

## 5.1 Probability Distributions and Expected Value



**discrete random variable**—a variable that assumes a unique value for each outcome



### Think about Step C

Why are there 36 possible rolls of the two dice? Refer to the chart on page 226.

**probability distribution**—a table, formula, or graph that provides the probabilities of a discrete random variable assuming any of its possible values

Should you close down a manufacturing plant because 5% of the drills produced during a specific shift were defective? Should you follow a prescribed medical treatment if you are not sure that the test results correctly identified your illness?

Probability distributions (theoretical and experimental) and expected value are significant pieces of information that can be used in making these kinds of decisions.

### INVESTIGATION 1: DISTRIBUTION OF DICE SUMS

Consider the experiment in which two six-sided dice are rolled. Suppose one die is red and the other is green.

#### Purpose

To compute the experimental probability for the sum resulting from the roll of two six-sided dice.

#### Procedure

The sum on the dice is a *discrete random variable*. It is the result of a random event and can take on any whole number from 2 to 12.

- Create a table or spreadsheet to record the possible sums that result from rolling the dice. Roll the dice 50 times to record 50 sums.
- Draw a histogram that indicates the frequency for each sum.
- Compute the experimental probability (relative frequency) for each possible sum.

#### Discussion Questions

- Why do all the experimental probabilities from Step C add to 1?
- Does a sum of 2 have the same probability as a sum of 12? Why?
- Which sum has the greatest probability of occurring? Explain why this is so.
- If you were to roll the dice 360 times, how often would you expect a sum of 7 to occur? How often would you expect a sum of 8? Explain your answers.

### PROBABILITY DISTRIBUTION OF A DISCRETE RANDOM VARIABLE

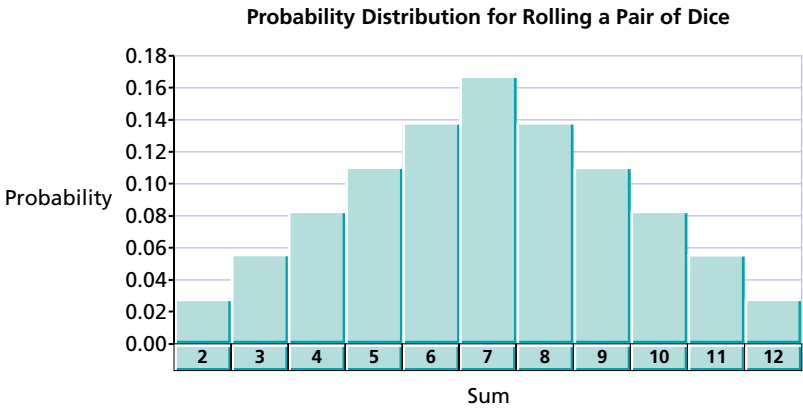
A table of probabilities for the possible sums of the dice appears in the margin on the following page. This table is one way to represent the **probability distribution** of the possible sums. After a number of trials of the experiment, the

**relative frequency distribution**—a table that shows the ratio of the number of times each value of a discrete random variable occurs to the total number of trials

Sum	Theoretical Probability
2	$\frac{1}{36}$
3	$\frac{2}{36}$
4	$\frac{3}{36}$
5	$\frac{4}{36}$
6	$\frac{5}{36}$
7	$\frac{6}{36}$
8	$\frac{5}{36}$
9	$\frac{4}{36}$
10	$\frac{3}{36}$
11	$\frac{2}{36}$
12	$\frac{1}{36}$

number of times each sum occurs is represented in a frequency distribution. The frequency distribution can be converted into a **relative frequency distribution** by dividing each frequency by the total number of trials. Each relative frequency is the observed experimental probability of a particular sum and is an estimate of the theoretical probability of that sum.

The graph below is another way to represent the probability distribution. This graph also provides the probability of each sum occurring when a pair of dice is rolled.



### Probability Distribution of a Discrete Random Variable

The probability distribution of a discrete random variable,  $X$ , is a function that provides the probability of each possible value of  $X$ . This function may be presented as a table of values, a graph, or a mathematical expression.

### ? Think about Mean and the Expected Value

The expected value of a discrete random variable is often referred to as its **mean**. Why is this an appropriate term?

### EXPECTED VALUE OF A DISCRETE RANDOM VARIABLE

An informal definition for the *expected value* of a random variable was given in the previous chapter. After many repetitions of the experiment, the average value of the random variable tends toward the expected value. A more mathematically formal definition of expected value will be introduced in this section.

By analyzing the probabilities of all possible sums when two dice are rolled, the probability distribution above was constructed.

You can see that the theoretical probability of a sum of 7 is  $\frac{6}{36} = \frac{1}{6}$ . In other words, after many trials of the experiment, you would expect to see 7 close to one-sixth of the time. Similarly, each of the other possible sum values should occur, in the long run, with a relative frequency close to its probability. The expected value of the sum is the result of adding each possible sum value multiplied by its expected relative frequency, or probability.

Therefore, the expected value for the sum of two dice is as follows:

$$\begin{aligned}
 E(\text{Sum}) &= 2P(\text{Sum} = 2) + 3P(\text{Sum} = 3) + 4P(\text{Sum} = 4) + \dots \\
 &\quad + 12P(\text{Sum} = 12) \\
 &= 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + \dots + 12 \times \frac{1}{36} \\
 &= \frac{252}{36} = 7
 \end{aligned}$$

### Expected Value of a Discrete Random Variable

The expected value of a discrete random variable,  $X$ , is the sum of the terms of the form  $X \cdot P(X)$  for all possible values of  $X$ . In other words, if  $X$  takes on the values  $x_1, x_2, \dots, x_n$ , then the expected value of  $X$  is given by

$$\begin{aligned}
 E(X) &= x_1P(X = x_1) + x_2P(X = x_2) + \dots + x_nP(X = x_n) \\
 &= \sum_{i=1}^n x_iP(X = x_i)
 \end{aligned}$$

where  $n$  represents the number of terms in the sum.

### Example 1 Tossing Three Coins

Suppose you were to toss three coins.

- What is the likelihood that you would observe at least two heads?
- What is the expected number of heads?

### Solution

The following table shows the theoretical probability distribution for this experiment. The discrete random variable,  $X$ , represents the number of heads observed.

$X$	0 heads	1 head	2 heads	3 heads
$P(X) = x$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

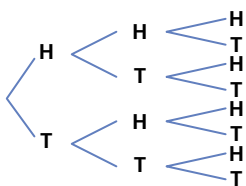
- The probability that at least two heads are observed is

$$\begin{aligned}
 P(X = 2) + P(X = 3) &= \frac{3}{8} + \frac{1}{8} \\
 &= \frac{1}{2}
 \end{aligned}$$

- The expected number of heads is

$$\begin{aligned}
 E(X) &= 0P(X = 0) + 1P(X = 1) + 2P(X = 2) + 3P(X = 3) \\
 &= 0\left(\frac{1}{8}\right) + 1\left(\frac{3}{8}\right) + 2\left(\frac{3}{8}\right) + 3\left(\frac{1}{8}\right) \\
 &= \frac{0}{8} + \frac{3}{8} + \frac{6}{8} + \frac{3}{8} \\
 &= 1\frac{1}{2}
 \end{aligned}$$

Toss 1    Toss 2    Toss 3



### ? Think about The Expected Value

- Why is  $1\frac{1}{2}$  heads in the toss of three coins not surprising?
- What does this result mean?
- Why is this expected value not a whole number?

### Example 2 Selecting a Committee

Suppose you want to select a committee consisting of three people. The group from which the committee members can be selected consists of four men and three women.

- (a) What is the probability that at least one woman is on the committee?
- (b) What is the expected number of women on the committee?

### Solution

There are  $C(7, 3)$  or 35 possible committees. Of these,  $C(4, 3)$  have no women;  $C(4, 2) \times C(3, 1)$  have one woman;  $C(4, 1) \times C(3, 2)$  have two women; and  $C(3, 3)$  have three women. The probability distribution for the number of women on the committee appears below. The discrete random variable,  $X$ , represents the number of women on the committee.

$X$	0 women	1 woman	2 women	3 women
$P(X)$	$\frac{4}{35}$	$\frac{18}{35}$	$\frac{12}{35}$	$\frac{1}{35}$

- (a) The probability that at least one woman will be on the committee is

$$\begin{aligned}P(X = 1) + P(X = 2) + P(X = 3) &= \frac{18}{35} + \frac{12}{35} + \frac{1}{35} \\&= \frac{31}{35}\end{aligned}$$

- (b) The expected number of women on the committee is

$$\begin{aligned}E(X) &= 0P(X = 0) + 1P(X = 1) + 2P(X = 2) + 3P(X = 3) \\&= 0\left(\frac{4}{35}\right) + 1\left(\frac{18}{35}\right) + 2\left(\frac{12}{35}\right) + 3\left(\frac{1}{35}\right) \\&= \frac{0}{35} + \frac{18}{35} + \frac{24}{35} + \frac{3}{35} \\&\doteq 1.3\end{aligned}$$

### KEY IDEAS

**discrete random variable**—a variable that assumes a unique value for each outcome

**probability distribution**—a function that gives the probability of each possible value of a discrete random variable,  $X$ ; this function may be presented as a table of values, a graph, or a mathematical expression

**relative frequency distribution**—a table that shows the ratio of the frequency of each value of a discrete random variable to the number of trials

**expected value**—the sum of the terms of the form  $X \cdot P(X)$  for all possible values of a discrete random variable,  $X$

$$E(X) = x_1P(X = x_1) + x_2P(X = x_2) + \dots + x_nP(X = x_n)$$

## 5.1 Exercises



- A**
- An experiment is conducted in which two coins are tossed and the number of heads is recorded.
    - Explain why the number of heads is a discrete random variable.
    - Create a probability distribution table for the number of heads.
    - Compute the expected number of heads per trial.
  - Identify which of the following are discrete random variables. Explain each answer.
    - the number of job applications received each week by a restaurant
    - the time it takes a student to complete math homework
    - the number of defective parts in a sample taken from a factory
    - the life span of a light bulb observed during a quality-control test
    - the total amount of money earned at a movie theatre during a day
  - Why is the sum of all the probabilities in a probability distribution for a discrete random variable equal to 1?
  - Knowledge and Understanding** Which of the following are valid probability distributions? Explain your answers.
 

(a)

$X$	$P(X)$
0	0.5
1	0.25
2	0.25

(b)

$X$	$P(X)$
0.5	0.2
0.2	0.3
0.3	0.25

(c)

$X$	$P(X)$
0	0.3
1	0.25
2	0.25
3	0.2
  - What is the expected value of the random variable representing the number observed on a single roll of a six-sided die? Explain why the answer is *not* an integer.
  - What is the expected value of the random variable representing the number observed on a single roll of an eight-sided die?

**B**


7. Look back on page 274 at the probability distribution for the possible sums that result from rolling two six-sided dice. Use the probability distribution to find the following probabilities.

- (a)  $P(\text{sum} \leq 5)$  (b)  $P(\text{sum is even and sum} > 8)$   
 (c)  $P(\text{sum is not } 7)$  (d)  $P(\text{sum is even and sum} \leq 8)$


8. Create a probability distribution for the results of rolling a single die with eight faces numbered 1 through 8.

9. Create a probability distribution for all the sums of two eight-sided dice.

10. A customer randomly selects two RAM modules from a shipment of six known to contain two defective modules.

- (a) Find the probability distribution for  $n$ , the number of defective modules in the purchase.  
 (b) Use labelled pieces of paper to simulate this situation. Carry out 50 trials and compare your simulation's experimental probability distribution with the theoretical distribution you obtained in part (a).  
 (c) Create a simulation using technology and compare your results for 100 or more trials with the theoretical probability distribution.  
 (d) Compute the expected number of defective RAM modules the customer would purchase.

11. **Application** A study of consumer habits indicates that 20% of all shoppers in a certain city read the unit-pricing labels on product packages before deciding which product to buy. Two customers are observed shopping in a food store in the city. Consider the discrete random variable,  $u$ , the number of shoppers who use the unit-pricing information.

- (a) Use five pieces of paper, cards, or coloured marbles to create a simulation of the situation with 30 trials. Record the experimental probability distribution for  $u$  from your results.  
 (b) Create a simulation using appropriate technology and compare your results for 1000 trials with your answer in part (a).  
 (c) Create the theoretical probability distribution for  $u$  and compare the results of your simulations to the theoretical values.  
 (d) Compute the expected number of shoppers in the sample of two customers who would use the unit-pricing information.

12. **Communication** If a construction company wins a bid for a project, it will earn \$50 000. The bid preparation will cost the company \$5000. The president feels that the company's probability of winning the bid is 0.4. The company has a policy of submitting a bid on any project for which the expected return is \$12 000 or more. Should the president submit a bid?

13. Suppose a marketing manager wishes to know which of two package designs, A or B, the general public will prefer. He decides to survey 20 randomly selected people. Assume that the probability of someone selecting design A is  $\frac{1}{2}$ . Consider the value  $a$ , the number of people in a group of 20 who select package design A.





- (a) Construct a non-technology-based simulation of the situation. Generate an experimental probability distribution for  $a$  based on several trials of your simulation. Estimate the probability that more than 15 people in a random sample of 20 would choose package design A.
- (b) Use a technology-based simulation to generate more trials and refine the estimated values for the probability distribution of  $a$ .
- (c) Determine the theoretical probability distribution for  $a$  and compare the values with the results of your simulations in parts (a) and (b).
- (d) What is the probability that, in a group of 20 randomly selected people, 15 or more will select package design A?

**C**



**Think about  
Question 14**

How would you find the expected value of the insurance company's costs on the policy?

- 14. Thinking, Inquiry, Problem Solving** An insurance company's statistical records indicate that a particular type of automobile accident has occurred nearly three times per 10 000 drivers. The company wants to break even on policies that pay out \$100 000 in the event of this particular type of accident. Find the premium that the insurance company must charge customers for insurance against this type of accident.

**ADDITIONAL ACHIEVEMENT CHART QUESTIONS**

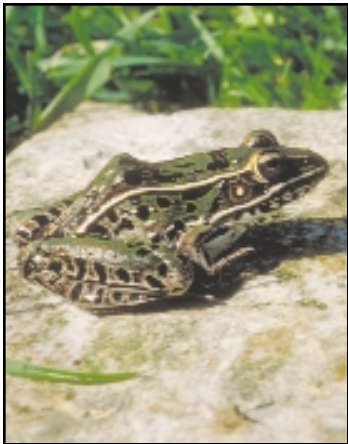
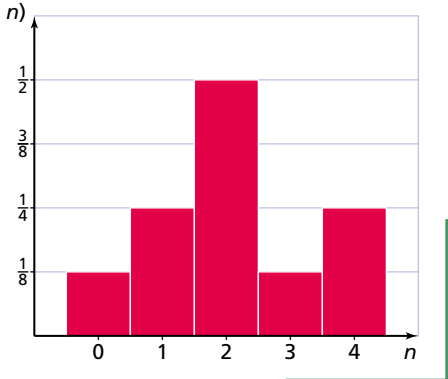
- 15. Knowledge and Understanding** The manager of a telemarketing firm conducted a time study to analyze the length of time his employees spent engaged in a typical sales-related phone call. The results are shown in the table below, where time has been rounded to the nearest minute.

Time (min)	1	2	3	4	5	6	7	8	9	10
Frequency	15	12	18	22	13	10	5	2	2	1

- (a) Define the random variable  $X$ .
- (b) Create a probability distribution for these data.
- (c) Determine the expected length of a typical sales-related call.
- 16. Application** A drawer contains four red socks and two blue socks. Three socks are drawn from the drawer without replacement.
- (a) Create a probability distribution in which the random variable represents the number of red socks.
- (b) Determine the expected number of red socks if three are drawn from the drawer without replacement.
- 17. Thinking, Inquiry, Problem Solving** Do some research to find a set of data that meets the following two conditions:
- The data are represented by a frequency table.
  - A discrete random variable can be used to represent the outcomes.
- (a) Create a probability distribution for your data set.
- (b) Use your data set to determine the expected value.



18. **Communication** The graph to the right shows the probabilities of a variable,  $N$ , for the values  $N = 0$  to  $N = 4$ . Is this the graph of a valid probability distribution? Explain.



## Chapter Problem

### Are Frog Populations Declining?

- CP1.** Imagine that a wetland you are studying has a population of 5000 frogs, including bullfrogs, spring peepers, and mink frogs. The populations of the three species are shown below.

Species	Percent of Total Marsh Population	Species Gender Ratio	
		Males	Females
Bullfrog	30%	60%	40%
Spring peeper	50%	55%	45%
Mink frog	20%	52%	48%

Determine the probability that the first two frogs captured are

- (a) bullfrogs (b) female bullfrogs  
(c) females of any species (d) frogs that are not spring peepers

- CP2.** (a) Design and describe a simulation that will allow you to construct a probability distribution for the number of bullfrogs in a sample of 30 frogs. Use it to predict the probability that the sample will have eight or more bullfrogs.  
(b) Compute the theoretical probability distribution for the number of bullfrogs in the sample, and then compare your simulation results with the predicted theoretical values.  
(c) Design and describe a simulation that will allow you to construct a probability distribution for the number of female frogs of each species in a sample of 30 frogs. Use it to predict the probability that the sample will have eight or more females.  
(d) Compute the theoretical probability distribution for the number of females in the sample, and then compare your simulation results with the predicted theoretical values.