

## GRADES 9–12

Copyright © 2009 by the National Council of Teachers of Mathematics, Inc. www.nctm.org. All rights reserved. This material may not be copied or distributed electronically or in other formats without written permission from NCTM.

"All students should ... understand the



concepts of conditional probability and independent events." (NCTM 2000, p. 324)

# NAVIGATING through PROBABILITY

Chapter 3 Independence and Conditional Probabilities

A group of high school students developed the following questions for research:

- Of the student drivers at Waldo High School, are those whose parents have set a midnight curfew less likely to have received a ticket for a traffic violation than those whose parents have not set such a curfew?
- Are students who listen regularly to classical music likely to have higher grade-point averages than those who do not?
- Are male teenaged drivers more likely than female teenaged drivers to have been involved in a traffic accident?
- Do students with above-average SAT scores have a higher probability than students with below-average SAT scores of being successful in college?
- Is a child of parents who smoke more likely to have asthma than is a child of parents who do not smoke?
- Are left-handed people more likely than right-handed people to be interested in the arts?

This chapter provides a framework for organizing data and interpreting probabilities to explore these types of questions. To conduct such investigations, students need an understanding of the concept of independence.

As discussed in chapter 2, two events are *independent* if the occurrence of one of them does not change the likelihood of the occurrence of the

Recall from chapter 2 that two outcomes are independent if the occurrence of one does not change the probability of the other.

"Students can ... use the sample space to answer conditional probability questions." (NCTM 2000, p. 332) other. Whether or not two events are independent can be determined by interpreting probabilities.

For example, in the case of the first research question, consider a student driver selected at random from all the Waldo High School students with parents who have set a midnight curfew. What is the likelihood that this selected student has received a ticket for a traffic violation? Suppose that 15 percent of all the student drivers at Waldo High School have received tickets, but only 3 percent of the student drivers with curfews have received them. Then the likelihood that a selected Waldo High School student driver *with* a curfew has received a traffic ticket is not the same as the likelihood that a selected Waldo High School student driver, *with or without a curfew*, has received a traffic ticket.

The second event that we are interested in, "selecting a student driver with a ticket," is not independent of the occurrence of the first event, "selecting a student driver with a curfew." The selection of a Waldo High School student driver with a curfew affects—in this case, *decreases*—the likelihood that the student driver has received a ticket. Thus, the events are not independent.

If the events *were* independent, the proportion of Waldo High School student drivers whose parents had set curfews and who also had received tickets would be equal to the proportion of Waldo High School student drivers, with or without curfews, who had received tickets. In other words, *if* the likelihood of selecting a student driver who had a curfew and had received a ticket were *the same as* the likelihood of selecting a student driver who had received a ticket, without regard to the curfew, then the events would be independent.

Probabilities indicate whether or not events are independent. This chapter probes this idea while addressing the following three questions:

- 1. When are two events considered independent?
- 2. What is a conditional probability?
- 3. How can conditional probabilities be used to tell if two events are independent or not independent?

The chapter's investigation of independent events begins with the activity Abby's Kennels, which presents data on dogs' sizes (large or small) and their results in an obedience course (passed or did not pass). Students examine conditional probabilities as a way of determining whether the events "selecting a dog that is large [or small]" and "selecting a dog that passed [or did not pass] the course" are independent.

# Abby's Kennels

#### Goal

• Use conditional probabilities to understand when two events are independent

### Materials and Equipment

- A copy of the activity pages for each student
- A set of three mystery bags for each student or group of students working together:
  - (a) 3 small paper bags labeled "X," "Y," and "Z"
  - (b) 50 chips—30 red and 20 blue—for each bag
  - (c) 15 stars (or other stickers) for each bag

#### Discussion

This activity presents students with a scenario about dogs that have just finished an obedience course at an establishment called Abby's Kennels. Of 50 dogs enrolled in the five-session course, 30 dogs were large (as measured by weight), and 20 dogs were small. A total of 15 dogs passed the course.

The activity lets students explore ways of thinking about the probability of two events—"selecting a dog of a particular size" (large or small) and "selecting a dog with a particular course result" (passed or did not pass)—occurring in the selection of a single dog from the population of 50 dogs that were enrolled in the obedience course at Abby's Kennels. The students use simulation, relative frequencies, and conditional probabilities to investigate the probabilities of selecting a dog of a particular size with particular course results.

The simulation requires three "mystery bags" that you will need to prepare in advance for each student or small group of students who will be working together, depending on how you decide to have students do the activity. Discovering the composition of the mystery bags helps students develop the idea of independence.

To make each set of mystery bags, label three small opaque bags "X," "Y," and "Z," and put 50 chips in each—30 red chips to represent the large dogs and 20 blue chips to represent the small dogs in the course at Abby's Kennels. To complete the bags—

- place a star (or other sticker) on any 15 red chips in bag X;
- place a star (or other sticker) on 9 red chips and 6 blue chips in bag Y;
- place a star (or other sticker) on 15 blue chips in bag Z.

The stars or stickers will represent the dogs that passed the obedience course.

Whether you have your students work independently or in groups, you should have each student collect data individually, as the activity pages direct, and record the data in the appropriate tables and charts. (Having the students work in small groups will reduce the number of





mystery bags that you will need and will also facilitate the students' collection of data.)

In the activity, students select chips to simulate selecting dogs at random from the 50 dogs that completed the obedience course at Abby's Kennels. After simulating the selection of a dog, the students record its size (large or small) and its obedience course result (passed or did not pass). The scenario supplies the numbers of small and large dogs, as well as the total number of dogs that passed the obedience course. Other characteristics of the group of dogs, such as the number of large dogs that passed the course, are not known. (Table 3.1 shows the data from the scenario entered in a two-way table such as students work with in the activity.)

#### Table 3.1

Obedience Course Results for Large and Small Dogs at Abby's Kennels

	Passed the Course	Did Not Pass the Course	Total
Large dogs			30
Small dogs			20
Total	15	35	50

As students select chips from each bag, the activity guides them in using relative frequencies to make conjectures about the make-up of each bag. Students are encouraged to articulate their observations about their data, which they summarize in tables (figure 3.1 shows a sample).

-	Descriptors of Dogs Represented by Chips Selected from Bag Y				
g s g	Selection Number	Large Dog (Red Chip)	Small Dog (Blue Chip)	Dog That Passed the Obedience Course (Chip with Star)	
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				
	10				

Summaries that indicate that all the dogs that passed the obedience course are large (bag X), or that none of the dogs that passed the obedience course is large (bag Z), indicate events that are not independent. That is, for bag X, the first event "selecting a large dog" affects—in this case, *increases*—the likelihood of the second event "selecting a dog that passed the obedience course." This likelihood—15/30, or 1/2—is not the same as the likelihood of selecting a dog that passed the obedience course if the first event happened to be "selecting a small dog." For bag X, *this* first event would *decrease* the likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a small dog." For bag X, *this* first event would *decrease* the likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a small dog." For bag X, *this* first event would *decrease* the likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a dog that passed the obedience course." For bag X, *this* first event would *decrease* the likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a dog that passed the obedience course." This likelihood of the second event "selecting a dog that passed the obedience course." This likelihood the second event "selecting a dog that passed the obedience course." This likelihood the second event "selecting a dog that passed the obedience course." This likelihood the second event "

#### Fig. **3.1**

A sample activity table for summarizing descriptors of dogs represented by chips selected from a mystery bag When students investigate bag Y, they should begin to offer statements that indicate their recognition that the events involving the size of a dog and its obedience course results are independent. They should discover that the likelihood of selecting a dog that passed the obedience course is "not noticeably different" for large or small dogs. That is, the likelihood of the event "selecting a dog that passed the obedience course" does not depend on the prior occurrence of the event "selecting a large [or a small] dog."

For bag Y, the exact proportion for selecting a chip that represents a dog that passed the course, given the occurrence of the event "selecting a large dog," is 9/30, or 3/10. This proportion is equal to the proportion for selecting a chip that represents a dog that passed the course, given the occurrence of the event "selecting a small dog" (6/20, or 3/10). The equality of these proportions indicates that the event "selecting a dog that passed the course" and the event "selecting a dog of a particular size" are independent. This proportion, 3/10, is thus also equal to the proportion for selecting a dog that passed the course without regard to size: 15/50, or 3/10.

Bag Y presents students with the activity's central challenge. It is unlikely that the data that they collect from this bag will demonstrate the exact proportions for selecting a dog that is large and that has passed the obedience course. Students should, however, observe that they are obtaining similar ratios for both large and small dogs that passed or did not pass the course. Question 10 formally presents the composition of the bag to the students through a two-way table.

The activity, which began with the task of estimating probabilities from relative frequencies, concludes with a more formal consideration of conditional probabilities and independence. Conditional probability is explained with an example that calls on students to consider the likelihood of selecting a dog that passed the obedience course *given that* the dog (selected from bag Y) is large. Students are told that this probability is represented symbolically as P(A | B), where, in the example, A stands for "selecting a dog that passed the course," and B stands for "selecting a large dog."

Questions 13 and 14 ask the students to use the data that they tabulated for bag Y to determine the probability that a randomly selected dog passed the course, given the size of the dog. Deriving conditional probabilities from a two-way table provides the students with a mechanism for determining whether or not the events are independent. According to the definition of independence supplied to the students in question 12, two events A and B are considered *independent* if the probability of both A and B occurring in a single selection from a population is the product of the probability of the first event times the probability of the second event.

This relationship is represented symbolically as  $P(A \cap B) = P(A) \bullet P(B)$ . If A and B are independent events, then P(A | B) = P(A). This result can be derived by applying the multiplication principle to substitute  $P(A) \bullet P(B)$  for  $P(A \cap B)$  in the statement of conditional probabilities and then simplifying, as shown in figure 3.2.

The activity prompts students to think about what conditional probabilities tell them about the independence of events and to consider what independence means. In the data that students gather on the dogs If all the dogs that passed the obedience course are large, or if none of the dogs that passed the obedience course is large, then the events "selecting a dog that is large" and "selecting a dog that passed the course" are not independent.

If P(A) represents the probability that event A will occur and P(B)represents the probability that event B will occur, then the conditional probability, P(A|B), which denotes the probability of event A given event B, is defined in the following way:

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

Recall from chapter 2 that for independent events A and B, P(A and B) is given by the multiplication principle:  $P(A \cap B) = P(A) \cdot P(B).$ 

> Fig. **3.2.** Derivation of *P*(A | B = *P*(A) for independent events

Additional lessons that develop ideas



about compound events have been reprinted by permission (in both teacher and student editions) from Hopfensperger, Kranendonk, and Scheaffer (1999) on the accompanying CD-ROM. enrolled in obedience training at Abby's Kennels, independence would mean that the probability of selecting a dog that passed the course (event A) from the population of large dogs (given event B) would be the same as the probability of selecting a dog that passed the obedience course from the total population of dogs with no condition given about the dog's size.

If A and B are independent events, then the probability of A, given the occurrence of B, is equal to the probability of A:

$P(A   B) = \frac{P(A \cap B)}{P(B)}$	Statement of conditional probabilities
$P(A \cap B) = P(A) \bullet P(B)$	Multiplication principle for independent events
$P(\mathbf{A} \mid \mathbf{B}) = \frac{P(\mathbf{A}) \bullet P(\mathbf{B})}{P(\mathbf{B})}$	Substitution
$\frac{P(A) \bullet P(B)}{P(B)} = P(A)$	Simplification
$\therefore P(A   B) = P(A)$	Substitution

Students use the specific values organized in the two-way table for bag Y to summarize the independence of the course results and the sizes of the dogs. For example, if A, the event "selecting a dog that passed the course," and B, the event "selecting a large dog," are independent, then the multiplication principle may be used to demonstrate that  $P(A \cap B)$  equals  $P(A) \bullet P(B)$ . In addition, students may use the two-way table to demonstrate that each conditional probability is equal to the probability of selecting a dog that passed the course from the total population of 50 dogs. This information helps them answer concluding questions about how they would expect a trainer at Abby's Kennels to respond if asked whether dogs of different sizes could be expected to have different results in an obedience training course.

The next activity, Independent or Not Independent? presents the results of a high school survey on asthma and cigarette smoke. The activity gives students an opportunity to examine a population in which two events are not independent.