Lesson 4: Calculating Conditional Probabilities and Evaluating Independence Using Two-Way Tables

Student Outcomes

- Students use a hypothetical 1000 two-way table to calculate probabilities of events.
- Students calculate conditional probabilities given a two-way data table or using a hypothetical 1000 two-way table.
- Students use two-way tables (data tables or hypothetical 1000 two-way tables) to determine if two events are independent.
- Students interpret probabilities, including conditional probabilities, in context.

Lesson Notes

This lesson builds on students' previous work with conditional probabilities to define independent events. In previous lessons, conditional probabilities were used to investigate whether or not there is a connection between two events. This lesson formalizes this idea and introduces the concept of independence.

Classwork

Exercise 1 (5 minutes)

Consider asking students to respond to each question independently in writing and then share answers as a class. Allow multiple student responses for each question.

Exercises

In previous lessons, conditional probabilities were used to investigate whether or not there is a connection between two events. This lesson formalizes this idea and introduces the concept of independence.

1. **Several questions are posed below. Each question is about a possible connection between two events. For each question, identify the two events and indicate whether or not you think that there would be a connection. Explain your reasoning.**
   a. **Are high school students whose parents or guardians set a midnight curfew less likely to have a traffic violation than students whose parents or guardians have not set such a curfew?**

      Responses vary.

      The two events are parents set a midnight curfew, and students have a traffic violation. Anticipate that students may indicate either that students with a curfew are less likely to have a traffic violation or that there is no connection. Either answer is acceptable as long as the student explains her reasoning.

Scaffolding:

Consider focusing on only one of these examples. Also consider revising examples to be more representative of your students' experiences and interests. Supplying a visual to accompany the situations may also increase accessibility.

b. **Are left-handed people more likely than right-handed people to be interested in the arts?**

Responses vary.

The two events are that people are right-handed, and people are interested in the arts. Students may argue that there is a connection or that there is not a connection between being right-handed and interest in the arts. Look for some reasoning that indicates that if one of the descriptions is met (selecting a right-handed person), it is more (or less) likely this person is also interested in the arts. Students may indicate that they do not think that there is a connection, which is acceptable as long as they explain their reasoning.

c. **Are students who regularly listen to classical music more likely to be interested in mathematics than students who do not regularly listen to classical music?**

Responses vary.

The two events are that students regularly listen to classical music, and students are interested in mathematics. Students may argue that there is a connection or there is not a connection between classical music and an interest in mathematics. Look for some reasoning that indicates that if one of the descriptions is met, it is more likely the other description will or will not occur.

d. **Are people who play video games more than 10 hours per week more likely to select football as their favorite sport than people who do not play video games more than 10 hours per week?**

Responses vary.

The two events are people who play video games more than 10 hours per week, and people whose favorite sport is football. Students may argue that there is a connection or there is not a connection between video games and an interest in football, which is acceptable as long as they explain their reasoning.

Ask students to write down the answer to the last question—"How can conditional probabilities be used to tell if two events are independent or not independent?"—and share their conjectures with a partner.

MP.3

Two events are independent when knowing that one event has occurred does not change the likelihood that the second event has occurred. How can conditional probabilities be used to tell if two events are independent or not independent?

Exercises 2–6 (15 minutes)

A two-way frequency table is used in these exercises in order to calculate probabilities, which helps students answer the types of questions posed in Exercise 1. The conditional probabilities are used to develop an understanding of the concept of independence. The data and several of the calculations in these exercises were first introduced in Lesson 3. This is intentional as the following exercises formalize ideas developed in students' previous work.

Recall the hypothetical 1000 two-way frequency table that was used to classify students at Rufus King High School according to gender and whether or not they participated in the after-school athletic program.

Table 1

Participation in after-school athletic program (Yes or No) of males and females

	Participate in after-school athletic program	Do not participate in after-school athletic program	Total
Females	232	348	580
Males	168	252	420
Total	400	600	$1,000$

2. **For each of the following, indicate whether the probability described is one that can be calculated using the values in Table 1. Also indicate whether or not it is a conditional probability.**
   a. **The probability that a randomly selected student participates in the after-school athletic program.**

      This probability can be calculated from the values in the table. It is not a conditional probability. The probability is based on the entire school population.

   b. **The probability that a randomly selected student who is female participates in the after-school athletic program.**

      This probability can be calculated from the values in the table. It is a conditional probability because it is a probability based on only the female students at the school.

   c. **The probability that a randomly selected student who is male participates in the after-school athletic program.**

      This probability can be calculated from the values in the table. It is a conditional probability because it is based on only the male students at the school.

3. **Use Table 1 to calculate each of the probabilities described in Exercise 2.**
   a. **The probability that a randomly selected student participates in the after-school athletic program.**

      $\frac{400}{1000} = 0.400$ *(This means that 40% of all students participate in the after-school athletic program.)*

   b. **The probability that a randomly selected student who is female participates in the after-school athletic program.**

      $\frac{232}{580} = 0.400$ *(This means that 40% of female students participate in the after-school athletic program.)*

   c. **The probability that a randomly selected student who is male participates in the after-school athletic program.**

      $\frac{168}{420} = 0.400$ *(This means that 40% of male students participate in the after-school athletic program.)*

MP.2

Exercise 4 is an important question as it provides students an understanding of the meaning of independent events. Students are expected to connect the concept of conditional probabilities with the definition of independent events. This exercise also provides students an opportunity to express their reasoning abstractly and quantitatively. Students use numerical probabilities to make inferences about real-world situations. Allow students to work individually on this question for a few minutes, and then discuss as a whole group. Use this question to define two independent events.

- Two events are independent when knowing that one event has occurred does not change the likelihood that the second event has occurred. How can conditional probabilities be used to tell if two events are independent or not independent?
  - *Allow for multiple responses. Then read through the definition, which is also provided for students, so that they can use it to complete this lesson.*

4. Would your prediction of whether or not a student participates in the after-school athletic program change if you knew the gender of the student? Explain your answer.

No. Based on the conditional probabilities, the probability that a student participates in the after-school athletic program is 0.400 for both males and females.

Two events are *independent* if knowing that one event has occurred does not change the probability that the other event has occurred. For example, consider the following two events:

F: the event that a randomly selected student is female

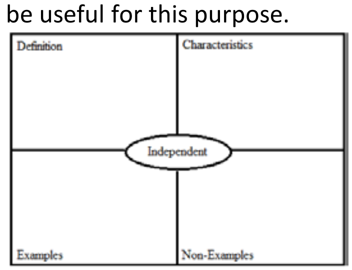
S: the event that a randomly selected student participates in after-school athletics.

F and S would be independent if the probability that a randomly selected student participates in after-school athletics is equal to the probability that a randomly selected student who is female participates in after-school athletics. If this were the case, knowing that a randomly selected student is female does not change the probability that the selected student participates in after-school athletics. Then F and S would be independent.

Exercise 5 and the definition of independence provide the opportunity to have students explore independence using the conditional probabilities based on columns. Exercise 6 is designed for students to observe that the probability that a randomly selected student who participates in the after-school athletic program is female is equal to the probability that a randomly selected student (from the entire school population) is female. The probability that a randomly selected student who participates in the after-school athletic program is male is equal to the probability that a randomly selected student (from the entire school population) is male. Point out that when this happens, then the events are independent. Also point out that conditional probabilities in either row or column can be used to decide if events are independent.

Scaffolding:

Students may have learned the term *independence* or *independent* in other contexts (such as social studies) and, as such, may need opportunities to add to their understandings. A Frayer model diagram may be useful for this purpose.



5. Based on the definition of independence, are the events randomly selected student is female and randomly selected student participates in after-school athletics independent? Explain.

Yes, they are independent because knowing that a randomly selected student is female does not change the probability that the selected student participates in the after-school athletic program. The probability that a randomly selected student participates in after-school athletics is 0.40, and the probability that a randomly selected student who is female participates is also 0.40.

6. A randomly selected student participates in the after-school athletic program.

a. What is the probability this student is a female?

$$\frac{232}{400} = 0.58$$

This is equal to the probability that a randomly selected student is female.

b. **Using only your answer from part (a), what is the probability that this student is a male? Explain how you arrived at your answer.**

The probability is 0.42.

The probability in part (a) can be interpreted as 58% of the students that participate in after-school athletics are female. The rest must be male, so the probability that a randomly selected student who participates in after-school athletics is male is $1 - 0.58 = 0.42$.

Exercises 7–11 (15 minutes)

The following exercises also provide examples of events that are *not independent*. Students may have already expressed this idea in their previous work. They now have a more formal way to express the relationship between two events.

Consider the data below.

	No household member smokes	At least one household member smokes	Total
Student indicates he or she has asthma	69	113	182
Student indicates he or she does not have asthma	473	282	755
Total	542	395	937

7. **You are asked to determine if the two events *a randomly selected student has asthma* and *a randomly selected student has a household member who smokes* are independent. What probabilities could you calculate to answer this question?**

   Students could indicate that the probability of selecting a student who has a household member who smokes is the same for students who have asthma as for those who do not have asthma. Or, students could indicate that the probability of selecting a student who has a household member who smokes from the students who have asthma would need to be equal to the probability of selecting a student who has a household member who smokes from all of the students. Or, students could indicate that the probability of selecting a student who has a household member who smokes from the students who do not have asthma would need to be equal to the probability of selecting a student who has a household member who smokes from all of the students.

   Students could also indicate that conditional probabilities based on the columns would have to be equal for the events to be independent. The probability of selecting a student who has asthma from the students who have no household member who smokes would need to be equal to the probability of selecting a student who has asthma from the students who have at least one household member who smokes.

8. **Calculate the probabilities you described in Exercise 7.**

   The row conditional probabilities described are as follows:

   $$\frac{69}{182} \approx 0.379$$

   $$\frac{473}{755} \approx 0.626$$

   The column conditional probabilities described are as follows:

   $$\frac{69}{542} \approx 0.127$$

   $$\frac{113}{395} \approx 0.286$$

9. **Based on the probabilities you calculated in Exercise 8, are these two events independent or not independent? Explain.**

   No, the conditional probabilities need to be equal for the events to be independent.

10. **Is the probability that a randomly selected student who has asthma and who has a household member who smokes the same as or different than the probability that a randomly selected student who does not have asthma but does not have a household member who smokes? Explain your answer.**

    The probabilities are different as the events are not independent. The probability of a randomly selected student who has asthma having a household member who smokes is $\frac{113}{182}$ *or approximately* 0.62. *The probability of a randomly selected student who does not have asthma having a household member who smokes is* $\frac{282}{755}$ *or approximately* 0.37.

11. **A student is selected at random. The selected student indicates that he or she has a household member who smokes. What is the probability that the selected student has asthma?**

    The probability is $\frac{113}{395}$ *or approximately* 0.29. *This is the conditional probability that a randomly selected student who has a household family member who smokes has asthma.*

Closing (5 minutes)

- Ask students to summarize the main ideas of the lesson in writing or with a neighbor. Use this as an opportunity to informally assess comprehension of the lesson. The Lesson Summary below offers some important ideas that should be included.
- It is important that students do not conclude that just because two events are not independent, there is a relationship between the events that indicates one event causes the other. Students often make this mistake. (Causation is covered in Algebra I.)
- Explain to students that when two events are not independent, it is important to remember that this does not mean that one event causes the other (***causation***). The asthma research was not a statistical experiment. There may be other possible explanations for why the events *has asthma* and *has a household member who smokes* might not be independent.
- If you know the probability that a randomly selected student from your school plans to attend a college or university after graduation, and you also know the probability that a randomly selected student from your school has a job, what would it mean for these two events to be independent?
  - *Students should indicate that knowing one event has occurred (for example that the selected student plans to attend college) does not change the probability that the second event occurred (for example that the selected student has a job).*

Lesson Summary

Data organized in a two-way frequency table can be used to calculate conditional probabilities.

Two events are independent if knowing that one event has occurred does not change the probability that the second event has occurred.

Probabilities calculated from two-way frequency tables can be used to determine if two events are independent or not independent.

Exit Ticket (5 minutes)

Name ______________________________ Date ________________

Lesson 4: Calculating Conditional Probabilities and Evaluating Independence Using Two-Way Tables

Exit Ticket

1. The following hypothetical 1000 two-way table was introduced in the previous lesson.

	Plan to vacation in New York within the next year	Do not plan to vacation in New York within the next year	Total
Watched the online ad	300	450	750
Did not watch the online ad	100	150	250
Total	400	600	1,000

Are the events *a randomly selected person watched the online ad* and *a randomly selected person plans to vacation in New York within the next year* independent or not independent? Justify your answer using probabilities calculated from information in the table.

2. A survey conducted at a local high school indicated that 30% of students have a job during the school year. If having a job and being in the 11^{th} grade are not independent, what do you know about the probability that a randomly selected student who is in the 11^{th} grade would have a job? Justify your answer.

3. Eighty percent of the dogs at a local kennel are in good health. If the events *a randomly selected dog at this kennel is in good health* and *a randomly selected dog at this kennel weighs more than* 30 *pounds* are independent, what do you know about the probability that a randomly selected dog that weighs more than 30 pounds will be in good health? Justify your answer.

Exit Ticket Sample Solutions

1. The following hypothetical 1000 two-way table was introduced in the previous lesson.

	Plan to vacation in New York within the next year	Do not plan to vacation in New York within the next year	Total
Watched the online ad	300	450	750
Did not watch the online ad	100	150	250
Total	400	600	$1,000$

Are the events *a randomly selected person watched the online ad* and *a randomly selected person plans to vacation in New York within the next year* independent or not independent? Justify your answer using probabilities calculated from information in the table.

The conditional probabilities that could be used to evaluate if the events are independent are the probability that given the selected member watched the online ad, the member plans to vacation in New York, and the probability that given the selected member did not watch the ad, the member plans to vacation in New York. Because the probabilities are equal, the events are independent.

$\frac{300}{750} = 0.40$ *and* $\frac{100}{250} = 0.40$

2. A survey conducted at a local high school indicated that 30% of students have a job during the school year. If having a job and being in the 11^{th} grade are not independent, what do you know about the probability that a randomly selected student who is in the 11^{th} grade would have a job? Justify your answer.

The probability that a student selected from the 11^{th} grade has a job would not be equal to 0.30. Not independent means that knowing the selected student is in the 11^{th} grade changes the probability that the student has a job.

3. Eighty percent of the dogs at a local kennel are in good health. If the events *a randomly selected dog at this kennel is in good health* and *a randomly selected dog at this kennel weighs more than 30 pounds* are independent, what do you know about the probability that a randomly selected dog that weighs more than 30 pounds will be in good health? Justify your answer.

As the events are independent, knowing that the selected dog weighs more than 30 pounds does not change the probability that the dog is in good health. This means that the probability that a large dog is in good health is also 0.80.

Problem Set Sample Solutions

1. Consider the following questions.

   a. A survey of the students at a Midwest high school asked the following questions:

   "Do you use a computer at least 3 times a week to complete your school work?"
   "Are you taking a mathematics class?"

   Do you think the events *a randomly selected student is taking a mathematics class* and *a randomly selected student uses a computer at least 3 times a week* are independent or not independent? Explain your reasoning.

   Anticipate students indicate that using a computer at least 3 times per week and taking a mathematics class are not independent. However, it is also acceptable for students to make a case for independence. Examine the explanation a student provides to see if they understand the meaning of independence.

b. The same survey also asked students the following:

"Do you participate in any extracurricular activities at your school?"
"Do you know what you want to do after high school?"

Do you think the events *a randomly selected student participates in extracurricular activities* and *a randomly selected student knows what he or she wants to do after completing high school* are independent or not independent? Explain your reasoning.

Answers will vary. Anticipate students indicate that students involved in extracurricular activities are often students who want to attend college. It is likely the events are not independent.

c. People attending a professional football game in 2013 completed a survey that included the following questions:

"Do you think football is too violent?"
"Is this the first time you have attended a professional football game?"

Do you think the events *a randomly selected person who completed the survey is attending a professional football game for the first time* and *a randomly selected person who completed the survey thinks football is too violent* are independent or not independent? Explain your reasoning.

Answers will vary. Anticipate that students indicate that people who attend football more often are more likely to not think the game is too violent. It is likely the events are not independent. Again, examine the explanation a student provides if they indicate the events are not independent.

2. Complete the table below in a way that would indicate the two events *uses a computer* and *is taking a mathematics class* are independent.

	Uses a computer at least 3 times a week for school work	Does not use a computer at least 3 times a week for school work	Total
In a mathematics class	420	280	700
Not in a mathematics class	180	120	300
Total	600	400	$1,000$

The values in the table were based on 0.60 of the students not in mathematics use a computer at least 3 times a week for school (0.60×300). Also, 0.60 of the students in mathematics use a computer at least 3 times a week for school (0.60×700).

3. Complete the following hypothetical 1000 table. Are the events *participates in extracurricular activities* and *know what I want to do after high school* independent or not independent? Justify your answer.

	Participate in extracurricular activities	Do not participate in extracurricular activities	Total
Know what I want to do after high school	550	250	800
Do not know what I want to do after high school	50	150	200
Total	600	400	$1,000$

The events that student participates in extracurricular activities and student knows what I want to do after high school are not independent. Students could indicate that the events are not independent in several ways. For example, $\frac{50}{200}$ (the probability that a randomly selected student who does not know what he or she wants to do after high school participates in extracurricular activities) does not equal $\frac{550}{800}$ (the probability that a randomly selected student who does know what he or she wants to do after high school participates in extracurricular activities).

4. The following hypothetical 1000 table is from Lesson 2.

	No household member smokes	At least one household member smokes	Total
Student indicates he or she has asthma	73	120	193
Student indicates he or she does not have asthma	506	301	807
Total	579	421	$1,000$

The actual data from the entire population is given in the table below.

	No household member smokes	At least one household member smokes	Total
Student indicates he or she has asthma	69	113	182
Student indicates he or she does not have asthma	473	282	755
Total	542	395	937

a. Based on the hypothetical 1000 table, what is the probability that a randomly selected student who has asthma has at least one household member who smokes?

$$\frac{120}{193} \approx 0.622$$

b. Based on the actual data, what is the probability that a randomly selected student who has asthma has at least one household member who smokes (round your answer to 3 decimal places)?

$$\frac{113}{182} \approx 0.621$$

c. Based on the hypothetical 1000 table, what is the probability that a randomly selected student who has no household member who smokes has asthma?

$$\frac{73}{579} \approx 0.126$$

d. Based on the actual data, what is the probability that a randomly selected student who has no household member who smokes has asthma?

$$\frac{69}{542} \approx 0.127$$

e. What do you notice about the probabilities calculated from the actual data and the probabilities calculated from the hypothetical 1000 table?

The conditional probabilities differ only due to rounding in constructing the hypothetical 1000 table from probability information based on the actual data. When an actual data table is available, it can be used to calculate probabilities. When only probability information is available, constructing a hypothetical 1000 table from that information and using it to compute other probabilities will give the same answers as if the actual data were available.

5. As part of the asthma research, the investigators wondered if students who have asthma are less likely to have a pet at home than students who do not have asthma. They asked the following two questions:

   "Do you have asthma?"
   "Do you have a pet at home?"

   Based on the responses to these questions, you would like to set up a two-way table that you could use to determine if the following two events are independent or not independent:

   Event 1: a randomly selected student has asthma
   Event 2: a randomly selected student has a pet at home.

   a. What would you use to label the rows of the two-way table?

      Anticipate students will indicate for the rows "Has asthma" for the first row and "Does not have asthma" for the second row. Students might use these labels for the columns rather than the rows, which is also acceptable.

   b. What would you use to label the columns of the two-way table?

      Anticipate students will indicate for the columns "Has a pet" and "Does not have a pet."

   c. What probabilities would you calculate to determine if Event 1 and Event 2 are independent?

      Answers may vary. Row conditional probabilities or column conditional probabilities would have to be equal if the events are independent. For column conditional probabilities (based on the definition of rows and columns above), this would mean that the probability that a randomly selected student who has a pet has asthma is equal to the probability that a randomly selected student who does not have a pet has asthma.