

Student Outcomes for this Chapter

Section 4.1: Contingency Tables

Students will be able to:

- Relate Venn diagrams and contingency tables
- Calculate percentages from a contingency table
- Calculate *and* empirical probabilities
- Calculate *or* empirical probabilities
- Calculate conditional probabilities
- Determine whether two characteristics are independent

Section 4.2: Theoretical Probability

Students will be able to:

- Write the sample space for theoretical probability situations
- Identify certain and impossible events
- Calculate the theoretical probability of a complement
- Determine the difference between empirical and theoretical probability
- Explain the Law of Large Numbers
- Identify independent and dependent events
- Calculate *and* theoretical probabilities
- Identify overlapping and disjoint sets
- Calculate *or* theoretical probabilities
- Calculate probability values for simple games

Section 4.3: Expected Value

Students will be able to:

- Make a probability model
- Calculate the expected value for a probability model
- Determine whether a game is fair

Section 4.1 Contingency Tables

When we looked at categorical data in the previous chapter, it was related to a single variable, or characteristic of interest, such as favorite movie or car color. To illustrate the data, we made a frequency table and used it to create a pie chart or bar chart. But what if we want to illustrate the relationship between two categorical variables? To do this, we can use a contingency table.

Contingency Tables

A **contingency table** summarizes all the possible combinations for two categorical variables. Each value in the table represents the number of times a particular combination of outcomes occurs. For example, suppose we randomly select 250 households from the greater Portland area and ask whether they have a cat and whether they have a dog. In this case, “have a cat” and “have a dog” are the two variables, and each variable has two categories: Yes and No.

To create the contingency table, we make columns for the categories of one variable, and rows for the categories of the other variable. We also add a row and column for the subtotals of each category. Each cell of the resulting table contains the number of outcomes having the characteristics of the intersecting row and column categories. For our dog and cat example, the table would look like this:

	Yes Dog	No Dog	Total
Yes Cat	<i>Yes Cat and Yes Dog</i>	<i>Yes Cat and No Dog</i>	<i>Yes Cat Total</i>
No Cat	<i>No Cat and Yes Dog</i>	<i>No Cat and No Dog</i>	<i>No Cat Total</i>
Total	<i>Yes Dog Total</i>	<i>No Dog Total</i>	<i>Grand Total</i>

Suppose that of the 250 households surveyed, 180 said they have a cat, 95 said they have a dog, and 52 said they have both a cat and a dog. We can use this information to fill in the cells of the table.

The first cell we can fill in is the **grand total**, which is the total number of subjects in the study. In this case, there are 250 households participating in the survey. The next two cells we can fill in are the total number of households that have a cat, 180, and the total number of households that have a dog, 95. The final cell we can fill in from the given information is the intersection of the having a dog column and a having a cat row, which is 52 households.

	Yes Dog	No Dog	Total
Yes Cat	52		180
No Cat			
Total	95		250

Since each row and column must sum to their totals, we can use subtraction to find the missing numbers as shown below.

	Yes Dog	No Dog	Total
Yes Cat	52	$180 - 52 = 128$	180
No Cat	$95 - 52 = 43$	$155 - 128 = 27$ or $70 - 43 = 27$	$250 - 180 = 70$
Total	95	$250 - 95 = 155$	250

Now that we have our contingency table completed, notice that the numbers in the central four cells add to the grand total as shown in the table on the left. The total row and the total column also add to the grand total as shown in the right table.

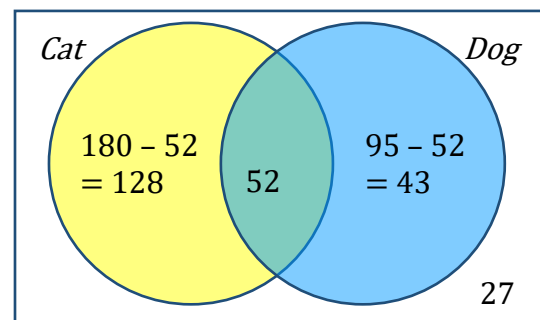
	Yes Dog	No Dog	Total
Yes Cat	52	128	180
No Cat	43	27	70
Total	95	155	250

	Yes Dog	No Dog	Total
Yes Cat	52	128	180
No Cat	43	27	70
Total	95	155	250

Contingency Tables and Venn Diagrams

If the subtractions we just did seem familiar, they should! This is very similar to what we did for reporting data with a Venn diagram. The Venn diagram for this data is shown to the right.

We also subtracted the intersection (green region) from the total of the cat and dog owners to find numbers in the crescent regions.



Notice that the numbers in the four regions of the Venn diagram are the same as the four cells in the center of the contingency table and add to the grand total.

“And” Statements

Now we can use the contingency table or the Venn diagram to determine the percentage of households that meet certain conditions. For instance, what percent of those surveyed own a cat **and do not** own a dog? In the Venn diagram, this is 128 households in the cat only (yellow) region.

In the contingency table we see the 128 households at the intersection of the row of households who own a cat and the column of households who do not own a dog. As a percentage, the total number of households surveyed, is

$\frac{128}{250} = 0.512$ or 51.2% that have a cat and no dog.

	Yes Dog	No Dog	Total
Yes Cat	52	128	180
No Cat	43	27	70
Total	95	155	250

“Or” Statements

How about the percentage of households surveyed that have a cat **or** a dog? We know from Venn diagrams that the inclusive *or* includes the number of households who own a cat only (yellow) region, a dog only (blue) region, and both a cat and a dog (green) region, or $128 + 52 + 43 = 223$ households. As a percentage of the total surveyed, we get

$\frac{223}{250} = 0.892$ or 89.2% of households in the sample have a dog **or** a cat.

	Yes Dog	No Dog	Total
Yes Cat	52	128	180
No Cat	43	27	70
Total	95	155	250

We can get the same answer from the contingency table. by adding the cells for households who have a cat and not a dog, a dog and not a cat, and the households that have both a cat and a dog. This also gives us 223 households.

There is another way to calculate an *or* statements from a contingency table. We could add the row and column totals for having a cat and having a dog, but then we have counted the 52 households in the intersection twice. We can subtract that number to get $180 + 95 - 52 = 223$ households with a dog or a cat, which we know is 89.2% of those surveyed.

Conditional Statements

Another question we can answer using a contingency table is what percentage of dog owning households also own a cat? In this case the group that we are interested in isn’t every household surveyed (the grand total), but just those households that own a dog.

We call this a **conditional** statement because we are only considering the households with a certain condition. If we focus on the column representing the households that own a dog, we see that there is a total of 95 households with a dog, and that 52 of those 95 households also have a cat. Therefore,

	Yes Dog	No Dog	Total
Yes Cat	52	128	180
No Cat	43	27	70
Total	95	155	250

$\frac{52}{95} = 0.547$ or 54.7% of the households with a dog also have a cat. Another way to phrase this conditional statement is, “What percent of households have a cat **given** they have a dog.” You will see the word given quite a bit in this chapter and that makes the denominator change. It is also possible to find this conditional percentage using the Venn diagram by taking the number in the intersection and dividing it by the total in the whole dog circle.

Contingency Tables with More Than Two Categories

When there are only two categories for each variable, like yes/no questions, Venn diagrams and contingency tables provide basically the same information and can be used interchangeably. A Venn diagram works well for yes/no variables since a subject is either inside the circle (has the characteristic) or outside the circle (does not have the characteristic). If we have more than two possibilities for any of the variables, though, we cannot use a Venn diagram. We can use a contingency table, though. Here is an example where one variable has four categories and the other has three categories.

Example 1: 910 randomly sampled registered voters from Tampa, FL were asked if they thought workers who have illegally entered the US should (i) be allowed to keep their jobs and apply for US citizenship, (ii) be allowed to keep their jobs as temporary guest workers but not be allowed to apply for US citizenship, or (iii) lose their jobs and have to leave the country. Not sure was also an option (iv). The results of the survey by political ideology are shown below¹. Use the contingency table to answer the questions.

	Conservative	Moderate	Liberal	Total
(i) Apply for citizenship	57	120	101	278
(ii) Guest worker	121	113	28	262
(iii) Leave the country	179	126	45	350
(iv) Not sure	15	4	1	20
Total	372	363	175	910

- What percent of the sampled Tampa, FL voters identified themselves as conservatives? To answer this question, we find the conservative column and look to the bottom cell for the total number of conservative voters and divide that by the total number of voters surveyed. This gives us $\frac{372}{910} = 0.409$ or 41% of the Tampa, FL voters who identify as conservative.
- What percent of the sampled voters are in favor of the citizenship option? For this question we find the apply for citizenship row, look across to find the total, and divide this by the total number of voters surveyed. We get

¹ SurveyUSA, [News Poll #18927](#), data collected Jan 27-29, 2012. Example adapted from [Open Intro: Advanced High School Statistics](#), by Diez et al, used under [CC-BY-SA 3.0](#).

$\frac{278}{910} = 0.305$ or 31% of these voters are in favor of the citizenship option.

- c. What percent of the sampled voters identify themselves as conservatives **and** are in favor of the citizenship option? For this question we are looking for the cell that is the intersection of those who identify as conservative and those who are in favor of the citizen option. This cell has 57 voters, so we divide that by the total number of voters. This gives us

$\frac{57}{910} = 0.063$ or 6% of these voters identify as conservatives and are in favor of the citizenship option.

- d. What percent of the sampled voters identify themselves as liberal **or** are in favor of the leaving the country option? The *or* in this question is inclusive, so we need to determine the number of voters who identify as liberal, who are in favor of the leaving the country option, or both.

	Conservative	Moderate	Liberal	Total
(i) Apply for citizenship	57	120	101	278
(ii) Guest worker	121	113	28	262
(iii) Leave the country	179	126	45	350
(iv) Not sure	15	4	1	20
Total	372	363	175	910

In terms of the individual cells, the number of voters who have the specified characteristics is the sum $179 + 126 + 101 + 28 + 45 + 1 = 480$, which we can divide by the total number of voters surveyed to get the percent. So, we have

$\frac{480}{910} = 0.527$ or 53% of the voters identify as liberal or are in favor of the

leave the country option.

Another way to calculate this is to add the total number who identify as liberal (175 voters) and the total number who are in favor of the leave the country option (350 voters), then subtract the double counted cell (45 voters) who are liberal and in favor of the leave the country option:

$$175 + 350 - 45 = 480.$$

- e. What percent of the sampled voters who identify as conservatives are also in favor of the citizenship option? What percent of moderate and liberal voters share this view? As we saw before, these are conditional statements. For the first part of this question, we want to focus just on those voters who identify as conservatives, and from among that group determine the percent in favor of the citizenship option. We calculate that

$\frac{57}{372} = 0.153$ or 15% of conservative voters are in favor of the citizenship

option.

For the second part, we want to focus on just those voters who identify as moderate, and from among that group determine the percent in favor of the citizenship option. Then we have

$$\frac{120}{363} = 0.331 \text{ or } 33\% \text{ of moderate voters are in favor of the citizen option.}$$

Finally, we want to focus on just those voters who identify as liberal, and from among that group determine the percent in favor of the citizenship option. We calculate

$$\frac{101}{175} = 0.577 \text{ or } 58\% \text{ of liberal voters are in favor of the citizenship option.}$$

Looking at these three percentages, it is clear that support of the citizenship option **depends** on political ideology. If support of the citizenship option were the same across political ideologies, then we would say that favoring the citizenship option and political ideology were **independent** of each other.

Empirical Probability

If our sample is representative of the population, then we can also interpret a percentage we calculate from a contingency table as a **probability**, or the likelihood that something will happen. Since a contingency table is constructed from data collected through sampling or an experiment, we call it an **empirical** or **experimental** probability. This is different from a **theoretical** probability which we will look at in the next section.

Finding Empirical Probabilities with a Contingency Table

Suppose that 60% of students in our class have a summer birthday (June, July, or August). Now suppose everyone's name and birth month are written on slips of paper and thrown into a bag. If we pull a slip of paper out of the bag at random, what is the probability that the selected student has a summer birthday? If you think there should be a 60% chance, you are right! The relative frequency of the characteristic of interest will be equal to its empirical probability. To write this as a probability statement, it would look like

$$P(\text{summer birthday}) = 60\%$$

Probability is a function named P , and the function is applied to what follows in the parentheses. Let's look at another example where we write probability statements and find empirical probabilities.

Example 2: A survey of licensed drivers asked whether they had received a speeding ticket in the last year and whether their car is red. The results of the survey are shown in the contingency table to the right.

	Speeding Ticket	No Speeding Ticket	Total
Red Car	15	135	150
Not Red Car	45	470	515
Total	60	605	665

Find the probability that a randomly selected survey participant:

- has a red car.

- b. has had a speeding ticket in the last year.
- c. has a red car and has not had a speeding ticket in the last year.
- d. has a red car or has had a speeding ticket in the last year.
- e. has had a speeding ticket in the last year given they have a red car.
- f. who has received a speeding ticket in the last year also has a red car.
- g. What do the answers to b and e suggest about the relationship between owning a red car and getting a speeding ticket?

Here are the solutions:

- a. To find $P(\text{red car})$, we divide the number of participants who own a red car by the total number of people surveyed:

$$P(\text{red car}) = \frac{150}{665} = 0.226 \text{ or } 22.6\%$$

	Speeding Ticket	No Speeding Ticket	Total
Red Car	15	135	150
Not Red Car	45	470	515
Total	60	605	665

- b. To find $P(\text{speeding ticket})$, we divide the number of participants who got a speeding ticket in the last year by the total number of people surveyed:

$$P(\text{speeding ticket}) = \frac{60}{665} = 0.090 \text{ or } 9\%$$

	Speeding Ticket	No Speeding Ticket	Total
Red Car	15	135	150
Not Red Car	45	470	515
Total	60	605	665

- c. To find $P(\text{red and no ticket})$, we find the intersection of the red car category and the no ticket category and divide by the total number of participants:

$$P(\text{red and no ticket}) = \frac{135}{665} = 0.203 \text{ or } 20.3\%$$

	Speeding Ticket	No Speeding Ticket	Total
Red Car	15	135	150
Not Red Car	45	470	515
Total	60	605	665

- d. To find $P(\text{red or ticket})$, we need to add those who drive a red car and did not have a speeding ticket (just red), those who had a speeding ticket and do not drive a red car (just ticket) and those who drive a red car and had a speeding ticket (both), and divide by the total number of participants:

$$\begin{aligned}
 P(\text{red or ticket}) &= \frac{135 + 45 + 15}{665} \\
 &= \frac{195}{665} \\
 &= 0.293 \text{ or } 29.3\%
 \end{aligned}$$

	Speeding Ticket	No Speeding Ticket	Total
Red Car	15	135	150
Not Red Car	45	470	515
Total	60	605	665

Recall from our earlier discussion that we could also calculate the *or* probability as:

$$\begin{aligned}
 P(\text{red or ticket}) &= P(\text{red}) + P(\text{speeding ticket}) - P(\text{red and speeding ticket}) \\
 &= \frac{150}{665} + \frac{60}{665} - \frac{15}{665} \\
 &= \frac{195}{665}
 \end{aligned}$$

which gives us the same answer as counting the individual cells.

- e. The probability $P(\text{speeding ticket given red car})$ is a conditional probability as we have seen before since it is conditional on the given characteristic occurring. In this problem, the given characteristic is owning a red car, so we isolate our attention to just the row of 150 red car owners and see how many have had a speeding ticket in the last year. Looking at the table, we see that there were 15 red car owners who had a speeding ticket in the last year, so we calculate:

	Speeding Ticket	No Speeding Ticket	Total
Red Car	15	135	150
Not Red Car	45	470	515
Total	60	605	665

$$\begin{aligned}
 P(\text{speeding ticket given red car}) &= \frac{15}{150} \\
 &= 0.10 \text{ or } 10\%
 \end{aligned}$$

- f. This question is also asking for a conditional probability, $P(\text{red car given speeding ticket})$, but it is phrased more like we would say it. In this case the given characteristic is that the person has received a speeding ticket, so we will isolate our attention to just the speeding ticket column. Among the 60 people who had a speeding ticket in the last year, we see that 15 also drove a red car. Now we can calculate the probability:

	Speeding Ticket	No Speeding Ticket	Total
Red Car	15	135	150
Not Red Car	45	470	515
Total	60	605	665

$$P(\text{red car given speeding ticket}) = \frac{15}{60} \\ = 0.25 \text{ or } 25\%$$

Notice that compared with part e, when we change the conditional characteristic, we change the denominator of the fraction.

- g. In part b, we determined that there was a 9% chance of randomly selecting a participant who had received a speeding ticket in the last year. However, in part e we found that there was a 25% chance of receiving a ticket in the last year if the person had a red car. This seems to suggest that there is a higher likelihood of getting a speeding ticket if you own a red car. This means that getting a speeding ticket is **dependent** on whether the person drives a red car, since that increases the probability of getting a ticket. We cannot say, however, whether driving a red car makes you speed or whether people who tend to drive faster buy red cars.

Conditional Probabilities

We have mentioned conditional probabilities, which we find by isolating our attention to the given row or column. Here is another example of finding conditional probabilities.

Example 3: A home pregnancy test was given to a sample of 93 women, and their pregnancy was then verified by a blood test. The contingency table below shows the home pregnancy test and whether or not they were actually pregnant as determined by the blood test. Find the probability that a randomly selected woman in the sample

- was not pregnant given the home test was positive.
- had a positive home pregnancy test given they were not pregnant.

	Positive Test	Negative Test	Total
Pregnant	70	4	74
Not Pregnant	5	14	19
Total	75	18	93

Here are the solutions:

- Since we are given the home test result was positive, we are limited to the 75 women in the positive test column, of which 5 were not pregnant. This gives:

	Positive Test	Negative Test	Total
Pregnant	70	4	74
Not Pregnant	5	14	19
Total	75	18	93

$$P(\text{not pregnant given positive test}) = \frac{5}{75} \\ = 0.067 \text{ or } 6.7\%$$

- b. Since we are given the woman is not pregnant, we are limited to the 19 women in the not pregnant row, of which 5 had a positive test. This gives:

	Positive Test	Negative Test	Total
Pregnant	70	4	74
Not Pregnant	5	14	19
Total	75	18	93

$$P(\text{positive test given not pregnant}) = \frac{5}{19} \\ = 0.263 \text{ or } 26.3\%$$

This result is referred to as a false positive: A positive test result when the woman is not actually pregnant.

In this section we have learned about empirical probability. In the next section we will discuss another kind of probability that you may be familiar with – theoretical probability.

Exercises 4.1

A professor gave a test to students in a morning class and the same test to the afternoon class. The grades are summarized below. Use the table for questions 1-2.

- If one student was chosen at random, find each probability:
 - P(in the morning class)
 - P(earned a C)
 - P(earned an A and was in the afternoon class)
 - P(earned an A given the student was in the morning class)

	A	B	C	Total
Morning Class	8	18	13	39
Afternoon Class	10	4	12	26
Total	18	22	25	65

2. What is the probability that a student who earned a B was in the afternoon class? What's the probability that a student in the afternoon class earned a B? Explain the difference between these two quantities.
3. The contingency table below shows the number of credit cards owned by a group of individuals below the age of 35 and above the age of 35.

	Zero	One	Two or more	Total
Between the ages of 18-35	9	5	19	33
Over age 35	18	10	20	48
Total	27	15	39	81

If one person was chosen at random, find each probability:

- a. $P(\text{had no credit cards})$
 - b. $P(\text{had one credit card})$
 - c. $P(\text{had zero and is over 35})$
 - d. $P(\text{had zero credit cards given that the person under 35})$
4. After reviewing the data, it was decided that more detail should be shown. The category of 35+ is divided into two groups and "Two or more" is turned into "Two" and "3 or more." Fill out the missing information.

	Zero	One	Two	Three or More	Total
Between the ages of 18-35	9	5		9	33
35-65	10	6	10		27
Over 65				8	
Total	27	15			81

5. Fill out the missing values in the contingency table below. This table will also be used for question 6.

	Heads	Tails	Total
Coin A	60		100
Coin B		2	2
Coin C	6,000		10,000
Total			

6. Using the table above to answer the following:
 - a. Find the percentage of heads for coins A, B and C.
 - b. If you knew one coin was weighted, which coin would you most suspect and why?

Section 4.2 Theoretical Probability

As we saw in the last section, the **probability** of a specified event is the chance or likelihood that it will occur. We calculated empirical or experimental probabilities using contingency tables. In this section, we will focus on **theoretical** probability and compare the two types.

Basic Probability Concepts

Let's begin with a brief introduction to some of the language and basic concepts of theoretical probability.

Experiment

If you roll a die, pick a card from a deck of playing cards, or randomly select a person and observe their hair color, you are conducting an **experiment**.

Events and Outcomes

The result of an experiment is an **outcome**, and a particular outcome, like rolling a five on a die, is called an **event**. An event can be a **simple event** or combination of outcomes, called a **compound event**.

Sample Space

The **sample space** is the set of all possible outcomes. For example, if we roll a six-sided die, the sample space S is the set $S = \{1, 2, 3, 4, 5, 6\}$.

Example 1: If we roll an eight-sided die, describe the sample space and give at least two examples of simple events and compound events.

The sample space is the set of all possible outcomes, or equivalently, all simple events: $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$

Examples of simple events are rolling a 1, rolling a 5, rolling a 6, and so on.

Examples of compound events include rolling an even number, rolling a 5 or a 3, and rolling a number that is at least 4.

Equally Likely Outcomes

When the outcomes of an experiment are equally likely, we can calculate the probability of an event as the number of ways it can happen out of the total number of outcomes.

Theoretical Probability

$$P(E) = \frac{\text{number of outcomes corresponding to the event } E}{\text{total number of equally-likely outcomes}}$$

We can write the result as a simplified fraction or as a decimal or percent.

Example 2: Write the sample space for the sum of two six-sided dice and determine whether the outcomes are equally likely.

The sample space for the sum of two six-sided dice is $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. The different sums, however, are not equally likely. If we look at a table of the different possible outcomes when rolling two dice, we see that there are 36 possible combinations. We will summarize this in a table by listing each outcome in the sample space. Then we find the probabilities by counting the number of ways each sum can occur and dividing it by 36.

	1	2	3	4	5	6	
1	1+1 = 2	1+2 = 3	1+3 = 4	1+4 = 5	1+5 = 6	1+6 = 7	Sum
2	2+1 = 3	2+2 = 4	2+3 = 5	2+4 = 6	2+5 = 7	2+6 = 8	Probability
3	3+1 = 4	3+2 = 5	3+3 = 6	3+4 = 7	3+5 = 8	3+6 = 9	2
4	4+1 = 5	4+2 = 6	4+3 = 7	4+4 = 8	4+5 = 9	4+6 = 10	1/36
5	5+1 = 6	5+2 = 7	5+3 = 8	5+4 = 9	5+5 = 10	5+6 = 11	3
6	6+1 = 7	6+2 = 8	6+3 = 9	6+4 = 10	6+5 = 11	6+6 = 12	2/36
							4
							3/36
							5
							4/36
							6
							5/36
							7
							6/36
							8
							5/36
							9
							4/36
							10
							3/36
							11
							2/36
							12
							1/36

From the probability table we can see that rolling a sum of 7 has the highest probability and rolling a 2 or a 12 have the lowest probabilities.

Example 3: Suppose we roll a fair six-sided die. Calculate the probability of:

- rolling a 6.
- rolling a number that is at least 4.
- rolling an even number.
- rolling a 5 or a 3.

Recall that the sample space is $S = \{1, 2, 3, 4, 5, 6\}$. Since each of the outcomes in the sample space is equally likely, we can find the probability of each event by counting the number of outcomes corresponding to the event and dividing by 6, the total number of equally likely outcomes.

- There is only one way to roll a 6, so

$$P(\text{rolling a 6}) = \frac{1}{6} \text{ or } 16.7\%$$

There is a 16.7% chance of rolling a 6.

- b. In probability we will often come across the phrases “at least” and “at most.” **At least** means that value or greater. **At most** means that value or less. Since we are looking for the probability of rolling a number that is at least 4, we need the number of outcomes that are 4 or greater. There are 3 values that meet this condition: 4, 5, and 6. The probability is

$$P(\text{Rolling a number that is at least 4}) = \frac{3}{6} = \frac{1}{2} \text{ or } 50\%$$

- c. Half of the numbers on a die are even, so we calculate:

$$P(\text{rolling an even number}) = \frac{3}{6} = \frac{1}{2} \text{ or } 50\%$$

- d. There are two ways to roll a 5 or a 3, so

$$P(\text{Rolling a 5 or a 3}) = \frac{2}{6} = \frac{1}{3} \text{ or } 33.3\%$$

There is a 33.3% chance of rolling a 5 or a 3.

Example 4: Suppose you have a bag containing 14 sweet cherries and 6 sour cherries. If you pick a cherry at random, what is the probability it will be sweet?

Each of the cherries are equally likely to be selected since our selection is random and we can assume there is no way to distinguish one cherry from another. This means that the probability of selecting a sweet cherry will be equal to the number of sweet cherries in the bag divided by the total number of cherries in the bag. Since there are 14 sweet cherries and a total of 20 cherries in the bag, we have:

$$P(\text{sweet}) = \frac{14}{20} = \frac{7}{10} \text{ or } 70\%$$

There is a 70% chance of selecting a sweet cherry from the bag.

Certain and Impossible Events

A probability is always a value between 0 and 1, or from 0% to 100%. If the probability of an event is 0 there are no outcomes that correspond with that event and we say it is **impossible**. If the probability of an event is 1 then every outcome corresponds to that event and we say it is **certain**.

Example 4:

- a. What is the probability of rolling an odd or even number on a six-sided die? Since all the numbers are either even or odd, this event includes all of the outcomes in the sample space. This event is certain.

$$P(\text{odd or even}) = \frac{6}{6} = 1.$$

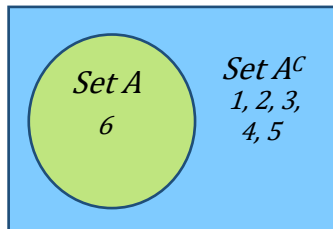
- b. What is the probability of rolling an 8 on a six-sided die? Since 8 is not one of the outcomes in the sample space, the event is impossible.

$$P(\text{roll an 8}) = \frac{0}{6} = 0.$$

Complementary Events

Just as we saw in the logic chapter, the **complement** of an event A means A does not happen. We can refer to the complement as *not A* or A^c . For example, consider the experiment of rolling a six-sided die and the simple event $A =$ rolling a 6. The complement of event A is everything in the sample space that is not a 6: $A^c = \{1, 2, 3, 4, 5\}$. Recall that we can illustrate the complement using a Venn diagram as shown below.

Notice that the outcomes from set A and the outcomes from set A^c will together equal the universal set, which is the sample space in probability. The probabilities must add up to 1 or 100%. Therefore, we can use subtraction to find the probability of a complement.



Complement of an Event

$$P(A^c) = 1 - P(A)$$

Example 5: If you roll an eight-sided die, what's the probability you don't get a 6?

Not rolling a 6 is the complement of rolling a 6, which is easier to calculate. Since there are 8 possible numbers to roll, we have:

$$\begin{aligned} P(\text{not rolling a 6}) &= 1 - P(\text{rolling a 6}) \\ &= 1 - \frac{1}{8} \\ &= \frac{7}{8} \text{ or } 0.875 \end{aligned}$$

There is an 87.5% chance of not rolling a 6.

Experimental vs. Theoretical Probability

Now that we have calculated experimental and theoretical probabilities, we can compare them. When we flip a coin, we say there is a 50% chance of getting heads. This is a theoretical probability because there are two equally likely outcomes – heads and tails – so we expect to get heads half of the time. But if you flip a coin, say 100 times, will you get heads exactly 50 times? Maybe, but you are more likely to get some number around 50 times. The number of heads you actually observe out of the total number of times you flip the coin is the experimental probability.

Example 6: The table shows the numbers that came up after rolling a six-sided die 10 times. What is the experimental probability of rolling a 6? What is the theoretical probability of rolling a 6?

Roll	1	2	3	4	5	6	7	8	9	10
Outcome	3	1	4	6	6	6	1	3	5	1

To find the experimental probability of rolling a 6, it would be helpful to change this into a frequency table. We list all the possible outcomes and count how many times each occurred.

According to our frequency table, we see that a 6 was rolled three times, so the experimental probability of rolling a 6 is

$$P(\text{roll } 6) = \frac{3}{10} \text{ or } 30\%.$$

Outcome	Frequency
1	3
2	0
3	2
4	1
5	1
6	3

Theoretically, however, we would expect the number 6 to come up 1 out of 6 times since there are 6 equally likely outcomes. Thus, the theoretical probability of rolling a 6 is

$$P(\text{roll } 6) = \frac{1}{6} \text{ or } 16.7\%$$

The Law of Large Numbers

As we saw in the previous example, theoretical and experimental probabilities are not necessarily equal. However, experimental probability will eventually approach theoretical probability as we conduct more and more trials. This phenomenon is called the **Law of Large Numbers**. This means if you flip a coin a small number of times, the experimental probability is likely to be different each time and could be very different from the theoretical probability. But if you flip a coin a large number of times, the experimental probability becomes very close to the theoretical probability of 50%.

The Law of Large Numbers is extremely powerful in that it allows us to approximate the theoretical probability of complex events – like changes in beliefs and opinions, likelihood of natural disasters, climate change effects – through repeated sampling and simulation.

Probability of Compound Events

Now that we have the basics in place, let's look at some compound probability problems that we will be studying in this course.

“And” Probabilities

As we saw with truth tables, the event ***A and B*** refers to an event where both *A* and *B* occur. These events may occur at the same time or they could happen in a sequence such as ***A and then B***. How we calculate the theoretical probability of the event ***A and B*** (or ***A and then B***) depends on whether the two events are independent or dependent.

Independent and Dependent Events

Two events *A* and *B* are **independent** if the probability of *B* occurring is the same whether or not *A* occurs. If the probability of *B* is affected by the occurrence of *A*, then we say that the events are **dependent**.

Coin flips and die rolls are common examples of independent events – flipping heads does not change the probability of flipping heads the next time, nor does rolling a six change the probability that the next roll will be a six.

Another type of event is a selection event, such as randomly selecting or drawing items from a bag, etc. These are also independent if we **draw with replacement**. By replacing the item, we reset the probability back to what it was before we made the selection. Since the probability of each selection is the same as the first selection, the events are independent.

If we draw **without replacement**, however, like selecting multiple people for a committee, we change the total number of possible outcomes, thereby changing the probability of subsequent selections. Therefore, if we draw without replacement, the events will be **dependent**.

Example 7: Determine whether the following events are independent or dependent.

- Flipping a coin twice and getting heads both times.
- Selecting a president and then a vice president at random from a pool of five equally qualified individuals.
- The event that it will rain in Portland tomorrow and the event that it will rain in Beaverton tomorrow.
- Wearing your lucky socks and getting an A on your exam.

Here are the answers:

- The probability of getting heads on the first flip is 0.5 or 50%. After flipping heads, the probability of getting heads on the second flip is still 0.5 or 50%. Since the probability of flipping heads on the second flip did not change because we flipped heads on the first flip, the events are **independent**.
- Since two different people will be put in the role of president and vice president, we are drawing without replacement and the events are therefore **dependent**.
- If it is raining in Portland it is more likely that it will rain in Beaverton, so the events are **dependent**.
- Although there may be some sort of placebo effect at play in terms of confidence and persistence, the socks you wear do not have a direct effect on how well you do on your exam, so these events are **independent**.

To calculate *and* probabilities we multiply, but we need to determine whether the events are independent or dependent. If they are independent, then we can multiply the individual probability of each event because one does not affect the other. If the events are dependent, then we need to multiply by the conditional probability based on what has previously happened. Here is a summary of this.

“And” Probabilities

If events A and B are independent, then $P(A \text{ and } B) = P(A) \cdot P(B)$

If events A and B are dependent, then $P(A \text{ and } B) = P(A) \cdot P(B \text{ given } A)$

The probability of B given A is called a **conditional probability** since it depends, or is conditional, on A occurring. We saw examples of conditional probability when we looked at contingency tables in the previous section.

Example 8: Suppose you have a bag containing 6 red Legos, 4 green Legos, and 3 black Legos. What is the probability of selecting

- two red Legos in a row if we put the first red Lego back in the bag?
- two red Legos in a row if we don't put the first Lego back in the bag?
- a red Lego and then a green Lego if we do not put the red Lego back in the bag?

Here are the solutions:

- Since the outcomes are equally likely, the probability of selecting a red Lego is the number of red Legos divided by the total number of Legos, or

$$P(\text{red}) = \frac{6}{13}.$$

If we replace the red Lego we selected (selections are independent), we go back to having 6 red Legos in the bag of 13 Legos total. Therefore,

$$P(\text{red and then red}) = P(\text{red}) \cdot P(\text{red})$$

$$= \frac{6}{13} \cdot \frac{6}{13}$$

$$= 0.213 \text{ or } 21.3\%$$

- If we do not replace the first red Lego (selections are dependent), then on our second draw there will only be 5 red Legos remaining, and 12 Legos in total. Therefore,

$$P(\text{red and then red}) = P(\text{red}) \cdot P(\text{red given red taken out})$$

$$= \frac{6}{13} \cdot \frac{5}{12}$$

$$= 0.192 \text{ or } 19.2\%$$

- The probability of selecting a red Lego on the first draw is the same as in parts a and b. Since we are not putting the red Lego back into the bag, we will have only 12 Lego left in total, of which 4 are green. Therefore,

$$P(\text{red and then green}) = P(\text{red}) \cdot P(\text{green given red taken out})$$

$$= \frac{6}{13} \cdot \frac{4}{12}$$

$$= \frac{6}{13} \cdot \frac{1}{3}$$

$$= 0.154 \text{ or } 15.4\%$$

Let's look at an example where we repeat an event many times.

Example 9: Suppose there is a 6% chance you will receive a citation if you ride the MAX train without a ticket. What is the probability that you get away without a single citation if you ride without purchasing a ticket for 20 days this month?

The first thing we want to recognize is that this question is essentially asking for the probability of no citation and no citation and no citation.... twenty times (one for each day you ride without buying a ticket). Since the outcomes are connected by an “and”, we know we will be multiplying the probabilities. In this case the the outcomes are independent (you are not more or less likely to get a citation if you already received a citation). Therefore,

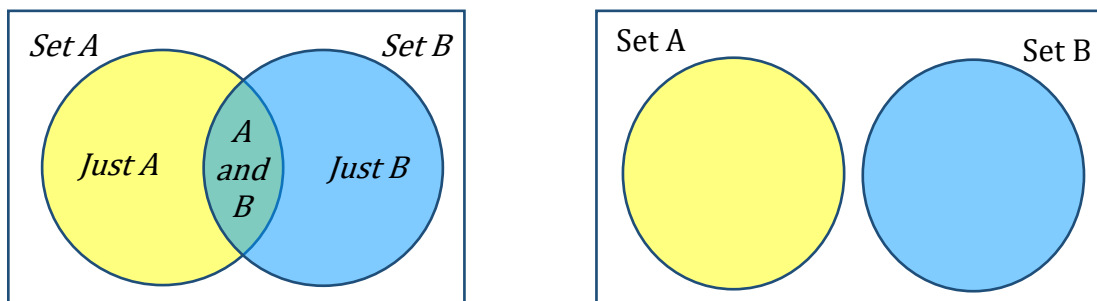
$$\begin{aligned} P(\text{no citation in 20 rides}) &= P(\text{no citation on a single ride})^{20} \\ &= (1 - 0.06)^{20} \\ &= (0.96)^{20} \\ &= 0.442 \text{ or } 44.2\% \end{aligned}$$

“Or” Probabilities

The event ***A or B*** refers to an event that includes the outcomes of *A* or *B* or *both*. We have seen the inclusive *or* both in terms of sets and logic, and in terms of contingency tables. The way we calculate the probability of *A or B* depends on whether the events have characteristics that are overlapping or disjoint.

Overlapping or Disjoint Sets

Recall that **disjoint** means the same thing as **not overlapping**. Just like we saw in the logic and sets chapter, the set diagram on the left shows overlapping sets and the set diagram on the right shows disjoint sets.



To apply this to probability, we will look at an example of events that have overlapping characteristics, such as color and shape.

Example 10: A prize machine is filled with 10 yellow erasers, 6 green erasers, 4 red pencil sharpeners, 8 yellow pencil sharpeners, and 5 red bouncy balls. Each prize is inside a plastic sphere, and the spheres are well mixed in the prize machine. Each game will get you just one prize. Determine the probability of

- getting a yellow prize.
- getting a red or yellow prize.
- getting a prize that is yellow or an eraser.

Here are the answers:

- a. Since yellow is a single event, we just need to know how many prizes there are in total, and how many of the prizes are yellow. The yellow prizes include the 10 yellow erasers and the 8 yellow pencil sharpeners.

$$P(\text{yellow}) = \frac{18}{33}.$$

- b. For a red or yellow prize, the set of red and the set of yellow do not overlap. They are disjoint sets, so we will add the probability of getting a red prize to the probability of getting a yellow prize.

$$P(\text{red or yellow}) = P(\text{red}) + P(\text{yellow})$$

$$\begin{aligned} &= \frac{9}{33} + \frac{18}{33} \\ &= \frac{27}{33} \end{aligned}$$

- c. To find the probability of getting a prize that is yellow or an eraser, we need to be careful because these are overlapping sets. There are two ways to calculate this, and it is a lot like what we did with contingency tables. The first way is to add all the items separately, being careful not to double count.

$$P(\text{yellow or eraser}) = P(\text{yellow eraser}) + P(\text{yellow pencil sharpener}) + P(\text{green eraser})$$

$$\begin{aligned} &= \frac{10}{33} + \frac{8}{33} + \frac{6}{33} \\ &= \frac{24}{33} \end{aligned}$$

The second way is to count the total of yellow items and the total of erasers, but the yellow erasers are in both sets, or the overlap. We would be counting them twice and so we subtract their probability.

$$P(\text{yellow or eraser}) = P(\text{yellow}) + P(\text{eraser}) - P(\text{yellow and eraser})$$

$$\begin{aligned} &= \frac{18}{33} + \frac{16}{33} - \frac{10}{33} \\ &= \frac{24}{33} \end{aligned}$$

Here is a summary of how we found the *or* probabilities.

***“Or”* Probabilities**

If the sets are disjoint, $P(A \text{ or } B) = P(A) + P(B)$

If the sets are overlapping, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

We could also use the overlapping formula as a general formula, because in the case of disjoint sets, there is no intersection and $P(A \text{ and } B) = 0$. Here is another example with overlapping events.

Example 11: What is the probability of rolling two dice and getting a pair or a sum of 6?

For complicated events it's a good idea to list all of the outcomes. Looking at the table of outcomes, we see that there are 6 outcomes that are pairs out of the 36 possible outcomes, and there are 5 outcomes that add to 6. We can also see that there is one outcome that is both a pair and a sum of 6, so the events are overlapping.

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

As in the last example, there are two ways to do this.

If we add all of the shaded squares without double counting, we get:

$$P(\text{pair or sum of 6}) = P(\text{pair}) + P(\text{sum of 6 that haven't been counted})$$

$$\begin{aligned} &= \frac{6}{36} + \frac{4}{36} \\ &= \frac{10}{36} \\ &= \frac{5}{18} \end{aligned}$$

To use the subtraction method, we need to add the probability of rolling a pair to the probability of rolling a sum of 6 and subtract the overlap. Thus we have:

$$P(\text{pair or sum of 6}) = P(\text{pair}) + P(\text{sum of 6}) - P(\text{pair and a sum of 6})$$

$$\begin{aligned} &= \frac{6}{36} + \frac{5}{36} - \frac{1}{36} \\ &= \frac{10}{36} \\ &= \frac{5}{18} \end{aligned}$$

Now that we have looked at empirical and theoretical probability, we will be able to use them for something very important in the next section – expected value.

Exercises 4.2

1. A ball is drawn randomly from a jar containing 6 red marbles, 2 white marbles, and 5 yellow marbles. Find the probability of:
 - a. Drawing a white marble.
 - b. Drawing a red marble.
 - c. Drawing a green marble.
 - d. Drawing two yellow marbles if you draw with replacement.
 - e. Drawing first a red marble then a white marble if marbles are drawn without replacement.
2. Compute the probability of tossing a six-sided die and getting
 - a. an even number.
 - b. a number less than 3.
3. Compute the probability of rolling a 12-sided die and getting
 - a. a number other than 8.
 - b. a 2 or 7.
4. A six-sided die is rolled twice. What is the probability of getting
 - a. a 6 on both rolls?
 - b. a 5 on the first roll and an even number on the second roll?
5. Suppose that 21% of people own dogs. If you pick two people at random, what is the probability that neither own a dog?
6. At some random moment, you look at your clock and note the minutes reading.
 - a. What is probability the minutes reading is 15?
 - b. What is the probability the minutes reading is 15 or less?
7. What is the probability of flipping a coin three times
 - a. and getting a head each time?
 - b. not getting a head at all?
8. What is the probability of rolling two six-sided dice
 - a. and getting a sum greater than or equal to 7?
 - b. getting an even sum or a sum greater than 7?

9. A box contains four black pieces of cloth, two striped pieces, and six dotted pieces. A piece is selected randomly and then placed back in the box. A second piece is selected randomly. What is the probability that
- both pieces are dotted?
 - the first piece is black and the second piece is dotted?
 - one piece is black and one piece is striped?

Section 4.3 Expected Value

Expected value is one of the useful probability concept we will discuss. It has many applications, from insurance policies to making financial decisions, and it's one thing that the casinos and government agencies that run gambling operations and lotteries may hope most people never learn about.

Expected Value

The **expected value** is the average gain or loss of an event if the procedure is repeated many times. To help get a better understanding of what expected value is and how it is used, consider the following scenario:

You are commissioned to design a game for a local carnival. Your proposed game will have players roll a six-sided die. If it comes up 6, they win \$10. If not, they get to roll again. If they get a 6 on the second roll, then they win \$3. If they do not get a 6 on the second roll, they lose. With the game design complete, you now need to decide how much the carnival game owner should charge players in order to make a profit over the long run.

To make a profit, the carnival needs to know how much they will pay in winnings, on average, over the long run and charge more than that. In other words, they must charge more than the expected value of the game.

One way to organize the outcomes and probabilities is with a probability model. A **probability model** or **probability distribution** is a table listing the possible outcomes and their corresponding probabilities. The outcomes will be the amounts a player can win, and we will calculate the probabilities using what we have learned about theoretical probability.

As we have seen with complements, probabilities in a probability distribution must add to 1, so that is important to check. Here is the probability model for the carnival game:

Outcome (\$ won)	Rolling Event	Probability
\$10	Roll a 6 on the first roll	$P(\text{roll a 6}) = \frac{1}{6}$
\$3	Roll not a 6 on the first roll and a 6 on the second roll	$P(\text{roll not a 6 then roll a 6}) = \frac{5}{6} \cdot \frac{1}{6} = \frac{5}{36}$
\$0	Roll not a 6 on the first roll and not a 6 on the second roll	$P(\text{roll not a 6 then not a 6}) = \frac{5}{6} \cdot \frac{5}{6} = \frac{25}{36}$

Think of the expected value as a weighted average. We could take the average of \$10, \$3, and \$0, but they are not all equally likely. It is much more likely to win \$0 than to win \$10. So, to find the average, we multiply each outcome by the chance it will happen and add the products together.

Expected Value

Multiply each outcome by its probability and add up the products

In this case we have:

$$\begin{aligned}\text{Expected Winnings} &= \$10\left(\frac{1}{6}\right) + \$3\left(\frac{5}{36}\right) + \$0\left(\frac{25}{36}\right) \\ &= \$2.08\end{aligned}$$

This tells us that over the long run, players can expect to win \$2.08 per game. This also means that the carnival owner will be paying out an average of \$2.08 per game! Since the carnival owner would rather not lose money by paying players over the long run we need to make sure to charge players enough to offset the average payout.

If the carnival owner charges exactly \$2.08 to play, the game is considered a **fair game** since the expected winnings would be \$0. In a fair game, the player isn't expected to win anything, nor is the owner expected to earn anything over the long run. However, if the carnival owner charges the player more than \$2.08 to play, they will earn money over the long run.

Suppose you suggest charging \$5 to play. We can determine the **net** winnings by subtracting the \$5 the player has to pay from their expected winnings. This gives us:

$$\begin{aligned}\text{Net player winnings} &= \$2.08 - \$5.00 \\ &= \$ - 2.92\end{aligned}$$

This means that over the long run, players can expect to lose an average of \$2.92 each game they play, and the carnival owner can expect to earn an average of \$2.92 per game over the long run. Here's another example.

Example 1: Pick4 is a game by the Oregon Lottery that costs \$1 to play. In this game you pick 4 numbers in a specific pattern. If you get the exact sequence, you can in theory earn a lot of money. Suppose that the payouts are as follows. Determine the player's expected net winnings.

This table is not quite a complete probability distribution since it is missing one important outcome: when the player loses. In that case the prize is \$0. We need to add a line for this. The prize for this missing outcome is \$0, and since losing is the complement to winning *something*, the probability will be:

Prize (\$)	Probability
\$250	1/417
\$500	1/1833
\$1,000	1/1667
\$1,500	1/2500

$$\begin{aligned}
 P(\text{win } \$0) &= 1 - \left(\frac{1}{417} + \frac{1}{1833} + \frac{1}{1667} + \frac{1}{2500} \right) \\
 &= 1 - 0.0039 \\
 &= 0.9961
 \end{aligned}$$

Adding this information to the table gives a complete probability distribution. Now we can see that players are going to lose more than 99% of the time, so the expected value will be heavily weighted toward winning \$0.

Prize (\$)	Probability
\$250	1/417
\$500	1/1833
\$1,000	1/1667
\$1,500	1/2500
\$0	0.9961

$$\begin{aligned}
 \text{Expected Winnings} &= 250 \left(\frac{1}{417} \right) + 500 \left(\frac{1}{1833} \right) + 1000 \left(\frac{1}{1667} \right) + 1500 \left(\frac{1}{2500} \right) + 0(0.9961) \\
 &= \$2.07
 \end{aligned}$$

Therefore, the player's expected winnings are \$2.07, on average, over the long run. To find the expected **net** winnings, we subtract the cost to play. Since it costs \$1 to play,

$$\begin{aligned}
 \text{Net Expected Winnings} &= \$2.07 - \$1.00 \\
 &= \$1.07
 \end{aligned}$$

Assuming the given payouts are correct, this would be one game you would want to play for investment purposes since you can expect to earn \$1.07 per game, on average, over the long run. Play a million times, and you just might become a millionaire!

In general, if the expected value of a game is negative, it is not a good idea to play, since in the long run you will lose money. It would be better to play a game with a positive expected value (good luck trying to find one!), although keep in mind that even if the average winnings are positive it could be the case that most people lose money and one very fortunate individual wins a great deal of money.

Not surprisingly, the expected value for casino games is always negative for the player, and therefore positive for the casino. It must be positive for the casino, or they would go out of business! Players just need to keep in mind that when they play a game repeatedly, they should expect to lose money. That is fine so long as you enjoy playing the game and think it is worth the cost, but it would be wrong to expect to come out ahead.

Expected value is not only used to determine the average amount won and lost at casinos and carnivals, it also has applications in business and insurance, just to name a few. Let's look at a couple of those applications.

Example 2: For 3 months, a coffee shop tracked their morning sales of coffee, between 6am and 10am. The following results were recorded:

Number of cups sold	145	150	155	160	170
Probability	0.15	0.22	0.37	0.19	0.07

How many cups of coffee should they expect to sell each morning?

In this case the table tells us that 15% of the time they sell 145 cups of coffee between 6am and 10am, 22% of the time they sell 150 cups, 37% of the time they sell 155 cups, 19% of the time they sell 160 cups, and 7% of the time they sell 170 cups. Since the highest probability is associated with 155 cups, the expected value should lie somewhat close to this.

To find the expected number of coffees sold, we multiply each number of cups of coffee by its respective probability and then add the products.

$$\begin{aligned} \text{Expected Number of Coffees Sold} &= 145(0.15) + 150(0.22) + 155(0.37) + 160(0.19) + 170(0.07) \\ &= 154.4 \text{ cups of coffee} \end{aligned}$$

This means that over the long run, the coffee shop can expect, on average, to sell around 154 cups of coffee each morning. This is an important tool for businesses since it helps inform them how much stock they should keep on hand.

Example 3: On average, a 40-year-old man in the US has a 0.242% chance of dying in the next year². An insurance company charges \$275 annually for a life insurance policy that pays a \$100,000 death benefit. What is the expected value for the insurance company on this policy?

The first thing we want to do is organize the probabilities and outcomes in a probability distribution table. There are two outcomes – either the policy holder dies, and the insurance company pays the benefit, or the policy holder does not die, and they do not pay anything.

The probability of paying the death benefit is equal to the chance of the person dying in the next year, and the probability of paying nothing is equal to the complement of the chance of dying in the next year.

Insurance Payout	Probability
\$100,000	0.00242
\$0	$1 - 0.00242 = 0.99758$

² According to the estimator at <http://www.numericalexample.com/index.php?view=article&id=91>

Then we can calculate:

$$\begin{aligned}\text{Expected Payout} &= \$100,000(0.00242) + \$0(0.99758) \\ &= \$242\end{aligned}$$

So, the expected payout for the insurance company is \$242, but they are charging \$275 for the policy. Their net revenue would be

$$\begin{aligned}\text{Net Value to Insurance Company} &= \$275 - \$242 \\ &= \$33\end{aligned}$$

The insurance company is making, on average, \$33 per policy per year. It shouldn't be too surprising because there are overhead costs and the insurance company can only afford to offer policies if they, on average, make money on them. But how much money should they make? As a consumer it is important to know about expected value.

Exercises 4.3

1. A bag contains 3 gold marbles, 6 silver marbles, and 28 black marbles. Someone offers to play this game: You randomly select on marble from the bag. If it is gold, you win \$3. If it is silver, you win \$2. If it is black, you lose \$1.
 - a. Make a probability model for this game.
 - b. What is your expected value if you play this game?
2. A friend devises a game that is played by rolling a single six-sided die once. If you roll a 6, he pays you \$3; if you roll a 5, he pays you nothing; if you roll a number less than 5, you pay him \$1.
 - a. Make a probability model for this game.
 - b. Compute the expected value for this game.
 - c. Should you play this game?
3. A company wants to offer a 2-year extended warranty in case their product fails after the original warranty period but within 2 years of the purchase. They estimate that 0.7% of their products will fail during that time, and it will cost them \$350 to replace a failed product. If they charge \$48 for the extended warranty, what is the company's expected profit or loss on each warranty sold?
4. An insurance company estimates the probability of an earthquake in the next year to be 0.0013. The average damage done by an earthquake it estimates to be \$60,000. If the company offers earthquake insurance for \$100, what is their expected value of the policy?
5. You purchase a raffle ticket to help out a charity. The raffle ticket costs \$5. The charity is selling 2000 tickets. One of them will be drawn and the person holding the ticket will be given a prize worth \$4000. Compute the expected value for this raffle.

6. At the local fair there is a game in which folks are betting where a chicken will poop on a 5 by 5-foot grid. (There are 25, 1 by 1 squares to choose from) You can buy a 1 by 1-foot square for \$10 and if the chicken poops on your square you win \$100. Find the expected value for this game.
7. Create a problem using the concept of expected value. Possible topics include insurance policies, financial decisions, gambling and lotteries. Determine the expected value of the situation you created.