CHAPTER 7

Probability Distributions

Specific Expectations	Section
Identify examples of discrete random variables.	7.1
Construct a discrete probability distribution function by calculating the probabilities of a discrete random variable.	7.1
Calculate expected values and interpret them within applications as averages over a large number of trials.	7.1, 7.2, 7.3, 7.4
Determine probabilities, using the binomial distribution.	7.2
Interpret probability statements, including statements about odds, from a variety of sources.	7.1, 7.2, 7.3, 7.4
Identify the advantages of using simulations in contexts.	7.1, 7.2, 7.3, 7.4
Design and carry out simulations to estimate probabilities in situations for which the calculation of the theoretical probabilities is difficult or impossible.	7.1, 7.2, 7.3, 7.4
Assess the validity of some simulation results by comparing them with the theoretical probabilities, using the probability concepts developed in the course.	7.1, 7.2, 7.3, 7.4

Chapter Problem

Collecting Cards

The Big K cereal company has randomly placed seven different collector cards of the world's great mathematicians into its Krakked Korn cereal boxes. Each box contains one card and each card is equally likely.

1. If you buy seven boxes of the cereal, what is the probability that you will get all seven different cards?
2. What is the probability that you will get seven copies of the same card?
3. Estimate the number of boxes you would have to buy to have a 50% chance of getting a complete set of the cards.
4. Design a simulation to estimate how many boxes of Krakked Korn you would have to buy to be reasonably sure of collecting an entire set of cards.

In this chapter, you will learn how to use probability distributions to calculate precise answers to questions like these.

Review of Prerequisite Skills

If you need help with any of the skills listed in purple below, refer to Appendix A.

1. Order of operations Evaluate.

a) $\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)$

b) $\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)$

c) $\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)^2$

d) $1.2\left(\frac{1}{5}\right) + 3.1\left(\frac{2}{5}\right) + 2.4\left(\frac{3}{5}\right) + 4.2\left(\frac{4}{5}\right)$

e) $0.2 + (0.8)(0.2) + (0.8)^2(0.2) + (0.8)^3(0.2)$

2. Sigma notation Write the following in sigma notation.

a) $t_1 + t_2 + \ldots + t_{12}$

b) $(0)({}_9C_0) + (1)({}_9C_1) + (2)({}_9C_2) + \ldots + (9)({}_9C_9)$

c) $\frac{2}{3} + \frac{3}{4} + \frac{4}{5} + \frac{5}{6} + \frac{6}{7} + \frac{7}{8}$

d) $\frac{a_0 + a_1 + a_2 + a_3 + a_4 + a_5}{6}$

3. Sigma notation Expand and simplify.

a) $\sum_{k=1}^{6} k^2$

b) $\sum_{m=1}^{15} b_{m-1}$

c) $\sum_{i=0}^{7} {}_7C_i$

d) $\sum_{x=0}^{8} (0.3)^x(0.7)$

4. Binomial theorem (section 5.5) Use the binomial theorem to expand and simplify.

a) $(x + y)^6$

b) $(0.4 + 0.6)^4$

c) $\left(\frac{1}{3} + \frac{2}{3}\right)^5$

d) $(p + q)^n$

5. Probability (Chapter 6) When rolling two dice,

a) what is the probability of rolling a sum of 7?

b) what is the probability of rolling a 3 and a 5?

c) what is the probability of rolling a 3 or a 5?

d) what is the probability of rolling a sum of 8?

e) what is the probability of rolling doubles?

6. Probability (Chapter 6) In a family of four children, what is the probability that all four are girls?

7. Probability (Chapter 6) Three people each select a letter of the alphabet.

a) What is the probability that they select the same letter?

b) What is the probability that they select different letters?

7.1 Probability Distributions

In Chapter 6, the emphasis was on the probability of individual outcomes from experiments. This chapter develops models for distributions that show the probabilities of all possible outcomes of an experiment. The distributions can involve outcomes with equal or different likelihoods. Distribution models have applications in many fields including science, game theory, economics, telecommunications, and manufacturing.

INVESTIGATE & INQUIRE: Simulating a Probability Experiment

For project presentations, Mr. Fermat has divided the students in his class into five groups, designated A, B, C, D, and E. Mr. Fermat randomly selects the order in which the groups make their presentations. Develop a simulation to compare the probabilities of group A presenting their project first, second, third, fourth, or fifth.

Method 1: Selecting by Hand

1. Label five slips of paper as A, B, C, D, and E.
2. Randomly select the slips one by one. Set up a table to record the order of the slips and note the position of slip A in the sequence.
3. Repeat this process for a total of ten trials.
4. Combine your results with those from all of your classmates.
5. Describe the results and calculate an empirical probability for each of the five possible outcomes.
6. Reflect on the results. Do you think they accurately represent the situation? Why or why not?

Method 2: Selecting by Computer or Graphing Calculator

1. Use a computer or graphing calculator to generate random numbers between 1 and 5. The generator must be programmed to not repeat a number within a trial. Assign A = 1, B = 2, C = 3, D = 4, and E = 5.
2. Run a series of trials and tabulate the results. If you are skilled in programming, you can set the calculator or software to run a large number of trials and tabulate the results for you. If you run fewer than 100 trials, combine your results with those of your classmates.

See Appendix B for details on software and graphing calculator functions you could use in your simulation.

3. Calculate an empirical probability for each possible outcome.
4. Reflect on the results. Do you think they accurately represent the situation? Why or why not?

The methods in the investigation on page 369 can be adapted to simulate any type of probability distribution:

Step 1 Choose a suitable tool to simulate the random selection process. You could use software, a graphing calculator, or manual methods, such as dice, slips of paper, and playing cards. (See section 1.4.) Look for simple ways to model the selection process.

Step 2 Decide how many trials to run. Determine whether you need to simulate the full situation or if a sample will be sufficient. You may want several groups to perform the experiment simultaneously and then pool their results.

Step 3 Design each trial so that it simulates the actual situation. In particular, note whether you must simulate the selected items being *replaced* (independent outcomes) or *not replaced* (dependent outcomes).

Step 4 Set up a method to record the frequency of each outcome (such as a table, chart, or software function). Combine your results with those of your classmates, if necessary.

Step 5 Calculate empirical probabilities for the simulated outcomes. The sum of the probabilities in the distribution must equal 1.

Step 6 Reflect on the results and decide if they accurately represent the situation being simulated.

Many probability experiments have numerical outcomes—outcomes that can be counted or measured. A **random variable**, $\boldsymbol{X}$, has a single value (denoted x) for each outcome in an experiment. For example, if X is the number rolled with a die, then x has a different value for each of the six possible outcomes. Random variables can be discrete or continuous. **Discrete variables** have values that are separate from each other, and the number of possible values can be small. **Continuous variables** have an infinite number of possible values in a continuous interval. This chapter describes distributions involving discrete random variables. These variables often have integer values.

Usually you select the property or attribute that you want to measure as the random variable when calculating probability distributions. The probability of a random variable having a particular value x is represented as $P(X = x)$, or $P(x)$ for short.

Project Prep

The difference between a discrete random variable and a continuous random variable will be important for your probability distributions project.

Example 1 Random Variables

Classify each of the following random variables as discrete or continuous.

a) the number of phone calls made by a salesperson

b) the length of time the salesperson spent on the telephone

c) a company's annual sales

d) the number of widgets sold by the company

e) the distance from Earth to the sun

Solution

a) Discrete: The number of phone calls must be an integer.

b) Continuous: The time spent can be measured to fractions of a second.

c) Discrete: The sales are a whole number of dollars and cents.

d) Discrete: Presumably the company sells only whole widgets.

e) Continuous: Earth's distance from the sun varies continuously since Earth moves in an elliptical orbit around the sun.

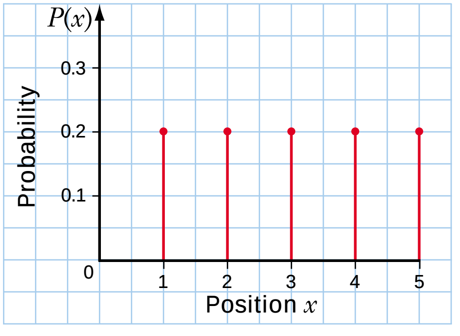
Example 2 Uniform Probability Distribution

Determine the probability distribution for the order of the group presentations simulated in the investigation on page 369.

Solution

Rather than considering the selection of a group to be first, second, and so on, think of each group randomly choosing its position in the order of presentations. Since each group would have an equal probability for choosing each of the five positions, each probability is $\frac{1}{5}$.

Random Variable, x	Probability, $P(x)$
Position 1	$\frac{1}{5}$
Position 2	$\frac{1}{5}$
Position 3	$\frac{1}{5}$
Position 4	$\frac{1}{5}$
Position 5	$\frac{1}{5}$



Observe that all outcomes in this distribution are equally likely in any single trial. A distribution with this property is a **uniform probability distribution**. The sum of the probabilities in this distribution is 1. In fact, all probability distributions must sum to 1 since they include all possible outcomes.

All outcomes in a uniform probability distribution are equally likely. So, for all values of x,

Probability in a Discrete Uniform Distribution

$$P(x) = \frac{1}{n},$$

where n is the number of possible outcomes in the experiment.

An **expectation** or **expected value, $E(X)$**, is the *predicted* average of all possible outcomes of a probability experiment. The expectation is equal to the sum of the products of each outcome (random variable $= x_i$) with its probability, $P(x_i)$.

Expectation for a Discrete Probability Distribution

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

Recall that the capital sigma, $\sum$, means "the sum of." The limits below and above the sigma show that the sum is from the first term ($i = 1$) to the nth term.

Example 3 Dice Game

Consider a simple game in which you roll a single die. If you roll an even number, you gain that number of points, and, if you roll an odd number, you lose that number of points.

a) Show the probability distribution of points in this game.

b) What is the expected number of points per roll?

c) Is this a fair game? Why?

Solution

a) Here the random variable is the number of points scored, not the number rolled.

Number on Upper Face	Points, x	Probability, $P(x)$
1	–1	$\frac{1}{6}$
2	2	$\frac{1}{6}$
3	–3	$\frac{1}{6}$
4	4	$\frac{1}{6}$
5	–5	$\frac{1}{6}$
6	6	$\frac{1}{6}$

b) Since each outcome occurs $\frac{1}{6}$ of the time, the expected number of points per roll is

$$E(X) = \left(-1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \left(-3 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right) + \left(-5 \times \frac{1}{6}\right) + \left(6 \times \frac{1}{6}\right)$$
$$= (-1 + 2 - 3 + 4 - 5 + 6) \times \frac{1}{6}$$
$$= 0.5$$

You would expect that the score in this game would average out to 0.5 points per roll.

c) The game is not fair because the points gained and lost are not equal. For a game to be **fair**, the expected outcome must be 0.

Example 4 Canoe Lengths

A summer camp has seven 4.6-m canoes, ten 5.0-m canoes, four 5.2-m canoes, and four 6.1-m canoes. Canoes are assigned randomly for campers going on a canoe trip.

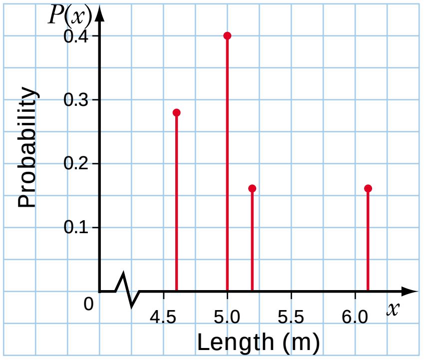
a) Show the probability distribution for the length of an assigned canoe.

b) What is the expected length of an assigned canoe?

Solution

a) Here the random variable is the canoe length.

Length of Canoe (m), x	Probability, $P(x)$
4.6	$\frac{7}{25}$
5.0	$\frac{10}{25}$
5.2	$\frac{4}{25}$
6.1	$\frac{4}{25}$



Observe that the sum of the probabilities is again 1, but the probabilities are not equal. This distribution is not uniform.

b) $$E(X) = (4.6)\left(\frac{7}{25}\right) + (5.0)\left(\frac{10}{25}\right) + (5.2)\left(\frac{4}{25}\right) + (6.1)\left(\frac{4}{25}\right)$$
$$= 5.1$$

The expected length of the canoe is 5.1 m.

Key Concepts

- A random variable, X, has a single value for each outcome in the experiment. Discrete random variables have separated values while continuous random variables have an infinite number of outcomes along a continuous interval.
- A probability distribution shows the probabilities of all the possible outcomes of an experiment. The sum of the probabilities in any distribution is 1.
- Expectation, or the predicted average of all possible outcomes of a probability experiment, is

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

- The expected outcome in a fair game is 0.
- The outcomes of a uniform probability distribution all have the same probability, $P(x) = \frac{1}{n}$, where n is the number of possible outcomes in the experiment.
- You can simulate a probability distribution with manual methods, calculators, or computer software.

Communicate Your Understanding

1. Explain the principal differences between the graphs of the probability distributions in Example 2 and Example 4.
2. In the game of battleship, you select squares on a grid and your opponent tells you if you scored a "hit." Is this process a uniform distribution? What evidence can you provide to support your position?

Practise

A

1. Classify each of the following random variables as discrete or continuous.
   a) number of times you catch a ball in a baseball game
   b) length of time you play in a baseball game
   c) length of a car in centimetres
   d) number of red cars on the highway
   e) volume of water in a tank
   f) number of candies in a box
2. Explain whether each of the following experiments has a uniform probability distribution.
   a) selecting the winning number for a lottery
   b) selecting three people to attend a conference
   c) flipping a coin

d) generating a random number between 1 and 20 with a calculator

e) guessing a person's age

f) cutting a card from a well-shuffled deck

g) rolling a number with two dice

3. Given the following probability distributions, determine the expected values.

a)

x	$P(x)$
5	0.3
10	0.25
15	0.45

b)

x	$P(x)$
1 000	0.25
100 000	0.25
1 000 000	0.25
10 000 000	0.25

c)

x	$P(x)$
1	$\frac{1}{6}$
2	$\frac{1}{5}$
3	$\frac{1}{4}$
4	$\frac{1}{3}$
5	$\frac{1}{20}$

4. A spinner has eight equally-sized sectors, numbered 1 through 8.

a) What is the probability that the arrow on the spinner will stop on a prime number?

b) What is the expected outcome, to the nearest tenth?

Apply, Solve, Communicate

B

5. A survey company is randomly calling telephone numbers in your exchange.

a) Do these calls have a uniform distribution? Explain.

b) What is the probability that a particular telephone number will receive the next call?

c) What is the probability that the last four digits of the next number called will all be the same?

6. a) Determine the probability distribution for the sum rolled with two dice.

b) Determine the expected sum of two dice.

c) Repeat parts a) and b) for the sum of three dice.

7. There are only five perfectly symmetrical polyhedrons: the tetrahedron (4 faces), the cube (6 faces), the octahedron (8 faces), the dodecahedron (12 faces), and the icosahedron (20 faces). Calculate the expected value for dies made in each of these shapes.

8. A lottery has a \$1 000 000 first prize, a \$25 000 second prize, and five \$1000 third prizes. A total of 2 000 000 tickets are sold.

a) What is the probability of winning a prize in this lottery?

b) If a ticket costs \$2.00, what is the expected profit per ticket?

9. **Communication** A game consists of rolling a die. If an even number shows, you receive double the value of the upper face in points. If an odd number shows, you lose points equivalent to triple the value of the upper face.

a) What is the expectation?

b) Is this game fair? Explain.

10. **Application** In a lottery, there are 2 000 000 tickets to be sold. The prizes are as follows:

Prize (\$)	Number of Prizes
1 000 000	1
50 000	5
1 000	10
50	50

What should the lottery operators charge per ticket in order to make a 40% profit?

11. In a family with two children, determine the probability distribution for the number of girls. What is the expected number of girls?

12. A computer has been programmed to draw a rectangle with perimeter of 24 cm. The program randomly chooses integer lengths. What is the expected area of the rectangle?

13. Suppose you are designing a board game with a rule that players who land on a particular square must roll two dice to determine where they move next. Players move ahead five squares for a roll with a sum of 7 and three squares for a sum of 4 or 10. Players move back n squares for any other roll.

a) Develop a simulation to determine the value of n for which the expected move is zero squares.

b) Use the probability distribution to verify that the value of n from your simulation does produce an expected move of zero squares.

14. **Inquiry / Problem Solving** Cheryl and Fatima each have two children. Cheryl's oldest child is a boy, and Fatima has at least one son.

a) Develop a simulation to determine whether Cheryl or Fatima has the greater probability of having two sons.

b) Use the techniques of this section to verify the results of your simulation.

15. Chapter Problem Suppose you buy four boxes of the Krakked Korn cereal. Remember that each box has an equal probability of containing any one of the seven collector cards.

a) What is the probability of getting

i) four identical cards?

ii) three identical cards?

iii) two identical and two different cards?

iv) two pairs of identical cards?

v) four different cards?

b) Sketch a probability distribution for the number of different cards you might find in the four boxes of cereal.

c) Is the distribution in part b) uniform?

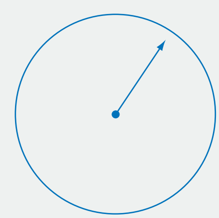
ACHIEVEMENT CHECK

Knowledge / Understanding	Thinking / Inquiry / Problem Solving	Communication	Application

16. A spinner with five regions is used in a game. The probabilities of the regions are

$P(1) = 0.3$
$P(2) = 0.2$
$P(3) = 0.1$
$P(4) = 0.1$
$P(5) = 0.3$

a) Sketch and label a spinner that will generate these probabilities.



b) The rules of the game are as follows: If you spin and land on an even number, you receive double that number of points. If you land on an odd number, you lose that number of points. What is the expected number of points a player will win or lose?

c) Sketch a graph of the probability distribution for this game.

d) Show that this game is not fair. Explain in words.

e) Alter the game to make it fair. Prove mathematically that your version is fair.

17. **Application** The door prizes at a dance are gift certificates from local merchants. There are four \$10 certificates, five \$20 certificates, and three \$50 certificates. The prize envelopes are mixed together in a bag and are drawn at random.

a) Use a tree diagram to illustrate the possible outcomes for selecting the first two prizes to be given out.

b) Determine the probability distribution for the number of \$20 certificates in the first two prizes drawn.

c) What is the probability that exactly three of the first five prizes selected will be \$10 certificates?

d) What is the expected number of \$10 certificates among the first five prizes drawn?

C

18. Most casinos have roulette wheels. In North America, these wheels have 38 slots, numbered 1 to 36, 0, and 00. The 0 and 00 slots are coloured green. Half of the remaining slots are red and the other half are black. A ball rolls around the wheel and players bet on which slot the ball will stop in. If a player guesses correctly, the casino pays out according to the type of bet.

a) Calculate the house advantage, which is the casino's profit, as a percent of the total amount wagered for each of the following bets. Assume that players place their bets randomly.

i) single number bet, payout ratio 35:1

ii) red number bet, payout ratio 1:1

iii) odd number bet, payout ratio 1:1

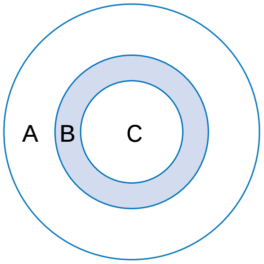
iv) 6-number group, payout ratio 5:1

v) 12-number group, payout ratio 2:1

b) Estimate the weekly profit that a roulette wheel could make for a casino. List the assumptions you have to make for your calculation.

c) European roulette wheels have only one zero. Describe how this difference would affect the house advantage.

19. **Inquiry / Problem Solving** Three concentric circles are drawn with radii of 8 cm, 12 cm, and 20 cm. If a dart lands randomly on this target, what are the probabilities of it landing in each region?



20. A die is a random device for which each possible value of the random variable has a probability of $\frac{1}{6}$. Design a random device with the probabilities listed below and determine the expectation for each device. Use a different type of device in parts a) and b).

a) $P(0) = \frac{1}{4}$

$P(1) = \frac{1}{6}$

$P(2) = P(3) = \frac{1}{8}$

$P(4) = P(5) = P(6) = P(7) = \frac{1}{12}$

b) $P(0) = \frac{1}{6}$

$P(1) = P(2) = \frac{1}{4}$

$P(3) = \frac{1}{3}$

21. **Communication** Explain how the population mean, μ, and the expectation, $E(X)$, are related.

7.2 Binomial Distributions

A manufacturing company needs to know the expected number of defective units among its products. A polling company wants to estimate how many people are in favour of a new environmental law. In both these cases, the companies can view the individual outcomes in terms of "success" or "failure." For the manufacturing company, a success is a product without defects; for the polling company, a success is an interview subject who supports the new law. Repeated independent trials measured in terms of such successes or failures are **Bernoulli trials**, named after Jacob Bernoulli (1654–1705), a Swiss mathematician who published important papers on logic, algebra, geometry, calculus, and probability.

WEB CONNECTION

www.mcgrawhill.ca/links/MDM12

To learn more about Bernoulli trials, visit the above web site and follow the links. Describe an application of Bernoulli trials that interests you.

INVESTIGATE & INQUIRE: Success/Failure Simulation

The Choco-Latie Candies company makes candy-coated chocolates, 40% of which are red. The production line mixes the candies randomly and packages ten per box. Develop a simulation to determine the expected number of red candies in a box.

1. Determine the key elements of the selection process and choose a random-number generator or other tool to simulate the process.
2. Decide whether each trial in your simulation must be independent. Does the colour of the first candy in a box affect the probability for the next one? Describe how you can set up your simulation to reflect the way in which the candies are packed into the boxes.
3. Run a set of ten trials to simulate filling one of the boxes and record the number of red candies. How could you measure this result in terms of successes and failures?
4. Simulate filling at least nine more boxes and record the number of successes for each box.

5. Review your results. Do you think the ten sets of trials are enough to give a reasonable estimate of the expected number of red candies? Explain your reasoning. If necessary, simulate additional sets of trials or pool your results with those of other students in your class.
6. Summarize the results by calculating empirical probabilities for each possible number of red candies in a box. The sum of these probabilities must equal 1. Calculate the expected number of red candies per box based on these results.
7. Reflect on the results. Do they accurately represent the expected number of red candies in a box?
8. Compare your simulation and its results with those of the other students in your class. Which methods produced the most reliable results?

The probabilities in the simulation above are an example of a **binomial distribution**. For such distributions, all the trials are *independent* and have only two possible outcomes, success or failure. The probability of success is the same in every trial—the outcome of one trial does not affect the probabilities of any of the later trials. The random variable is the number of successes in a given number of trials.

Example 1 Success/Failure Probabilities

A manufacturer of electronics components produces precision resistors designed to have a tolerance of ±1%. From quality-control testing, the manufacturer knows that about one resistor in six is actually within just 0.3% of its nominal value. A customer needs three of these more precise resistors. What is the probability of finding exactly three such resistors among the first five tested?

Solution

You can apply the concept of Bernoulli trials because the tolerances of the resistors are independent of each other. A success is finding a resistor with a tolerance of ±0.3% or less.

For each resistor, the probability of success is about $\frac{1}{6}$ and the probability of failure is about $\frac{5}{6}$.

You can choose three resistors from the batch of five in ${}_5C_3$ ways. Since the outcomes are independent events, you can apply the product rule for independent events. The probability of success with all three resistors in each of these combinations is the product of the probabilities for the individual resistors.

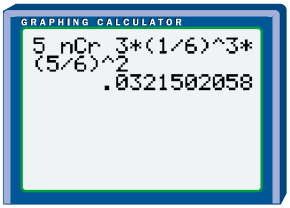
$$\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^3$$

The probability for the three successful trials is equal to the number of ways they can occur multiplied by the probability for each way: ${}_5C_3\left(\frac{1}{6}\right)^3$.

Similarly, there are ${}_2C_2$ ways of choosing the other two resistors and the probability of a failure with both these resistors is $\left(\frac{5}{6}\right)^2$. Thus, the probability for the two failures is ${}_2C_2\left(\frac{5}{6}\right)^2 = \left(\frac{5}{6}\right)^2$ since ${}_nC_n = 1$.

Now, apply the product rule for independent events to find the probability of having three successes and two failures in the five trials.

$$\begin{aligned} P(x = 3) &= {}_5C_3\left(\frac{5}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 \\ &= 10\left(\frac{1}{216}\right)\left(\frac{25}{36}\right) \\ &= 0.032150 \ldots \end{aligned}$$



The probability that exactly three of the five resistors will meet the customer's specification is approximately 0.032.

You can apply the method in Example 1 to show that the probability of x successes in n Bernoulli trials is

Probability in a Binomial Distribution

$P(x) = {}_nC_x p^x q^{n-x}$,
where p is the probability of success on any individual trial and $q = 1 - p$ is the probability of failure.

Since the probability for all trials is the same, the expectation for a success in any one trial is p. The expectation for n independent trials is

Expectation for a Binomial Distribution

$E(X) = np$

Example 2 Expectation in a Binomial Distribution

Tan's family moves to an area with a different telephone exchange, so they have to get a new telephone number. Telephone numbers in the new exchange start with 446, and all combinations for the four remaining digits are equally likely. Tan's favourite numbers are the prime numbers 2, 3, 5, and 7.

a) Calculate the probability distribution for the number of these prime digits in Tan's new telephone number.

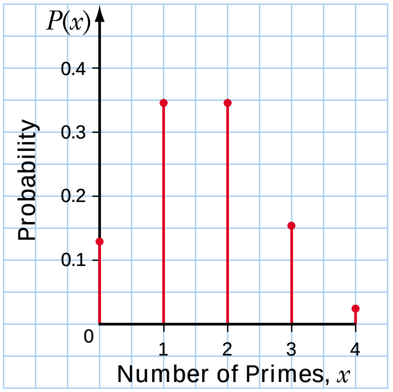
b) What is the expected number of these prime digits in the new telephone number?

Solution 1 Using Pencil and Paper

a) The probability of an individual digit being one of Tan's favourite numbers is $\frac{4}{10} = 0.4$.

So,
$p = 0.4$ and $q = 1 - 0.4$
$= 0.6$

Number of Primes, *x*	Probability, *P*(*x*)
0	${}_4C_0(0.4)^0(0.6)^4 = 0.1296$
1	${}_4C_1(0.4)^1(0.6)^3 = 0.3456$
2	${}_4C_2(0.4)^2(0.6)^2 = 0.3456$
3	${}_4C_3(0.4)^3(0.6)^1 = 0.1536$
4	${}_4C_4(0.4)^4(0.6)^0 = 0.0256$



b) You can calculate the expectation in two ways.

Using the equation for the expectation of any probability distribution,

$$E(X) = 0(0.1296) + 1(0.3456) + 2(0.3456) + 3(0.1536) + 4(0.0256)$$
$$= 1.6$$

Using the formula for a binomial expectation,

$$E(X) = np$$
$$= 4(0.4)$$
$$= 1.6$$

On average, there will be 1.6 of Tan's favourite digits in telephone numbers in his new exchange.

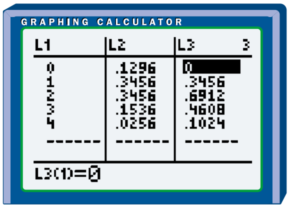
Solution 2 Using a Graphing Calculator

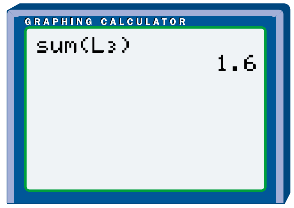
a) Check that lists L1, L2, and L3 are clear. Enter all the possible values for x into L1.

Use the **binompdf(function** from the DISTR menu to calculate the probabilities for each value of x. Binompdf stands for binomial probability density function.

Enter the formula binompdf(4,.4,L1) into L2 to find the values of $P(x)$.

b) In L3, calculate the value of $xP(x)$ using the formula L1 × L2. Then, use the **sum(function** in the LIST OPS menu to calculate the expected value.





For more details on software and graphing calculator functions, see Appendix B.

Solution 3 Using a Spreadsheet

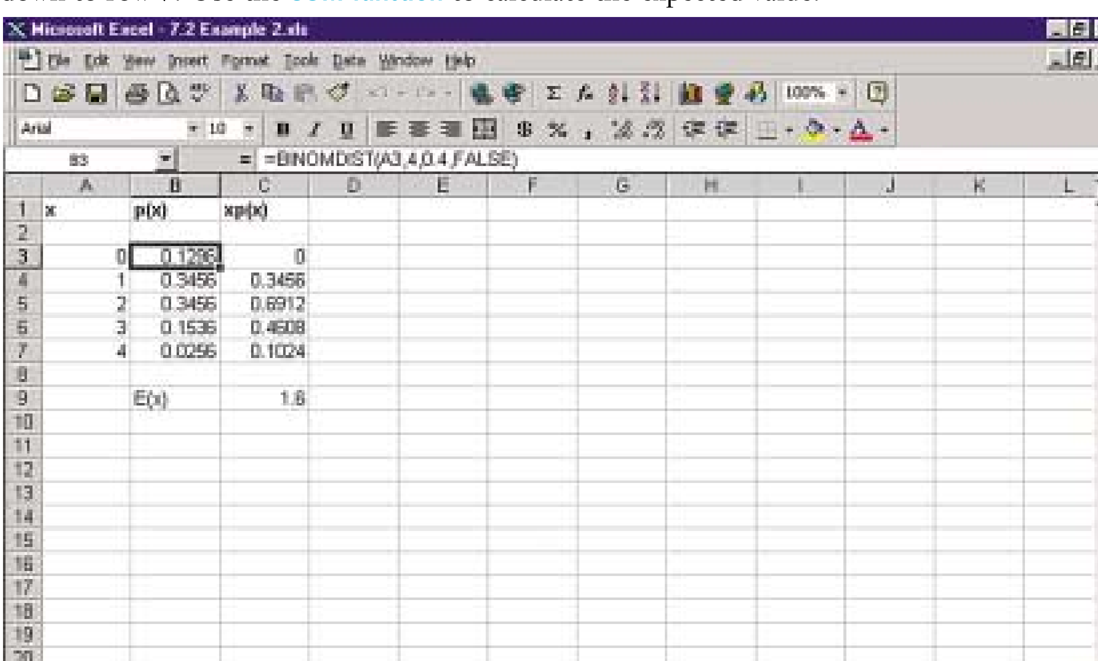
a) Open a new spreadsheet. Create headings for x, $P(x)$, and $xP(x)$ in columns A to C. Enter the possible values of the random variable x in column A. Use the **BINOMDIST function** to calculate the probabilities for the different values of x. The syntax for this function is

BINOMDIST(*number of successes*, *trials*, *probability of success*, *cumulative*)

Enter this formula in cell B3 and then copy the formula into cells B4 through B7.

Note that you must set the cumulative feature to 0 or FALSE, so that the function calculates the probability of exactly 4 successes rather than the probability of up to 4 successes.

b) To calculate $xP(x)$ in column C, enter the formula A3*B3 in C3 and then copy cell C3 down to row 7. Use the **SUM function** to calculate the expected value.



Solution 4 Using Fathom™

a) Open a new Fathom™ document. Drag a new **collection** box to the work area and name it Primes. Create five new cases.

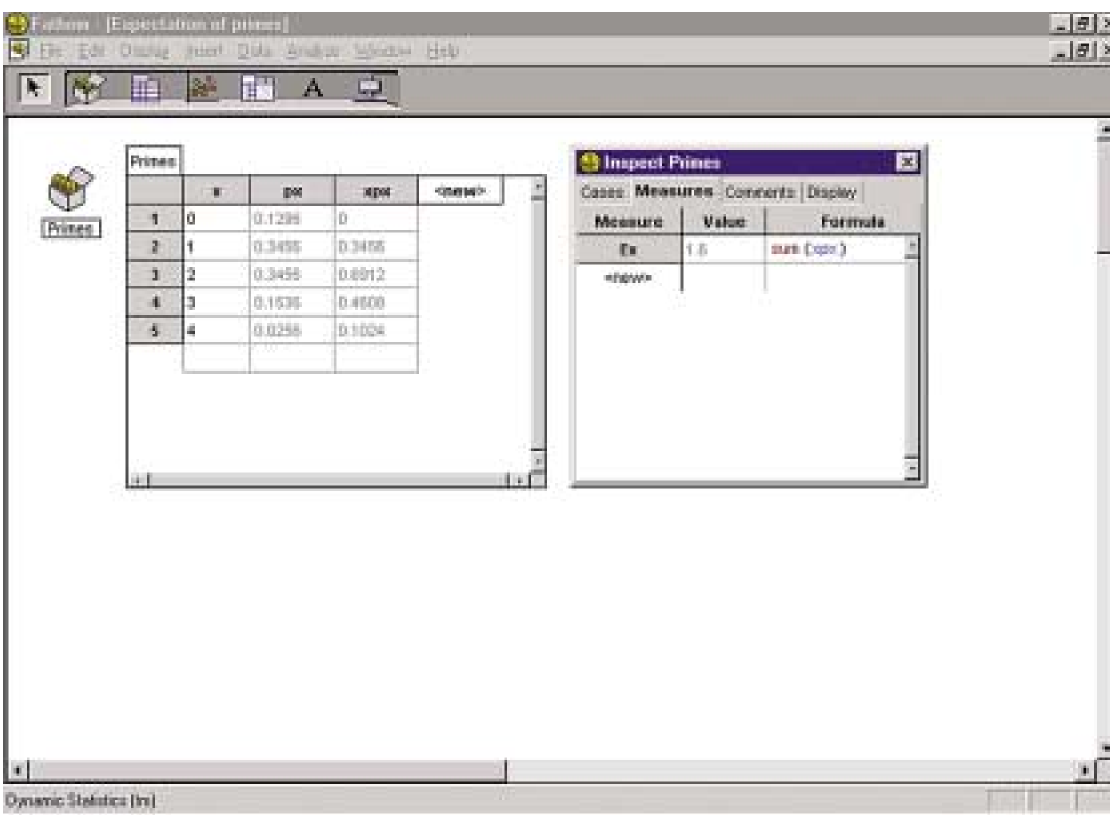
Drag a new **case table** to the work area. Create three new attributes: x, px, and xpx. Enter the values from 0 to 4 in the x attribute column. Then, use the **binomialProbability function** to calculate the probabilities for the different values of x. The syntax of this function is

binomialProbability(*number of successes*, *trials*, *probability of success*)

Right-click on the px attribute and then select Edit Formula / Functions / Distributions / Binomial to enter binomialProbability(x,4,.4).

b) You can calculate $xP(x)$ with the formula x*px.

Next, double-click on the **collection** box to open the **inspector**. Select the Measures tab, and name a new measure Ex. Right-click on Ex and use the **sum function**, located in the Functions / Statistical / One Attribute menu, to enter the formula sum(x*px).



Example 3 Counting Candies

Consider the Choco-Latie candies described in the investigation on page 378.

a) What is the probability that at least three candies in a given box are red?

b) What is the expected number of red candies in a box?

Solution

a) A success is a candy being red, so $p = 0.4$ and $q = 1 - 0.4 = 0.6$. You could add the probabilities of having exactly three, four, five, …, or ten red candies, but it is easier to use an indirect method.

$$\begin{aligned} P(\geq 3 \text{ red}) &= 1 - P(<3) \\ &= 1 - P(0 \text{ red}) - P(1 \text{ red}) - P(2 \text{ red}) \\ &= 1 - {}_{10}C_0(0.4)^0(0.6)^{10} - {}_{10}C_1(0.4)^1(0.6)^9 - {}_{10}C_2(0.4)^2(0.6)^8 \\ &\doteq 0.8327 \end{aligned}$$

The probability of at least three candies being red is approximately 0.8327.

b)
$$\begin{aligned} E(X) &= np \\ &= 10(0.4) \\ &= 4 \end{aligned}$$

The expected number of red candies in a box is 4.

Key Concepts

- A binomial distribution has a specified number of independent trials in which the outcome is either success or failure. The probability of a success is the same in each trial.
- The probability of x successes in n independent trials is
  $P(x) = {}_nC_x\, p^x\, q^{n-x}$,
  where p is the probability of success on an individual trial and q is the probability of failure on that same individual trial ($p + q = 1$).
- The expectation for a binomial distribution is $E(X) = np$.
- To simulate a binomial experiment,
  - choose a simulation method that accurately reflects the probabilities in each trial
  - set up the simulation tool to ensure that each trial is independent
  - record the number of successes and failures in each experiment
  - summarize the results by calculating the probabilities for r successes in n trials (the sum of individual probabilities must equal 1)

Communicate Your Understanding

1. Consider this question: If five cards are dealt from a standard deck, what is the probability that two of the cards are the ace and king of spades?
   a) Explain why the binomial distribution is not a suitable model for this scenario.
   b) How could you change the scenario so that it does fit a binomial distribution? What attributes of a binomial distribution would you use in your modelling?
2. Describe how the graph in Example 2 differs from the graph of a uniform distribution.
3. Compare your results from the simulation of the Choco-Latie candies at the beginning of this section with the calculated values in Example 3. Explain any similarities or differences.

Practise

A

1. Which of the following situations can be modelled by a binomial distribution? Justify your answers.
   a) A child rolls a die ten times and counts the number of 3s.
   b) The first player in a free-throw basketball competition has a free-throw success rate of 88.4%. A second player takes over when the first player misses the basket.
   c) A farmer gives 12 of the 200 cattle in a herd an antibiotic. The farmer then selects 10 cattle at random to test for infections to see if the antibiotic was effective.
   d) A factory producing electric motors has a 0.2% defect rate. A quality-control inspector needs to determine the expected number of motors that would fail in a day's production.
2. Prepare a table and a graph for a binomial distribution with
   a) $p = 0.2$, $n = 5$
   b) $p = 0.5$, $n = 8$

Apply, Solve, Communicate

B

3. Suppose that 5% of the first batch of engines off a new production line have flaws. An inspector randomly selects six engines for testing.
   a) Show the probability distribution for the number of flawed engines in the sample.
   b) What is the expected number of flawed engines in the sample?
4. **Application** Design a simulation to predict the expected number of 7s in Tan's new telephone number in Example 2.
5. The faces of a 12-sided die are numbered from 1 to 12. What is the probability of rolling 9 at least twice in ten tries?
6. **Application** A certain type of rocket has a failure rate of 1.5%.
   a) Design a simulation to illustrate the expected number of failures in 100 launches.
   b) Use the methods developed in this section to determine the probability of fewer than 4 failures in 100 launches.
   c) What is the expected number of failures in 100 launches of the rocket?

7. Suppose that 65% of the families in a town own computers. If eight families are surveyed at random,
   a) what is the probability that at least four own computers?
   b) what is the expected number of families with computers?

8. **Inquiry / Problem Solving** Ten percent of a country's population are left-handed.
   a) What is the probability that 5 people in a group of 20 are left handed?
   b) What is the expected number of left-handed people in a group of 20?
   c) Design a simulation to show that the expectation calculated in part b) is accurate.

9. **Inquiry / Problem Solving** Suppose that Bayanisthol, a new drug, is effective in 65% of clinical trials. Design a problem involving this drug that would fit a binomial distribution. Then, provide a solution to your problem.

10. Pythag-Air-US Airlines has determined that 5% of its customers do not show up for their flights. If a passenger is bumped off a flight because of overbooking, the airline pays the customer $200. What is the expected payout by the airline, if it overbooks a 240-seat airplane by 5%?

11. A department-store promotion involves scratching four boxes on a card to reveal randomly printed letters from A to F. The discount is 10% for each A revealed, 5% for each B revealed, and 1% for the other four letters. What is the expected discount for this promotion?

12. a) Expand the following binomials.
    i) $(p + q)^6$ ii) $(0.2 + 0.8)^5$
    b) Use the expansions to show how the binomial theorem is related to the binomial probability distribution.

ACHIEVEMENT CHECK

Knowledge / Understanding	Thinking / Inquiry / Problem Solving	Communication	Application

13. Your local newspaper publishes an Ultimate Trivia Contest with 12 extremely difficult questions, each having 4 possible answers. You have no idea what the correct answers are, so you make a guess for each question.
    a) Explain why this situation can be modelled by a binomial distribution.
    b) Use a simulation to predict the expected number of correct answers.
    c) Verify your prediction mathematically.
    d) What is the probability that you will get at least 6 answers correct?
    e) What is the probability that you will get fewer than 2 answers correct?
    f) Describe how the graph of this distribution would change if the number of possible answers for each question increases or decreases.

C

14. The French mathematician Simeon-Denis Poisson (1781–1840) developed what is now known as the *Poisson distribution*. This distribution can be used to approximate the binomial distribution if p is very small and n is very large. It uses the formula
$$P(x) = \frac{e^{-np}(np)^x}{x!},$$
where e is the irrational number 2.718 28 … (the base for the natural logarithm).

Use the Poisson distribution to approximate the following situations. Compare the results to those found using the binomial distribution.

a) A certain drug is effective in 98% of cases. If 2000 patients are selected at random, what is the probability that the drug was ineffective in exactly 10 cases?

b) Insurance tables indicate that there is a probability of 0.01 that a driver of a specific model of car will have an accident requiring hospitalization within a one-year period. If the insurance company has 4500 policies, what is the probability of fewer than 5 claims for accidents requiring hospitalization?

c) On election day, only 3% of the population voted for the Environment Party. If 1000 voters were selected at random, what is the probability that fewer than 8 of them voted for the Environment Party?

15. Communication Suppose heads occurs 15 times in 20 tosses of a coin. Do you think the coin is fair? Explain your reasoning.

16. Inquiry / Problem Solving

a) Develop a formula for $P(x)$ in a "trinomial" distribution that has three possible outcomes with probabilities p, q, and r, respectively.

b) Use your formula to determine the probability of rolling a 3 twice and a 5 four times in ten trials with a standard die.

17. Communication A judge in a model-airplane contest says that the probability of a model landing without damage is 0.798, so there is only "one chance in five" that any of the seven models in the finals will be damaged. Discuss the accuracy of the judge's statement.

Career Connection

Actuary

Actuaries are statistics specialists who use business, analytical, and mathematical skills to apply mathematical models to insurance, pensions, and other areas of finance. Actuaries assemble and analyse data and develop probability models for the risks and costs of accidents, sickness, death, pensions, unemployment, and so on. Governments and private companies use such models to determine pension contributions and fair prices for insurance premiums. Actuaries may also be called upon to provide legal evidence on the value of future earnings of an accident victim.

Actuaries must keep up-to-date on social issues, economic trends, business issues, and the law. Most actuaries have a degree in actuarial science, statistics, or mathematics and have studied statistics, calculus, algebra, operations research, numerical analysis, and interest theory. A strong background in business or economics is also useful.

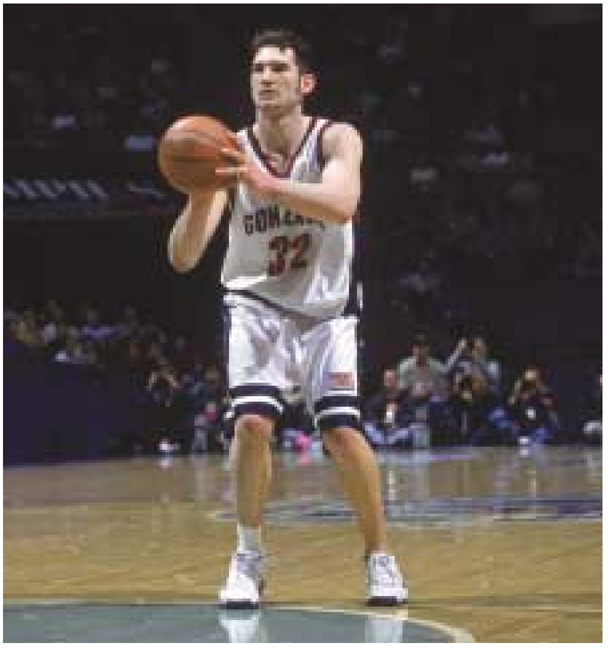
Actuaries work for insurance companies, pension-management firms, accounting firms, labour unions, consulting groups, and federal and provincial governments.

WEB CONNECTION

www.mcgrawhill.ca/links/MDM12

For more information about a career as an actuary, visit the above web site and follow the links.

7.3 Geometric Distributions



In some board games, you cannot move forward until you roll a specific number, which could take several tries. Manufacturers of products such as switches, relays, and hard drives need to know how many operations their products can perform before failing. In some sports competitions, the winner is the player who scores the most points before missing a shot. In each of these situations, the critical quantity is the **waiting time** or **waiting period**—the number of trials before a specific outcome occurs.

INVESTIGATE & INQUIRE: Simulating Waiting Times

To get out of jail in the game of MONOPOLY®, you must either roll doubles or pay the bank \$50. Design a simulation to find the probabilities of getting out of jail in x rolls of the two dice.

1. Select a random-number generator to simulate the selection process.
2. Decide how to simplify the selection process. Decide, also, whether the full situation needs to be simulated or whether a proportion of the trials would be sufficient.
3. Design each trial so that it simulates the actual situation. Determine whether your simulation tool must be reset or replaced after each trial to properly correspond to rolling two dice.
4. Set up a method to record the frequency of each outcome. Record the number of failures before a success finally occurs.
5. Combine your results with those of your classmates, if necessary.
6. Use these results to calculate an empirical probability for each outcome and the expected waiting time before a success.
7. Reflect on the results. Do they accurately represent the expected number of failures before success?
8. Compare your simulation and its results with those of other students in your class. Which simulation do you think worked best? Explain the reasons for your choice.

The simulation above models a **geometric distribution**. Like binomial distributions, trials in a geometric distribution have only two possible outcomes, success or failure, whose probabilities do not change from one trial to the next. However, the random variable for a geometric distribution is the *waiting time*, the number of unsuccessful independent trials before success occurs. Having different random variables causes significant differences between binomial and geometric distributions.

Example 1 Getting out of Jail in MONOPOLY®

a) Calculate the probability distribution for getting out of jail in MONOPOLY® in x rolls of the dice.

b) Estimate the expected number of rolls before getting out of jail.

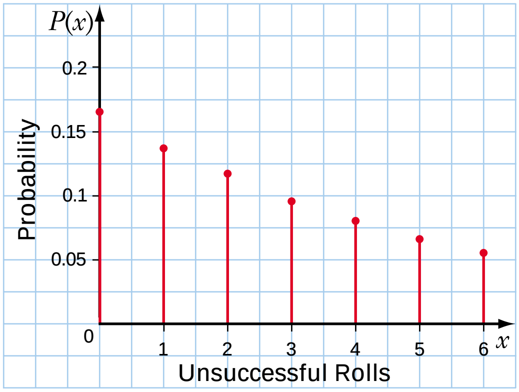
Solution 1 Using Pencil and Paper

a) The random variable is the number of unsuccessful rolls before you get out of jail. You can get out of jail by rolling doubles, and $P(\text{doubles}) = \frac{6}{36}$.
So, for each independent roll,

$$p = \frac{6}{36} \qquad \text{and} \qquad q = 1 - \frac{1}{6}$$
$$= \frac{1}{6} \qquad\qquad\qquad = \frac{5}{6}$$

You can apply the product rule to find the probability of successive independent events (see section 6.3). Thus, each unsuccessful roll preceding the successful one adds a factor of $\frac{5}{6}$ to the probability.

Unsuccessful Rolls (Waiting Time), *x*	Probability, *P(x)*
0	$\frac{1}{6} = 0.166\ 66\ \ldots$
1	$\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) = 0.138\ 88\ \ldots$
2	$\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right) = 0.115\ 74\ \ldots$
3	$\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right) = 0.096\ 45\ \ldots$
4	$\left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right) = 0.080\ 37\ \ldots$
…	…



This distribution theoretically continues forever since one possible outcome is that the player never rolls doubles. However, the probability for a waiting time decreases markedly as the waiting time increases. Although this distribution is an infinite geometric series, its terms still sum to 1 since they represent the probabilities of all possible outcomes.

b) Calculate the first six terms. If these terms approach zero rapidly, the sum of these terms will give a rough first approximation for the expectation.

$$\begin{aligned}E(X) &= \sum_{x=0}^{\infty} xP(x)\\ &= (0)(0.16666\ldots) + (1)(0.13888\ldots) + (2)(0.11574\ldots) + 3(0.09645\ldots)\\ &\quad + (4)(0.08037\ldots) + (5)(0.06698\ldots) + \ldots\\ &> 0 + 0.13888 \quad + 0.23148 \quad + 0.28935 + 0.32150 + 0.33489 + \ldots\\ &> 1.3\end{aligned}$$

Clearly, the six terms are not approaching zero rapidly. All you can conclude is that the expected number of rolls before getting out of jail in MONOPOLY® is definitely more than 1.3.

Solution 2 Using a Graphing Calculator

a) You can use the calculator's lists to display the probabilities. Clear lists L1 to L4. For a start, enter the integers from 0 to 39 into L1. The **seq(function** in the LIST OPS menu provides a convenient way to enter these numbers.

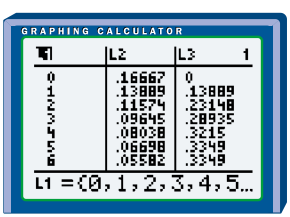
seq(A,A,0,39)→L1

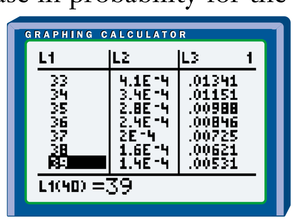
Next, use the geometric probability density function in the DISTR menu to calculate the probability of each value of x. The **geometpdf(function** has the syntax

geometpdf(*probability of success*, *number of trial on which first success occurs*)

Since x is the number of trials *before* success occurs, the number of the trial on which the first success occurs is $x + 1$. Therefore, enter geometpdf(1/6,L1+1)→L2.

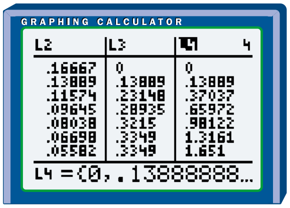
Notice the dramatic decrease in probability for the higher values of x.

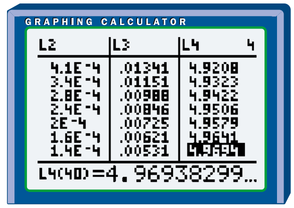




b) In L3, calculate the value of $xP(x)$ with the simple formula L1 × L2. You can sum these values to get a reasonable estimate for the expectation. To determine the accuracy of this estimate, it is helpful to look at the cumulative or running total of the $xP(x)$ values in L3.

To do this, in L4, select 6:cumSum(from the LIST OPS menu and type L3).





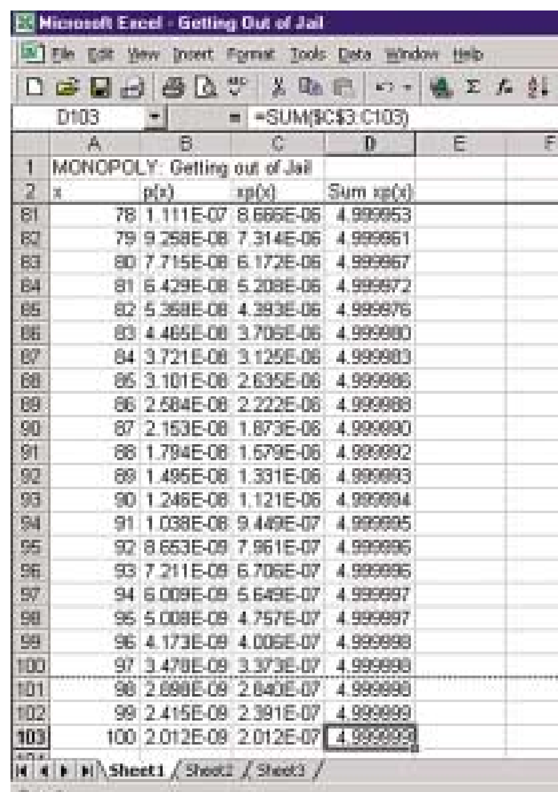
Note how the running total of $xP(x)$ in L4 increases more and more slowly toward the end of the list, suggesting that the infinite series for the expectation will total to a little more than 4.97. Thus, 5.0 would be a reasonable estimate for the number of trials before rolling doubles. You can check this estimate by performing similar calculations for waiting times of 100 trials or more.

Solution 3 Using a Spreadsheet

a) Open a new spreadsheet. Enter the headings x, p(x), and xp(x) in columns A, B, and C. Use the **Fill feature** to enter a sequence of values of the random variable x in column A, starting with 0 and going up to 100. Calculate the probability $P(x)$ for each value of the random variable x by entering the formula (5/6)^A3*(1/6) in cell B3 and then copying it down the rest of the column.

b) You can calculate $xP(x)$ in column C by entering the formula A3*B3 in cell C3 and then copying it down the rest of the column. Next, calculate the cumulative expected values in column D using the **SUM function** (with absolute cell references for the first cell) in cell D3 and copying it down the column.

Note that for $x = 50$ the cumulative expected value is over 4.99. By $x = 100$, it has reached 4.999 999 and is increasing extremely slowly. Thus, 5.000 00 is an accurate estimate for the expected number of trials before rolling doubles.



You can use the method in Example 1 to show that the probability of success after a waiting time of x failures is

Probability in a Geometric Distribution

$P(x) = q^x p$,
where p is the probability of success in each single trial and q is the probability of failure.

The expectation of a geometric distribution is the sum of an infinite series. Using calculus, it is possible to show that this expectation converges to a simple formula.

Expectation for a Geometric Distribution

$$E(X) = \sum_{x=0}^{\infty} xP(x)$$
$$= \frac{q}{p}$$

Project Prep

Techniques for calculating expected values will be useful for your probability distributions project.

Example 2 Expectation of Geometric Distribution

Use the formula for the expectation of a geometric distribution to evaluate the accuracy of the estimates in Example 1.

Solution

$$E(X) = \frac{q}{p}$$
$$= \frac{\frac{5}{6}}{\frac{1}{6}}$$
$$= 5$$

For this particular geometric distribution, the simple manual estimate in Example 1 is accurate only to an order of magnitude. However, the calculator and the spreadsheet estimates are much more accurate.

Example 3 Basketball Free Throws

Jamaal has a success rate of 68% for scoring on free throws in basketball. What is the expected waiting time before he misses the basket on a free throw?

Solution

Here, the random variable is the number of trials before Jamaal misses on a free throw. For calculating the waiting time, a success is Jamaal *failing* to score. Thus,

$$q = 0.68 \quad \text{and} \quad \begin{aligned} p &= 1 - 0.68 \\ &= 0.32 \end{aligned}$$

Using the expectation formula for the geometric distribution,

$$\begin{aligned} E(X) &= \frac{q}{p} \\ &= \frac{0.68}{0.32} \\ &= 2.1 \end{aligned}$$

The expectation is that Jamaal will score on 2.1 free throws before missing.

Example 4 Traffic Management

Suppose that an intersection you pass on your way to school has a traffic light that is green for 40 s and then amber or red for a total of 60 s.

a) What is the probability that the light will be green when you reach the intersection at least once a week?

b) What is the expected number of days before the light is green when you reach the intersection?

Solution

a) Each trial is independent with

$$\begin{aligned} p &= \frac{40}{100} \\ &= 0.40 \end{aligned} \quad \text{and} \quad q = 0.60$$

There are five school days in a week. To get a green light on one of those five days, your waiting time must be four days or less.

$$\begin{aligned} P(0 \le x \le 4) &= 0.40 + (0.60)(0.40) + (0.60)^2(0.40) + (0.60)^3(0.40) + (0.60)^4(0.40) \\ &= 0.92 \end{aligned}$$

The probability of the light being green when you reach the intersection at least once a week is 0.92.

b) $E(X) = \frac{q}{p}$

$= \frac{0.60}{0.40}$

$= 1.5$

The expected waiting time before catching a green light is 1.5 days.

Key Concepts

- A geometric distribution has a specified number of independent trials with two possible outcomes, success or failure. The random variable is the number of unsuccessful outcomes before a success occurs.
- The probability of success after a waiting time of x failures is $P(x) = q^x p$, where p is the probability of success in each single trial and q is the probability of failure.
- The expectation of a geometric distribution is $E(X) = \frac{q}{p}$.
- To simulate a geometric experiment, you must ensure that the probability on a single trial is accurate for the situation and that each trial is independent. Summarize the results by calculating probabilities and the expected waiting time.

Communicate Your Understanding

1. Describe how the graph in Example 1 differs from those of the uniform and binomial distributions.
2. Consider this question: What is the expected number of failures in 100 launches of a rocket that has a failure rate of 1.5%? Explain why this problem does not fit a geometric distribution and how it could be rewritten so that it does.

Practise

A

1. Which of the following situations is modelled by a geometric distribution? Explain your reasoning.
   - **a)** rolling a die until a 6 shows
   - **b)** counting the number of hearts when 13 cards are dealt from a deck
   - **c)** predicting the waiting time when standing in line at a bank
   - **d)** calculating the probability of a prize being won within the first 3 tries
   - **e)** predicting the number of successful launches of satellites this year

Data in Action

In 2000, there were 82 successful launches of satellites and 3 failures. Launching a satellite usually costs between \$75 million and \$600 million.

2. Prepare a table and a graph for six trials of a geometric distribution with

a) $p = 0.2$ b) $p = 0.5$

Apply, Solve, Communicate

B

3. For a 12-sided die,

a) what is the probability that the first 10 will be on the third roll?

b) what is the expected waiting time until a 1 is rolled?

4. The odds in favour of a Pythag-Air-US Airlines flight being on time are 3:1.

a) What is the probability that this airline's next eight flights will be on time?

b) What is the expected waiting time before a flight delay?

5. **Communication** To finish a board game, Sarah needed to land on the last square by rolling a sum of 2 with two dice. She was dismayed that it took her eight tries. Should she have been surprised? Explain.

6. In a TV game show, the grand prize is randomly hidden behind one of three doors. On each show, the finalist gets to choose one of the doors. What is the probability that no finalists will win a grand prize on four consecutive shows?

7. **Application** A teacher provides pizza for his class if they earn an A-average on any test. The probability of the class getting an A-average on one of his tests is 8%.

a) What is the probability that the class will earn a pizza on the fifth test?

b) What is the probability that the class will not earn a pizza for the first seven tests?

c) What is the expected waiting time before the class gets a pizza?

8. Minh has a summer job selling replacement windows by telephone. Of the people he calls, nine out of ten hang up before he can give a sales pitch.

a) What is the probability that, on a given day, Minh's first sales pitch is on his 12th call?

b) What is the expected number of hang-ups before Minh can do a sales pitch?

9. Despite its name, Zippy Pizza delivers only 40% of its pizzas on time.

a) What is the probability that its first four deliveries will be late on any given day?

b) What is the expected number of pizza deliveries before one is on time?

10. A poll indicated that 34% of the population agreed with a recent policy paper issued by the government.

a) What is the probability that the pollster would have to interview five people before finding a supporter of the policy?

b) What is the expected waiting time before the pollster interviews someone who agrees with the policy?

11. Suppose that 1 out of 50 cards in a scratch-and-win promotion gives a prize.

a) What is the probability of winning on your fourth try?

b) What is the probability of winning within your first four tries?

c) What is the expected number of cards you would have to try before winning?

12. A top NHL hockey player scores on 93% of his shots in a shooting competition.

a) What is the probability that the player will not miss the goal until his 20th try?

b) What is the expected number of shots before he misses?

13. **Application** A computer manufacturer finds that 1.5% of its chips fail quality-control testing.

a) What is the probability that one of the first five chips off the line will be defective?

b) What is the expected waiting time for a defective chip?

14. **Inquiry/Problem Solving** Three friends of an avid golfer have each given her a package of 5 balls for her birthday. The three packages are different brands. The golfer keeps all 15 balls in her golf bag and picks one at random at the start of each round.

a) Design a simulation to determine the waiting time before the golfer has tried all three brands. Assume that the golfer does not lose any of the balls.

b) Use the methods described in this section to calculate the expected waiting time before the golfer has tried the three different brands.

15. The Big K cereal company has randomly placed one of a set of seven different collector cards in each box of its Krakked Korn cereal. Each card is equally likely.

Chapter Problem

a) Design a simulation to estimate how many boxes of Krakked Korn you would have to buy to get a complete set of cards.

b) Use the methods described in this section to calculate the average number of boxes of Krakked Korn you would have to buy to get a complete set of cards.

C

16. **Inquiry/Problem Solving** Consider a geometric distribution where the random variable is the number of the trial with the first success instead of the number of failures before a success. Develop formulas for the probabilities and expectation for this distribution.

17. **Communication** A manufacturer of computer parts lists a mean time before failure (MTBF) of 5.4 years for one of its hard drives. Explain why this specification is different from the expected waiting time of a geometric distribution. Could you use the MTBF to calculate the probability of the drive failing in any one-year period?

18. In a sequence of Bernoulli trials, what is the probability that the second success occurs on the fifth trial?

19. **Communication** The rack behind a coat-check counter collapses and 20 coats slip off their numbered hangers. When the first person comes to retrieve one of these coats, the clerk brings them out and holds them up one at a time for the customer to identify.

a) What is the probability that the clerk will find the customer's coat

i) on the first try?

ii) on the second try?

iii) on the third try?

iv) in fewer than 10 tries?

b) What is the expected number of coats the clerk will have to bring out before finding the customer's coat?

c) Explain why you cannot use a geometric distribution to calculate this waiting time.

7.4 Hypergeometric Distributions

When choosing the starting line-up for a game, a coach obviously has to choose a different player for each position. Similarly, when a union elects delegates for a convention or you deal cards from a standard deck, there can be no repetitions. In such situations, each selection reduces the number of items that could be selected in the next trial. Thus, the probabilities in these trials are dependent. Often we need to calculate the probability of a specific number of successes in a given number of dependent trials.

INVESTIGATE & INQUIRE: Choosing a Jury

In Ontario, a citizen can be called for jury duty every three years. Although most juries have 12 members, those for civil trials in Ontario usually require only 6 members. Suppose a civil-court jury is being selected from a pool of 18 citizens, 8 of whom are men. Develop a simulation to determine the probability distribution for the number of women selected for this jury.

1. Select a random-number generator to simulate the selection process.
2. Decide how to simplify the selection process. Decide, also, whether the full situation needs to be simulated or whether a proportion of the trials would be sufficient.
3. Design each trial so that it simulates the actual situation. Ensure that each trial is dependent by setting the random-number generator so that there are no repetitions within each series of trials.
4. Set up a method to record the number of successes in each experiment. Pool your results with those of other students in your class, if necessary.
5. Use the results to estimate the probabilities of x successes (women) in r trials (selections of a juror).
6. Reflect on the results. Do they accurately represent the probability of x women being selected?
7. Compare your simulation and its results with those of your classmates. Which are the better simulations? Explain why.

Data in Action

The cost of running the criminal, civil, and family courts in Ontario was about $310 million for 2001. These courts have the equivalent of 3300 full-time employees.

The simulation in the investigation models a **hypergeometric distribution**. Such distributions involve a series of *dependent* trials, each with success or failure as the only possible outcomes. The probability of success changes as each trial is made. The random variable is the number of successful trials in an experiment. Calculations of probabilities in a hypergeometric distribution generally require formulas using combinations.

Example 1 Jury Selection

a) Determine the probability distribution for the number of women on a civil-court jury selected from a pool of 8 men and 10 women.

b) What is the expected number of women on the jury?

Solution 1 Using Pencil and Paper

a) The selection process involves dependent events since each person who is already chosen for the jury cannot be selected again. The total number of ways the 6 jurors can be selected from the pool of 18 is

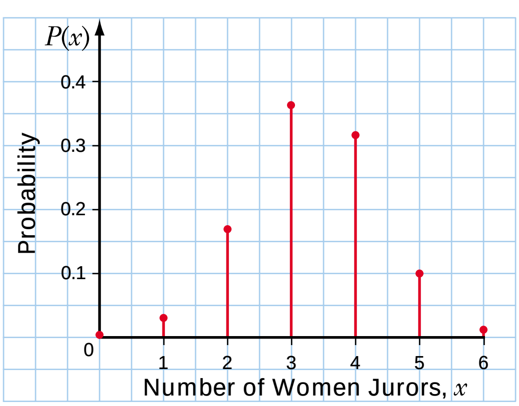
$$n(S) = {}_{18}C_6$$
$$= 18\ 564$$

This combination could also be written as $C(18, 6)$ or $\binom{18}{6}$.

There can be from 0 to 6 women on the jury. The number of ways in which x women can be selected is ${}_{10}C_x$. The men can fill the remaining $6 - x$ positions on the jury in ${}_8C_{6-x}$ ways. Thus, the number of ways of selecting a jury with x women on it is ${}_{10}C_x \times {}_8C_{6-x}$ and the probability of a jury with x women is

$$P(x) = \frac{n(x)}{n(S)}$$
$$= \frac{{}_{10}C_x \times {}_8C_{6-x}}{{}_{18}C_6}$$

Number of Women, x	Probability, $P(x)$
0	$\frac{{}_{10}C_0 \times {}_8C_6}{{}_{18}C_6} \doteq 0.001\ 51$
1	$\frac{{}_{10}C_1 \times {}_8C_5}{{}_{18}C_6} \doteq 0.030\ 17$
2	$\frac{{}_{10}C_2 \times {}_8C_4}{{}_{18}C_6} \doteq 0.169\ 68$
3	$\frac{{}_{10}C_3 \times {}_8C_3}{{}_{18}C_6} \doteq 0.361\ 99$
4	$\frac{{}_{10}C_4 \times {}_8C_2}{{}_{18}C_6} \doteq 0.316\ 74$
5	$\frac{{}_{10}C_5 \times {}_8C_1}{{}_{18}C_6} \doteq 0.108\ 60$
6	$\frac{{}_{10}C_6 \times {}_8C_0}{{}_{18}C_6} \doteq 0.011\ 31$



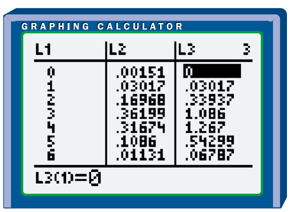
b) $$E(X) = \sum_{i=0}^{6} x_i P(x_i)$$
$$\doteq (0)(0.001\ 51) + (1)(0.030\ 17) + (2)(0.169\ 68) + (3)(0.361\ 99) + (4)(0.316\ 74) + (5)(0.108\ 60) + (6)(0.011\ 31)$$
$$\doteq 3.333\ 33$$

The expected number of women on the jury is approximately 3.333.

Solution 2 Using a Graphing Calculator

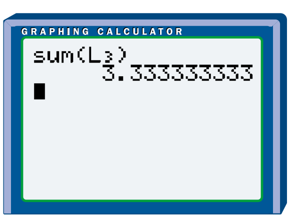
a) Enter the possible values for x, 0 to 6, in L1. Then, enter the formula for $P(x)$ in L2:

(10 nCr L1) × (8 nCr (6–L1)) ÷ (18 nCr 6)



b) Calculate $xP(x)$ in L3 using the formula L1 × L2.

QUIT to the home screen. You can find the expected number of women by using the sum(function in the LIST MATH menu.



The expected number of women on the jury is approximately 3.333.

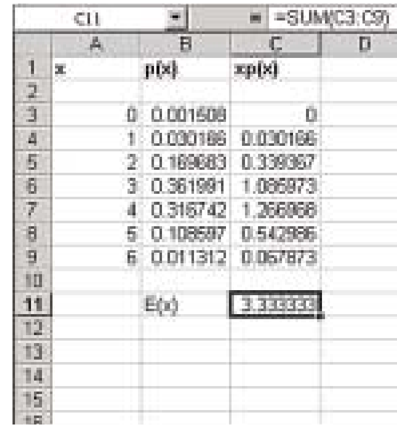
Solution 3 Using a Spreadsheet

a) Open a new spreadsheet. Create titles x, p(x), and xp(x) in columns A to C.

Enter the values of the random variable x in column A, ranging from 0 to 6. Next, use the combinations function to enter the formula for $P(x)$ in cell B3 and copy it to cells B4 through B9.

b) Calculate $xP(x)$ in column C by entering the formula A3*B3 in cell C3 and copying it to cells C4 through C9. Then, calculate the expected value using the **SUM function**.

The expected number of women on the jury is approximately 3.333.



C11 = =SUM(C3:C9)

	A	B	C	D
1	x	p(x)	xp(x)	
2				
3	0	0.001508	0	
4	1	0.030166	0.030166	
5	2	0.169683	0.339367	
6	3	0.361991	1.085973	
7	4	0.316742	1.266968	
8	5	0.108597	0.542986	
9	6	0.011312	0.067873	
10				
11		E(x)	3.333333	

Solution 4 Using Fathom™

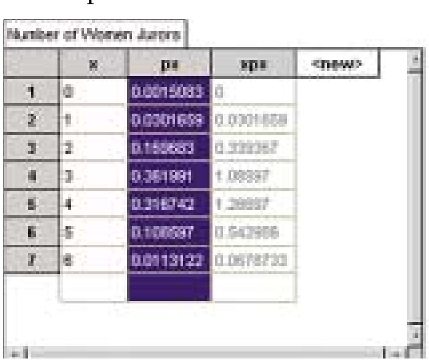
Open a new Fathom™ document. Drag a new **collection** box to the work area and name it Number of Women Jurors. Create seven new cases.

Drag a new **case table** to the work area. Create three new attributes: x, px, and xpx. Enter the values from 0 to 6 for the x attribute. Right-click on the px attribute, select Edit Formula, and enter

combinations(10,x)*combinations(8,6-x)/combinations(18,6)

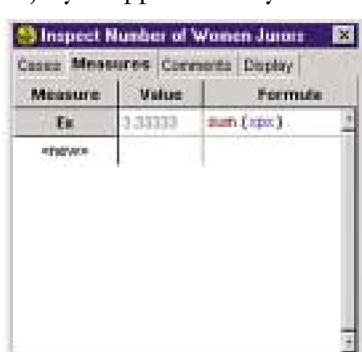
Similarly, calculate $xP(x)$ using the formula x*px. Next, double-click on the **collection** box to open the **inspector**. Select the Measures tab, and name a new measure Ex. Right-click on Ex and use the **sum function** to enter the formula sum(x*px).

The expected number of women on the jury is approximately 3.333.



Number of Women Jurors

	x	px	xpx	<new>
1	0	0.0015083	0	
2	1	0.0301659	0.0301659	
3	2	0.169683	0.339367	
4	3	0.361991	1.08597	
5	4	0.316742	1.26697	
6	5	0.108597	0.542986	
7	6	0.0113122	0.0678733	



Measure	Value	Formula
Ex	3.33333	sum (xpx)
<new>		

You can generalize the methods in Example 1 to show that for a hypergeometric distribution, the probability of x successes in r dependent trials is

Probability in a Hypergeometric Distribution

$$P(x) = \frac{{}_aC_x \times {}_{n-a}C_{r-x}}{{}_nC_r},$$

where a is the number of successful outcomes among a total of n possible outcomes.

Although the trials are dependent, you would expect the *average* probability of a success to be the same as the ratio of successes in the population, $\frac{a}{n}$. Thus, the expectation for r trials would be

Expectation for a Hypergeometric Distribution

$$E(X) = \frac{ra}{n}$$

This formula can be proven more rigorously by some challenging algebraic manipulation of the terms when $P(x) = \frac{{}_aC_x \times {}_{n-a}C_{r-x}}{{}_nC_r}$ is substituted into the equation for the expectation of any probability distribution, $E(X) = \sum_{i=1}^{n} x_i P(x_i)$.

Example 2 Applying the Expectation Formula

Calculate the expected number of women on the jury in Example 1.

Solution

$$\begin{aligned} E(X) &= \frac{ra}{n} \\ &= \frac{6 \times 10}{18} \\ &= 3.3\overline{3} \end{aligned}$$

The expected number of women jurors is $3.3\overline{3}$.

Example 3 Expectation of a Hypergeometric Distribution

A box contains seven yellow, three green, five purple, and six red candies jumbled together.

a) What is the expected number of red candies among five candies poured from the box?

b) Verify that the expectation formula for a hypergeometric distribution gives the same result as the general equation for the expectation of any probability distribution.

Solution

a) $n = 7 + 3 + 5 + 6 \quad r = 5 \quad a = 6$
$= 21$

Using the expectation formula for the hypergeometric distribution,

$$E(X) = \frac{ra}{n}$$
$$= \frac{5 \times 6}{21}$$
$$= 1.4285\ldots$$

One would expect to have approximately 1.4 red candies among the 5 candies.

b) Using the general formula for expectation,

$$E(X) = \sum xP(x)$$
$$= (0)\frac{{}_6C_0 \times {}_{15}C_5}{{}_{21}C_5} + (1)\frac{{}_6C_1 \times {}_{15}C_4}{{}_{21}C_5} + (2)\frac{{}_6C_2 \times {}_{15}C_3}{{}_{21}C_5} + (3)\frac{{}_6C_3 \times {}_{15}C_2}{{}_{21}C_5} + (4)\frac{{}_6C_4 \times {}_{15}C_1}{{}_{21}C_5} + (5)\frac{{}_6C_5 \times {}_{15}C_0}{{}_{21}C_5}$$
$$= 1.4285\ldots$$

Again, the expected number of red candies is approximately 1.4.

Example 4 Wildlife Management

In the spring, the Ministry of the Environment caught and tagged 500 raccoons in a wilderness area. The raccoons were released after being vaccinated against rabies. To estimate the raccoon population in the area, the ministry caught 40 raccoons during the summer. Of these 15 had tags.

a) Determine whether this situation can be modelled with a hypergeometric distribution.

b) Estimate the raccoon population in the wilderness area.

WEB CONNECTION

www.mcgrawhill.ca/links/MDM12

To learn more about sampling and wildlife, visit the above web site and follow the links. Write a brief description of some of the sampling techniques that are used.

Solution

a) The 40 raccoons captured during the summer were all different from each other. In other words, there were no repetitions, so the trials were dependent. The raccoons were either tagged (a success) or not (a failure). Thus, the situation does have all the characteristics of a hypergeometric distribution.

b) Assume that the number of tagged raccoons caught during the summer is equal to the expectation for the hypergeometric distribution. You can substitute the known values in the expectation formula and then solve for the population size, n.

Here, the number of raccoons caught during the summer is the number of trials, so $r = 40$. The number of tagged raccoons is the number of successes in the population, so $a = 500$.

$$E(X) = \frac{ra}{n}, \quad \text{so} \quad 15 \doteq \frac{40 \times 500}{n}$$

$$n \doteq \frac{40 \times 500}{15}$$

$$n \doteq 1333.3$$

The raccoon population in the wilderness area is approximately 1333.

Alternatively, you could assume that the proportion of tagged raccoons among the sample captured during the summer corresponds to that in the whole population. Then, $\frac{15}{40} = \frac{500}{n}$, which gives the same estimate for n as the calculation shown above.

Key Concepts

- A hypergeometric distribution has a specified number of dependent trials having two possible outcomes, success or failure. The random variable is the number of successful outcomes in the specified number of trials. The individual outcomes cannot be repeated within these trials.
- The probability of x successes in r dependent trials is $P(x) = \frac{{}_aC_x \times {}_{n-a}C_{r-x}}{{}_nC_r}$, where n is the population size and a is the number of successes in the population.
- The expectation for a hypergeometric distribution is $E(X) = \frac{ra}{n}$.
- To simulate a hypergeometric experiment, ensure that the number of trials is representative of the situation and that each trial is dependent (no replacement or resetting between trials). Record the number of successes and summarize the results by calculating probabilities and expectation.

Communicate Your Understanding

1. Describe how the graph in Example 1 differs from the graphs of the uniform, binomial, and geometric distributions.
2. Consider this question: What is the probability that 5 people out of a group of 20 are left handed if 10% of the population is left-handed? Explain why this situation does not fit a hypergeometric model. Rewrite the question so that you can use a hypergeometric distribution.

Practise

A

1. Which of these random variables have a hypergeometric distribution? Explain why.
   a) the number of clubs dealt from a deck
   b) the number of attempts before rolling a six with a die
   c) the number of 3s produced by a random-number generator
   d) the number of defective screws in a random sample of 20 taken from a production line that has a 2% defect rate
   e) the number of male names on a page selected at random from a telephone book
   f) the number of left-handed people in a group selected from the general population
   g) the number of left-handed people selected from a group comprised equally of left-handed and right-handed people

2. Prepare a table and a graph of a hypergeometric distribution with
   a) $n = 6, r = 3, a = 3$ b) $n = 8, r = 3, a = 5$

Apply, Solve, Communicate

B

3. There are five cats and seven dogs in a pet shop. Four pets are chosen at random for a visit to a children's hospital.
   a) What is the probability that exactly two of the pets will be dogs?
   b) What is the expected number of dogs chosen?

4. **Communication** Earlier this year, 520 seals were caught and tagged. On a recent survey, 30 out of 125 seals had been tagged.
   a) Estimate the size of the seal population.
   b) Explain why you cannot calculate the exact size of the seal population.

5. Of the 60 grade-12 students at a school, 45 are taking English. Suppose that 8 grade-12 students are selected at random for a survey.
   a) Develop a simulation to determine the probability that 5 of the selected students are studying English.
   b) Use the formulas developed in this section to verify your simulation results.

6. **Inquiry/Problem Solving** In a study of Canada geese, 200 of a known population of 1200 geese were caught and tagged. Later, another 50 geese were caught.
   a) Develop a simulation to determine the expected number of tagged geese in the second sample.
   b) Use the formulas developed in this section to verify your simulation results.

7. **Application** In a mathematics class of 20 students, 5 are bilingual. If the class is randomly divided into 4 project teams,
   a) what is the probability that a team has fewer than 2 bilingual students?
   b) what is the expected number of bilingual students on a team?

8. In a swim meet, there are 16 competitors, 5 of whom are from the Eastern Swim Club.
   a) What is the probability that 2 of the 5 swimmers in the first heat are from the Eastern Swim Club?
   b) What is the expected number of Eastern Swim Club members in the first heat?

9. The door prizes at a dance are four \$10 gift certificates, five \$20 gift certificates, and three \$50 gift certificates. The prize envelopes are mixed together in a bag, and five prizes are drawn at random.
   a) What is the probability that none of the prizes is a \$10 gift certificate?
   b) What is the expected number of \$20 gift certificates drawn?

10. A 12-member jury for a criminal case will be selected from a pool of 14 men and 11 women.

a) What is the probability that the jury will have 6 men and 6 women?

b) What is the probability that at least 3 jurors will be women?

c) What is the expected number of women?

11. Seven cards are dealt from a standard deck.

a) What is the probability that three of the seven cards are hearts?

b) What is the expected number of hearts?

12. A bag contains two red, five black, and four green marbles. Four marbles are selected at random, without replacement. Calculate

a) the probability that all four are black

b) the probability that exactly two are green

c) the probability that exactly two are green and none are red

d) the expected numbers of red, black, and green marbles

ACHIEVEMENT CHECK

Knowledge / Understanding	Thinking / Inquiry / Problem Solving	Communication	Application

13. A calculator manufacturer checks for defective products by testing 3 calculators out of every lot of 12. If a defective calculator is found, the lot is rejected.

a) Suppose 2 calculators in a lot are defective. Outline two ways of calculating the probability that the lot will be rejected. Calculate this probability.

b) The quality-control department wants to have at least a 30% chance of rejecting lots that contain only one defective calculator. Is testing 3 calculators in a lot of 12 sufficient? If not, how would you suggest they alter their quality-control techniques to achieve this standard? Support your answer with mathematical calculations.

C

14. Suppose you buy a lottery ticket for which you choose six different numbers between 1 and 40 inclusive. The order of the first five numbers is not important. The sixth number is a bonus number. To win first prize, all five regular numbers and the bonus number must match, respectively, the randomly generated winning numbers for the lottery. For the second prize, you must match the bonus number plus four of the regular numbers.

a) What is the probability of winning first prize?

b) What is the probability of winning second prize?

c) What is the probability of not winning a prize if your first three regular numbers match winning numbers?

15. **Inquiry / Problem Solving** Under what conditions would a binomial distribution be a good approximation for a hypergeometric distribution?

16. **Inquiry / Problem Solving** You start at a corner five blocks south and five blocks west of your friend. You walk north and east while your friend walks south and west at the same speed. What is the probability that the two of you will meet on your travels?

17. A research company has 50 employees, 20 of whom are over 40 years old. Of the 22 scientists on the staff, 12 are over 40. Compare the expected numbers of older and younger scientists in a randomly selected focus group of 10 employees.

Review of Key Concepts

7.1 Probability Distributions

Refer to the Key Concepts on page 374.

1. Describe the key characteristics of a uniform distribution. Include an example with your description.

2. James has designed a board game that uses a spinner with ten equal sectors numbered 1 to 10. If the spinner stops on an odd number, a player moves forward double that number of squares. However, if the spinner stops on an even number, the player must move back half that number of squares.

a) What is the expected move per spin?

b) Is this rule "fair"? Explain why you might want an "unfair" spinner rule in a board game.

3. Suppose a lottery has sold 10 000 000 tickets at $5.00 each. The prizes are as follows:

Prize	Number of Prizes
$2 000 000	1
$1 000	500
$100	10 000
$5	100 000

Determine the expected value of each ticket.

4. An environmental artist is planning to construct a rectangle with 36 m of fencing as part of an outdoor installation. If the length of the rectangle is a randomly chosen integral number of metres, what is the expected area of this enclosure?

5. Which die has the higher expectation?

a) an 8-sided die with its faces numbered 3, 6, 9, and so on, up to 24

b) a 12-sided die with its faces numbered 2, 4, 6, and so on, up to 24

7.2 Binomial Distributions

Refer to the Key Concepts on page 384.

6. Describe the key characteristics of a binomial distribution. Include an example with your description.

7. Cal's Coffee prints prize coupons under the rims of 20% of its paper cups. If you buy ten cups of coffee,

a) what is the probability that you would win at least seven prizes?

b) what is your expected number of prizes?

8. Use a table and a graph to display the probability distribution for the number of times 3 comes up in five rolls of a standard die.

9. A factory produces computer chips with a 0.9% defect rate. In a batch of 100 computer chips, what is the probability that

a) only 1 is defective?

b) at least 3 are defective?

10. A dart board contains 20 equally-sized sectors numbered 1 to 20. A dart is randomly tossed at the board 10 times.

a) What is the probability that the dart lands in the sector labelled 20 a total of 5 times?

b) What is the expected number of times the dart would land in a given sector?

11. Each question in a 15-question multiple-choice quiz has 5 possible answers. Suppose you guess randomly at each answer.

a) Show the probability distribution for the number of correct answers.

b) Verify the formula, $E(X) = np$, for the expectation of the number of correct answers.

7.3 Geometric Distributions

Refer to the Key Concepts on page 394.

12. Describe the key characteristics of a geometric distribution. Include an example with your description.

13. Your favourite TV station has ten minutes of commercials per hour. What is the expected number of times you could randomly select this channel without hitting a commercial?

14. A factory making printed-circuit boards has a defect rate of 2.4% on one of its production lines. An inspector tests randomly selected circuit boards from this production line.

a) What is the probability that the first defective circuit board will be the sixth one tested?

b) What is the probability that the first defective circuit board will be among the first six tested?

c) What is the expected waiting time until the first defective circuit board?

15. A computer has been programmed to generate a list of random numbers between 1 and 25.

a) What is the probability that the number 10 will not appear until the 6th number?

b) What is the expected number of trials until a 10 appears?

16. In order to win a particular board game, a player must roll, with two dice, the exact number of spaces remaining to reach the end of the board. Suppose a player is two spaces from the end of the board. Show the probability distribution for the number of rolls required to win, up to ten rolls.

7.4 Hypergeometric Distributions

Refer to the Key Concepts on page 403.

17. Describe the key characteristics of a hypergeometric distribution. Include an example with your description.

18. Of the 15 students who solved the challenge question in a mathematics contest, 8 were enrolled in mathematics of data management. Five of the solutions are selected at random for a display.

a) Prepare a table and graph of the probability distribution for the number of solutions in the display that were prepared by mathematics of data management students.

b) What is the expected number of solutions in the display that were prepared by mathematics of data management students?

19. Seven cards are randomly dealt from a standard deck. Show the probability distribution of the number of cards dealt that are either face cards or aces.

20. One summer, conservation officials caught and tagged 98 beavers in a river's flood plain. Later, 50 beavers were caught and 32 had been tagged. Estimate the size of the beaver population.

21. Suppose that 48 of the tagged beavers in question 20 were males.

a) Develop a simulation to estimate the probabilities for the number of tagged males among the 32 beavers captured a second time.

b) Verify the results of your simulation with mathematical calculations.

Chapter Test

ACHIEVEMENT CHART				
Category	Knowledge / Understanding	Thinking / Inquiry / Problem Solving	Communication	Application
Questions	All	5, 10, 12	1, 12	3, 4, 6–8, 10–12

1. Determine if a uniform, binomial, geometric, or hypergeometric distribution would be the best model for each of the following experiments. Explain your reasoning.

a) drawing names out of a hat without replacement and recording the number of names that begin with a consonant

b) generating random numbers on a calculator until it displays a 5

c) counting the number of hearts in a hand of five cards dealt from a well-shuffled deck

d) asking all students in a class whether they prefer cola or ginger ale

e) selecting the winning ticket in a lottery

f) predicting the expected number of heads when flipping a coin 100 times

g) predicting the number of boys among five children randomly selected from a group of eight boys and six girls

h) determining the waiting time before picking a winning number in a lottery.

2. A lottery ticket costs \$2.00 and a total of 4 500 000 tickets were sold. The prizes are as follows:

Prize	Number of Prizes
\$500 000	1
\$50 000	2
\$5 000	5
\$500	20
\$50	100

Determine the expected value of each ticket.

3. Of 25 people invited to a birthday party, 5 prefer vanilla ice cream, 8 prefer chocolate, and 4 prefer strawberry. The host surveys 6 of these people at random to determine how much ice cream to buy.

a) What is the probability that at least 3 of the people surveyed prefer chocolate ice cream?

b) What is the probability that none prefer vanilla?

c) What is the expected number of people who prefer strawberry?

d) What is the expected number of people who do not have a preference for any of the three flavours?

4. Suppose you randomly choose an integer n between 1 and 5, and then draw a circle with a radius of n centimetres. What is the expected area of this circle to the nearest hundredth of a square centimetre?

5. At the Statsville County Fair, the probability of winning a prize in the ring-toss game is 0.1.

a) Show the probability distribution for the number of prizes won in 8 games.

b) If the game will be played 500 times during the fair, how many prizes should the game operators keep in stock?

6. a) What is the probability that a triple will occur within the first five rolls of three dice?

b) What is the expected waiting time before a triple?

7. In July of 2000, 38% of the population of Canada lived in Ontario. Design a simulation to estimate the expected number of residents of Ontario included in a random survey of 25 people in Canada.

8. A multiple-choice trivia quiz has ten questions, each with four possible answers. If someone simply guesses at each answer,
   a) what is the probability of only one or two correct guesses?
   b) what is the probability of getting more than half the questions right?
   c) what is the expected number of correct guesses?

9. In an experiment, a die is rolled repeatedly until all six faces have finally shown.
   a) What is the probability that it only takes six rolls for this event to occur?
   b) What is the expected waiting time for this event to occur?

10. The Burger Barn includes one of three different small toys with every Barn Burger. Each of the toys is equally likely to be included with a burger. Design a simulation to determine the number of Barn Burger purchases necessary for 3:1 odds of collecting all three different toys.

11. To determine the size of a bear population in a provincial park, 23 bears were caught and fitted with radio collars. One month later, 8 of 15 bears sighted had radio collars. What is the approximate size of the bear population?

ACHIEVEMENT CHECK

Knowledge / Understanding	Thinking / Inquiry / Problem Solving	Communication	Application

12. Louis inserts a 12-track CD into a CD player and presses the random play button. This CD player's random function chooses each track independently of any previously played tracks.
   a) What is the probability that the CD player will select Louis's favourite track first?
   b) What is the probability that the second selection will not be his favourite track?
   c) What is the expected waiting time before Louis hears his favourite track?
   d) Sketch a graph of the probability distribution for the waiting times.
   e) Explain how having a different number of tracks on the CD would affect the graph in part d).
   f) If Louis has two favourite tracks, what is the expected waiting time before he hears both tracks?