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: 1-5 all, 7, 8, 10-14, 17, 21, 23, 29, 31.
4.2: 1, 2, 3, 5, 7, 9, 15, 19.
4.3: 1, 5, 7, 21, 23, 25, 29, 31, 33, 35, 47.

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.

---

Example 4.0.0

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

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

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, e.g., $\{1, 4, 5\} = \{4, 5, 1\} = \{5, 1, 4\}$. 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 $\emptyset$ or $\varnothing$ (it depends on the author’s preference).
Note that $S$ is a subset of itself, and $\emptyset$ is a subset of $S$, so an event can be all of $S$ or it can be empty.

**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 $D$.

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

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!).

**Example 4.0.2**

Suppose $E = \{\text{all red cars in the US}\}$. The complement of $E$ is $E^c = \{\text{all cars in the US that are not red}\}$.

From Example 4.0.1, the complement of $A$ is $A^c = \{1, 3, 5\}$, and the complement of $B$ is $B^c = \{1, 2, 3\}$.

Note that the complement of $\emptyset$ is $S$ and the complement of $S$ is $\emptyset$. 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, i.e., their intersection. Common notation for the **and** operator is $\cap$, the intersect symbol. We read $E \cap F$ as “$E$ intersect $F$”. Notice that if $S$ is our sample space and $E$ is an event, then we have $E \cap E = \emptyset$ and $E \cap S = E$.

The event “$E$ or $F$” means that either $E$ or $F$ or both happen. It’s the combination of the sample spaces for both events, i.e., their union. Common notation for the **or** operator is $\cup$, the union symbol. We read $E \cup F$ as “$E$ union $F$”. Notice that if $S$ is our sample space and $E$ is an event, then $E \cup E = S$. 
Example 4.0.3

Using A and B from above, we see that \((A \land B) = \{4, 6\}\), and \((A \lor B) = \{2, 4, 5, 6\}\). These can be nicely represented using Venn Diagrams if you’re a visual person. Below is a Venn Diagram for the events A and B. Notice that the whole box is S, our sample space (this is also called our universe, sometimes). One circle contains A while the other contains B, and their intersection is where the two circles overlap. Try to determine the following from the Venn Diagram (answers are below the diagram):

(a) \(\overline{A} \lor B\)   
(b) \(A \land \overline{B}\)   
(c) \(\overline{A} \lor \overline{B}\)   
(d) \(\overline{A} \land B\)

\[
\begin{array}{c}
\text{Venn Diagram} \\
\hline
S & A & B \\
2 & 6 & 4 \\
\hline
3 & 1 \\
\end{array}
\]

(a) \(\{1,3,4,5,6\}\)   
(b) \(\{2\}\)   
(c) \(\{1,2,3,5\}\)   
(d) \(\{5\}\)

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, e.g., 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 \(\frac{1}{4}\).

- The probability of the entire sample space is 1, i.e., \(P(S)=1\), and the probability of the empty event is 0, i.e., \(P(\{\})=0\). 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 die, the probability of rolling any given number is \(\frac{1}{6}\). There are six numbers in the sample space, i.e., there are six possible outcomes, so the sum of the probabilities of all of them is 1:

\[
\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6} = 1.
\]

- The probability 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

\[ P(\{2,4,6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}. \]

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. Equally likely experiments are very nice experiments.

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 \( S \), i.e.,

**Equation 4.1.0**

\[ \frac{\# \text{outcomes in } E}{\# \text{outcomes in } S}. \]

So, for \( E=\{\text{roll an even number}\} \), there are three possible outcomes in \( E \). There are a total of six possible outcomes in \( S \). Thus we have \( P(\{\text{roll an even number}\}) = \frac{3}{6} \), as we found above.

---

**Example 4.1.1**

Suppose you have a box with three snickers, three milky ways, and two 3 musketeers. Suppose also that you cannot 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 \( A = \{ \text{draw a snickers} \} \), \( B = \{ \text{draw a milky way} \} \), and \( C = \{ \text{draw a 3 musketeers} \} \). Find the following probabilities.

(a) \( P(A) \), \( P(B) \), \( P(C) \) and  
(b) \( P(A \text{ or } B) \), \( P(A \text{ or } C) \), \( P(B \text{ or } C) \)

This is an equally likely experiment! According to what’s above, we can find the probability of event \( A \) by adding up all the probabilities comprising the event. Well, there are 3 snickers in the box and each has a \( \frac{1}{9} \) chance of being drawn, so we get

\[
P(A) = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{3}{9} = \frac{1}{3}.
\]

Note that there are 3 outcomes in event \( A \) and there are 9 outcomes in the sample space. Since this experiment is an equally likely experiment, we could just use Equation 4.1.0 above to get

\[
P(A) = \frac{\# \text{outcomes in } E}{\# \text{outcomes in } S} = \frac{3}{9} = \frac{1}{3}.
\]

It should be clear that \( P(A) = P(B) = P(C) \).

To find \( P(A \text{ or } B) \), we do the same thing. We add up the probabilities of all outcomes that make up the event “\( A \) or \( B \).” This event consists of drawing either a snickers or a milky way. Well, there are 6 outcomes in this event, and there are 9 outcomes in the sample space. So we get

\[
P(A \text{ or } B) = \frac{\# \text{outcomes in } E}{\# \text{outcomes in } S} = \frac{6}{9} = \frac{2}{3}.
\]

And it should be clear that \( P(A \text{ or } B) = P(A \text{ or } C) = P(B \text{ or } C) \).

Not all experiments are equally likely so we can’t always use this 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:

<table>
<thead>
<tr>
<th>( s )</th>
<th>( P(s) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.35</td>
</tr>
<tr>
<td>1</td>
<td>.05</td>
</tr>
<tr>
<td>2</td>
<td>.05</td>
</tr>
<tr>
<td>3</td>
<td>.25</td>
</tr>
<tr>
<td>4</td>
<td>.25</td>
</tr>
<tr>
<td>5</td>
<td>.05</td>
</tr>
</tbody>
</table>

So, 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, etc.
Use the table above to find the following probabilities:

(a) $P(\{\text{no more than 2 seeds germinated}\})$

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,

$$P(\{\text{no more than 2 seeds germinated}\}) = P(\{0,1,2\}) = .35 + .05 + .05 = .45.$$

(b) $P(\{\text{at least 4 seeds germinated}\})$

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

$$P(\{4,5\}) = .25 + .05 = .30$$

Notice that the complement of $\{4,5\}$ is $\{0,1,2,3\}$ and $P(\{0,1,2,3\}) = 1 - P(\{4,5\}) = .70$ (more on this next).

---

**Definition 4.1.1** Two events are said to be **mutually exclusive** or **disjoint** events if they cannot both occur together.

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 disjoint, then their intersection is empty, i.e., $A \cap B = \emptyset$. It follows that $P(A \text{ and } B) = 0$ if and only if $A$ and $B$ are disjoint. If and only if, denoted $\iff$, means that if $P(A \text{ and } B) = 0$, then $A$ and $B$ are disjoint, and if $A$ and $B$ are disjoint, then $P(A \text{ and } B) = 0$. (This is for finite discrete probability only.)

Sometimes, finding the probability of the complement of an event is less computationally tedious than calculating the probability of the event itself. Note that for any event $A$, we have that $A \cup \bar{A} = S$, where $A$ and $\bar{A}$ are disjoint (an event and its complement are disjoint by definition). So, $P(S) = P(A) + P(\bar{A})$. But we know $P(S) = 1$, so substitute 1 in for $P(S)$, do a little basic algebra, and we have the following formula for the **probability of the complement of an event**:

**Equation 4.1.1**

$$P(\bar{A}) = 1 - P(A)$$

or equivalently, $P(A) = 1 - P(\bar{A})$.

---

**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.

For *mutually exclusive* events (defined above) A and B, the **addition rule** is

**Equation 4.1.2** \[ P(A \, \text{or} \, B) = P(A) + P(B). \]

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 \( P(A \, \text{or} \, B) = .05 + .05 = .10 \).

If events are *not* mutually exclusive, the **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 \( A \cup B = \{1,2,3,2,3,4\} = \{1,2,3,4\} \). When calculating probabilities of union of events, we account for this ‘double counting’ by subtracting the probability of the intersection. So, our formula becomes

**Equation 4.1.3** \[ P(A \, \text{or} \, B) = P(A) + P(B) - P(A \, \text{and} \, B). \]

---

**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 a 5.

Symbolically, this looks like \( P(\{\text{diamond or 5}\}) \). We know that there are 13 diamonds and four 5's in the deck. We also know that one of the 5's is both a diamond and a 5, so there is a non-empty intersection between the events ‘draw a diamond’ and ‘draw a 5’, i.e., these are *not* mutually exclusive events. We must add the probability of the events and subtract out their intersection. So, we have

\[
P(\{\text{diamond or 5}\}) = P(\{\text{diamond }\}) + P(\{5\}) - P(\{\text{diamond and 5}\})
\]

\[
= \frac{13}{52} + \frac{4}{52} - \frac{1}{52}
\]

\[
= \frac{16}{52}
\]

\[
= \frac{4}{13}.
\]

---

**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 following table:
Each shaded cell has a probability of $\frac{1}{36}$ as each outcome is equally likely. Find the following probabilities:

(a) $P(\text{roll a sum of 8 or a sum of 5})$

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

$$P(\{8,5\}) = P(\{8\}) + P(\{5\}) = \frac{5}{36} + \frac{4}{36} = \frac{1}{4}.$$  

(b) $P(\text{sum is even or sum is a multiple of 3})$

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

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

$$= \frac{18}{36} + \frac{12}{36} - \frac{6}{36}$$

$$= \frac{2}{3}.$$  

Now we move on to somewhat trickier stuff.

4.2 Conditional Probability

Conditional probability is denoted $P(A|B)$ and it is the probability of A given that B has already happened or will happen before A. 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). Following are pictures of what we just stated.

Here’s the original picture:
Here’s the picture using the knowledge that B will happen or has already happened:

![Venn Diagram](image)

So, when we say the probability of A given B, we mean we want the shaded area in the following diagram:

![Venn Diagram](image)

Provided that P(B) is not zero, the formula is

**Equation 4.2.0**

\[
P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{P(A \cap B)}{P(B)},
\]

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.

**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 \( P(A \text{ and } B) \) and \( P(B) \). We have

\[
P(A \text{ and } B) = P(\text{sum is even and sum is 6}) = P(\text{sum is 6}) = P(B).
\]

Thus, we have  \( P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{P(B)}{P(B)} = 1. \)
(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. Like before, we have

\[
P(A \text{ and } B) = P(\text{sum is even and sum is 6}) = P(\text{sum is 6}) = P(B) = \frac{5}{36}.
\]

We also have that \( P(A) = \frac{18}{36} \). So, employing the formula, we get

\[
P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{\frac{5}{36}}{\frac{18}{36}} = \frac{5}{18}.
\]

Checking with the table from the example, we see that 18 outcomes are even numbers, and of those, 5 are the number 6. So, given that we’ve rolled an even number, the probability of it being a 6 is 5/18.

Example 4.2.1

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

Intuition helps a lot with this particular problem. 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, the probability of drawing a green ball on the second draw is simply \( \frac{2}{5} \).

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

This problem is a bit messier. Employing the formula gives us

\[
P(\text{red 1}\text{st} \text{ given green 2}\text{nd}) = \frac{P(\text{red 1}\text{st} \text{ and green 2}\text{nd})}{P(\text{green 2}\text{nd})}.
\]

But, what is \( P(\text{red 1}\text{st} \text{ and green 2}\text{nd}) \)? And what is \( P(\text{green 2}\text{nd}) \)? To determine these probabilities, we will need a couple of more tools.
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**\[ P(A \text{ and } B) = P(A) \cdot P(B \mid A) \text{ or equivalently, } P(A \text{ and } B) = P(B) \cdot P(A \mid B). \]

Note that this formula is derived directly from Equation 4.2.0!

---

**Example 4.2.2**

Consider part (b) of Example 4.2.1. Let \( A = \{\text{draw red 1}^{\text{st}}\} \) and let \( B = \{\text{draw green 2}^{\text{nd}}\} \). Then, using the multiplication rule, we have the following information:

\[
P(A) = P(\{\text{draw red 1}^{\text{st}}\}) = \frac{3}{6} = \frac{1}{2}, \text{ and from part (a), } P(B \mid A) = \frac{2}{5}.
\]

Thus, we have

\[
P(A \text{ and } B) = P(A) \cdot P(B \mid A) = \frac{1}{2} \cdot \frac{2}{5} = \frac{1}{5}.
\]

Now all we need to complete part (b) of Example 15 is \( P(B) \). First, let’s look at the sample space. Since we are drawing twice without replacement, our sample space is

\[
S = \{\text{RR, RG, RY, YR, YG, GG, GR, GY}\}
\]

= \{ ○○, ○●, ●○, ○●, ●●, ●●, ●●, ●○, ●●, ●● \}

where RR stands for draw red 1\(^{\text{st}}\) and draw red 2\(^{\text{nd}}\), RG stands for draw red 1\(^{\text{st}}\) and draw green 2\(^{\text{nd}}\), etc.

It is clear that \( B=\{\text{RG, GG, YG}\} \), i.e., \( B \) is the set of all outcomes where a green ball is drawn 2\(^{\text{nd}}\). So, we need to find the probabilities of each of these outcomes and then add them to get \( P(B) \) (these are not mutually exclusive events). We already know (from above) that \( P(\{\text{RG}\})=\frac{1}{5} \). In the same manner, we find that \( P(\{\text{GG}\})=\frac{1}{15} \) and \( P(\{\text{YG}\})=\frac{1}{15} \). So, we have that

\[
P(B) = \frac{1}{5} + \frac{1}{15} + \frac{1}{15} = \frac{5}{15} = \frac{1}{3}.
\]

Finally, we can compute

\[
P(A \mid B) = \frac{1}{5} / \frac{1}{3} = \frac{3}{5} = \frac{3}{5}.
\]
Are you thinking ‘what a pain in the butt!’? I agree. Some probability problems can be exceptionally tedious. So, here’s a tool to help you through these more tedious problems.

### 4.3 Tree Diagrams

Tree diagrams are a tool used to help calculate various probabilities for a given experiment. The following tree diagram is for the experiment of Example 4.2.2. The first set of three ‘branches’ indicates all possible outcomes of the first draw and the second set of branches indicates the second draw. Notice that there are only two branches coming off the first 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.

From this tree diagram, we can write our sample space and the probabilities of all the outcomes in the sample space. So, instead of going through the tedium of the previous page, we could have just drawn this tree and used it accordingly. Let’s try it!

Although the sample space was 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 (as before):

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

Now, here’s the really cool part: we can just multiply down the branches to get the probabilities of all the outcomes in the sample space. We get:

\[
\begin{align*}
P(\{RR\}) &= \frac{3}{6} \cdot \frac{2}{5} = \frac{6}{30} = \frac{1}{5}, & P(\{RG\}) &= \frac{3}{6} \cdot \frac{2}{5} = \frac{6}{30} = \frac{1}{5}, & P(\{RY\}) &= \frac{3}{6} \cdot \frac{1}{5} = \frac{3}{30} = \frac{1}{10}, \\
P(\{GR\}) &= \frac{2}{6} \cdot \frac{3}{5} = \frac{6}{30} = \frac{1}{5}, & P(\{GG\}) &= \frac{2}{6} \cdot \frac{1}{5} = \frac{2}{30} = \frac{1}{15}, & P(\{GY\}) &= \frac{2}{6} \cdot \frac{1}{5} = \frac{2}{30} = \frac{1}{15}, \\
P(\{YR\}) &= \frac{1}{6} \cdot \frac{3}{5} = \frac{3}{30} = \frac{1}{10}, & P(\{YG\}) &= \frac{1}{6} \cdot \frac{2}{5} = \frac{2}{30} = \frac{1}{15}.
\end{align*}
\]

Let’s check this result. Recall that the probabilities of all outcomes in the sample space should sum to 1.

\[ \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{10} + \frac{1}{10} + \frac{1}{15} + \frac{1}{15} + \frac{1}{15} = 1 \]

It checks! If you compare this with the results of Examples 4.2.1 and 4.2.2, hopefully you see that all the numbers we need to answer the posed questions are right before our eyes.
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 1\textsuperscript{st} draw, Blue on 2\textsuperscript{nd} draw and Red on 3\textsuperscript{rd} draw.

Find the following probabilities:

a. \( P(\text{BBR}) \)

b. \( P(\text{RRB}) \)

c. \( P(\text{exactly 2 reds out of the 3 draws}) \)

d. \( P(\text{R from box 2 } | \text{ two blue were drawn from box 1}) \)

e. \( P(\text{exactly one R from box 1 } | \text{ R was drawn from box 2}) \)

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.

1. **Draw the tree diagram for this experiment.**

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

```
                B  R
               /   \
           3/4 1/4

B  R
```

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

```
                B  R
               /   \
           3/4 1/4
            /   \
        2/3 1/3 3/3

B  R  B
```
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:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability of Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBB</td>
<td>1/3 = 8/24</td>
</tr>
<tr>
<td>BBR</td>
<td>1/6 = 4/24</td>
</tr>
<tr>
<td>BRB</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>BRR</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>RBB</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>RBR</td>
<td>1/8 = 3/24</td>
</tr>
</tbody>
</table>

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

b. This can’t happen, so \( P(\text{RRB}) = 0 \).

c. 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 \( P(\text{exactly 2 reds out of the 3 draws}) = 6/24 = 1/4 \).

d. 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 \( P(R \text{ from box 2} | \text{two blue were drawn from box 1}) = 2/6 = 1/3 \).

We could also employ Equation 4.2.0 to solve part (d).

Let \( A = \{\text{draw R from Box 2}\} \) and \( B = \{\text{draw two B from Box 1}\} \).
P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{1/6}{1/3 + 1/6} = \frac{1/6}{3/6} = \frac{1}{3}.

e. 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

\[ P(\text{one } R \text{ from box 1 } | \ R \text{ from box 2}) = \frac{P(\text{one } R \text{ from box 1} \text{ and } R \text{ from box 2})}{P(R \text{ from box 2}).} \]

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

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBB</td>
<td>1/3 = 8/24</td>
</tr>
<tr>
<td>BBR</td>
<td>1/6 = 4/24</td>
</tr>
<tr>
<td>BRB</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>BRR</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>RBB</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>RBR</td>
<td>1/8 = 3/24</td>
</tr>
</tbody>
</table>

So the \( P(A \text{ and } B) = 3/24 + 3/24 = 6/24. \)

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

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBB</td>
<td>1/3 = 8/24</td>
</tr>
<tr>
<td>BBR</td>
<td>1/6 = 4/24</td>
</tr>
<tr>
<td>BRB</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>BRR</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>RBB</td>
<td>1/8 = 3/24</td>
</tr>
<tr>
<td>RBR</td>
<td>1/8 = 3/24</td>
</tr>
</tbody>
</table>

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

The answer is then \( P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{6/24}{10/24} = \frac{6}{24} \times \frac{24}{10} = \frac{6}{10} = \frac{3}{5}. \)

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 that re 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 \( P(A \text{ and } B) = P(A) P(B) \). Note that this formula generalizes to more than two events, i.e., if three events, A, B, and C, are pairwise independent, then \( P(A \text{ and } B \text{ and } C) = P(A)P(B)P(C) \).

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 \( \iff P(A \text{ and } B) = 0 \) whereas independent \( \iff P(A \text{ and } B) = P(A)P(B) \).

**Example 4.4.0**

Suppose your experiment is to draw 3 cards from a standard 52-card deck with replacement. 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). Note also that a tree diagram would be very messy for this particular problem since there are three draws and four suits (that’s \( 4^3 \) branches!). Find the probability of getting at least one diamond.

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 \( \frac{13}{52} = \frac{1}{4} \). From this, we can deduce the probability of drawing anything other than a diamond is \( \frac{39}{52} = \frac{3}{4} \). 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 probability of getting at least one diamond is \( P(\{1,2,3\}) \). So, we can find the probability of each of these and add them. Let’s start with the probability of drawing exactly one diamond. There are three different ways this can happen; we can draw a diamond first, second, or third. So, we find the probability of one of these and multiply by 3 and we get

\[ P(\{1\}) = 3 \times \left( \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} \right) = 3 \times \frac{9}{64} = \frac{27}{64}. \]

In a similar manner, we get

\[ P(\{2\}) = 3 \times \left( \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} \right) = 3 \times \frac{3}{64} = \frac{9}{64}, \text{ and } P(\{3\}) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{64}. \]

So, we have \( P(\{1,2,3\}) = \frac{27}{64} + \frac{9}{64} + \frac{1}{64} = \frac{37}{64} \). Hopefully, you noticed it would be easier to calculate the probability of the complement, \( P(\{0\}) \) and subtract from 1, as follows:

\[ P(\{0\}) = \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{27}{64}, \text{ so } P(\{1,2,3\}) = 1 - \frac{27}{64} = \frac{37}{64}. \]