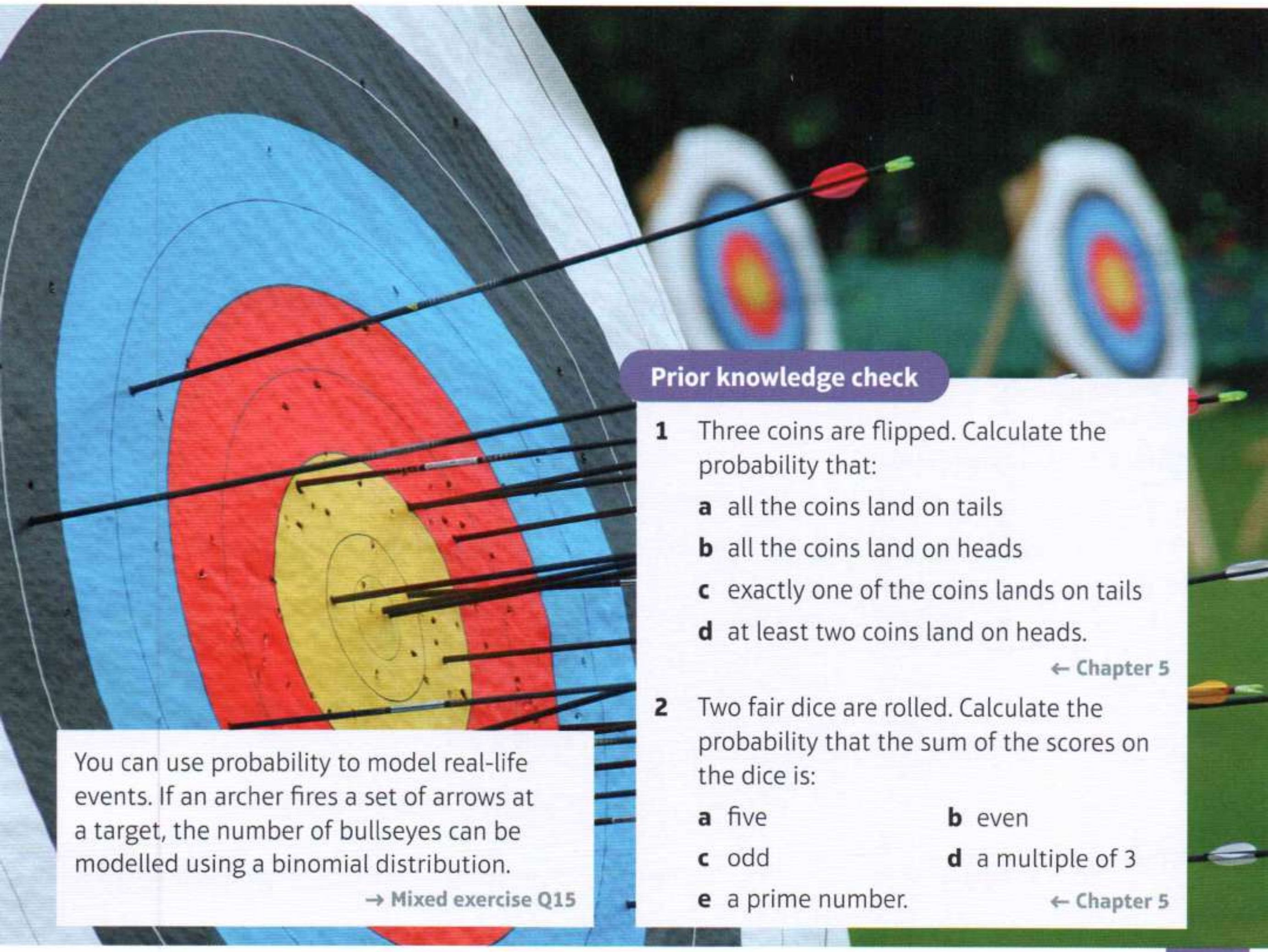
Statistical distributions

6

Objectives

After completing this chapter you should be able to:

- Understand and use simple discrete probability distributions including the discrete uniform distribution → pages 84–88
- Understand the binomial distribution as a model and comment on appropriateness → page 88
- Calculate individual probabilities for the binomial distribution → pages 89–91
- Calculate cumulative probabilities for the binomial distribution → pages 91–94



You can use probability to model real-life events. If an archer fires a set of arrows at a target, the number of bullseyes can be modelled using a binomial distribution.

→ Mixed exercise Q15

Prior knowledge check

1 Three coins are flipped. Calculate the probability that:

a all the coins land on tails

b all the coins land on heads

c exactly one of the coins lands on tails

d at least two coins land on heads.

← Chapter 5

2 Two fair dice are rolled. Calculate the probability that the sum of the scores on the dice is:

a five **b** even

c odd **d** a multiple of 3

e a prime number.

← Chapter 5

6.1 Probability distributions

A **random variable** is a variable whose value depends on the outcome of a random event.

- The range of values that a random variable can take is called its **sample space**.
- A **variable** can take any of a range of specific values.
- The variable is **discrete** if it can only take *certain* numerical values.
- The variable is **random** if the outcome is not known until the experiment is carried out.

Notation Random variables are written using upper case letters, for example X or Y.

The particular values the random variable can take are written using equivalent lower case letters, for example x or y.

The probability that the random variable X takes a particular value x is written as $P(X = x)$.

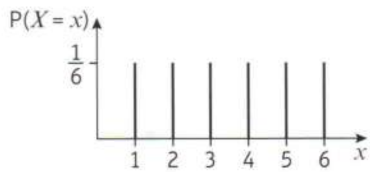
- **A probability distribution fully describes the probability of any outcome in the sample space.**

The probability distribution for a discrete random variable can be described in a number of different ways. For example, take the random variable X = 'score when a fair dice is rolled'. It can be described:

- as a **probability mass function**: $P(X = x) = \frac{1}{6}$, $x = 1, 2, 3, 4, 5, 6$
- using a table:

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

- using a diagram:



All of these representations show the probability that the random variable takes any given value in its sample space.

When all of the probabilities are the same, as in this example, the distribution is known as a **discrete uniform distribution**.

Example 1

Three fair coins are tossed.

a Write down all the possible outcomes when the three coins are tossed.

A random variable, X, is defined as the number of heads when the three coins are tossed.

b Write the probability distribution of X as:

i a table **ii** a probability mass function.

a HHH, HHT, HTH, HTT, THH, THT, TTH, TTT

These are the outcomes of the experiment.

b i

No. of heads, x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

X is the number of heads, so the sample space of X is {0, 1, 2, 3}.

These are the values of the random variable.

ii

$$P(X = x) = \begin{cases} \frac{1}{8} & x = 0, 3 \\ \frac{3}{8} & x = 1, 2 \\ 0 & \text{otherwise} \end{cases}$$

These are the values the random variable can take.

- **The sum of the probabilities of all outcomes of an event add up to 1. For a random variable X, you can write $\sum P(X = x) = 1$ for all x.**

Example 2

A biased four-sided dice with faces numbered 1, 2, 3 and 4 is rolled. The number on the bottom-most face is modelled as a random variable X.

Given that $P(X = x) = \frac{k}{x}$:

a Find the value of k.

b Give the probability distribution of X in table form.

c Find the probability that:

i $X > 2$ **ii** $1 < X < 4$ **iii** $X > 4$

a The probability distribution will be:

x	1	2	3	4
$P(X = x)$	$\frac{k}{1}$	$\frac{k}{2}$	$\frac{k}{3}$	$\frac{k}{4}$

$$\frac{k}{1} + \frac{k}{2} + \frac{k}{3} + \frac{k}{4} = 1$$

Since this is a probability distribution, $\sum P(X = x) = 1$

$$k\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right) = 1$$

$$k\left(\frac{12 + 6 + 4 + 3}{12}\right) = 1$$

$$k = \frac{12}{25}$$

Problem-solving

Write an equation and solve it to find the value of k. Then substitute this value of k into $P(X = x) = \frac{k}{x}$ for each x to find the probabilities.

b The probability distribution is:

x	1	2	3	4
$P(X = x)$	$\frac{12}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{3}{25}$

c i $X > 2$ is the same as getting 3 or 4 so

Consider all the values of X that satisfy this condition. Add the probabilities to find $P(X > 2)$.

$$P(X > 2) = \frac{4}{25} + \frac{3}{25} = \frac{7}{25}$$

ii $1 < X < 4$ is the same as getting 2 or 3

$$P(1 < X < 4) = \frac{6}{25} + \frac{4}{25} = \frac{10}{25} = \frac{2}{5}$$

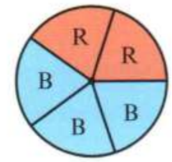
iii There are no elements in the sample space that satisfy $X > 4$ so

$$P(X > 4) = 0$$

Watch out This random variable only **models** the behaviour of the dice. The outcomes from experiments in real life will never exactly fit the model, but the model provides a useful way of analysing possible outcomes.

Example 3

This spinner is spun until it lands on red or has been spun four times in total. Find the probability distribution of the random variable S, the number of times the spinner is spun.



Problem-solving

Read the definition of the random variable carefully. Here it is the number of spins.

$P(S = 1)$ is the probability that the spinner lands on red the first time:

$P(S = 1) = \frac{2}{5}$

On any given spin, $P(\text{Red}) = \frac{2}{5}$ and $P(\text{Blue}) = \frac{3}{5}$.

If the spinner is spun twice, it must have landed on blue on its first spin:

$P(S = 2) = \frac{3}{5} \times \frac{2}{5} = \frac{6}{25}$

Each spin is an independent event so P(Blue then red) = P(Blue) × P(Red)

Likewise for landing on red on the third spin:

$P(S = 3) = \frac{3}{5} \times \frac{3}{5} \times \frac{2}{5} = \frac{18}{125}$

B, B, R is the only outcome for which $S = 3$.

The experiment stops after 4 spins so:

$P(S = 4) = 1 - \left(\frac{2}{5} + \frac{6}{25} + \frac{18}{125}\right) = \frac{27}{125}$

The sample space of S is {1, 2, 3, 4}. So $P(S = 4) = 1 - P(S = 1, 2 \text{ or } 3)$.

s	1	2	3	4
$P(S = s)$	$\frac{2}{5}$	$\frac{6}{25}$	$\frac{18}{125}$	$\frac{27}{125}$

You have found $P(S = s)$ for all values in the sample space, so you have found the complete probability distribution. You can summarise it in a table.

Exercise 6A

1 Write down whether or not each of the following is a discrete random variable. Give a reason for your answer.

a The height, X cm, of a seedling chosen randomly from a group of plants.

b The number of times, R, a six is rolled when a fair dice is rolled 100 times.

c The number of days, W, in a given week.

2 A fair dice is thrown four times and the number of times it falls with a 6 on the top, Y, is noted. Write down the sample space of Y.

3 A bag contains two discs with the number 2 on them and two discs with the number 3 on them. A disc is drawn at random from the bag and the number noted. The disc is returned to the bag. A second disc is then drawn from the bag and the number noted.

a Write down all the possible outcomes of this experiment.

The discrete random variable X is defined as the sum of the two numbers.

b Write down the probability distribution of X as:

i a table **ii** a probability mass function.

4 A discrete random variable X has the probability distribution shown in the table.

Find the value of k.

x	1	2	3	4
$P(X = x)$	$\frac{1}{3}$	$\frac{1}{3}$	k	$\frac{1}{4}$

P **5** The random variable X has a probability function

$$P(X = x) = kx \qquad x = 1, 2, 3, 4.$$

Show that $k = \frac{1}{10}$. **(2 marks)**

P **6** The random variable X has a probability function

$$P(X = x) = \begin{cases} kx & x = 1, 3 \\ k(x - 1) & x = 2, 4 \end{cases}$$

where k is a constant.

a Find the value of k. **(2 marks)**

b Find $P(X > 1)$. **(2 marks)**

P **7** The discrete random variable X has a probability function

$$P(X = x) = \begin{cases} 0.1 & x = -2, -1 \\ \beta & x = 0, 1 \\ 0.2 & x = 2 \end{cases}$$

a Find the value of β.

b Construct a table giving the probability distribution of X.

c Find $P(-1 \leqslant X < 2)$.

P **8** A discrete random variable has the probability distribution shown in the table.

Find the value of a.

x	0	1	2
$P(X = x)$	$\frac{1}{4} - a$	a	$\frac{1}{2} + a$

P **9** The random variable X can take any integer value from 1 to 50. Given that X has a discrete uniform distribution, find:

a $P(X = 1)$

b $P(X \geqslant 28)$

c $P(13 < X < 42)$

E **10** A discrete random variable X has the probability distribution shown in this table.

Find:

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$

a $P(1 < X \leqslant 3)$ **(1 mark)**

b $P(X < 2)$ **(1 mark)**

c $P(X > 3)$ **(1 mark)**

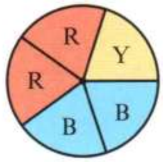
(E/P) **11** A biased coin is tossed until a head appears or it is tossed four times.

If $\text{P(Head)} = \frac{2}{3}$:

a Write down the probability distribution of S, the number of tosses, in table form. **(4 marks)**

b Find $\text{P}(S > 2)$. **(1 mark)**

(P) **12** A fair five-sided spinner is spun.



Given that the spinner is spun five times, write down, in table form, the probability distributions of the following random variables:

a X, the number of times red appears

b Y, the number of times yellow appears.

The spinner is now spun until it lands on blue, or until it has been spun five times. The random variable Z is defined as the number of spins in this experiment.

c Find the probability distribution of Z.

(E/P) **13** Marie says that a random variable X has a probability distribution defined by the following probability mass function:

$$\text{P}(X = x) = \frac{2}{x^2}, \quad x = 2, 3, 4$$

a Explain how you know that Marie's function does not describe a probability distribution. **(2 marks)**

b Given that the correct probability mass function is in the form $\text{P}(X = x) = \frac{k}{x^2}, \quad x = 2, 3, 4$ where k is a constant, find the exact value of k. **(2 marks)**

Challenge

The independent random variables X and Y have probability distributions

$$\text{P}(X = x) = \frac{1}{8}, \quad x = 1, 2, 3, 4, 5, 6, 7, 8 \qquad \text{P}(Y = y) = \frac{1}{y}, \quad y = 2, 3, 6$$

Find $\text{P}(X > Y)$.

Hint X and Y are independent so the value taken by one does not affect the probabilities for the other.

6.2 The binomial distribution

When you are carrying out a number of trials in an experiment or survey, you can define a random variable X to represent the number of **successful trials**.

- **You can model X with a binomial distribution, B(n, p), if:**
  - **there are a fixed number of trials, n**
  - **there are two possible outcomes (success and failure)**
  - **there is a fixed probability of success, p**
  - **the trials are independent of each other**

Hint If P(Success) = p and there are only two outcomes then P(Failure) = $1 - p$.

- **If a random variable X has the binomial distribution $B(n, p)$ then its probability mass function is given by**

$$P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r}$$

n is sometimes called the **index** and p is sometimes called the **parameter**.

You can use your calculator to work out binomial probabilities. You can either use the rule given above, together with the ${}^nC_r$ function, or use the binomial probability distribution function directly.

Notation You write $X \sim B(n, p)$

Links $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

It is sometimes written as nCr or ${}^nC_r$. It represents the number of ways of selecting r successful outcomes from n trials.

← Pure Year 1, Chapter 8

Example 4

The random variable $X \sim B\left(12, \frac{1}{6}\right)$. Find:

a $P(X = 2)$ **b** $P(X = 9)$ **c** $P(X \leqslant 1)$

a $P(X = 2) = \binom{12}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{10} = \frac{12!}{2!10!}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{10}$

$= 0.296\,09\ldots$

$= 0.296$ (3 s.f.)

Use the formula with $n = 12$, $p = \frac{1}{6}$ and $r = 2$.

b $P(X = 9) = \binom{12}{9}\left(\frac{1}{6}\right)^9\left(\frac{5}{6}\right)^3$

$= 0.000\,012\,63\ldots$

$= 0.000\,012\,6$ (3 s.f.)

Use the formula with $n = 12$, $p = \frac{1}{6}$ and $r = 9$.

c $P(X \leqslant 1) = P(X = 0) + P(X = 1)$

$= \left(\frac{5}{6}\right)^{12} + \binom{12}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^{11}$

$= 0.112\,156\ldots + 0.269\,17\ldots$

$= 0.381\,33\ldots$

$= 0.381$ (3 s.f.)

A binomial distribution can take any value from 0 up to n inclusive. So there are two possible outcomes that satisfy the inequality: $X = 0$ and $X = 1$.

Online Use the ${}^nC_r$ function on your calculator to work out binomial probabilities.

Example 5

The probability that a randomly chosen member of a reading group is left-handed is 0.15. A random sample of 20 members of the group is taken.

a Suggest a suitable model for the random variable X, the number of members in the sample who are left-handed. Justify your choice.

b Use your model to calculate the probability that:

i exactly 7 of the members in the sample are left-handed

ii fewer than two of the members in the sample are left-handed.

a The random variable can take two values, left-handed or right-handed.
There are a fixed number of trials, 20, and a fixed probability of success: 0.15.
Assuming each member in the sample is independent, a suitable model is
$X \sim B(20, 0.15)$

b i $P(X = 7) = \binom{20}{7} \times (0.15)^7(0.85)^{13}$
$= 0.016\,01\ldots$
$= 0.0160$ (3 s.f.)

ii $P(X < 2) = P(X = 0) + P(X = 1)$
$= 0.038\,75\ldots + 0.136\,79\ldots$
$= 0.176$ (3 s.f.)

A binomial model is a suitable choice. State the assumptions that are necessary for the binomial model, and make sure that you specify the values of n and p.

Online Work this out directly using the binomial probability distribution function on your calculator and entering $x = 7$, $n = 20$ and $p = 0.15$.

In this situation, 'fewer than two' means 0 or 1.

Exercise 6B

1 The random variable $X \sim B\left(8, \frac{1}{3}\right)$. Find:

a $P(X = 2)$ **b** $P(X = 5)$ **c** $P(X \leqslant 1)$

2 The random variable $T \sim B\left(15, \frac{2}{3}\right)$. Find:

a $P(T = 5)$ **b** $P(T = 10)$ **c** $P(3 \leqslant T \leqslant 4)$

3 A student suggests using a binomial distribution to model the following situations.
Give a description of the random variable, state any assumptions that must be made and give possible values for n and p.

a A sample of 20 bolts from a large batch is checked for defects. The production process should produce 1% of defective bolts.

b Some traffic lights have three phases: stop 48% of the time, wait or get ready 4% of the time, and go 48% of the time. Assuming that you only cross a traffic light when it is in the go position, model the number of times that you have to wait or stop on a journey passing through 6 sets of traffic lights.

c When Stephanie plays tennis with Timothy, on average one in eight of her serves is an 'ace'. How many 'aces' does Stephanie serve in the next 30 serves against Timothy?

4 State which of the following can be modelled with a binomial distribution and which cannot. Give reasons for your answers.

a Given that 15% of people have blood that is Rhesus negative (Rh^-), model the number of pupils in a statistics class of 14 who are Rh^-.

b You are given a fair coin and told to keep tossing it until you obtain 4 heads in succession. Model the number of tosses you need.

c A certain car manufacturer produces 12% of new cars in the colour red, 8% in blue, 15% in white and the rest in other colours. You make a note of the colour of the first 15 new cars of this make. Model the number of red cars you observe.

(P) **5** A balloon manufacturer claims that 95% of his balloons will not burst when blown up. If you have 20 of these balloons to blow up for a birthday party:

a What is the probability that none of them burst when blown up?

b Find the probability that exactly 2 balloons burst.

(E/P) **6** The probability of a switch being faulty is 0.08. A random sample of 10 switches is taken from the production line.

a Define a suitable distribution to model the number of faulty switches in this sample, and justify your choice. **(2 marks)**

b Find the probability that the sample contains 4 faulty switches. **(2 marks)**

(E/P) **7** A particular genetic marker is present in 4% of the population.

a State any assumptions that are required to model the number of people with this genetic marker in a sample of size n as a binomial distribution. **(2 marks)**

b Using this model, find the probability of exactly 6 people having this marker in a sample of size 50. **(2 marks)**

(E/P) **8** A dice is biased so that the probability of it landing on a six is 0.3. Hannah rolls the dice 15 times.

a State any assumptions that are required to model the number of sixes as a binomial distribution. State the distribution. **(2 marks)**

b Find the probability that Hannah rolls exactly 4 sixes. **(2 marks)**

c Find the probability that she rolls two or fewer sixes. **(3 marks)**

6.3 Cumulative probabilities

A **cumulative probability function** for a random variable X tells you the sum of all the individual probabilities up to and including the given value x in the calculation for $P(X \leqslant x)$.

For the binomial distribution $X \sim B(n, p)$ there are tables in the formula book giving $P(X \leqslant x)$ for various values of n and p.

An extract from the tables is shown below:

$p =$	0.05	0.10	0.15	0.20	0.25	0.30
$n = 5, x = 0$	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681
1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282
2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369
3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692
4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976
$n = 6, x = 0$	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176

For the binomial distribution $X \sim B(5, 0.3)$, this tells you that $P(X \leqslant 2) = 0.8369$ to 4 d.p.

You can also use the binomial cumulative probability function on your calculator to find $P(X \leqslant x)$ for any values of x, n and p.

Example 6

The random variable $X \sim B(20, 0.4)$. Find:

a $P(X \leqslant 7)$ **b** $P(X < 6)$ **c** $P(X \geqslant 15)$

a $P(X \leqslant 7) = 0.4159$

Use $n = 20$, $p = 0.4$ and $x = 7$. You can use tables or your calculator.

Online Use the binomial **cumulative distribution** function on your calculator. You want to find $P(X \leqslant 7)$, not $P(X = 7)$. On some calculators, this is labelled 'Binomial CD'.

b $P(X < 6) = P(X \leqslant 5)$
$= 0.1256$

X can only take whole number values, so $P(X < 6) = P(X \leqslant 5)$.

c $P(X \geqslant 15) = 1 - P(X \leqslant 14)$
$= 1 - 0.9984$
$= 0.0016$

When questions are set in context there are different phrases that can be used to ask for probabilities. The correct interpretation of these phrases is critical, especially when dealing with cumulative probabilities. The table below gives some examples.

Phrase	Meaning	Calculation
… greater than 5 …	$X > 5$	$1 - P(X \leqslant 5)$
… no more than 3 …	$X \leqslant 3$	$P(X \leqslant 3)$
… at least 7 …	$X \geqslant 7$	$1 - P(X \leqslant 6)$
… fewer than 10 …	$X < 10$	$P(X \leqslant 9)$
… at most 8 …	$X \leqslant 8$	$P(X \leqslant 8)$

Example 7

A spinner is designed so that the probability it lands on red is 0.3. Jane has 12 spins. Find the probability that Jane obtains:

a no more than 2 reds

b at least 5 reds.

Jane decides to use this spinner for a class competition, where students win a prize if they reach a certain number of reds.

c Find how many reds are needed to win a prize.

Let X = the number of reds in 12 spins.

$X \sim B(12, 0.3)$

a $P(X \leqslant 2) = 0.2528$

'no more than 2' means $X \leqslant 2$.

b $P(X \geqslant 5) = 1 - P(X \leqslant 4)$
$= 1 - 0.7237$
$= 0.2763$

'at least 5' means $X \geqslant 5$.

c Let r = the smallest number of reds needed to win a prize.

Require: $P(X \geqslant r) < 0.05$

Form a probability statement to represent the condition for winning a prize.

Problem-solving

When you are looking for the first probability that is greater or less than a given threshold, it is sometimes quicker to look on the tables rather than use your calculator.

From tables:

$P(X \leqslant 5) = 0.8822$

$P(X \leqslant 6) = 0.9614$

$P(X \leqslant 7) = 0.9905$

So: $P(X \leqslant 6) = 0.9614$ implies that

$P(X \geqslant 7) = 1 - 0.9614$
$= 0.0386 < 0.05$

Since $x = 6$ gives the first value > 0.95, use this probability and find $r = 7$.

So 7 or more reds will win a prize.

Always make sure that your final answer is related back to the context of the original question.

Online Explore the cumulative probabilities for the binomial distribution for this example using technology.

Exercise 6C

1 The random variable $X \sim B(9, 0.2)$. Find:

a $P(X \leqslant 4)$ **b** $P(X < 3)$ **c** $P(X \geqslant 2)$ **d** $P(X = 1)$

2 The random variable $X \sim B(20, 0.35)$. Find:

a $P(X \leqslant 10)$ **b** $P(X > 6)$

c $P(X = 5)$ **d** $P(2 \leqslant X \leqslant 7)$

Hint $P(2 \leqslant X \leqslant 7) = P(X \leqslant 7) - P(X \leqslant 1)$

3 The random variable $X \sim B(40, 0.47)$. Find:

a $P(X < 20)$ **b** $P(X > 16)$

c $P(11 \leqslant X \leqslant 15)$ **d** $P(10 < X < 17)$

Watch out For questions **3** and **4** the values of n and p aren't given in the tables. Use the binomial cumulative distribution function on your calculator.

4 The random variable $X \sim B(37, 0.65)$. Find:

a $P(X > 20)$ **b** $P(X \leqslant 26)$

c $P(15 \leqslant X < 20)$ **d** $P(X = 23)$

(P) **5** Eight fair coins are tossed and the total number of heads showing is recorded. Find the probability of:

a no heads **b** at least 2 heads **c** more heads than tails.

(P) **6** For a particular type of plant 25% have blue flowers. A garden centre sells these plants in trays of 15 plants of mixed colours. A tray is selected at random.
Find the probability that the number of plants with blue flowers in this tray is:

a exactly 4

b at most 3

c between 3 and 6 (inclusive).

(E/P) **7** The random variable $X \sim B(50, 0.40)$. Find:

a the largest value of k such that $P(X \leqslant k) < 0.05$ **(1 mark)**

b the smallest number r such that $P(X > r) < 0.01$. **(2 marks)**

(E/P) **8** The random variable $X \sim B(40, 0.10)$. Find:

a the largest value of k such that $P(X < k) < 0.02$ **(1 mark)**

b the smallest number r such that $P(X > r) < 0.01$ **(2 marks)**

c $P(k \leqslant X \leqslant r)$. **(2 marks)**

(E/P) **9** In a town, 30% of residents listen to the local radio.
Ten residents are chosen at random.
X = the number of these ten residents that listen to the local radio.

a Suggest a suitable distribution for X and comment on any necessary assumptions. **(2 marks)**

b Find the probability that at least half of these 10 residents listen to local radio. **(2 marks)**

c Find the smallest value of s so that $P(X \geqslant s) < 0.01$. **(2 marks)**

(E/P) **10** A factory produces a component for the motor trade and 5% of the components are defective. A quality control officer regularly inspects a random sample of 50 components. Find the probability that the next sample contains:

a fewer than 2 defectives **(1 mark)**

b more than 5 defectives. **(2 marks)**

The officer will stop production if the number of defectives in the sample is greater than a certain value d. Given that the officer stops production less than 5% of the time:

c find the smallest value of d. **(2 marks)**

Mixed exercise 6

(E) **1** The random variable X has probability function

$$P(X = x) = \frac{x}{21} \qquad x = 1, 2, 3, 4, 5, 6.$$

a Construct a table giving the probability distribution of X.

b Find $P(2 < X \leqslant 5)$.

E **2** The discrete random variable X has the probability distribution shown.

x	−2	−1	0	1	2	3
P($X = x$)	0.1	0.2	0.3	r	0.1	0.1

a Find r. **(1 mark)**

b Calculate $P(-1 \leqslant x < 2)$. **(2 marks)**

E/P **3** The random variable X has probability function

$$P(X = x) = \frac{(3x - 1)}{26} \qquad x = 1, 2, 3, 4.$$

a Construct a table giving the probability distribution of X. **(2 marks)**

b Find $P(2 < X \leqslant 4)$. **(2 marks)**

E **4** Sixteen counters are numbered 1 to 16 and placed in a bag. One counter is chosen at random and the number, X, recorded.

a Write down one condition on selecting a counter if X is to be modelled as a discrete uniform distribution. **(1 mark)**

b Find:

i $P(X = 5)$ **(1 mark)**

ii $P(X \text{ is prime})$ **(2 marks)**

iii $P(3 \leqslant X < 11)$ **(2 marks)**

E/P **5** The random variable Y has probability function

$$P(Y = y) = \frac{y}{k} \qquad y = 1, 2, 3, 4, 5$$

a Find the value of k. **(2 marks)**

b Construct a table giving the probability distribution of Y. **(2 marks)**

c Find $P(Y > 3)$. **(1 mark)**

E/P **6** Stuart rolls a biased dice four times. $P(\text{six}) = \frac{1}{4}$. The random variable T represents the number of times he rolls a six.

a Construct a table giving the probability distribution of T. **(3 marks)**

b Find $P(T < 3)$. **(2 marks)**

He rolls the dice again, this time recording the number of rolls required to roll a six.
He rolls the dice a maximum of five times. Let the random variable S stand for the number of times he rolls the dice.

c Construct a table giving the probability distribution of S. **(3 marks)**

d Find $P(S > 2)$. **(2 marks)**

E **7** The discrete random variable $X \sim B(30, 0.73)$. Find:

a $P(X = 20)$ **(1 mark)**

b $P(X \leqslant 13)$ **(1 mark)**

c $P(11 < X \leqslant 25)$ **(2 marks)**

(P) **8** A coin is biased so that the probability of a head is $\frac{2}{3}$. The coin is tossed repeatedly. Find the probability that:

a the first tail will occur on the sixth toss

b in the first 8 tosses there will be exactly 2 tails.

(P) **9** Records kept in a hospital show that 3 out of every 10 patients who visit the accident and emergency department have to wait more than half an hour. Find, to 3 decimal places, the probability that of the first 12 patients who come to the accident and emergency department:

a none

b more than 2

will have to wait more than half an hour.

(E/P) **10** **a** State clearly the conditions under which it is appropriate to assume that a random variable has a binomial distribution. **(2 marks)**

A door-to-door canvasser tries to persuade people to have a certain type of double glazing installed. The probability that his canvassing at a house is successful is 0.05.

b Find the probability that he will have at least 2 successes out of the first 10 houses he canvasses. **(2 marks)**

c Calculate the smallest number of houses he must canvass so that the probability of his getting at least one success exceeds 0.99. **(4 marks)**

(P) **11** A completely unprepared student is given a true/false-type test with 10 questions. Assuming that the student answers all the questions at random:

a find the probability that the student gets all the answers correct.

It is decided that a pass will be awarded for 8 or more correct answers.

b Find the probability that the student passes the test.

(P) **12** A six-sided die is biased. When the die is thrown the number 5 is twice as likely to appear as any other number. All the other faces are equally likely to appear. The die is thrown repeatedly. Find the probability that:

a the first 5 will occur on the sixth throw

b in the first eight throws there will be exactly three 5s.

(E/P) **13** A manufacturer produces large quantities of plastic chairs. It is known from previous records that 15% of these chairs are green. A random sample of 10 chairs is taken.

a Define a suitable distribution to model the number of green chairs in this sample. **(1 mark)**

b Find the probability of at least 5 green chairs in this sample. **(3 marks)**

c Find the probability of exactly 2 green chairs in this sample. **(3 marks)**

(E/P) **14** A bag contains a large number of beads of which 45% are yellow. A random sample of 20 beads is taken from the bag. Use the binomial distribution to find the probability that the sample contains:

a fewer than 12 yellow beads **(2 marks)**

b exactly 12 yellow beads. **(3 marks)**

E/P **15** An archer hits the bullseye with probability 0.6. She shoots 20 arrows at a time.

a Find the probability that she hits the bullseye with at least 50% of her arrows. **(3 marks)**

She shoots 12 sets of 20 arrows.

b Find the probability that she hits the bullseye with at least 50% of her arrows in 7 of the 12 sets of arrows. **(2 marks)**

c Find the probability that she hits the bullseye with at least 50% of her arrows in fewer than 6 of the 12 sets of arrows. **(2 marks)**

Challenge

A driving theory test has 50 questions. Each question has four answers, of which only one is correct.

Annabelle is certain she got 32 answers correct, but she guessed the remaining answers. She needs to get 43 correct answers to pass the test.

Find the probability that Annabelle passed the test.

Summary of key points

1 A **probability distribution** fully describes the probability of any outcome in the sample space.

2 The sum of the probabilities of all outcomes of an event add up to 1. For a random variable X, you can write $\sum P(X = x) = 1$ for all x.

3 You can model X with a **binomial distribution**, **B(n, p)**, if:

- there are a fixed number of trials, n
- there are two possible outcomes (success or failure)
- there is a fixed probabillity of success, p
- the trials are independent of each other.

4 If a random variable X has the binomial distribution $B(n, p)$ then its probability mass function is given by

$$P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r}$$