# 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.

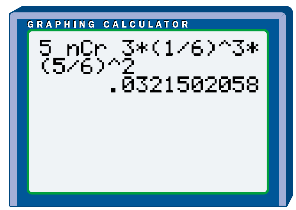
$$\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^3$$

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.

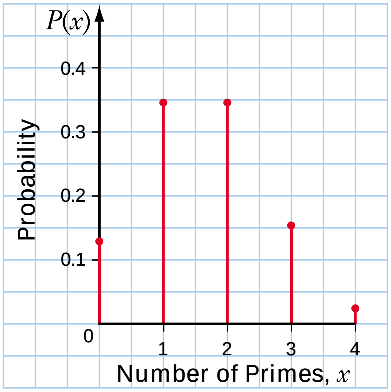
**b)** What is the expected number of these prime digits in the new telephone number?

### *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,
$$p = 0.4 \quad \text{and} \quad q = 1 - 0.4 = 0.6$$

| 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.

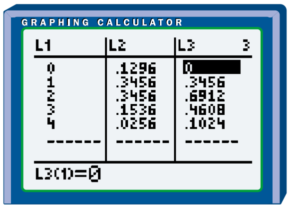
### *Solution 2 Using a Graphing Calculator*

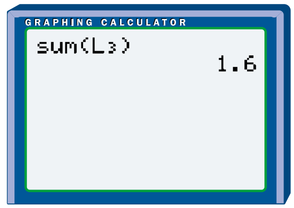
**a)** Check that lists L1, L2, and L3 are clear. Enter all the possible values for $x$ into L1.

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*

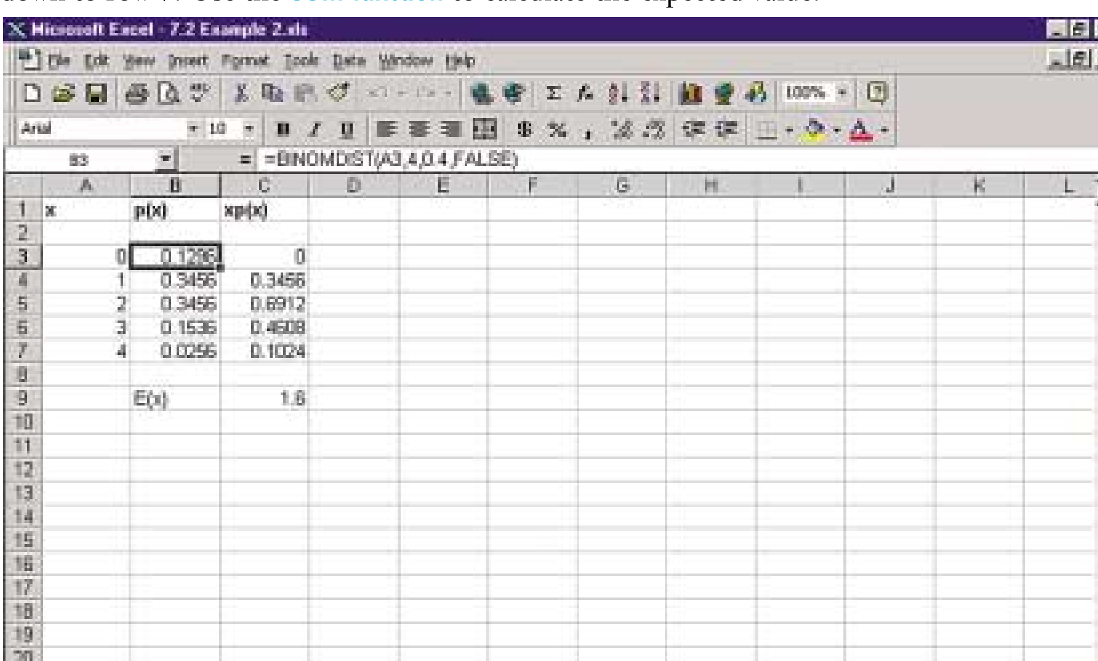
**a)** Open a new spreadsheet. Create headings for $x$, $P(x)$, and $xP(x)$ in columns A to C. Enter the possible values of the random variable $x$ in column A. Use the BINOMDIST function to calculate the probabilities for the different values of $x$. The syntax for this function is

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.

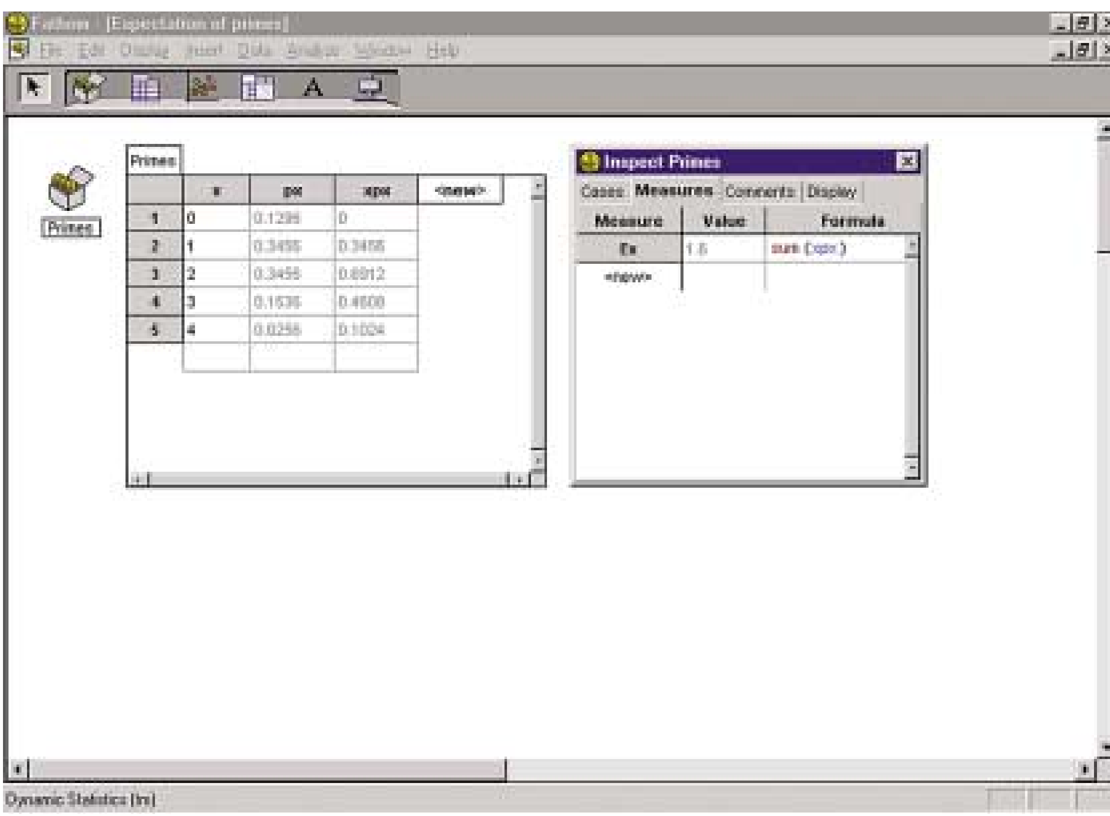
Drag a new **case table** to the work area. Create three new attributes: x, px, and xpx. Enter the values from 0 to 4 in the x attribute column. Then, use the **binomialProbability function** to calculate the probabilities for the different values of $x$. The syntax of this function is

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.