CHAPTER 11
The binomial distribution

eBook*plus*

DIGITAL DOC
doc-9226
10 Quick Questions

CHAPTER CONTENTS

studyon

Units:	3 & 4
AOS:	4
Topic:	2
Concept:	1

Concept summary
Read a summary of this concept.

See more
Watch a video about binomial distributions.

11A The binomial distribution

The binomial distribution is an example of a particular type of *discrete probability distribution.* It has relevance and importance in many real-life everyday applications. This particular branch of mathematics moves away from the textbook and the classroom and into the areas of medical research, simulation activities and business applications such as quality control.

The binomial distribution may be referred to as a *Bernoulli distribution*, and the trials conducted are known as *Bernoulli trials*. They were named in honour of the Swiss mathematician Jakob Bernoulli (1654–1705).

Bernoulli trials and sequences

A Bernoulli trial is an experiment in which the outcome is either a *success* or a *failure*.
A Bernoulli sequence is a sequence of Bernoulli trials in which:
1. the probability of each possible outcome is *independent* of the results of the previous trial
2. the probability of each possible outcome is the *same* for each trial.
Note: In a Bernoulli sequence, the number of successes follows the binomial distribution.

WORKED EXAMPLE 1

Determine which of the following sequences can be defined as Bernoulli sequences.
a Rolling an 8-sided die numbered 1 to 8 forty times and recording the number of 6s obtained
b Drawing a card from a fair deck with replacement and recording the number of aces
c Rolling a die 60 times and recording the number that is obtained

THINK

a Check that all the characteristics have been satisfied for a Bernoulli sequence.

WRITE

a Yes, this is an example of a Bernoulli sequence, as there are two possible outcomes for each trial (success is 'obtaining a 6' and failure is 'not obtaining a 6'), the outcome of each trial is independent of the outcome of previous trials, and the probability of success is the same for each trial.

b Check that all the characteristics have been satisfied for a Bernoulli sequence.

b Yes, this is an example of a Bernoulli sequence, as there are two possible outcomes for each trial (success is 'obtaining an ace' and failure is 'not obtaining an ace'), the outcome of each trial is independent of the outcome of previous trials, and the probability of success is the same for each trial (as each time a card is drawn it is replaced).

c Check that all the characteristics have been satisfied for a Bernoulli sequence.

c This is not an example of a Bernoulli sequence, since 'success' has not been defined.

If X represents a random variable that has a binomial distribution, then it can be expressed as:

$$X \sim \text{Bi}(n, p) \text{ or } X \sim \text{B}(n, p).$$

Translated into words, $X \sim \text{Bi}(n, p)$ means that X follows a binomial distribution with parameters n (the number of trials) and p (the probability of success).

Consider the experiment where a fair die is rolled four times. If X represents the number of times a 3 appears uppermost, then X is a binomial variable. Obtaining a 3 will represent a success and all other values will represent a failure. The die is rolled four times so the number of trials, n, equals 4 and the probability, p, of obtaining a 3 is equal to $\frac{1}{6}$. Using the shorthand notation, $X \sim \text{Bi}(n, p)$ becomes $X \sim \text{Bi}(4, \frac{1}{6})$.

We will now determine the probability of a 3 appearing uppermost 0, 1, 2, 3 and 4 times. Obtaining 3 is defined as a success and is denoted by S. All other numbers are defined as a failure and are denoted by F. The possible outcomes are listed in the table below.

Note that $q = 1 - p$; the probability of a failure.

Occurrence of 3	Possible outcomes	Probability	
0	*FFFF*	$1 \times \left(\frac{5}{6}\right)^4 = \frac{625}{1296}$	q^4
1	*SFFF FSFF FFSF FFFS*	$4 \times \left(\frac{5}{6}\right)^3 \left(\frac{1}{6}\right) = \frac{500}{1296}$	$4q^3p$
2	*SSFF SFSF SFFS FSSF FSFS FFSS*	$6 \times \left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right)^2 = \frac{150}{1296}$	$6q^2p^2$
3	*SSSF SSFS SFSS FSSS*	$4 \times \left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^3 = \frac{20}{1296}$	$4qp^3$
4	*SSSS*	$1 \times \left(\frac{1}{6}\right)^4 = \frac{1}{1296}$	p^4

It is interesting to note that the binomial probability distribution is closely related to the binomial theorem (see far right-hand column). Furthermore, if we examine the coefficients of the terms — that is, 1, 4, 6, 4, 1 — it is evident that they are the entries of Pascal's triangle.

This procedure for determining the individual probabilities can become tedious, particularly once the number of trials increases. Hence if X is a binomial random variable, its probability distribution is defined as follows.

$\Pr(X = x) = {}^nC_x p^x q^{n-x}$ where $x = 0, 1, 2, \ldots n$. That is:

x = the occurrence of the successful outcome.

The formula may also be written as:

$\Pr(X = x) = {}^nC_x\, p^x (1 - p)^{n-x}$ where $x = 0, 1, 2, \ldots n$.

Here, the probability of failure, q, is replaced by $1 - p$.

${}^nC_x$ represents the number of ways that x different outcomes can be obtained from n trials. It can also be written as $\binom{n}{x}$ and has the formula ${}^nC_x = \dfrac{n!}{x!(n-x)!}$.

Since this is a probability distribution, we would expect that the sum of the probabilities is 1. Therefore, for the previous example:

$$\Pr(X = x) = \Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2) + \Pr(X = 3) + \Pr(X = 4)$$
$$= \frac{625}{1296} + \frac{500}{1296} + \frac{150}{1296} + \frac{20}{1296} + \frac{1}{1296}$$
$$= 1.$$

WORKED EXAMPLE 2

eBook*plus*

TUTORIAL
eles-1191
Worked example 2

A binomial variable, X, has the probability function $\Pr(X = x) = {}^6C_x(0.4)^x(0.6)^{6-x}$ where $x = 0, 1, \ldots 6$. Find:

a n, the number of trials

b p, the probability of success

c the probability distribution for x as a table.

THINK	WRITE
a Obtain the relevant information from the given function. The number of trials, n, is the value of the number located at the top left-hand corner of C.	**a** $n = 6$
b Obtain the relevant information from the given function. The probability of success, p, is the value in the first bracket.	**b** $p = 0.4$
c 1 Write the rule of the given probability function.	**c** $\Pr(X = x) = {}^6C_x(0.4)^x(0.6)^{6-x}$
2 Substitute $x = 0$ into the rule.	$\Pr(X = 0) = {}^6C_0(0.4)^0(0.6)^6$
3 Evaluate.	$= 1 \times 1 \times 0.046\,656$ $= 0.046\,656$
4 Substitute $x = 1$ into the rule.	$\Pr(X = 1) = {}^6C_1(0.4)(0.6)^5$
5 Evaluate.	$= 6 \times 0.4 \times 0.077\,76$ $= 0.186\,624$
6 Substitute $x = 2$ into the rule.	$\Pr(X = 2) = {}^6C_2(0.4)^2(0.6)^4$
7 Evaluate.	$= 15 \times 0.16 \times 0.1296$ $= 0.311\,04$
8 Substitute $x = 3$ into the rule.	$\Pr(X = 3) = {}^6C_3(0.4)^3(0.6)^3$
9 Evaluate.	$= 20 \times 0.064 \times 0.216$ $= 0.276\,48$
10 Substitute $x = 4$ into the rule.	$\Pr(X = 4) = {}^6C_4(0.4)^4(0.6)^2$
11 Evaluate.	$= 15 \times 0.0256 \times 0.36$ $= 0.138\,24$
12 Substitute $x = 5$ into the rule.	$\Pr(X = 5) = {}^6C_5(0.4)^5(0.6)$
13 Evaluate.	$= 6 \times 0.010\,24 \times 0.6$ $= 0.036\,864$
14 Substitute $x = 6$ into the rule.	$\Pr(X = 6) = {}^6C_6(0.4)^6(0.6)^0$
15 Evaluate.	$= 1 \times 0.004\,096 \times 1$ $= 0.004\,096$
16 Place values in a table. *Note:* The table at right clearly displays the probability distribution of the given function. It also shows that the individual probabilities sum to 1.	

x	0	1	2	3	4	5	6
$\Pr(X = x)$	0.0467	0.1866	0.3110	0.2765	0.1382	0.0369	0.0041

WORKED EXAMPLE 3

A fair die is rolled five times. Find the probability of obtaining:

a exactly four 5s
b exactly two even numbers
c all results greater than 3
d a 5 on the first roll only
e a 5 on the second and third roll only.

THINK	WRITE
a 1 Check that all the characteristics have been satisfied for a binomial distribution.	**a** Binomial distribution — n independent trials and two outcomes, fixed p and q.
2 Write the rule for the binomial probability distribution.	$\Pr(X = x) = {}^nC_x p^x q^{n-x}$
3 Define and assign values to variables. The number of 5s obtained is exactly four.	$n = 5$ Let X = the number of 5s obtained. That is, $x = 4$ p = probability of a 5 $= \frac{1}{6}$ $q = \frac{5}{6}$
4 Substitute the values into the rule.	$\Pr(X = 4) = {}^5C_4 \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)$
5 Evaluate.	$= 5 \times \frac{1}{1296} \times \frac{5}{6}$ $= \frac{25}{7776}$
b 1 Define and assign values to variables. Two even numbers means we have $x = 2$.	**b** $n = 5$ Let X = the number of even numbers. That is, $x = 2$ p = probability of an even number $= \frac{1}{2}$ $q = \frac{1}{2}$
2 Substitute the values into the rule.	$\Pr(X = 2) = {}^5C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3$
3 Evaluate.	$= 10 \times \frac{1}{4} \times \frac{1}{8}$ $= \frac{10}{32}$
4 Simplify.	$= \frac{5}{16}$
c 1 Define and assign values to variables. We require 5 occasions when results are greater than 3.	**c** $n = 5$ Let X = values greater than 3. That is, $x = 5$ Three outcomes — 4, 5, 6 out of six are greater than 3. p = probability of number greater than 3 $= \frac{1}{2}$ $q = \frac{1}{2}$
2 Substitute the values into the rule.	$\Pr(X = 5) = {}^5C_5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^0$
3 Evaluate.	$= 1 \times \frac{1}{32} \times 1$ $= \frac{1}{32}$

d Since a specific order is required here, the binomial rule is not required.

d Pr(5 on the first roll only)

$$= \Pr(SFFFF)$$
$$= \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6}$$
$$= \frac{625}{7776}$$

e Since a specific order is required here, the binomial rule is not required.

e Pr(5 on the second and third roll only)

$$= \Pr(FSSFF)$$
$$= \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6}$$
$$= \frac{125}{7776}$$

Note: If the rule for the binomial probability distribution were to be used in part **d**, it would provide an answer of $5 \times \frac{625}{7776} = \frac{3125}{7776}$. This answer gives the probability of obtaining a 5 once on any of the five trials, not necessarily on the first roll only.

Hence, if a specific order is required the rule for the binomial probability distribution should not be used.

Similarly in part **e** the rule would produce an answer of $10 \times \frac{125}{7776} = \frac{1250}{7776}$, giving the probability of obtaining a 5 twice on any of the five trials.

WORKED EXAMPLE 4

A new drug for hay fever is known to be successful in 40% of cases. Ten hay fever sufferers take part in the testing of the drug. Find the probability, correct to 4 decimal places, that:

a four people are cured **b no people are cured**

c all 10 are cured.

THINK	WRITE
a 1 Check that all the characteristics have been satisfied for a binomial distribution.	**a** This is a binomial distribution with n independent trials and two outcomes, p and q.
2 Write the rule for the binomial probability distribution.	$\Pr(X = x) = {}^nC_x p^x q^{n-x}$
3 Define and assign values to variables.	$n = 10$, $p = 0.4$, $q = 0.6$ Let X = the number of people cured, therefore $x = 4$
4 Substitute the values into the rule.	$\Pr(X = 4) = {}^{10}C_4(0.4)^4\,(0.6)^6$
5 Evaluate.	$= 210 \times 0.0256 \times 0.046\,656$ $= 0.250\,822\,656$
6 Round the answer to 4 decimal places.	≈ 0.2508
7 A CAS calculator can also be used to calculate the probability for a particular x-value. Define the variables.	$n = 10$ $p = 0.4$ $X \sim \text{Bi}(10, 0.4)$ $x = 4$
8 Use the binomial Pdf feature of a CAS calculator. Enter $n = 10$, $p = 0.4$, $x = 4$.	binom Pdf (10, 0.4, 4)
9 Record the result.	0.250 823
10 Answer the question.	The probability that exactly 4 people are cured is 0.2508.

b 1 In this case $n = 10$, $p = 0.4$, $x = 0$ Use a CAS calculator to determine the probability.	**b** binom Pdf (10, 0.4, 0)
2 Record the result.	0.006 046 6
3 Answer the question.	The probability that exactly 0 people are cured is 0.0060
c 1 In this case $n = 10$, $p = 0.4$, $x = 10$ Use a CAS calculator to determine the probability.	**c** binom Pdf (10, 0.4, 10)
2 Record the result.	0.000 104 86
3 Answer the question.	The probability that exactly 10 people are cured is 0.0001.

WORKED EXAMPLE 5

eBook*plus*

TUTORIAL
eles-1232
Worked example 5

Grant is a keen darts player and knows that his chance of scoring a bullseye on any one throw is 0.3.

a If Grant takes 6 shots at the target, find the probability, correct to 4 decimal places, that he:
i misses the bullseye each time **ii hits the bullseye at least once.**

b Find the number of throws Grant would need to ensure a probability of more than 0.8 of scoring at least one bullseye.

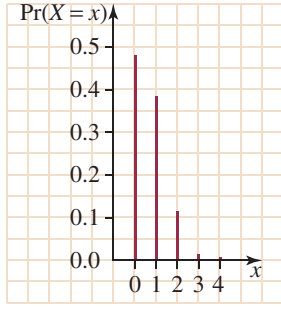
THINK	WRITE
a i 1 Check that all the characteristics have been satisfied for a binomial distribution.	**a i** This is a binomial distribution with n independent trials and two outcomes, p and q.
2 Write the rule for the binomial probability distribution.	$\Pr(X = x) = {}^nC_x p^x q^{n-x}$
3 Define and assign values to variables.	$n = 6$ Let X = the number of bullseyes, therefore $x = 0, 1, 2, 3, 4, 5, 6$ $p = 0.3$ $q = 0.7$
4 Substitute the values into the rule.	$\Pr(X = x) = {}^nC_x p^x q^{n-x}$
5 Evaluate and round answer to 4 decimal places. *Note:* Check the answer using binomPdf(6,0.3,0).	$\Pr(X = 0) = {}^6C_0(0.3)^0(0.7)^6$ $= 1 \times 1 \times 0.117\,649$ $= 0.117\,649$ ≈ 0.1176
6 Answer the question.	The probability that Grant misses the bullseye each time is 0.1176.
ii 1 Define and assign values to variables. *Note:* $\Pr(X \geq 1)$ would involve adding probabilities from $\Pr(X = 1)$ to $\Pr(X = 6)$. Using the fact that $\Pr(X \geq 1) = 1 - \Pr(X = 0)$ allows us to solve the problem using fewer terms.	**ii** $n = 6$ Let X = the number of bullseyes, therefore $x = 0, 1, 2, 3, 4, 5, 6$ $p = 0.3$ $q = 0.7$
2 Substitute the values into the rule.	$\Pr(X \geq 1) = \Pr(X = 1) + \Pr(X = 2)$ $+ \ldots + \Pr(X = 6)$ $= 1 - \Pr(X = 0)$ $= 1 - {}^6C_0(0.3)^0(0.7)^6$
3 Evaluate and round answer to 4 decimal places. *Note:* Check the answer using 1 – binomPdf(6,0.3,0) on a CAS calculator.	$= 1 - 0.117\,649$ $= 0.882\,351$ ≈ 0.8824

4	Answer the question.	The probability that Grant hits the bullseye at least once is 0.8824.
b 1	Define and assign values to variables.	**b** $n = ?$ Let X = the number of bullseyes, therefore $x = 0, 1, 2, 3, 4, 5, 6$ $p = 0.3$ $q = 0.7$ $\Pr(X \geq 1) > 0.8$
2	Write the rule as required.	$\Pr(X \geq 1) = \Pr(X = 1) + \Pr(X = 2) + \ldots + \Pr(X = 6)$ $= 1 - \Pr(X = 0)$ $\therefore\ 1 - \Pr(X = 0) > 0.8$
3	Substitute the values into the rule.	$1 - {}^nC_0(0.3)^0(0.7)^n > 0.8$
4	Evaluate by solving for n.	$1 - 0.7^n > 0.8$ $1 - 0.8 > 0.7^n$ $0.2 > 0.7^n$ $\log_e(0.2) > \log_e(0.7^n)$ $\log_e(0.2) > n \times \log_e(0.7)$ $\dfrac{\log_e(0.2)}{\log_e(0.7)} < n \quad$ (since $\log_e(0.7) < 0$) $n > 4.512\,338\,026$
5	Interpret the results and answer the question.	Grant would need to take 5 shots to ensure a probability of 0.8 of scoring at least one bullseye.
6	Alternatively, a CAS calculator can be used to solve the inequation. $1 - (0.7)^n > 0.8$ Use the calculator's solve feature.	solve $(1 - (0.7)^n > 0.8, n)$
7	Write the result.	$n > 4.51\,234$
8	Write the solution.	Solving $1 - 0.7^n > 0.8$ for n implies $n > 4.51\,234$
9	Interpret the results and answer the question.	Grant would need to take 5 shots to ensure a probability of 0.8 of scoring at least one bullseye.

Graphs of the binomial distribution

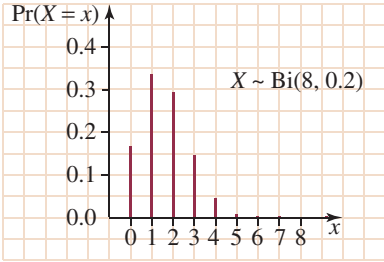
We will now consider the graph of a binomial distribution. If we refer to the example of obtaining a 3 when rolling a die four times (see the table on page 516) we note that $X \sim \text{Bi}(4, \frac{1}{6})$. The probability distribution of the random variable, X, is given in the table and graph over the page.

x	$\Pr(X = x)$
0	0.4823
1	0.3858
2	0.1157
3	0.0154
4	0.0008

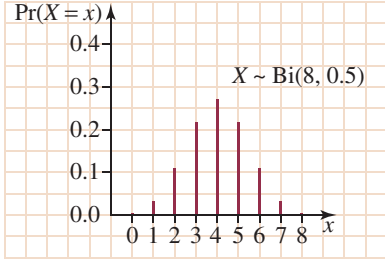


The effect of changing *n* and *p* on binomial distribution graphs

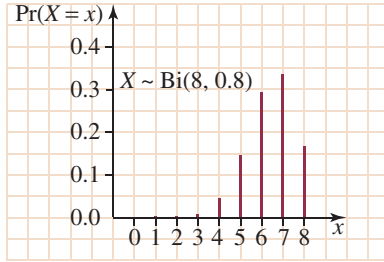
The effect of increasing p can be seen in the series of graphs below.



The graph is positively skewed.



The graph is symmetrical.



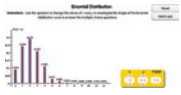
The graph is negatively skewed.

study on

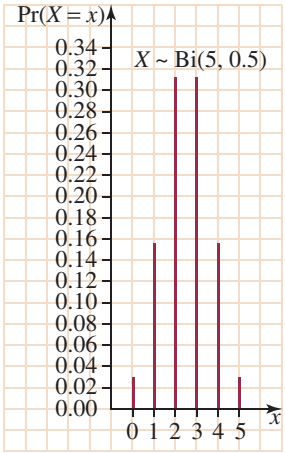
Units:	3 & 4
AOS:	4
Topic:	2
Concept:	3

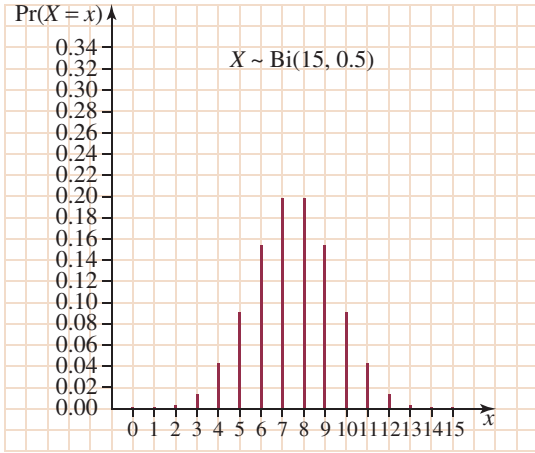
Concept summary
Read a summary of this concept.

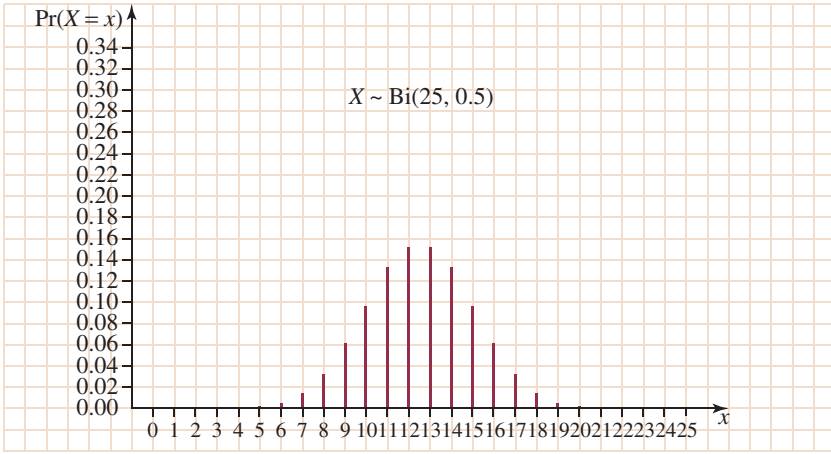
Do more
Interact with binomial distributions.



We will now keep p fixed and vary n.





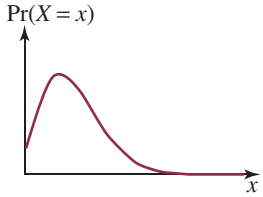


From the preceding graphs it can be seen that:

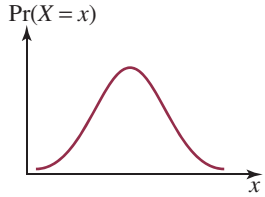
1. when $p = 0.5$, the graph is symmetrical
2. as n increases, and the interval between the vertical columns decreases, the graph approximates a smooth hump or bell shape.

The effects that the parameters n and p have on the binomial probability distribution curve can be summarised in the following way.

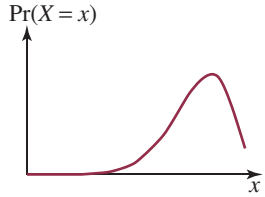
If $p < 0.5$, the graph is positively skewed.



If $p = 0.5$, the graph is symmetrical.



If $p > 0.5$, the graph is negatively skewed.



When n is very large and $p = 0.5$, the vertical columns are closer together and the line graph becomes a bell-shaped curve or a normal distribution curve.

Exercise 11A The binomial distribution

1 **WE1** Determine which of the following sequences can be defined as a Bernoulli sequence:
- **a** Rolling a die 10 times and recording the number that comes up
- **b** Rolling a die 10 times and recording the number of 3s that come up
- **c** Spinning a spinner numbered 1 to 10 and recording the number that is obtained
- **d** Tossing a coin 15 times and recording the number of tails obtained
- **e** Drawing a card from a fair deck, without replacement, and recording the number of picture cards
- **f** Drawing a card from a fair deck, with replacement, and recording the number of black cards
- **g** Selecting three marbles from a jar containing three yellow marbles and two black marbles, without replacement.

2 Evalute the following, correct to 4 decimal places:
- **a** ${}^{7}C_2\,(0.4)^2\,(0.6)^5$
- **b** ${}^{9}C_3\,(0.1)^3\,(0.9)^6$
- **c** ${}^{10}C_5\,(0.5)^5\,(0.5)^5$
- **d** ${}^{8}C_5\,(0.2)^5\,(0.8)^3$
- **e** ${}^{9}C_7\left(\frac{1}{3}\right)^7\left(\frac{2}{3}\right)^2$
- **f** ${}^{10}C_0\,(0.15)^0\,(0.85)^{10}$.

eBook plus
DIGITAL DOC
doc-9227
Spreadsheet
Binomial distribution

3 **WE2** A binomial variable, X, has the probability function $\Pr(X = x) = {}^{5}C_x\,(0.3)^x\,(0.7)^{5-x}$ where $x = 0, 1, \ldots, 5$. Find:
- **a** n, the number of trials
- **b** p, the probability of success
- **c** the probability distribution for x as a table.

4 Twenty per cent of items made by a certain machine are defective. The items are packed and sold in boxes of 5. What is the probability of 4 items being defective in a box?

5 Alex lives so close to where she works that she only has a 0.1 chance of being late. What is the probability that she is late on 3 out of 4 days?

6 Ange has four chances to knock an empty can off a stand by throwing a ball. On each throw, the probability of success is $\frac{1}{3}$. Find the probability that she will knock the empty can off the stand:
- **a** once
- **b** twice
- **c** at least once.

7 **WE3** A fair coin is tossed four times. Find the probability of obtaining:
- **a** heads on the first two tosses and tails on the second two
- **b** heads on every roll
- **c** two heads and two tails.

8 Peter is quite poor at doing crossword puzzles and the probability of him completing one is 0.2. Find the probability that:
- **a** of the next three crossword puzzles that he attempts, he is successful in completing two
- **b** he successfully completes the first three crossword puzzles that he tries, but has no luck on the next one
- **c** he successfully completes the first three crossword puzzles that he tries
- **d** he successfully completes the first three puzzles that he tries given that he was successful in completing his first two.

9 A weighted coin is biased such that a tail comes up 60% of the time. The coin is tossed five times. Find the probability of obtaining:
- **a** tails on the first four tosses only
- **b** four tails.

10 **WE4** Fifty-five per cent of the local municipality support the local council. If eight people are selected at random, find the probability, correct to 4 decimal places, that:

a half support the council
b all eight support the council
c five support the council
d three oppose the council.

11 The probability of Colin beating Maria at golf is 0.4. If they play once a week throughout the entire year and the outcome of each game is independent of any other, find the probability that they will have won the same number of matches, correct to 4 decimal places.

12 It is known that 5 out of every 8 people eat Superflakes for breakfast. Find the probability that half of a random sample of 20 people surveyed eat Superflakes, correct to 4 decimal places.

13 On a certain evening, during a ratings period, two television stations put their best shows on against each other. The ratings showed that 39% of people watched Channel 6, while only 30% of people watched Channel 8. The rest watched other channels. A random sample of 10 people were surveyed the next day. Find the probability, correct to 4 decimal places, that exactly:

a six watched Channel 6
b four watched Channel 8.

14 Three per cent of items produced by a certain machine are defective. A random sample of 10 items is taken. Find the probability that exactly 10% are defective, correct to 4 decimal places.

15 A new drug being trialled gives 8% of the patients a violent reaction. If 200 patients trial the drug, find the probability that 12 patients have a violent reaction to the drug, correct to 4 decimal places.

16 A pen company produces 30 000 pens per week. However, of these 30 000 pens, 600 are defective. Pens are sold in boxes of 20. Find the probability, correct to 4 decimal places, that:

a two defective pens are found in one box
b a box is free from defective pens.

17 A box contains 5 red marbles, 3 blue marbles and 2 yellow marbles. A marble is chosen at random and replaced. This selection process is completed eight times. Find the probability, correct to 4 decimal places, that:

a exactly 4 reds are selected
b exactly 2 blues are selected
c no yellows are selected.

18 **MC** Claire's position in the netball team is goal shooter. The probability of her shooting a goal is 78%. If Claire has 10 attempts at scoring, the probability she will score fewer than 3 goals is:

A ${}^{10}C_3\,(0.78)^3(0.22)^7$
B ${}^{10}C_3\,(0.78)^3(0.22)^7 + {}^{10}C_4\,(0.78)^4(0.22)^6 + \ldots + (0.78)^{10}$
C ${}^{10}C_2\,(0.78)^2(0.22)^8 + {}^{10}C_1\,(0.78)^1(0.22)^9 + (0.22)^{10}$
D ${}^{10}C_2\,(0.78)^2(0.22)^8 + {}^{10}C_1\,(0.78)^1(0.22)^9 + (0.78)^{10}$
E ${}^{10}C_2\,(0.22)^2(0.78)^8 + {}^{10}C_1\,(0.22)^1(0.78)^9 + (0.78)^{10}$

19 **MC** The probability that the temperature in Melbourne will rise above 25 °C on any given summer day, independent of any other summer day, is 0.6. The probability that four days in a week reach in excess of 25 °C is:

A $0.6^4 \times 0.4^3$
B $7 \times 0.6^4 \times 0.4^3$
C $\frac{4}{7} \times 0.4^3$
D 0.6^4
E $35 \times 0.6^4 \times 0.4^3$

20 **MC** Rachel sits a multiple-choice test containing 20 questions, each with four possible answers. If she guesses every answer, the probability of Rachel getting 11 questions correct is:

A ${}^{20}C_{11}\left(\frac{1}{4}\right)^{20}\left(\frac{3}{4}\right)^9$
B ${}^{20}C_{11}\left(\frac{1}{4}\right)^{11}\left(\frac{3}{4}\right)^9$
C ${}^{20}C_{10}\left(\frac{1}{4}\right)^{11}\left(\frac{3}{4}\right)^9$
D ${}^{20}C_{11}\left(\frac{1}{4}\right)^{9}\left(\frac{3}{4}\right)^{11}$
E ${}^{20}C_{10}\left(\frac{1}{4}\right)^{9}\left(\frac{3}{4}\right)^{11}$

21 **MC** A smoke detector has a probability of failing 2% of the time. If a shopping complex has installed 40 of these smoke detectors, the probability that at least one fails is given by:

A $1 - {}^{40}C_1\,(0.98)^1(0.02)^{39}$
B $1 - {}^{40}C_1\,(0.02)^1(0.98)^{39}$
C $1 - (0.02)^{40}$
D $1 - (0.98)^{40}$
E $(0.98)^{40}$

22 MC It is found that 3 out of every 10 cars are unroadworthy. Ten cars are selected at random. The probability that 3 are unroadworthy is:

A 0.009 **B** 0.2601 **C** 0.2668
D 0.5 **E** 1

23 WE5 Grant is a keen darts player and knows that his chance of scoring a bullseye on any one throw is 0.6.

a If Grant takes 5 shots at the target find the probability that he:
  i misses the bullseye each time
  ii hits the bullseye at least once.

b Find the number of throws Grant would need to ensure a probability of more than 0.7 of scoring at least one bullseye.

eBook*plus*

DIGITAL DOC
doc-9269
SkillSHEET 11.1
Solving indicial equations

24 The chance of winning a major prize in a raffle is 0.05. Find the number of tickets required to ensure a probability of more than 0.6 of winning a major prize at least once.

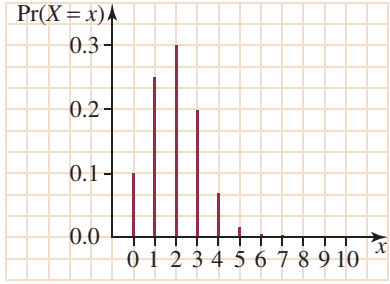
25 A darts player knows that her chance of scoring a bullseye on any one throw is 0.1. Find the number of turns she would need to ensure a probability of 0.9 of scoring at least one bullseye.

26 In Tattslotto, your chance of winning first division is $\frac{1}{8145\,060}$. Find:

a the number of games you would need to play if you wanted to ensure a more than 50% chance of winning first division at least once
b the number of tickets you would need to buy for part **a** if there are 16 games on each ticket
c the cost of buying these tickets, if they cost \$4.10 each.

27 The following probability distribution is for $p = 0.2$ and $n = 10$.

a Find the most likely outcome for x.
b Describe the plot.



28 From the following binomial distribution tables:

i draw a graph of the probability distribution
ii describe the skewness of each graph.

a $X \sim \text{Bi}(10, 0.3)$
b $X \sim \text{Bi}(10, 0.5)$
c $X \sim \text{Bi}(10, 0.8)$

eBook*plus*

DIGITAL DOC
doc-9227
Spreadsheet
Binomial distribution

x	$\Pr(X = x)$
0	0.028 25
1	0.121 06
2	0.233 47
3	0.266 83
4	0.200 12
5	0.102 92
6	0.036 76
7	0.009 00
8	0.001 45
9	0.000 01
10	5.9×10^{-6}

x	$\Pr(X = x)$
0	0.000 98
1	0.009 77
2	0.043 95
3	0.117 19
4	0.205 08
5	0.246 09
6	0.205 08
7	0.117 19
8	0.043 95
9	0.009 77
10	0.000 98

x	$\Pr(X = x)$
0	1×10^{-7}
1	4.1×10^{-6}
2	0.000 07
3	0.000 79
4	0.005 51
5	0.026 42
6	0.088 08
7	0.201 33
8	0.301 99
9	0.268 44
10	0.107 37

29 **a** Describe the plots of the following binomial probability distributions, without drawing the graphs.

i $n = 25, p = 0.1$ **ii** $n = 50, p = 0.5$ **iii** $n = 30, p = 0.9$

b What effect does p have on the graph of a binomial probability distribution?

30 **a** Describe the plot of the binomial probability distribution, $X \sim \text{Bi}(60, 0.5)$, without drawing the graph.

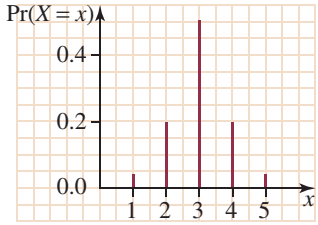
b Suggest how the graph might look for a binomial probability distribution with the same p, but double the value of n.

31 **a** Describe the plot of the binomial probability distribution, $X \sim \text{Bi}(100, 0.4)$, without drawing the graph.

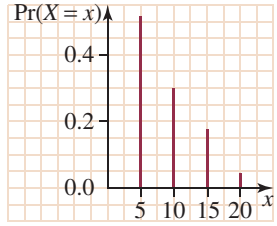
b Suggest how the graph might look for a binomial probability distribution with the same n, but double the value of p.

32 Describe the skewness of the graphs of the following binomial probability distributions.

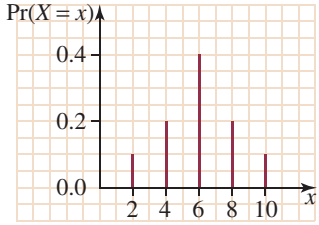
a



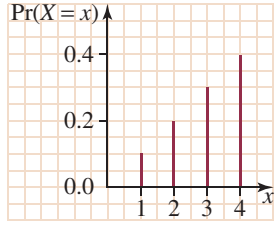
b



c



d



eBook*plus*

DIGITAL DOC
doc-9228
WorkSHEET 11.1

11B Problems involving the binomial distribution for multiple probabilities

We shall now work with problems involving the binomial distribution for multiple probabilities.

WORKED EXAMPLE 6

The binomial variable, X, has the following probability table.

x	0	1	2	3	4	5
$\Pr(X = x)$	0.2311	0.3147	0.3321	0.1061	0.0112	0.0048

a $\Pr(X > 3)$
b $\Pr(X \le 4)$.

THINK	WRITE
a 1 Pr $(X > 3)$ means $\Pr(X = 4)$ or $\Pr(X = 5)$. Add these probabilities.	**a** $\Pr(X > 3) = \Pr(X = 4) + \Pr(X = 5)$ $= 0.0112 + 0.0048$
2 Evaluate.	$= 0.0160$
b 1 $\Pr(X \le 4)$ would involve adding the probabilities from $\Pr(X = 0)$ to $\Pr(X = 4)$. Using the fact that $\Pr(X \le 4) = 1 - \Pr(X > 4)$ allows us to solve the problem using fewer terms.	**b** $\Pr(X \le 4) = \Pr(X = 0) + \ldots + \Pr(X = 4)$ $= 1 - \Pr(X = 5)$
2 Substitute the value into the rule.	$= 1 - 0.0048$
3 Evaluate.	$= 0.9952$

WORKED EXAMPLE 7

Find $\Pr(X \geq 3)$ if X has a binomial distribution with the probability of success, p, and the number of trials, n, given by $p = 0.3$, $n = 5$.

THINK	WRITE
1 State the probability distribution.	$X \sim \text{Bi}(n, p)$ $X \sim \text{Bi}(5, 0.3)$
2 Write what is required.	$\Pr(X \geq 3) = \Pr(X = 3) + \Pr(X = 4) + \Pr(X = 5)$
3 Substitute the values for n, p and q into the rule $\Pr(X = x) = {}^nC_x p^x q^{n-x}$.	$= {}^5C_3(0.3)^3(0.7)^2 + {}^5C_4(0.3)^4(0.7) + {}^5C_5(0.3)^5(0.7)^0$
4 Evaluate.	$= 10 \times 0.027 \times 0.49 + 5 \times 0.0081 \times 0.7 + 1 \times 0.002\,43 \times 1$ $= 0.1323 + 0.028\,35 + 0.002\,43$ $= 0.163\,08$

WORKED EXAMPLE 8

So Jung has a bag containing 4 red and 3 blue marbles. She selects a marble at random and then replaces it. She does this 7 times. Find the probability, correct to 4 decimal places, that:

a at least 5 marbles are red
b greater than 3 are red
c no more than 2 are red.

eBook*plus*
TUTORIAL
eles-1233
Worked example 8

THINK	WRITE
a 1 State the probability distribution and define and assign values to variables.	a Let X = number of red marbles selected $n = 7$ $p = \frac{4}{7}$ $\therefore X \sim \text{Bi}(7, \frac{4}{7})$ As X = number of red marbles selected, therefore $x = 5$. We want at least 5 red marbles, $\therefore \Pr(X \geq 5)$.
2 Use the binom Cdf feature of a CAS calculator to find $\Pr(x \geq 5)$. Enter n, p.	$n = 7$ $f = \frac{4}{7}$ binom Cdf $(7, \frac{4}{7}, 5, 7)$
3 Record the result. Round to 4 decimal places.	0.359 345 $\Pr(x \geq 5) = 0.3593$
b 1 Repeat step **2** to find $\Pr(x > 3)$. Note that $x > 3$ is the save as $x \geq 4$.	b binom Cdf $(7, \frac{4}{7}, 4, 7)$
2 Record the result. Give answer correct to 4 decimal places.	0.65 310 008 $\Pr(x > 3) = 0.6531$
c 1 Repeat step **2** to find $\Pr(x \leq 2)$.	c binom Cdf $(7, \frac{4}{7}, 0, 2)$
2 Record the result. Round to 4 decimal places.	0.126 584 $\Pr(x \leq 2) = 0.1266$

WORKED EXAMPLE 9

X follows a binomial distribution with $n = 9$, $p = 0.4$. Find, correct to 4 decimal places:

a **$\Pr(X \geq 7)$**

b **the probability that X is greater than 7 given it is greater than 5; that is, $\Pr(X > 7 \mid X > 5)$.**

THINK	WRITE
a 1 State the probability distribution.	**a** $X \sim \text{Bi}(n, p)$ $X \sim \text{Bi}(9, 0.4)$
2 Write what is required.	$\Pr(X \geq 7) = \Pr(X = 7) + \Pr(X = 8) + \Pr(X = 9)$
3 Substitute the values for n, p and q into the rule $\Pr(X = x) = {}^nC_x\, p^x q^{n-x}$.	$\Pr(X \geq 7) = {}^9C_7(0.4)^7(0.6)^2 + {}^9C_8(0.4)^8(0.6) + {}^9C_9(0.4)^9(0.6)^0$
4 Evaluate and round the answer to 4 decimal places.	$= 36 \times 1.6384 \times 10^{-3} \times 0.36 + 9 \times 6.5536 \times 10^{-4} \times 0.6 + 1 \times 2.621\,44 \times 10^{-4} \times 1$ $= 0.021\,23 + 0.003\,54 + 2.621\,44 \times 10^{-4}$ $= 0.025\,034\,752$ ≈ 0.0250
Part **a** can also be solved with a CAS calculator. (Refer to Worked example 8).	$\Pr(X \geq 7) = \text{binomCdf}(9, 0.4, 7, 9)$ $= 0.0250$
b 1 Write what is required.	**b** $\Pr(X > 7 \mid X > 5) = \dfrac{\Pr((X > 7) \cap (X > 5))}{\Pr(X > 5)}$ $= \dfrac{\Pr(X > 7)}{\Pr(X > 5)}$
2 Evaluate each probability individually. $\Pr(X > 7)$ can be obtained from results in **a**.	$\Pr(X > 7) = \Pr(X = 8) + \Pr(X = 9)$ $= 3.801\,088 \times 10^{-3}$ $\Pr(X > 5) = \Pr(X = 6) + \Pr(X = 7) + \Pr(X = 8) + \Pr(X = 9)$ $= {}^9C_6(0.4)^6(0.6)^3 + 0.025\,034\,752$ $= 84 \times 4.096 \times 10^{-3} \times 0.216 + 0.025\,034\,752$ $= 0.993\,525\,76$
3 Evaluate $\dfrac{\Pr(X > 7)}{\Pr(X > 5)}$.	$\dfrac{\Pr(X > 7)}{\Pr(X > 5)} = \dfrac{3.801\,088 \times 10^{-3}}{0.993\,525\,76}$ $= 0.038\,258\,75$
4 Round the answer to 4 decimal places. Part **b** can also be solved with a CAS calculator.	≈ 0.0383 $\Pr(X > 7 \mid X > 5) = \dfrac{\Pr(X > 7)}{\Pr(X > 5)}$ $= \dfrac{\text{binomCdf}(9, 0.4, 7, 9)}{\text{binomCdf}(9, 0.4, 5, 9)}$ $= 0.0383$

WORKED EXAMPLE 10

Seventy per cent of all scheduled trains through Westbourne station arrive on time. If 10 trains go through the station every day, find, correct to 4 decimal places:

a **the probability that at least 8 trains are on time**

b **the probability that at least 8 trains are on time for 9 out of the next 10 days.**

eBook*plus*

TUTORIAL
eles-1234
Worked example 10

THINK	WRITE
a 1 State the probability distribution.	**a** $X \sim \text{Bi}(n, p)$ $X \sim \text{Bi}(10, 0.7)$
2 Write what is required.	$\Pr(X \geq 8) = \Pr(X = 8) + \Pr(X = 9) + \Pr(X = 10)$
3 Substitute the values of n, p and q into the rule $\Pr(X = x) = {}^nC_x p^x q^{n-x}$.	$= {}^{10}C_8(0.7)^8(0.3)^2 + {}^{10}C_9(0.7)^9(0.3)^1 + {}^{10}C_{10}(0.7)^{10}(0.3)^0$
4 Evaluate and round the answer to 4 decimal places.	$= 45 \times 0.057\,648\,01 \times 0.09 + 10 \times 0.040\,353\,607 \times 0.3 + 1 \times 0.028\,247\,524\,9 \times 1$ $= 0.233\,474\,440\,5 + 0.121\,060\,821 + 0.028\,247\,524\,9$ $= 0.382\,782\,786\,4$ ≈ 0.3828
Part **a** can also be solved with a CAS calculator. (Refer to Worked example 8).	$\Pr(X \geq 8) = \text{binomCdf}(10, 0.7, 8, 10)$ $= 0.3828$
5 Answer the question.	The probability that at least 8 trains are on time is 0.3828.
b 1 State the probability distribution.	**b** $X \sim \text{Bi}(n, p)$ $X \sim \text{Bi}(10, 0.3828)$
2 Write what is required.	$\Pr(X = 9) = {}^{10}C_9(0.3828)^9(0.6172)$
3 Substitute the values of n, p and q into the rule $\Pr(X = x) = {}^nC_x p^x q^{n-x}$.	
4 Evaluate and round the answer to 4 decimal places.	$= 10 \times 0.000\,176\,501\,08 \times 0.6172$ $= 0.001\,089\,364\,7$ ≈ 0.0011
Part **b** can also be solved with a CAS calculator.	$\Pr(X = 9) = \text{binomPdf}(10, 0.3828, 9)$ $= 0.0011$
5 Answer the question.	The probability that at least 8 trains are on time for 9 out of the next 10 days is 0.0011.

Exercise 11B Problems involving the binomial distribution for multiple probabilities

1 Find:

a ${}^4C_3(0.4)^3\,(0.6) + {}^4C_4(0.4)^4\,(0.6)^0$

b ${}^5C_3(0.6)^3\,(0.4)^2 + {}^5C_4(0.6)^4\,(0.4) + {}^5C_5(0.6)^5\,(0.4)^0$.

2 WE6 The binomial variable, X, has the following probability table.

x	0	1	2	3	4	5
$\Pr(X = x)$	0.15	0.3	0.1	0.22	0.15	0.08

Find:

a $\Pr(X \geq 4)$

b $\Pr(X > 0)$

c $\Pr(X \leq 4)$

d $\Pr(X < 2)$.

3 A binomial variable, X, has the probability function $\Pr(X = x) = {}^5C_x(0.3)^x\,(0.7)^{5-x}$, where x is the probability of success and $x = 0, 1, \ldots, 5$.

Find:

a $\Pr(X \geq 2)$

b $\Pr(X < 4)$.

eBook*plus*

DIGITAL DOC
doc-9270
SkillSHEET 11.2
Multiple probabilities

4 **WE7** Find $\Pr(X \geq 4)$, correct to 4 decimal places, if X has a binomial distribution with the probability of success, p, and the number of trials, n, given by:

a $p = 0.6, n = 5$ **b** $p = 0.5, n = 6$ **c** $p = 0.2, n = 7$.

5 A fair coin is tossed 5 times. Calculate the probability of obtaining:

a at least one tail
b greater than three tails
c greater than three tails, given that at least one tail was obtained.

6 Marco has a faulty alarm clock and the probability that it sounds in the morning is $\frac{1}{3}$. Calculate the probability, for the next 4 mornings, that his alarm clock:

a works at least 3 times **b** works fewer than 2 times
c works at least 3 times, given that it works at least once.

7 Generally, 10% of people who enter a modelling contest are male. For a particular competition, three winners were chosen. What is the probability that less than two females were chosen?

8 **WE8** A bag contains 4 red and 2 blue marbles. A marble is selected at random and replaced. The experiment is repeated 6 times. Find the probability that:

a all 6 selections are red **b** at least 2 are red **c** not more than 1 is red.

9 It is known that 40% of Victorians play sport regularly. Ten people are selected at random. Calculate the probability, correct to 4 decimal places, that:

a at least half play sport regularly **b** at least nine don't play sport regularly.

10 Surveys have shown that 8 out of 10 VCE students study every night. Six VCE students are selected at random. Find the probability, correct to 4 decimal places, that, on any one day:

a at least 50% of these students study every night
b less than 3 students study every night.

11 A die is weighted such that $\Pr(X = 6) = \frac{1}{2}$, $\Pr(X = 2) = \Pr(X = 4) = \frac{1}{6}$ and $\Pr(X = 1) = \Pr(X = 3) = \Pr(X = 5) = \frac{1}{18}$. The die is rolled five times. Calculate the probability, correct to 4 decimal places, of obtaining:

a at least three 6s **b** at least two even numbers
c a maximum of two odd numbers.

12 **WE9** If X is binomially distributed with $n = 8$ and $p = 0.7$, find, correct to 4 decimal places:

a $\Pr(X \geq 7)$ **b** $\Pr(X > 7 \mid X > 5)$.

13 A survey shows that 49% of the public support the current government. Twelve people are selected at random. Calculate, correct to 4 decimal places:

a the probability that at least 8 support the government
b the probability that at least 10 support the government, given that at least 8 do.

14 **MC** When Graeme kicks for goal, the probability of his kicking a goal is 0.7. If he has five kicks at goal, the probability that he will score fewer than two goals is:

A ${}^5C_1(0.7)^1\,(0.3)^4 + (0.3)^5$
B ${}^5C_2(0.7)^2\,(0.3)^3$
C ${}^5C_2(0.7)^2\,(0.3)^3 + {}^5C_1(0.7)^1\,(0.3)^4$
D ${}^5C_2(0.7)^2\,(0.3)^3 + {}^5C_1(0.7)^1\,(0.3)^4 + (0.3)^5$
E $1 - {}^5C_2(0.7)^2\,(0.3)^3$

15 **MC** The proportion of patients who suffer a violent reaction from a new drug being trialled is p. If 80 patients trial the drug, the probability that one-quarter of the patients have a violent reaction is:

A ${}^{80}C_{25}\,(p)^{25}(1-p)^{55}$ **B** ${}^{80}C_{20}\,(1-p)^{20}(p)^{60}$ **C** ${}^{80}C_{25}\,(1-p)^{25}(p)^{55}$
D ${}^{80}C_{20}\,(p)^{20}(1-p)^{60}$ **E** ${}^{80}C_{20}\,(p)^{20}$

16 **MC** If X is a random variable, binomially distributed with $n = 10$ and $p = k$, $\Pr(X \geq 1)$ is:

A $1 - (1-k)^{10}$ **B** $(1-k)^{10}$ **C** $10(k)(1-k)^9$ **D** $(k)^{10}$ **E** $1 - (k)^{10}$

17 **MC** Three per cent of items made by a certain machine are defective. The items are packed and sold in boxes of 10. If 3 or more are defective, the box can be returned and money refunded. The chance of being eligible for a refund is:

A 0 **B** 0.0002 **C** 0.0036 **D** 0.0028 **E** 0.9972

18 **MC** Long-term statistics show that Silvana wins 60% of her tennis matches. The probability that she will win at least 80% of her next 10 matches is:

A 0.0061 **B** 0.0464 **C** 0.1673 **D** 0.8327 **E** 0.9536

19 **MC** Nineteen out of every 20 cricketers prefer 'Boundary' cricket gear. A squad of 12 cricketers train together. The probability that at least 11 use Boundary gear, given that at least 10 use it, is:

A 0.5404 **B** 0.6129 **C** 0.8816
D 0.8992 **E** 0.9804

20 A school council, comprising 15 members of the school community, requires a minimum two-thirds majority to pass a motion. It is known that 50% of the school community favour a new uniform. Calculate the probability that the school council will pass a motion in favour of a new uniform, correct to 4 decimal places.

21 A car insurance salesman knows that he has a good chance of finding customers in the age group from 18 to 20, as people often buy their first car at this age. Five per cent of all people in this age group are looking to purchase a car. The salesman questions 30 people in this age group. Calculate the probability, correct to 4 decimal places, that he will get:

a no more than 3 sales
b at least 3 sales.

22 Police radar camera tests have shown that 1% of all cars drive at over 30 km/h above the speed limit, 2% between 10 km/h and 30 km/h above the limit and 4% below 10 km/h over the limit. In one particular hour, a radar camera tests 50 cars. Calculate the probability, correct to 4 decimal places, that:

a at most, one car is over 30 km/h above the limit
b at most, two cars are between 10 km/h and 30 km/h above the limit
c at most, two cars are below 10 km/h above the limit
d at most, three cars are above the limit.

23 An Australian cricketer scores 50 or more runs in one-third of all his test match innings. The Australian selectors are aiming to predict his next 10 innings. Calculate, correct to 4 decimal places, the probability that he will score 50 or more runs on:

a no occasions
b exactly four occasions
c at least two occasions.

24 Two dice are rolled simultaneously and their difference is recorded. Find the probability, correct to 4 decimal places, that in 5 rolls:

a a difference of zero occurs at least once
b a difference of 1 occurs at least twice
c a difference of 5 occurs at least once.

25 **WE10** Eighty per cent of all scheduled trains through Westbourne station arrive on time. If 10 trains go through the station every day, find:

a the probability that at least 8 trains are on time
b the probability that at least 8 trains are on time for 9 out of the next 10 days.

26 Seventy-five per cent of all scheduled trains through Westbourne station arrive on time. If 15 trains go through the station every day, find, correct to 4 decimal places:

a the probability that at least 10 trains are on time
b the probability that at least 10 trains are on time for 8 out of the next 10 days.

27 An experiment involves rolling a die 6 times. Calculate, correct to 4 decimal places:

a the probability of obtaining at least four prime numbers
b the probability of obtaining at least four prime numbers on 5 occasions if the experiment is repeated 8 times.

28 Tennis balls are packed in cans of 6. Five per cent of all balls are made too flat (that is, they don't bounce high enough). The cans are then packed in boxes of two dozen. Calculate the probability, correct to 4 decimal places, that:

a a can contains, at most, one flat ball
b a box contains at least 22 cans with a maximum of one flat ball.

eBook*plus*

INTERACTIVITY
int-0256
Markov chains and transition matrices

study on

Units:	3 & 4
AOS:	4
Topic:	2
Concept:	4

Concept summary
Read a summary of this concept.

See more
Watch a video about Markov Chains.

study on

Units:	3 & 4
AOS:	4
Topic:	2
Concept:	5

Concept summary
Read a summary of this concept.

11C Markov chains and transition matrices

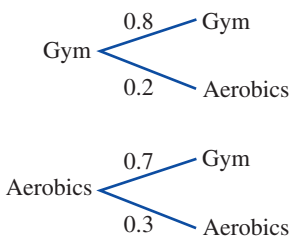
Andrei Markov was a Russian mathematician whose name is given to a technique that calculates probability associated with the state of various transitions. It answers questions such as, 'What is the probability that James will be late to work today given that he was late yesterday?' or 'What can be said about the long-term behaviour of James' punctuality?'.

A Markov chain is a sequence of repetitions of an experiment in which:

1. The probability of a particular outcome in an experiment is *conditional* only on the outcome of the experiment immediately before it.
2. The *conditional probabilities* of each outcome in a particular experiment are the *same* every single time.

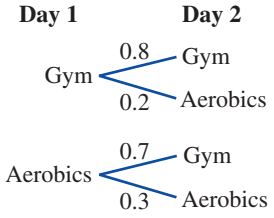
A two-state Markov chain is one in which there are only two possible outcomes for each experiment.

Consider a leisure centre that offers aerobics classes and has a gym. Records show that 20% of the members who use the gym on a particular day will participate in an aerobics class the next day and 70% of the members who participate in an aerobics class on a particular day will use the gym the next day. It is also known that 200 members use the leisure centre each day and they all participate in aerobics classes or use the gym, but not both. On a particular day 150 members use the gym and 50 members attend an aerobics class. The possible outcomes may be illustrated on a tree diagram as shown below.



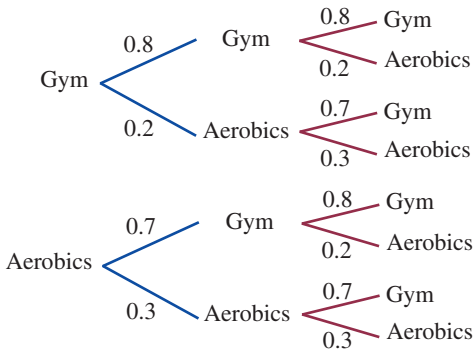
Outcome	Interpretation
GG	The member uses the gym on day 2 given the member used the gym on day 1.
GA	The member attends an aerobics class on day 2 given the member used the gym on day 1.
AG	The member uses the gym on day 2 given the member attended an aerobics class on day 1.
AA	The member attends an aerobics class on day 2 given the member attended an aerobics class on day 1.

The tree diagram can also be used to calculate how many members use the gym or attend an aerobics class. From the tree diagram below, it can be seen that on the second day 155 members use the gym and 45 attend an aerobics class.



Outcome	Number of members
GG	$150 \times 0.8 = 120$
GA	$150 \times 0.2 = 30$
AG	$50 \times 0.7 = 35$
AA	$50 \times 0.3 = 15$

The tree diagram may be extended to display the possible outcomes and their respective probabilities for the third day.



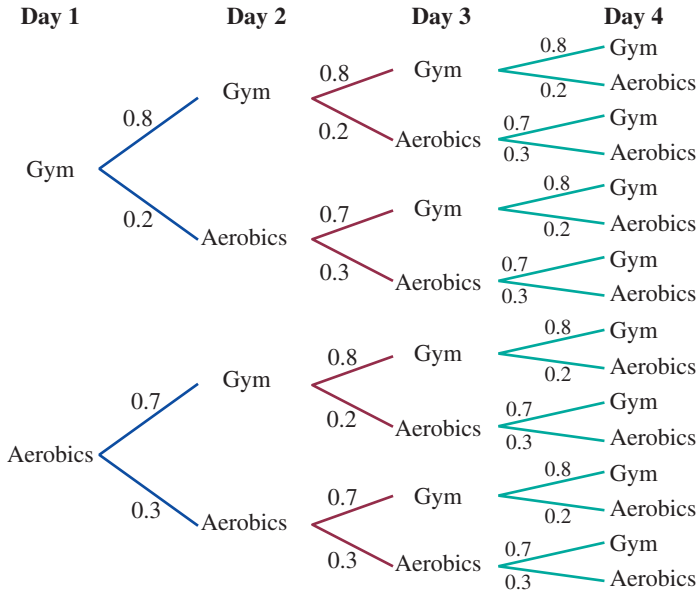
Outcome	Probability
GGG	$0.8 \times 0.8 = 0.64$
GGA	$0.8 \times 0.2 = 0.16$
GAG	$0.2 \times 0.7 = 0.14$
GAA	$0.2 \times 0.3 = 0.06$
AGG	$0.7 \times 0.8 = 0.56$
AGA	$0.7 \times 0.2 = 0.14$
AAG	$0.3 \times 0.7 = 0.21$
AAA	$0.3 \times 0.3 = 0.09$

GGG represents the member using the gym on all three days; Pr(GGG) = 0.64.

If you want to find out the probability of the member using the gym on two out of three days, three outcomes need to be considered: GGA, GAG and AGG.

$\therefore$ Pr(gym 2 out of 3 days) = Pr(GGA) + Pr(GAG) + Pr(AGG)
$= 0.16 + 0.14 + 0.56$
$= 0.86$

The tree diagram may be further extended to display the possible outcomes and their respective probabilities on the fourth day.



Outcome	Probability
GGGG	$0.8 \times 0.8 \times 0.8 = 0.512$
GGGA	$0.8 \times 0.8 \times 0.2 = 0.128$
GGAG	$0.8 \times 0.2 \times 0.7 = 0.112$
GGAA	$0.8 \times 0.2 \times 0.3 = 0.048$
GAGG	$0.2 \times 0.7 \times 0.8 = 0.112$
GAGA	$0.2 \times 0.7 \times 0.2 = 0.028$
GAAG	$0.2 \times 0.3 \times 0.7 = 0.042$
GAAA	$0.2 \times 0.3 \times 0.3 = 0.018$
AGGG	$0.7 \times 0.8 \times 0.8 = 0.448$
AGGA	$0.7 \times 0.8 \times 0.2 = 0.112$
AGAG	$0.7 \times 0.2 \times 0.7 = 0.098$
AGAA	$0.7 \times 0.2 \times 0.3 = 0.042$
AAGG	$0.3 \times 0.7 \times 0.8 = 0.168$
AAGA	$0.3 \times 0.7 \times 0.2 = 0.042$
AAAG	$0.3 \times 0.3 \times 0.7 = 0.063$
AAAA	$0.3 \times 0.3 \times 0.3 = 0.027$

As we extend the analysis to cover a greater number of days, setting up the tree diagram becomes very messy and time consuming.

The original information can be set up as a pair of recurrence relationships. That is, if 20% of the members using the gym attend an aerobics class the next day, 80% will use the gym the next day. Furthermore, if 70% of the members attending an aerobics class use the gym the next day, 30% will attend an aerobics class the next day.

Let g_i = the number of members who use the gym on day i.
Let a_i = the number of members who attend an aerobics class on day i.

$g_{i+1} = 0.8g_i + 0.7a_i$ and $a_{i+1} = 0.2g_i + 0.3a_i$

Number of gym users

Day i

$g_i = 150$

Day $i + 1$

$g_{i+1} = 0.8g_i + 0.7a_i$
$= 0.8 \times 150 + 0.7 \times 50$
$= 155$

Day $i + 2$

$g_{i+2} = 0.8g_{i+1} + 0.7a_{i+1}$
$= 0.8 \times 155 + 0.7 \times 45$
$= 155.5$

Day $i + 3$

$g_{i+3} = 0.8g_{i+2} + 0.7a_{i+2}$
$= 0.8 \times 155.5 + 0.7 \times 44.5$
$= 155.55$

Number of aerobics participants

Day i

$a_i = 50$

Day $i + 1$

$a_{i+1} = 0.2g_i + 0.3a_i$
$= 0.2 \times 150 + 0.3 \times 50$
$= 45$

Day $i + 2$

$a_{i+2} = 0.2g_{i+1} + 0.3a_{i+1}$
$= 0.2 \times 155 + 0.3 \times 45$
$= 44.5$

Day $i + 3$

$a_{i+3} = 0.2g_{i+2} + 0.3a_{i+2}$
$= 0.2 \times 155.5 + 0.3 \times 44.5$
$= 44.45$

Day $i + 4$

$$g_{i+4} = 0.8g_{i+3} + 0.7a_{i+3}$$
$$= 0.8 \times 155.55 + 0.7 \times 44.45$$
$$= 155.555$$

Day $i + 4$

$$a_{i+4} = 0.2g_{i+3} + 0.3a_{i+3}$$
$$= 0.2 \times 155.55 + 0.3 \times 44.45$$
$$= 44.445$$

This method allows us to clearly see how many members (when rounded to integer values) are using the gym or attending an aerobics class each day.

Often we are interested in the long-term behaviour (or the steady state, as it is often called) of a particular situation, in this case how many members will use the gym or attend an aerobics class. We can determine this by using the following information.

Let g = the number of members who use the gym.
Let a = the number of members who attend an aerobics class.
Total number of members = 200

This gives the equation $\quad g + a = 200$

which when rearranged is equal to $\quad a = 200 - g \quad [1]$

Also $\quad g = 0.8g + 0.7a \quad [2]$

and $\quad a = 0.2g + 0.3a \quad [3]$

Rearranging equation [2]

$$g - 0.8g = 0.7a$$
$$0.2g = 0.7a$$
$$\frac{0.2g}{0.2} = \frac{0.7a}{0.2}$$
$$g = 3.5a \quad [4]$$

Substituting [4] into [1] and solving gives

$$a = 200 - 3.5a$$
$$a + 3.5a = 200$$
$$4.5a = 200$$
$$a = \frac{400}{9}$$
$$a = 44.4444 \text{ (correct to 4 decimal places)}$$

Substituting $a = 44.4444$ gives $g = 155.5556$.

In the long term, 156 members will use the gym and 44 members will attend an aerobics class.

WORKED EXAMPLE 11

The Nee Islands are very wet. If it is raining on a particular day, the chance that it will rain the next day is 60%. If it is not raining on a particular day, the chance that it will rain on the following day is 45%.

a If it is raining on Tuesday, draw a tree diagram to represent the next two days.

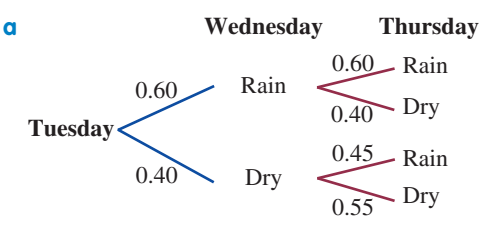
b Extending the tree diagram, calculate the probability that, if it is raining on Tuesday, it will also be raining on Friday of the same week.

THINK

a Draw a tree diagram labelling each branch and place the appropriate probability along the relevant branch.

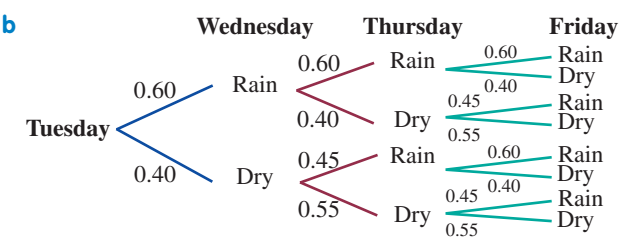
WRITE/DRAW

a



b 1 Draw a tree diagram labelling each branch and place the appropriate probability along the relevant branch. *Note:* The tree diagram will start at Tuesday and extend to Friday.

b



2 List each of the outcomes and calculate the probability for each individual outcome.

Outcome	Probability
RRR	$\mathbf{0.6 \cdot 0.6 \cdot 0.6 = 0.216}$
RRD	$0.6 \times 0.6 \times 0.4 = 0.144$
RDR	$\mathbf{0.6 \cdot 0.4 \cdot 0.45 = 0.108}$
RDD	$0.6 \times 0.4 \times 0.55 = 0.132$
DRR	$\mathbf{0.4 \cdot 0.45 \cdot 0.6 = 0.108}$
DRD	$0.4 \times 0.45 \times 0.4 = 0.072$
DDR	$\mathbf{0.4 \cdot 0.55 \cdot 0.55 = 0.099}$
DDD	$0.4 \times 0.55 \times 0.55 = 0.121$

3 Add the probabilities of the required outcomes.

$0.216 + 0.108 + 0.108 + 0.099 = 0.531$

4 Interpret the answer.

The probability that it will be raining on Friday if it is raining on Tuesday is 0.531. That is, there is a 53.1% chance of raining on Friday given it will rain on Tuesday.

WORKED EXAMPLE 12

eBook*plus*

TUTORIAL
eles-1235
Worked example 12

Commuters travelling into the centre of Trenchtown use either the bus or the train. Research has shown that each month 20% of those using the bus switch to train travel and 30% of those using the train switch to bus travel.

If, at the beginning of January, 4800 people were using the bus and 3600 were using the train to get into the city, calculate:

a the number of people using the train at the beginning of May

b the number of people using the bus and train in the long term.

THINK | **WRITE**

a 1 Interpret the given information.

a Each month 20% of bus commuters switch to train travel and 80% will not switch.
Each month 30% of train commuters switch to bus travel and 70% will not switch.

2 Define the variables.

Let t_i = the number of people using the train at the beginning of month i.
Let b_i = the number of people using the bus at the beginning of month i.

3 Set up a pair of recurrence relationships.

$t_{i+1} = 0.7t_i + 0.2b_i$ and $b_{i+1} = 0.3t_i + 0.8b_i$

4 Substitute the given values of bus and train commuters into the recurrence relationships.

Train travellers

January (month i)
$t_i = 3600$

February (month $i + 1$)
$t_{i+1} = 0.7t_i + 0.2b_i$
$= 0.7 \times 3600 + 0.2 \times 4800$
$= 3480$

March (month $i + 2$)
$t_{i+2} = 0.7t_{i+1} + 0.2b_{i+1}$
$= 0.7 \times 3480 + 0.2 \times 4920$
$= 3420$

Bus travellers

January (month i)
$b_i = 4800$

February (month $i + 1$)
$b_{i+1} = 0.3t_i + 0.8b_i$
$= 0.3 \times 3600 + 0.8 \times 4800$
$= 4920$

March (month $i + 2$)
$b_{i+2} = 0.3t_{i+1} + 0.8b_{i+1}$
$= 0.3 \times 3480 + 0.8 \times 4920$
$= 4980$

April (month $i + 3$)

$t_{i+3} = 0.7t_{i+2} + 0.2b_{i+2}$
$= 0.7 \times 3420 + 0.2 \times 4980$
$= 3390$

May (month $i + 4$)

$t_{i+4} = 0.7t_{i+3} + 0.2b_{i+3}$
$= 0.7 \times 3390 + 0.2 \times 5010$
$= 3375$

April (month $i + 3$)

$b_{i+3} = 0.3t_{i+2} + 0.8b_{i+2}$
$= 0.3 \times 3420 + 0.8 \times 4980$
$= 5010$

May (month $i + 4$)

$b_{i+4} = 0.3t_{i+3} + 0.8b_{i+3}$
$= 0.3 \times 3390 + 0.8 \times 5010$
$= 5025$

5. Evaluate the number of bus and train commuters at the beginning of February.

6. Use the values obtained in step 5 to calculate the number of bus and train commuters at the beginning of March.

7. Repeat the processes involved in step 5 and 6 until the number of bus and train commuters at the beginning of May is obtained.

8. Answer the question.

At the beginning of May, the number of people using the bus and train respectively is 5025 and 3375.

b 1. Set up and number equations that relate to:
(i) the number of train commuters
(ii) the number of bus commuters
(iii) the total number of commuters.
Note: Maintain the variables defined in part **a**.

b Let t = number of people using the train.
Let b = number of people using the bus.
Total = 3600 + 4800

$t + b = 8400$ [1]
$t = 0.7t + 0.2b$ [2]
$b = 0.3t + 0.8b$ [3]

2. Rearrange equation [1] so that t is the subject.
Note: Either variable may be the subject.

$t + b = 8400$
$t = 8400 - b$ [4]

3. Rearrange equation [3] so that b can be expressed in terms of t.
Note: Again either variable may be transposed.

$b = 0.3t + 0.8b$
$b - 0.8b = 0.3t$
$0.2b = 0.3t$
$\frac{0.2b}{0.2} = \frac{0.3t}{0.2}$
$b = 1.5t$ [5]

4. Substitute equation [5] into [4] and solve for t.

$t = 8400 - b$
$t = 8400 - 1.5t$
$t + 1.5t = 8400$
$2.5t = 8400$
$\frac{2.5t}{2.5} = \frac{8400}{2.5}$
$t = 3360$

5. Substitute t into equation [5] and solve for b.

$b = 1.5t$
$b = 1.5 \times 3360$
$= 5040$

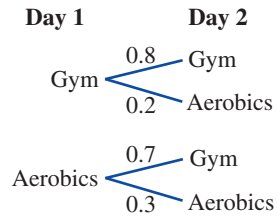
6. Answer the question.

In the long term, 3360 commuters will travel by train and 5040 will travel by bus.

Transition matrices

Transition matrices are another technique for solving *some* Markov chain problems. Transition matrices are specifically used for problems where you are only interested in the probability of the final outcome (for example, the probability that a person goes to the gym on the seventh day). Tree diagrams must still be used to solve problems that take into account the number of different ways of reaching the final outcome (e.g. the probability that a person goes to the gym 4 out of the next 7 days).

Consider again the previous scenario, from page 32, of the probability of using the gym or attending an aerobics class.



This can also be described in a transition probability table:

	Gym today	Aerobics today
Gym tomorrow	0.8	0.7
Aerobics tomorrow	0.2	0.3

or

	Aerobics today	Gym today
Aerobics tomorrow	0.3	0.2
Gym tomorrow	0.7	0.8

Note: Each column should add up to one, as do the pairs of branches in the given tree diagram.

A transition matrix can be used to simplify calculations involving Markov chains. As its name suggests, the matrix assists in the calculation of the transition from one state to the next.

We can convert the above transition probability tables into a transition matrix, T, as shown below.

$$T = \begin{array}{c} \\ G_2 \\ A_2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} G_1 & A_1 \end{array} \\ \begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix} \end{array} \quad \text{or} \quad T = \begin{array}{c} \\ A_2 \\ G_2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} A_1 & G_1 \end{array} \\ \begin{bmatrix} 0.3 & 0.2 \\ 0.7 & 0.8 \end{bmatrix} \end{array}$$

The proportion of the population that use the gym or attend the aerobics class is called the state of the system, S. It can either relate to specific populations, or the probabilities associated with the different outcomes.

For our example, the initial system, S_0 is given by:

$$\begin{bmatrix} n(\text{using the gym on day 1}) \\ n(\text{doing aerobics on day 1}) \end{bmatrix} = \begin{bmatrix} 150 \\ 50 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} n(\text{doing aerobics on day 1}) \\ n(\text{using the gym on day 1}) \end{bmatrix} = \begin{bmatrix} 50 \\ 150 \end{bmatrix}$$

At any time, the state is represented by the column matrix $\begin{bmatrix} n(G_i) \\ n(A_i) \end{bmatrix}$ or $\begin{bmatrix} n(A_i) \\ n(G_i) \end{bmatrix}$, where $n(G_i)$ is the number of gym users and $n(A_i)$ is the number of aerobics participants.

There are two different transition and initial state matrices for this problem, so which one should we choose to use? It does not matter which one we choose, as long as the correct pairing of T and S_0 is selected.

For example, if we choose $T = \begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix}$, S_0 must be $\begin{bmatrix} 150 \\ 50 \end{bmatrix}$.

The element in the 1st row of S_0 must represent the initial value for the element in the 1st row, 1st column of the transition matrix.

These elements must match (i.e. the probability about gym users matches with the initial number of gym users).

Similarly, these elements must match (i.e. the probability about aerobics participants matches with the initial number of aerobics participants).

i.e. when multiplying T and S: $\begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix} \times \begin{bmatrix} 150 \\ 50 \end{bmatrix}$

They are then multiplied together as part of the 1st element for S_0.

$$= \begin{bmatrix} 0.8 \times 150 + 0.7 \times 50 \\ 0.2 \times 150 + 0.3 \times 50 \end{bmatrix}$$

$$= \begin{bmatrix} 155 \\ 45 \end{bmatrix}$$

Hence, the state after 1 day, S_1 is $\begin{bmatrix} n(\mathrm{G}_1) \\ n(\mathrm{A}_1) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix} \begin{bmatrix} 150 \\ 50 \end{bmatrix}$

$$= \begin{bmatrix} 0.8 \times 150 + 0.7 \times 50 \\ 0.2 \times 150 + 0.3 \times 50 \end{bmatrix}$$

$$= \begin{bmatrix} 155 \\ 45 \end{bmatrix}$$

Note: These are the same values as we calculated earlier.

From the above calculation, we can see that $S_1 = T \times S_0$.

If the conditional probabilities remain the same, then a similar equation will express the transition from any particular state to the next.

Therefore, in general, $S_{n+1} = T \times S_n$

Also, as, $S_1 = T \times S_0$ and, $S_2 = T \times S_1$,

then, $S_2 = T \times T \times S_0 = T^2 \times S_0$

and, $S_3 = T \times T \times T \times S_0 = T^3 \times S_0$

$\therefore$ **In general, $S_n = T^n \times S_0$**

To find the long-term behaviour, or steady state, choose n to be a large number, for example $n = 50$, and find S_{50}.

T can be written from the generalised transition probability table below:

	Event A_1	**Event A_1'**
Event A_2	$\Pr(\mathrm{A}_2 \mid \mathrm{A}_1)$	$\Pr(\mathrm{A}_2 \mid \mathrm{A}_1')$
Event A_2'	$\Pr(\mathrm{A}_2' \mid \mathrm{A}_1)$	$\Pr(\mathrm{A}_2' \mid \mathrm{A}_1')$

$$T = \begin{bmatrix} \Pr(\mathrm{A}_2 \mid \mathrm{A}_1) & \Pr(\mathrm{A}_2 \mid \mathrm{A}_1') \\ \Pr(\mathrm{A}_2' \mid \mathrm{A}_1) & \Pr(\mathrm{A}_2' \mid \mathrm{A}_1') \end{bmatrix} \text{ and } S_0 = \begin{bmatrix} n(\mathrm{A}_1)_0 \\ n(\mathrm{A}_2)_0 \end{bmatrix} \text{ or } \begin{bmatrix} \Pr(\mathrm{A}_1)_0 \\ \Pr(\mathrm{A}_2)_0 \end{bmatrix}.$$

WORKED EXAMPLE 13

Using the above data for attending the gym or aerobics class, find:

a the proportion of people attending the gym and aerobics class on the 5th day

b the number of people attending the gym or aerobics class in the long term.

THINK	WRITE
a 1 Write down the transition matrix.	**a** $T = \begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix}$

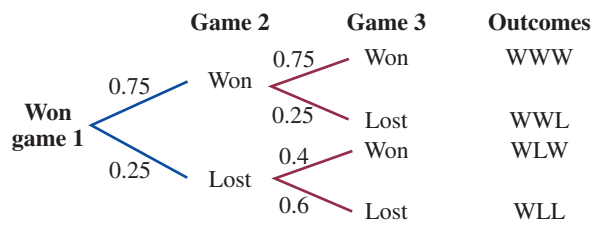
2 Write down a suitable initial state matrix. In this case, it is the initial numbers of people attending the gym and aerobics class.	$S_0 = \begin{bmatrix} 150 \\ 50 \end{bmatrix}$
3 Identify which state matrix is required.	As S_0 corresponds to day 1, therefore day 5 corresponds to the state matrix S_4.
4 Using a CAS calculator define the transition matrix T.	$\begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix} \rightarrow t$
5 Define the initial state in terms of numbers of people attending gym (150) and aerobics (50).	$\begin{bmatrix} 150 \\ 50 \end{bmatrix} \rightarrow s_0$
6 Calculate the proportion of people attending gym or aerobics on the 5th day. $S_4 = T^4 \times S_0$.	$t^4 S_0$
7 Write the answer.	$\begin{bmatrix} 155.555 \\ 44.445 \end{bmatrix}$
8 Write the solution.	$S_4 = T^4 \times S_0$ $= \begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix}^4 \times \begin{bmatrix} 150 \\ 50 \end{bmatrix}$ $= \begin{bmatrix} 155.555 \\ 44.445 \end{bmatrix}$
9 Answer the question (rounding to the nearest whole number).	On the 5th day, there will be 156 people at the gym and 44 people attending the aerobics class.
b 1 Repeat step 6 for the 'long term' — choose $n = S_0$.	**b** $t^{50} S_0$
2 Record the result.	$\begin{bmatrix} 155.556 \\ 44.444 \end{bmatrix}$
3 Write the solution.	$S_{50} = T^{50} \times S_0$ $= \begin{bmatrix} 0.8 & 0.7 \\ 0.2 & 0.3 \end{bmatrix}^{50} \times \begin{bmatrix} 150 \\ 50 \end{bmatrix}$ $= \begin{bmatrix} 155.555 \\ 44.445 \end{bmatrix}$
4 Answer the question (rounding to the nearest whole number).	In the long-term, 156 people attend the gym and 44 people go to aerobics class.

WORKED EXAMPLE 14

The chance of Jo's netball team winning a given game depends on how the team performs in the previous game. If her team wins, then the chance that it will win the next game is 0.75. If her team loses, there is only a 0.4 chance that they will win the next match. Given the team wins their first match, find the probability, that:

a they win two out of their first three games

b they win the 7th game they play, correct to 4 decimal places.

THINK	WRITE/DRAW
a 1 Draw a tree diagram labelling each branch and place the appropriate probability along the relevant branch. List the outcomes.	**a** 
2 Select the appropriate outcomes for the required probability.	Pr(win 2 out of 3 games) = Pr(WWL) + Pr(WLW)
3 Calculate the probabilities by multiplying along the branches.	Pr(win 2 out of 3 games) $= 0.75 \times 0.25 + 0.25 \times 0.4$ $= 0.1875 + 0.1$ $= 0.2875$
b 1 Write down the transition matrix, T.	**b** $T = \begin{array}{c} \\ W_2 \\ L_2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} W_1 & L_1 \end{array} \\ \begin{bmatrix} 0.75 & 0.4 \\ 0.25 & 0.6 \end{bmatrix} \end{array}$ $T = \begin{bmatrix} 0.75 & 0.4 \\ 0.25 & 0.6 \end{bmatrix}$
2 Write down a suitable initial state matrix, S_0. As the value of the 1st row, 1st column of T is the probability associated with winning consecutive games, the first element of S_0 will be the probability that the team has won. We know for certain that Jo's team won the first match, so the probability must be 1. Therefore, the probability of the second element must be 0.	$S_0 = \begin{bmatrix} \text{Pr(Jo's team wins 1st game)} \\ \text{Pr(Jo's team loses 1st game)} \end{bmatrix}$ $S_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
3 Identify which state matrix is required.	As S_0 corresponds to game 1, therefore game 7 corresponds to the state matrix S_6.
4 Using a CAS calculator, define the transition matrix, T.	$\begin{bmatrix} 0.75 & 0.4 \\ 0.25 & 0.6 \end{bmatrix} \rightarrow t$
5 Define the initial state matrix, S_0.	$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow S_0$
6 Calculate the state matrix for the 7th day (S_6) ($S_6 = T^5 \times S_0$).	$t^6 S_0$
7 Write your answer.	$\begin{bmatrix} 0.616\,092 \\ 0.383\,908 \end{bmatrix}$
8 Write the solution.	Pr(winning the 7th game) $= T^6 \times S_0$ $= \begin{bmatrix} 0.75 & 0.4 \\ 0.25 & 0.6 \end{bmatrix}^6 \times \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $= \begin{bmatrix} 0.616\,092 \\ 0.383\,908 \end{bmatrix}$
9 As the answer required is the probability of winning the 7th game, we need the first element of the state matrix. Round the answer to 4 decimal places.	Pr(winning the 7th game) = 0.6161

WORKED EXAMPLE 15

People from the suburb of Balwyn have easy access to two modes of public transport for travelling to work locally or into the city. The modes of transport are buses and trams. When 3000 local people were interviewed, it was found that 40% used buses on a regular basis and 60% used trams on a regular basis. It was also found on a day-to-day basis, that if a person caught a bus one day there was a 75% chance of them catching a bus again the following day. Also, if a person caught a tram one day there was a 65% chance of them catching a tram again the following day.

a For a particular five day working week, find the proportion of people using the bus and the tram on Tuesday.

b For a particular five day working week, find the proportion of people using the bus and the tram at the end of the working week, namely Friday.

c The proportion of people from the Balwyn area using the bus and the tram in the long term.

THINK | **WRITE**

a 1 Draw up a transition probability table.

a

	Bus day 1	Tram day 1
Bus next day	0.75	0.35
Tram next day	0.25	0.65
	1.00	1.00

2 Write down the transition matrix T and the initial state matrix S_0.

$\begin{bmatrix} 0.75 & 0.35 \\ 0.25 & 0.65 \end{bmatrix}$ is the transition matrix, T.

$\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$ is the initial state matrix, S_0.

3 If $S_n = T \times S_{n-1}$ then find S_2 to determine the proportion of people using each mode of transport on Tuesday.

Note: The sum of the resultant two probabilities MUST add to one.

Monday:

$$S_1 = \begin{bmatrix} 0.75 & 0.35 \\ 0.25 & 0.65 \end{bmatrix}\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

$$= \begin{bmatrix} 0.75 \times 0.4 + 0.35 \times 0.6 \\ 0.25 \times 0.4 + 0.65 \times 0.6 \end{bmatrix}$$

$$= \begin{bmatrix} 0.51 \\ 0.49 \end{bmatrix}$$

Tuesday:

$$S_2 = \begin{bmatrix} 0.75 & 0.35 \\ 0.25 & 0.65 \end{bmatrix}\begin{bmatrix} 0.51 \\ 0.49 \end{bmatrix}$$

$$= \begin{bmatrix} 0.75 \times 0.51 + 0.35 \times 0.49 \\ 0.25 \times 0.51 + 0.65 \times 0.49 \end{bmatrix}$$

$$= \begin{bmatrix} 0.554 \\ 0.446 \end{bmatrix}$$

4 State the result.

On Tuesday 55.4% of the people from Balwyn used the bus and 44.6% of the people used the tram.

b 1 Continue using $S_n = T \times S_{n-1}$ to determine the proportion of people using each mode of transport on Friday.

b Wednesday:

$$S_3 = \begin{bmatrix} 0.75 & 0.35 \\ 0.25 & 0.65 \end{bmatrix}\begin{bmatrix} 0.554 \\ 0.446 \end{bmatrix}$$

$$= \begin{bmatrix} 0.75 \times 0.554 + 0.35 \times 0.446 \\ 0.25 \times 0.554 + 0.65 \times 0.446 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5716 \\ 0.4284 \end{bmatrix}$$

Thursday:

$$S_4 = \begin{bmatrix} 0.75 & 0.35 \\ 0.25 & 0.65 \end{bmatrix}\begin{bmatrix} 0.5716 \\ 0.4284 \end{bmatrix}$$
$$= \begin{bmatrix} 0.75\times0.5716+0.35\times0.4284 \\ 0.25\times0.5716+0.65\times0.4284 \end{bmatrix}$$
$$= \begin{bmatrix} 0.57864 \\ 0.42136 \end{bmatrix}$$

Friday:

$$S_5 = \begin{bmatrix} 0.75 & 0.35 \\ 0.25 & 0.65 \end{bmatrix}\begin{bmatrix} 0.57864 \\ 0.42136 \end{bmatrix}$$
$$= \begin{bmatrix} 0.75\times0.57\,864+0.35\times0.42\,136 \\ 0.25\times0.57\,864+0.65\times0.42\,136 \end{bmatrix}$$
$$= \begin{bmatrix} 0.581\,456 \\ 0.418\,544 \end{bmatrix}$$

2 State the result.
Note: If the results for Thursday and Friday were both given correct to 2 decimal places, the final matrix values would be the same, namely 0.58 and 0.42. This is because the result is quickly approaching a steady state.

On Friday 58.1% of the people from Balwyn used the bus and 41.8% of the people used the tram.

c **1** For the long term solution define the steady state probability.

c $$\begin{bmatrix} 0.75 & 0.35 \\ 0.25 & 0.65 \end{bmatrix}\begin{bmatrix} B \\ T \end{bmatrix} = \begin{bmatrix} B \\ T \end{bmatrix}$$

where $\begin{bmatrix} B \\ T \end{bmatrix}$ represents the steady state.

It can be seen in part (b), as n becomes larger and larger the S_n matrix is changed by only very very small amounts. This is because it is approaching a steady state.

2 Write simultaneous equations for the long term probability.

$0.75B + 0.35T = B$ [1]
$0.25B + 0.65T = T$ [2]

3 Simplify each equation.

From [1]
$0.35T = 0.25B$
From [2]
$0.25B = 0.35T$

4 Both equations are identical, so express B in terms of T.

$\frac{1}{4}B = \frac{7}{20}T$

$B = \frac{7}{5}T$

5 Use the fact that $B + T = 1$, always.

$B = \frac{7}{5}T$ [1]
$B + T = 1$ [2]
And rearranging [2], we have
$B = 1 - T$ [3]
Substitute [3] into (1)
$1 - T = \frac{7}{5}T$

$1 = \frac{7}{5}T + T$

$1 = \frac{12}{5}T$

$\frac{5}{12} = T$

6 State the result.

In the long run $\frac{5}{12}$ or approximately 41.7% of the Balwyn people use the tram and $\frac{7}{12}$ or approximately 58.3% use the bus.

Exercise 11C Markov chains and transition matrices

1 **WE11** Kelly has developed a method for predicting whether or not the surf will be good on a particular day. If it is good today, there is a 65% chance it will be good tomorrow. If it is poor today, there is a 45% chance it will be poor tomorrow.

- **a** If the surf is good on Thursday, draw a tree diagram to represent the next two days.
- **b** Extending the tree diagram, calculate the probability that, if the surf is good on Thursday, it will also be good on Sunday.

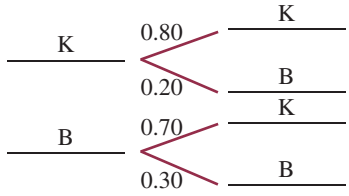
2 Mio Custor's films have either been hugely successful or have failed miserably at the box office. It is known that there is a 25% chance of Mio's next film being successful if his previous film was a success. It is also known there is a 62% chance of his next film failing if the previous film was a failure. He is currently filming a trilogy that will be released over a period of 3 years.

- **a** What is the probability that the second instalment of the trilogy will be a failure if his latest film before the trilogy was a success?
- **b** What is the probability that the final instalment of the trilogy will be a failure if his latest film before the trilogy was a success? Answer to 4 decimal places.

3 In the local cricket competition, teams can use either of two types of ball — Kingfisher or Best Match. At the end of each season, clubs sometimes decide to change the ball they use. Research suggests that 80% of those using Kingfisher stay with that ball for the next season.

An incomplete tree diagram representing this situation is shown at right. Complete the diagram and then answer the following questions.

- **a** If a club chooses Best Match one season, what is the probability it will choose Best Match the following season?
- **b** If a club chooses Best Match one season, what is the probability it will be using Kingfisher in 3 years time?



4 The probability that Alicia is on time to school in the morning is dependent upon if she is asleep by 10 pm. If she is asleep before 10 pm, the probability of her being on time to school is 0.8. If she stays up until after 10 pm the night before, the probability of her being on time to school is only 0.4. The probability that she is asleep before 10 pm is 0.6.

- **a** Calculate the probability that she is on time to school on any given day.
- **b** Given that she was on time to school, find the probability that Alicia went to bed later than 10 pm.

5 The Lo Schiavo family take a holiday every summer. They choose between a resort in Cairns and visiting relatives in Tasmania. If they fly to Cairns one year, the probability of returning to Cairns the next year is 0.4. If they decide to visit their relatives one summer, there is only a 0.3 chance of a repeat visit the following year. In a particular year, the Lo Schiavos visited their relatives in Tasmania. What is the probability that they were holidaying in Cairns two years later?

6 Every Sunday night, Priya gets takeaway. She only selects from Chinese takeaway or fish and chips. If she eats Chinese takeaway one week, the probability of her eating fish and chips the week after is 0.7. If she ate fish and chips one Sunday, the probability that she eats Chinese the next Sunday is 0.5. Given that she eats Chinese on a particular Sunday, calculate the probability that she eats Chinese takeaway on only one of the next three Sundays.

7 **WE12** Commuters travelling into the centre of Trenchtown use either the bus or train. Research has shown that each month 30% of those using the bus switch to train travel and 60% of those using the train revert to bus travel.

If, at the beginning of January, 5600 people were using the bus and 4900 were using the train to get into the city, calculate:

- **a** the number of people using the train at the beginning of May
- **b** the number of people using the bus and train in the long term.

8 **WE13** Residents of Trenchtown purchase their petrol from either Pete's Premium Petrol or Slick Sam's Servo. Research has shown that each month 10% of Pete's customers switch to Slick Sam's Servo and 20% of Sam's customers switch to Pete's Premium Petrol.

If, at the beginning of June, 2800 customers purchased their petrol from Pete and 3100 customers purchased their petrol from Sam, calculate:

- **a** the number of people purchasing petrol from Pete at the beginning of October
- **b** the number of people purchasing petrol from Pete and Sam in the long term.

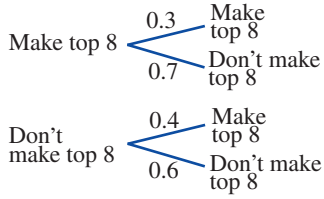
9 A leisure centre offers aerobics classes and has a gym. Records show that 10% of the members who use the gym on a particular day will participate in an aerobics class the next day, and 70% of the members who participate in an aerobics class on a particular day will use the gym the next day. It is also known that 300 members use the leisure centre each day and they all participate in aerobics classes or use the gym, but not both. On a particular day 200 members use the gym and 100 members attend an aerobics class.

a In the long term, how many people will use the gym?

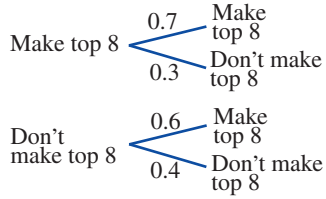
b In the long term, how many people will attend an aerobics class?

10 **MC** Records show that if a local football team make the top eight in a particular year, the chance that they make the top eight in the following year is 70%. If they don't make the top eight in a particular year, the chance that they make the top eight in the following year is 40%. The tree diagram that best represents this situation is:

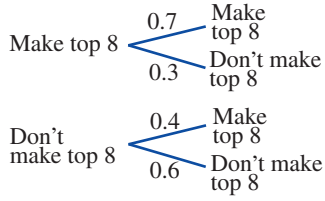
A



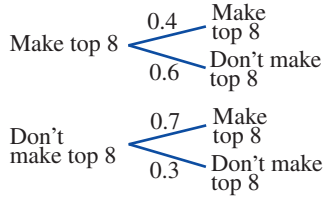
B



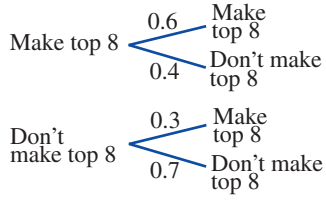
C



D



E



11 Miya prefers to shop at either Southland or Chadstone each weekend. The place she shops at depends only on where she shopped the previous time. If she visited Southland one weekend, the next weekend she goes shopping, the probability of her returning to Southland is $\frac{1}{4}$. The transition matrix for the probabilities of Miya visiting either Chadstone or Southland given the shopping centre she visited the weekend before is $\begin{bmatrix} \frac{1}{4} & \frac{2}{5} \\ \frac{3}{4} & \frac{3}{5} \end{bmatrix}$.

a If Miya shops at Chadstone one weekend, what is the probability she shops at Southland the following weekend?

b Miya does visit Chadstone on a particular weekend. What is the probability, correct to 3 decimal places, that she will be at Chadstone again in four weekend's time?

c In the long term, what proportion of weekends, correct to 3 decimal places, does Miya spend at Chadstone?

12 **WE14** The chance of Paul hitting a bullseye in darts is dependent on the success of his previous throw. If he hits a bullseye, then the chance that his next throw will also be a bullseye is 0.65. If he misses, though, there is only a 0.15 chance that he will get a bullseye on his next throw. Given that Paul's first throw is a bullseye, find the probability, correct to 4 decimal places, that:

a he hits the bullseye on his next two throws, but misses on the third

b on his tenth throw, he gets a bullseye.

eBook*plus*

DIGITAL DOC
doc-9229
WorkSHEET 11.2

13 Markov chains have a long history in computer generated music. They work by analysing the change of a given pitch going directly to any other pitch. Suppose it is known that C is the most common pitch to start a composition with a 65% rating. C is followed by A with an 18% rating and then by B with a 17% rating. When composing music, for each possible current state, there are three possible next states. Each column in the transition matrix below indicates the relative probability of going to the next state.

$$T = \begin{array}{c} \\ \text{A} \\ \text{B} \\ \text{C} \end{array} \begin{array}{c} \begin{array}{ccc} \text{A} & \text{B} & \text{C} \end{array} \\ \begin{bmatrix} 0.24 & 0.35 & 0.70 \\ 0.48 & 0.20 & 0.14 \\ 0.28 & 0.45 & 0.16 \end{bmatrix} \end{array}$$

If you are currently on A, there is a 0.24 probability of repeating A, a 0.48 probability of going to B and a 0.28 probability of going to C.

a Define the initial state matrix.

b In a computer music composition, what are the respective probabilities for the fourth note to be either A, B or C? Give answers correct to 4 decimal places.

14 **WE15** An investigation, at the beginning of 2007, into the movement between inner city living (I) and suburban living (S) of major cities in Australia found that on an annual basis, 40% of the inner city population moved to the suburbs and 30% of the suburban population moved to the inner city. It was also found that on average 55% of the city's population were inner city dwellers and 45% were suburban dwellers.

a Define the initial state matrix and the transition matrix.

b If another urban study was taken at the beginning of 2012, what proportion of people would be expected to be living in the inner city and in the suburbs after the 3 year period? Give your answers correct to 1 decimal place.

c Find the proportion of the population who will be city dwellers and suburban dwellers in the long run. Give your answers correct to 2 decimal places.

15 Auto-immune diseases affect around 1 in 20 people and are one of the important health issues in Australia. The causes of auto-immune diseases are not yet known, however, in many cases it appears that there is some inherited tendency to develop auto-immune diseases. Some examples of auto-immune diseases are multiple sclerosis, rheumatoid arthritis, psoriasis and lupus.

A researcher investigating auto-immune diseases, constructed the following transition matrix, to further investigate the possibility of an off-spring from a union developing an auto-immune disease.

Parent's auto-immune disease history

$$T = \text{off-spring's history} \begin{array}{c} \\ \text{A} \\ \text{A}' \end{array} \begin{array}{c} \begin{array}{cc} \text{A} & \text{A}' \end{array} \\ \begin{bmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{bmatrix} \end{array}$$

A = has auto-immune disease A′ = does not have auto-immune disease

a Define the initial state matrix.

b Determine the probability of a child having or not having an auto-immune disease. Give your answers correct to 3 decimal places.

c For the long term solution, define the steady state probability.

d Find the proportions (as fractions) of family descendants who will or will not suffer from an auto-immune disease in the long run.

16 The following study was the result of a socio-economic investigation completed in the USA. The study investigated if the income of an offspring from a legal or de facto union, when he/she was an adult, depended on the income of the occupation of the parents. The following transition matrix was given:

Parent's income

$$T = \text{offspring's income} \begin{array}{c} \\ \text{L} \\ \text{M} \\ \text{H} \end{array} \begin{array}{c} \begin{array}{ccc} \text{L} & \text{M} & \text{H} \end{array} \\ \begin{bmatrix} 0.6 & 0.15 & 0.05 \\ 0.2 & 0.6 & 0.3 \\ 0.2 & 0.25 & 0.65 \end{bmatrix} \end{array}$$

L = Low income, M = Medium income, H = High income

The initial state matrix of $S_0 = \begin{array}{c} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \text{L} \\ \text{M} \\ \text{H} \end{array}$ also allows us to investigate the descendants of low income parents.

a Explain each of the entries in the first row of the transition matrix.

b What is the probability that the great grandchild of a low income earner family will earn at a high income level? Give your answers correct to 2 decimal places.

c In the long run, what proportions of the offspring of low income earners will be earning at a low, medium or high income levels? Give your answers correct to 1 decimal place.

17 Consider again the socio-economic study described in question **16**.

a Define the initial state matrix appropriate to enable us to investigate the descendents of a high income earners.

b What is the probability that the great great grandchild of high income earning parents will earn at:

i a high level?

ii medium level?

iii low level?

Give answers correct to 2 decimal places.

18 A delivery truck delivers goods to/between three adjacent municipalities: Birmingham (B), Kurumburra (K) and Redhorse (R). It is known that 35% of all deliveries go to Birmingham, 35% of all deliveries go to Kurumburra and 30% of all deliveries go to Redhorse. The transition matrix for the truck deliveries is as follows:

$$T = T_0 \quad \begin{array}{c} \\ \text{B} \\ \text{K} \\ \text{R} \end{array} \begin{array}{c} \text{From} \\ \begin{array}{ccc} \text{B} & \text{K} & \text{R} \end{array} \\ \begin{bmatrix} 0.1 & 0.2 & 0.5 \\ 0.5 & 0.2 & 0.5 \\ 0.4 & 0.6 & 0.0 \end{bmatrix} \end{array}$$

a Define the initial state matrix.

b Find the probability that the fifth delivery for the day is to Redhorse. Give your answers correct to 4 decimal places.

c Find the percentage of the deliveries that go to each of the municipalities in the long run. Give your answers correct to 1 decimal place.

19 Each morning a young woman leaves her house and goes for a run. She has three pairs of running shoes that she can choose to wear. One pair has red trimmings, another pair has blue trimmings and a third pair has yellow trimmings. As the pair with red trimmings (R) are the oldest and the most comfortable, she chooses them 60% of the time, while the blue trim shoes (B) are chosen 25% of the time and the yellow trim shoes (Y) are chosen 15% of the time. The transition matrix for the choice of running shoes is as follows:

$$T = \text{Next day} \quad \begin{array}{c} \\ \text{R} \\ \text{B} \\ \text{Y} \end{array} \begin{array}{c} \text{Day One} \\ \begin{array}{ccc} \text{R} & \text{B} & \text{Y} \end{array} \\ \begin{bmatrix} 0.5 & 0.4 & 0.4 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.2 & 0.3 \end{bmatrix} \end{array}$$

a Define the initial state matrix.

b Let day 1 be Monday. What is the probability that she will be wearing her favourite shoes with red trimmings on the following Friday? Give your answers correct to 2 decimal places.

c Find the long term likelihood that she would wear each of the pairs of running shoes. Give your answers correct to 1 decimal place.

20 A person can buy their lunch from the company canteen. They have a choice from the following options.

- A sandwich with fruit (S) chosen by 25% of the workers on a regular basis.
- A chicken salad with fruit (C) chosen by 30% of the workers on a regular basis.
- Lasagne with fruit (L) chosen by 45% of the workers on a regular basis.

The workers who buy a sandwich one day have a 40% chance of buying it again the next day and a 30% chance of buying a chicken salad the next day. The workers who buy a chicken salad one day have a 50% chance of buying it again the next day and a 40% chance of buying

a lasagne the next day. The workers who buy a lasagne one day have a 50% chance of buying it again the next day and a 10% chance of buying a sandwich the next day.

a Define the initial state matrix and the transition matrix.

b What percentage of people chose a chicken salad on the Thursday of a working week? Give your answers correct to 1 decimal place.

c Find the percentage of people who consume each of the lunch options in the long run.

Give your answers correct to 2 decimal places.

21 Two hundred members of a particular gym were questioned about their attendance patterns. 20% said their attendance was 3 or more times per week (H), 35% said their attendance was 1–2 times per week (M) and the remainder said their attendance was 0–1 times per week (L). It was also discovered that if a person attended 3 or more times per week for one week, they had a 50% chance of maintaining that pattern the next week but a 10% chance of dropping to 1–0 times per week the next week. If a person attended 1–2 times per week for one week, they had a 35% chance of maintaining that attendance pattern the next week and a 40% chance of upgrading the attendance to 3 or more times the following week. Finally, if a person attended 0–1 times for one week, they had a 80% chance of maintaining that attendance pattern the following week and only a 5% chance of improving their attendance pattern to 3 or more time per week the following week.

a Define the initial state matrix and the transition matrix.

b What percentage of the gym members attended 1–2 times per week after 4 continuous weeks? Give your answers correct to 2 decimal places.

c Find the percentage of the gym members who attended at H, M and L levels in the long run. Give your answers correct to 1 decimal place.

11D Expected value, variance and standard deviation of the binomial distribution

studyon

Units:	3 & 4
AOS:	4
Topic:	2
Concept:	2

Concept summary
Read a summary of this concept.

When working with the binomial probability distribution, (like other distributions) it is very useful to know the expected value (mean), variance and the standard deviation.

The random variable, X, is such that $X \sim \text{Bi}(8, 0.3)$ and has the following probability distribution.

x	0	1	2	3	4	5	6	7	8
$\Pr(X = x)$	0.057 65	0.197 65	0.296 48	0.254 12	0.136 14	0.046 68	0.010 00	0.001 22	0.000 07

In chapter 10 we saw that the expected value, $\text{E}(X)$ was defined as $\text{E}(X) = \sum x \Pr(X = x)$. Hence, the expected value for the above table is:

$$\begin{aligned}\text{E}(X) &= \sum x \Pr(X = x).\\ &= 0 \times 0.057\,65 + 1 \times 0.197\,65 + 2 \times 0.296\,48 + 3 \times 0.254\,12 + 4 \times 0.136\,14 + 5 \times 0.046\,68 + 6 \times 0.010\,00 + 7 \times 0.001\,22 + 8 \times 0.000\,07\\ &= 0 + 0.197\,65 + 0.592\,96 + 0.762\,36 + 0.544\,56 + 0.233\,400 + 0.060\,00 + 0.008\,54 + 0.000\,56\\ &= 2.4\end{aligned}$$

The variance was defined by the rule $\text{Var}(X) = \text{E}(X^2) - [\text{E}(X)]^2$. Hence, the variance for the above table is:

$$\begin{aligned}\text{Var}(X) &= \text{E}(X^2) - [\text{E}(X)]^2\\ &= 0^2 \times 0.057\,65 + 1^2 \times 0.197\,65 + 2^2 \times 0.29\,648 + 3^2 \times 0.254\,12 + 4^2 \times 0.136\,14 + 5^2 \times 0.046\,68 + 6^2 \times 0.010\,00 + 7^2 \times 0.001\,22 + 8^2 \times 0.000\,07 - (2.4)^2\\ &= 0 + 0.19\,765 + 1.185\,92 + 2.287\,08 + 2.178\,24 + 1.166\,89 + 0.360\,01 + 0.060\,011 + 0.004\,194 - (2.4)^2\\ &= 7.44 - (2.4)^2\\ &= 1.68\end{aligned}$$

The standard deviation was defined by the rule $\text{SD}(X) = \sqrt{\text{Var}(X)}$. Hence the standard deviation for the above table is:

$$\text{SD}(X) = \sqrt{\text{Var}(X)} = \sqrt{1.68} = 1.30$$

Since this method for obtaining the expected value, variance and the standard deviation is tedious and time consuming, a quicker method has been developed to calculate these terms. It can be shown that if $X \sim \text{Bi}(n, p)$ then:

$$\text{E}(X) = np \qquad \text{Var}(X) = npq \qquad \text{SD}(X) = \sqrt{npq}$$

To check that these agree with the previous example, we will substitute the values into the given rules. When $X \sim \text{Bi}(8, 0.3)$, we obtain the following.

The expected value:
$$\begin{aligned}\text{E}(X) &= np \\ &= 8 \times 0.3 \\ &= 2.4\end{aligned}$$

The variance:
$$\begin{aligned}\text{Var}(X) &= npq \\ &= 8 \times 0.3 \times 0.7 \\ &= 1.68\end{aligned}$$

The standard deviation:
$$\begin{aligned}\text{SD}(X) &= \sqrt{npq} \\ &= \sqrt{1.68} \\ &= 1.30\end{aligned}$$

As can be seen, these values correspond to those obtained earlier. A great deal of time is saved using these rules and the margin for making mistakes is reduced.

Hence, if X is a random variable and $X \sim \text{Bi}(n, p)$ then:

$$\text{E}(X) = np \qquad \text{Var}(X) = npq \qquad \text{SD}(X) = \sqrt{npq}$$

Note: The distribution must be binomial for these rules to apply.

WORKED EXAMPLE 16

The random variable X follows a binomial distribution such that $X \sim \text{Bi}(40, 0.25)$. Determine the:
a expected value **b variance and standard deviation.**

THINK	WRITE
a 1 Write the rule for the expected value.	**a** $\text{E}(X) = np$
2 List the values for n and p.	$n = 40, p = 0.25$
3 Substitute the values into the rule.	$\text{E}(X) = 40 \times 0.25$
4 Evaluate.	$= 10$
b 1 Write the rule for the variance.	**b** $\text{Var}(X) = npq$
2 Write the values for n, p and q.	$n = 40, p = 0.25, q = 0.75$
3 Substitute the values into the rule.	$\text{Var}(X) = 40 \times 0.25 \times 0.75$
4 Evaluate.	$= 7.5$
5 Write the rule for the standard deviation.	$\text{SD}(X) = \sqrt{npq}$
6 Substitute the value obtained for the variance and take the square root.	$= \sqrt{7.5}$
7 Evaluate.	$= 2.74$

WORKED EXAMPLE 17

A fair die is rolled 15 times. Find:
a the expected number of 3s rolled
b the probability of obtaining more than the expected number of 3s.

THINK	WRITE
a 1 Write the rule for the expected value.	**a** $\text{E}(X) = np$
2 List the values for n and p.	$n = 15, p = \frac{1}{6}$

3 Substitute the values into the rule.	$E(X) = 15 \times \frac{1}{6}$
4 Evaluate.	$= 2.5$
b 1 State the probability distribution.	**b** $X \sim \text{Bi}(n, p)$ $X \sim \text{Bi}\left(15, \frac{1}{6}\right)$
2 Write what is required. *Note:* Since X represents the number of 3s, X can have only whole number values.	$\Pr(X > 2.5) = \Pr(X \geq 3)$ $= 1 - \Pr(X < 3)$ $= 1 - [\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)]$
3 Write the rule for the binomial probability distribution. Then substitute the values of n, p and q into the rule.	$= 1 - \left[{}^{15}C_0\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{15} + {}^{15}C_1\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^{14} + {}^{15}C_2\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{13}\right]$
4 Evaluate and round the answer to 4 decimal places.	$= 1 - (0.064\,905\,471\,5 + 0.194\,716\,414\,6$ $+ 0.272\,602\,980\,4)$ $= 1 - 0.532\,224\,866\,5$ $= 0.467\,775\,133\,5$ ≈ 0.4678
Part **b** can also be solved with a CAS calculator.	$\Pr(X > 2.5) = \Pr(X \geq 3)$ $= \text{binomCdf}(15, \frac{1}{6}, 3, 15)$ $= 0.4678$

WORKED EXAMPLE 18

A binomial random variable has an expected value of 14.4 and a variance of 8.64. Find:
a the probability of success, p
b the number of trials, n.

THINK	WRITE
a 1 Write what is known and what is required.	**a** $E(X) = 14.4$ so $np = 14.4$ $\text{Var}(X) = 8.64$ so $npq = 8.64$ $p = ?$
2 Substitute the value of np into the variance equation.	$npq = 8.64$ $14.4q = 8.64$
3 Transpose the equation to make q the subject.	$q = \frac{8.64}{14.4}$
4 Evaluate.	$q = 0.6$
5 Solve for p using the relationship $q = 1 - p$.	$q = 1 - p$ $0.6 = 1 - p$ $p = 0.4$
b 1 Write what is known and what is required.	**b** $np = 14.4$ where $p = 0.4$ $n = ?$
2 Substitute the values into the equation.	$n \times 0.4 = 14.4$
3 Transpose the equation to make n the subject.	$n = \frac{14.4}{0.4}$ $= 36$

WORKED EXAMPLE 19

eBook*plus*

TUTORIAL
eles-1236
Worked example 19

A new test designed to assess the reading ability of students entering high school showed that 10% of the students displayed a reading level that was inadequate to cope with high school.

a If 400 students are selected at random, find the expected number of students whose reading level is inadequate to cope with high school.

b Determine the standard deviation of students whose reading level is inadequate to cope with high school and hence calculate $\mu \pm 2\sigma$.

c Discuss the results obtained in part b.

THINK	WRITE
a 1 Write the rule for the expected value.	a $E(X) = np$
2 Write the values for n and p.	$n = 400,\ p = 0.1$
3 Substitute the values into the rule.	$E(X) = 400 \times 0.1$
4 Evaluate.	$= 40$
5 Interpret the answer.	It is expected that 40 students will have a reading level that is inadequate to cope with high school.
b 1 Write the rule for the variance.	b $Var(X) = npq$
2 Write the values for n, p and q.	$n = 400,\ p = 0.1,\ q = 0.9$
3 Substitute the values into the rule.	$Var(X) = 400 \times 0.1 \times 0.9$
4 Evaluate.	$= 36$
5 Write the rule for the standard deviation.	$SD(X) = \sqrt{npq}$
6 Substitute the value obtained for the variance and take the square root.	$= \sqrt{36}$
7 Evaluate.	$= 6$
8 Calculate $\mu - 2\sigma$.	$\mu - 2\sigma = 40 - 2 \times 6$ $= 28$
9 Calculate $\mu + 2\sigma$.	$\mu + 2\sigma = 40 + 2 \times 6$ $= 52$
c Interpret the results obtained in part b. *Note:* Recall that for many random variables, approximately 0.95 of the probability distribution lies within two standard deviations of the mean.	c There is a probability of 0.95 that between 28 to 52 students (inclusive) will have a reading level that is inadequate to cope with high school.

Exercise 11D Expected value, variance and standard deviation of the binomial distribution

eBook*plus*

DIGITAL DOC
doc-9227
Spreadsheet
Binomial distribution

In questions **1**, **2** and **3**, assume we have a binomial distribution with number of trials, n and probability of success, p, as given.

1 WE16a Determine the mean if:

a $n = 10$ and $p = 0.6$
b $n = 8$ and $p = 0.2$
c $n = 100$ and $p = 0.5$
d $n = 50$ and $p = \frac{3}{4}$

2 WE16b Determine the variance if:

a $n = 20$ and $p = 0.6$
b $n = 15$ and $p = 0.9$
c $n = 25$ and $p = 0.4$
d $n = 20$ and $p = \frac{1}{4}$

3 **WE16b** Determine the standard deviation if:

a $n = 10$ and $p = 0.2$
b $n = 30$ and $p = 0.5$
c $n = 50$ and $p = 0.7$
d $n = 72$ and $p = \frac{2}{4}$

4 A fair coin is tossed 10 times. Find:

a the expected number of heads
b the variance for the number of heads
c the standard deviation for the number of heads.

5 A card is selected at random from a standard playing pack of 52 and then replaced. This procedure is completed 20 times. Find:

a the expected number of picture cards
b the variance for the number of picture cards
c the standard deviation for the number of picture cards.

6 Six out of every 10 cars manufactured are white. Twenty cars are randomly selected. Find:

a the expected number of white cars
b the variance for the number of white cars
c the standard deviation for the number of white cars.

7 **WE17** A fair die is rolled 10 times. Find:

a the expected number of 2s rolled
b the probability of obtaining more than the expected number of 2s.

8 A fair die is rolled 15 times. Find:

a the expected number of prime values
b the probability of obtaining more than the expected number of prime values.

9 A mathematics exam contains 27 multiple-choice questions, each with 5 possible answers. Find:

a the expected number of correct answers if a student guessed each question
b the probability of obtaining more than the expected number of correct answers.

10 A new drug that is being trialled has a success rate of 35%. Of the next 80 patients who trial the drug, find:

a the expected number of patients who will be cured
b the probability that more than the expected number of patients will be cured.

11 Eighty per cent of rabbits that contract a certain disease will die. If a group of 120 rabbits contract the disease, how many would you expect to:

a die?
b live?

12 **WE18** A binomial random variable has a mean of 10 and a variance of 5. Find:

a the probability of success, p
b the number of trials, n.

13 A binomial random variable has a mean of 12 and a variance of 3. Find:

a the probability of success, p
b the number of trials, n.

14 X is a binomial random variable with a mean of 9 and a variance of 6. Find:

a the probability of success, p
b the number of trials, n
c $\Pr(X = 10)$.

15 If X is a binomial random variable with $\mathrm{E}(X) = 3$ and $\mathrm{Var}(X) = 2.4$, find:

a the probability of success, p
b the number of trials, n
c $\Pr(X = 10)$
d $\Pr(X \leq 2)$.

16 **MC** The expected number of heads in 20 tosses of a fair coin is:

A $\frac{1}{2}$
B 5
C 10
D 15
E 20

17 **MC** Jenny is a billiards player who knows from experience that 7 out of every 10 shots she plays will result in a ball being potted. If she has 40 shots, the number of balls she expects to pot is:

A 7
B 14
C 21
D 25
E 28

18 MC The variance of the number of balls that Jenny pots from her 40 shots in question **17** is:

A 2.898 **B** 7.3 **C** 8.4 **D** 22.2 **E** 28

19 MC Eighty per cent of children are immunised against a certain disease. A sample of 200 children is taken. The mean and variance of the number of immunised children is:

A 80 and 5.66 respectively **B** 80 and 32 respectively **C** 100 and 50 respectively
D 160 and 5.66 respectively **E** 160 and 32 respectively

20 MC A binomial random variable has a mean of 10 and a variance of 6. The values of n and p respectively are:

A 5 and $\frac{2}{5}$ **B** 5 and $\frac{3}{5}$ **C** 20 and $\frac{3}{5}$ **D** 25 and $\frac{2}{5}$ **E** 25 and $\frac{3}{5}$

21 MC A Bernoulli trial has a probability of success, p. If 5 trials are conducted, the probability of three successes is:

A p^3q^2 **B** p^2q^3 **C** $10p^2q^3$ **D** $10p^3(1-p)^2$ **E** p^3

22 A binomial experiment is completed 20 times, with the expected number of successes being 16. Find:

a the probability of success, p **b** the variance
c the standard deviation.

23 A multiple-choice test has 20 questions with five different choices of answer for each question. If the answers to each question are guessed, find:

a the probability of getting 50% of the questions correct
b the probability of getting at least three correct.

24 Four per cent of pens made at a certain factory do not work. If pens are sold in boxes of 25, find the probability that a box contains more than the expected number of faulty pens.

25 A biased coin is tossed 10 times. Let X be the random variable representing the number of tails obtained. If X has an expectation of 3, find:

a the probability of obtaining exactly two tails
b the probability of obtaining no more than two tails.

26 Eighty per cent of Melbourne households have DVD players. A random sample of 500 households is taken. How many would you expect to have DVD players?

27 The executive committee of a certain company contains 15 members. Find the probability that more females than males will hold positions if:

a males and females are equally likely to fill any position
b females have a 55% chance of holding any position
c females have a 45% chance of holding any position.

28 A statistician estimates the probability that a spectator at a Carlton versus Collingwood AFL match barracks for Carlton is $\frac{1}{2}$. At an AFL grand final between these two teams there are 10 000 spectators. Find:

a the expected number of Carlton supporters
b the variance of the number of Carlton supporters
c the standard deviation of the number of Carlton supporters.

29 Thirty children are given five different yoghurts to try. The yoghurts are marked A to E, and each child has to select his or her preferred yoghurt. Each child is equally likely to select any brand. The company running the tests manufactures yoghurt B.

a How many children would the company expect to pick yoghurt B?
b The tests showed that half of the children selected yoghurt B as their favourite. What does this tell the company manufacturing this product?

30 The proportion of defective fuses made by a certain company is 0.02. A sample of 30 fuses is taken for quality control inspection.

a Find the probability that there are no defective fuses in the sample.
b Find the probability that there is only one defective fuse in the sample.
c How many defective fuses would you expect in the sample?
d The hardware chain that sells the fuses will accept the latest batch for sale only if, upon inspection, there is at most one defective fuse in the sample of 30. What is the probability that they accept the batch?
e Ten quality control inspections are conducted monthly for the hardware chain. Find the probability that all of these inspections will result in acceptable batches.

31 **WE19** A new test designed to assess the reading ability of students entering high school showed that 10% of the students displayed a reading level that was inadequate to cope with high school.

- **a** If 1600 students are selected at random, find the expected number of students whose reading level is inadequate to cope with high school.
- **b** Determine the standard deviation of students whose reading level is inadequate to cope with high school, and hence calculate $\mu \pm 2\sigma$.
- **c** Discuss the results obtained in part **b**.

32 The success rate of a new drug that is being trialled is 70%.

- **a** If 1800 patients are selected at random, find the expected number of patients cured.
- **b** Determine the standard deviation of patients cured, and hence calculate $\mu \pm 2\sigma$.
- **c** Discuss the results obtained in part **b**.

eBook*plus*

DIGITAL DOC
doc-9230
Investigation
Winning at racquetball!

Summary

The binomial distribution

- The binomial distribution is an example of a discrete probability distribution.
- The binomial distribution may be referred to as a Bernoulli distribution, and the trials conducted are known as Bernoulli trials.
- For a sequence to be defined as a Bernoulli sequence, each of the following characteristics must be satisfied.
  1. n independent trials must be conducted.
  2. Only two possible outcomes must exist for each trial, that is, success (p) and failure (q).
  3. The probability of success, p, is fixed for each trial.
- If X represents a random variable that has a binomial distribution, then it can be expressed as $X \sim \text{Bi}(n, p)$ or $X \sim \text{B}(n, p)$. This means that X follows a binomial distribution with parameters n (the number of trials) and p (the probability of success).
- If X is a binomial random variable its probability is defined as:

$$\Pr(X = x) = {}^nC_x p^x q^{n-x} \text{ where } x = 0, 1, 2, \ldots n.$$

The effects of *n* and *p* on binomial distribution graphs

- The parameters n and p affect the binomial probability distribution curve as follows.
  1. If $p < 0.5$, the graph is positively skewed.
  2. If $p = 0.5$, the graph is symmetrical or is a normal distribution curve.
  3. If $p > 0.5$, the graph is negatively skewed.
  4. When n is very large and $p = 0.5$, the interval between the vertical columns decreases and the graph approximates a smooth hump or bell shape.

Problems involving the binomial distribution for multiple probabilities

- When solving problems dealing with the binomial distribution for multiple probabilities, always:
  1. define the distribution
  2. write what is required
  3. write the rule for the binomial probability distribution
  4. substitute the values into the given rule and evaluate.

Markov chains and transition matrices

- A Markov chain is a sequence of repetitions of an experiment in which:
  1. The probability of a particular outcome in an experiment is *conditional* only on the result of the outcome immediately before it.
  2. The *conditional probabilities* of each outcome in a particular experiment are the *same* every single time.
- A two-state Markov chain is where there are only two possible outcomes for each experiment.
- Problems involving two-state Markov chains can be solved using:
  1. tree diagrams where the probabilities are multiplied along the branches
  2. recurrence relationships
  3. transition matrices where $T = \begin{bmatrix} \Pr(A_2 \mid A_1) & \Pr(A_2 \mid A_1') \\ \Pr(A_2' \mid A_1) & \Pr(A_2' \mid A_1') \end{bmatrix}$,

$$S_0 = \text{the initial state}, \begin{bmatrix} n(A_1)_0 \\ n(A_2)_0 \end{bmatrix} \text{or} \begin{bmatrix} \Pr(A_1)_0 \\ \Pr(A_2)_0 \end{bmatrix}$$

$$S_n = T^n \times S_0$$

To find the long-term proportion or steady state, let n = a large number, relative to the problem, and solve for S_n.

Expected value, variance and standard deviation of the binomial distribution

- If X is a random variable and $X \sim \text{Bi}(n, p)$ then:

$$\text{E}(X) = np$$
$$\text{Var}(X) = npq$$
$$\text{SD}(X) = \sqrt{npq}$$

This applies only to a binomial distribution.

Chapter review

SHORT ANSWER

1 Zoe tosses a fair coin 5 times. Calculate:
- **a** the probability of obtaining 3 heads
- **b** the probability of obtaining at least 1 head
- **c** the probability of obtaining 4 heads then a tail.

2 Matt is a dedicated football player and knows that his chance of scoring a goal on any one kick is 0.9. If he has three shots at goal in a row, calculate the probability that:
- **a** he scores two goals
- **b** all of his shots miss
- **c** after missing the first, he scores a goal on each of his next two shots.

3 Nisha is pulling marbles out of a bag containing 3 green and 9 red ones. After each selection she replaces the marble. In total she selects four marbles. What is the probability she selects:
- **a** two red marbles?
- **b** two red marbles, given that at least one red marble was selected?

4 One out of three people read the *Bugle* newspaper, while one out of five read the *Headline*. If four people are sampled, what is the probability that:
- **a** one of them reads the *Headline*?
- **b** at least three of them read the *Bugle*?

5 A cafe prepares its macchiatos with either skinny or regular milk. It is known that 20% of customers who have a regular macchiato on a particular day will have a skinny macchiato the next day. Also, 70% of customers who have a skinny macchiato on a particular day will have a regular macchiato the next day. Given a particular customer has a skinny macchiato on Monday, calculate the probability:
- **a** they had a regular macchiato on Tuesday and a skinny macchiato on Wednesday
- **b** they had a skinny macchiato on Wednesday.

6 Hilary loves going to the movies every Friday night, and prefers either comedy or action films. If she watches a comedy one week, the probability that she sees a comedy the following week is 0.1. If she sees an action movie on a particular Friday, there is a probability of 0.4 that she will see an action movie the following Friday. Given that Hilary saw a comedy on the first Friday in March, calculate the probability:
- **a** she sees an action film on the third Friday of the month
- **b** she saw a comedy on the 2nd Friday, given that she saw an action film on the third Friday of the month.

7 Maria and Patrick love to play a game of tennis each week. The probability of Maria beating Patrick if she wins the previous week is 0.7; however, if she loses a match, the chance of her beating him the following week is only 0.2. Given that Maria beats Patrick in their first match of the year, calculate the probability that she will beat him exactly one more time in the next three weeks.

8 Calculate:
- **i** the mean
- **ii** the variance

for the binomial random variables with n and p given by:
- **a** $n = 100$ and $p = 0.5$
- **b** $n = 50$ and $p = 0.8$.

9 A binomial random variable has a mean of 10 and variance of 8. Calculate:
- **a** the probability of success, p
- **b** the number of trials, n.

10 A mathematics exam contains 40 multiple-choice questions, each with 5 possible answers. Calculate:
- **a** the expected number of correct answers if a student guessed each question
- **b** the variance of the number of correct answers if a student guessed each question.

MULTIPLE CHOICE

1 Three out of every 7 students ride bikes to school. Twenty students are randomly selected. The probability that 8 of these rode to school today is:
A 0.0951 **B** 0.1426 **C** 0.1738 **D** 0.3178 **E** 0.4916

2 An unbiased 8-sided die is rolled 12 times. The probability of obtaining three results greater than 5 is:
A 0.1135 **B** 0.1688 **C** 0.2188 **D** 0.2279 **E** 0.2824

3 Which of the following does not represent a Bernoulli sequence?
- **A** Rolling a die 10 times and recording the number of 5s
- **B** Rolling a die 50 times and recording the results

C Spinning a 7-sided spinner and recording the number of 4s obtained
D Surveying 100 people and asking them if they eat 'Superflakes' cereal
E Drawing a card with replacement and recording the number of red cards obtained

4 The probability that the temperature in Melbourne will rise above 25 °C on any given summer day, independent of any other summer day, is 0.65. The probability that 3 days in a week reach in excess of 25 °C is:

A 0.65^3 **B** $\frac{3}{7} \times 0.35^4$ **C** $7 \times 0.65^3 \times 0.4^4$
D $35 \times 0.65^3 \times 0.35^4$ **E** $0.65^3 \times 0.35^4$

5 A fair die is rolled 10 times. The probability of obtaining 4 even numbers is:
A 0.0004 **B** 0.2051 **C** 0.0009 **D** 0.7949 **E** 0.2461

6 If the same die from question **5** is rolled 10 times, the probability of obtaining at least 8 even numbers is:
A 0.0107 **B** 0.0390 **C** 0.0547 **D** 0.9453 **E** 0.9893

7 The probability of Sam beating Abby in a game of cards is 0.36. Abby and Sam decide to play a game every day for n days. What is the fewest number of games they need to play to ensure the probability of Sam winning at least once is greater than 0.85?
A 3 **B** 4 **C** 5 **D** 6 **E** 7

8 One in every 50 apples sold at Grubby Granny's Greengrocers has worms in it. If I buy a box of 100 apples, the probability of at least three apples containing worms is:
A 0.3233 **B** 0.3781 **C** 0.5 **D** 0.6219 **E** 0.6867

9 X is a random variable, binomially distributed with $n = 20$ and $p = \frac{3}{7}$. $\Pr(X \geq 1)$ is:

A $1 - \left(\frac{4}{7}\right)^{20}$ **B** $1 - \left(\frac{3}{7}\right)^{20}$ **C** $20\left(\frac{3}{7}\right)\left(\frac{4}{7}\right)^{19}$ **D** $\left(\frac{3}{7}\right)^{20}$ **E** $\left(\frac{4}{7}\right)^{20}$

10 Claire's position in the netball team is goal shooter. The probability of her shooting a goal is 67%. If she has 15 attempts at scoring, the probability she will score at least 7 goals is:
A ${}^{15}C_7\,(0.67)^7(0.33)^8$
B ${}^{15}C_7\,(0.33)^7(0.67)^8 + {}^{15}C_8\,(0.33)^8(0.67)^7 + \ldots + (0.33)^{15}$
C ${}^{15}C_8\,(0.67)^8(0.33)^7 + {}^{15}C_9\,(0.67)^9(0.33)^6 + \ldots + (0.67)^{15}$
D ${}^{15}C_7\,(0.67)^7(0.33)^8 + {}^{15}C_8\,(0.67)^8(0.33)^7 + \ldots + (0.67)^{15}$
E $(0.33)^{15} + {}^{15}C_1\,(0.67)(0.33)^{14} + \ldots + {}^{15}C_7\,(0.67)^7(0.33)^8$

11 The proportion of patients that suffer a violent reaction from a new drug being trialled is k. If 60 patients trial the drug, the probability that one-fifth of the patients have a violent reaction is:
A ${}^{60}C_5\,(1-k)^5(k)^{55}$ **B** ${}^{60}C_{12}\,(k)^{12}$ **C** ${}^{60}C_5\,(k)^5(1-k)^{55}$
D ${}^{60}C_{12}\,(1-k)^{12}(k)^{48}$ **E** ${}^{60}C_{12}\,(k)^{12}(1-k)^{48}$

12 Theo passes three sets of traffic lights on his way to school each morning. The lights at each intersection operate independently of each other. The probability of him having to stop for a red light is $\frac{2}{3}$ and the probability of passing through the intersection without stopping is $\frac{1}{3}$. If Theo encountered a red light at least once during a particular trip to school, what is the probability that he had to stop at exactly two intersections?

A $\frac{2}{13}$ **B** $\frac{6}{19}$ **C** $\frac{3}{5}$ **D** $\frac{6}{13}$ **E** $\frac{2}{9}$

13 Consider a Markov chain where the nth state is given by $S_n = T^n \times S_0$. Which one of the following could represent the transition matrix, T, and the initial state matrix, S_0, respectively?

A $\begin{bmatrix} 0.1 & 0.9 \\ 0.6 & 0.4 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ **B** $\begin{bmatrix} 0.6 & 0.1 \\ 0.4 & 0.9 \end{bmatrix}$ and $\begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}$ **C** $\begin{bmatrix} 15 \\ 30 \end{bmatrix}$ and $\begin{bmatrix} 0.6 & 0.1 \\ 0.4 & 0.9 \end{bmatrix}$

D $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0.25 & 0.75 \\ 0.4 & 0.6 \end{bmatrix}$ **E** $\begin{bmatrix} 0.25 & 0.45 \\ 0.85 & 0.65 \end{bmatrix}$ and $\begin{bmatrix} 10 \\ 20 \end{bmatrix}$

14 Kelly has developed a method for predicting whether or not the surf will be good on a particular day. If it is good today, there is an 83% chance it will be good tomorrow. If it is poor today, there is a 65% chance it will be poor tomorrow. The probability that the surf will be poor on Thursday given it was poor on Tuesday is:
A 0.4225 **B** 0.4820 **C** 0.0595 **D** 0.6620 **E** 0.6500

15 Two fair coins are tossed simultaneously 60 times. The number of times that they both show heads is expected to be:

A 10 **B** 15 **C** 20 **D** 30 **E** 45

16 The variance in the number of heads obtained from 50 tosses of a fair coin is:

A $\frac{1}{50}$ **B** $\frac{2}{25}$ **C** $\frac{1}{4}$ **D** $\frac{25}{2}$ **E** 3.53

17 The expected number of heads in 50 tosses of a fair coin is:

A $\frac{1}{2}$ **B** 6.25 **C** 10 **D** 2.5 **E** 25

18 A variance of 1.35 occurs when:

A $n = 20$ and $p = 0.6$ **B** $n = 15$ and $p = 0.9$ **C** $n = 25$ and $p = 0.4$

D $n = 20$ and $p = \frac{1}{4}$ **E** $n = 50$ and $p = 0.3$

EXTENDED RESPONSE

1 Speedy Saverio's Pizza House claims to cook and deliver 90% of pizzas within 15 minutes of the order being placed. If your pizza is not delivered within this time, it is free. On one busy Saturday night, Saverio has to make 150 deliveries.

a How many deliveries are expected to be made within 15 minutes of placing the order?

b What is the probability of receiving a free pizza on this night?

c If Saverio loses an average of \$4 for every late delivery, how much would he expect to lose on late deliveries this night?

2 Ten per cent of all Olympic athletes are tested for drugs at the conclusion of their event. One per cent of all athletes use performance enhancing drugs. Of the 1000 Olympic wrestlers competing from all over the world, Australia sends 10. Find:

a the expected number of Australian wrestlers who are tested for drugs

b the probability that half the Australian wrestlers are tested for drugs

c the probability that at least two Australian wrestlers are tested for drugs

d the expected number of drug users among all wrestlers.

3 Five per cent of watches made at a certain factory are defective. Watches are sold to retailers in boxes of 20. Find:

a the expected number of defective watches in each box

b the probability that a box contains more than the expected number of defective watches per box

c the probability of a 'bad batch', if a 'bad batch' entails more than a quarter of the box being defective.

4 Aiko is a keen basketball player and knows that her chance of scoring a goal on any one throw is 0.65.

a If Aiko takes 6 shots for a goal, find the probability, correct to 4 decimal places, that she:

i misses each time

ii scores a goal at least three times

iii scores a goal five times, given that she scored a goal at least three times.

b Find the number of throws Aiko would need to ensure a probability of more than 0.9 of scoring at least one goal.

5 Sixty-eight per cent of all scheduled trains through Westbourne station arrive on time. If 20 trains go through the station every day, calculate the probability, correct to 4 decimal places, that:

a no more than 10 trains are on time

b at least 12 trains are on time

c at least 12 trains are on time for 9 out of the next 10 days.

6 The success rate of a new drug being trialled is 60%.

a If 2400 patients are selected at random, find the expected number of patients cured.

b Determine the standard deviation of patients cured and hence calculate $\mu \pm 2\sigma$.

c Interpret the results found in part **b**.

7 Phil is running a stall at the local Primary School Fair involving lucky dips. It costs $2 to have a go, and contained in a large box are 80 lucky dips from which to choose. Phil claims that one in 5 lucky dips contains a prize. By the end of the day, all 80 have been sold. Calculate the probability, correct to 4 decimal places, that:

- **a** the first four people to select a lucky dip don't win a prize, but the next two do
- **b** there are at least 10 winners
- **c** there are no more than 18 prize winners, given that at least 10 people won a prize.

8 Every afternoon Anna either goes for a run or a walk. If she goes for a walk one afternoon, the probability that she goes for a run the next is 0.45, and if she decides to run one afternoon, then the probability of her going for a walk the next afternoon is 0.8. On Wednesday, Anna decides to go for a walk around the park.

- **a** What is the probability that she goes for a run on each of the next three afternoons?
- **b** What is the probability, correct to 4 decimal places, that over the next three afternoons, she goes for a run at least once?
- **c** What is the probability, correct to 4 decimal places, that on the following Wednesday Anna will decide to go for a run?
- **d** In the long term, on what proportion of afternoons will she choose to go for a run?

9 A small town has two restaurants — Kaz's Kitchen and Al's Fine Dining. Records show that 30% of customers who eat at Kaz's Kitchen one week will visit Al's Fine Dining the next. Also, 40% of customers who eat at Al's Fine Dining during the week will dine at Kaz's Kitchen the following week. There are 400 members of the town who regularly dine out once a week. During the first week of July, 300 people visit Kaz's Kitchen and Al has 100 customers at his restaurant.

- **a** How many customers will visit each restaurant in the first week of August?
- **b** In the long run, how many customers will Kaz and Al each get per week?

10 Two companies are competing for the mobile phone market. At the end of January, market research revealed the following patterns in the subscriptions of mobile phone users.

Of the 400 Tellya customers who were interviewed, 340 were staying with Tellya and 60 were changing to Yodacall.

Of the 100 Yodacall customers who were interviewed, 90 were staying with Yodacall and 10 were changing to Tellya.

- **a** Set up a pair of recurrence relationships that describes the given patterns.
- **b** What is the original state of the companies in terms of market share?
- **c** What is the state of each company at the end of the next month (February)?
- **d** What is the state, in terms of market share, of each company 4 months later (May)?
- **e** What is the state, in terms of market share, of each company 7 months later (August)?
- **f** What is the state, in terms of market share, of each company at the end of the following January?
- **g** A company will fail to be viable if its market share falls below 25%. Which, if either, of these companies will not achieve this market share in the long run?

11 The winter months in many states of Australia can be cold, windy, mild and sunny all in a matter of days. A study was made of the winter months of June, July and August during which it was found that it was cold and/or raining (R) on 45% of the days, cloudy and/or windy (W) on 35% of the days and sunny and mild (S) on 20% of the days. It was also found that the proportion of consecutive days of cold/rain was 45%, the proportion of consecutive days of cloud/wind was 55% and the proportion of consecutive days of sunny/mild was 45%. A rainy/cold day followed by a cloudy/windy day happened 40% of the time, a sunny/mild day followed by a cold/rainy day happened 25% of the time and a cloudy/windy day followed by a sunny/mild day happened 5% of the time.

- **a** Define the initial state matrix and the transition matrix.
- **b** What is the probability that the third day of three consecutive days is sunny and mild? Give your answer correct to 3 decimal places.
- **c** In the long run, find the percentage of days that will be R, W or S over the winter months. Give your answers correct to 2 decimal places.

studyON

Units:	3 & 4
AOS:	4
Topic:	2

Practice VCE exam questions

Use StudyON to access all exam questions on this topic since 2002.

eBook*plus*

DIGITAL DOC
doc-9231
Test Yourself
Chapter 11

eBookplus ICT activities

Chapter opener

DIGITAL DOC

- 10 Quick Questions doc-9226: Warm up with ten quick questions on applications of the binomial distribution. *(page 515)*

11A The binomial distribution

TUTORIALS

- WE2 eles-1191: Watch a worked example on indentifying the number of trials in an experiment. *(page 517)*
- WE5 eles-1232: Watch a worked example on constructing a probability distribution table. *(page 520)*

DIGITAL DOCS

- Spreadsheet doc-9227: Investigate the binomial distribution. *(pages 523 and 525)*
- SkillSHEET 11.1 doc-9269: Practise solving indicial equations. *(page 525)*
- WorkSHEET 11.1 doc-9228: Recognise Bernoulli sequences and calculate cumulative and non-cumulative probabilities. *(page 526)*

11B Problems involving the binomial distribution for multiple probabilities

TUTORIALS

- WE8 eles-1233: Watch a worked example on calculating probabilities using a CAS calculator. *(page 527)*
- WE 10 eles-1234: Watch a worked example on calculating probabilities using the cumulative binomial distribution. *(page 528)*

DIGITAL DOC

- SkillSHEET 11.2 doc-9270: Practise multiple probabilities. *(page 529)*

11C Markov chains and transition matrices

INTERACTIVITY int-0256

- The binomial distribution: Consolidate your understanding of the binomial distribution. *(page 532)*

TUTORIAL

- WE 12 eles-1235: Watch a worked example on calculating long-term probabilities. *(page 535)*

DIGITAL DOC

- WorkSHEET 11.2 doc-9229: Identify and perform appropriate techniques to calculate probabilities. *(page 544)*

11D Expected value, variance and standard deviation of the binomial distribution

TUTORIAL

- WE 19 eles-1236: Watch a worked example on calculating the mean, median and mode of a probability density function. *(page 550)*

DIGITAL DOCS

- Spreadsheet doc-9227: Investigate the binomial distribution. *(page 550)*
- Investigation doc-9230: Winning at racquetball *(page 553)*

Chapter review

DIGITAL DOC

- Test Yourself doc-9231: Take the end-of-chapter test to test your progress. *(page 558)*

To access eBookPLUS activities, log on to www.jacplus.com.au

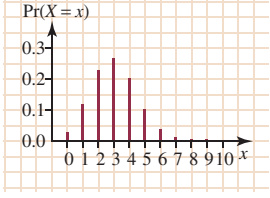
THE BINOMIAL DISTRIBUTION

Exercise 11A The binomial distribution

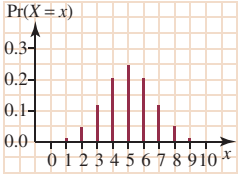
1 b, d and f constitute a Bernoulli trial.
2 a 0.2613 b 0.0446 c 0.2461
d 0.0092 e 0.0073 f 0.1969
3 a 5 b 0.3
c

x	$\Pr(X = x)$
0	0.168 07
1	0.360 15
2	0.3087
3	0.1323
4	0.028 35
5	0.002 43

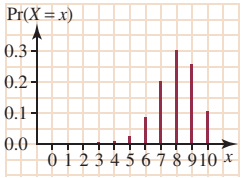
4 $\frac{4}{625}$
5 0.0036
6 a $\frac{32}{81}$ b $\frac{8}{27}$ c $\frac{65}{81}$
7 a $\frac{1}{16}$ b $\frac{1}{16}$ c $\frac{3}{8}$
8 a 0.096 b 0.0064
c 0.008 d 0.2
9 a 0.0518 b 0.2592
10 a 0.2627 b 0.0084
c 0.2568 d 0.2568
11 0.0381
12 0.0924
13 a 0.1023 b 0.2001
14 0.2281
15 0.0653
16 a 0.0528 b 0.6676
17 a 0.2734 b 0.2965 c 0.1678
18 C
19 E
20 B
21 D
22 C
23 a i 0.0102 ii 0.9898
b 2
24 18 tickets
25 22
26 a 5 645 726 b 352 858
c $1 446 717.80
27 a 2 b Positively skewed
28 a i



ii Positively skewed
b i



ii Symmetrical
c i



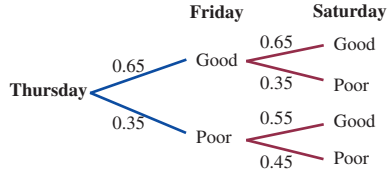
ii Negatively skewed
29 a i Positively skewed
ii Symmetrical
iii Negatively skewed
b Controls skewness
30 a Symmetrical
b The curve would resemble more of a bell shape that is tall and narrow.
31 a Positively skewed
b Negatively skewed
32 a Symmetrical
b Positively skewed
c Symmetrical
d Negatively skewed

Exercise 11B Problems involving the binomial distribution for multiple probabilities

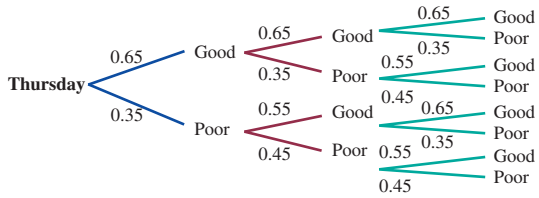
1 a 0.1792 b 0.6826
2 a 0.23 b 0.85
c 0.92 d 0.45
3 a 0.4718 b 0.9692
4 a 0.3370 b 0.3438 c 0.0333
5 a $\frac{31}{32}$ b $\frac{3}{16}$ c $\frac{6}{31}$
6 a $\frac{1}{9}$ b $\frac{48}{81}$ c $\frac{9}{65}$
7 0.028
8 a 0.0878 b 0.9822 c 0.0178
9 a 0.3669 b 0.0464
10 a 0.9830 b 0.0170
11 a 0.5000 b 0.9967
c 0.9645
12 a 0.2553 b 0.1045
13 a 0.1751 b 0.0930
14 A 15 D 16 A
17 D 18 C 19 D
20 0.1509
21 a 0.9392 b 0.1878
22 a 0.9106 b 0.9216
c 0.6767 d 0.5327
23 a 0.0173 b 0.2276 c 0.8960
24 a 0.5981 b 0.4256 c 0.2486
25 a 0.6778 b 0.0973
26 a 0.8516 b 0.2741
27 a 0.3438 b 0.0760
28 a 0.9672 b 0.9573

Exercise 11C Markov chains and transition matrices

1 a



b 0.6115



2 a 0.6525 b 0.6652
3 a 0.3

b 0.777
4 a 0.64 b 0.25
5 0.49
6 0.525
7 a 3500
b 3500 (train), 7000 (bus)
8 a 3661 b 3933 (Pete), 1967 (Sam)
9 a 265 b 35
10 C
11 a $\frac{2}{5}$ b 0.652 c 0.652
12 a 0.1479 b 0.3014
13 a

$$S_0 = \begin{array}{c} \begin{bmatrix} 0.18 \\ 0.17 \\ 0.65 \end{bmatrix} \end{array} \begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \end{array}$$

b The fourth note is A with a probability of 0.4111, B with a probability of 0.2967 and C with a probability of 0.2922.
14 a

$$S_0 = \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix} \begin{array}{c} \text{I} \\ \text{S} \end{array}$$

$$T = \begin{array}{c} \\ \text{T} \\ \text{S} \end{array} \begin{array}{c} \begin{array}{cc} \text{I} & \text{S} \end{array} \\ \begin{bmatrix} 0.7 & 0.4 \\ 0.3 & 0.4 \end{bmatrix} \end{array}$$

b After 3 years, 57.1% of the population would be inner city dwellers and 42.9% of the population would be suburban dwellers.
c In the long run, 57.14% of the population will be inner city dwellers and 42.86% will be suburban dwellers.
15 a

$$S_0 = \begin{bmatrix} 0.05 \\ 0.95 \end{bmatrix} \begin{array}{c} \text{A} \\ \text{A} \end{array}$$

Note: 1 in 20 Australians have an auto-immune disease.
b A child has a probability of 0.315 of having an auto-immune disease and a probability of 0.685 of not having an auto-immune disease.
c

$$\begin{bmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{bmatrix} \begin{bmatrix} \text{A} \\ \text{A}' \end{bmatrix} = \begin{bmatrix} \text{A} \\ \text{A}' \end{bmatrix}$$

d In the long run, a descendant has a $\frac{3}{7}$ chance of inheriting the auto-immune disease and a $\frac{4}{7}$ chance of not inheriting the auto-immune disease.

16 a 0.6 indicates the probability that if the parents are low income earners then the offspring will also be low income earners.

0.15 indicates the probability that even though parents are medium income earners the offspring will be low income earners.

0.05 indicates the probability that even though parents are high income earners the offspring will be low income earners.

b The probability that the great-grandchild of a low income earner earns at a high level is 0.35.

c In the long run, a member of the population, who descends from a low earning family, will be earning at a low level with a probability of 0.2, will be earning at a medium level with a probability of 0.4 and will be earning at a high level with a probability of 0.4.

17 a

$$S_0 = \begin{array}{c} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array} \begin{array}{c} \text{L} \\ \text{M} \\ \text{H} \end{array}$$

b The great great grandchild in a high income earning family has:

i a probability of 0.42 of earning a high income

ii a probability of 0.41 of earning a medium income

iii a probability of 0.17 of earning a low income.

18 a

$$S_0 = \begin{bmatrix} 0.35 \\ 0.35 \\ 0.30 \end{bmatrix} \begin{array}{c} \text{B} \\ \text{K} \\ \text{R} \end{array}$$

b The fifth delivery of the day has a probability of 0.3405 of going to the Redhorse municipality.

c In the long run, 38.5% of deliveries go to Kurumburra, 27.5% go to Birmingham and 34.1% go to Redhorse.

19 a

$$S_0 = \begin{bmatrix} 0.60 \\ 0.25 \\ 0.15 \end{bmatrix} \begin{array}{c} \text{R} \\ \text{B} \\ \text{Y} \end{array}$$

b There is a probability of 0.44 that she will wear her favourite running shoes on Friday.

c In the long run, the shoe with red trimmings would be worn 44.4% of the time, the shoes with blue trimmings would be 33.3% of the time and the shoes with yellow trimmings would be worn 22.2% of the time.

20 a

$$S_0 = \begin{bmatrix} 0.25 \\ 0.30 \\ 0.45 \end{bmatrix} \begin{array}{c} \text{S} \\ \text{C} \\ \text{L} \end{array}$$

$$T = \text{Next day} \begin{array}{c} \text{First day} \\ \begin{array}{ccc} \text{S} & \text{C} & \text{L} \end{array} \\ \begin{bmatrix} 0.4 & 0.1 & 0.1 \\ 0.3 & 0.5 & 0.4 \\ 0.3 & 0.4 & 0.5 \end{bmatrix} \end{array} \begin{array}{c} \\ \text{S} \\ \text{C} \\ \text{L} \end{array}$$

b 42.8% of the workers chose chicken and fruit for lunch on Thursday of a working week.

c In the long run, 14.29% of the workers order a sandwich and fruit, 42.86% order chicken salad and fruit and 42.86% order lasagne and fruit.

21 a

$$S_0 = \begin{bmatrix} 0.2 \\ 0.35 \\ 0.45 \end{bmatrix} \begin{array}{c} \text{H} \\ \text{M} \\ \text{L} \end{array}$$

$$T = \text{Next week} \begin{array}{c} \text{First week} \\ \begin{array}{ccc} \text{H} & \text{M} & \text{L} \end{array} \\ \begin{bmatrix} 0.50 & 0.40 & 0.05 \\ 0.40 & 0.35 & 0.15 \\ 0.10 & 0.25 & 0.80 \end{bmatrix} \end{array} \begin{array}{c} \\ \text{H} \\ \text{M} \\ \text{L} \end{array}$$

b After 4 weeks, 26.96% of the members would be expected to attend the gym 1–2 times per week.

c In the long run, 26.2% of the gym members have a high attendance, 27.0% have a medium attendance and 46.8% have a low attendance.

Exercise 11D Expected value, variance and standard deviation of the binomial distribution

1 a 6 **b** 1.6
c 50 **d** $37\frac{1}{2}$
2 a 4.8 **b** 1.35
c 6 **d** $3\frac{3}{4}$
3 a 1.26 **b** 2.74
c 3.24 **d** 4.16
4 a 5 **b** 2.5 **c** 1.58
5 a 4.62 **b** 3.55 **c** 1.88
6 a 12 **b** 4.8 **c** 2.19
7 a 1.67 **b** 0.5155
8 a 7.5 **b** 0.5
9 a 5.4 **b** 0.4613
10 a 28 **b** 0.4488
11 a 96 **b** 24
12 a $\frac{1}{2}$ **b** 20
13 a $\frac{3}{4}$ **b** 16
14 a $\frac{1}{3}$ **b** 27 **c** 0.1450
15 a $\frac{1}{5}$ **b** 15
c 0.0001 **d** 0.3980
16 C **17** E **18** C
19 E **20** D **21** D
22 a $\frac{4}{5}$ **b** 3.2 **c** 1.79
23 a 0.0020 **b** 0.7939
24 0.2642
25 a 0.2335 **b** 0.3828
26 400
27 a 0.5000 **b** 0.6535 **c** 0.3465
28 a 5000 **b** 2500 **c** 50
29 a 6
b The company has an extremely popular product.
30 a 0.5455 **b** 0.3340 **c** 0.6
d 0.8795 **e** 0.2769
31 a 160 students
b $\text{SD}(X) = 12$, $\mu - 2\sigma = 136$, $\mu + 2\sigma = 184$
c There is a probability of 0.95 that between 136 to 184 students (inclusive) will have a reading level that is inadequate to cope with high school.
32 a 1260 patients
b $\text{SD}(X) = 19.44$, $\mu - 2\sigma = 1221.12$, $\mu + 2\sigma = 1298.88$
c There is a probability of 0.95 that between 1222 to 1298 patients (inclusive) will be cured.

CHAPTER REVIEW

SHORT ANSWER

1 a $\frac{5}{16}$ **b** $\frac{31}{32}$ **c** $\frac{1}{32}$
2 a 0.243 **b** 0.001 **c** 0.081
3 a $\frac{27}{128}$ **b** $\frac{54}{255}$
4 a $\frac{256}{625}$ **b** $\frac{1}{9}$
5 a 0.14 **b** 0.23
6 a 0.45 **b** 0.2
7 0.234
8 a i 50 **ii** 25
b i 40 **ii** 8
9 a $p = \frac{1}{5}$ **b** $n = 50$
10 a 8 **b** 6.4

MULTIPLE CHOICE

1 C **2** B **3** B **4** D
5 B **6** C **7** C **8** A
9 A **10** D **11** E **12** D
13 B **14** B **15** B **16** D
17 E **18** B

EXTENDED RESPONSE

1 a 135 **b** 0.1 **c** \$60
2 a 1 **b** 0.0015
c 0.2639 **d** 10
3 a 1 **b** 0.2642 **c** 0.0003
4 a i 0.0018 **ii** 0.8826 **iii** 0.2761
b 3
5 a 0.0719 **b** 0.8432 **c** 0.3378
6 a 1440
b $\sigma = 24$, $\mu - 2\sigma = 1392$, $\mu + 2\sigma = 1488$
c This means that there is a probability of about 0.95 that between 1392 and 1488 of the 2400 people selected will be cured by the drug.
7 $X \sim \text{Bi}(80, \frac{1}{5})$
a 0.0164 **b** 0.9713 **c** 0.7550

8 a 0.018 **b** 0.8336
c 0.3600 **d** 0.36

9 a In the first week of August, 229 customers will visit Kaz's Kitchen and 171 people will visit Al's fine dining.
b In the long run, 229 people will dine at Kaz's Kitchen each week and Al will get 171 customers each week.

10 a Let t_i = the number of Tellya customers at the end of January month i.
Let y_i = the number of Yodacall customers at the end of January month i.
$t_{i+1} = 0.85t_i + 0.1y_i$ and
$y_{i+1} = 0.15t_i + 0.9y_i$
b Tellya has 400 customers or 80% of the market; Yodacall has 100 customers or 20% of the market.
c Tellya has 350 customers or 70% of the market, Yodacall has 150 customers or 30% of the market.
d Tellya has 52.6% of the market, Yodacall has 47.4% of the market.
e Tellya has 45.4% of the market, Yodacall has 54.6% of the market.
f Tellya has 41.2% of the market, Yodacall has 58.8% of the market.
g Tellya has 40% of the market, Yodacall has 60% of the market. Both companies will be viable, but in the long term Tellya's market share will drop by half, while Yodacall's share will increase (in fact, will triple).

11 a

$$S_0 = \begin{array}{c} \begin{bmatrix} 0.45 \\ 0.35 \\ 0.20 \end{bmatrix} \end{array} \begin{array}{c} \text{R} \\ \text{W} \\ \text{S} \end{array}$$

$$T = \text{Next day} \begin{array}{c} \text{First day} \\ \begin{array}{ccc} \text{R} & \text{W} & \text{S} \end{array} \\ \begin{bmatrix} 0.45 & 0.40 & 0.25 \\ 0.40 & 0.55 & 0.30 \\ 0.15 & 0.05 & 0.45 \end{bmatrix} \end{array} \begin{array}{c} \\ \text{R} \\ \text{W} \\ \text{S} \end{array}$$

b The probability of the third consecutive day being sunny is 0.153.
c In the long run, 39.74% of the winter month days were rainy, 45.30% were windy and 14.96% were sunny.