Chapter 15

# 15 Bernoulli sequences and the binomial distribution

## Objectives

- To define a **Bernoulli sequence** and a **Bernoulli random variable**.
- To define the **binomial probability distribution**.
- To investigate the shape of the graph of the binomial probability distribution for different values of the parameters.
- To calculate and interpret the **mean**, **variance** and **standard deviation** for the binomial probability distribution.
- To use the binomial probability distribution to solve problems.

The binomial distribution is important because it has very wide application. It is concerned with situations where there are two possible outcomes, and many 'real life' scenarios of interest fall into this category. For example:

- A political poll of voters is carried out. Each polled voter is asked whether or not they would vote for the present government.
- A poll of Year 12 students in Australia is carried out. Each student is asked whether or not they watch the ABC on a regular basis.
- The effectiveness of a medical procedure is tested by selecting a group of patients and recording whether or not it is successful for each patient in the group.
- Components for an electronic device are tested to see if they are defective or not.

The binomial distribution has application in each of these examples.

We will use the binomial distribution again in Chapter 18, where we further develop our understanding of sampling.

This chapter covers Unit 4 Topic 3: Discrete random variables 2.

# 15A Introduction to Bernoulli sequences and the binomial distribution

## ▶ Bernoulli sequences

An experiment often consists of repeated trials, each of which may be considered as having only two possible outcomes. For example, when a coin is tossed, the two possible outcomes are ‘head’ and ‘tail’. When a die is rolled, the two possible outcomes are determined by the random variable of interest for the experiment. If the event of interest is a ‘six’, then the two outcomes are ‘six’ and ‘not a six’.

A **Bernoulli sequence** is the name used to describe a sequence of repeated trials with the following properties:

- Each trial results in one of two outcomes, which are usually designated as either a success, $S$, or a failure, $F$.
- The probability of success on a single trial, $p$, is constant for all trials (and thus the probability of failure on a single trial is $1 - p$).
- The trials are independent (so that the outcome of any trial is not affected by the outcome of any previous trial).

### Example 1

Suppose that a netball player has a probability of $\frac{1}{3}$ of scoring a goal each time she attempts to goal. She repeatedly has shots for goal. Is this a Bernoulli sequence?

**Solution**

In this example:

- Each trial results in one of two outcomes, goal or miss.
- The probability of scoring a goal $(\frac{1}{3})$ is constant for all attempts, as is the probability of a miss $(\frac{2}{3})$.
- The trials are independent (so that the outcome of any trial is not affected by the outcome of any previous trial).

Thus, the player’s shots at goal can be considered a Bernoulli sequence.

## ▶ Bernoulli random variables

The outcome from a Bernoulli trial is represented by a **Bernoulli random variable**, which is a random variable that takes only the values 1 (indicating a success) and 0 (indicating a failure). Thus a Bernoulli random variable $Y$ has a probability distribution of the form:

| $y$ | 0 | 1 |
|---|---|---|
| $\Pr(Y = y)$ | $1 - p$ | $p$ |

In Example 1, each shot at goal can be modelled by a Bernoulli random variable with $p = \frac{1}{3}$, where 1 represents a goal and 0 represents a miss.

## The binomial probability distribution

We can think of a Bernoulli random variable as counting the number of successes in a single Bernoulli trial. What happens if we count the number of successes in a Bernoulli sequence?

The number of successes in a Bernoulli sequence of $n$ trials is called a **binomial random variable** and is said to have a **binomial probability distribution**.

For example, consider rolling a fair six-sided die three times. Let the random variable $X$ be the number of 3s observed. Let T represent a 3, and let N represent not a 3. Each roll meets the conditions of a Bernoulli trial. Thus $X$ is a binomial random variable.

Now consider all the possible outcomes from the three rolls and their probabilities.

| Outcome | Number of 3s | Probability | |
|---|---|---|---|
| TTT | $X = 3$ | $\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6}$ | $\Pr(X = 3) = (\frac{1}{6})^3$ |
| TTN | $X = 2$ | $\frac{1}{6} \times \frac{1}{6} \times \frac{5}{6}$ | |
| TNT | $X = 2$ | $\frac{1}{6} \times \frac{5}{6} \times \frac{1}{6}$ | $\Pr(X = 2) = 3 \times (\frac{1}{6})^2 \times \frac{5}{6}$ |
| NTT | $X = 2$ | $\frac{5}{6} \times \frac{1}{6} \times \frac{1}{6}$ | |
| TNN | $X = 1$ | $\frac{1}{6} \times \frac{5}{6} \times \frac{5}{6}$ | |
| NTN | $X = 1$ | $\frac{5}{6} \times \frac{1}{6} \times \frac{5}{6}$ | $\Pr(X = 1) = 3 \times \frac{1}{6} \times (\frac{5}{6})^2$ |
| NNT | $X = 1$ | $\frac{5}{6} \times \frac{5}{6} \times \frac{1}{6}$ | |
| NNN | $X = 0$ | $\frac{5}{6} \times \frac{5}{6} \times \frac{5}{6}$ | $\Pr(X = 0) = (\frac{5}{6})^3$ |

Thus the probability distribution of $X$ is given by the following table.

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $\Pr(X = x)$ | $\frac{125}{216}$ | $\frac{75}{216}$ | $\frac{15}{216}$ | $\frac{1}{216}$ |

Instead of listing all the outcomes to find the probability distribution, we can use our knowledge of selections from Mathematical Methods Units 1 & 2 (revised in Appendix A).

Consider the probability that $X = 1$, that is, when exactly one 3 is observed. We can see from the table that there are three ways this can occur. Since the 3 could occur on the first, second or third roll of the die, we can consider this as selecting one object from a group of three, which can be done in $\binom{3}{1}$ ways.

Consider the probability that $X = 2$, that is, when exactly two 3s are observed. Again from the table there are three ways this can occur. Since the two 3s could occur on any two of the three rolls of the die, we can consider this as selecting two objects from a group of three, which can be done in $\binom{3}{2}$ ways.

This leads us to a general formula for this probability distribution:

$$\Pr(X = x) = \binom{3}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{3-x} \qquad x = 0, 1, 2, 3$$

This is an example of the binomial distribution.

If the random variable $X$ is the number of successes in $n$ independent trials, each with probability of success $p$, then $X$ has a **binomial distribution** and the rule is

$$\Pr(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \qquad x = 0, 1, \ldots, n$$

where $\binom{n}{x} = \dfrac{n!}{x!\,(n - x)!}$

## Example 2

Find the probability of obtaining exactly three heads when a fair coin is tossed seven times, correct to four decimal places.

### Solution

Obtaining a head is considered a success here, and the probability of success on each of the seven independent trials is 0.5.

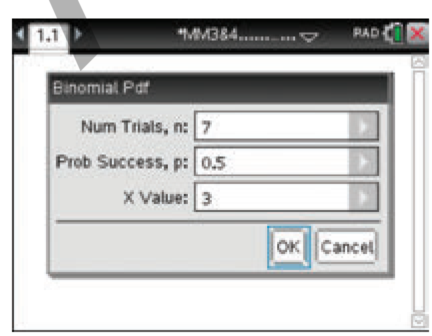
Let $X$ be the number of heads obtained. In this case, the parameters are $n = 7$ and $p = 0.5$.

$$\Pr(X = 3) = \binom{7}{3} (0.5)^3 (1 - 0.5)^{7-3}$$

$$= 35 \times (0.5)^7 = 0.2734$$

### Using the TI-Nspire CX non-CAS

Use (menu) > **Probability** > **Distributions** > **Binomial Pdf** and complete as shown.



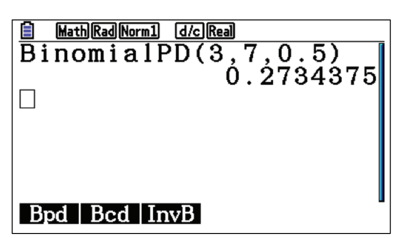

Use (tab) or ▼ to move between cells.

The result is shown below.


Note: You can also type in the command and the parameter values directly if preferred.

### Using the Casio

- In **Run-Matrix** mode, go to the **Statistics** menu (OPTN) (F5).
- For the binomial probability distribution, select **Distributions** (F3), **Binomial** (F5), **Bpd** (F1).
- Complete by entering: 3, 7, 0.5)
- Press (EXE).




Note: The syntax for the binomial probability distribution is:

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

## Example 3

The probability that a person currently in prison has ever been imprisoned before is 0.72. Find the probability that of five prisoners chosen at random at least three have been imprisoned before, correct to four decimal places.

### Solution

If $X$ is the number of prisoners who have been imprisoned before, then

$$\Pr(X = x) = \binom{5}{x}(0.72)^x(0.28)^{5-x} \qquad x = 0, 1, \dots, 5$$

and so

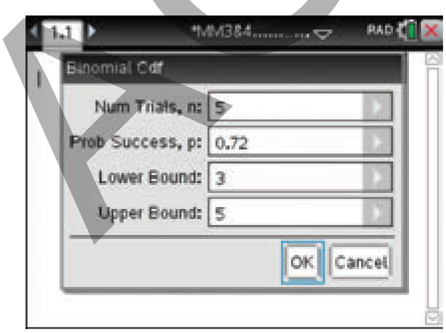
$$\begin{aligned}\Pr(X \geq 3) &= \Pr(X = 3) + \Pr(X = 4) + \Pr(X = 5)\\ &= \binom{5}{3}(0.72)^3(0.28)^2 + \binom{5}{4}(0.72)^4(0.28)^1 + \binom{5}{5}(0.72)^5(0.28)^0\\ &= 0.8624\end{aligned}$$

### Using the TI-Nspire CX non-CAS

Use menu > **Probability** > **Distributions** > **Binomial Cdf** and complete as shown.

Use tab or ▼ to move between cells.

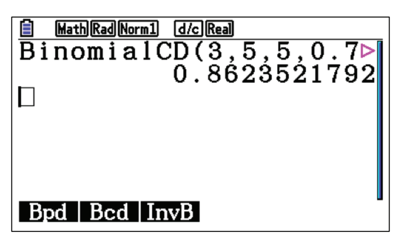
The result is shown below.


Note: You can also type in the command and the parameter values directly if preferred.

### Using the Casio

- In **Run-Matrix** mode, go to the **Statistics** menu OPTN F5.
- For the binomial cumulative distribution, select **Distributions** F3, **Binomial** F5, **Bcd** F2.
- Complete by entering: 3, 5, 5, 0.72)
- Press EXE.




Note: The syntax for the binomial cumulative distribution is:
BinomialCD(*lower bound*, *upper bound*, *number of trials*, *probability of success*)

## The binomial distribution and conditional probability

We can use the binomial distribution to solve problems involving conditional probabilities.

### Example 4

The probability of a netballer scoring a goal is 0.3. Find the probability that out of six attempts the netballer scores a goal:

**a** four times  **b** four times, given that she scores at least one goal.

#### Solution

Let $X$ be the number of goals scored.
Then $X$ has a binomial distribution with $n = 6$ and $p = 0.3$.

**a**
$$\Pr(X = 4) = \binom{6}{4}(0.3)^4(0.7)^2$$
$$= 15 \times 0.0081 \times 0.49$$
$$= 0.059535$$

**b**
$$\Pr(X = 4 \mid X \geq 1) = \frac{\Pr(X = 4 \cap X \geq 1)}{\Pr(X \geq 1)}$$
$$= \frac{\Pr(X = 4)}{\Pr(X \geq 1)}$$
$$= \frac{0.059535}{1 - 0.7^6} \qquad \text{since } \Pr(X \geq 1) = 1 - \Pr(X = 0)$$
$$= 0.0675$$

### Section summary

- A **Bernoulli sequence** is a sequence of trials with the following properties:
  - Each trial results in one of two outcomes, which are usually designated as either a success, $S$, or a failure, $F$.
  - The probability of success on a single trial, $p$, is constant for all trials (and thus the probability of failure on a single trial is $1 - p$).
  - The trials are independent (so that the outcome of any trial is not affected by the outcome of any previous trial).
- A **Bernoulli random variable** describes the outcome from a Bernoulli trial; it has a probability distribution of the form $\Pr(Y = 1) = p$ and $\Pr(Y = 0) = 1 - p$.
- The number of successes, $X$, in a Bernoulli sequence of $n$ trials is called a **binomial random variable** and has a **binomial probability distribution**:

$$\Pr(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \qquad x = 0, 1, \dots, n$$

where $\binom{n}{x} = \dfrac{n!}{x!\,(n - x)!}$

## Exercise 15A

Skillsheet

Example 1

SF

**1** Which of the following describes a Bernoulli sequence?

**A** tossing a fair coin many times

**B** drawing balls from an urn containing five red and three black balls, replacing the chosen ball each time

**C** selecting people at random from the population and noting their age

**D** selecting people at random from the population and noting their sex, male or female

Example 2

**2** Find the probability of obtaining exactly four heads when a fair coin is tossed seven times, correct to four decimal places.

**3** For a binomial distribution with $n = 4$ and $p = 0.2$, find the probability of:

**a** three successes **b** four successes.

**4** For a binomial distribution with $n = 5$ and $p = 0.4$, find the probability of:

**a** no successes **b** three successes **c** five successes.

**5** Suppose that a fair coin is tossed three times, and the number of heads observed.

**a** Write down a general rule for the probability distribution of the number of heads.

**b** Use the rule to calculate the probability of observing two heads.

CF

**6** Suppose that $X$ is the number of male children born into a family of six children. Assume that the distribution of $X$ is binomial, with probability of success 0.48.

**a** Write down a general rule for the probability distribution of the number of male children.

**b** Use the rule to calculate the probability that a family with six children will have exactly two male children.

Example 3

**7** A fair die is rolled six times and the number of 2s noted. Find the probability of:

**a** exactly three 2s **b** more than three 2s **c** at least three 2s.

**8** Jo knows that each ticket has a probability of 0.1 of winning a prize in a lucky ticket competition. Suppose she buys 10 tickets.

**a** Write down a general rule for the probability distribution of the number of winning tickets.

**b** Use the rule to calculate the probability that Jo has:

**i** no wins **ii** at least one win.

**9** Suppose that the probability that a person selected at random is left-handed is always 0.2. If 11 people are selected at random for the cricket team:

**a** Write down a general rule for the probability distribution of the number left-handed people on the team.

**b** Use the rule to calculate the probability of selecting:

**i** exactly two left-handers **ii** no left-handers **iii** at least one left-hander.

CF

**10** In a particular city, the probability of rain falling on any given day is $\frac{1}{5}$.

**a** Write down a general rule for the probability distribution of the number of days of rain in a week.

**b** Use the rule to calculate the probability that in a particular week rain will fall:

**i** every day **ii** not at all **iii** on two or three days.

**11** The probability of a particular drug causing side effects in a person is 0.2. What is the probability that at least two people in a random sample of 10 people will experience side effects?

**12** Records show that $x\%$ of people will pass their driver's licence on the first attempt. If six students attempt their driver's licence, write down in terms of $x$ the probability that:

**a** all six students pass **b** only one fails **c** no more than two fail.

**13** A supermarket has four checkouts. A customer in a hurry decides to leave without making a purchase if all the checkouts are busy. At that time of day the probability of each checkout being free is 0.25. Assuming that whether or not a checkout is busy is independent of any other checkout, calculate the probability that the customer will make a purchase.

**14** A fair die is rolled 50 times. Find the probability of observing:

**a** exactly 10 sixes **b** no more than 10 sixes **c** at least 10 sixes.

**15** Find the probability of getting at least nine successes in 100 trials for which the probability of success is $p = 0.1$.

**16** A fair coin is tossed 50 times. If $X$ is the number of heads observed, find:

**a** $\Pr(X = 25)$ **b** $\Pr(X \leq 25)$ **c** $\Pr(X \leq 10)$ **d** $\Pr(X \geq 40)$

**17** A survey of the population in a particular city found that 40% of people regularly participate in sport. What is the probability that fewer than half of a random sample of six people regularly participate in sport?

**18** An examination consists of six multiple-choice questions. Each question has four possible answers. At least three correct answers are required to pass the examination. Suppose that a student guesses the answer to each question.

**a** What is the probability the student guesses every question correctly?

**b** What is the probability the student will pass the examination?

Example 4 **19** The manager of a shop knows from experience that 60% of her customers will use a credit card to pay for their purchases. Find the probability that:

**a** the next three customers will use a credit card, and the three after that will not

**b** three of the next six customers will use a credit card

**c** at least three of the next six customers will use a credit card

**d** exactly three of the next six customers will use a credit card, given that at least three of the next six customers use a credit card.

CF

**20** A multiple-choice test has eight questions, each with five possible answers, only one of which is correct. Find the probability that a student who guesses the answer to every question will have:

- **a** no correct answers
- **b** six or more correct answers
- **c** every question correct, given they have six or more correct answers.

**21** The probability that a full forward in Australian Rules football will kick a goal from outside the 50-metre line is 0.15. If the full forward has 10 kicks at goal from outside the 50-metre line, find the probability that he will:

- **a** kick a goal every time
- **b** kick at least one goal
- **c** kick more than one goal, given that he kicks at least one goal.

# 15B The graph, expectation and variance of a binomial distribution

We looked at the properties of discrete probability distributions in Chapter 14. We now consider these properties for the binomial distribution.

## The graph of a binomial probability distribution

As discussed in Chapter 14, a probability distribution may be represented as a rule, a table or a graph. We now investigate the shape of the graph of a binomial probability distribution for different values of the parameters $n$ and $p$.

A method for plotting a binomial distribution with a graphics calculator can be found in the calculator appendices in the Interactive Textbook.

### Example 5

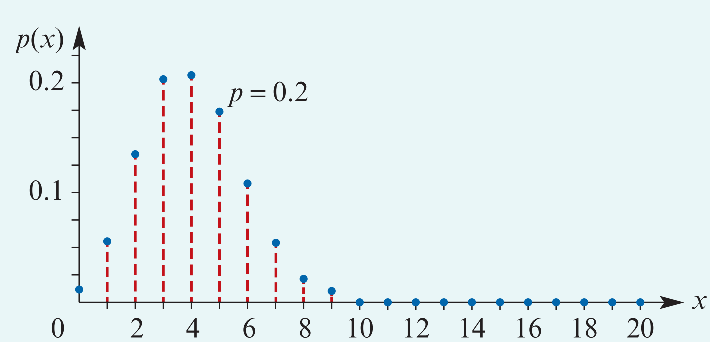
Construct and compare the graph of the binomial probability distribution for 20 trials ($n = 20$) with probability of success:

**a** $p = 0.2$ **b** $p = 0.5$ **c** $p = 0.8$

**Solution**

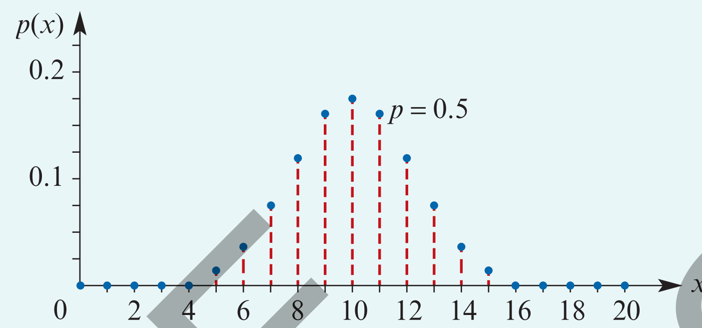
**a** For $p = 0.2$, the graph is positively skewed.

Mostly from 1 to 8 successes will be observed in 20 trials.



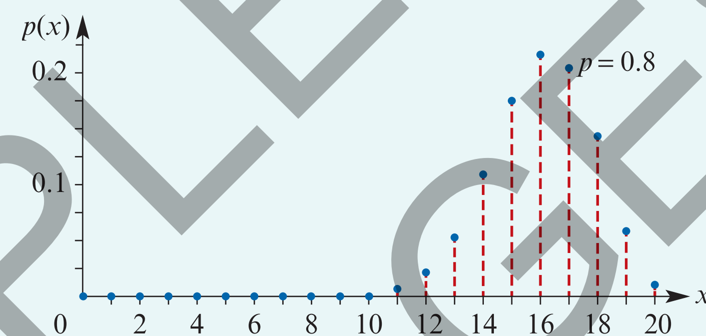
**b** For $p = 0.5$, the graph is symmetrical (as the probability of success is the same as the probability of failure).

Mostly from 6 to 14 successes will be observed in 20 trials.




**c** For $p = 0.8$, the graph is negatively skewed.

Mostly from 12 to 19 successes will be observed in 20 trials.




## ▶ Expectation and variance for a Bernoulli random variable

The table on the right shows the probability distribution for a Bernoulli random variable.

| $y$ | 0 | 1 |
|---|---|---|
| $\Pr(Y = y)$ | $1 - p$ | $p$ |

Thus $\quad \mathrm{E}(Y) = 0 \times (1 - p) + 1 \times p = p$

and $\quad \mathrm{E}(Y^2) = 0^2 \times (1 - p) + 1^2 \times p = p$

so $\quad \mathrm{Var}(Y) = p - p^2 = p(1 - p)$

Hence, if $Y$ is a Bernoulli random variable with probability of success $p$, then

$$\mathrm{E}(Y) = p$$

$$\mathrm{Var}(Y) = p(1 - p)$$

## ▶ Expectation and variance for a binomial random variable

How many heads would you expect to obtain, on average, if a fair coin was tossed 10 times?

While the exact number of heads in the 10 tosses would vary, and could theoretically take values from 0 to 10, it seems reasonable that the long-run average number of heads would be 5. It turns out that this is correct. That is, for a binomial random variable $X$ with $n = 10$ and $p = 0.5$,

$$\mathrm{E}(X) = \sum_x x \cdot \Pr(X = x) = 5$$

In general, the expected value of a binomial random variable is equal to the number of trials multiplied by the probability of success. The variance can also be calculated from the parameters $n$ and $p$.

If $X$ is the number of successes in $n$ trials, each with probability of success $p$, then the expected value and the variance of $X$ are given by

$$E(X) = np$$
$$Var(X) = np(1 - p)$$

Note: These formulas are consistent with those for a Bernoulli random variable, which is a special case of a binomial random variable where $n = 1$.

While it is not necessary in this course to be familiar with the derivations of these formulas, they are included for completeness in the final section of this chapter.

## Example 6

An examination consists of 30 multiple-choice questions, each question having three possible answers. A student guesses the answer to every question. Let $X$ be the number of correct answers.

**a** How many will she expect to get correct? That is, find $E(X) = \mu$.

**b** Find $Var(X)$.

### Solution

The number of correct answers, $X$, is a binomial random variable with parameters $n = 30$ and $p = \frac{1}{3}$.

**a** The student has an expected result of $\mu = np = 10$ correct answers. (This is not enough to pass if the pass mark is 50%.)

**b** $Var(X) = np(1 - p)$

$$= 30 \times \frac{1}{3} \times \frac{2}{3} = \frac{20}{3}$$

## Section summary

If $X$ is the number of successes in $n$ trials, each with probability of success $p$, then the expected value and the variance of $X$ are given by

- $E(X) = np$
- $Var(X) = np(1 - p)$

## Exercise 15B

Example 5 **1** Plot the graph of the probability distribution

$$\Pr(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \qquad x = 0, 1, \dots, n$$

for $n = 8$ and $p = 0.25$.

SF

SF

**2** Plot the graph of the probability distribution

$$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \dots, n$$

for $n = 12$ and $p = 0.35$.

**3** **a** Plot the graph of the probability distribution

$$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \dots, n$$

for $n = 10$ and $p = 0.2$.

**b** On the same axes, plot the graph of

$$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \dots, n$$

for $n = 10$ and $p = 0.8$, using a different plotting symbol.

**c** Compare the two distributions.

**d** Comment on the effect of the value of $p$ on the shape of the distribution.

Example 6

**4** Find the mean and variance of the binomial random variables with parameters:

**a** $n = 25,\ p = 0.2$ **b** $n = 10,\ p = 0.6$

**c** $n = 500,\ p = \frac{1}{3}$ **d** $n = 40,\ p = 20\%$

**5** A fair die is rolled six times.

**a** Find the expected value for the number of sixes obtained.

**b** Find the probability that more than the expected number of sixes is obtained.

CF

**6** The survival rate for a certain disease is 75%. Of the next 50 people who contract the disease, how many would you expect would survive?

**7** Records show that 60% of the students in a certain state attend government schools. If a group of 200 students are to be selected at random, find the mean and standard deviation of the number of students in the group who attend government schools.

**8** A binomial random variable $X$ has mean 12 and variance 9. Find the parameters $n$ and $p$, and hence find $\Pr(X = 7)$.

**9** A binomial random variable $X$ has mean 30 and variance 21. Find the parameters $n$ and $p$, and hence find $\Pr(X = 20)$.

# 15C Finding the sample size

While we can never be absolutely certain about the outcome of a random experiment, sometimes we are interested in knowing what size sample would be required to observe a certain outcome. For example, how many times do you need to roll a die to be reasonably sure of observing a six, or how many lotto tickets must you buy to be reasonably sure that you will win a prize?

## Example 7

The probability of winning a prize in a game of chance is 0.48.

**a** What is the least number of games that must be played to ensure that the probability of winning at least once is more than 0.95?

**b** What is the least number of games that must be played to ensure that the probability of winning at least twice is more than 0.95?

### Solution

Since the probability of winning each game is the same each time the game is played, this is an example of a binomial distribution, with the probability of success $p = 0.48$.

**a** The required answer is the smallest value of $n$ such that $\Pr(X \geq 1) > 0.95$.

$$\begin{aligned} & \Pr(X \geq 1) > 0.95 \\ \Leftrightarrow\quad & 1 - \Pr(X = 0) > 0.95 \\ \Leftrightarrow\quad & \Pr(X = 0) < 0.05 \\ \Leftrightarrow\quad & 0.52^n < 0.05 \quad \text{since } \Pr(X = 0) = 0.52^n \end{aligned}$$

This can be solved by taking logarithms of both sides:

$$\ln(0.52^n) < \ln(0.05)$$

$$n \ln(0.52) < \ln(0.05)$$

$$\therefore \quad n > \frac{\ln(0.05)}{\ln(0.52)} \approx 4.58$$

Thus the game must be played at least five times to ensure that the probability of winning at least once is more than 0.95.

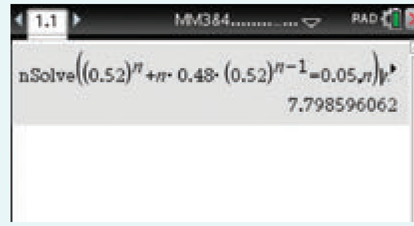
**b** The required answer is the smallest value of $n$ such that $\Pr(X \geq 2) > 0.95$, or equivalently, such that $\Pr(X < 2) < 0.05$. We have

$$\begin{aligned} \Pr(X < 2) &= \Pr(X = 0) + \Pr(X = 1) \\ &= \binom{n}{0} 0.48^0 0.52^n + \binom{n}{1} 0.48^1 0.52^{n-1} \\ &= 0.52^n + 0.48n(0.52)^{n-1} \end{aligned}$$

So the answer is the smallest value of $n$ such that

$$0.52^n + 0.48n(0.52)^{n-1} < 0.05$$

This equation cannot be solved algebraically; but a graphics calculator can be used to find the solution $n > 7.7985\ldots$.
Thus the game must be played at least eight times to ensure that the probability of winning at least twice is more than 0.95.

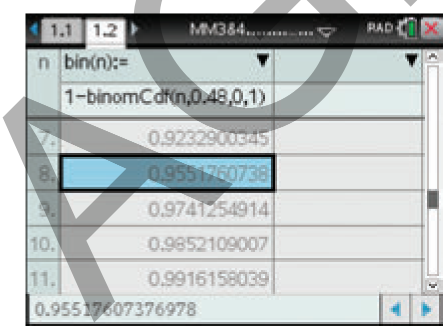



The following calculator inserts give a solution to part **b** of Example 7. Similar techniques can be used for part **a**.

## Using the TI-Nspire CX non-CAS

To find the smallest value of $n$ such that $\Pr(X \geq 2) > 0.95$, where $p = 0.48$:

- Define the binomial CDF as shown.
  The last two parameters are the lower and upper bounds (inclusive) of the $X$ value.
- Insert a **Lists & Spreadsheet** page. Press ctrl T to show the table of values.
- Scroll through the table to find where the probability is greater than 0.95. Hence $n = 8$.

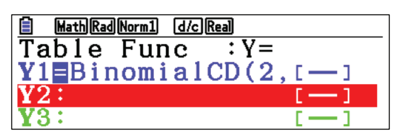




## Using the Casio

- Press MENU 7 to select **Table** mode.
- Select the binomial cumulative distribution:

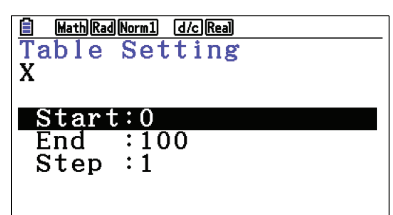
  OPTN F6 F3 F1 F5 F2

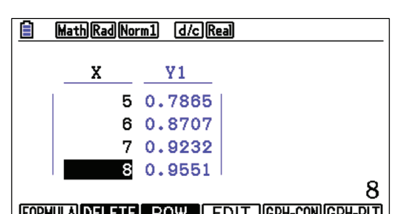
- Complete the rule as follows: $2, x, x, 0.48)$
- Press EXE.




Note that this rule gives the probability of winning between 2 and $x$ times, when $x$ games are played.

- Select **Set** F5 and adjust the Table Settings as shown. Press EXIT to return to the function list.
- Select **Table** F6 and scroll down until the probability exceeds 0.95.
- Hence, at least eight games must be played to ensure that the probability of winning at least twice is more than 0.95.



## Exercise 15C

Skillsheet
Example 7

1 The probability of a target shooter hitting the bullseye on any one shot is 0.2.

CF

a If the shooter takes five shots at the target, find the probability of:

i missing the bullseye every time

ii hitting the bullseye at least once.

b What is the smallest number of shots the shooter should make to ensure a probability of more than 0.95 of hitting the bullseye at least once?

c What is the smallest number of shots the shooter should make to ensure a probability of more than 0.95 of hitting the bullseye at least twice?

2 The probability of winning a prize with a lucky ticket on a wheel of fortune is 0.1.

a If a person buys 10 lucky tickets, find the probability of:

i winning twice

ii winning at least once.

b What is the smallest number of tickets that should be bought to ensure a probability of more than 0.7 of winning at least once?

3 Rex is shooting at a target. His probability of hitting the target is 0.6. What is the minimum number of shots needed for the probability of Rex hitting the target exactly five times to be more than 25%?

4 Janet is selecting chocolates at random out of a box. She knows that 20% of the chocolates have hard centres. What is the minimum number of chocolates she needs to select to ensure that the probability of choosing exactly three hard centres is more than 10%?

5 The probability of winning a prize in a game of chance is 0.35. What is the fewest number of games that must be played to ensure that the probability of winning at least twice is more than 0.9?

6 Geoff has determined that his probability of hitting '4' off any ball when playing cricket is 0.07. What is the fewest number of balls he must face to ensure that the probability of hitting more than one '4' is more than 0.8?

7 Monique is practising goaling for netball. She knows from past experience that her chance of making any one shot is about 70%. Her coach has asked her to keep practising until she scores 50 goals. How many shots would she need to attempt to ensure that the probability of scoring at least 50 goals is more than 0.99?

# 15D Proofs for the expectation and variance

In this section we give proofs of three important results on the binomial distribution.

The probabilities of a binomial distribution sum to 1.

**Proof** The binomial theorem, discussed in Appendix A, states that

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Now, using the binomial theorem, the sum of the probabilities for a binomial random variable $X$ with parameters $n$ and $p$ is given by

$$\sum_{x=0}^{n} \Pr(X = x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$
$$= ((1-p)+p)^n = (1)^n = 1$$

**Expected value**

If $X$ is a binomial random variable with parameters $n$ and $p$, then $\mathrm{E}(X) = np$.

**Proof** By the definition of expected value:

$$\mathrm{E}(X) = \sum_{x=0}^{n} x \cdot \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{by the distribution formula}$$
$$= \sum_{x=0}^{n} x \cdot \left(\frac{n!}{x!\,(n-x)!}\right) p^x (1-p)^{n-x} \qquad \text{expanding } \tbinom{n}{x}$$
$$= \sum_{x=1}^{n} x \cdot \left(\frac{n}{x!\,(n-x)!}\right) p^x (1-p)^{n-x} \qquad \text{since the } x = 0 \text{ term is zero}$$
$$= \sum_{x=1}^{n} x \cdot \left(\frac{n!}{x(x-1)!\,(n-x)!}\right) p^x (1-p)^{n-x} \qquad \text{since } x! = x(x-1)!$$
$$= \sum_{x=1}^{n} \left(\frac{n!}{(x-1)!\,(n-x)!}\right) p^x (1-p)^{n-x} \qquad \text{cancelling the } x\text{s}$$

This expression is very similar to the probability function for a binomial random variable, and we know the probabilities sum to 1. Taking out factors of $n$ and $p$ from the expression and letting $z = x - 1$ gives

$$\mathrm{E}(X) = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} (1-p)^{n-x}$$
$$= np \sum_{z=0}^{n-1} \binom{n-1}{z} p^z (1-p)^{n-1-z}$$

Note that this sum corresponds to the sum of all the values of the probability function for a binomial random variable $Z$, which is the number of successes in $n - 1$ trials each with probability of success $p$. Therefore the sum equals 1, and so

$$\mathrm{E}(X) = np$$

**Variance**

If $X$ is a binomial random variable with parameters $n$ and $p$, then $\text{Var}(X) = np(1 - p)$.

**Proof** The variance of the binomial random variable $X$ may be found using

$$\text{Var}(X) = \text{E}(X^2) - \mu^2, \qquad \text{where } \mu = np$$

Thus, to find the variance, we need to determine $\text{E}(X^2)$:

$$\begin{aligned}\text{E}(X^2) &= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} x^2 \left(\frac{n!}{x!\,(n-x)!}\right) p^x (1-p)^{n-x}\end{aligned}$$

But $x^2$ is not a factor of $x!$ and so we cannot proceed as in the previous proof for expected value.

The strategy used here is to determine $\text{E}[X(X-1)]$:

$$\begin{aligned}\text{E}[X(X-1)] &= \sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} x(x-1) \left(\frac{n!}{x!\,(n-x)!}\right) p^x (1-p)^{n-x} \\ &= \sum_{x=2}^{n} x(x-1) \left(\frac{n!}{x!\,(n-x)!}\right) p^x (1-p)^{n-x}\end{aligned}$$

since the first and second terms of the sum equal zero (when $x = 0$ and $x = 1$).

Taking out a factor of $n(n-1)p^2$ and letting $z = x - 2$ gives

$$\begin{aligned}\text{E}[X(X-1)] &= n(n-1)p^2 \sum_{x=2}^{n} \left(\frac{(n-2)!}{(x-2)!\,(n-x)!}\right) p^{x-2} (1-p)^{n-x} \\ &= n(n-1)p^2 \sum_{z=0}^{n-2} \binom{n-2}{z} p^z (1-p)^{n-2-z}\end{aligned}$$

Now the sum corresponds to the sum of all the values of the probability function for a binomial random variable $Z$, which is the number of successes in $n - 2$ trials each with probability of success $p$, and is thus equal to 1. Hence

$$\text{E}[X(X-1)] = n(n-1)p^2$$

$$\therefore \quad \text{E}(X^2) - \text{E}(X) = n(n-1)p^2$$

$$\therefore \quad \begin{aligned}\text{E}(X^2) &= n(n-1)p^2 + \text{E}(X) \\ &= n(n-1)p^2 + np\end{aligned}$$

This is an expression for $\text{E}(X^2)$ in terms of $n$ and $p$, as required. Thus

$$\begin{aligned}\text{Var}(X) &= \text{E}(X^2) - \mu^2 \\ &= n(n-1)p^2 + np - (np)^2 \\ &= np(1-p)\end{aligned}$$

## Chapter summary

AS Nrich

- A **Bernoulli sequence** is a sequence of trials with the following properties:
  - Each trial results in one of two outcomes, which are usually designated as either a success, $S$, or a failure, $F$.
  - The probability of success on a single trial, $p$, is constant for all trials (and thus the probability of failure on a single trial is $1 - p$).
  - The trials are independent (so that the outcome of any trial is not affected by the outcome of any previous trial).
- A **Bernoulli random variable** describes the outcome from a Bernoulli trial; it has a probability distribution of the form $\Pr(Y = 1) = p$ and $\Pr(Y = 0) = 1 - p$.
- If $X$ is the number of successes in $n$ Bernoulli trials, each with probability of success $p$, then $X$ is called a **binomial random variable** and is said to have a **binomial probability distribution** with parameters $n$ and $p$. The probability of observing $x$ successes in the $n$ trials is given by

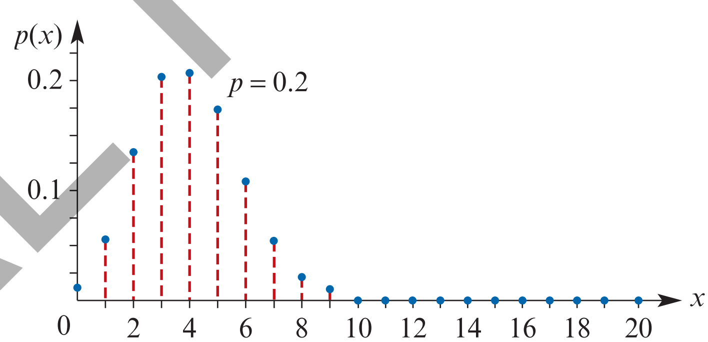
$$\Pr(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \qquad x = 0, 1, \dots, n$$

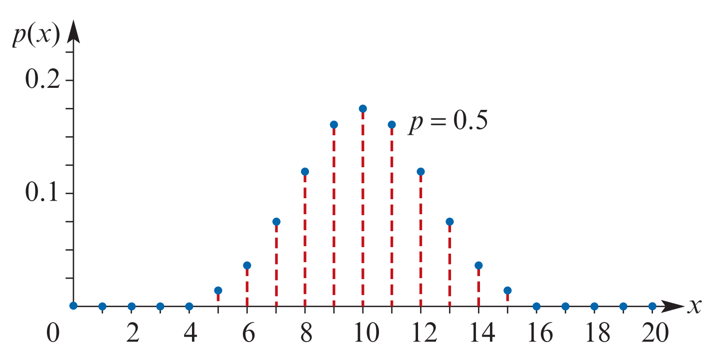
where $\binom{n}{x} = \dfrac{n!}{x!\,(n-x)!}$

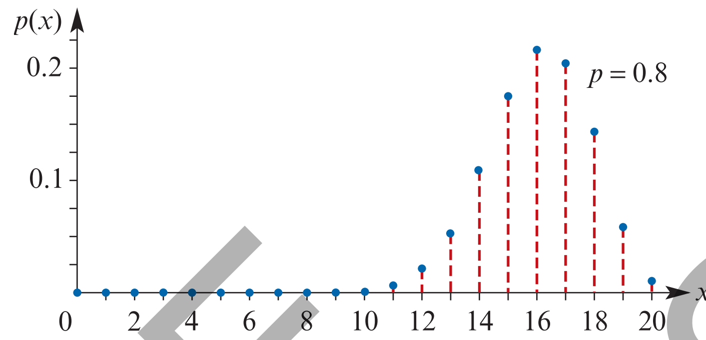
- If $X$ has a binomial probability distribution with parameters $n$ and $p$, then

$$\mathrm{E}(X) = np$$
$$\mathrm{Var}(X) = np(1 - p)$$

- The shape of the graph of a binomial probability function depends on the values of $n$ and $p$.








## Technology-free questions

**1** If $X$ is a binomial random variable with parameters $n = 4$ and $p = \frac{1}{3}$, find:

**a** $\Pr(X = 0)$ **b** $\Pr(X = 1)$ **c** $\Pr(X \le 1)$ **d** $\Pr(X \ge 1)$

SF

CF

**2** A salesperson knows that 60% of the people who enter a particular shop will make a purchase. What is the probability that of the next three people who enter the shop exactly two will make a purchase?

**3** If 10% of patients fail to improve on a certain medication, find the probability that of five patients selected at random one or more will fail to show improvement.

**4** A machine has a probability of 0.1 of manufacturing a defective part.

**a** What is the expected number of defective parts in a random sample of 20 parts manufactured by the machine?

**b** What is the standard deviation of the number of defective parts?

**5** An experiment consists of four independent trials. Each trial results in either a success or a failure. The probability of success in a trial is $p$. Find the probability of each of the following in terms of $p$:

**a** no successes **b** one success **c** at least one success

**d** four successes **e** at least two successes.

**6** A coin is tossed 10 times. The probability of three heads is $m \times (\frac{1}{2})^{10}$. State the value of $m$.

**7** An experiment consists of five independent trials. Each trial results in either a success or a failure. The probability of success in a trial is $p$. Find, in terms of $p$, the probability of exactly one success given at least one success.

**8** A die is rolled five times. What is the probability of obtaining an even number on the uppermost face on exactly three of the rolls?

**9** In a particular city, the probability of rain on any day in June is $\frac{1}{5}$. What is the probability of it raining on three of five days?

## Multiple-choice questions

**1** A coin is biased such that the probability of a head is 0.6. The probability that exactly three heads will be observed when the coin is tossed five times is

**A** $0.6 \times 3$ **B** $(0.6)^3$ **C** $(0.6)^3(0.4)^2$ **D** $10 \times (0.6)^3(0.4)^2$ **E** $\binom{5}{3}(0.6)^5$

**2** The probability that the 8:25 train arrives on time is 0.35. What is the probability that the train is on time at least once during a working week (Monday to Friday)?

**A** $1 - (0.65)^5$ **B** $(0.35)^5$ **C** $1 - (0.35)^5$
**D** $5 \times (0.35)^1(0.65)^4$ **E** $(0.65)^5$

**3** A fair die is rolled four times. The probability that a number greater than 4 is observed on two occasions is

**A** $\frac{1}{4}$ **B** $\frac{16}{81}$ **C** $\frac{1}{9}$ **D** $\frac{1}{81}$ **E** $\frac{8}{27}$

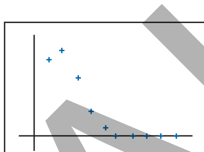
**4** The probability that a person in a certain town has a tertiary education is 0.4. What is the probability that, if 80 people are chosen at random from this town, less than 30 will have a tertiary education?

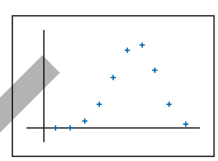
**A** 0.7139 **B** 0.2861 **C** 0.0827 **D** 0.3687 **E** 0.3750

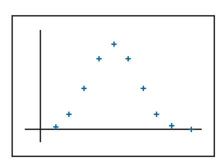
**5** If $X$ is a binomial random variable with parameters $n = 18$ and $p = \frac{1}{3}$, then the mean and variance of $X$ are closest to

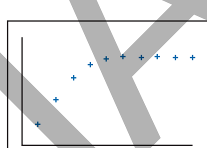
**A** $\mu = 6,\ \sigma^2 = 4$ **B** $\mu = 9,\ \sigma^2 = 4$ **C** $\mu = 6,\ \sigma^2 = 2$
**D** $\mu = 6,\ \sigma^2 = 16$ **E** $\mu = 18,\ \sigma^2 = 6$

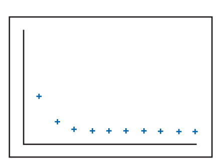
**6** Which one of the following best represents the shape of the probability distribution of a binomial random variable $X$ with 10 independent trials and probability of success 0.7?

**A**


**B**


**C**


**D**


**E**


**7** Suppose that $X$ is a binomial random variable with mean $\mu = 10$ and standard deviation $\sigma = 2$. The probability of success, $p$, in any trial is

**A** 0.4 **B** 0.5 **C** 0.6 **D** 0.7 **E** 0.8

**8** Suppose that $X$ is the number of heads observed when a coin known to be biased towards heads is tossed 10 times. If $\text{Var}(X) = 1.875$, then the probability of a head on any one toss is

**A** 0.25 **B** 0.55 **C** 0.75 **D** 0.65 **E** 0.80

*Questions 9 and 10 refer to the following information.*

The probability of Thomas beating William in a set of tennis is 0.24, and Thomas and William decide to play a set of tennis every day for $n$ days.

**9** What is the fewest number of days on which they should play to ensure that the probability of Thomas winning at least one set is more than 0.95?

**A** 7 **B** 8 **C** 9 **D** 10 **E** 11

**10** What is the fewest number of days on which they should play to ensure that the probability of Thomas winning at least two sets is more than 0.95?

**A** 12 **B** 18 **C** 17 **D** 21 **E** 14

## Extended-response questions

CU

**1** In a test to detect learning disabilities, a child is asked 10 questions, each of which has possible answers labelled $A$, $B$ and $C$. Children with a disability of type 1 almost always answer $A$ or $B$ on every question, while children with a disability of type 2 almost always answer $C$ on every question. Children without either disability have an equal chance of answering $A$, $B$ or $C$ for each question.

**a** What is the probability that the answers given by a child without either disability will be all $A$s and $B$s, thereby indicating a type 1 disability?

**b** A child is further tested for type 2 disability if he or she answers $C$ five or more times. What is the probability that a child without either disability will test positive for type 2 disability?

**2** An inspector takes a random sample of 10 items from a very large batch. If none of the items is defective, he accepts the batch; otherwise, he rejects the batch. What is the probability that a batch is accepted if the fraction of defective items is 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1? Plot these probabilities against the corresponding fraction defective. Is the inspection method a good one or not?

**3** It has been found in the past that 4% of the CDs produced in a certain factory are defective. A sample of 10 CDs is drawn randomly from each hour's production and the number of defective CDs is noted.

**a** What percentage of these hourly samples would contain at least two defective CDs?

**b** Find the mean and standard deviation of the number of defective CDs in a sample, and calculate $\mu \pm 2\sigma$.

**c** A particular sample is found to contain three defective CDs. Would this cause you to have doubts about the production process?

**4** A pizza company claims that they deliver 90% of orders within 30 minutes. In a particular 2-hour period, the supervisor notes that there are 67 orders, and of these 12 orders are delivered late. If the company claim is correct, and 90% of orders are delivered on time, what is the probability that at least 12 orders are delivered late?

CU

**5** **a** A sample of six objects is to be drawn from a large population in which 20% of the objects are defective. Find the probability that the sample contains:

**i** three defectives **ii** fewer than three defectives.

**b** Another large population contains a proportion $p$ of defective items.

**i** Write down an expression in terms of $p$ for $P$, the probability that a sample of six items contains exactly two defectives.

**ii** By differentiating to find $\frac{dP}{dp}$, show that $P$ is greatest when $p = \frac{1}{3}$.

**6** Groups of six people are chosen at random and the number, $x$, of people in each group who normally wear glasses is recorded. The table gives the results from 200 groups.

| Number wearing glasses, $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| Number of occurrences | 17 | 53 | 65 | 45 | 18 | 2 | 0 |

**a** Calculate, from the above data, the mean value of $x$.

**b** Assuming that the situation can be modelled by a binomial distribution having the same mean as the one calculated above, state the appropriate values for the binomial parameters $n$ and $p$.

**c** Calculate the theoretical frequencies corresponding to those in the table.

**7** A sampling inspection scheme is devised as follows. A sample of size 10 is drawn at random from a large batch of articles and all 10 articles are tested. If the sample contains fewer than two faulty articles, the batch is accepted; if the sample contains three or more faulty articles, the batch is rejected; but if the sample contains exactly two faulty articles, a second sample of size 10 is taken and tested. If this second sample contains no faulty articles, the batch is accepted; but if it contains any faulty articles, the batch is rejected. Previous experience has shown that 5% of the articles in a batch are faulty.

**a** Find the probability that the batch is accepted after the first sample is taken.

**b** Find the probability that the batch is rejected.

**c** Find the expected number of articles to be tested.

**8** Assume that dates of birth in a large population are distributed such that the probability of a randomly chosen person's birthday being in any particular month is $\frac{1}{12}$.

**a** Find the probability that of six people chosen at random exactly two will have a birthday in January.

**b** Find the probability that of eight people at least one will have a birthday in January.

**c** $N$ people are chosen at random. Find the least value of $N$ such that the probability that at least one will have a birthday in January exceeds 0.9.

**9** Suppose that, in flight, aeroplane engines fail with probability $q$, independently of each other, and that a plane will complete the flight successfully if at least half of its engines are still working. For what values of $q$ is a two-engine plane to be preferred to a four-engine one?