

1 Binomial distribution

In this chapter you will learn

- what the binomial distribution is
- how to recognise situations where the binomial model applies
- how to find probabilities for a given binomial distribution, by calculation and from tables
- how to find the mean, variance and standard deviation of a binomial distribution

Key points from Statistics 1

- A discrete random variable takes individual values (usually integers), each with a given probability.
- The probability function $p(x)$ of a discrete random variable X is the function that gives the probability for each possible value of X , so $p(x) = P(X = x)$.
- The cumulative distribution function F is defined by

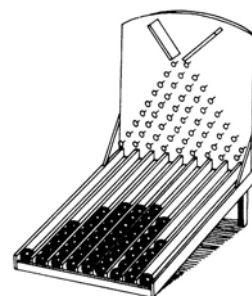
$$F(a) = P(X \leq a) = \sum_{x \leq a} p(x)$$

- The mean of a discrete random variable X is denoted by $E(X)$ or μ . It is defined as $\sum x \times P(X = x)$ or $\sum x_i p_i$, where p_i is the probability of the value x_i .
- The variance σ^2 is defined by $\sigma^2 = \text{Var}(X) = \sum (x_i - \mu)^2 p_i$. An equivalent form of this equation is $\sigma^2 = \sum x_i^2 p_i - \mu^2$. This can also be written as $\text{Var}(X) = E(X^2) - [E(X)]^2$.

A Pascal's triangle (answers p 93)

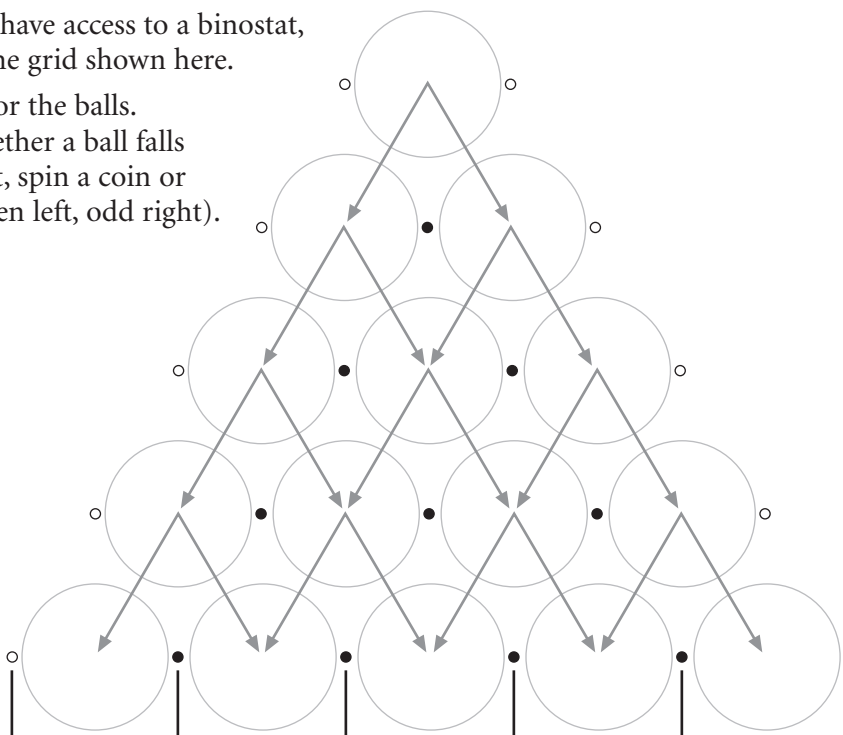
The picture on the right shows a binostat. It is like a pinball machine; balls are fed in at the top, fall through a triangular grid and collect in a series of slots. Each time a ball hits a pin, it is equally likely to fall to the left or the right.

The picture shows the distribution in the slots after a number of balls have passed through the grid.



If you do not have access to a binostat,
 you can use the grid shown here.

Use pennies for the balls.
 To decide whether a ball falls
 to left or right, spin a coin or
 roll a dice (even left, odd right).

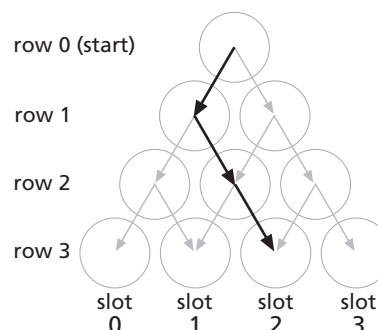


- A1** Put 20 balls through the grid and see how many end up in each of the slots at the base. Collect together the results for the whole class. Does any pattern emerge?
- A2** Can you explain why a ball is more likely to end up in one of the middle slots than in one of the end slots?

To see in more detail what is going on, consider this simple binostat. The rows and slots have been labelled for reference. You will see later why it is helpful to label the slots starting with 0 rather than 1.

Each time a ball falls from one row to the next, the probabilities of falling left or right are each $\frac{1}{2}$.

The heavy arrows show one possible route for a ball that ends up in slot 2.



- A3** (a) How many different routes are there from the start to slot 2?
 (b) What is the probability of each of these routes?
 (c) So what is the probability that a ball ends in slot 2?
- A4** The number of the slot that a ball ends in is a discrete random variable. Call it S . In the previous question you found $P(S = 2)$. Find $P(S = 0)$, $P(S = 1)$ and $P(S = 3)$. Hence make a table for the probability distribution of S :

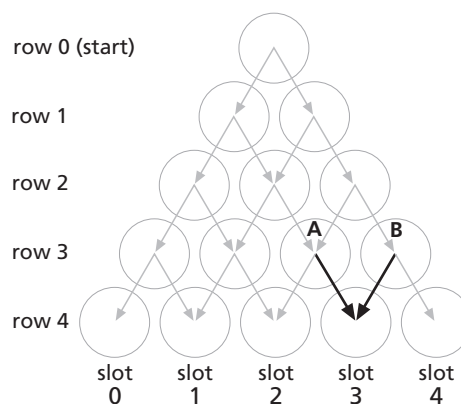
s	0	1	2	3
$P(S = s)$				

In question A3, you found the probability of ending in slot 2 like this:

- Find the number of routes to slot 2.
- Find the probability of each route.
- Multiply the number of routes by the probability of each route.

For a larger binostat the method is the same. The probability of each route is $(\frac{1}{2})^n$, where n is the last row's number. But you also need to find the number of routes to a particular slot. The next questions are about this.

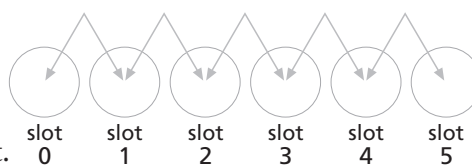
- A5** This binostat has an extra row. Think about the routes that lead to slot 3. They will have to go through either position A or position B.
- In question A3 (a), you found that there are 3 routes to position A. How many are there to position B? So how many are there to slot 3?



- A6** (a) Find the number of routes to each of the other slots in this binostat.
 (b) Find the probability of ending up in each slot.

- A7** Suppose the binostat has another row added.

- (a) Use a similar argument to find the number of routes to each slot.
 (b) Find the probability of ending up in each slot.

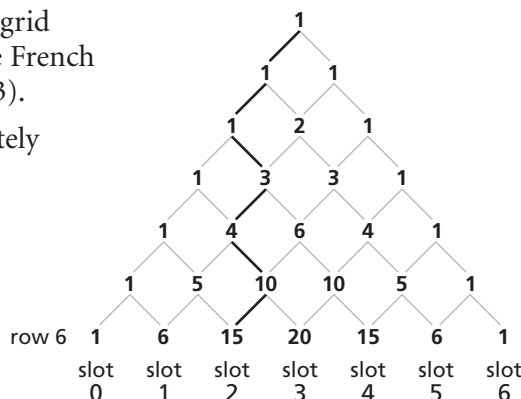


The number of routes to each position in the binostat grid is shown in this table, called **Pascal's triangle**, after the French mathematician and philosopher Blaise Pascal (1623–63).

Each number is the sum of the two numbers immediately to left and right in the previous row.

The heavy lines show one of the 15 routes leading to row 6, slot 2.

Let L stand for a fall to the left and R for a fall to the right. The route shown is LLRLRL.



- A8** Check that each of these routes leads to row 6 slot 2:
 LRLLL RLLRL LLLRR RLLLL

The routes to slot 2 consist of all the different ways of arranging two Rs and four Ls.

The numbers in Pascal's triangle have another important use. Suppose you have 6 objects A, B, C, D, E, F and you want to know how many ways there are of picking 2 of them.

Each possible way corresponds to saying 'yes' (Y) to 2 objects and 'no' (N) to the other 4. There are 15 ways, each corresponding to one of the ways of arranging two Ys and four Ns:

A B C D E F Y Y N N N N	A B C D E F Y N Y N N N	A B C D E F Y N N Y N N	A B C D E F Y N N N Y N	A B C D E F Y N N N N Y
A B C D E F N Y Y N N N	A B C D E F N Y N Y N N	A B C D E F N Y N N Y N	A B C D E F N Y N N N Y	A B C D E F N N Y Y N N
A B C D E F N N Y N Y N	A B C D E F N N Y N N Y	A B C D E F N N N Y Y N	A B C D E F N N N Y N Y	A B C D E F N N N N Y Y

The number of ways of selecting 2 objects out of 6 is denoted by $\binom{6}{2}$.

From Pascal's triangle, $\binom{6}{2} = 15$.

K The number of ways of selecting r objects from n is denoted by $\binom{n}{r}$.

A9 (a) Write down rows 7 and 8 of Pascal's triangle.

(b) Hence write down the value of

(i) $\binom{7}{3}$

(ii) $\binom{7}{5}$

(iii) $\binom{8}{4}$

A10 (a) Explain why $\binom{10}{3} = \binom{10}{7}$

(b) Explain why $\binom{n}{n-r} = \binom{n}{r}$

There is a formula for $\binom{n}{r}$. (The explanation can be found in Core 2, pages 44–46.)

K $\binom{n}{1} = n$ $\binom{n}{2} = \frac{n(n-1)}{1 \times 2}$ $\binom{n}{3} = \frac{n(n-1)(n-2)}{1 \times 2 \times 3}$

$\binom{n}{r} = \frac{n(n-1)(n-2)\dots}{1 \times 2 \times 3 \times \dots \times r}$ ← Continue until there are r factors here.

For example, $\binom{9}{4} = \frac{9 \times 8 \times 7 \times 6}{1 \times 2 \times 3 \times 4} = 126$

Another notation for $\binom{n}{r}$ is ${}_nC_r$ (or nC_r). (The C stands for 'combinations', another word for selections.) Your calculator should have have an ${}_nC_r$ key.

To find, say, $\binom{9}{3}$, enter 9, press ${}_nC_r$, enter 3, then press =.

The next example shows how values of $\binom{n}{r}$ are used to work out probabilities.

Example 1

A coin is tossed 7 times. Find the probability of getting exactly 3 heads.

Solution

One way of getting 3 heads is HHHTTTT. The probability of this is $(\frac{1}{2})^7$.

Other ways are HTHHTTT, HHTHTTT, and so on. Each has the probability $(\frac{1}{2})^7$.

There are $\binom{7}{3}$ ways altogether.

$$\text{So } P(3 \text{ heads}) = \binom{7}{3} \times (\frac{1}{2})^7 = \frac{7 \times 6 \times 5}{1 \times 2 \times 3} \times \frac{1}{128} = 35 \times \frac{1}{128} = \frac{35}{128}.$$

Exercise A (answers p 93)

- 1 A coin is tossed six times.
 - (a) How many ways are there of getting exactly four heads?
 - (b) What is the probability of each of these ways (for example HTHTHH)?
 - (c) What is the probability of getting exactly four heads?

- 2 An ordinary pack of 52 playing cards is cut at random seven times. The outcome can be 'red' (diamonds or hearts) or 'black' (clubs or spades). Find the probability that the outcome is red on exactly four occasions.

- 3 Assume that boy and girl births are equally likely and that the sex of a child is independent of the sex of previous children. Find the probability that in a family of five children

(a) none is a boy	(b) one is a boy	(c) two are boys
(d) three are boys	(e) four are boys	(f) there are more girls than boys

- 4 A coin is tossed six times.
 - (a) Find, as a fraction, the probability of getting

(i) no heads	(ii) one head	(iii) two heads
--------------	---------------	-----------------
 - (b) The number of heads, H , is a discrete random variable which can take the values 0, 1, 2, 3, 4, 5, 6. Copy and complete this table to show the probability distribution of H .

h	0	1	2	3	4	5	6
$P(H = h)$							

B Unequal probabilities

In every case looked at so far, there have been two possible outcomes (left or right, yes or no, and so on) each with probability $\frac{1}{2}$.

In this section we look at cases where there are still two possible outcomes, but with unequal probabilities. For example, if an ordinary dice is rolled, we can consider the possible outcomes as either a six, with probability $\frac{1}{6}$, or not a six, with probability $\frac{5}{6}$.

Example 2

An ordinary dice is rolled 5 times. Find the probability that it shows six exactly 3 times.

Solution

Let S stand for 'six' and N for 'not a six'. One way of getting 3 sixes is SSSNN.

The probability of this particular outcome is $(\frac{1}{6})^3 \times (\frac{5}{6})^2$.

The probability of each other outcome with 3 sixes and 2 non-sixes is also $(\frac{1}{6})^3 \times (\frac{5}{6})^2$.

There are $\binom{5}{3}$ different ways of getting 3 sixes and 2 non-sixes.

So $P(3 \text{ sixes in } 5 \text{ rolls}) = \binom{5}{3} \times (\frac{1}{6})^3 \times (\frac{5}{6})^2 = 0.0322$ (to 3 s.f.)

Exercise B (answers p 93)

- An ordinary dice is rolled three times. Find the probability of getting
 - a six on the first two rolls but not on the third
 - a six on any two rolls and a non-six on the other
- A fair tetrahedral dice has four faces. One is red and the other three are black. The dice is rolled five times. Find the probability that the dice shows red
 - three times
 - four times
 - at least three times
- A box contains two red counters and three white counters. A counter is picked at random from the box and then replaced. This is done six times. Find the probability that a red counter is picked
 - three times
 - four times
 - five times
- An ordinary pack of 52 playing cards is cut at random eight times.
 - Find the probability of cutting diamonds
 - never
 - once
 - twice
 - three times
 - Hence find the probability of cutting diamonds more than three times.
- An ordinary dice is rolled five times. The number of sixes, S , is a discrete random variable.

Copy and complete this table for the probability distribution of S .

s	0	1	2	3	4	5
$P(S = s)$						

C Using the binomial distribution (answers p 93)

The cases we have looked at so far can be described as follows.

A given number of 'trials' are carried out (for example, rolling a dice five times).

There are two possible outcomes of each trial: call them 'success' and 'failure' (for example, 'success' could mean 'getting a six').

The trials are independent of each other and the probability of success is the same in every trial.

Suppose the number of trials is n and the probability of success in each trial is p . When you know the values of n and p , you can work out the probability of getting a given number of successes.

Example 3

Given that the number n of independent trials is 5, and the probability p of success in each trial is 0.4, find the probability of getting 3 successes.

Solution

The probability of failure = $1 - p = 0.6$.

One way to get 3 successes is SSSFF (S = success and F = failure).

The probability of this particular outcome is $(0.4)^3 \times (0.6)^2$.

The probability of each other outcome with 3 successes and 2 failures is also $(0.4)^3 \times (0.6)^2$.

There are $\binom{5}{3}$ different ways of getting 3 successes and 2 failures.

So $P(3 \text{ successes}) = \binom{5}{3} \times (0.4)^3 \times (0.6)^2 = 0.230$ (to 3 s.f.)

C1 The number of successes in the example above is a discrete random variable, X , which can take the values 0, 1, 2, 3, 4, 5.

You have already seen that $P(X = 3) = 0.230$ (to 3 s.f.).

Find the other values of $P(X = x)$ and show the probability distribution of X in a table.

x	0	1	2	3	4	5
$P(X = x)$				0.230		

(The probabilities may not add up to 1 exactly because of rounding errors.)

C2 Change the values of n and p to $n = 4$ and $p = 0.8$.

Let X denote the number of successes.

Calculate the values of $P(X = x)$

for $x = 0, 1, 2, 3$ and 4 and show the probability distribution of X in a table.

x	0	1	2	3	4
$P(X = x)$					

D C3 What is the general formula for the probability of getting x successes in n trials when the probability of success in each trial is p ?

K The set-up we have been working with is called the **binomial probability model**. It is defined as follows.

- A given number, n , of independent trials are carried out.
- Each trial has two possible outcomes: success (probability p) or failure (probability $1 - p$).

The number of successes, X , is a discrete random variable whose probability distribution is given by the formula

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

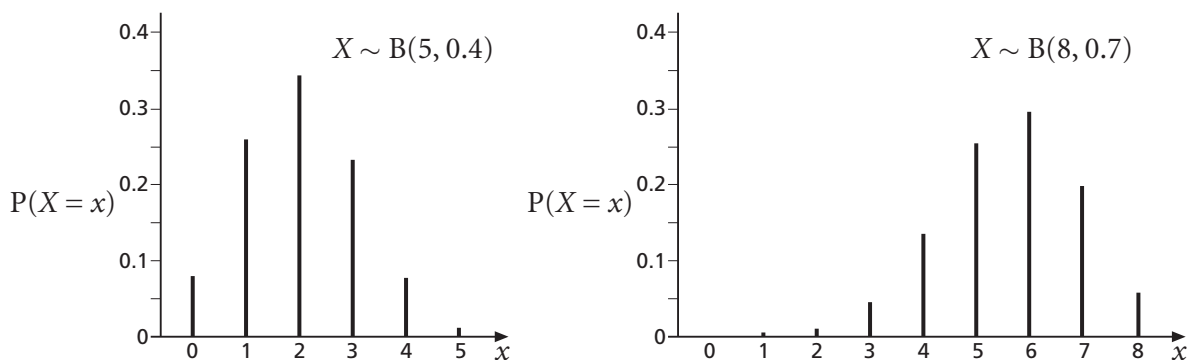
This distribution is called the **binomial distribution with index n and parameter p** .

A 'parameter' is a value which needs to be given in order to define a distribution completely.

The words in bold are abbreviated as $B(n, p)$.

' X follows the binomial distribution with $n = 7$ and $p = 0.2$ ' is written ' $X \sim B(7, 0.2)$ '.

Here are two examples of binomial distributions shown graphically.



Example 4

The probability that a Glo-brite light bulb is still working after 1000 hours of use is 0.3. Six new Glo-brite bulbs are installed. Find the probability that exactly four of them are working after 1000 hours of use. (Assume that the bulbs behave independently.)

Solution

Each of the six bulbs is a 'trial', so $n = 6$.

'Success' means 'still working after 1000 hours', so $p = 0.3$.

Let X be the number of bulbs working after 1000 hours. Then X is $B(6, 0.3)$.

$$\text{So } P(X = x) = \binom{6}{x} \times (0.3)^x \times (0.7)^{6-x}$$

$$P(X = 4) = \binom{6}{4} \times (0.3)^4 \times (0.7)^2 = 0.0595 \text{ (to 3 s.f.)}$$

Example 5

In a large population of animals it is known that the proportion of males is 45%.
 A random sample of 10 animals is selected.
 Calculate the probability that the sample contains exactly 3 males.

Solution

The population is described as 'large'. This means you can assume that the proportion of males is altered so little when a few animals are taken out that you can treat it as constant. So the probability of picking a male is always 0.45.

The number of males in the sample is $B(10, 0.45)$.

$$\text{So } P(3 \text{ males}) = \binom{10}{3}(0.45)^3(0.55)^7 = 0.166 \text{ (to 3 s.f.)}$$

Exercise C (answers p 93)

- 1 A six-sided dice is loaded (biased) so that the probability of getting a six is 0.1. The dice is rolled 5 times.
Calculate the probability of getting three sixes.
- 2 In a large population of animals of a certain type, it is known that 35% have black fur and the rest brown fur.
Four of the animals are caught at random.
Find the probability that two have black fur and two brown fur.
(You may assume that the population is so large that the percentage with black fur is not changed when a few animals are taken out.)
- 3 A test has eight questions, with yes or no answers.
James takes the test but spins a coin to decide each of his answers, so the probability of a correct answer to each question is $\frac{1}{2}$.
Calculate the probability that he gets exactly four correct answers.
- 4 Records show that on the island of Notascorcha it rains on 70% of days.
Why might it be inappropriate to use the binomial model to find the probability that it rains on five out of the seven days in a week?
- *5 In a large population of animals, males and females are in the ratio 3:2.
30% of the animals are susceptible to a certain disease.
This susceptibility is independent of the animal's sex.
 - (a) What is the probability that an animal chosen at random is a male that is susceptible to the disease?
 - (b) A sample of ten animals is selected at random from the population.
Calculate the probability that three of these animals are males susceptible to the disease.

D Using tables of the binomial distribution (answers p 94)

Often in problems it is necessary to find the sum of a number of individual probabilities. For example, suppose you want to find the probability of getting 3 or fewer successes in 6 trials where the probability of success is 0.3.

Here is the binomial distribution for $n = 6$ and $p = 0.3$.

$n = 6$	x	0	1	2	3	4	5	6
$p = 0.3$	$P(X = x)$	0.1176	0.3026	0.3241	0.1852	0.0596	0.0102	0.0007

The probability of getting 3 or fewer successes is found by adding:

$$P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0.1176 + 0.3026 + 0.3241 + 0.1852 = 0.9295$$

This is a **cumulative** probability. It is the probability that $X \leq 3$ or, in symbols, $P(X \leq 3)$.

Here is the complete table of cumulative probabilities.

$n = 6$	x	0	1	2	3	4	5	6
$p = 0.3$	$P(X \leq x)$	0.1176	0.4202	0.7443	0.9295	0.9891	0.9993	1.0000

$p =$	0.05	0.30	0.35	0.40	0.45	0.50
$n = 6, x = 0$	0.7351	0.1176	0.0754	0.0467	0.0277	0.0156
1	0.9672	0.4202	0.3191	0.2333	0.1636	0.1094
2	0.9978	0.7443	0.6471	0.5443	0.4415	0.3438
3	0.9999	0.9295	0.8826	0.8208	0.7447	0.6563
4	1.0000	0.9891	0.9777	0.9590	0.9308	0.8906
5	1.0000	0.9993	0.9982	0.9959	0.9917	0.9844

Tables of cumulative binomial probabilities are provided in a booklet for AS/A2 exams and are on pages 85–89 of this book. There is a separate table for each value of n . Values of p are printed across the page and the cumulative probabilities down the page.

You can use these tables to find $P(X = x)$ for individual values of x , or probabilities like $P(2 \leq X \leq 4)$, as the next example shows.

Example 6

Use the table of cumulative probabilities for $n = 6, p = 0.3$ to find

- (a) $P(X = 3)$ (b) $P(2 \leq X \leq 4)$ (c) $P(X \geq 4)$

Solution

- (a) The table tells you that $P(X = 0, 1, 2 \text{ or } 3) = 0.9295$ and that $P(X = 0, 1 \text{ or } 2) = 0.7443$. $P(X = 3)$ is the difference between these.

$$P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0.9295 - 0.7443 = 0.1852$$

- (b) The inequality $2 \leq X \leq 4$ means the values 2, 3, 4. The table tells you that $P(X = 0, 1, 2, 3 \text{ or } 4) = 0.9891$ and that $P(X = 0 \text{ or } 1) = 0.4202$.

$P(2 \leq X \leq 4)$ is the difference between these.

$$P(2 \leq X \leq 4) = P(X \leq 4) - P(X \leq 1) = 0.9891 - 0.4202 = 0.5689$$

- (c) The inequality $X \geq 4$ means the values 4, 5, 6. The table tells you that

$P(X = 0, 1, 2, 3, 4, 5 \text{ or } 6) = 1$ (which is obvious!) and that $P(X = 0, 1, 2 \text{ or } 3) = 0.9295$.

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.9295 = 0.0705.$$