

Mathematical Statistics

Gregg Waterman
Oregon Institute of Technology

©2016 Gregg Waterman



This work is licensed under the Creative Commons Attribution 4.0 International license. The essence of the license is that

You are free to:

- **Share** copy and redistribute the material in any medium or format
- **Adapt** remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

- **Attribution** You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to the web page below.

To view a full copy of this license, visit <https://creativecommons.org/licenses/by/4.0/legalcode>.

Contents

1	Probability Basics	5
1.1	Experiments, Outcomes and Events	6
1.2	Probability	10
1.3	Counting	13
1.4	The Addition Rule	16
1.5	Conditional Probability and Independence	19
1.6	The Multiplication Rule	22
1.7	Bayes' Theorem	24
1.8	Chapter 1 Exercises	27
H	Solutions to Exercises	131
H.1	Chapter 1 Solutions	131

1 Probability Basics

Performance Criteria:

1. (a) Given a verbal description of an experiment and an event, give sets representing the sample space of the experiment and the event. Be able to do this for experiments involving repeated trials, with selection either with or without replacement.
- (b) Determine whether two events are mutually exclusive or complementary.
- (c) Determine the probability of an event using the classical definition of probability.
- (d) Solve counting problems using combinatorial methods.
- (e) Use the addition rule to determine a probability.
- (f) Use a Venn diagram to model a probability problem.
- (g) Compute a conditional probability.
- (h) Apply the multiplication rule to determine probabilities.
- (i) Determine whether two events are independent.
- (j) Use Bayes' Theorem to determine probabilities.

In this chapter we will introduce the methods of “classical” probability that have been in use for hundreds of years, along with some useful combinatorial principles. These things will provide the underpinnings for our study of more “modern” (1900s) probability theory in the rest of the book.

1.1 Experiments, Outcomes and Events

Performance Criteria:

1. (a) Given a verbal description of an experiment and an event, give sets representing the sample space of the experiment and the event. Be able to do this for experiments involving repeated trials, with selection either with or without replacement.
- (b) Determine whether two events are mutually exclusive or complementary.

The main objective of this course is to understand the concept of probability. The idea of probability is to attach numbers to things that could happen that indicate their likelihoods of happening. A common example is when a weatherman (or woman) says there is a 30% chance of snow they are giving a probability that it will snow. We might be interested in other probabilities, like the probability of getting more than 100 hits on a web page in a 5 minute period, or the probability that a part for something we are making is within some tolerance of a desired value.

In order to study probability, we need some common language that we all use so that we can communicate our ideas clearly and precisely. We will be making use of the following definitions.

- An **experiment** is an act whose result can be summarized by some sort of observation.
- When an experiment is conducted, the results that are observed are called **outcomes** of the experiment.
- The *set* of all possible outcomes of an experiment is called the **sample space** of the experiment, denoted by S . This set will usually play the role of the universal set defined in Appendix C.4.
- Any *subset* of the sample space of an experiment is called an **event**. Note that a subset could contain just one outcome, so in a sense an outcome is also an event; the converse is not necessarily true.

◇ **Example 1.1(a):** An experiment consists of flipping a coin and rolling a die. (A **die** is a single small cube with one to six dots on each face. **Dice** are a pair of die.) The outcomes could be denoted by giving the result on the coin followed by the result on the die: H1, H2, ... The sample space is then $S = \{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6\}$. One possible event would be tails on the coin and an odd number on the die. In set notation, this event is the set $\{T1, T3, T5\}$

◇ **Example 1.1(b):** The number of hits on a website during a 24 hour period is observed (the experiment). *Theoretically*, the sample space is $S = \{0, 1, 2, 3, \dots\}$. The event of at least 500 hundred hits is the set $\{500, 501, 502, \dots\}$.

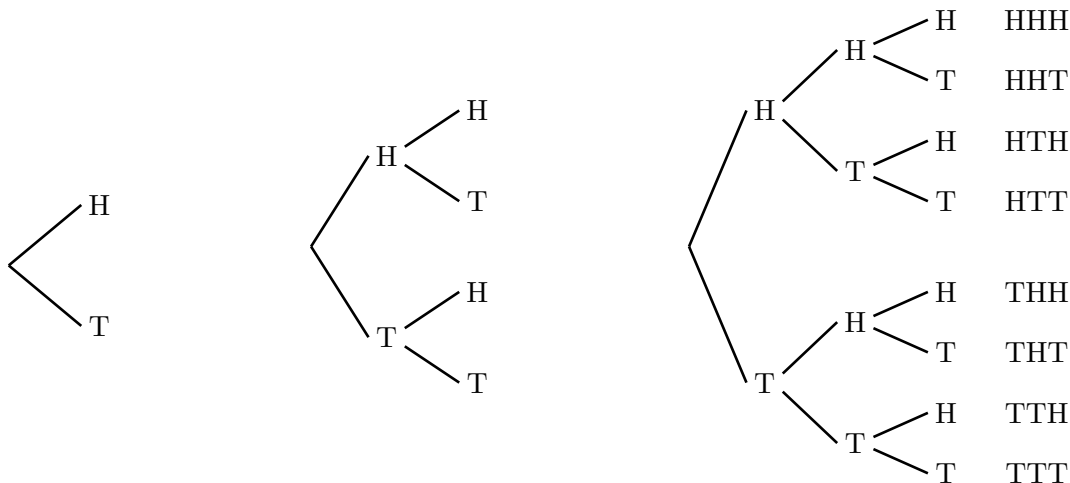
- ◇ **Example 1.1(c):** The lengths of machine bolts are recorded. Assuming that we could measure them with any degree of accuracy we wished, the sample space would be all real numbers greater than zero. Using interval notation, $S = (0, \infty)$. (In reality, the lengths would fall in some shorter interval, but the interval given is certainly the safest one to give.)
-

A sample space containing a finite or countable number of outcomes is called a **discrete sample space**; the sample spaces in Examples 1.1(a) and 1.1(b) above are discrete sample spaces. A sample space that is an interval of the real line, or a region in the plane or in three-dimensional space (or any higher dimensional space) is called a **continuous sample space**. The sample space for Example 1.1(c) is continuous.

1. For each of the following exercises, an experiment is given, followed by an event. Give the sets that represent the sample space of the experiment and the event.
 - (a) *Experiment:* A coin is flipped once. *Event:* Heads is observed.
 - (b) *Experiment:* A coin is flipped three times. *Event:* At least two of the flips result in heads.
 - (c) *Experiment:* A coin is flipped repeatedly until a head is obtained. *Event:* A head is obtained in *less than* five flips.
 - (d) *Experiment:* The three letters A, B and C are arranged in all possible ways, using each letter exactly once. *Event:* The first letter is B.
2. Consider again the experiment of flipping a coin and rolling a die.
 - (a) Give the event of a tail on the coin and a number less than three on the die.
 - (b) Give the event of a tail on the coin or a number less than three on the die.
 - (c) Give the event of four or more on the die. (Since there is no mention of the coin, we assume that it can be either heads or tails.)
3. A pair of dice is rolled. Assume that the two die can be distinguished from one another; perhaps one is red and the other green.
 - (a) How many outcomes will there be for this experiment?
 - (b) Each outcome can be denoted by an ordered pair of numbers. Give the event that the sum of the numbers on the dice is eight *as a set of ordered pairs*. (This event has more than three outcomes!)
4. The actual resistance R of a 10 ohm resistor is measured, to the nearest hundredth of an ohm. Give the event that the resistance is greater than 10.12 ohms, using set notation.
5. The amount of time t (in years, allowing parts of years, like 13.4915 years) from when an electrical component starts to be used until it fails is observed. If we assume that the component is functional to begin with, we can describe the sample set as $\{t \in \mathbb{R} \mid t > 0\}$. Give the set notation for the event that the component lasts at least 10 years but no more than 15 years.

There is a tool called a **tree diagram** that will be very useful when considering experiments consisting of a sequence of steps. Consider the experiment consisting of flipping a coin three times in a row, from Exercise 1(b). At the first step, one of two things can happen: we can get heads, or tails. So our tree starts with two branches, one for getting heads on the first flip, one for tails.

See the leftmost picture at the top of the next page. At the second step we can get heads or tails regardless of whether we got heads or tails on the first flip. So each of the original two branches has two branches coming off of it, as shown in the middle picture on the next page. The tree is then finished by giving the branches for the third flip and writing the individual outcomes at the ends of the branches, as shown.



6. (a) Draw a tree diagram for the experiment from Exercise 1(d). *Remember that each letter is only used once.*
- (b) Draw a tree diagram for the experiment from Exercise 2.
- (c) It should be clear that the tree diagram for the experiment from Exercise 1(c) is infinite. Draw the diagram for the first three “branchings”, putting $\cdots$ at the ends of any branches that continue on.

Consider again the experiment of flipping a coin until a head is obtained. It is very likely that you represented the sample space of that experiment with the set $\{H, TH, TTH, TTTH, \dots\}$. Note that instead we could represent the sample space with the set $\{1, 2, 3, 4, \dots\}$, with each number representing the number of the flip on which the first head is obtained. This set could also be represented by $\{n \in \mathbb{N} \mid n > 0\}$. *Note that this situation is completely analogous to the situation in Exercise 5, except that the sample space from Exercise 1(c) is a countable set, and the sample space from Exercise 5 is an uncountable set.*

When the same process is repeated over and over, like repeatedly flipping a coin, we call each time it is done a **trial**. When a process has two possible outcomes we often refer to one of the outcomes as **success** and the other outcome as **failure**. *We will attach no judgement to these words; we might call something a success when in fact it is a failure in the “real world” sense.* Note that if we were to consider getting a head on the coin to be a failure, then each element of the sample space $\{1, 2, 3, \dots\}$ for Exercise 1(c) represents the number of trials to failure. Similarly, each real number in the set $\{t \in \mathbb{R} \mid t > 0\}$ of possible times to failure of the electrical component represents the amount of time to failure.

Now look at your sample space from Exercise 1(d), where you found the possible arrangements of a set of objects. Such arrangements are called **permutations** of the objects. Note that this concept comes into play when considering situations like the following.

7. The four members of a famous rock band are Trent, Chance, John and Colin. They are trying to decide what order to come out of the dressing room for a concert in front of thousands of adoring fans. How many ways can they do this? If you feel ambitious, use the letters T, Ch, J and Co to represent the band members and give all possible orders in which they can go onstage.

Two events are called **mutually exclusive** if they are disjoint sets. (Remember that this means they have no elements in common.)

8. For each of the following, an experiment and two events are given. Determine whether the events are mutually exclusive; if they are not, give an outcome that is in both events.
 - (a) A coin is flipped three times. The events are (i) a head is obtained on at least two of the tosses and (ii) a tail is obtained on exactly one of the tosses.
 - (b) A coin is flipped three times. The events are (i) a head is obtained on at least two of the tosses and (ii) a tail is obtained on more than one of the tosses.
 - (c) A pair of dice is rolled. The events are (i) the sum of the numbers on the dice is at least ten and (ii) the number on at least one of the die is a two.
9.
 - (a) An experiment consists of selecting two of the four letters A, B, C and D. A letter cannot be used more than once and different orders of the letters are not considered distinct from each other. Give the sample space of this experiment.
 - (b) Consider the same experiment, but with the first letter replaced before selecting the second, so that the same letter can be selected twice. Give the sample space.

It is important that we make a distinction between the two selection processes from the last exercise. When we select items from a group and no item can be selected more than once we say we are selecting **without replacement**. When an item is replaced in the group after it has been selected, we say we are selecting **with replacement**.

10.
 - (a) For the experiment from Exercise 9(a), give the event that at least one of the letters selected is an A.
 - (b) For the experiment from Exercise 9(b), give the event that both letters selected are the same.
 - (c) For the experiment from Exercise 9(b), give the event that the two letters selected *are not* the same.

Note that the two events from Exercises 10(b) and (c) are mutually exclusive, and that their union is the entire sample space. Such events are called **complementary events**. As sets, they are complementary sets.

11. Five men, Adam, Bill, Clint, Derek and Ed, are broken into two groups, one with two men and one with three. Different orders within each group are not considered distinct. One possible outcome is that Adam and Derek are grouped together, and Bill, Clint and Ed are grouped together. We will denote this outcome by (AD, BCE) . Using this notation, give the sample space for this experiment.

NOTE: The act of breaking a set up into smaller sets is called **partitioning** the set. (See Appendix C.8.) Note that

- each of the smaller sets is a subset of the original set,
 - no object is in more than one of the subsets,
 - the *union* of all of the subsets is the original set.
12. Consider the experiment consisting of partitioning the same set of five men from Exercise 11 into two sets of two and a set of one. Give the sample space of this experiment.

1.2 Probability

Performance Criteria:

1. (c) Determine the probability of an event using the classical definition of probability.

As stated previously, the general idea of **probability** is to assign to an event A a number, denoted by $P(A)$, that measures the “likelihood” that the event will happen. (Note the use of function notation here; probability is a function that takes an event and gives out a number.) To do this we will need to devise some scheme for assigning such numbers, called probabilities. Of course if there are two events A and B with A more likely to happen than B , we would probably like to have $P(A) > P(B)$. Therefore we need some range of values for probabilities, and it would be logical to make the smallest possible value be zero, for events that can’t happen. We will take one to be the upper limit of probabilities, the probability of an event that is sure to happen. Since any event of a sample space S is a subset of S and S is the largest subset of itself, we then should have $P(S) = 1$.

Now consider the experiment of flipping a coin twice in a row, with event A denoting the event of obtaining heads on at least one of the flips. Let A_1 be the event of obtaining heads on *exactly one* flip and let A_2 be the event of obtaining heads on *both* flips. Then A_1 and A_2 are mutually exclusive events with $A = A_1 \cup A_2$. Intuitively, we would all probably agree that it should be the case that $P(A) = P(A_1) + P(A_2)$. This should extend to more mutually exclusive sets, so that if $A = A_1 \cup A_2 \cup \dots \cup A_n$, with all of the sets A_i mutually exclusive with all the others, then $P(A) = P(A_1) + P(A_2) + \dots + P(A_n)$. This can even be extended to a *countable* sequence of mutually exclusive sets $A_1, A_2, A_3, \dots$. (Note that we are then dealing with an infinite sum, requiring some knowledge of the theory of infinite series.)

We will then build our theory of probability on a set of basic rules, called **postulates** or **axioms**, which we take to be true without proof. These axioms are a precise summary of the previous discussion.

Axiom 1: For any event A , $0 \leq P(A) \leq 1$.

Axiom 2: For an experiment with sample space S , $P(S) = 1$.

Axiom 3: If $A_1, A_2, A_3, \dots, A_n$ is a finite sequence of mutually exclusive events whose union is A , then

$$P(A) = P(A_1) + P(A_2) + \dots + P(A_n)$$

and if $A_1, A_2, A_3, \dots$ is a countably infinite sequence of mutually exclusive events whose union is A , then

$$P(A) = P(A_1) + P(A_2) + P(A_3) + \dots$$

Mathematicians generally prefer to build a theory from the fewest axioms necessary, and there is a little bit of redundancy in the first two items above. We will not fuss about that, but will just go on our merry way from here! Once the axioms are established we then attempt to build

a structure of consequences that logically follow from the axioms. The first consequence of our axioms can be seen as follows. Each outcome of an experiment is itself an event (well, the singleton set containing that outcome is an event), and those outcomes are obviously mutually exclusive. If we then have a discrete sample space, those outcomes constitute a countable set of events as well, so we can apply Axiom 3 to get the following.

Theorem 1.1: If an experiment has discrete sample space S , then the probability of any event A is the sum of the probabilities of the individual outcomes in A .

Any “rule” that can be deduced logically from axioms or other previously established “rules” is called a **theorem**. (We will number theorems $*.*$, with the ones digit indicating the chapter in which the theorem can be found, and the tenth’s digit indicating which theorem in that chapter it is.) A very useful application of Theorem 1.1 occurs when we have a finite sample space with each outcome having equal probability.

- ◇ **Example 1.2(a):** If a coin is flipped three times in a row, what is the probability of getting heads exactly once?

The sample space for this experiment is

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

and it would seem that all outcomes should be equally likely. Furthermore, each outcome alone can be considered an event as well, and all such events are mutually exclusive. Thus the probability of each outcome must be $\frac{1}{8}$. (Do you see why that is, based on the axioms?) The event $A = \{HTT, THT, TTH\}$ of getting heads exactly once then has probability $\frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$.

Note that in this example $|A| = 3$, $|S| = 8$ and $P(A) = \frac{3}{8}$. Although this does not constitute proof of the following, it does seem to indicate its truth.

Theorem 1.2: Suppose that all outcomes of an experiment with *finite* sample space S are equally likely (have equal probability as events themselves). Then the probability of an event A is

$$P(A) = \frac{|A|}{|S|}.$$

NOTE: For the time being every probability should be given in *exact form*. This means fractions or decimals that *have not been rounded*. Unlike in your past mathematical experience, it is not necessary to reduce fractions; leaving them unreduced will often be more revealing about how they were obtained.

1. A coin is flipped three times. Find the probability that a head is obtained at least twice.
2. A pair of dice is rolled.
 - (a) How many outcomes does the sample space contain?

- (b) Suppose that we want to know the probability that the sum of the numbers on the dice is four. Make a table with the outcomes from one die along the top and the outcomes from the other along the left side. Where a row and column meet, put the sum of the two outcomes for that row and column.
- (c) From your table you can now determine the probability that the sum of the numbers on the dice is four. Do this.
- (d) What is the probability that the sum of the numbers on the dice is four or less?
- (e) *In words*, what is the complement of the event that the sum of the numbers on the dice is four or less? What is the probability of that event?

In answering 2(e) you should have seen (or maybe even used unknowingly) the following.

Theorem 1.3: For any event A with complement A' ,

$$P(A') = 1 - P(A).$$

As a final note, let's re-examine what probability is in a slightly more formal way. First, recall that the set of all subsets of a set U is called the **power set** of U . Probability is a function defined on the power set of the sample space, that assigns to any subset of the sample space a real number between zero and one. If we let $\mathcal{P}(S)$ be the power set of the sample space, we can denote this idea symbolically by

$$P : \mathcal{P}(S) \rightarrow [0, 1]$$

1.3 Counting

Performance Criteria:

1. (d) Solve counting problems using combinatorial methods.

When computing probabilities using Theorem 1.2 it is necessary to know how many elements are in a sample space, and how many elements are in an event. This is often a challenging problem. The area of mathematics that deals with how many ways things can be done is called **combinatorics** or, more simply, **counting**. In this section you will learn the fundamental operations of counting.

1. Suppose that Mr. Waterman has two pairs of pants, a blue pair and a khaki pair. Suppose also that he has three shirts, one blue, one green and one red. Suppose also that he has no qualms about wearing any color with any other, and that he puts on his pants before his shirt in the mornings.
 - (a) How many ways can Mr. W dress?
 - (b) Mr. Waterman's wife tells him that he looks ridiculous when he wears the blue shirt with the blue pants. If Mr. W always selects his shirt and pants randomly, what is the probability that he will dress in a way that his wife does not approve of?

The above exercise illustrates the following.

Theorem 1.4: Suppose that a sequential process consists of selecting one of n_1 choices, then selecting one of n_2 choices. Then there are $n_1 \cdot n_2$ ways to complete the process.

Note that it would not have made any difference whether Mr. W put on his shirt before his pants, there would have been the same number of ways for him to dress. This is, of course, because multiplication is commutative! The same idea can, of course, be extended to situations where there are more choices to be made.

Theorem 1.5: Suppose that a sequential process with k steps consists of selecting one of n_1 choices, one of n_2 choices, ... , and one of n_k choices. Then there are $n_1 \cdot n_2 \cdot n_3 \cdots n_k$ ways to complete the process.

2. A state's license plate numbers consist of two letters, followed by three digits. (A letter or digit may be re-used.) For whatever reason, they do not want the last digit to be zero.
 - (a) How many different license plates can they make with this scheme?
 - (b) How many plates begin with AB and end with the digit 7?
 - (c) What is the probability that a randomly selected plate begins with a vowel and ends with an even number? (Remember that the vowels are a, e, i, o and u .)

3. The **permutations** of a set of distinct objects are all the ways that those objects can be arranged (from a first object to a last object). When thinking about the following, it might be helpful to think of tree diagrams and how many branches there are at each selection step.
- (a) How many permutations of two objects are there?
 - (b) How many permutations of three objects are there?
 - (c) How many permutations of four objects are there?
 - (d) How many permutations of n objects are there?

For $n \geq 1$, the quantity $n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$ is denoted by $n!$, spoken as n **factorial**. The factorial operation is extremely important in combinatorics, and proves useful in other areas of mathematics as well. Because it makes things work out well, we *define* $0!$ to have value one. Because this is a definition, it is not up for debate - that's just the way it is! You will soon see that all the formulas involving factorials make sense with this definition, and would not with any other. Using this notation, the result of the previous exercise is as follows.

Theorem 1.6: The number of permutations of n distinct objects is $n!$.

Often we are interested in situations where we want to know the number of permutations of a few objects selected from a larger group of objects. The following exercise examines such situations.

4. (a) Four contestants, Adam, Bill, Clint and Derek run a race. The top two finishers are recorded, *in order*. How many possible finishes are there?
- (b) A horse race has eight contestants, and only the first three spots are recorded (win, place or show). How many possible top three finishes are possible in a horse race?
- (c) The language we use for the values you found in this exercise is “permutations of four things taken two at a time” and “permutations of eight things taken three at a time”. Can you see how to find the number of permutations of 18 things taken 5 at a time?
- (d) Can you see, in general, how to find the number of permutations of n things taken r at a time? Try to express your answer as a factorial divided by another factorial.

In part (d) above you should have come up with this:

Theorem 1.7: The number of permutations of n distinct objects taken r at a time ($0 \leq r \leq n$) is

$${}_nP_r = n \cdot (n-1) \cdots (n-r+1) = \frac{n!}{(n-r)!}.$$

If $r = n$ then we just have the number of permutations of n objects.

We will now look at similar sorts of situations, but for which *the order in which the individuals chosen doesn't matter*.

5. (a) Two of the four men from Exercise 4(a) are chosen randomly to help carry a heavy object. Clearly it makes no difference who is chosen first and who is chosen second, they both have to carry the load! How many ways can two of the four be chosen?

- (b) Suppose that instead of a horse race, we are selecting three students from a group of eight to give presentations on a given day. It is of no concern who will go first, just who the three are. How many ways can three be chosen? (**Hint:** How many ways could three be chosen *if the order was significant*? Note that this gives the same groups of people several times over. How many times?)

We call the answer to (b) the number of **combinations** of eight things taken three at a time. Another way you might hear this stated is “eight choose three”. The hint should have suggested the following to you:

- The number of permutations of eight objects taken three at a time would tell us how many groups of three could be chosen *if order did matter*.
- Since any group of three individuals can be arranged in $3!$ ways, the same three individuals will show up together in the permutations six times, so we need to divide the number of permutations by $3!$ to get the number of combinations of eight people taken three at a time.

In general, we have the following theorem.

Theorem 1.8: The number of combinations of n distinct objects taken r at a time ($0 \leq r \leq n$) is

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

There are situations in which we wish to compute n choose n or n choose 0 . Since we have defined $0! = 1$, this causes no problems. You will note that r is allowed to be zero in both of the above theorems.

Theorem 1.9: Suppose that there are n objects of k different types; n_1 of the objects are of one type, n_2 of another, and so on. Then there are

$$\frac{n!}{n_1! n_2! \cdots n_k!}$$

distinct permutations of the objects. (Here two arrangements in which two objects of the same type are exchanged are *not* considered distinct.)

6. An urn containing five blue marbles, two yellow marbles and nine red marbles. How many different arrangements of the the marbles are there, if we can't distinguish marbles of the same color from each other?
7. A variation of this that will be important later is this: Suppose that an urn contains red and blue marbles, say at least ten of each. How many ways can we draw three red marbles and five blue marbles, assuming that we can't distinguish between marbles of the same color?

1.4 The Addition Rule

Performance Criteria:

1. (e) Use the addition rule to determine a probability.
- (f) Use a Venn diagram to model a probability problem.

1. Consider the experiment of drawing a card from a deck of cards, with Event F being that a face card (Jack, Queen or King) is drawn and Event E being that an even numbered card is drawn.
 - (a) What is the probability of Event F ?
 - (b) What is the probability of Event E ?
 - (c) What is the probability of the event $F \cup E$?
 - (d) Let A and B be two events for the same experiment. Based on what you have just seen, how do you think $P(A \cup B)$ relates to $P(A)$ and $P(B)$?
2. Again consider the experiment of drawing a random card from a deck of cards. Let Event C be the event that a club is drawn, with Event F again being that a face card is drawn.
 - (a) Are Events C and F mutually exclusive? If not, you know what you must do!
 - (b) Find the probabilities of each of the two events.
 - (c) What is the probability of event $C \cup F$?
 - (d) Go back to your answer to 1(d). Do you think it is correct, in general? If not, amend it so that it is. Does the new version hold for the situations of both this exercise *AND* Exercise 1?

In 2(d) you should have arrived at the following, which we sometimes call the **addition rule**.

Theorem 1.11: For events A and B from the same sample space S ,

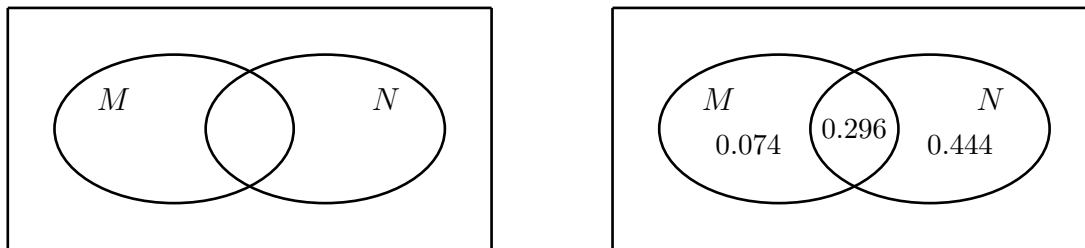
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

3. The table below shows the distribution of 27 students in one of Mr. Waterman's statistics classes relative to gender and smoking habits. (For example, the number 2 means that two of the students are males who smoke.)

	Male	Female	
Smoker	2	5	7
Non-Smoker	8	12	20
	10	17	27

- (a) Find the probability of randomly selecting a male or a non-smoker, showing clearly how you obtained your answer.
- (b) Can you find other ways to obtain the same answer? There are several more...

Something called a **Venn diagram** can be a useful tool for solving probability problems. (See additional discussion of Venn diagrams in Appendix C.) Consider the data from Exercise 3, with the experiment of randomly selecting a student from the class. If we let M be the event of selecting a male and N be the event of selecting a non-smoker, we can draw the Venn diagram shown below and to the left. The rectangle represents the sample space (the entire class) and the two ovals represent the males and non-smokers. The region where the two ovals overlap represents the individuals who are both male and non-smokers, and the region outside both ovals represents the individuals who are not male and who smoke.



Above and to the right the Venn diagram has been modified by putting in each region the probability of the event it represents. Note that the probability of 0.074 is for male smokers. That is, *it is the probability of being in the oval representing the smokers and NOT in the oval representing the non-smokers*. Similarly the probability of 0.444 is for the non-smoking females, not all of the non-smokers. The probability of 0.296 is for the male non-smokers.

The remaining probability outside both ovals can easily be found and written into the appropriate part of the Venn diagram - this is left to the reader. Putting the probabilities in using this method has the advantage that, once labelled, we can obtain any probability relating to the Venn diagram by simply adding some of the numbers.

4. Use the Venn diagram to find the probability of selecting an individual who
 - (a) is a female non-smoker;
 - (b) who is female;
 - (c) who is a male or a non-smoker, but not both.
5. Draw a Venn diagram for a situation in which $P(A \cup B) = P(A)$. Explain what is happening here in terms of Theorem 1.11.
6. Draw a Venn diagram for a situation in which $P(A \cup B) = P(A) + P(B)$. Explain what is happening here in terms of Theorem 1.11.
7. For an experiment with Events A and B , $P(A) = 0.42$, $P(B) = 0.17$ and $P(A - B) = 0.35$. Find each of the following.
 - (a) $P(A \cup B)$
 - (b) $P(A \cap B)$
 - (c) $P[(A \cup B)^c]$
8. On a Friday night vehicles are stopped at a roadblock and a police officer has a brief conversation with the driver to try to determine whether they are drunk. For the experiment of randomly selecting a driver, let Event D be that the driver is drunk and Event B be that the

officer believes the driver is drunk and chooses to administer a breathalyzer test. Clearly we can have any combination of these events. We are given that

$$P(D) = 0.14, \quad P(B) = 0.227, \quad P(D \cap B) = 0.098.$$

Find the probability of each of the following:

- (a) The driver is not drunk.
- (b) The driver is drunk *OR* the officer administers a breathalyzer test.
- (c) The driver is not drunk *AND* a breathalyzer test is not administered.

1.5 Conditional Probability and Independence

Performance Criteria:

1. (g) Compute a conditional probability.
- (h) Determine whether two events are independent.

1. Consider again the data from Exercise 3 of the last section, shown below, and assume that one student will be drawn at random.

	Male	Female	
Smoker	2	5	7
Non-Smoker	8	12	20
	10	17	27

- (a) What is the probability that the student drawn is a smoker?
- (b) Suppose that a student is drawn at random, and we know they are male. What is the probability that the student drawn is a smoker?

The probability that you computed in 1(b) is called a **conditional probability** (because you found the probability of getting a smoker under the condition that the person selected was a male). The language we often use for talking about conditional probabilities is a little bit special; the question of 1(b) will be phrased “What is the probability that a randomly selected student is a smoker, *given that they are a male?*”

2. A die is rolled.

- (a) What is the probability of getting a two, given that a number less than or equal to four was obtained?
- (b) What is the probability of getting a two, given that an even number was obtained?
- (c) What is the probability of getting an even number, given that a two was obtained?

3. Use the data from Exercise 1 above to answer the following.

- (a) What is the probability of selecting a smoker, given that you have selected a female?
- (b) What is the probability of selecting a male, given that you have selected a non-smoker?
- (c) What is the probability of selecting a non-smoker, given that you have selected a male?

Suppose that an experiment has events A and B . We denote the probability of B given A by $P(B|A)$. Note that the last two parts of each of the above exercises illustrate that $P(B|A)$ and $P(A|B)$ are not necessarily the same. (Can you think of an experiment with two events A and B where they *ARE* the same?) We will now consider the question of how $P(B|A)$ is related to $P(A)$, $P(B)$, or other probabilities.

Think about how you computed your answer to 1(b). You probably considered that there are ten males and two of those are smokers, so the probability of selecting a smoker given that you

have selected a male is $\frac{2}{10} = \frac{1}{5}$. Suppose that Sm is the event of selecting a smoker (I'm not using S here because it is reserved for the sample space) and M is the event of selecting a male. With just a little thought we can see that the probability of selecting a smoker, given that we've selected a male is

$$\frac{2}{10} = \frac{2/27}{10/27} = \frac{P(Sm \cap M)}{P(M)}.$$

We denote this probability by $P(Sm|M)$. In general, we have this definition.

Let A and B be events with the same sample space, and with $P(A) \neq 0$. Then the **conditional probability of B given A** is

$$P(B|A) = \frac{P(B \cap A)}{P(A)}.$$

Clearly there are times, like in Exercises 1-3, when conditional probabilities can be determined without this definition. Sometimes the definition is necessary, however.

4. Recall the information from Exercise 8 of the previous section: On a Friday night vehicles are stopped at a roadblock and a police officer has a brief conversation with the driver to try to determine whether they are drunk. For the experiment of randomly selecting a driver, let Event D be that the driver is drunk and Event B be that the officer believes the driver is drunk and chooses to administer a breathalyzer test. Clearly we can have any combination of these events. We are given that

$$P(D) = 0.14, \quad P(B) = 0.227, \quad P(D \cap B) = 0.098.$$

- (a) What is the probability that a person is given the breathalyzer test, given that they are drunk?
 - (b) What is the probability that a person is not drunk, given that they are given the breathalyzer test?
 - (c) Which of the above probabilities would the police want to be high, and why? Which would the police want to be low, and why?
5. Find $P(A|B)$ and $P(B|A)$ for the situation of Exercise 7 of the previous section.
 6. A card is drawn from a standard deck of cards.
 - (a) What is the probability that the card is a king?
 - (b) What is the probability of selecting a king, given that you have selected a face card?
 - (c) What is the probability of selecting a king given that you have selected a heart?

Note that the probabilities for Exercises 6(a) and (c) are the same. When the probability of an event is the same as the probability of that event given a second event, we say that the two events are **independent**. We might then define two events A and B to be independent if $P(B|A) = P(B)$, but this presents a problem with the above definition if A happens to be the null set. We get around this as follows. If two events A and B are independent we then have $P(B|A) = P(B)$, so from the definition of conditional probability we have

$$P(B) = \frac{P(B \cap A)}{P(A)} \implies P(A \cap B) = P(A)P(B).$$

This leads us to make the following definition.

Two events A and B are **independent events** if

$$P(A \cap B) = P(A)P(B).$$

We can sometimes sense intuitively whether or not two events are independent, by simply asking whether one event occurring should influence the probability that the other will occur. We must still use the above definition to show formally that the two events are independent, however.

7. Consider the experiment of flipping a coin and rolling a die.
 - (a) Let H be the event of getting a head on the coin and A be the event of getting a two on the die. Do you think these events are independent? Find $P(H \cap A)$, $P(H)$ and $P(A)$. Are the events of getting a head on the coin and a two on the die independent?
 - (b) Let A still be the event of getting a two on the die and let B be the event of getting an even on the die. Do you think these events are independent? Use the definition to determine whether they actually are.
8. A pair of dice are rolled. Let A be the event of getting a two on either or both die. Let B be the event that the sum of the numbers on the die is 7. Do you think these events are independent? Use the definition to find out whether they are.
9. Suppose that an urn contains two blue marbles and three yellow marbles, and two marbles are to be drawn, without replacement. Find the probability of the event of drawing two yellow marbles as follows.
 - (a) Let B_1 and B_2 stand for the two blue marbles and Y_1 , Y_2 and Y_3 stand for the three yellow marbles. Draw a tree diagram of all possible outcomes for drawing two marbles without replacement.
 - (b) Of course each possible outcome has the same probability, so you can determine the desired probability from your tree diagram. Do so.

1.6 The Multiplication Rule

Performance Criteria:

1. (i) Apply the multiplication rule to determine probabilities.

Recall the last exercise from the previous section, in which two marbles are drawn, without replacement, from an urn containing two blue marbles and three yellow marbles. You were asked to find the probability of the event of drawing two yellow marbles. The method used there for finding the desired probability was not difficult, but it would be a bit cumbersome if there have been 17 yellow marbles and 12 blue marbles! In this section we will see a method for working with such probabilities in a simpler way.

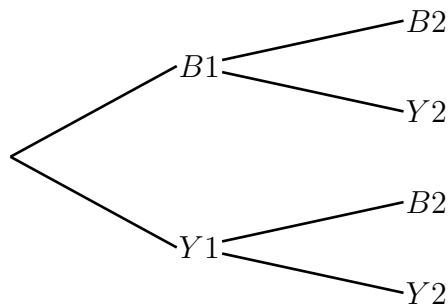
1. Suppose again that we have an urn with two blue marbles and three yellow marbles, and that we are again going to select two marbles without replacement.
 - (a) Let A be the event of getting a yellow marble on the first selection, and B be the event of getting a yellow marble on the second selection. Do you think the events are independent? Verify your answer using the definition of independent events.
 - (b) Find $P(A)$. You should also be able to find the probability $P(B|A)$ without using the formula in the definition. Do so.
 - (c) Note that the event of getting two yellow marbles is the event $A \cap B$. Thus the probability of getting two yellow marbles is $P(A \cap B)$. Insert your answers to (b) into the formula for the definition of conditional probability and solve for $P(A \cap B)$.

The previous exercise illustrates the following, which we call the **multiplication rule**.

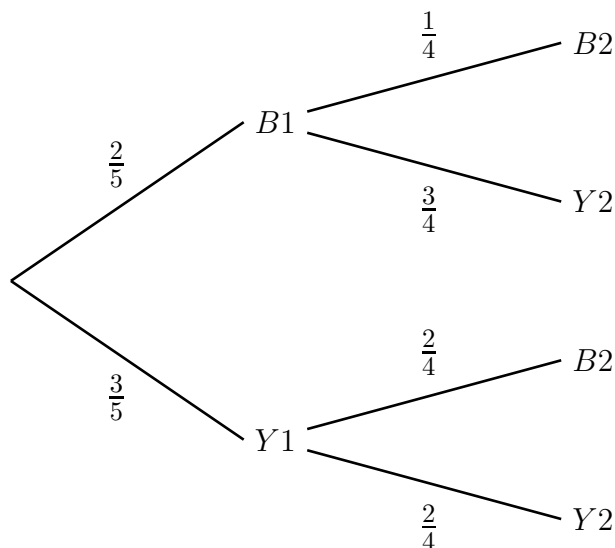
Theorem 1.12: For any two events A and B from the same sample space and with $P(A) \neq 0$,

$$P(A \cap B) = P(A)P(B|A).$$

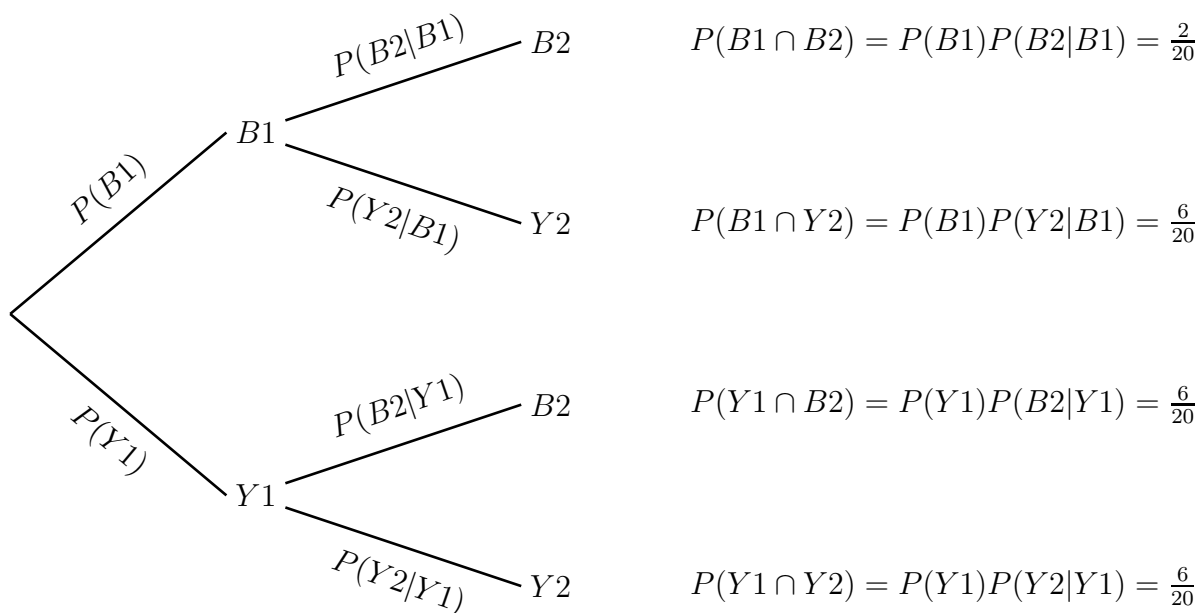
We can use the multiplication rule to make a type of tree diagram that is more efficient for finding probabilities like the one desired in Exercise 1. Such a diagram can be used in any situation where experiments are performed in sequence, or can be thought of as being performed in sequence. For our situation we begin with the following tree diagram that shows what can happen at each step, selecting the first marble and selecting the second. Here $B1$ means getting a blue marble on the first selection, $Y2$ means yellow on the second selection, and so on.



We then put the probability of following each branch of the tree *given that the previous branch has already been followed* along each branch, as shown here:



By the multiplication rule we can then obtain the probability of arriving at the end of any final branch of the tree by multiplying the probabilities along all of the branches that lead to that point:



2. Use the type of tree diagram just demonstrated to find the probability of selecting two yellow marbles when selecting two marbles without replacement from an urn containing 17 yellow marbles and 12 blue marbles.

1.7 Bayes' Theorem

Performance Criteria:

1. (j) Use Bayes' Theorem to determine probabilities.

Let's begin this section with a classical and interesting problem: Suppose that it is known that 2 out of every 1000 adult Americans is afflicted with a particular disease. There is a test to determine whether a person has the disease, but it is not perfect. In 99% of the cases where a person has the disease, the test will say that they have it. (We will say the person "tests positive.") It is also known to give a positive result for 3% of the people who do not have the disease. Suppose now that you are diagnosed as having the disease. Due to the imperfections of the test, it is possible that you really do have the disease, but it is also possible that you do not. (When a test indicated that a person has a disease when in fact they don't, it is called a **false positive**.) Our eventual goal is to compute the probability that you really do have the disease.

1. Consider the experiment of randomly selecting an adult American. Let A be the event that a person has the disease and let B be the event that a person tests positive for the disease.
 - (a) There are three probabilities given above. Give each of them in terms of the events A and B .
 - (b) In terms of the events A and B , what probability is it that we wish to compute? Give the correct "formula" for computing that probability.
 - (c) Think of this experiment as a two-step process: At the first step the randomly selected person either has the disease or they don't, and at the second step they either test positive for the disease or they don't. Draw a tree diagram for this situation and label each of the branches with the probability of "travelling" that branch, as we have already done. (For an example, see the tree diagrams in the previous section.)
 - (d) Find $P(B)$, the probability that a person tests positive for the disease. Don't round your answer.
 - (e) Find the desired probability to three places past the decimal, which is the probability that you really have the disease, given that you tested positive. Do you have cause for immediate concern?
2. The objective of this exercise is to obtain a formula for determining the probability you just found, in terms of the known probabilities.
 - (a) Redraw your tree diagram, but instead of labelling the branches whose probabilities were originally known with the probabilities themselves, label them with the symbolic representations of those probabilities, like the second tree diagram in the previous section.
 - (b) Use the tree diagram to give an equation for $P(B)$ in terms of the known probabilities. (Note that even though the probability of *not* having the disease was not explicitly given, it is easily found, so consider it given.)
 - (c) We were given $P(B|A)$. Give the formula for this probability, and solve it for $P(A \cap B)$.
 - (d) The probability that you were looking for was $P(A|B)$, which by definition is $P(A|B) = \frac{P(A \cap B)}{P(B)}$. Substitute your results from parts (b) and (c) into this to get the desired formula for $P(A|B)$.

- (e) Use your formula from (D) to compute the desired probability, showing a step where each of the known probabilities is substituted into the formula. You should, of course, get the same result as you did for exercise 1!

Your final result from the previous exercise should have been the following.

Theorem 1.13 (Bayes' Theorem): For any two events A and B from the same sample space with $P(A) \neq 0$ and $P(B) \neq 0$,

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A')P(B|A')}.$$

Note that if $P(A) = 0$ then $P(A|B) = 0$, and if $P(B) = 0$ then $P(A|B)$ makes no sense.

3. It is known that about 73% of all messages sent to a person's e-mail address are "spam." (I made this number up, so don't repeat it! If anyone knows the correct figure, please let me know.) A particular spam filter detects and removes 97% of all spam messages, but it also removes 5% of the non-spam messages. Determine the probability that a removed message is spam. Do this by either using a tree, or by defining events and applying Bayes' Theorem. If you choose the second method, you might try to define events A and B in such a way that your events correspond to those of the theorem.
4. A small manufacturing operation employs three people Ann, Bob and Cathy for assembly of Widgets. The three work at different speeds, so out of every 1000 Widgets assembled, Ann assembles 355, Bob assembles 314 and Cathy assembles 331. Based on past data, it is known that each of the three assembles the following percentage incorrectly: Ann, 4.7%, Bob 3.1% and Cathy 2.5%.
 - (a) What is the probability that a Widget is assembled incorrectly? (Use a tree diagram?)
 - (b) Let A , B and C be the events that a randomly selected Widget is assembled is assembled by Ann, Bob or Cathy, respectively, and let I be the event that a Widget is assembled incorrectly. Label each of the given probabilities in terms of this notation. For example, $P(A) = \frac{355}{1000}$.
 - (c) Using the notation from (b), give a formula for how you computed $P(I)$ in terms of the given probabilities.

The last exercise illustrated a simple version of the following theorem, which is sometimes called the **total probability rule**.

Theorem 1.14: Let $A_1, A_2, \dots, A_n$ be a partition of the sample space of an experiment. Then for any event B for the same experiment,

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(A_n)P(B|A_n).$$

5. Draw a tree for this theorem, indicating missing branches with $\vdots$. Label the branches that are relevant to the theorem with their probabilities, and put the probabilities at the ends of the relevant branches.

6. Consider the situation of Exercise 4. Suppose that a randomly selected Widget is assembled incorrectly. What is the probability that it was assembled by Cathy?

Note that the sets A and A' of Bayes' Theorem partition the sample space S . In the previous exercise the events A , B and C also partition the sample space, so whether you knew it or not you were actually using Bayes' Theorem to do the last exercise, but you were using a "bigger" version. I won't give it here because it is very confusing notationally, but it can be found in most books on probability or mathematical statistics.

1.8 Chapter 1 Exercises

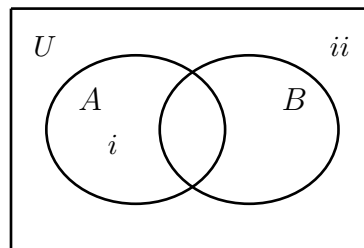
1. The Venn diagram below and to the right is for the experiment of rolling a pair of dice. Event A is the event that at least one of the die is a three and Event B is that the sum of the numbers on the dice is at least seven.

(a) It should be clear that the two events are not mutually exclusive. Give the event indicated with the letter i on the diagram, *as a set*.

(b) Give the event indicated by ii on the diagram, *as a set*.

(c) Give the event from (a) as a union or intersection of two of A , B , A' or B' .

(d) Give the event from (b) as a union or intersection of A and B or their complements.



2. Suppose that three letters are to be selected from the letters A, B, C, D and E, without regard to order and without replacement.

(a) Determine how many different ways this can be done by using one of the counting formulas from the previous section. Which theorem did you use?

(b) Which other theorem also applies?

(c) Draw a tree diagram illustrating all possible outcomes of this experiment. Write each outcome at the end of the branch of the tree resulting in that outcome.

3. Suppose that we were to instead select three letters from the same set as in the previous exercise (still without replacement), but with different orders of the same three letters being distinct. We wish to determine how many outcomes this experiment has.

(a) Which theorem from the previous section applies? Use that formula to determine the number of outcomes.

(b) If you were to draw a tree showing all the outcomes of this experiment it would be rather large. How many branches would it have at the first level? At the second level? at the third?

(c) There are actually two theorems from the previous section that apply to this scenario. Find the other one and apply it to find the number of outcomes.

4. Determine how many arrangements of three letters from the same list there are if we allow each letter to be used more than once. This is equivalent to asking how many ways can we draw three of the five letters with replacement, if different orders of the same letters are considered distinct. This exercise can be answered using one of the counting formulas.

5. Determine how many arrangements of three letters from the same list there are if we allow each letter to be used more than once, but we don't distinguish between different arrangements of the same three letters?. I don't think that this question can easily be answered using the counting formulas. Try using a tree diagram to answer this question.

6. A class has 31 students. Suppose that I decide to give away \$120 to three students in the class.
 - (a) Suppose that I decide to give \$40 to each of three students. How many ways can this be done?
 - (b) Suppose that I decide to give \$60 to one student, \$40 to another, and \$20 to a third. How many ways can this be done?
7. In a particular state, licence plates are to consist of a three digit number, followed by a letter. *The first digit of the number is not allowed to be zero.*
 - (a) How many different number/letter arrangements are there? (Different orders of the same digits are considered distinct, since they can be distinguished from each other.)
 - (b) How many of number/letter arrangements consist of an odd number followed by a consonant?
 - (c) A witness to a bank robbery gets a glimpse of the getaway car as it speeds off. They remember that the letter is an L and that the digits of the number are 0, 5 and 8, but they can't remember what order they were in. How many plates do the police need to trace?
8. An urn contains three blue marbles and two yellow marbles. A single marble is drawn repeatedly, *with replacement*, until yellow has been obtained twice.
 - (a) This experiment has an infinite, but discrete, sample space. Write it by listing the "first" six events in the sample space, using the notation BYBBY, YBY, etc.
 - (b) Find the probability of obtaining a yellow marble twice within the first three draws of a marble. (There are a couple of ways that I can think of to do this. One is to use a probability tree of the sort that we have been working with this week.)
 - (c) Suppose now that there is a new urn, containing 30 blue marbles and 20 yellow marbles, and that marbles are drawn *without replacement* until two yellow marbles have been drawn. Find the probability of obtaining a yellow marble twice within the first three draws of a marble.
9. Consider the experiment that a coin is flipped and a four-sided die is rolled.
 - (a) Give the sample space for this experiment.
 - (b) Suppose that we assign the value one to a flip of the coin resulting in a head, and zero to a flip resulting in a tail. Give the event that the sum of the numbers from the coin and the die is even.
 - (c) Determine the probability that the sum of the numbers from the coin and the die is less than three.
 - (d) Determine the probability that the sum of the numbers of the coin and the die is greater than three.
10.
 - (a) Give two events A and B from the experiment of the previous exercise for which $P(A \cup B) = P(A) + P(B)$ is true. You may describe the events in words OR with set notation, but you need NOT do both.
 - (b) Give two events C and D from the experiment of the previous exercise for which $P(C \cup D) = P(C) + P(D)$ is *NOT* true.

11. Consider the following experiment: A coin is flipped once. If it comes up heads, then a four-sided die (with equal probability of getting any of the values from 1 to 4) is rolled and the experiment is over. If the coin comes up tails, then it is flipped two more times and the experiment is over. So the possible outcomes are

$$H1, H2, H3, H4, TTT, TTH, THT, THH,$$

and *all eight outcomes are equally likely*. Consider also the following events:

Event A: An even number is obtained on the die.

Event B: A tail is obtained on the second flip of the coin.

Event C: A head is obtained on the first flip of the coin.

- (a) Draw a probability tree for the diagram with all probabilities labelled appropriately. It should show that all eight outcomes do in fact have equal probabilities.
 - (b) Find the probability of each of the following.
 - (i) $P(B) =$ _____
 - (ii) $P(B \cup C) =$ _____
 - (iii) $P(B \cap C) =$ _____
 - (iv) $P(A|C) =$ _____
 - (v) $P(C|A) =$ _____
 - (c) Show/explain *mathematically* why Events A and C are not independent.
12. An experiment consists of flipping a coin repeatedly until a head is obtained; The possible outcomes are

$$H, TH, TTH, TTTH, TTTTH, TTTTTH, \dots$$

- (a) What is the probability of obtaining a head on the first flip?
 - (b) What is the probability of obtaining a head on the second flip?
 - (c) What is the probability of obtaining a head on the third flip?
 - (d) What is the probability that it will take at least three flips to obtain the first head?
13. A coin is flipped once.
- (a) Give the sample space for the experiment.
 - (b) List all possible outcomes. Remember that an outcome is simply a subset of the sample space.
14. A coin is flipped twice.
- (a) Give the sample space for the experiment.
 - (b) List all possible outcomes. Remember that an outcome is simply a subset of the sample space.
15. A coin is flipped three times.
- (a) Give the sample space for the experiment.
 - (b) Give the event that exactly one head is obtained, as a set.
 - (c) Give the event that at least one head is obtained.

16. The digits 3, 4 and 7 are used to create three digit numbers, using each of those three digits exactly once.
 - (a) Give the sample space.
 - (b) Give the event that the number is less than 350.
 - (c) Give the event that the number is even.

17. Two marbles are drawn randomly, with replacement, from an urn containing five blue marbles and seven red marbles. In this exercise you will determine the probability that exactly one of the marbles is red by using a Venn diagram.
 - (a) What is the probability of selecting a red marble on the first draw?
 - (b) What is the probability of selecting a red marble on the second draw?
 - (c) What is the probability of selecting a red marble on both draws, and why?
 - (d) Let Event A be that a red marble is selected on the first draw (but not necessarily on the first draw only), and let Event B be that a red marble is selected on the second draw.
 - (e) Sketch a Venn diagram illustrating the sample space and events given. Fill each region with its probability as an event.
 - (f) Use your Venn diagram to determine the probability of getting exactly one red marble.
 - (g) Use your Venn diagram to determine the probability of getting no red marbles.

18. Two marbles are drawn randomly, with replacement, from an urn containing five blue marbles and seven red marbles. In this exercise you will determine the probability that exactly one of the marbles is red by using a tree diagram.
 - (a) Draw a tree diagram for this experiment, labeling each branch with its probability.
 - (b) Give the outcome at the end of each branch (like RB) and give its probability.
 - (c) Use your tree diagram to determine the probability of getting exactly one red marble.
 - (d) Use your tree diagram to determine the probability of getting no red marbles.

19. Two marbles are drawn randomly, with replacement, from an urn containing five blue marbles and seven red marbles. In this exercise you will determine the probability that exactly one of the marbles is red by using Theorems 1.4 and 1.2.
 - (a) Assuming that the 12 marbles being used can all be distinguished from each other, how many different ways can two marbles be selected? *Remember that marbles are being drawn WITH replacement.* Use Theorem 1.4.
 - (b) How many of those ways consist of a red marble followed by a blue marble?
 - (c) How many of those ways consist of a blue marble followed by a red marble?
 - (d) Use Theorem 1.2 to determine the probability of selecting exactly one red marble.

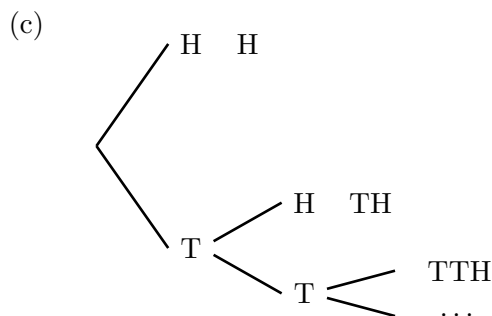
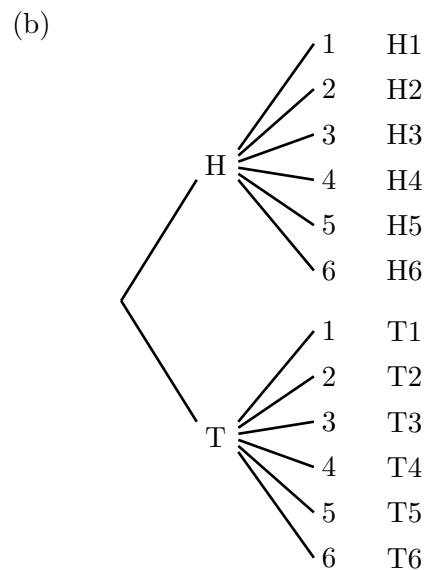
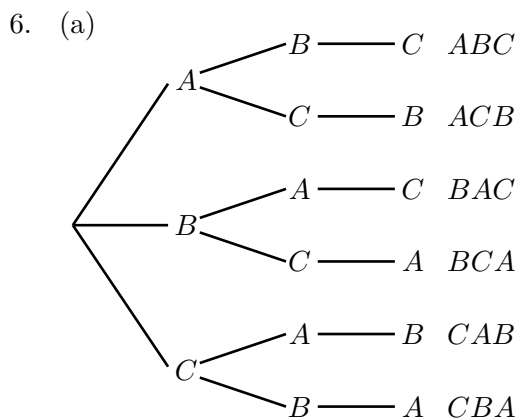
20. An urn contains three blue marbles and two red marbles. An experiment consists of drawing two marbles *without replacement*. What is the probability of having gotten blue for the first marble, given that you got blue for the second?

H Solutions to Exercises

H.1 Chapter 1 Solutions

Section 1.1

1. (a) $\{H, T\}, \{H\}$
 (b) $\{TTT, HTT, THT, TTH, HHT, HTH, THH, HHH\}, \{HHT, HTH, THH, HHH\}$
 (c) $\{H, TH, TTH, TTTH, \dots\}, \{H, TH, TTH, TTTH\}$
 (d) $\{ABC, ACB, BAC, BCA, CAB, CBA, BAC, BCA\}, \{BAC, BCA\}$
2. (a) $\{T1, T2\}$ (b) $\{T1, T2, T3, T4, T5, T6, H1, H2\}$
 (c) $\{H4, H5, H6, T4, T5, T6\}$
3. (a) There are 36 outcomes possible, because there are six possible outcomes on one die and, for each of those outcomes, another six on the second die.
 (b) $\{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$
4. $(10.12, \infty)$ or $\{x \in \mathbb{R} \mid x > 10.12\}$
5. $[10, 15]$ or $\{t \in \mathbb{R} \mid 10 \leq t \leq 15\}$



7. There are 24 ways: TChJCo, TChCoJ, TJChCo, TJCoCh, TCoJCh, TCoChJ, JTChCo, JTCChCo, JChTCo, JChCoT, JCoTCh, JCoChT, ChTJCo, ChTCoJ, ChJTCo, ChJCoT, ChCoTJ, ChCoJT, CoTChJ, CoTJCh, CoChTJ, CoChJT, CoJTCh, CoJChT

8. (a) The events are not mutually exclusive; both events contain HHT, HTH and THH.
 (b) The events are mutually exclusive.
 (c) The events are mutually exclusive.
9. (a) $\{AB, AC, AD, BC, BD, CD\}$
 (b) $\{AA, AB, AC, AD, BB, BC, BD, CC, CD, DD\}$
10. (a) $\{AB, AC, AD\}$ (b) $\{AA, BB, CC, DD\}$
 (c) $\{AB, AC, AD, BC, BD, CD\}$
11. $\{(AB, CDE), (AC, BDE), (AD, BCE), (AE, BCD), (BC, ADE), (BD, ACE), (BE, ACD), (CD, ABE), (CE, ABD), (DE, ABC)\}$
12. $\{(AB, CD, E), (AB, CE, D), (AB, DE, C), (AC, BD, E), (AC, BE, D), (AC, DE, B), (AD, BC, E), (AD, BE, C), (AD, CE, B), (AE, BC, D), (AE, BD, C), (AE, CD, B), (BC, DE, A), (BD, CE, A), (BE, CD, A)\}$

Section 1.2

1. $\frac{4}{8}$
2. (a) 36 (c) $\frac{3}{36}$ (d) $\frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{6}{36}$
 (e) The sum of the numbers is five or more. The probability of this is $1 - \frac{6}{36} = \frac{30}{36}$.

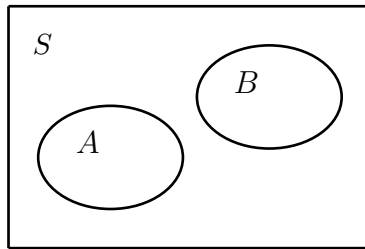
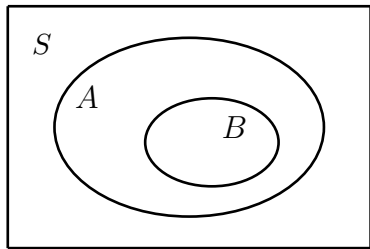
Section 1.3

1. (a) Mr. W can dress six ways, BB, BG, BR, KB, KG, KR (b) $\frac{1}{6}$
2. (a) $26 \cdot 26 \cdot 10 \cdot 10 \cdot 9 = 608,400$ (b) $1 \cdot 1 \cdot 10 \cdot 10 \cdot 1 = 100$ (c) $\frac{52000}{608400}$
3. (a) 2 (b) $3 \cdot 2 \cdot 1 = 6$ (c) $4 \cdot 3 \cdot 2 \cdot 1 = 24$ (d) $n(n-1)(n-2) \cdots 2 \cdot 1$
4. (a) $4 \cdot 3 = 12$ (b) $8 \cdot 7 \cdot 6 = 336$ (c) $18 \cdot 17 \cdot 16 \cdot 15 \cdot 14 = 1,028,160$
5. (a) 6 (b) 56
6. $\frac{16!}{5!2!9!} = 240,240$ 7. $\frac{8!}{3!5!} = 56$ 8. $\frac{5!}{2!3!} = 10$ 9. $\frac{5!}{2!2!1!} = 30$

Section 1.4

1. (a) $P(F) = \frac{12}{52} = \frac{3}{13}$ (b) $P(E) = \frac{20}{52} = \frac{5}{13}$ (c) $P(F \cup E) = \frac{32}{52} = \frac{8}{13}$
 (d) $P(A \cup B) = P(A) + P(B)$
2. (a) C and F are not mutually exclusive because the Jack, Queen and King of clubs are in both events.
 (b) $P(C) = \frac{13}{52}$, $P(F) = \frac{12}{52} = \frac{1}{4}$ (c) $P(A \cup C) = \frac{22}{52} = \frac{11}{26}$
3. (a) $P(\text{male or non-smoker}) = P(\text{male}) + P(\text{smoker}) - P(\text{male and smoker})$
 $= \frac{10}{27} + \frac{20}{27} - \frac{8}{27} = \frac{22}{27}$

4. (a) $P(M' \cap N) = 0.444$ (b) $P(M') = 0.186 + 0.444 = 0.630$
 (c) $0.074 + 0.444 = 0.518$
5. See Venn diagram below and to the left. Since $B \subset A$, $A \cap B = B$, which then gives $P(B) - P(A \cap B) = 0$.



6. See Venn diagram above and to the right. Since $A \cap B = \emptyset$, $P(A \cap B) = 0$.
7. (a) $P(A \cup B) = 0.17 + 0.35 = 0.52$ (b) $P(A \cap B) = 0.42 - 0.35 = 0.07$
 (c) $P[(A \cup B)'] = 1 - 0.52 = 0.48$
8. (a) $P(D') = 1 - 0.14 = 0.86$ (b) $P(D \cup B) = 0.14 + 0.227 - 0.098 = 0.269$
 (c) $P(D' \cap B') = P[(D \cup B)'] = 1 - 0.269 = 0.731$ Note the use of DeMorgan's Law!

Section 1.5

1. (a) $\frac{7}{27}$ (b) $\frac{2}{10}$
2. (a) $\frac{1}{4}$ (b) $\frac{1}{3}$ (c) 1
3. (a) $\frac{5}{17}$ (b) $\frac{8}{20}$ (c) $\frac{8}{10}$
4. (a) $\frac{0.098}{0.14} = 0.70$ (b) $\frac{0.129}{0.227} = 0.568$
5. $P(A|B) = \frac{0.07}{0.17} = 0.41$ $P(B|A) = \frac{0.07}{0.42} = 0.17$
6. (a) $\frac{4}{52} = \frac{1}{13}$ (b) $\frac{4}{12}$ (c) $\frac{1}{13}$
7. (a) $P(H \cap A) = \frac{1}{12}$, $P(H) = \frac{6}{12} = \frac{1}{2}$, $P(A) = \frac{212}{6}$ The events are independent because $P(H \cap A) = P(H)P(A)$.
 (b) $P(A) = \frac{1}{6}$, $P(B) = \frac{6}{12} = \frac{1}{2}$, $P(A \cap B) = \frac{1}{6}$ The events are not independent.
8. The events are not independent.

Section 1.6

1. (a) The events should not be independent. If we get a yellow on the first selection, then we are less likely to get a yellow on the second. $P(A) = \frac{3}{5}$, $P(B) = \frac{12}{20}$, $P(A \cap B) = \frac{6}{20}$. Since $P(A \cap B) \neq P(A)P(B)$, the events are not independent.
 (b) $P(A) = \frac{3}{5}$, $P(B|A) = \frac{2}{4}$
 (c) $P(B|A) = \frac{P(B \cap A)}{P(A)} \Rightarrow \frac{2}{4} = \frac{P(B \cap A)}{\frac{3}{5}} \Rightarrow P(B \cap A) = \frac{2}{4} \cdot \frac{3}{5} = \frac{6}{20}$
2. $\frac{17}{29} \cdot \frac{16}{28} = \frac{68}{203}$

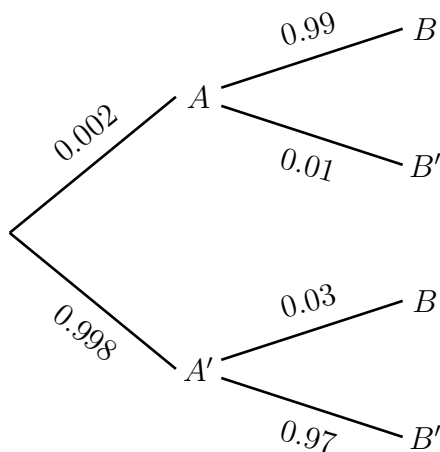
Section 1.7

1. (a) $P(A) = \frac{2}{1000} = 0.002$, $P(B|A) = 0.99$, $P(B|A') = 0.03$

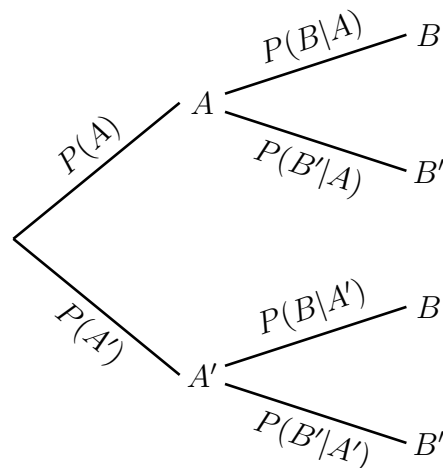
(b) $P(A|B) = \frac{P(A \cap B)}{P(B)}$ (c) See below.

(d) $P(B) = (0.002)(0.99) + (0.998)(0.03) = 0.03192$

(e) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{(0.002)(0.99)}{0.03192} = 0.062 = 6.2\%$



Exercise 1(c)



Exercise 2(a)

2. (a) See above. (b) $P(B) = P(A)P(B|A) + P(A')P(B|A')$

(c) $P(B|A) = \frac{P(A \cap B)}{P(A)} \Rightarrow P(A \cap B) = P(A)P(B|A)$

(d) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A')P(B|A')}$

(e) $P(A|B) = \frac{(0.002)(0.99)}{(0.002)(0.99) + (0.998)(0.03)} = 0.062$

3. **Experiment:** Select a random e-mail message sent to a person.

Event A: The message is spam. **Event B:** The message is removed.

$P(A) = 0.73$, $P(B|A) = 0.97$, $P(B|A') = 0.05$

$P(A|B) = \frac{(0.73)(0.97)}{(0.73)(0.97) + (0.27)(0.05)} = 0.981$

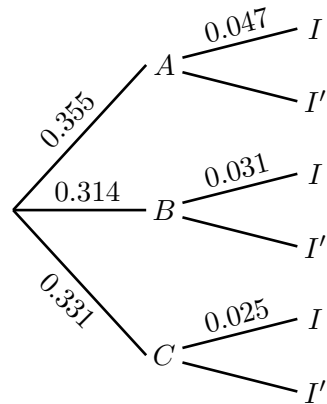
4. (a) $P(I) = (0.355)(0.047) + (0.314)(0.031) + (0.331)(0.025) = 0.034694$

See tree diagram below and to the right.

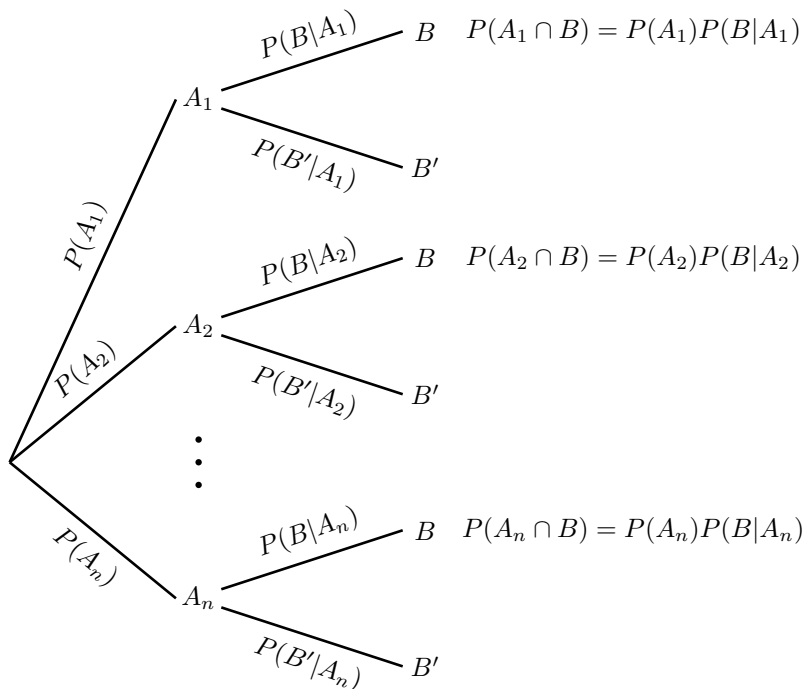
(b) $P(A) = 0.355, P(B) = 0.314, P(C) = 0.331$

$P(I|A) = 0.047, P(I|B) = 0.031, P(I|C) = 0.025$

(c) $P(I) = P(A)P(I|A) + P(B)P(I|B) + P(C)P(I|C)$



5.



6. $P(C|I) = \frac{P(C \cap I)}{P(I)} = \frac{(0.331)(0.025)}{0.034694}$