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.

Project Prep

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

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.

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.

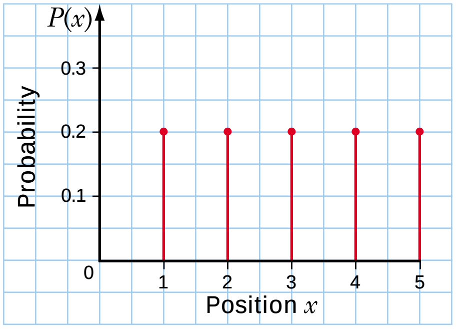
Example 2 Uniform Probability Distribution

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

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

Recall that the capital sigma, Σ, 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

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

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.

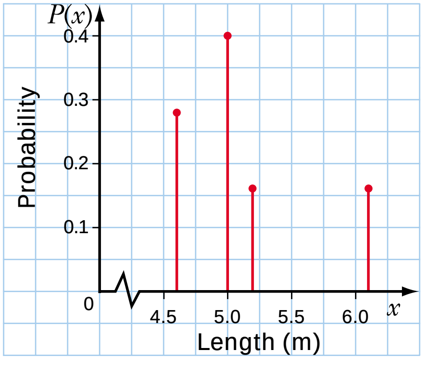
a) Show the probability distribution for the length of an assigned canoe.

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

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

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

- 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?

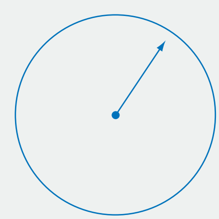
ACHIEVEMENT CHECK

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

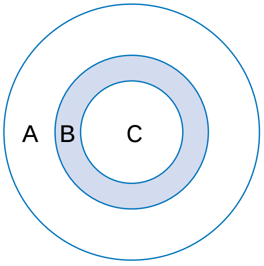
iv) 6-number group, payout ratio 5:1

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.