You can use tree diagrams in any problem in which there is a clear sequence to the outcomes, including problems which are not necessarily to do with selection of objects.

Example 4.4.1

Weather records indicate that the probability that a particular day is dry is $\frac{3}{10}$. Arid F.C. is a football team whose record of success is better on dry days than on wet days. The probability that Arid win on a dry day is $\frac{3}{8}$, whereas the probability that they win on a wet day is $\frac{3}{11}$. Arid are due to play their next match on Saturday.

(a) What is the probability that Arid will win?

(b) Three Saturdays ago Arid won their match. What is the probability that it was a dry day?

Here the sequence involves first the type of weather and then the result of the football match. The tree diagram in Fig. 4.6 illustrates the information.

Notice that the probabilities in bold type were not given in the statement of the question. They have been calculated by using equations like $P(\text{wet}) + P(\text{dry}) = 1$.

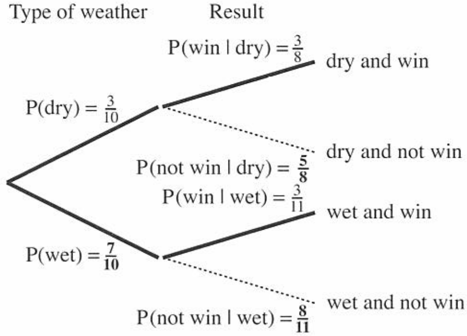


Fig. 4.6. Tree diagram for football results.

$$\begin{aligned}\text{(a) } P(\text{win}) &= P\big((\text{dry and win}) \text{ or } (\text{wet and win})\big)\\ &= P(\text{dry and win}) + P(\text{wet and win})\\ &= P(\text{dry}) \times P(\text{win} \mid \text{dry}) + P(\text{wet}) \times P(\text{win} \mid \text{wet})\\ &= \tfrac{3}{10} \times \tfrac{3}{8} + \tfrac{7}{10} \times \tfrac{3}{11} = \tfrac{9}{80} + \tfrac{21}{110} = \tfrac{267}{880}.\end{aligned}$$

(b) In this case you have been asked to calculate a conditional probability. However, here the sequence of events has been reversed and you want to find $P(\text{dry} \mid \text{win})$.

$$P(\text{dry} \mid \text{win}) = \frac{P(\text{dry and win})}{P(\text{win})} = \frac{9/80}{267/880} = \frac{99}{267}.$$

You can think of $P(\text{dry} \mid \text{win})$ *as being the proportion of times that the weather is dry out of all the times that Arid win.*

4.5 Independent events

Consider again a jar containing seven red discs and four white discs. Two discs are selected, but this time with replacement. This means that the first disc is returned to the jar before the second disc is selected.

Let R_1 be the event that the first disc is red, R_2 be the event that the second disc is red, W_1 be the event that the first disc is white and W_2 be the event that the second disc is white. You can represent the selection of the two discs with Fig. 4.7, a tree diagram similar to Fig. 4.3 but with different probabilities on the second 'layer'.

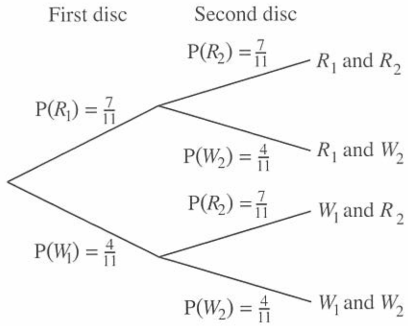


Fig 4.7. Tree diagram to show the outcomes when two discs are drawn with replacement from a jar.

The probability $P(R_2)$ that the second disc is red can also be found using the addition and multiplication laws.

$$\begin{aligned} P(R_2) &= P((R_1 \text{ and } R_2) \text{ or } P(W_1 \text{ and } R_2)) \\ &= P(R_1 \text{ and } R_2) + P(W_1 \text{ and } R_2) = \tfrac{7}{11}\times\tfrac{7}{11}+\tfrac{4}{11}\times\tfrac{7}{11}=\tfrac{7}{11}. \end{aligned}$$

In this case $P(R_2)=P(R_2 \mid R_1)$, which means that the first disc's being red has no effect on the chance of the second disc being red. This is what you would expect, since the first disc was replaced before the second was removed. Two events A and B for which $P(B \mid A)=P(B)$ are called **independent**. Independent events have no effect upon one another.

Recall also that, from the definition of conditional probability, $P(B \mid A)=\dfrac{P(A \text{ and } B)}{P(A)}$.

So when you equate the two expressions for $P(B \mid A)$ for independent events, you get

$$\frac{P(A \text{ and } B)}{P(A)} = P(B), \quad \text{which when rearranged gives} \quad P(A \text{ and } B) = P(A)\times P(B).$$

Independent events are events which have no effect on one another.
For two independents events A and B,

$$P(A \text{ and } B) = P(A)\times P(B). \qquad (4.5)$$

This result is called the **multiplication law for independent events.**

Example 4.5.1
In a carnival game, a contestant has to first spin a fair coin and then roll a fair cubical dice whose faces are numbered 1 to 6. The contestant wins a prize if the coin shows heads and the dice score is below 3. Find the probability that a contestant wins a prize.

$$P(\text{prize won}) = P((\text{coin shows heads}) \text{ and } (\text{dice score is lower than 3})).$$

The event that the coin shows heads and the event that the dice score is lower than 3 are independent, because the score on the dice can have no effect on the result of the spin of the coin. Therefore the multiplication law for independent events can be used.

$$\begin{aligned} P(\text{prize won}) &= P((\text{coin shows heads}) \text{ and } (\text{dice score is lower than 3})) \\ &= P(\text{coin shows heads}) \times P(\text{dice score is lower than 3}) = \tfrac{1}{2}\times\tfrac{2}{6}=\tfrac{1}{6}. \end{aligned}$$

The law of multiplication for independent events can be extended to more than two events, provided they are all independent of one another.

If $A_1, A_2, \ldots, A_n$ are n independent events, then

$$\mathrm{P}(A_1 \text{ and } A_2 \text{ and} \ldots \text{and } A_n) = \mathrm{P}(A_1) \times \mathrm{P}(A_2) \times \ldots \times \mathrm{P}(A_n). \qquad (4.6)$$

Example 4.5.2

A fair cubical dice with faces numbered 1 to 6 is thrown four times. Find the probability that three of the four throws result in a 6.

You can use the addition law of mutually exclusive events and the multiplication law of independent events to break the event {three of the four scores are 6} down into smaller sub-events whose probabilities you can easily determine:

$$\mathrm{P}(\text{three of the scores are 6s}) = \mathrm{P}\begin{pmatrix} (6_1 \text{ and } 6_2 \text{ and } 6_3 \text{ and } N_4) \text{ or} \\ (6_1 \text{ and } 6_2 \text{ and } N_3 \text{ and } 6_4) \text{ or} \\ (6_1 \text{ and } N_2 \text{ and } 6_3 \text{ and } 6_4) \text{ or} \\ (N_1 \text{ and } 6_2 \text{ and } 6_3 \text{ and } 6_4) \end{pmatrix},$$

where, for example, 6_1 means that the first score was a 6 and N_3 means that the third score was not a 6.

Using the addition and multiplication laws,

$$\begin{aligned}
\mathrm{P}(\text{three of the scores are 6s}) &= \mathrm{P}\begin{pmatrix} (6_1 \text{ and } 6_2 \text{ and } 6_3 \text{ and } N_4) \text{ or} \\ (6_1 \text{ and } 6_2 \text{ and } N_3 \text{ and } 6_4) \text{ or} \\ (6_1 \text{ and } N_2 \text{ and } 6_3 \text{ and } 6_4) \text{ or} \\ (N_1 \text{ and } 6_2 \text{ and } 6_3 \text{ and } 6_4) \end{pmatrix} \\
&= \mathrm{P}(6_1 \text{ and } 6_2 \text{ and } 6_3 \text{ and } N_4) \\
&\quad + \mathrm{P}(6_1 \text{ and } 6_2 \text{ and } N_3 \text{ and } 6_4) \\
&\quad\quad + \mathrm{P}(6_1 \text{ and } N_2 \text{ and } 6_3 \text{ and } 6_4) \\
&\quad\quad\quad + \mathrm{P}(N_1 \text{ and } 6_2 \text{ and } 6_3 \text{ and } 6_4) \\
&= \mathrm{P}(6_1) \times \mathrm{P}(6_2) \times \mathrm{P}(6_3) \times \mathrm{P}(N_4) \\
&\quad + \mathrm{P}(6_1) \times \mathrm{P}(6_2) \times \mathrm{P}(N_3) \times \mathrm{P}(6_4) \\
&\quad\quad + \mathrm{P}(6_1) \times \mathrm{P}(N_2) \times \mathrm{P}(6_3) \times \mathrm{P}(6_4) \\
&\quad\quad\quad + \mathrm{P}(N_1) \times \mathrm{P}(6_2) \times \mathrm{P}(6_3) \times \mathrm{P}(6_4) \\
&= \left(\tfrac{1}{6} \times \tfrac{1}{6} \times \tfrac{1}{6} \times \tfrac{5}{6}\right) + \left(\tfrac{1}{6} \times \tfrac{1}{6} \times \tfrac{5}{6} \times \tfrac{1}{6}\right) \\
&\quad + \left(\tfrac{1}{6} \times \tfrac{5}{6} \times \tfrac{1}{6} \times \tfrac{1}{6}\right) + \left(\tfrac{5}{6} \times \tfrac{1}{6} \times \tfrac{1}{6} \times \tfrac{1}{6}\right) \\
&= 4 \times \left(\tfrac{1}{6}\right)^3 \times \tfrac{5}{6} = \tfrac{5}{324}.
\end{aligned}$$

4.6 Practical activities

1 Cylinder When you throw a coin it is very unlikely to land on its edge (or curved surface). However, if you were to drop a soft-drinks can on the floor there is quite a good chance that it will land on its curved surface. Both the coin and the can are (nearly) cylindrical in shape.
(a) Find several cylinders for which the ratio of the height to the radius is different, and investigate how the ratio of height to radius affects the chance of a cylinder landing on its curved surface.
(b) Throw each cylinder 50 times and work out the experimental probability that the cylinder lands on its curved surface. Plot a graph of this experimental probability against the ratio of height to radius.
(c) For what ratio of height to radius would you estimate that the cylinder was equally likely to land on its curved surface as it is to land on one of its plane faces?

2 Buffon's needle For this experiment you require a needle (or a similar object). On a sheet of paper rule parallel lines, separated by a distance equal to half the length of the needle. Place the sheet of paper on a table and throw the needle on to the paper 'at random'. Note whether the needle falls across a line. Repeat this 100 times in all. From your results find the experimental probability that the needle will fall across a line.

This experiment is named after the Comte de Buffon (1707–1788) who carried it out in order to obtain an estimate for π. It can be shown that the theoretical probability of the needle falling across a line is $\frac{2}{\pi}$. (For a proof see, for example, *Practical Statistics* by Mary Rouncefield and Peter Holmes, Macmillan, 1993). Use your results to obtain an estimate for π.

3 Cards
(a) Shuffle a pack of cards and pick a card at random. Record the identity of the card. Repeat this to give 50 selections in all, replacing the card and shuffling after every selection. From your data calculate the experimental probabilities of picking
(i) a red card, (ii) an even score $\{2, 4, 6, 8, 10\}$, (iii) a picture card $\{J, Q, K\}$.
Check whether the following statements are true for your experimental probabilities:
(b) P(even or picture) = P(even) + P(picture),
(c) P(even or red) = P(even) + P(red), (d) P(even and red) = P(even) $\times$ P(red).
Compare with what you would expect theoretically.

Exercise 4B

1 A bag contains six red and four green counters. Two counters are drawn, without replacement. Use a carefully labelled tree diagram to find the probabilities that

(a) both counters are red, (b) both counters are green, (c) just one counter is red,

(d) at least one counter is red, (e) the second counter is red.

Compare your answers with those to Exercise 4A Question 5.

Does it make any difference to your answers to parts (a), (b), (c) and (d) if the two counters are drawn simultaneously rather than one after the other?

2 Two cards are drawn, without replacement, from an ordinary pack. Find the probabilities that

(a) both are picture cards (K, Q, J), (b) neither is a picture card,

(c) at least one is a picture card, (d) at least one is red.

3 Events A, B and C satisfy these conditions:

$\mathrm{P}(A) = 0.6$, $\mathrm{P}(B) = 0.8$, $\mathrm{P}(B \mid A) = 0.45$, $\mathrm{P}(B \text{ and } C) = 0.28$.

Calculate

(a) $\mathrm{P}(A \text{ and } B)$, (b) $\mathrm{P}(C \mid B)$, (c) $\mathrm{P}(A \mid B)$.

4 A class consists of seven boys and nine girls. Two different members of the class are chosen at random. A is the event {the first person is a girl}, and B is the event {the second person is a girl}. Find the probabilities of

(a) $B \mid A$, (b) $B' \mid A$, (c) $B \mid A'$

(d) $B' \mid A'$, (e) B.

Is it true that

(f) $\mathrm{P}(B \mid A) + \mathrm{P}(B' \mid A) = 1$, (g) $\mathrm{P}(B \mid A) + \mathrm{P}(B \mid A') = 1$?

5 A weather forecaster classifies all days as wet or dry. She estimates that the probability that 1 June next year is wet is 0.4. If any particular day in June is wet, the probability that the next day is wet is 0.6; otherwise the probability that the next day is wet is 0.3. Find the probability that, next year,

(a) the first two days of June are both wet,

(b) June 2nd is wet,

(c) at least one of the first three days of June is wet.

6 Two chess players, K1 and K2, are playing each other in a series of games. The probability that K1 wins the first game is 0.3. If K1 wins any game, the probability that he wins the next is 0.4; otherwise the probability is 0.2. Find the probability that K1 wins

(a) the first two games, (b) at least one of the first two games,

(c) the first three games, (d) exactly one of the first three games.

The result of any game can be a win for K1, a win for K2, or a draw. The probability that any one game is drawn is 0.5, independent of the result of all previous games. Find the probability that, after two games,

(e) K1 won the first and K2 the second, (f) each won one game,

(g) each has won the same number of games.

7 A fair cubical dice is thrown four times. Find the probability that

(a) all four scores are 4 or more, (b) at least one score is less than 4,

(c) at least one of the scores is a 6.

8 The Chevalier du Meré's Problem. A seventeenth-century French gambler, the Chevalier du Meré, had run out of takers for his bet that, when a fair cubical dice was thrown four times, at least one 6 would be scored. (See Question 7(c).) He therefore changed the game to throwing a pair of fair dice 24 times. What is the probability that, out of these 24 throws, at least one is a double 6?

9 The Birthday Problem. What is the probability that, out of 23 randomly chosen people, at least two share a birthday? Assume that all 365 days of the year are equally likely and ignore leap years. (Hint: find the probabilities that two people have different birthdays, that three people have different birthdays, and so on.)

10 Given that $\mathrm{P}(A)=0.75$, $\mathrm{P}(B \mid A)=0.8$ and $\mathrm{P}(B \mid A')=0.6$, calculate $\mathrm{P}(B)$ and $\mathrm{P}(A \mid B)$.

11* For any events A and B, write $\mathrm{P}(A \text{ and } B)$ and $\mathrm{P}(B)$ in terms of $\mathrm{P}(A)$, $\mathrm{P}(A')$, $\mathrm{P}(B \mid A)$ and $\mathrm{P}(B \mid A')$. Deduce Bayes' theorem: $\mathrm{P}(A \mid B)=\dfrac{\mathrm{P}(A)\mathrm{P}(B \mid A)}{\mathrm{P}(A)\mathrm{P}(B \mid A)+\mathrm{P}(A')\mathrm{P}(B \mid A')}$.

12* The Doctor's Dilemma. It is known that, among all patients displaying a certain set of symptoms, the probability that they have a particular rare disease is 0.001. A test for the disease has been developed. The test shows a positive result on 98% of the patients who have the disease and on 3% of patients who do not have the disease.

The test is given to a particular patient displaying the symptoms, and it records a positive result. Find the probability that the patient has the disease. Comment on your answer.

13* The Prosecutor's Fallacy. An accused prisoner is on trial. The defence lawyer asserts that, in the absence of further evidence, the probability that the prisoner is guilty is 1 in a million. The prosecuting lawyer produces a further piece of evidence and asserts that, if the prisoner were guilty, the probability that this evidence would be obtained is 999 in 1000, and if he were not guilty it would be only 1 in 1000; in other words, $\mathrm{P}(\text{evidence} \mid \text{guilty})=0.999$, and $\mathrm{P}(\text{evidence} \mid \text{not guilty})=0.001$. Assuming that the court admits the legality of the evidence, and that both lawyers' figures are correct, what is the probability that the prisoner is guilty? Comment on your answer.

Miscellaneous exercise 4

1 Bag A contains 1 red ball and 1 black ball, and bag B contains 2 red balls; all four balls are indistinguishable apart from their colour. One ball is chosen at random from A and is transferred to B. One ball is then chosen at random from B and is transferred to A.

(a) Draw a tree diagram to illustrate the possibilities for the colours of the balls transferred from A to B and then from B to A.

(b) Find the probability that, after both transfers, the black ball is in bag A. (OCR)

2 The probability that an event A occurs is $\mathrm{P}(A)=0.3$. The event B is independent of A and $\mathrm{P}(B)=0.4$.

(a) Calculate $\mathrm{P}(A \text{ or } B \text{ or both occur})$.

Event C is defined to be the event that neither A nor B occurs.

(b) Calculate $\mathrm{P}(C \mid A')$, where A' is the event that A does not occur. (OCR, adapted)

3 Two cubical fair dice are thrown, one red and one blue. The scores on their faces are added together. Determine which, if either, is greater:

(a) the probability that the total score will be 10 or more given that the red dice shows a 6,

(b) the probability that the total score will be 10 or more given that at least one of the dice shows a 6. (OCR)

4 Half of the A-level students in a community college study science and 30% study mathematics. Of those who study science, 40% study mathematics.

(a) What proportion of the A-level students study both mathematics and science?

(b) Calculate the proportion of those students who study mathematics but do not study science. (OCR)

5 Three friends, Ahmed, Benjamin and Chi, live in a town where there are only three cafés. They arrange to meet at a café one evening but do not specify the name of the café. The probabilities that they will each choose a particular café are independent. Ahmed lives close to Café Expresso and so the probability that he will choose to go there is $\frac{5}{9}$ whereas Café Kola and Café Pepi have equal chances of being visited by him.

Benjamin lives a long distance from Café Kola and the probability that he will choose this one is $\frac{1}{7}$, but he will choose either of the other two cafés with equal probability.

Each café has an equal chance of being visited by Chi.

(a) Show that the probability that the three friends meet at Café Expresso is $\frac{5}{63}$.

(b) Calculate the probability that

(i) the three friends will meet at the same café,

(ii) at most two friends will meet at the same café. (OCR)

6 Two events A and B are such that $P(A)=\frac{3}{4}$, $P(B \mid A)=\frac{1}{5}$ and $P(B' \mid A')=\frac{4}{7}$. By use of a tree diagram, or otherwise, find

(a) $P(A \text{ and } B)$, (b) $P(B)$, (c) $P(A \mid B)$. (OCR, adapted)

7 Students have to pass a test before they are allowed to work in a laboratory. Students do not retake the test once they have passed it. For a randomly chosen student, the probability of passing the test at the first attempt is $\frac{1}{3}$. On any subsequent attempt, the probability of failing is half the probability of failing on the previous attempt. By drawing a tree diagram, or otherwise,

(a) show that the probability of a student passing the test in 3 attempts or fewer is $\frac{26}{27}$,

(b) find the conditional probability that a student passed at the first attempt, given that the student passed in 3 attempts or fewer. (OCR)

8 The probability of event A occurring is $P(A)=\frac{13}{25}$. The probability of event B occurring is $P(B)=\frac{9}{25}$. The conditional probability of A occurring given that B has occurred is $P(A \mid B)=\frac{5}{9}$.

(a) Determine the following probabilities.

(i) $P(A \text{ and } B)$ (ii) $P(B \mid A)$ (iii) $P(A \text{ or } B \text{ or both})$ (iv) $P(A' \mid B')$.

(b) Determine $P(A \text{ occurs or } B \text{ does not occur})$ showing your working.

(OCR, adapted)

9 (a) The probability that an event A occurs is $P(A)=0.4$. B is an event independent of A and $P(A \text{ or } B \text{ or both})=0.7$. Find $P(B)$.

(b) C and D are two events such that $P(D \mid C)=\frac{1}{5}$ and $P(C \mid D)=\frac{1}{4}$. Given that $P(C \text{ and } D \text{ occur})=p$, express in terms of p (i) $P(C)$, (ii) $P(D)$.

(c) Given also that $P(C \text{ or } D \text{ or both occur})=\frac{1}{5}$, find the value of p. (OCR, adapted)

10 A batch of forty tickets for an event at a stadium consists of ten tickets for the North stand, fourteen tickets for the East stand and sixteen tickets for the West stand. A ticket is taken from the batch at random and issued to a person, X. Write down the probability that X has a ticket for the North stand.

A second ticket is taken from the batch at random and issued to Y. Subsequently a third ticket is taken from the batch at random and issued to Z. Calculate the probability that

(a) both X and Y have tickets for the North stand,

(b) X, Y and Z all have tickets for the same stand,

(c) two of X, Y and Z have tickets for one stand and the other of X, Y and Z has a ticket for a different stand. (OCR)

11 In a lottery there are 24 prizes allocated at random to 24 prize-winners. Ann, Ben and Cal are three of the prize-winners. Of the prizes, 4 are cars, 8 are bicycles and 12 are watches. Show that the probability that Ann gets a car and Ben gets a bicycle or a watch is $\frac{10}{69}$.

Giving each answer either as a fraction or as a decimal correct to 3 significant figures, find

(a) the probability that both Ann and Ben get cars, given that Cal gets a car,

(b) the probability that either Ann or Cal (or both) gets a car,

(c) the probability that Ann gets a car and Ben gets a car or a bicycle,

(d) the probability that Ann gets a car given that Ben gets either a car or a bicycle. (OCR)

12 In a certain part of the world there are more wet days than dry days. If a given day is wet, the probability that the following day will also be wet is 0.8. If a given day is dry, the probability that the following day will also be dry is 0.6.

Given that Wednesday of a particular week is dry, calculate the probability that

(a) Thursday and Friday of the same week are both wet days,

(b) Friday of the same week is a wet day.

In one season there were 44 cricket matches, each played over three consecutive days, in which the first and third days were dry. For how many of these matches would you expect that the second day was wet? (OCR)

13 A study of the numbers of male and female children in families in a certain population is being carried out.

(a) A simple model is that each child in any family is equally likely to be male or female, and that the sex of each child is independent of the sex of any previous children in the family. Using this model calculate the probability that, in a randomly chosen family of 4 children,

(i) there will be 2 males and 2 females,

(ii) there will be exactly 1 female given that there is at least 1 female.

(b) An alternative model is that the first child in any family is equally likely to be male or female, but that, for any subsequent children, the probability that they will be of the same sex as the previous child is $\frac{3}{5}$. Using this model, calculate the probability that, in a randomly chosen family of 4 children,

(i) all four will be of the same sex,

(ii) no two consecutive children will be of the same sex,

(iii) there will be 2 males and 2 females. (OCR, adapted)

14 A dice is known to be biased in such a way that, when it is thrown, the probability of a 6 showing is $\frac{1}{4}$. This biased dice and an ordinary fair dice are thrown. Find the probability that

(a) the fair dice shows a 6 and the biased dice does not show a 6,

(b) at least one of the two dice shows a 6,

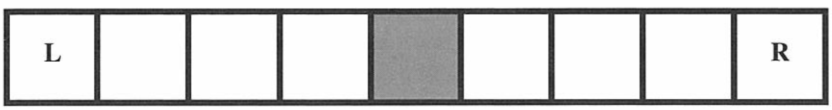
(c) exactly one of the two dice shows a 6, given that at least one of them shows a 6.

(OCR)

15 Spares of a particular component are produced by two firms, Bestbits and Lesserprod. Tests show that, on average, 1 in 200 components produced by Bestbits fail within one year of fitting, and 1 in 50 components produced by Lesserprod fail within one year of fitting. Given that 20 per cent of the components sold and fitted are made by Bestbits and 80 per cent by Lesserprod, what is the probability that a component chosen at random from those sold and fitted will fail within a year of fitting?

Find the proportion of components sold and fitted that would need to be made by Bestbits for this probability to be 0.01. (OCR)

16 A game is played using a regular 12-faced fair dice, with faces labelled 1 to 12, a coin and a simple board with nine squares as shown in the diagram.



Initially, the coin is placed on the shaded rectangle. The game consists of rolling the dice and then moving the coin one rectangle towards **L** or **R** according to the outcome on the dice. If the outcome is a prime number (2, 3, 5, 7 or 11) the move is towards **R**, otherwise it is towards **L**. The game stops when the coin reaches either **L** or **R**. Find, giving your answers correct to 3 decimal places, the probability that the game

(a) ends on the fourth move at **R**,

(b) ends on the fourth move,

(c) ends on the fifth move,

(d) takes more than six moves. (OCR)

17 You have n identical black balls, and n identical white balls, identical with the black balls except for their colour.

Box X contains 3 of the black balls, and $n-3$ white balls. Box Y contains $n-3$ black balls, and 3 white balls.

A ball is taken at random from box X and put into box Y. A ball is then taken at random from box Y.

(a) Calculate in terms of n the probability that the ball taken from box Y is white.

(b) Calculate in terms of n the probability that the first ball is black, given that the second ball is white. (OCR)

5 Permutations and combinations

This chapter is about numbers of arrangements of different objects, and the number of ways you can choose different objects. When you have completed it, you should

- know what a permutation is, and be able to calculate with permutations
- know what a combination is, and be able to calculate with combinations
- be able to apply permutations and combinations to probability.

5.1 Permutations

In the last chapter you often had to count the number of outcomes in a sample space. When the number of outcomes is fairly small, this is quite straightforward, but suppose, for example, that you were to deal hands of 5 playing cards from a standard pack of 52 cards. The number of different hands which you might receive is very large and it would not be sensible to try to list all of them. It is necessary to find a method of counting the number of possible hands which avoids writing out a complete list.

It is useful to start with an easier situation. Suppose that you have the three letters A, B and C, one on each of three separate cards, and that you are going to arrange them in a line to form 'words'. How many three-letter words are there?

In this case the number of words is small enough for you to write them out in full.

$ABC \quad ACB \quad BCA \quad BAC \quad CAB \quad CBA$

You could also show the possible choices by using a tree diagram, as in Fig. 5.1.

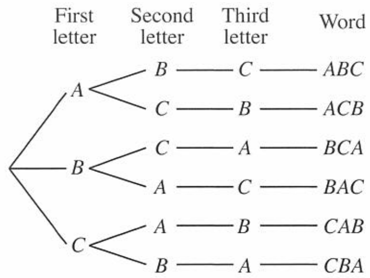


Fig. 5.1. Possible words made from the letters A, B and C.

You have 3 choices for the first letter:

either A, B or C.

Having chosen the first letter, you then have just 2 choices for the second letter:

B or C if A has been used,
C or A if B has been used,
A or B if C has been used.

Having chosen the first two letters, you then have only 1 choice for the third letter:

only C is left if A and B have been used,
only B is left if C and A have been used,
only A is left if B and C have been used.

So altogether there are $3 \times 2 \times 1 = 6$ possible words you can make with three letters.

Using a similar argument you can find the number of words which you can make from four letters A, B, C and D.

> You have 4 possibilities for the first letter.
> Having chosen the first letter, you then have 3 possibilities for the second letter.
> Having chosen the first two letters, you then have 2 possibilities for the third letter.
> You then have only 1 possibility for the last letter.

Therefore there are $4\times3\times2\times1=24$ possible words.

Each of these 24 possibilities is listed in Fig. 5.2.

ABCD	*BACD*	*CABD*	*DABC*
ABDC	*BADC*	*CADB*	*DACB*
ACBD	*BCAD*	*CBAD*	*DBAC*
ACDB	*BCDA*	*CBDA*	*DBCA*
ADBC	*BDAC*	*CDAB*	*DCAB*
ADCB	*BDCA*	*CDBA*	*DCBA*

Fig. 5.2. All possible arrangements of the letters A, B, C and D.

You can now generalise this result to the case where there are n distinct letters.

> When you arrange n distinct letters in a line, the number of different 'words' you can make is
>
> $$n\times(n-1)\times(n-2)\times\ldots\times2\times1.$$

The argument and the result given above apply whenever n distinct objects are arranged in a line. The objects need not be letters. The different arrangements of the objects are called **permutations**.

It is useful to have a concise way of writing the expression

$$n\times(n-1)\times(n-2)\times\ldots\times2\times1.$$

The expression $n\times(n-1)\times(n-2)\times\ldots\times2\times1$ is called n **factorial** and written as $n!$.

> The number of permutations of n distinct objects is $n!$, where
>
> $$n!=n\times(n-1)\times(n-2)\times\ldots\times2\times1. \qquad (5.1)$$

Table 5.3 shows how the number of permutations increases as the number of objects being arranged gets larger.

Number of objects, n	Number of permutations
1	$1=1$
2	$2\times1=2$
3	$3\times2\times1=6$
4	$4\times3\times2\times1=24$
5	$5\times4\times3\times2\times1=120$

Table 5.3. Number of permutations as n increases.

As you can see, the number of permutations increases very rapidly as the number of objects being arranged gets larger. For example,

$$8!=8\times7\times6\times5\times4\times3\times2\times1=40\,320.$$

Your calculator will probably give you values of $n!$, but it can only approximate by using standard index form as n gets larger.

Suppose now that you have more letters than you need to make a word. For example, suppose that you have the seven letters A, B, C, D, E, F and G but that you want to make a four-letter word.

You have 7 choices for the first letter.
Having chosen the first letter, you have 6 choices for the second letter.
Having chosen the first two letters, you have 5 choices for the third letter.
Having chosen the first three letters, you have 4 choices for the fourth (and last) letter.

The number of permutations of 4 letters chosen from 7 letters (that is, the number of four-letter words) is therefore $7\times6\times5\times4$.

This result can be written concisely in terms of factorials.

$$7\times6\times5\times4=\frac{7\times6\times5\times4\times3\times2\times1}{3\times2\times1}=\frac{7!}{3!}=\frac{7!}{(7-4)!}.$$

This is an illustration of a general rule.

The number of different permutations of r objects which can be made from n distinct objects is $\frac{n!}{(n-r)!}$. This number is usually given the special symbol ^nP_r.

> The number ^nP_r of different permutations of r objects which can be made from n distinct objects is given by $^nP_r=\frac{n!}{(n-r)!}$. (5.2)

On some calculators, this is written as $_nP_r$.

Example 5.1.1
Eight runners are hoping to take part in a race, but the track has only six lanes. In how many ways can six of the eight runners be assigned to lanes?

Using Equation 5.2, the number of permutations is

$$^{8}P_{6} = \frac{8!}{(8-6)!} = \frac{8!}{2!} = \frac{8\times7\times6\times5\times4\times3\times2\times1}{2\times1} = 20\,160.$$

Notice that, if you try to use the formula $^{n}P_{r} = \dfrac{n!}{(n-r)!}$ to find the number of permutations of the eight letters $A, B, C, \dots, H$, you find that

$$^{8}P_{8} = \frac{8!}{(8-8)!} = \frac{8!}{0!}.$$

But $0!$ cannot be defined using the relationship $n! = n\times(n-1)\times(n-2)\times\ldots\times2\times1$, so what does $0!$ mean? The answer is that it can be defined to be any convenient value. Recall that $^{8}P_{8}$ denotes the number of permutations of 8 objects when you have selected them from 8 distinct objects. But this has already been shown to be $8!$.

Therefore $\dfrac{8!}{0!}$ should equal $8!$.

This will only be true if $0! = 1$. The value of $0!$ is defined to be 1 to make the formula for $^{n}P_{r}$ consistent when $r = n$.

Example 5.1.2
Eight people, $A, B, \dots, H$, are arranged randomly in a line. What is the probability that
(a) A and B are next to each other, (b) A and B are not next to each other?

(a) There is a neat trick which helps to solve this problem. Imagine that A and B are stuck together in the order AB. There are then $7!$ ways to arrange the people in line. There are also another $7!$ ways to arrange them if A and B are stuck together in the order BA. There are therefore $2\times7!$ ways of arranging the eight people in line with A and B next to each other.

However, there are $8!$ ways of arranging the eight people in line if there are no restrictions.

The required probability is therefore $\dfrac{2\times7!}{8!} = \frac{2}{8} = \frac{1}{4}$.

(b) $\text{P}(A \text{ and } B \text{ are not together}) = 1 - \text{P}(A \text{ and } B \text{ are together})$,

so $\text{P}(A \text{ and } B \text{ are not together}) = 1 - \frac{1}{4} = \frac{3}{4}$.

If you need to find the probability that two objects are not together, it is usually a good idea to find first the probability that they are together.

In this section emphasis is placed on the fact that the objects being arranged have to be distinct. That is, you have to be able to identify each object uniquely. The next section shows how to tackle the permutation problem when the objects are not distinct.

5.2 Permutations when the objects are not distinct

Recall that there were $4! = 24$ permutations of the letters A, B, C and D. These were listed earlier in Fig. 5.2. Imagine that the A remains but that the letters B, C and D are all replaced by the letter Z. How many permutations will there be now?

Suppose that B is temporarily replaced by Z_1, C is replaced by Z_2 and D is replaced by Z_3 so that you can tell the Zs apart. Fig. 5.4 shows the permutations of A, Z_1, Z_2 and Z_3 using a different procedure from that used to write out Fig. 5.2.

Write the first permutation of A, Z_1, Z_2 and Z_3 at the top of the first column. Any permutation can be the first permutation. Leave A in the same position, but write the other permutations of Z_1, Z_2 and Z_3 underneath. Write a permutation not already used at the top of the next column, and repeat writing the other permutations of Z_1, Z_2 and Z_3 underneath while keeping A in the same position. Keep going until you have written all the permutations of A, Z_1, Z_2 and Z_3.

$AZ_1Z_2Z_3$	$Z_1AZ_2Z_3$	$Z_1Z_2AZ_3$	$Z_1Z_2Z_3A$
$AZ_1Z_3Z_2$	$Z_1AZ_3Z_2$	$Z_1Z_3AZ_2$	$Z_1Z_3Z_2A$
$AZ_2Z_1Z_3$	$Z_2AZ_1Z_3$	$Z_2Z_1AZ_3$	$Z_2Z_1Z_3A$
$AZ_2Z_3Z_1$	$Z_2AZ_3Z_1$	$Z_2Z_3AZ_1$	$Z_2Z_3Z_1A$
$AZ_3Z_1Z_2$	$Z_3AZ_1Z_2$	$Z_3Z_1AZ_2$	$Z_3Z_1Z_2A$
$AZ_3Z_2Z_1$	$Z_3AZ_2Z_1$	$Z_3Z_2AZ_1$	$Z_3Z_2Z_1A$

Fig. 5.4. All permutations of the letters A, Z_1, Z_2 and Z_3.

There are 4! arrangements in Fig. 5.4. Each column has all the permutations of Z_1, Z_2 and Z_3, 3! in all, so there must be $\frac{4!}{3!} = 4$ columns altogether. Now replace Z_1, Z_2 and Z_3 by Z and you have the permutations of A, Z, Z and Z in the top row. These are

$$AZZZ, \quad ZAZZ, \quad ZZAZ, \quad ZZZA.$$

You can generalise this argument. Suppose that you have n objects and r of them are identical. Then the number of arrangements in the table equivalent to Fig. 5.4 will be $n!$. When you write down the permutations in the columns corresponding to the arrangement at the top of the column, you find that there are $r!$ of them, so the table will have $r!$ rows. The number of columns (that is, the number of distinct permutations) is therefore $\frac{n!}{r!}$.

This result also generalises.

> The number of distinct permutations of n objects, of which p are identical to each other, and then q of the remainder are identical, and r of the remainder are identical, and so on is
>
> $$\frac{n!}{p! \times q! \times r! \times \ldots}, \quad \text{where} \quad p + q + r + \ldots = n.$$

Example 5.2.1
Find the number of distinct permutations of the letters of the word *MISSISSIPPI*.

The number of letters is 11, of which there are 4 *S*s, 4 *I*s, 2 *P*s and 1 *M*. The number of distinct permutations of the letters is therefore

$$\frac{11!}{4! \times 4! \times 2! \times 1!} = 34\ 650\ .$$

Exercise 5A

1 Seven different cars are to be loaded on to a transporter truck. In how many different ways can the cars be arranged?

2 How many numbers are there between 1245 and 5421 inclusive which contain each of the digits 1, 2, 4 and 5 once and once only?

3 An artist is going to arrange five paintings in a row on a wall. In how many ways can this be done?

4 Ten athletes are running in a 100-metre race. In how many different ways can the first three places be filled?

5 By writing out all the possible arrangements of $D_1E_1E_2D_2$, show that there are $\frac{4!}{2!2!} = 6$ different arrangements of the letters of the word *DEED*.

6 A typist has five letters and five addressed envelopes. In how many different ways can the letters be placed in each envelope without getting every letter in the right envelope? If the letters are placed in the envelopes at random what is the probability that each letter is in its correct envelope?

7 How many different arrangements can be made of the letters in the word *STATISTICS*?

8 (a) Calculate the number of arrangements of the letters in the word *NUMBER*.

(b) How many of the arrangements in part (a) begin and end with a vowel?

9 How many different numbers can be formed by taking one, two, three and four digits from the digits 1, 2, 7 and 8, if repetitions are not allowed?

One of these numbers is chosen at random. What is the probability that it is greater than 200?

5.3 Combinations

In the last section you considered permutations (arrangements), for which the order of the objects is significant when you count the number of different possibilities. In some circumstances, however, the order of selection does not matter. For example, if you were dealt a hand of 13 cards from a standard pack of 52 playing cards, you would not be interested in the order in which you received the cards. When a selection is made from a set of objects and the order of selection is unimportant it is called a **combination**.

To see the difference between combinations and permutations consider what happens when you select three letters from the four letters A, B, C and D.

Here is a procedure for finding all the combinations. It starts by considering permutations, and gives you a method of counting the combinations.

Start with any permutation of three letters from A, B, C and D, and write it at the top of the first column. Write the other permutations of the same three letters underneath it. Write a permutation not already used at the top of the next column, and write the other permutations of the letters underneath. Keep on until you have used all the permutations of three letters from A, B, C and D.

The results are shown in Fig. 5.5.

ABC	*ABD*	*ACD*	*BCD*
ACB	*ADB*	*ADC*	*BDC*
BAC	*BAD*	*CAD*	*CBD*
BCA	*BDA*	*CDA*	*CDB*
CAB	*DAB*	*DAC*	*DBC*
CBA	*DBA*	*DCA*	*DCB*

Fig. 5.5. Procedure for finding the number of combinations.

Each column then corresponds to a single combination because the elements in any one column differ only in the order in which the letters are written. The permutations are all different, but they all give rise to the same combination at the head of the column. To count the combinations, it is sufficient to count the columns.

There are ${}^4P_3$ permutations of 3 objects from 4 objects, so there are ${}^4P_3$ elements in total in Fig. 5.5.

Each column has $3!$ elements, so, by dividing, you find that there must be $\frac{{}^4P_3}{3!}$ columns, which means $\frac{{}^4P_3}{3!}$ combinations. As ${}^4P_3 = \frac{4!}{(4-3)!}$,

$$\frac{{}^4P_3}{3!} = \frac{4!}{(4-3)! \times 3!} = \frac{4!}{1! \times 3!} = \frac{4 \times 3 \times 2 \times 1}{1 \times (3 \times 2 \times 1)} = 4.$$

So there are 4 combinations of three letters from the four letters A, B, C and D.

You can apply this reasoning and this calculation to finding the number of combinations of r objects taken from n objects. In the table which corresponds to Fig. 5.5, there would be ${}^nP_r$ elements in total, and each column would have $r!$ elements. There would therefore be $\frac{{}^nP_r}{r!}$ columns, which corresponds to $\frac{{}^nP_r}{r!}$ combinations of r objects taken from n objects.

Writing ${}^{n}P_r$ in factorials as $\frac{n!}{(n-r)!}$ leads to a simpler expression to remember: the number of combinations of r objects taken from n objects is $\frac{n!}{(n-r)! \times r!}$.

The number of combinations of r objects chosen from n distinct objects is given the symbol $\binom{n}{r}$, which is often read as 'n choose r'. The older symbols, ${}_nC_r$ and ${}^nC_r$ are also used, and your calculator probably uses one of them.

> A **combination** is a selection in which the order of the objects selected is unimportant.
>
> The number of different combinations of r objects selected from n distinct objects is $\binom{n}{r}$, where $\binom{n}{r} = \frac{n!}{(n-r)! \times r!}$.

Example 5.3.1

The manager of a football team has a squad of 16 players. He needs to choose 11 to play in a match. How many possible teams can be chosen?

This example is not entirely realistic because players will not be equally capable of playing in every position, but it does show how many possible teams there are. It is important to decide whether this question is about permutations or combinations. Clearly the important issue here is the people in the team and not their order of selection. Therefore this question is about combinations rather than permutations.

The number of teams is $\binom{16}{11} = \frac{16!}{(16-11)! \times 11!} = \frac{16!}{5! \times 11!} = 4368$.

The number of teams is surprisingly large.

You may notice in Example 5.3.1 that if you had chosen the 5 players to drop out of the squad of 16 players, you would in effect be selecting the 11 by another method. You can select the 5 players in $\binom{16}{5}$ ways, and

$$\binom{16}{5} = \frac{16!}{(16-5)! \times 5!} = \frac{16!}{11! \times 5!},$$

which is clearly equal to $\binom{16}{11}$.

When you come to calculate a number like $\binom{16}{11}$ or $\binom{16}{5}$, you can take a short cut. Since

$$16! = 16 \times 15 \times \ldots \times 12 \times 11 \times 10 \times \ldots \times 2 \times 1,$$

you can cancel the $11 \times 10 \times \ldots \times 2 \times 1$ in the numerator with the $11!$ in the denominator.

Therefore you can write down immediately that

$$\binom{16}{5} = \frac{16 \times 15 \times \ldots \times 12}{5!},$$

where you need to make sure that you multiply 5 numbers in the numerator if the denominator is $5!$.

In general,

$$\binom{n}{r} = \frac{\overbrace{n \times (n-1) \times \ldots \times (n-r+1)}^{r \text{ factors}}}{r!}.$$

Example 5.3.2
A team of 5 people, which must contain 3 men and 2 women, is chosen from 8 men and 7 women. How many different teams can be selected?

The number of different teams of 3 men which can be selected from 8 is $\binom{8}{3}$.

The number of different teams of 2 women which can be selected from 7 is $\binom{7}{2}$.

Any of the $\binom{8}{3}$ men's teams can join up with any of the $\binom{7}{2}$ women's teams to make an acceptable team of 5. Therefore you need to multiply these two quantities together to find the number of different teams possible.

The number of possible teams is

$$\binom{8}{3} \times \binom{7}{2} = \frac{8 \times 7 \times 6}{1 \times 2 \times 3} \times \frac{7 \times 6}{1 \times 2} = 56 \times 21 = 1176.$$

You can now apply some of these counting methods to probability examples.

Example 5.3.3
Five cards are dealt without replacement from a standard pack of 52 cards. Find the probability that exactly 3 of the 5 cards are hearts.

The sample space is very large. It would consist of a list of all possible sets of 5 cards which you could choose from the 52 cards in the pack. You do not need such a list, however. All that you need to know is how many different sets of cards the sample space contains. You are choosing 5 objects from 52, so the number of unrestricted choices is $\binom{52}{5}$, because the order of selection is irrelevant.

Let A be the event that exactly 3 cards of the 5 dealt out are hearts. The method used to find the number of outcomes in the event A is very similar to the technique used in Example 5.3.2.

There are $\binom{13}{3}$ 'teams' of 3 hearts.

There are $\binom{39}{2}$ 'teams' of 2 'non-hearts'.

Therefore the number of sets of 5 cards with exactly 3 hearts is $\binom{13}{3}\times\binom{39}{2}$

The probability that event A happens is $\dfrac{\binom{13}{3}\times\binom{39}{2}}{\binom{52}{5}}=0.0815$, correct to 3 significant figures.

Exercise 5B

1 How many different three-card hands can be dealt from a pack of 52 cards?

2 From a group of 30 boys and 32 girls, two girls and two boys are to be chosen to represent their school. How many possible selections are there?

3 A history exam paper contains eight questions, four in Part A and four in Part B. Candidates are required to attempt five questions. In how many ways can this be done if

(a) there are no restrictions,

(b) at least two questions from Part A and at least two questions from Part B must be attempted?

4 A committee of three people is to be selected from four women and five men. The rules state that there must be at least one man and one woman on the committee. In how many different ways can the committee be chosen?

Subsequently one of the men and one of the women marry each other. The rules also state that a married couple may not both serve on the committee. In how many ways can the committee be chosen now?

5 A box of one dozen eggs contains one that is bad. If three eggs are chosen at random, what is the probability that one of them will be bad?

6 In a game of bridge the pack of 52 cards is shared equally between all four players. What is the probability that one particular player has no hearts?

7 A bag contains 20 chocolates, 15 toffees and 12 peppermints. If three sweets are chosen at random, what is the probability that they are

(a) all different, (b) all chocolates,

(c) all the same, (d) all not chocolates?

8 Show that $\binom{n}{r}=\binom{n}{n-r}$.

9 Show that the number of permutations of n objects of which r are of one kind and $n-r$ are of another kind is $\binom{n}{r}$.

5.4 Applications of permutations and combinations

In Example 5.1.2 you were asked to find the number of ways that eight people could stand in a line when two people had to stand next to each other. This was an example in which you were asked to find the number of permutations or combinations of a set of objects with some extra condition included. This section will show you how to answer such questions.

Example 5.4.1
Find the number of ways of arranging 6 women and 3 men to stand in a row so that all 3 men are standing together.

You can make this problem simpler by thinking of the 3 men as a single unit. Imagine tying them together for example! You would then have 7 items (or units), the 6 individual women and the block of 3 men.

So one possible arrangement would be

$$W_1 \quad W_3 \quad W_5 \quad W_6 \quad W_2 \quad W_4 \quad M_1M_3M_2,$$

where W_3, for example, represents the third woman.

The number of permutations of these 7 units is $7!$. However, for each of these permutations the men could be arranged (inside the rope) in $3!$ different ways. Therefore the total number of permutations in which the 6 women and 3 men can be arranged so that the 3 men are standing together is $7! \times 3! = 30\,240$.

Example 5.4.2
Find the number of ways of arranging 6 women and 3 men in a row so that no two men are standing next to one another.

You can ensure that no two men stand next to one another in the following way.

Arrange the 6 women to stand in a line with a space between each pair of them and two extra spaces, one at each end of the line. One such arrangement is

Space 1		Space 2		Space 3		Space 4		Space 5		Space 6		Space 7
↓	W_1	↓	W_2	↓	W_5	↓	W_4	↓	W_6	↓	W_3	↓

There are $6!$ arrangements of the 6 women. For any of these $6!$ arrangements you can now pick a space in which to place the first man M_1. This can be done in 7 ways.

Here is the arrangement above with one of the men, M_1, placed in Space 2.

Space 1				Space 3		Space 4		Space 5		Space 6		Space 7
↓	W_1	M_1	W_2	↓	W_5	↓	W_4	↓	W_6	↓	W_3	↓

By using a similar argument you can see that there will be 6 choices for the position of M_2 and 5 choices for the position of M_3. Once the 3 men have been placed the remaining spaces can be 'closed up' or simply ignored. By using this method you can guarantee that no two men can stand next to one another. Also all possible arrangements will be counted using this method.

Therefore the number of permutations in which no two men stand next to one another is $6! \times 7 \times 6 \times 5 = 151\,200$.

It should be clear that you multiply $6!$ *by* 7, 6 *and* 5 *because for every one of the* $6!$ *arrangements of the women there will be* 7 *spaces to choose for* M_1, *and then* 6 *places to choose for* M_2, *and then* 5 *places to choose for* M_3.

It is also worth noting that the answers to Examples 5.4.1 and 5.4.2 when added together do not give $9!$, *which is the total number of arrangements of* 9 *people without any restriction at all. This is because there is a third possibility. If two men were standing together and the third man was separated from these two by some women, then it would not be the case that all the men were together but neither would it be the case that the three men were all apart from one another.*

Example 5.4.3

A group of 12 people consisting of 6 married couples is arranged at random in a line for a photograph. Find the probability that each wife is standing next to her husband.

The number of unrestricted arrangements is $12!$. Each of them is equally likely.

If each husband and wife 'couple' is to stand together, then you can consider each couple as a unit. There are therefore 6 such units.

The number of permutations of these units is $6!$.

But the first couple H_1W_1 can be arranged in $2!$ ways, either H_1W_1 or W_1H_1. This applies equally to couples 2, 3, 4, 5 and 6. Therefore the number of arrangements in which each couple stands together is $6! \times (2!)^6$.

Hence $\text{P(each couple stands together)} = \dfrac{6! \times (2!)^6}{12!} = \dfrac{1}{10\,395} = 9.62 \times 10^{-5}$, correct to 3 significant figures.

Example 5.4.4

Four letters are to be selected from the letters in the word *RIGIDITY*. How many different combinations are there?

The problem here is that the letters are not all distinct since there are three *I*s. Therefore the answer is not $\binom{8}{4}$. In order to deal with this problem a useful strategy is to split it into different cases depending on the number of *I*s chosen.

Case 1 Combinations with no *I*s

In this case you are selecting 4 letters from the 5 letters R, G, D, T, Y, so the number of combinations is $\binom{5}{4} = 5$.

These are the combinations with no *I*. Remember that since this is a problem about combinations you are not interested in the order. All that matters here is which 'team' of letters you choose. The possible teams are

RGDT *RGDY* *RGTY* *RDTY* *GDTY*.

Case 2 Combinations with one I

In this case you are selecting 3 letters from the 5 letters R, G, D, T, Y, together with one I, so the number of combinations is $\binom{5}{3} = 10$. Here are the combinations with one I:

RGDI RGTI RGYI RDTI RDYI
RTYI GDTI GDYI GTYI DTYI.

Case 3 Combinations with two Is

In this case you are selecting 2 letters from the 5 letters R, G, D, T, Y, together with two Is, so the number of combinations is $\binom{5}{2} = 10$. Here they are:

RGII RDII RTII RYII GDII
GTII GYII DTII DYII TYII.

Case 4 Combinations with three Is.

In this case you are selecting 1 letter from the 5 letters R, G, D, T, Y, together with three Is, so the number of combinations is $\binom{5}{1} = 5$. Here they are:

RIII GIII DIII TIII YIII .

The total number of distinct combinations of four letters selected from the letters of the word *RIGIDITY* is $5 + 10 + 10 + 5 = 30$. All 30 combinations have been listed above.

Exercise 5C

1 The letters of the word *CONSTANTINOPLE* are written on 14 cards, one on each card. The cards are shuffled and then arranged in a straight line.

(a) How many different possible arrangements are there?

(b) How many arrangements begin with P?

(c) How many arrangements start and end with O?

(d) How many arrangements are there where no two vowels are next to each other?

2 A coin is tossed 10 times.

(a) How many different sequences of heads and tails are possible?

(b) How many different sequences containing six heads and four tails are possible?

(c) What is the probability of getting six heads and four tails?

3 Eight cards are selected with replacement from a standard pack of 52 playing cards, with 12 picture cards, 20 odd cards and 20 even cards.

(a) How many different sequences of eight cards are possible?

(b) How many of the sequences in part (a) will contain three picture cards, three odd-numbered cards and two even-numbered cards?

(c) Use parts (a) and (b) to determine the probability of getting three picture cards, three odd-numbered cards and two even-numbered cards if eight cards are selected with replacement from a standard pack of 52 playing cards.

4 Eight women and five men are standing in a line.

(a) How many arrangements are possible if any individual can stand in any position?

(b) In how many arrangements will all five men be standing next to one another?

(c) In how many arrangements will no two men be standing next to one another?

5 Each of the digits 1, 1, 2, 3, 3, 4, 6 is written on a separate card. The seven cards are then laid out in a row to form a 7-digit number.

(a) How many distinct 7-digit numbers are there?

(b) How many of these 7-digit numbers are even?

(c) How many of these 7-digit numbers are divisible by 4?

(d) How many of these 7-digit numbers start and end with the same digit?

6 Three families, the Mehtas, the Mupondas and the Lams, go to the cinema together to watch a film. Mr and Mrs Mehta take their daughter Indira, Mr and Mrs Muponda take their sons Paul and John, and Mrs Lam takes her children Susi, Kim and Lee. The families occupy a single row with eleven seats.

(a) In how many ways could the eleven people be seated if there were no restriction?

(b) In how many ways could the eleven people sit down so that the members of each family are all sitting together?

(c) In how many of the arrangements will no two adults be sitting next to one another?

7 The letters of the word *POSSESSES* are written on nine cards, one on each card. The cards are shuffled and four of them are selected and arranged in a straight line.

(a) How many possible selections are there of four letters?

(b) How many arrangements are there of four letters?

Miscellaneous exercise 5

1 The judges in a 'Beautiful Baby' competition have to arrange 10 babies in order of merit. In how many different ways could this be done? Two babies are to be selected to be photographed. In how many ways can this selection be made?

2 In how many ways can a committee of four men and four women be seated in a row if

(a) they can sit in any position,

(b) no one is seated next to a person of the same sex?

3 How many distinct arrangements are there of the letters in the word *ABRACADABRA*?

4 Six people are going to travel in a six-seater minibus but only three of them can drive. In how many different ways can they seat themselves?

5 There are eight different books on a bookshelf: three of them are hardbacks and the rest are paperbacks.

(a) In how many different ways can the books be arranged if all the paperbacks are together and all the hardbacks are together?

(b) In how many different ways can the books be arranged if all the paperbacks are together?

6 Four boys and two girls sit in a line on stools in front of a coffee bar.

(a) In how many ways can they arrange themselves so that the two girls are together?

(b) In how many ways can they sit if the two girls are not together? (OCR)

7 Ten people travel in two cars, a saloon and a Mini. If the saloon has seats for six and the Mini has seats for four, find the number of different ways in which the party can travel, assuming that the order of seating in each car does not matter and all the people can drive. (OCR)

8 Giving a brief explanation of your method, calculate the number of different ways in which the letters of the word *TRIANGLES* can be arranged if no two vowels may come together. (OCR)

9 I have seven fruit bars to last the week. Two are apricot, three fig and two peach. I select one bar each day. In how many different orders can I eat the bars?

If I select a fruit bar at random each day, what is the probability that I eat the two apricot ones on consecutive days?

10 A class contains 30 children, 18 girls and 12 boys. Four complimentary theatre tickets are distributed at random to the children in the class. What is the probability that

(a) all four tickets go to girls,

(b) two boys and two girls receive tickets? (OCR)

11 (a) How many different 7-digit numbers can be formed from the digits 0, 1, 2, 2, 3, 3, 3 assuming that a number cannot start with 0 ?

(b) How many of these numbers will end in 0? (OCR)

12 Calculate the number of ways in which three girls and four boys can be seated on a row of seven chairs if each arrangement is to be symmetrical. (OCR)

13 Find the number of ways in which

(a) 3 people can be arranged in 4 seats,

(b) 5 people can be arranged in 5 seats.

In a block of 8 seats, 4 are in row A and 4 are in row B. Find the number of ways of arranging 8 people in the 8 seats given that 3 specified people must be in row A. (OCR)

14 Eight different cards, of which four are red and four are black, are dealt to two players so that each receives a hand of four cards.

Calculate

(a) the total number of different hands which a given player could receive,

(b) the probability that each player receives a hand consisting of four cards all of the same colour. (OCR)

15 A piece of wood of length 10 cm is to be divided into 3 pieces so that the length of each piece is a whole number of cm, for example 2 cm, 3 cm and 5 cm.

(a) List all the different sets of lengths which could be obtained.

(b) If one of these sets is selected at random, what is the probability that the lengths of the pieces could be lengths of the sides of a triangle? (OCR)

16 Nine persons are to be seated at three tables holding 2, 3 and 4 persons respectively. In how many ways can the groups sitting at the tables be selected, assuming that the order of sitting at the tables does not matter? (OCR)

17 (a) Calculate the number of different arrangements which can be made using all the letters of the word *BANANA*.

(b) The number of combinations of 2 objects from n is equal to the number of combinations of 3 objects from n. Determine n. (OCR)

18 A 'hand' of 5 cards is dealt from an ordinary pack of 52 playing cards. Show that there are nearly 2.6 million distinct hands and that, of these, 575 757 contain no card from the heart suit.

On three successive occasions a card player is dealt a hand containing no heart. What is the probability of this happening? What conclusion might the player justifiably reach?

(OCR)

19* Notice that $7! \times 6! = 10!$. Find three integers, m, n and r, where $r > 10$, for which $m! \times n! = r!$.

6 Probability distributions

This chapter introduces the idea of a discrete random variable. When you have completed it, you should

- understand what a discrete random variable is
- know the properties of a discrete random variable
- be able to construct a probability distribution table for a discrete random variable.

6.1 Discrete random variables

Most people have played board games at some time. Here is an example.

Game *A* A turn consists of throwing a dice and then moving a number of squares equal to the score on the dice.

'The number of squares moved in a turn' is a variable because it can take different values, namely $1, 2, 3, 4, 5$ and 6. However, the value taken at any one turn cannot be predicted, but depends on chance. For these reasons 'the number of squares moved in a turn' is called a 'random variable'.

A **random variable** is a quantity whose value depends on chance.

The 'number of squares moved in a turn' is a **discrete** random variable because there are clear steps between the different possible values it can take.

Although you cannot predict the result of the next throw of the dice, you do know that, if the dice is fair, the probability of getting each value is $\frac{1}{6}$. A convenient way of expressing this information is to let X stand for 'the number of squares moved in a turn'. Then, for example, $\mathrm{P}(X=3)=\frac{1}{6}$ means 'the probability that X takes the value 3 is $\frac{1}{6}$'. Generalising, $\mathrm{P}(X=x)$ means 'the probability that the variable X takes the value x'.

Note how the capital letter stands for the variable itself and the small letter stands for the value which the variable takes.

This notation is used in Table 6.1 to give the possible values for the number of squares moved and the probability of each value. This table is called the 'probability distribution' of X.

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

Table 6.1. Probability distribution of X, the number of squares moved in a turn for a single throw of a dice.

The **probability distribution** of a discrete random variable is a listing of the possible values of the variable and the corresponding probabilities.

In some board games, a dice is used in a more complicated way in order to decide how many squares a person should move. Here are two different examples.

Game *B* A person is allowed a second throw of the dice if a 6 is thrown, and, in this case, moves a number Y of squares equal to the sum of the two scores obtained.

Game *C* The dice is thrown twice and the number, W, of squares moved is the sum of the two scores.

Fig. 6.2 is a tree diagram illustrating Game B. As Y is the number of squares moved in a turn, it can take the values 1, 2, 3, 4, 5, 7, 8, 9, 10, 11 and 12. The probability of each of the first five values is $\frac{1}{6}$, as in the previous game. In order to score 7, you have to score 6 followed by 1. Since the two events are independent, the probability of scoring a 6 followed by a 1 is found by multiplying the two probabilities:

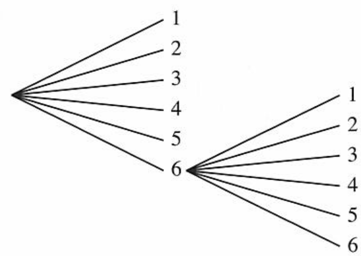


Fig. 6.2. Tree diagram for Game B.

$$\begin{aligned} P(Y = 7) &= P(6 \text{ on first throw}) \times P(1 \text{ on second throw}) \\ &= \tfrac{1}{6} \times \tfrac{1}{6} = \tfrac{1}{36}. \end{aligned}$$

The probability that Y takes each of the values 8, 9, 10, 11 and 12 will also be $\frac{1}{36}$. Table 6.3 gives the probability distribution of Y.

y	1	2	3	4	5	7	8	9	10	11	12	Total
$P(Y = y)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	1

Table 6.3. Probability distribution of Y, the number of squares moved in Game B.

The possible values of W in Game C can be found by constructing a table as shown in Fig. 6.4.

		First throw					
		1	2	3	4	5	6
	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
Second	3	4	5	6	7	8	9
throw	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Fig. 6.4. Possible total scores when two individual scores are added in Game C.

There are 36 outcomes in the table and they are all equally likely, so, for example, $P(W=6)=\frac{5}{36}$ and $P(W=7)=\frac{6}{36}$. Table 6.5 gives the probability distribution of W. The fractions could have been cancelled but in their present forms it is easier to see the shape of the distribution.

w	2	3	4	5	6	7	8	9	10	11	12	Total
$P(W=w)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

Table 6.5. Probability distribution of W, the number of squares moved in Game C.

Fig. 6.6 allows you to compare the probability distributions of X, Y and W.

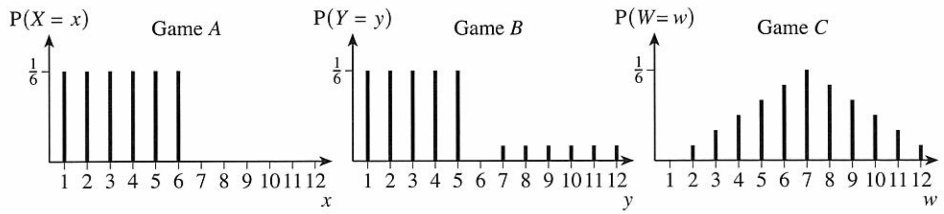


Fig. 6.6. Comparing the probability distributions of Games A, B and C.

Looking at Fig. 6.6, which method of scoring will take you round the board most quickly and which most slowly? (A method for finding the answer to this question by calculation is given in Example 8.1.1.)

Examples 6.1.1 and 6.1.2 illustrate some other probability distributions.

Example 6.1.1

A bag contains two red and three blue marbles. Two marbles are selected at random without replacement and the number, X, of blue marbles is counted. Find the probability distribution of X.

Fig. 6.7 is a tree diagram illustrating this situation. R_1 denotes the event that the first marble is red and R_2 the event that the second marble is red. Similarly B_1 and B_2 stand for the events that the first and second marbles respectively are blue. X can take the values 0, 1 and 2.

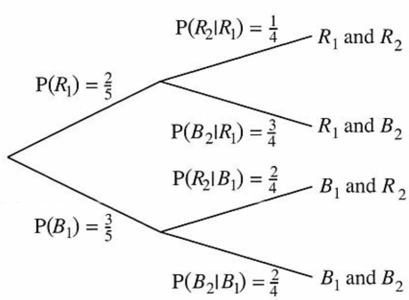


Fig. 6.7. Tree diagram for Example 6.1.1.

$$\begin{aligned}P(X=0) &= P(R_1 \text{ and } R_2)\\ &= P(R_1)\times P(R_2 \mid R_1)\\ &= \tfrac{2}{5}\times\tfrac{1}{4}=\tfrac{2}{20}=\tfrac{1}{10}.\end{aligned}$$

$$\begin{aligned}P(X=1) &= P(B_1 \text{ and } R_2)+P(R_1 \text{ and } B_2)\\ &= P(B_1)\times P(R_2 \mid B_1)+P(R_1)\times P(B_2 \mid R_1)\\ &= \tfrac{3}{5}\times\tfrac{2}{4}+\tfrac{2}{5}\times\tfrac{3}{4}=\tfrac{12}{20}=\tfrac{3}{5}.\end{aligned}$$

$$\begin{aligned}P(X=2) &= P(B_1 \text{ and } B_2) \\ &= P(B_1) \times P(B_2 \mid B_1) \\ &= \tfrac{3}{5} \times \tfrac{2}{4} = \tfrac{6}{20} = \tfrac{3}{10}.\end{aligned}$$

Here is the probability distribution of X.

x	0	1	2	Total
$P(X=x)$	$\frac{1}{10}$	$\frac{3}{5}$	$\frac{3}{10}$	1

Example 6.1.2

A random variable, X, has the probability distribution shown below.

x	1	2	3	4
$P(X=x)$	0.1	0.2	0.3	0.4

Two observations are made of X and the random variable Y is equal to the larger minus the smaller; if the two observations are equal, Y takes the value 0. Find the probability distribution of Y. Which value of Y is most likely?

The following table gives the values for X and the corresponding value of Y. Since the two observations of X are independent, you find the probability of each pair by multiplying the probabilities for the two X values, as shown in the last column of the table.

First value of X	Second value of X	Y	Probability
1	1	0	$0.1 \times 0.1 = 0.01$
1	2	1	$0.1 \times 0.2 = 0.02$
1	3	2	$0.1 \times 0.3 = 0.03$
1	4	3	$0.1 \times 0.4 = 0.04$
2	1	1	$0.2 \times 0.1 = 0.02$
2	2	0	$0.2 \times 0.2 = 0.04$
2	3	1	$0.2 \times 0.3 = 0.06$
2	4	2	$0.2 \times 0.4 = 0.08$
3	1	2	$0.3 \times 0.1 = 0.03$
3	2	1	$0.3 \times 0.2 = 0.06$
3	3	0	$0.3 \times 0.3 = 0.09$
3	4	1	$0.3 \times 0.4 = 0.12$
4	1	3	$0.4 \times 0.1 = 0.04$
4	2	2	$0.4 \times 0.2 = 0.08$
4	3	1	$0.4 \times 0.3 = 0.12$
4	4	0	$0.4 \times 0.4 = 0.16$

This table shows that Y takes the values 0, 1, 2 and 3. You can find the total probability for each value of Y by adding the individual probabilities:

$$P(Y=0) = 0.01 + 0.04 + 0.09 + 0.16 = 0.30,$$
$$P(Y=1) = 0.02 + 0.02 + 0.06 + 0.06 + 0.12 + 0.12 = 0.40,$$
$$P(Y=2) = 0.03 + 0.08 + 0.03 + 0.08 = 0.22,$$
$$P(Y=3) = 0.04 + 0.04 = 0.08.$$

So Y has the probability distribution shown below.

y	0	1	2	3	Total
$P(Y=y)$	0.30	0.40	0.22	0.08	1

The most likely value of Y is 1, because it has the highest probability.

Exercise 6A

1 A fair coin is thrown four times. The random variable X is the number of heads obtained. Tabulate the probability distribution of X.

2 Two fair dice are thrown simultaneously. The random variable D is the difference between the smaller and the larger score, or zero if they are the same. Tabulate the probability distribution of D.

3 A fair dice is thrown once. The random variable X is related to the number N thrown on the dice as follows. If N is even, then X is half N; otherwise X is double N. Tabulate the probability distribution of X.

4 Two fair dice are thrown simultaneously. The random variable H is the highest common factor of the two scores. Tabulate the probability distribution of H, combining together all the possible ways of obtaining the same value.

5 When a four-sided dice is thrown, the score is the number on the bottom face. Two fair four-sided dice, each with faces numbered 1 to 4, are thrown simultaneously. The random variable M is the product of the two scores multiplied together. Tabulate the probability distribution of M, combining together all the possible ways of obtaining the same value.

6 A bag contains six red and three green counters. Two counters are drawn from the bag, without replacement. Tabulate the probability distribution of the number of green counters obtained.

7 Satish picks a card at random from an ordinary pack. If the card is an ace, he stops; if not, he continues to pick cards at random, without replacement, until either an ace is picked, or four cards have been drawn. The random variable C is the total number of cards drawn. Construct a tree diagram to illustrate the possible outcomes of the experiment, and use it to calculate the probability distribution of C.

8* Obtain the probability distribution of the total score when three cubical dice are thrown simultaneously.

9 If you have access to a spreadsheet, use it to draw diagrams to illustrate the probability distributions of the total number of heads obtained when $3, 4, 5, \ldots$ fair coins are thrown.

6.2 An important property of a probability distribution

You have probably noticed that, in all the probability distributions considered so far, the sum of the probabilities is 1. This must always be the case, because one of the outcomes must happen. This is an important property of a probability distribution and a useful check that you have found the probabilities correctly.

> For any random variable, X, the sum of the probabilities is 1; that is, $\sum P(X = x) = 1$.

Example 6.2.1
The table below gives the probability distribution of the random variable T.
Find (a) the value of c, (b) $P(T \leqslant 3)$, (c) $P(T > 3)$.

t	1	2	3	4	5
$P(T = t)$	c	$2c$	$2c$	$2c$	c

(a) Since the probabilities must sum to 1,

$$c + 2c + 2c + 2c + c = 1, \quad \text{so} \quad 8c = 1, \quad \text{giving} \quad c = \tfrac{1}{8}.$$

(b) $P(T \leqslant 3) = P(T = 1) + P(T = 2) + P(T = 3) = c + 2c + 2c = 5c = \tfrac{5}{8}$.

(c) $P(T > 3) = P(T = 4) + P(T = 5) = 2c + c = 3c = \tfrac{3}{8}$.

Example 6.2.2
A computer is programmed to give single-digit numbers X between 0 and 9 inclusive in such a way that the probability of getting an odd digit $(1, 3, 5, 7, 9)$ is half the probability of getting an even digit $(0, 2, 4, 6, 8)$. Find the probability distribution of X.

Let the probability of getting an even digit be c. Then the probability of getting an odd digit is $\tfrac{1}{2}c$.

Since the probabilities must sum to 1,

$$\sum P(X = x) = c + \tfrac{1}{2}c + c + \tfrac{1}{2}c + c + \tfrac{1}{2}c + c + \tfrac{1}{2}c + c + \tfrac{1}{2}c = 1,$$

which gives $\tfrac{15}{2}c = 1$; that is, $c = \tfrac{2}{15}$.

The probability distribution of X is $P(X = x) = \tfrac{1}{15}$ for $x = 1, 3, 5, 7$ and 9 and $P(X = x) = \tfrac{2}{15}$ for $x = 0, 2, 4, 6$ and 8.

Exercise 6B

1 In the following probability distribution, c is a constant. Find the value of c.

x	0	1	2	3
$\mathrm{P}(X=x)$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{2}$	c

2 In the following probability distribution, d is a constant. Find the value of d.

x	0	1	2	3
$\mathrm{P}(X=x)$	d	0.1	0.2	0.4

3 In the following probability distribution, d is a constant. Find the value of d.

x	0	1	2	3	4
$\mathrm{P}(X=x)$	d	0.2	0.15	0.2	$2d$

4 The score S on a spinner is a random variable with distribution given by $\mathrm{P}(S=s)=k$ $(s=1,2,3,4,5,6,7,8)$, where k is a constant. Find the value of k.

5 A cubical dice is biased so that the probability of an odd number is three times the probability of an even number. Find the probability distribution of the score.

6 A cubical dice is biased so that the probability of any particular score between 1 and 6 (inclusive) being obtained is proportional to that score. Find the probability of scoring a 1.

7 For a biased cubical dice the probability of any particular score between 1 and 6 (inclusive) being obtained is inversely proportional to that score. Find the probability of scoring a 1.

8 In the following probability distribution, c is a constant. Find the value of c.

x	0	1	2	3
$\mathrm{P}(X=x)$	0.6	0.16	c	c^2

6.3 Using a probability distribution as a model

So far, the discussion of probability distributions in this chapter has been very mathematical. At this point it may be helpful to point out the practical application of probability distributions. Probability distributions are useful because they provide models for experiments. Consider again the random variable X, the score on a dice, whose probability distribution was given in Table 6.1. Suppose you actually threw a dice 360 times. Since the values 1, 2, 3, 4, 5 and 6 have equal probabilities, you would expect them to occur with approximately equal frequencies of, in this case, $\frac{1}{6}\times 360=60$. It is very unlikely that all the observed frequencies will be exactly equal to 60. However, if the model is a suitable one, the observed frequencies should be close to the expected values.

What conclusion would you draw about the dice if the observed frequencies were not close to the expected values?

Now look at the random variable Y, whose probability distribution was given in Table 6.3. For this variable the values are not equally likely and so you would not expect to observe approximately equal frequencies. In Section 4.1 you met the idea that

$$\text{relative frequency} = \frac{\text{frequency}}{\text{total frequency}} \approx \text{probability}.$$

You can rearrange this equation to give an expression for the frequencies you would expect to observe:

$$\text{frequency} \approx \text{total frequency} \times \text{probability}.$$

For 360 observations of Y, the expected frequencies will be about $360 \times \frac{1}{6} = 60$ for $y = 1, 2, 3, 4, 5$ and $360 \times \frac{1}{36} = 10$ for $y = 7, 8, 9, 10, 11$ and 12.

What will the expected frequencies be for 360 *observations of the random variable* W*, whose probability distribution is given in Table 6.5?*

Exercise 6C

1 A card is chosen at random from a pack and replaced. This experiment is carried out 520 times. State the expected number of times on which the card is

(a) a club,
(b) an ace,
(c) a picture card (K, Q, J)
(d) either an ace or a club or both,
(e) neither an ace nor a club.

2 The biased dice of Exercise 6B Question 5 is rolled 420 times. State how many times you would expect to obtain

(a) a one,
(b) an even number,
(c) a prime number.

3 The table below gives the cumulative probability distribution for a random variable R. 'Cumulative' means that the probability given is $P(R \leqslant r)$, not $P(R = r)$.

r	0	1	2	3	4	5
$P(R \leqslant r)$	0.116	0.428	0.765	0.946	0.995	1.000

One hundred observations of R are made. Calculate the expected frequencies of each outcome, giving each answer to the nearest whole number.

4* A random variable G has a probability distribution given by the following formulae:

$$P(G = g) = \begin{cases} 0.3 \times (0.7)^g & \text{for } g = 1, 2, 3, 4, \\ k & \text{for } g = 5, \\ 0 & \text{for all other values of } g. \end{cases}$$

Find the value of k, and find the expected frequency of the result $G = 3$ when 1000 independent observations of G are made.

Miscellaneous exercise 6

1 Three cards are selected at random, without replacement, from a shuffled pack of 52 playing cards. Using a tree diagram, find the probability distribution of the number of honours (A, K, Q, J, 10) obtained.

2 An electronic device produces an output of 0, 1 or 3 volts, each time it is operated, with probabilities $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{6}$ respectively. The random variable X denotes the result of adding the outputs for two such devices which act independently.

(a) Tabulate the possible values of X with their corresponding probabilities.

(b) In 360 independent operations of the device, state on how many occasions you would expect the outcome to be 1 volt. (OCR, adapted)

3 The probabilities of the scores on a biased dice are shown in the table below.

Score	1	2	3	4	5	6
Probability	k	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{2}$

(a) Find the value of k.

Two players, Hazel and Ross, play a game with this biased dice and a fair dice. Hazel chooses one of the two dice at random and rolls it. If the score is 5 or 6 she wins a point.

(b) Calculate the probability that Hazel wins a point.

(c) Hazel chooses a dice, rolls it and wins a point. Find the probability that she chose the biased dice. (OCR)

4 In an experiment, a fair cubical dice was rolled repeatedly until a six resulted, and the number of rolls was recorded. The experiment was conducted 60 times.

(a) Show that you would expect to get a six on the first roll ten times out of the 60 repetitions of the experiment.

(b) Find the expected frequency for two rolls correct to one decimal place.

(OCR, adapted)

5 The probability distribution of the random variable Y is given in the following table, where c is a constant.

y	1	2	3	4	5
$P(Y = y)$	c	$3c$	c^2	c^2	$\frac{15}{32}$

Prove that there is only one possible value of c, and state this value.

7 The binomial distribution

This chapter introduces you to a discrete probability distribution called the binomial distribution. When you have completed it, you should

- know the conditions necessary for a random variable to have a binomial distribution
- be able to calculate probabilities for a binomial distribution
- know what the parameters of a binomial distribution are.

7.1 The binomial distribution

The spinner in Fig. 7.1 is an equilateral triangle. When it is spun it comes to rest on one of its three edges. Two of the edges are white and one is black. In Fig. 7.1 the spinner is resting on the black edge. This will be described as 'showing black'.

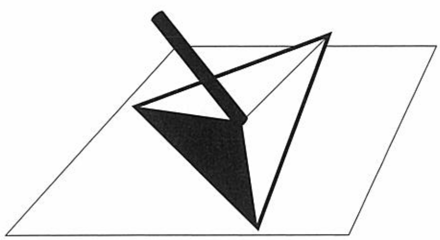


Fig. 7.1. A triangular spinner.

The spinner is fair, so the probability that the spinner shows black is $\frac{1}{3}$ and the probability that it shows white is $\frac{2}{3}$.

Suppose now that the spinner is spun on 5 separate occasions. Let the random variable X be the number of times out of 5 that the spinner shows black.

To derive the probability distribution of X, it is helpful to define some terms. The act of spinning the spinner once is called a **trial**. A simple way of describing the result of each trial is to call it a **success** (s) when the spinner shows black, and a **failure** (f) when the spinner shows white. So X could now be defined as the number of successes in the 5 trials.

The event $\{X=0\}$ would mean that the spinner did not show black on any of its 5 spins. The notation f_1 will be used to mean that the first trial resulted in a failure, f_2 will mean that the second trial resulted in a failure, and so on, giving

$$\mathrm{P}(X=0)=\mathrm{P}(\text{there are 5 failures})=\mathrm{P}(f_1 f_2 f_3 f_4 f_5).$$

Since the outcomes of the trials are independent,

$$\begin{aligned}\mathrm{P}(f_1 f_2 f_3 f_4 f_5)&=\mathrm{P}(f_1)\times\mathrm{P}(f_2)\times\mathrm{P}(f_3)\times\mathrm{P}(f_4)\times\mathrm{P}(f_5)\\&=\tfrac{2}{3}\times\tfrac{2}{3}\times\tfrac{2}{3}\times\tfrac{2}{3}\times\tfrac{2}{3}=\tfrac{32}{243}.\end{aligned}$$

The probability $\mathrm{P}(X=1)$ is more complicated to calculate. The event $\{X=1\}$ means that there is one success and also four failures. One possible sequence of a success and four failures is $s_1 f_2 f_3 f_4 f_5$ (where s_1 denotes the event that the first trial was a success).

The probability that the first trial is a success and the other four trials are failures is

$$\begin{aligned}P(s_1f_2f_3f_4f_5) &= P(s_1)\times P(f_2)\times P(f_3)\times P(f_4)\times P(f_5)\\ &= \tfrac{1}{3}\times\tfrac{2}{3}\times\tfrac{2}{3}\times\tfrac{2}{3}\times\tfrac{2}{3}\\ &= \left(\tfrac{1}{3}\right)\left(\tfrac{2}{3}\right)^4 = \tfrac{16}{243}.\end{aligned}$$

However there are four other possible sequences for the event $\{X=1\}$. They are

$$f_1s_2f_3f_4f_5 \quad f_1f_2s_3f_4f_5 \quad f_1f_2f_3s_4f_5 \quad f_1f_2f_3f_4s_5.$$

Therefore

$$\begin{aligned}P(X=1) &= P(s_1f_2f_3f_4f_5)+P(f_1s_2f_3f_4f_5)+P(f_1f_2s_3f_4f_5)\\ &\qquad +P(f_1f_2f_3s_4f_5)+P(f_1f_2f_3f_4s_5)\\ &= \left(\tfrac{1}{3}\right)\times\left(\tfrac{2}{3}\right)^4 + \left(\tfrac{2}{3}\right)\times\left(\tfrac{1}{3}\right)\times\left(\tfrac{2}{3}\right)^3 + \left(\tfrac{2}{3}\right)^2\times\left(\tfrac{1}{3}\right)\times\left(\tfrac{2}{3}\right)^2\\ &\qquad + \left(\tfrac{2}{3}\right)^3\times\left(\tfrac{1}{3}\right)\times\left(\tfrac{2}{3}\right) + \left(\tfrac{2}{3}\right)^4\times\left(\tfrac{1}{3}\right)\\ &= 5\times\left(\tfrac{1}{3}\right)\times\left(\tfrac{2}{3}\right)^4 = \tfrac{80}{243}.\end{aligned}$$

Notice that the probability for each individual sequence is the same as for any other sequence, namely $\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^4$, and that the 5 in the last line corresponds to the number of different sequences which give $X=1$. You could have counted the number of sequences by using an argument involving combinations. There are 5 positions in each sequence and one of the positions must be filled with a success (s) and the remaining four must all be failures (f). The choice of which position to place the s in can be made in any one of $\binom{5}{1}$ ways.

Therefore $P(X=1)=\binom{5}{1}\times\left(\frac{1}{3}\right)\times\left(\frac{2}{3}\right)^4$.

Similarly you could find that for $X=2$ there are sequences such as $s_1s_2f_3f_4f_5$ and $f_1s_2f_3s_4f_5$. To see how many of these sequences there are, consider the following argument. There are 5 positions which have to be filled with 2 *s*s and 3 *f*s. After you choose the places in which to put the 2 *s*s, there is then no choice as to where the *f*s go.

There are 5 places to choose from for the 2 *s*s. From Section 5.3, you have seen that the number of choices is $\binom{5}{2}$. Each of these choices has probability $\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^3$. Using these results $P(X=2)=\binom{5}{2}\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^3$.

Continuing in this way, you can find the distribution of X, given in Table 7.2.

x	$\mathrm{P}(X=x)$	
0	$\left(\frac{2}{3}\right)^5$	$=\frac{32}{243}$
1	$\binom{5}{1}\times\left(\frac{1}{3}\right)\times\left(\frac{2}{3}\right)^4$	$=\frac{80}{243}$
2	$\binom{5}{2}\times\left(\frac{1}{3}\right)^2\times\left(\frac{2}{3}\right)^3$	$=\frac{80}{243}$
3	$\binom{5}{3}\times\left(\frac{1}{3}\right)^3\times\left(\frac{2}{3}\right)^2$	$=\frac{40}{243}$
4	$\binom{5}{4}\times\left(\frac{1}{3}\right)^4\times\left(\frac{2}{3}\right)$	$=\frac{10}{243}$
5	$\binom{5}{5}\times\left(\frac{1}{3}\right)^5$	$=\frac{1}{243}$

Table 7.2. Probability distribution for the number of times out of 5 that the spinner shows black.

Notice that $\sum_{x=0}^{5}\mathrm{P}(X=x)$ is 1. This is a useful check for the probabilities in any distribution table.

Notice also that as $\left(\frac{1}{3}\right)^0=1$ and $\left(\frac{2}{3}\right)^0=1$ you could write $\mathrm{P}(X=0)$ as $\binom{5}{0}\times\left(\frac{1}{3}\right)^0\times\left(\frac{2}{3}\right)^5$, and $\mathrm{P}(X=5)$ as $\binom{5}{5}\times\left(\frac{1}{3}\right)^5\times\left(\frac{2}{3}\right)^0$. These results enable you to write $\mathrm{P}(X=x)$ as a formula:

$$\mathrm{P}(X=x)=\binom{5}{x}\times\left(\tfrac{1}{3}\right)^x\times\left(\tfrac{2}{3}\right)^{5-x}.$$

However the formula on its own is not sufficient, because you must also give the values for which the formula is defined. In this case x can take integer values from 0 to 5 inclusive. So a more concise definition of the distribution of X than Table 7.2 would be

$$\mathrm{P}(X=x)=\binom{5}{x}\times\left(\tfrac{1}{3}\right)^x\times\left(\tfrac{2}{3}\right)^{5-x}\qquad\text{for } x=0,1,2,\ldots,5.$$

Although the case of the spinner is not important in itself, it is an example of an important and frequently occurring situation.

- A single trial has just two possible outcomes (often called success, s, and failure, f).
- There is a fixed number of trials, n.
- The outcome of each trial is independent of the outcome of all the other trials.
- The probability of success at each trial, p, is constant.