Probability

Prerequisites

You should already be familiar with Venn diagrams and understand the use of the symbols $\cup$, $\cap$ and $\subseteq$ to denote relationships between sets.

Example (1)

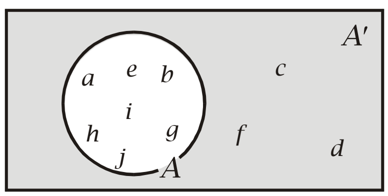
Domain of discourse, $S = \{a, b, c, d, e, f, g, h, i, j\}$
$A = \{a, b, e, g, h, i, j\}$
$B = \{b, f, g\}$

(*a*) Find $A', A \cup B, A \cap B, A' \cap B$ and show these sets using Venn diagrams.

(*b*) List all the subsets of *B*.

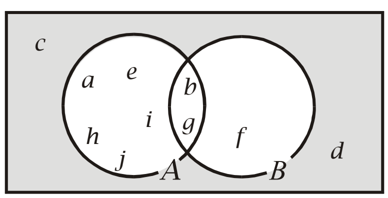
Solution

(*a*) (*i*) $A' = \{c, d, f\}$

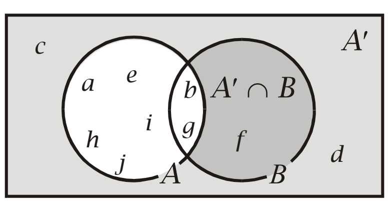


(*ii*) $A \cup B = \{a, b, e, f, g, h, i, j\}$

(*iii*) $A \cap B = \{b, g\}$



(*b*) (*iv*) $A' \cap B = \{f\}$



(b) $B = \{b,\ f,\ g\}$

Subsets $= \{b, f, g\}, \{b, f\}, \{b, g\}, \{f, g\}, \{b\}, \{f\}, \{g\}, \varnothing$

$\varnothing$ = the null set, the set with no members

What is probability?

Probability is concerned with the question – how likely is it that a certain event will occur? For example, how likely is it that a train will be on time? How likely is it that if someone shuffles a pack of cards and draws one card from it, that the card will be an ace? Answers to questions like these can be placed on a *possibility spectrum* at the one end of which we say that the event is *certain* and the other end we say that the event is *impossible*; in between we use terms such as *very likely, probable, unlikely* and so forth. Intuitively, we recognise and use such terms in every day life and often we use them in a very subjective way, based on feelings and guesses.

Mathematicians seek a precise definition of probability. Probability for them is concerned with events. The question in the most general sense is: what is the probability that this event will occur? And the answer for them is a number ranging from 1 to 0, where to say that the probability of an event is 1 means that it is certain that the event will occur (no other event could occur), whilst to say that the probability that an event will occur is 0 means that it is impossible that the event could occur. Decimals between 0 and 1 express various shades of possibility. Sometimes these numbers are expressed as percentages and sometimes they are expressed as ratios. We use a ratio when we say, for instance, that the probability that when someone draws one card from a pack, that the card will be a heart is 1 in 4.

Example (2)

Someone says, "The probability that I when I toss this coin that it will come up heads is 1 in 2." Express this probability

(*a*) As a decimal between 0 and 1.

(*b*) As a percentage.

(*c*) As a fraction

Solution

(*a*) 0.5

(*b*) 50%

(*c*) $\frac{1}{2}$

All of the responses in example (2) are mathematical statements about the probability of an event occurring.

Experimental probability

Let us take the question: how likely is it that the next train will be on time?

Example (3)
How would you determine by experiment the probability that a train will be on time?

Solution
The answer to this question is more involved than one might initially think. We would need to be precise about (*a*) what trains we are considering and (*b*) what our definition of a train being on time is. Once we have answered these questions we would then take a *sample*, meaning we would observe trains and count
n = how many trains you observed in total.
r = how many of these trains were on time.
We would then divide r by n to obtain the *experimental probability* for the event

$$\text{Experimental probability that the train is on time} = \frac{r}{n}.$$

Finally, we would *assume* that the probability that we have just determined by experiment can be used to *predict* the probability of future events.

Every time we make an observation in a case like this it is called a *trial*. When a trial has two possible *outcomes* we may call one of these outcomes a *success* and the other a *failure*. From the mathematical point-of-view this does not imply a value judgement – a success in this mathematical sense might be something "bad" and a failure something "good". In example (2) we might call the departure of a train on time a *success*, but this is just so that we can make a difference between that outcome and a *failure* when the train does not depart on time.

Example (4)
100 observations are made of trains departing from a certain station. Of these, 85 trains left at the scheduled time.

(*a*) What is the experimental probability that a train at this station will be on time?

(*b*) If a further 60 trains are observed at this station, how many trains may be expected to depart on time?

(*c*) Is your prediction in answer to part (*b*) certain to be fulfilled?

Solution

(*a*) Here the total number of observations is $n = 100$.

The number of "successful" observations, where the train left at the scheduled time was $r = 85$.

$$\text{Experimental probability that the train departs on time} = \frac{r}{n} = \frac{85}{100} = 0.85$$

(*b*) This is a prediction. We must base the prediction on something, and the obvious assumption to make here is that the future will be like the past. So the probability of a train departing on time is expected to be the same, that is 0.85 as in part (*a*). The number of trains expected to depart on time out of 60 is therefore

$$\begin{aligned}\text{Expectation} &= \text{number of trials} \times \text{probability}\\ &= 60 \times 0.85\\ &= 51\end{aligned}$$

(*c*) It is quite clear that there is no certainty that *exactly* 51 of the next 60 trains will depart on time. On the contrary we would expect some *variation* in the future. This might be just due to chance and it might be brought about by other causes. For example, if the management of the train company improved a greater proportion of trains might depart on time.

Probabilities when all events are equally likely

In many cases intuitively we have a fair idea of how to predict a future event.

Example (5)

A coin is tossed. This shall be called a trial. A success shall be if the coin comes up heads, and a failure shall be if the coin comes up tails. What is the probability that a single trial will result in a success? What assumption must you make in order to state this probability?

Solution

There is one trial with two possible outcomes – success = heads, failure = tails. The number of possible outcomes is 2, and the number of ways in which we can have a success (a head in this case) is 1. Therefore, the probability is $\frac{1}{2} = 0.5$. The assumption underlying this argument is that every possible outcome is *equally likely*.

If we are able to claim that every event is equally likely then we do not need an experiment in order to make a prediction about the probability of an event occurring. In such a case the probability of an event occurring shall be

$$\text{Probability} = \frac{\text{number of outcomes in which the event occurs}}{\text{total number of possible outcomes}}$$

This is called the *classical definition of probability*.

Example (6)

A fair coin is tossed three times. At each trial there are two possible outcomes H = heads, and T = tails.

(*a*) Make a list of all the possible outcomes. This shall be called the *sample space*. Let n be the number of possible outcomes. Find n.

(*b*) Make a list of all those outcomes where there is exactly one heads. Let r be the number of outcomes where there is exactly one head. Find r.

(*c*) What is the probability that when a coin is tossed three times there will be exactly one head?

Solution

(*a*) Let (H T T) mean, "heads followed by tails followed by tails".
Possible outcomes are
$(\text{HHH})\ (\text{HHT})\ (\text{HTH})\ (\text{HTT})\ (\text{THH})\ (\text{THT})\ (\text{TTH})\ (\text{TTT})$
$n = 8$

(*b*) $(\text{HTT})\ (\text{THT})\ (\text{TTH})$
$r = 3$

(*c*) The probability exactly 1 head out 3 $= \dfrac{r}{n} = \dfrac{3}{8} = 0.375$

The solution here, namely that the probability of exactly 1 head out of 3 trials, is $P = 0.375$ can be found without recourse to experiment on the assumption that each outcome is equally likely. Throughout the remainder of this chapter we shall continue to make this assumption.

In the solution to example (6) we allowed (H T T) to stand for the *outcome* "heads followed by tails followed by tails" in precisely that order. The collection all possible outcomes arising from n trials is a set, which is called the *sample space*. The sample space may also be called *possibility space* and each possible outcome in a sample space is called a *sample point*. An *event* is a set of sample points that is a subset of the sample space. A single outcome is also by this definition an event. Since the sample space is a set we use the convention that we represent it by enclosing all the possible outcomes (sample points) within curly $\{\ \}$ brackets. So in this example (5) the sample space is the set

$S = \{(\mathrm{HHH}), (\mathrm{HHT}), (\mathrm{HTH}), (\mathrm{HTT}), (\mathrm{THH}), (\mathrm{THT}), (\mathrm{TTH}), (\mathrm{TTT})\}$.

The phrase "The probability of event A" is represented by the mathematical expression $P(A)$.

For example, we could express the phrase, "The probability of exactly 1 head out of 3 is 0.375" by

$P(\text{exactly 1 head out of 3}) = 0.375$

or

$P(A) = 0.375$ where A = exactly 1 head out of 3.

Example (6) continued

A coin is tossed three times. At each trial there are two possible outcomes H = heads, and T = tails.

(*a*) Let B = the event that two or more outcomes are tails. Find B as a set of possible outcomes.

(*b*) Find $P(B)$.

(*c*) Let B' be the complement of B; that is, the event *not-B*. Find B' and $P(B')$.

(*d*) Draw a Venn diagram to illustrate the fact that B and B' are subsets of the set S, where S is the sample space.

Solution

(*a*) $B = \{(\mathrm{HTT}), (\mathrm{THH}), (\mathrm{TTH}), (\mathrm{TTT})\}$

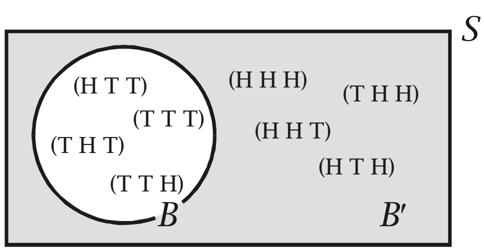
(*b*) $n = 8 \quad r = 4$

$$P(B) = \frac{r}{n} = \frac{4}{8} = \frac{1}{2}$$

(*c*) $B' = \{(\mathrm{HHH}), (\mathrm{HHT}), (\mathrm{HTH}), (\mathrm{THH})\}$

$$P(B') = \frac{4}{8} = \frac{1}{2}$$

(*d*)



The Venn diagram illustrates the fact that the union of any set A and its complement A' is equal to the entire sample space S.

$A \cup A' = S$

For this reason the probability of $A \cup A'$ is the probability of S which is 1.

$P(A \cup A') = P(S) = 1$

This represents the fact that it is certain that either A or *not-A* shall occur, whatever the event A. It also means that if for a given event A we already know $P(A)$ then the probability of *not-A* is given by $P(A') = 1 - P(A)$. For the following question we also remind you that the null set is defined to be the set that contains no members, and is represented by the symbol $\varnothing$.

Example (6) continued

A coin is tossed three times. At each trial there are two possible outcomes H = heads, and T = tails. Let

A = the event that exactly one outcome is heads

B = the event that two or more outcomes are tails

C = the event that exactly two outcomes are tails

D = the event that exactly three outcomes are tails

(*a*) State the relationship between A and C.

(*b*) Draw separate Venn diagrams to illustrate each of the following relationships between sets.

(*i*) $C \subseteq B$

(*ii*) $B = C \cup D$

(*iii*) $C \cap D = \varnothing$

(*c*) Find

(*i*) $P(C)$ (*ii*) $P(D)$ (*iii*) $P(C \cap D)$

(*iv*) $P(C \cup D)$ (*v*) $P(C')$

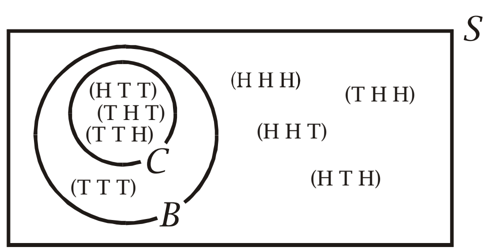
(*d*) Show that $P(C \cup D) = P(C) + P(D) - P(C \cap D)$.

Solution

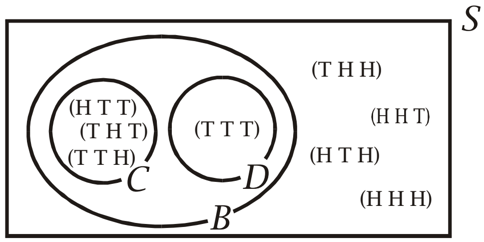
(*a*) $A = C$ The events A and C are different descriptions of the same event.

(*b*) (*i*) $C \subseteq B$

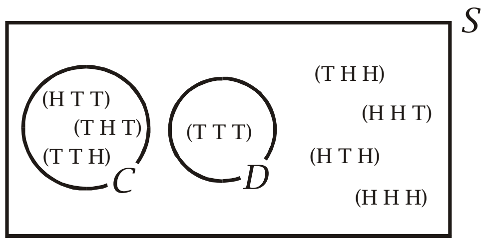
In the diagram note that HHH represents the outcome heads followed by heads followed by heads in precisely that order.



(*ii*) $B = C \cup D$



(iii) $C \cap D = \varnothing$



There is no overlap between the sets *C* and *D*.

(c) (i) $P(C) = \frac{3}{8}$

(ii) $P(D) = \frac{1}{8}$

(iii) $P(C \cap D) = 0$

(iv) $P(C \cup D) = P(B)$

$$= \frac{4}{8}$$

$$= \frac{1}{2}$$

(v) $P(C') = 1 - P(C)$

$$= 1 - \frac{3}{8}$$

$$= \frac{5}{8}$$

(d) $P(C \cup D) = \frac{1}{2}$

$P(C) + P(D) - P(C \cap D) = \frac{3}{8} + \frac{1}{8} - 0 = \frac{1}{2}$

Therefore

$P(C \cup D) = P(C) + P(D) - P(C \cap D)$

In this last example we had the events

C = the event that exactly two outcomes are tails

D = the event that exactly three outcomes are tails we saw that $C \cap D = \varnothing$. If event C occurs then event D cannot occur. In this case the events C and D are said to be *mutually exclusive.*

Mutually exclusive events

Two events A and B are said to be mutually exclusive if it is not possible for both to occur. This means that $A \cap B = \varnothing$ and $P(A \cap B) = 0$.

Another important relationship between events is when they are independent.

Independent events

Two events A and B are said to be independent if the fact that one occurs does not influence the probability of the other occurring. In order to define independent events we need the concept of *conditional probability.* This notion is introduced in a subsequent chapter. For the present, we shall state that when events are independent then their intersection is empty; that is $A \cap B \neq \varnothing$ and

$P(A \cap B) = P(A) \times P(B)$ provided that $P(A) \neq 0$ and $P(B) \neq 0$.

This may also be used to test whether events are independent, as illustrated by Example (7).

From these definitions we see immediately that if A and B are mutually exclusive then they cannot be independent.

Example (7)

A golf club has 120 members in the following categories.

	Membership type		
	Silver	Gold	Platinum
Male	20	30	20
Female	25	15	10

A guest at the club wishes to play golf with a member, who is selected at random. Let A denote the event that the selected member is female and B denote the event that the selected member is a gold member.

Evaluate

(a) $P(A)$

(b) $P(B')$

(c) $P(A \cup B)$

(d) $P(A \cap B)$

(*e*) (*i*) Show that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

(*ii*) Are events A and B mutually exclusive?

(*iii*) Are events A and B independent?

Solution

(*a*) $P(A) = \dfrac{\text{total number of females in the club}}{\text{total number of members}} = \dfrac{25+15+10}{120} = \dfrac{50}{120} = \dfrac{5}{12}$

(*b*) $P(B) = \dfrac{\text{total number of gold members}}{\text{total number of members}} = \dfrac{30+15}{120} = \dfrac{45}{120} = \dfrac{3}{8}$

$P(B') = 1 - P(B) = 1 - \dfrac{3}{8} = \dfrac{5}{8}$

(*c*) $$P(A \cup B) = \frac{\begin{pmatrix}\text{total number of female members} \\ + \text{ total number of male gold members}\end{pmatrix}}{\text{total number of members}} = \frac{50+30}{120} = \frac{80}{120} = \frac{2}{3}$$

(*d*) $$P(A \cap B) = \frac{\text{total number of female members that are also gold members}}{\text{total numberof members}} = \frac{15}{120} = \frac{1}{8}$$

(*e*) (*i*) $P(A) + P(B) - P(A \cap B) = \dfrac{5}{12} + \dfrac{3}{8} - \dfrac{1}{8} = \dfrac{10+9-3}{24} = \dfrac{16}{24} = \dfrac{2}{3}$

From part (c) $P(A \cup B) = \dfrac{2}{3}$. Therefore, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(*ii*) A and B are not mutually exclusive because $P(A \cap B) = \dfrac{1}{8} \neq 0$.

(*iii*) A and B are not independent because

$P(A) \times P(B) = \dfrac{5}{12} \times \dfrac{3}{8} = \dfrac{5}{32}$ $\qquad$ $P(A \cap B) = \dfrac{1}{8} \neq P(A) \times P(B)$.

Example (8)

Two fair 4-sided dice, the faces of which are both numbered 1, 2, 3 and 4, are thrown. Let A be the event that the sum of the scores of the two die is 6. Let B be the event that at least one of the two dice is a 1.

(*a*) List all the events of the sample space as a set.

(*b*) Find

(*i*) $P(A)$, $P(B)$ and $P(A \cap B)$.

(*ii*) State with a reason whether A and B are mutually exclusive events.

(*c*) Use the formula $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ to find $P(A \cup B)$.

Solution

(*a*) The sample space is the set

$$S=\left\{\begin{array}{l}(1,1),(1,2),(1,3),(1,4),(2,1),(2,2,),(2,3),(2,4),\\ \qquad\qquad (3,1),(3,2),(3,3,),(3,4),(4,1),(4,2),(4,3),(4,4)\end{array}\right\}$$

Here the symbol $(1,2)$ means a 1 followed by a 2 in that order.

(*b*) The sum of the scores of the two die is shown by the following table.

	First die				
Second die	**1**	**2**	**3**	**4**	
1	2	3	4	5	
2	3	4	5	6	***A***
3	4	5	6	7	
4	5	6	7	8	
	B				

In this diagram of *sample points* we have circled event A in bold and shaded event B in grey. The total number of events in the sample space is $n=16$.

(*i*) As the table shows there are $r_1=3$ ways in which the sum of the two die can be 6. Hence $P(A)=\frac{r_1}{n}=\frac{3}{16}$.

There are $r_2=7$ ways in which one of the two die can be a 1. Hence

$P(B)=\frac{r_2}{n}=\frac{7}{16}$.

The diagram also shows that there is no overlap between the two events

$A\cap B=\varnothing \qquad \Rightarrow \qquad P(A\cap B)=0$

(*ii*) Since $P(A\cap B)=0$, A and B are mutually exclusive events.

(*c*) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

$$=\frac{3}{16}+\frac{7}{16}-0=\frac{10}{16}=\frac{5}{8}$$

Some proofs

In the last section we introduced and used without proof three results about probability. We will now prove these.

(**1**) The probability of an event occurring is a number between 0 and 1 inclusive.

Proof

Let S be a sample space and let the number of sample points (outcomes) in sample space be n. Let A be an event comprising a subset of sample points of S, and let the number of sample points of A be r. By the definition of the probability of A

$$P(A) = \frac{r}{n}.$$

Since A is a subset of S

$$0 \le r \le n$$

This implies that

$$0 \le \frac{r}{n} \le 1$$

and hence

$$0 \le P(A) \le 1$$

(**2**) Let A be an event and A' denote the event that A does not occur. Then $P(A') = 1 - P(A)$.

Proof

Let S be a sample space and let the number of sample points (outcomes) in sample space be n. Let A be an event comprising a subset of sample points of S, and let the number of sample points of A be r. Let $n(A')$ denote the number of sample points of A'.

Since $S = A \cup A'$ we have

$$n(A') = n - r$$

Hence

$$\begin{aligned} P(A') &= \frac{n-r}{n} \\ &= 1 - P(A) \end{aligned}$$

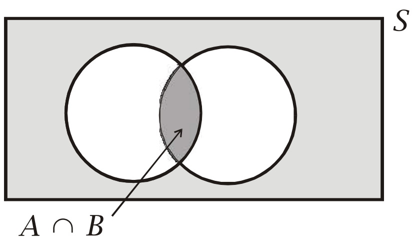
This also establishes that $P(A) + P(A') = 1$.

(**3**) Let A and B be events. Then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

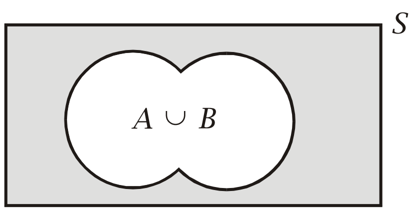
Proof

Let S be a sample space and let the number of sample points (outcomes) in sample space be n. The definition of $A \cup B$ means that if A and B are interpreted as events, then $A \cup B$

is the set of all sample points of S that are included in either A **or** B. Also the set $A \cap B$ is the set of sample points that are included in both A **and** B. Let $n(A)$ be the number of sample points in A, $n(B)$ be the number of sample points in B, $n(A \cup B)$ be the number of sample points in $A \cup B$, and $n(A \cap B)$ be the number of sample points in $A \cap B$. The following diagram shows that both the sets A and B include the set $A \cap B$.



Therefore $n(A)+n(B)$ counts the sample points of $A \cap B$ twice. On finding $n(A \cup B)$ we must subtract from $n(A)+n(B)$ the number $n(A \cap B)$.



Therefore

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

$$\begin{aligned} P(A \cup B) &= \frac{n(A)+n(B)-n(A \cap B)}{n} \\ &= \frac{n(A)}{n} + \frac{n(B)}{n} - \frac{n(A \cap B)}{n} \\ &= P(A) + P(B) + P(A \cap B) \end{aligned}$$

Probability tree

In order to solve problems involving probabilities we use a diagram called a *probability tree.*

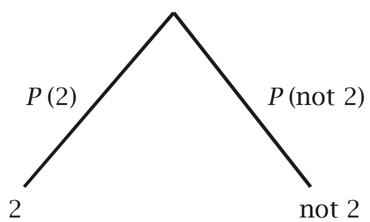
Example (9)

There are 52 cards in a normal pack of cards. The cards are arranged into four suits, and in each suit there is just one with the number 2 on it, so there are four 2s in a pack of 52 cards in all.

(*a*) What is the probability of drawing a 2 from a pack of 52 cards?

(*b*) What is the probability of not drawing a 2 from a pack of 52 cards?

(*c*) Represent this information as a probability tree as follows.



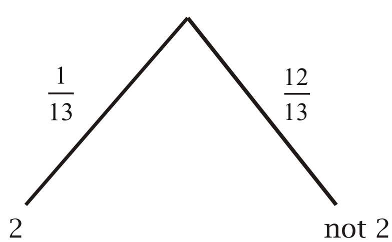
In this diagram where it is written $P(2)$ and $P(\text{not } 2)$ you should fill in the numerical values of these probabilities. Write these as fractions.

Solution

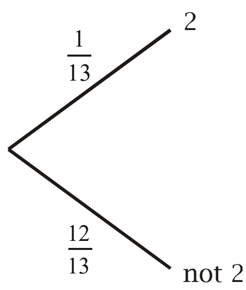
(*a*)
$$P(2) = \frac{\text{number of times 2 can occur}}{\text{total number of events}}$$
$$= \frac{\text{number of 2s in a pack of cards}}{\text{total number of cards in a pack}} = \frac{4}{52} = \frac{1}{13}$$

(*b*)
$$P(\text{not } 2) = 1 - P(2) = 1 - \frac{1}{13} = \frac{12}{13}$$

(*c*)



This probability tree shows the two possible outcomes of drawing one card from a pack of 52 cards as the end points to branches of a "tree". Events, here either a "2" or a "not 2", are placed at the end of the branches. Along the sides of the branches are placed the probabilities of those events occurring. Here we have drawn the probability tree going down the page, but it is equally possible to draw the tree going across the page.

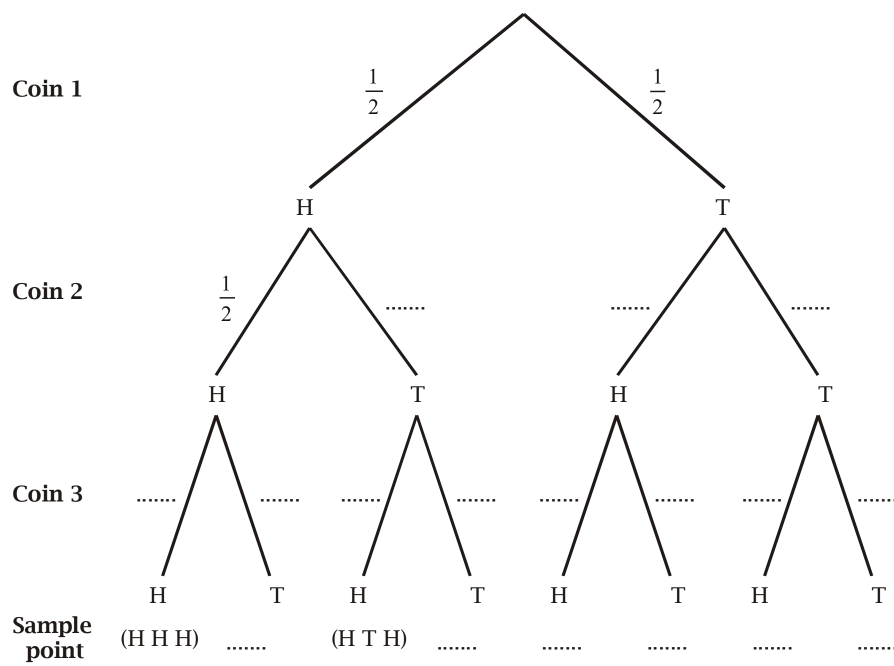


In this text we will draw the probability trees going down the page. In example (9) there is just one trial – only one card is drawn from the pack. Two or more trials are represented in a probability tree by successive branches.

Example (10)

In an experiment three fair coins (coin 1, coin 2 and coin 3) are tossed simultaneously. Let a success be defined to be a heads and let s denote the number of successes in this experiment. Let the outcome coin 1 is heads, coin 2 is heads and coin 3 is tails be represented by $(\mathrm{H\ H\ T})$ and so forth. The use of brackets indicates that the outcomes occur in a specific order; thus $(\mathrm{H\ H\ T})$ is a different outcome from $(\mathrm{H\ T\ H})$.

(*a*) Complete the following probability tree for this experiment.



(*b*) Using the definition of probability as

$$\text{Probability} = \frac{\text{number of outcomes in which the event occurs}}{\text{total number of possible outcomes}} = \frac{r}{n}$$

find each of the following

$P(s=0),\ P(s=1),\ P(s=2),\ P(s=3)$

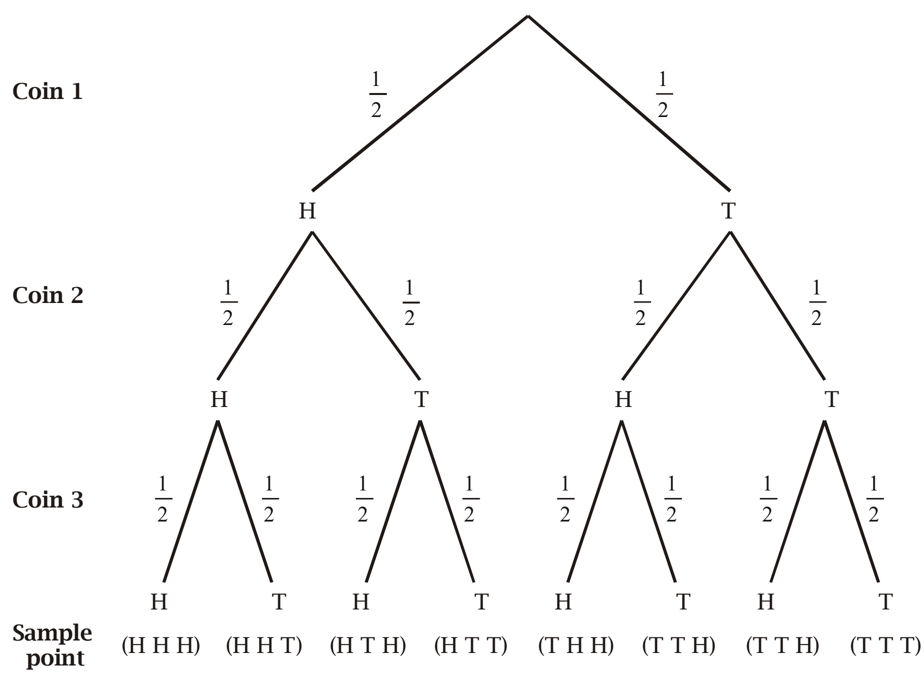
Remark

There is an assumption in this question – namely that the outcome of one trial does not in some way influence the outcome of another. This means that we are assuming, for instance, that if coin 1 turns up heads it will not change the likelihood of coin 2 turning up tails. The tossing of coin 2 is said to be *independent* of the tossing of coin 1. Thus it is also assumed that the events "coin 1 is a head" and "coin 2 is a tail" are *independent events.* When we say that a coin is fair, this means that we should apply the rule of classical probability – that all events are equally likely. When we toss any one coin there are two possible events – heads and tails. Since both are equally likely, the probability of each is $\frac{1}{2}$.

Also, when two events are independent we can simply multiply their probabilities. Here

$$\begin{aligned} P(s=0) &= P(\text{coin 1 is heads, coin 2 is heads, coin 3 is heads}) \\ &= P(\text{coin 1 is heads}) \times P(\text{coin 2 is heads}) \times P(\text{coin 3 is heads}) \\ &= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8} \end{aligned}$$

Solution to example (10)



(*b*) We examined a problem similar to this in example (6). The sample space is

$$S = \{(\mathrm{HHH}), (\mathrm{HHT}), (\mathrm{HTH}), (\mathrm{HTT}), (\mathrm{THH}), (\mathrm{THT}), (\mathrm{TTH}), (\mathrm{TTT})\}$$

As before, we have used the curved bracket $(\ )$ to indicate explicitly that the order in which the outcomes take place matter; that is $(\mathrm{H\,H\,T})$ is a different sample point from $(\mathrm{H\,T\,H})$ and $(\mathrm{H\,H\,T})$ means heads followed by heads followed by tails in precisely that order. The probabilities $P(s=0)$, $P(s=1)$, $P(s=2)$, $P(s=3)$ correspond to the following four events.

$P(s=0) \Leftrightarrow$ T T T in any order
$P(s=1) \Leftrightarrow$ T T H in any order
$P(s=2) \Leftrightarrow$ T H H in any order
$P(s=3) \Leftrightarrow$ H H H in any order

The sets of sample points for each of these events are

$P(s=0)$	$\Leftrightarrow$	T T T	in any order	$= \{(\mathrm{T\,T\,T})\}$	$r=1$
$P(s=1)$	$\Leftrightarrow$	T T H	in any order	$= \{(\mathrm{T\,T\,H}),(\mathrm{T\,H\,T}),(\mathrm{H\,T\,T})\}$	$r=3$
$P(s=2)$	$\Leftrightarrow$	T H H	in any order	$= \{(\mathrm{T\,H\,H}),(\mathrm{H\,T\,H}),(\mathrm{H\,H\,T})\}$	$r=3$
$P(s=3)$	$\Leftrightarrow$	H H H	in any order	$= \{(\mathrm{H\,H\,H})\}$	$r=1$

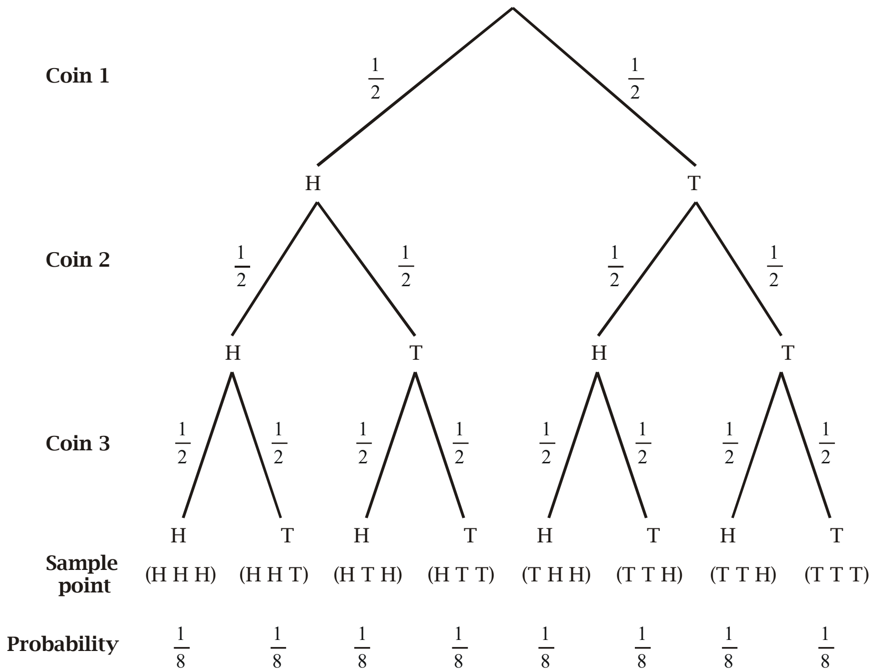
Thus $P(s=0)=\frac{1}{8}$ $\quad P(s=1)=\frac{3}{8}$ $\quad P(s=2)=\frac{3}{8}$ $\quad P(s=3)=\frac{1}{8}$

This example reflects an important distinction that has been implicit throughout this chapter. The symbol $(\mathrm{T\,H\,H})$ represents a specific order. It is called a *permutation*. A permutation is defined to be an ordered arrangement of items in a set. Each outcome (sample point) is a specific permutation. But we also had

the event there are 2 heads = T H H in any order $= \{(\mathrm{T\,H\,H}),(\mathrm{H\,T\,H}),(\mathrm{H\,H\,T})\}$

The set T H H in any order is a *combination*. A combination is simply a set or group of objects with no particular order chosen. In this example there are three permutations that give rise to the combination T H H in any order $= \{(\mathrm{T\,H\,H}),(\mathrm{H\,T\,H}),(\mathrm{H\,H\,T})\}$. We note also that in the probability tree the probability of each branch representing each outcome (sample point) is found by multiplying together each separate probability.

$$\begin{aligned} P(\text{perumation } (\mathrm{T\,H\,H})) &= P(\text{coin 1 is T}) \times P(\text{coin 2 is H}) \times P(\text{coin 3 is H}) \\ &= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \\ &= \frac{1}{8} \end{aligned}$$



Each outcome has the same probability of occurring = $\frac{1}{8}$. As there are three permutations belonging to the combination T H H in any order, the probability of this combination is

$$P(s=2) = 3 \times \frac{1}{8} = \frac{3}{8}$$

Thus the probability tree can be used to solve problems involving probabilities of successive trials. It provides a problem-solving device that avoids the need to explicitly find the sample points in a sample space and count the number of permutations corresponding to each event, which is a combination.

So where there are repeated trials the outcomes (sample points) correspond to permutations, and events are sets of outcomes corresponding to combinations. An important result about the probability of events is that the sum of the probabilities of all possible events is 1, which we prove as follows.

Proof

Let S be the sample space and let the number of sample points (outcomes) in sample space be n. Each outcome is a mutually exclusive sample point of the total sample space S. The total sample space is the sum of all the sample points. Let A_i represent an event; that is, it is a subset of S and is the union of mutually exclusive outcomes. Let $n(A_i)$ be the number of sample points in A_i. If

$A_1, A_2, ..., A_m$ are m mutually exclusive events such that $S = A_1 \cup A_2 \cup \ ... \ \cup A_m$ then $n(A_1) + n(A_2) + \ ... \ + n(A_m) = n(S)$ and

$$\begin{aligned} P(A_1) + P(A_2) + \ ... \ + P(A_m) &= \frac{n(A_1)}{n(S)} + \frac{n(A_2)}{n(S)} + \ ... \ + \frac{n(A_m)}{n(S)} \\ &= \frac{n(A_1) + n(A_2) + \ ... \ + n(A_m)}{n(S)} \\ &= \frac{n(S)}{n(S)} \\ &= 1 \end{aligned}$$

Problem-solving by means of probability trees

Problems with replacement

A problem *with replacement* is illustrated by the following example.

Example (11)

Suppose a card is drawn from a pack of cards and then replaced. Suppose after that another card is drawn.

(*a*) Draw the probability tree for this situation.

(*b*) What is the probability of obtaining two 2s?

(*c*) What is the probability of drawing just one 2?

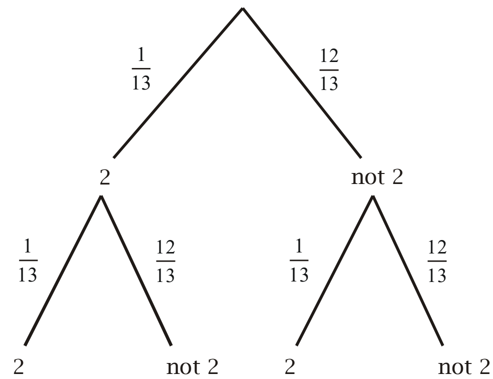
Remarks

(1) The first line of this example indicates clearly that it is a problem *with replacement*. Because the card drawn first is replaced it is assumed that the probability of drawing the second card is exactly the same as the probability of drawing the first card. There is a background assumption here that before drawing the second card the pack has been thoroughly reshuffled. This background assumption is equivalent to saying that both trials (drawing the first card, drawing the second card) are *independent events*.

(2) We remind you that there are 52 cards in a normal pack of cards. The cards are arranged into four suits, and in each suit there is just one with the number 2 on it, so there are four 2s in a pack of 52 cards in all.

Solution

(*a*)



(*b*) $P(2 \text{ followed by } 2) = P(\text{first card is a } 2) \times P(\text{second card is a } 2)$

$$= \frac{1}{13} \times \frac{1}{13} = \frac{1}{169}$$

(*c*) $P(\text{just one } 2) = P(2 \text{ followed by not } 2) + P(\text{not } 2 \text{ followed by } 2)$

$$= \left(\frac{1}{13} \times \frac{12}{13}\right) + \left(\frac{12}{13} \times \frac{1}{13}\right)$$

$$= 2 \times \frac{12}{169} = \frac{24}{169}$$

Problems without replacement

Example (12)

A card is drawn from a pack of 52 cards. It is not replaced. A second card is drawn.

(*a*) Find the probability of obtaining two 2s.

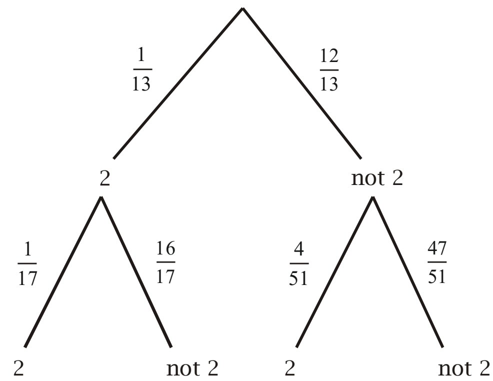
(*b*) Find the probability of drawing just one 2.

(*c*) Find the probability of obtaining no 2s.

Remark

After the first card is picked, there will be only 51 cards left. If a 2 was picked on the first occasion, then there will be just three 2s left and 48 other cards. If a 2 was not picked first, there will still be four 2s in the pack, but only 47 other cards. Thus, if a 2 was picked first, then the probability of a 2 on the second occasion is $P(\text{second card is a } 2) = \frac{3}{51} = \frac{1}{17}$. If a 2 was not picked first, then the probability of a 2 on the second occasion is $P(\text{second card is a } 2) = \frac{4}{51}$.

Solution



(*a*) $P(2 \text{ followed by } 2) = P(\text{first card is a } 2) \times P(\text{second card is a } 2)$

$$= \frac{1}{13} \times \frac{1}{17}$$
$$= \frac{1}{221}$$

(*b*) $P(\text{just one } 2) = P(2 \text{ followed by not } 2) + P(\text{not } 2 \text{ followed by } 2)$

$$= \left(\frac{1}{13} \times \frac{16}{17}\right) + \left(\frac{12}{13} \times \frac{4}{51}\right)$$
$$= \left(\frac{1}{13} \times \frac{16}{17}\right) + \left(\frac{4}{13} \times \frac{4}{17}\right) = 2 \times \frac{16}{221} = \frac{32}{221}$$

(*c*) $P(\text{not } 2 \text{ followed by not } 2) = P(\text{first card is not a } 2) \times P(\text{second card is not a } 2)$

$$= \frac{12}{13} \times \frac{47}{51}$$
$$= \frac{4}{13} \times \frac{47}{17} = \frac{188}{221}$$

If we now add all these probabilities together we obtain

$$\text{Total probability} = \frac{1}{221} + \frac{32}{221} + \frac{188}{221} = \frac{221}{221} = 1$$

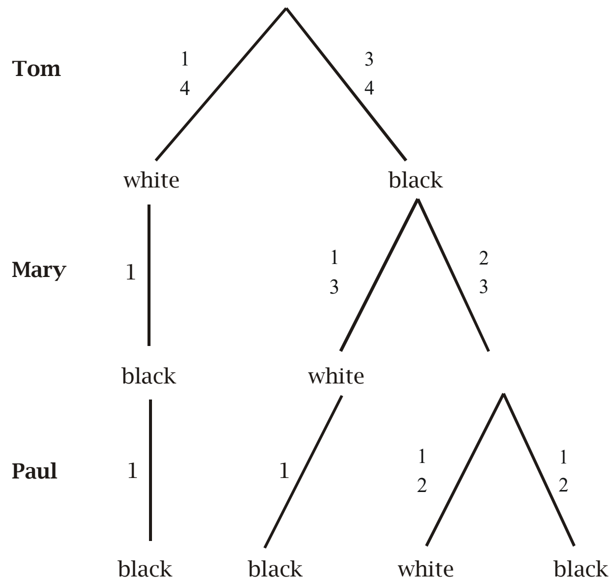
This illustrates the law of total probability that the sum of all the probabilities for all mutually exclusive possible events is always equal to 1.

Example (13)

A bag contains 4 balls, 3 are black and 1 is white. Tom takes one ball from the bag and keeps it. The Mary takes a ball from the bag and keeps it. Finally Paul takes a ball from the bag and keeps it. Draw a probability tree to illustrate this situation and find the probability that the white ball is picked by

(*a*) Tom (*b*) Mary (*c*) Paul (*d*) No one.

Solution



(*a*) $P(\text{Tom gets white}) = \frac{1}{4}$

(*b*) $P(\text{Mary gets white}) = \frac{3}{4} \times \frac{1}{3} = \frac{1}{4}$

(*c*) $P(\text{Paul gets white}) = \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{4}$

(*d*) $P(\text{no one gets white}) = 1 - \frac{3}{4} = \frac{1}{4}$

Problems in permutations and combinations

Whenever we draw a branch of a probability tree, this branch represents a possible outcome – a sample point of the total sample space. The probability of that outcome is determined by multiplying together the probabilities of the individual segments of that branch. An event is a set of outcomes and a subset of the sample space. To find the probability of an event we add the probabilities of the outcomes.

Example (14)

There are 11 books on a shelf, 4 of which are novels and 7 of which are textbooks. Three books are selected at random. Find the probability that

(*a*) 2 textbooks and 1 novel are selected.

(*b*) All 3 selected are either all textbooks or all novels.

Solution

Let us denote the outcome that a novel is chosen by N and the outcome that a textbook is chosen by T. This is a problem without replacement, since once a book has been chosen it is not put back on the shelf to be chosen again. So the same book cannot be chosen twice.

(*a*) It looks as if we need to draw the entire probability tree in order to solve this problem. However, on reflection this is not necessary. The problem requires us to evaluate the probability of T T N in *any* order. We see that there are three ways in which the symbols T T N in any order can be chosen. These are the specific permutations

(T T N) (T N T) (N T T)

Here, for example, (T N T) means T followed by N followed by T in precisely that order. These three permutations correspond to three branches of the probability tree representing the three possible ways in which the 2 textbooks and 1 novel can be selected. Each branch has the same probability.

The branch (T T N) has the probability $P(\text{T T N}) = \frac{7}{11} \times \frac{6}{10} \times \frac{4}{9} = \frac{28}{165}$.

The branch T N T has the same numbers but in a different permutation $P(\text{T N T}) = \frac{7}{11} \times \frac{4}{10} \times \frac{6}{9} = \frac{28}{165}$.

So the total probability of choosing 2 textbooks and 1 novel is

$$\begin{aligned} P(\text{2 textbooks and 1 novel}) &= 3 \times P(\text{T T N}) \\ &= 3 \times \left(\frac{7}{11} \times \frac{6}{10} \times \frac{4}{9}\right). \\ &= \frac{28}{55} \end{aligned}$$

(*b*) First observe that

$P(\text{All 3 selected are either all textbooks or all novels}) = P(\text{T T T}) + P(\text{N N N})$

In the case of the outcome T T T there is only one permutation corresponding to this combination, and only one branch of the probability tree. Likewise, there is only one permutation for the outcome N N N. Hence

$$
\begin{aligned}
P(\text{All 3 selected are either all textbooks or all novels}) &= P(\mathrm{T\ T\ T}) + P(\mathrm{N\ N\ N}) \\
&= \left(\frac{7}{11}\times\frac{6}{10}\times\frac{5}{9}\right) + \left(\frac{4}{11}\times\frac{3}{10}\times\frac{2}{9}\right) \\
&= \frac{234}{990} \\
&= \frac{13}{55}
\end{aligned}
$$