. There are 3 red balls, 2 green balls, and 1 yellow ball. Our experiment is to draw 2 balls without replacement. Without replacement means we do not put the ball back in the box after it is drawn. Find the following probabilities:

(a) P(drawing a green 2nd given that a red was drawn 1st)
(b) P(drawing a red 1st given that a green is drawn 2nd)

Let's use a tree diagram.

Tree diagrams are a tool used to help calculate various probabilities for a given experiment. The top set of three 'branches' indicates all possible outcomes of the first draw and the second set of branches indicates the second draw, which of course depends on the first draw. Notice that there are only two branches coming off the top yellow branch. This is because, if we draw a yellow on the first draw, there are none left in the box for the second draw. The fraction next to each branch is the probability associated with that particular outcome. For instance, the probability of drawing a red after drawing a yellow is 3/5. That is P(draw R 2^{nd} | draw a Y 1^{st}) = 3/5.

From this tree diagram, we can write our sample space and the



probabilities of all the outcomes in the sample space. Although the sample space is fairly easy to ascertain without the diagram, it's even less work for our brains with the diagram in front of us. Just read down each set of branches and we get:

S={RR, RG, RY, GR, GG, GY, YR, YG}.

Here RR stands for draw red 1st and draw red 2nd, RG stands for draw red 1st and draw green 2nd, etc.

To find the probabilities of each of these possible outcomes, we use the multiplication rule. That is, we just multiply down the branches. So, for example:

P(GG) = 2/6 * 1/5 = 2/30.
P(YG) = 1/6 * 2/5 = 2/30.
P(RG) = 3/6 * 2/5 = 6/30. (We'll need these later.)

Let's try to answer the problems now.

(a) P(drawing a green 2nd given that a red was drawn 1^{st})

We use intuition here. Since a red was already drawn (put your finger over a red ball to block it out), we know that there remains 2 red balls, 2 green balls, and 1 yellow ball in the box (a total of 5 balls). So,

P(drawing a green 2nd | red drawn 1^{st}) = 2/5

That's exactly why there is a 2/5 on the branch. That's where all the lower numbers on the diagram come from.

(b) P(drawing a red 1st given that a green is drawn 2nd)

This problem is a bit messier. Employing the definition of conditional probability gives us:

$$P(\text{red 1st} \mid \text{green 2nd}) = \frac{P(\text{red 1st and green 2nd})}{P(\text{green 2nd})} = \frac{P(RG)}{P(\text{green 2nd})}$$

We got the numerator above, P(RG) = 6/30 = 1/5.

To find P(G 2^{nd}), we look at the sample space. We're going to add all the outcomes where there is a green drawn on the 2^{nd} draw.

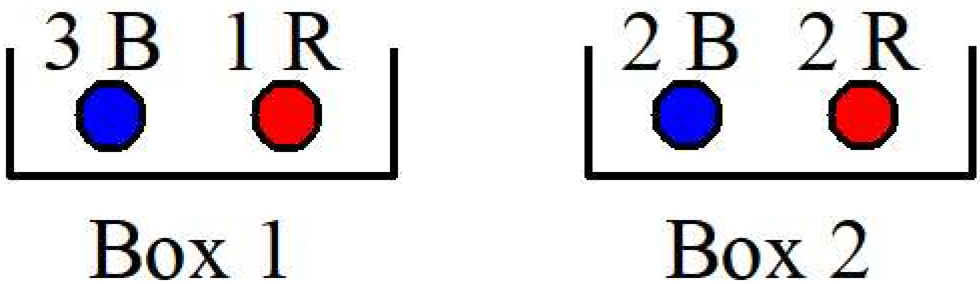
That would be P(GG) + P(YG) + P(RG) = 2/30 + 2/30 + 6/30 = 10/30 = 1/3.

Thus,

$$P(\text{red 1st} \mid \text{green 2nd}) = \frac{P(RG)}{P(\text{green 2nd})} = \frac{1/5}{1/3} = 3/5.$$

Example 4.3.0

Box 1 contains 4 marbles: 3 Blue, and 1 Red. Box 2 also contains 4 marbles: 2 Red and 2 Blue.



The experiment is multistage. There are a total of 3 draws. First, draw twice from box 1 without replacement. Now place the two marbles drawn from box 1 and put them into box 2. The third and final draw is made from box 2.

Notation: I'll write "**BBR**" to mean Blue on 1st draw, Blue on 2nd draw and Red on 3rd draw.

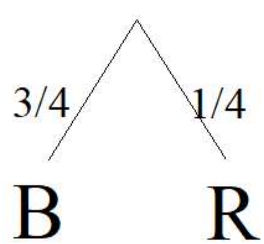
Find the following probabilities:

a. P(**BBR**)
b. P(**RRB**)
c. P(exactly 2 reds out of the 3 draws)
d. P(**R** from box 2 | two blue were drawn from box 1)
e. P(exactly one **R** from box 1 | **R** was drawn from box 2)

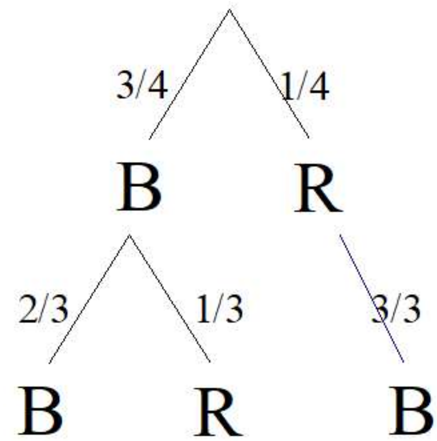
Ponder what these mean for a moment. Try to figure them out on your own if you'd like.

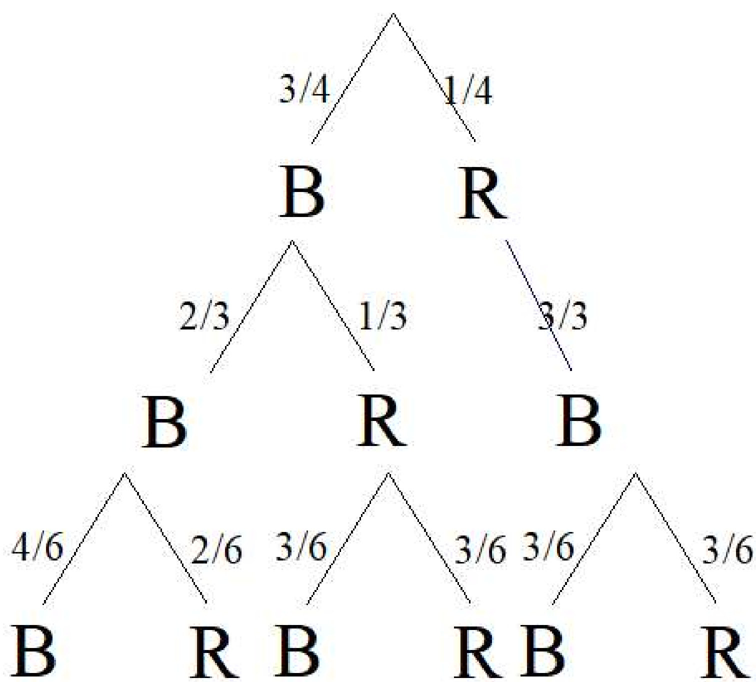
The easiest way to do this problem is to draw a tree diagram to get the sample space and probabilities of each outcome for the experiment.

Step 1. Here's a tree for the first draw from box 1, complete with the probabilities for each outcome.



Step 2. The second draw from box 1. Notice that all the new probabilities are conditional, based on the first draw.





Step 3: Now we add the draw from box 2 to the tree. Remember that we add the marbles from the first two draws to Box 2 before making the third and final draw.

Step 4: We can now see all the outcomes for this experiment and their probabilities. The outcomes are found by taking all the possible routes from the top to the bottom, and the probabilities are found by multiplying each branch's probability. From left to right we have:

Outcome	**Probability of Outcome**
BBB	3/4 * 2/3 * 4/6 = 1/3 = 8/24
BBR	3/4 * 2/3 * 2/6 = 1/6 = 4/24
BRB	3/4 * 1/3 * 3/6 = 1/8 = 3/24
BRR	3/4 * 1/3 * 3/6 = 1/8 = 3/24
RBB	1/4 * 3/3 * 3/6 = 1/8 = 3/24
RBR	1/4 * 3/3 * 3/6 = 1/8 = 3/24

(I have the probabilities reduced and then also all with a common denominator. This will ease some of the computations I'm going to do next, but it isn't technically necessary.)

You should verify the above probabilities. Now I attempt the exercises.

a. P(**BBR**)

BBR is a basic outcome, so we can read it right off the chart: P(BBR) = 1/6.

b. P(**RRB**)

This can't happen, so P(RRB) = 0.

c. P(exactly 2 reds out of the 3 draws)

The probability of an event is the sum of all the basic outcomes contained in that event. I check each outcome to see if it satisfies the event. There are two outcomes that satisfy the event, BRR and RBR. Adding the probabilities of these two outcomes gives us

P(exactly 2 reds out of the 3 draws) = 6/24 = 1/4.

d. P(**R** from box 2 | two blue were drawn from box 1)

If you know how to read this question correctly, it isn't too bad. In English it is asking "what is the probability of drawing a red from box 2 given that two blue marbles were drawn from box 1." We computed this to construct our tree in step 3. Think of it like this: we have already drawn or will draw two blue from Box 1 with probability 1. So we look at the tree, and we see P(R from box 2 | two blue were drawn from box 1) = 2/6 = 1/3.

e. P(exactly one **R** from box 1 | **R** was drawn from box 2)

This is an example of conditional probability where you almost need to rely on the formula. (Read it in English, it feels backwards.) We are seeking the probability that we got a red marble from Box 1 given that we got a red marble from Box 2. This is not intuitive at all (not for me anyway!). We are almost forced to use Equation 4.2.0.

Let A={one R from box 1} and B = {R from box 2}. Using Equation 4.2.0, we have

P(one R from box 1 | R from box 2) = P(one R from box 1 and R from box 2) / P(R from box 2).

The outcomes that satisfy 'one R from box 1 and R from box 2' are highlighted:

BBB	1/3 = 8/24
BBR	1/6 = 4/24
BRB	1/8 = 3/24
BRR	**1/8 = 3/24**
RBB	1/8 = 3/24
RBR	**1/8 = 3/24**

So the P(A and B) = 3/24 + 3/24 = 6/24.

Now I find P(B) in the same way:

BBB	1/3 = 8/24
BBR	**1/6 = 4/24**
BRB	1/8 = 3/24
BRR	**1/8 = 3/24**
RBB	1/8 = 3/24
RBR	**1/8 = 3/24**

Thus, P(B) = 4/24 + 3/24 + 3/24 = 10/24.

The answer is then P(A|B) = P(A and B)/P(B) = 6/24 ÷ 10/24 = 6/24 * 24/10 = 6/10 = 3/5.

You should note that tree diagrams are not always useful. They get very complicated very quickly, and so you should choose carefully when to use them. For example, try a box with 10 balls with 4 different colors, and make 5 draws. Draw the tree diagram for this. You should get a mess.

Finally, we are ready for one of the most powerful concepts in probability theory: independence of events.

4.4 Independence of Events

Definition 4.4.0 Two events, A and B, are independent events if A occurring does not affect the probability that B will occur, or vice versa.

Definition 4.4.1 A more formal definition is two events, A and B, are independent if and only if

$$P(A \text{ and } B) = P(A)\,P(B).$$

When events are independent, P(B|A) = P(B) since A occurring does not change the probability that B occurs. So, Equation 4.2.1 above becomes P(A and B) = P(A) P(B). Note that this formula generalizes to more than two events, i.e., if three events, A, B, and C, are pairwise independent, then P(A and B and C) = P(A)P(B)P(C).

I would like to take a moment to caution you about a common mistake students tend to make at this juncture. Do not confuse independent events with mutually exclusive events. They are very different concepts:

mutually exclusive means P(A and B) = 0   whereas   independent means P(A and B) =P(A)P(B).

Example 4.4.0

Suppose your experiment is to draw 3 cards from a standard 52-card deck with replacement, shuffling between draws. Note that the words "with replacement" indicate each draw is an independent event (since we're putting the card back each time, the probability of the next draw is unaffected by the one before). Find the probability of getting at least one diamond. Note also that a tree diagram would be very messy for this particular problem since there are three draws and four suits (that's 43 branches!).

Drawing a card is an equally likely experiment (each card has the same chance of being drawn). Since there are 13 diamonds in a deck, the probability of drawing a diamond is 13/52 = 1/4. From this, we can deduce the probability of drawing anything other than a diamond is 3/4. For each draw, either we get a diamond or we don't and since each draw is independent of all the others, we can simply multiply the probabilities.

The compliment of the event "get at least one diamond" is "get exactly no diamonds diamonds". We'll find this first.

$$\begin{aligned} P(\text{No D } 1^{st} \text{ and no D } 2^{nd} \text{ and no D } 3^{rd}) &= P(\text{No D } 1^{st})\, P(\text{no D } 2^{nd})\, P(\text{no D } 3^{rd}) \\ &= (3/4)(3/4)(3/4) \\ &= 27/64 \end{aligned}$$

So, the P(get at least one diamond) = 1 - P(get exactly no diamonds diamonds) = 1 – 27/64 = 37/64.