5 BINOMIAL DISTRIBUTION

Objectives

After studying this chapter you should

- be able to recognise when to use the binomial distribution;
- understand how to find the mean and variance of the distribution;
- be able to apply the binomial distribution to a variety of problems.

Note: Statistical tables can be found in many books and are also available online.

5.0   Introduction

'Bi' at the beginning of a word generally denotes the fact that the meaning involves 'two' and binomial is no exception. A random variable follows a binomial distribution when each trial has exactly **two possible outcomes**. For example, when Sarah, a practised archer, shoots an arrow at a target she either hits or misses each time.   If X is 'the number of hits Sarah scores in ten shots', then the probabilities associated with 0, 1, 2, ..., 10 hits can be expected to follow a particular pattern, known as the **binomial distribution**.

5.1   Finding the distribution

You have already met this type of distribution in Chapter 4, as can be seen in the following example.
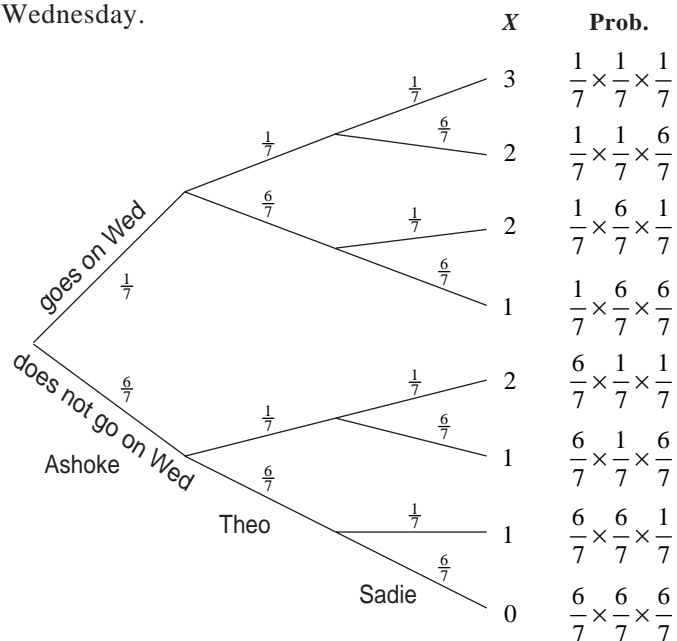
Example

Ashoke, Theo and Sadie will each visit the local leisure centre to swim on one evening next week but have made no arrangement between themselves to meet or go on any particular day. The random variable X is 'the number of the three who go to the leisure centre on Wednesday'. Find the probability distribution for X.

Solution

The probabilities of 0, 1, 2 or 3 people going on Wednesday can be found by using the tree diagram method covered in Section 1.5.

The following tree diagram shows probabilities for how many go on Wednesday.

	X	Prob.
	3	$\frac{1}{7} \times \frac{1}{7} \times \frac{1}{7}$
	2	$\frac{1}{7} \times \frac{1}{7} \times \frac{6}{7}$
	2	$\frac{1}{7} \times \frac{6}{7} \times \frac{1}{7}$
	1	$\frac{1}{7} \times \frac{6}{7} \times \frac{6}{7}$
	2	$\frac{6}{7} \times \frac{1}{7} \times \frac{1}{7}$
	1	$\frac{6}{7} \times \frac{1}{7} \times \frac{6}{7}$
	1	$\frac{6}{7} \times \frac{6}{7} \times \frac{1}{7}$
	0	$\frac{6}{7} \times \frac{6}{7} \times \frac{6}{7}$

(tree diagram with branches labelled: goes on Wed $\frac{1}{7}$, does not go on Wed $\frac{6}{7}$; Ashoke, Theo, Sadie; branch probabilities $\frac{1}{7}$ and $\frac{6}{7}$)

You can see that

$$P(X = 3) = \left(\frac{1}{7}\right)^3, \qquad P(X = 0) = \left(\frac{6}{7}\right)^3$$

and
$$P(X = 2) = \frac{1}{7} \times \frac{1}{7} \times \frac{6}{7} + \frac{1}{7} \times \frac{6}{7} \times \frac{1}{7} + \frac{6}{7} \times \frac{1}{7} \times \frac{1}{7}$$

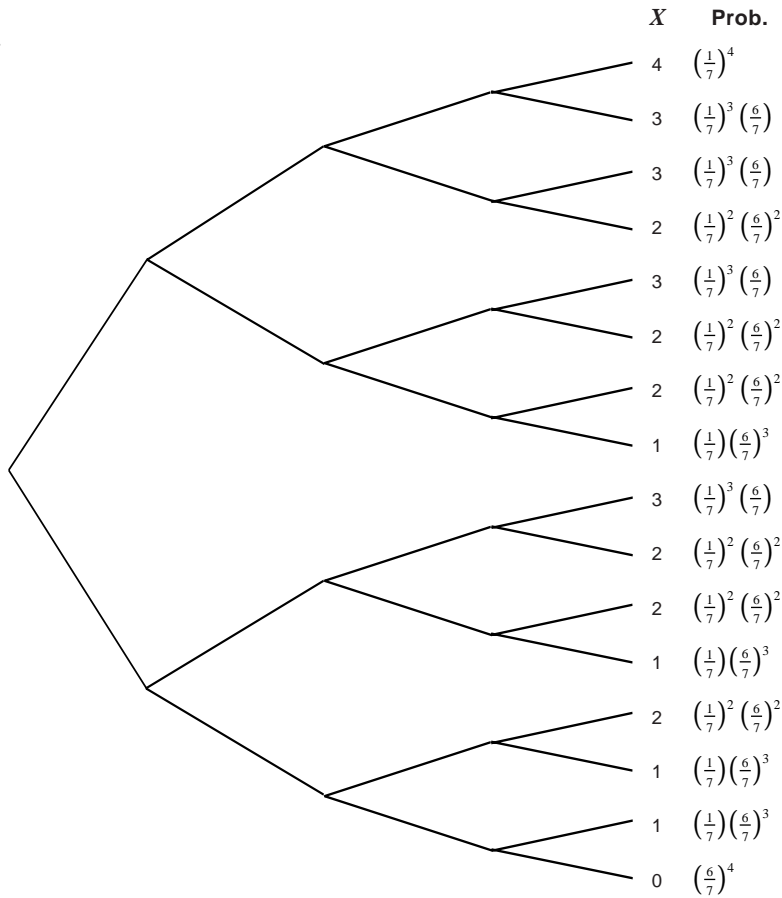$$= 3 \times \left(\frac{1}{7}\right)^2 \left(\frac{6}{7}\right)$$

$$P(X = 1) = \frac{1}{7} \times \frac{6}{7} \times \frac{6}{7} + \frac{6}{7} \times \frac{1}{7} \times \frac{6}{7} + \frac{6}{7} \times \frac{6}{7} \times \frac{1}{7}$$

$$= 3 \times \left(\frac{1}{7}\right)\left(\frac{6}{7}\right)^2.$$

This gives the table below:

x	0	1	2	3
$P(X = x)$	$\left(\frac{6}{7}\right)^3$	$3\left(\frac{1}{7}\right)\left(\frac{6}{7}\right)^2$	$3\left(\frac{1}{7}\right)^2\left(\frac{6}{7}\right)$	$\left(\frac{1}{7}\right)^3$
	$= \frac{216}{343}$	$= \frac{108}{343}$	$= \frac{18}{343}$	$= \frac{1}{343}$

The method used in the example above can be extended to a fourth person so that there will be sixteen branches to cover all the possibilities as shown in the diagram below, with X now 'the number of the four people who go on Wednesday'.

	X	Prob.

	4	$\left(\frac{1}{7}\right)^4$
	3	$\left(\frac{1}{7}\right)^3 \left(\frac{6}{7}\right)$
	3	$\left(\frac{1}{7}\right)^3 \left(\frac{6}{7}\right)$
	2	$\left(\frac{1}{7}\right)^2 \left(\frac{6}{7}\right)^2$
	3	$\left(\frac{1}{7}\right)^3 \left(\frac{6}{7}\right)$
	2	$\left(\frac{1}{7}\right)^2 \left(\frac{6}{7}\right)^2$
	2	$\left(\frac{1}{7}\right)^2 \left(\frac{6}{7}\right)^2$
	1	$\left(\frac{1}{7}\right)\left(\frac{6}{7}\right)^3$
	3	$\left(\frac{1}{7}\right)^3 \left(\frac{6}{7}\right)$
	2	$\left(\frac{1}{7}\right)^2 \left(\frac{6}{7}\right)^2$
	2	$\left(\frac{1}{7}\right)^2 \left(\frac{6}{7}\right)^2$
	1	$\left(\frac{1}{7}\right)\left(\frac{6}{7}\right)^3$
	2	$\left(\frac{1}{7}\right)^2 \left(\frac{6}{7}\right)^2$
	1	$\left(\frac{1}{7}\right)\left(\frac{6}{7}\right)^3$
	1	$\left(\frac{1}{7}\right)\left(\frac{6}{7}\right)^3$
	0	$\left(\frac{6}{7}\right)^4$

The resulting probability distribution is

x	0	1	2	3	4
$P(X=x)$	$\left(\dfrac{6}{7}\right)^4$	$4\left(\dfrac{1}{7}\right)\left(\dfrac{6}{7}\right)^3$	$6\left(\dfrac{1}{7}\right)^2\left(\dfrac{6}{7}\right)^2$	$4\left(\dfrac{1}{7}\right)^3\left(\dfrac{6}{7}\right)$	$\left(\dfrac{1}{7}\right)^4$

The fractions are as you might expect. For instance, looking at

$$P(X=2) = 6\left(\frac{1}{7}\right)^2\left(\frac{6}{7}\right)^2,$$

since you are interested in having two present and the

probability for each is $\frac{1}{7}$, the $\left(\frac{1}{7}\right)^2$ is explained, and as two are

not to attend this produces the $\left(\frac{6}{7}\right)^2$.

Explain the reason for the coefficient 6.

This has come from the second of the tree diagrams: there are six branches corresponding to two being present. This is because there are six ways of writing down two $\left(\frac{1}{7}\right)$s and two $\left(\frac{6}{7}\right)$s in a row and each produces a branch.

The six ways are shown below:

$$\frac{1}{7}\times\frac{1}{7}\times\frac{6}{7}\times\frac{6}{7}, \quad \frac{1}{7}\times\frac{6}{7}\times\frac{1}{7}\times\frac{6}{7}, \quad \frac{1}{7}\times\frac{6}{7}\times\frac{6}{7}\times\frac{1}{7},$$

$$\frac{6}{7}\times\frac{1}{7}\times\frac{1}{7}\times\frac{6}{7}, \quad \frac{6}{7}\times\frac{1}{7}\times\frac{6}{7}\times\frac{1}{7}, \quad \frac{6}{7}\times\frac{6}{7}\times\frac{1}{7}\times\frac{1}{7}.$$

Pascal's Triangle

There are four fractions to write down, two of each type, and the number of different ways of combining these is six. You don't want to draw a tree diagram every time so another method can be developed. For example, having ten people going to the leisure centre would need a tree with $2^{10} = 1024$ branches.

If you can produce the fractions in the probability distribution then all that is needed is a way of getting the right numbers to put with them. With three people the numbers were 1, 3, 3, 1 and with four they were 1, 4, 6, 4, 1.

You might recognise these as being rows of **Pascal's Triangle**, shown opposite.

```
              1    1
           1    2    1
         1    3    3    1
       1    4    6    4    1
     1    5   10   10    5    1
   1    6   15   20   15    6    1
 1    7   21   35   35   21    7    1
1    8   28   56   70   56   28    8    1
1   9   36   84  126  126   84   36   9    1
1  10   45  120  210  252  210  120   45   10    1
```

For three to be present you need $\left(\frac{1}{7}\right)^3$ and for seven to stay away the term is $\left(\frac{6}{7}\right)^7$. When writing down these fractions, how many different ways are there of combining them? According to Pascal's Triangle there are 120, so

$$P(X=3)=120\left(\frac{1}{7}\right)^3\left(\frac{6}{7}\right)^7$$

which is easier to calculate than drawing 1024 branches!   The number of ways of choosing 3 from 10 is often written

$$\binom{10}{3} \quad \text{or} \quad {}^{10}C_3,$$

so
$$p(X=3)=\binom{10}{3}\left(\frac{1}{7}\right)^3\left(\frac{6}{7}\right)^7$$

and the probability distribution is

$$P(X = x) = \binom{10}{x}\left(\frac{1}{7}\right)^x\left(\frac{6}{7}\right)^{10-x} \quad x = 0, 1, \ldots, 10.$$

Whilst the values needed can easily be read off Pascal's Triangle, there is an even easier way of working out the coefficients given in terms of factorials. Note that $n! = n(n-1)\ldots 2.1$ and, for example,

$$\binom{10}{3} = \frac{10!}{7!\,3!}$$

$$= \frac{10 \times 9 \times 8 \times (7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1)}{(7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1) \times (3 \times 2 \times 1)}$$

$$= \frac{10 \times 9 \times 8}{3 \times 2 \times 1} \quad \left(\text{since } \frac{10!}{7!} = 10 \times 9 \times 8\right)$$

$$= 120.$$

Activity 1

Check the values of

$$\binom{10}{1}, \quad \binom{10}{2}, \quad \binom{10}{4} \quad \text{and} \quad \binom{10}{5}$$

by using Pascal's Triangle and the factorial formula.

It should also be noted that most calculators have the facility to produce the coefficients. Another notation used for the number of different ways of combining three from ten is $^{10}C_3$, and the majority of calculators have a key labelled with one of ^nC_r or $\binom{n}{r}$. Your instruction book will tell you how to use this function.

In general, when the probability of success is p (instead of $\frac{1}{7}$), and the experiment is repeated n independent times, the probabilitry distribution for the number of successes is given by

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

The notation $X \sim B(n, p)$ is often used.

Activity 2   E.S.P. test

Take five cards numbered from one to five. Seat two people back to back and give the cards to one of them. This person selects a card at random and the other participant tries to identify it.

This is done five times and repeated with other pairs of people. Record the results.

Now, if X is 'the number predicted correctly out of five attempts', the probability distribution is

$$P(X = x) = \binom{5}{x}(0.2)^x (0.8)^{5-x} \quad \text{for } x = 0,\ 1,\ \dots\ 5.$$

So, for example, the probability of getting one correct is given by

$$P(X = 1) = \binom{5}{1}(0.2)(0.8)^4$$

$$= 5(0.2)(0.8)^4$$

$$= 0.4096.$$

Hence for 20 people, you would expect

$$20 \times 0.4096 = 8.192 \approx 8$$

to get one correct. A table of expected frequencies is shown opposite.

Compare your observed frequencies from Activity 2.

How closely do they match the expected frequencies?

You may want to discard results from anyone who you feel has not co-operated, perhaps saying the same number every time.

Number correct (x)	$p\,(x)$	Expected frequencies 20 $p\,(x)$
0	0.3277	6.554
1	0.4096	8.192
2	0.2048	4.096
3	0.0512	1.024
4	0.0064	0.128
5	0.0032	0.064

Example

If X is binomially distributed with 6 trials and a probability of success equal to $\frac{1}{4}$ at each attempt, what is the probability of:

(a)   exactly 4 successes   (b)   at least one success?

Solution

This question can be rewritten in the following way.

If $X \sim B\left(6, \frac{1}{4}\right)$, what is: (a) $P(X = 4)$ (b) $P(X \geq 1)$?

(a) $P(X = 4) = \begin{pmatrix} 6 \\ 4 \end{pmatrix}\left(\dfrac{1}{4}\right)^4\left(\dfrac{3}{4}\right)^2$

$$= 15 \times \dfrac{1}{256} \times \dfrac{9}{16}$$

$$= \dfrac{135}{4096} = 0.033 \ \left(\text{to 3 d.p.}\right).$$

(b) $P(X \geq 1) = 1 - P(X = 0)$

$$= 1 - \left(\dfrac{3}{4}\right)^6$$

$$= 1 - \dfrac{729}{4096}$$

$$= \dfrac{3367}{4096} = 0.822 \ \left(\text{to 3 d.p.}\right).$$

Note that tables giving cumulative binomial probabilities are given in the Appendix (p 253) and these can be used where appropriate.

Example

When an unbiased coin is tossed eight times what is the probability of obtaining:

(a)   less than 4 heads   (b)   more than five heads?

Solution

H = number of heads $\Rightarrow H \sim B(8, \ 0.5)$

(a)   Using the appropriate table in the Appendix you can simply   write down the answer,

$$P(H < 4) = P(H \leq 3)$$

$$= 0.3633.$$

Alternatively,

$$P(H \le 3) = P(H=0) + P(H=1) + P(H=2) + P(H=3)$$

$$= \left(\frac{1}{2}\right)^8 + \binom{8}{1}\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^7 + \binom{8}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^6$$

$$+ \binom{8}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^5$$

$$= \left(\frac{1}{2}\right)^8 + 8\left(\frac{1}{2}\right)^8 + 28\left(\frac{1}{2}\right)^8 + 56\left(\frac{1}{2}\right)^8$$

$$= 93\left(\frac{1}{2}\right)^8$$

$$= \frac{93}{256}$$

$$= 0.3633 \text{ (to 4 d.p.).}$$

(b) $$P(H > 5) = 1 - P(H \le 5)$$

$$= 1 - 0.8555 \quad \text{(from the table)}$$

$$= 0.1445.$$

or

$$P(H > 5) = P(H=6) + P(H=7) + P(H=8)$$

$$= \binom{8}{6}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^2 + \binom{8}{7}\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^8$$

$$= 28\left(\frac{1}{2}\right)^8 + 8\left(\frac{1}{2}\right)^8 + \left(\frac{1}{2}\right)^8$$

$$= 37\left(\frac{1}{2}\right)^8$$

$$= \frac{37}{256}$$

$$= 0.1445 \text{ (to 4 d.p.).}$$

Discuss why it is true that $\begin{pmatrix} 8 \\ 2 \end{pmatrix}$ is the same as $\begin{pmatrix} 8 \\ 6 \end{pmatrix}$.

Will it always be true that $\begin{pmatrix} n \\ r \end{pmatrix} = \begin{pmatrix} n \\ n-r \end{pmatrix}$?

Exercise 5A

Where decimals are used give answers correct to 3 significant figures.

1. If $X \sim B(6, \frac{1}{3})$ find:

    (a) $P(X=2)$ (b) $P(X<2)$ (c) $P(X \geq 1)$.

2. If $X \sim B(10, 0.3)$ find:

    (a) $P(X=9)$ (b) $P(X=0)$ (c) $P(X \leq 5)$.

3. A regular tetrahedron has three white faces and one red face. It is rolled four times and the colour of the bottom face is noted. What is the most likely number of times that the red face will end downwards?

4. If the probability that I get a lift to work on any morning is 0.6 what is the probability that in a working week of five days I will get a lift only twice?

5. When a consignment of pens arrives at the retailer's, ten of them are tested. The whole batch is returned to the wholesaler if more than one of those selected is found to be faulty. What is the probability that the consignment will be accepted if 2% of the pens are faulty?

5.2 The mean and variance of the binomial distribution

If you play ten games of table tennis against an opponent who, from past experience, you know only has a $\frac{1}{5}$ chance of winning a game with you, how many games do you expect him to win?

Most people would reply 'two' and would argue that since the opponent wins on average $\frac{1}{5}$ of the games he can expect to be successful in $\frac{1}{5} \times 10 = 2$.

Another way of writing this would be to say, if $X \sim B(10, \frac{1}{5})$, what is the value of $E(X)$? The answer then is
$E(X) = 10 \times \frac{1}{5} = 2$.

In general, if $X \sim B(n, p)$, then the expected value of X is given by

$$E(X) = n\,p$$

The formal proof of this result requires some work from pure maths, and in particular uses the result that

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

Note that it will **not** be examined in the AEB Statistics paper.

Now $E(X) = \sum_{0}^{n} x\, p(x)$

$$= 0\binom{n}{0}q^n + 1\binom{n}{1}pq^{n-1} + 2\binom{n}{2}p^2 q^{n-2}$$

$$+ 3\binom{n}{3}p^3 q^{n-3} + \dots + n\binom{n}{n}p^n, \text{ where } q = 1-p$$

$$= \frac{n!}{1!(n-1)!}pq^{n-1} + \frac{2\times n!}{2!(n-2!)}p^2 q^{n-2}$$

$$+ \frac{3\times n!}{3!(n-3)!}p^3 q^{n-3} + \dots + np^n$$

$$= np\left\{\frac{(n-1)!}{1!(n-1)!}q^{n-1} + \frac{(n-1)!}{1!(n-2)!}pq^{n-2}\right.$$

$$\left.+ \frac{(n-1)!}{2!(n-3)!}p^2 q^{n-3} + \dots + p^{n-1}\right\}$$

$$= np\left\{q^{n-1} + (n-1)q^{n-2}p + \frac{(n-1)(n-2)}{2}q^{n-3}p^2 +\right.$$

$$\left.\dots + p^{n-1}\right\}$$

$$= np\,(p+q)^{n-1}, \quad \text{using the Binomial Theorem}$$

$$= np, \text{ since } p+q=1.$$

The variance of X is given by

$$\boxed{V(X) = npq}$$

The proof is even more complex than the analysis above, so it is set in the next Activity, which is optional.

*Activity 3

Using the formula

$$V(X) = E(X^2) - [E(X)]^2$$

show that

$$V(X) = npq.$$

*Activity 4

Use a computer package designed to show graphs of binomial distributions for different values of n and p to look at a variety of binomial distributions. In particular, identify the most likely outcome by picking out the tallest bar and see, for example, how much more spread out are the outcomes for B (40, 0.5) than B (10, 0.5).

Example

A biased die is thrown thirty times and the number of sixes seen is eight. If the die is thrown a further twelve times find:

(a)   the probability that a six will occur exactly twice;

(b)   the expected number of sixes;

(c)   the variance of the number of sixes.

Solution

(a)   Let X be defined by 'the number of sixes seen in twelve throws'.

Then $X \sim B\,(12,\ p)$ where $p = \dfrac{8}{30} = \dfrac{4}{15}$.

Since $X \sim B\!\left(12, \tfrac{4}{15}\right)$,

$$P(X = 2) = \binom{12}{2}\left(\frac{4}{15}\right)^2\left(\frac{11}{15}\right)^{10}$$

$$= \frac{66 \times 4^2 \times 11^{10}}{15^{12}}$$

$$= 0.211 \quad (\text{to 3 d.p.}).$$

(b) $E(X) = np = 12 \times \dfrac{4}{15} = 3.2$.

(c) $V(X) = npq = 12 \times \dfrac{4}{15} \times \dfrac{11}{15} = 2\dfrac{26}{75} = 2.347$ (to 3 d.p.).

Example

A random variable X is binomially distributed with mean 6 and variance 4.2. Find $P(X \le 6)$.

Solution

Since X is a binomial distribution,

$$\text{mean} = np = 6,$$

$$\text{variance} = npq = 4.2.$$

Dividing, $q = \dfrac{4.2}{6} = 0.7$

and so $p = 1 - q = 0.3$.

This gives $0.3n = 6$

$\Rightarrow \quad n = 20$.

Hence $X \sim B(20, 0.3) \Rightarrow P(X \le 6) = 0.6080$ (from tables).

Activity 5   Binomial quiz

Ask your fellow students, and anyone else who will participate, whether the following statements are 'true' or 'false'.

1.   *The Portrait of a Lady* was written by Henry James.
2.   Psalms is the 20th book of the Old Testament.
3.   The equatorial diameter of Mercury is about 3032 miles.
4.   Mankoya is a place in Zambia.
5.   'The Potato Eaters' is a painting by Cezanne.
6.   The Battle of Sowton was fought in 1461.

Make a frequency table to show the number of correct answers out of six for those asked.

Is there any evidence from your results that people really know some or all of the answers?

If they are just guessing, the number of correct answers, C say,

should follow a binomial distribution,

$$C \sim B\left(6, \tfrac{1}{2}\right).$$

Work out $P(C=0)$, $P(C=1)$, ..., $P(C=6)$ and multiply them by the number of people asked to get the frequencies with which you would expect 0, 1,...., 6 correct answers to occur. Draw a diagram to show how your observed and expected frequencies compare.

Note that tables of **cumulative** binomial probabilities are available. For example, $B(10, 0.1)$ has entries as shown opposite.

So if X is distributed $B(10,0.1)$, then $P(X \leq 3) = 0.9872$.

But $\qquad P(X=3) = P(X \leq 3) - P(X \leq 2)$

$$= 0.9872 - 0.9298$$

$$= 0.0574.$$

x	$P(X \leq x)$
0	0.3487
1	0.7361
2	0.9298
3	0.9872
4	0.9984
5	0.9999
6	1.0000

This figure may be inaccurate in the last digit as it has come from two numbers which have both been rounded. Using tables usually saves having to do several calculations and the benefit is considerable in cases such as $Y \sim B(30, 0.1)$ when the value of quantities like $P(Y \leq 8)$ are needed.

Exercise 5B

1. On average a bowler takes a wicket every eight overs. What is the probability that he will bowl ten overs without succeeding in getting a wicket?

2. How many times must an unbiased coin be tossed so that the probability that at least one tail will occur is at least 0.99?

3. The random variable X has a binomial distribution $B(11, p)$. If $P(X=8) = P(X=7)$ find the value of p.

4. 100 families each with three children are found to have the following number of boys.

Number of boys	0	1	2	3
Frequency	13	34	40	13

   (a) Find the probability that a single baby born is a boy.

   (b) Calculate the number of families with three children you would expect to have two boys in a sample of 100 using your value from (a).

5. A multiple choice test has twenty questions and five possible answers for each one with only one correct per question. If X is 'the number of questions answered correctly' give:

   (a) the distribution of X;

   (b) the mean and variance of X;

   (c) the probability that a student will achieve a pass mark of 10 or more purely by guessing.

6. Investigate your results from the 'drawing pins' question in Exercise 1B to see if X, the number of times a pin finishes point upwards in 10 trials, follows a binomial distribution.

7. The probability that a student will pass a maths test is 0.8. If eighteen students take the test, give the distribution of X, 'the number of students who pass', and find its most likely value.

8. (Drunkard's Walk) A drunk is ten steps away from falling in the dock. Every step he takes is either directly towards or away from the dock and he is equally likely to move in either direction. Find the probability that he will fall in the dock on his

   (a) 10th step   (b) 12th step.

   Find also the probability that he is further from the dock after ten steps than he was at the start.

9. Find the probability that at most four heads will occur when a coin is tossed ten times.

10. If the probability that Don will hit a target on any shot is 0.2 and the probability for Yvette is 0.4, which of them is more likely to score at least three hits if Don has ten goes and Yvette has five goes?

5.3   Miscellaneous Exercises

1. (a) The probability that a certain type of vacuum tube will shatter during a thermal shock test is 0.15.

   What is the probability that if 25 such tubes are tested

   (i) 4 or more will shatter,

   (ii) between 16 and 20 (inclusive) will survive?

   Another type of tube is tested in samples of 30. It is observed that on 40% of occasions all 30 survive the test. What is the probability (assumed constant) of a single tube of this type surviving the test?

   (b) A monkey in a cage is rewarded with food if it presses a button when a light flashes. Say, giving reasons, whether or not it is likely that the following variables follow the binomial distribution:

   (i) Y is the number of times the light flashes before the monkey is twice successful in obtaining the food.

   (ii) Z is the number of times that the monkey obtains the food by the time the light has flashed 20 times.   (AEB)

2. A die is biased and the probability, p, of throwing a six is known to be less than $\frac{1}{6}$. An experiment consists of recording the number of sixes in 25 throws of the die. In a large number of experiments the standard deviation of the number of sixes is 1.5. Calculate the value of p and hence determine, to two places of decimals, the probability that exactly three sixes are recorded during a particular experiment.

3. (a) For each of the experiments described below, state, giving reasons, whether a binomial distribution is appropriate.

   Experiment 1   A bag contains black, white and red marbles which are selected at random, one at a time with replacement. The colour of each marble is noted.

   Experiment 2   This experiment is a repeat of Experiment 1 except that the bag contains black and white marbles only.

   Experiment 3   This experiment is a repeat of Experiment 2 except that marbles are not replaced after selection.

   (b) On average 20% of the bolts produced by a machine in a factory are faulty. Samples of 10 bolts are to be selected at random each day. Each bolt will be selected and replaced in the set of bolts which have been produced on that day.

   (i) Calculate, to 2 significant figures, the probability that, in any one sample, two bolts or less will be faulty.

   (ii) Find the expected value and the variance of the number of bolts in a sample which will not be faulty.

4. A crossword puzzle is published in *The Times* each day of the week, except Sunday. A man is able to complete, on average, 8 out of 10 of the crossword puzzles.

   (a) Find the expected value and the standard deviation of the number of completed crosswords in a given week.

   (b) Show that the probability that he will complete at least 5 in a given week is 0.655 (to 3 significant figures).

   (c) Given that he completes the puzzle on Monday, find, to 3 significant figures, the probability that he will complete at least 4 in the rest of the week.

   (d) Find, to 3 significant figures, the probability that, in a period of four weeks, he completes 4 or less in only one of the four weeks.

5. At a certain university in Cambford students attending a first course in statistics are asked by the lecturer, Professor Thomas Bayes, to complete 10 example sheets during the course. At the end of the course each student sits an examination as a result of which he either passes of fails. Assuming that

(i) the number, *N*, of example sheets completed by any student has a binomial distribution given by

$$P(N = n) = {}^{10}C_n \left(\frac{2}{3}\right)^n \left(\frac{1}{3}\right)^{(10-n)}$$
$$n = 0, 1, \ldots, 10$$

and

(ii) the probability of a student passing the examination given that he completed *n* sheets during the course, is *n*/10,

(a) what is the (unconditional) probability that a student passes the examination?

(b) What is the probability that a student selected at random from the examination pass list had in fact completed four example sheets or less? (AEB)

*6 Thatcher's Pottery produces large batches of coffee mugs decorated with the faces of famous politicians. They are considering adopting one of the following sampling plans for batch inspection.

Method A (single sample plan) Select 10 mugs from the batch at random and accept the batch if there are 2 or less defectives, otherwise reject batch.

Method B (double sample plan) Select 5 mugs from the batch at random and accept the batch if there are no defectives, reject the batch if there are 2 or more defectives, otherwise select another 5 mugs at random. When the second sample is drawn count the number of defectives in the combined sample of 10 and accept the batch if the number of defectives is 2 or less, otherwise reject the batch.

(a) If the proportion of defectives in a batch is *p*, find, in terms of *p*, for each method in turn, the probability that the batch will be accepted.

(b) Evaluate both the above probabilities for $p = 0.2$ and $p = 0.5$.

(c) Hence, or otherwise, decide which of these two plans is more appropriate, and why.
(AEB)