21. PROBABILITY

Probability is a branch of mathematics that deals with uncertainty, random or chance occurrences. Through the study of probability, we learn how to quantify the chances or likelihoods that are associated with various outcomes of an experiment.

Terms used in probability

The term experiment, in the topic of probability, has a precise meaning that may differ slightly from its everyday use as in science.

An **experiment** is the process by which an observation or measurement is made.

Examples of experiments in probability can be tossing a coin or die, selecting cards from a deck, conducting an opinion survey or testing an item to determine if it is defective.

Consider the experiment of tossing an ordinary fair die. The results that are possible when the experiment is performed are called **outcomes**. For example, the possible results, in this case, are 1 or 2 or 3 or 4 or 5 or 6. Note that there is some degree of uncertainty as to which of these outcomes will occur.

The **set** of all possible outcomes is called the **sample space, *S***.

Hence, the sample space is {1, 2, 3, 4, 5, 6}.

An **event, *E*** is any subset of outcomes contained in the sample space.
A **simple event** has exactly one outcome.
A **compound event** has more than one outcome and is a collection of simple events.

Consider the events**:**

E_1 = ***a six occurs***

E_1 is a **simple** event, it has only one outcome, $E_1=\{6\}$.

E_2 = ***an odd number occurs***

E_2 is a **compound event** as it consists of a set of simple events,

$E_2 = \{1, 3, 5\}$

In calculating probabilities, it is necessary to determine if the event of interest is a simple or compound event.

A **favourable outcome** is one that is desirable or required. In the experiment of tossing a die, if we are interested in the event,

E_3 = a prime number occurs, then the set

$E_3 = \{2,3,5\}$ is our set of favourable outcomes.

Equally likely outcomes

In performing experiments we may find that there are cases where all the outcomes have the same chance of occurring. For example, if we flip a fair coin or toss a fair die, then each outcome has the same chance of occurring.

These outcomes are **equally likely** or **equiprobable**. In the case of the coin, each outcome, Head or Tail, has the same probability. $P(H) = P(T) = \frac{1}{2}$.

In the tossing of the die, all six outcomes are equally likely events and may be written as: $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$.

So, events are equally likely if they have the same probability of occurring.

Defining Probability

For simple experiments with equally likely outcomes and a finite sample space, we calculate probabilities by using the following formula.

If E is an event in an experiment, then the probability that E occurs is written as $P(E)$, where

$$P(E) = \frac{Number\ of\ outcomes\ favourable\ to\ the\ event,\ E}{Total\ Number\ of\ outcomes\ in\ the\ sample\ space}$$

The simplest experiment to which probability applies is one in which there are two possible outcomes. For example, tossing a fair coin.

H is the event of getting 'a head'. $P(H) = \frac{1}{2}$

The numerator = 1 since there is only one head or one outcome which is favourable to the event.

Denominator = 2, since there are two possible outcomes, that is, head and tail.

The probability of obtaining a 'Head' when tossing a coin can be expressed as a decimal (0.5), a fraction (½), a percentage (50%) or in ratio form of 1:2. Expressing the probability in an exact form is desirable. As such, we find that expressing the probability as a numerical fraction is preferred. For instance, if $P(E) = \frac{3}{7}$, there is no exact decimal equivalent to this value.

Experimental and Theoretical Probability

The above definition of probability is theoretical in nature and relies on the construction of a sample space. Sometimes this method is not possible and we have to rely on past experiences to estimate probabilities. In such cases, we use data from observations based on actual occurrences of past events. Probabilities based on weather predictions are estimated by such methods.

When we use data from observations to calculate probabilities, we obtain results based on what actually happened. This is referred to as experimental probability. Data from frequency distributions are commonly used to calculate such probabilities. We refer to such probabilities as the relative frequency of an event, defined as follows:

$$\text{Relative Frequency} = \frac{\text{Number of times the event occurred}}{\text{Total number of trials}} = \frac{n(A)}{n}$$

As an example, we can calculate the probability of drawing an ace from a deck of 52 playing cards by setting up an experiment, drawing one card at a time, and recording the result as 'Ace' or 'not Ace'.
If after 100 trials we get 10 Aces, then

$P(Ace) = \frac{10}{100} = \frac{1}{10}$ by experiment.

We could have calculated the theoretical probability using the probability formula. There are 4 Aces in the deck and the number of favourable outcomes is therefore 4. The number of possible outcomes is 52.

$$P(Ace) = \frac{\text{No. of outcomes favourable to the event}}{\text{No. of possible outcomes}} = \frac{4}{52} = \frac{1}{13}$$

Notice that the relative frequency may not be the same as the theoretical value. If we were to increase the number of trials in the experiment, then the experimental probability would become closer and closer to the theoretical probability. We may think of the theoretical value as the result that would be obtained if an infinitely large number of trials were conducted in the experiment.

Basic Laws of Probability

The first law of probability is concerned with the range of values that the probability of an event can take. It is stated below.

1. The probability of an event, *P(A)* must lie in the interval $0 \le P(A) \le 1$.

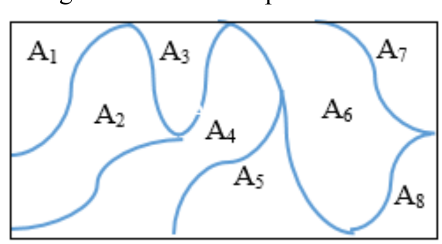
The smallest value of *P(A)* is zero. An **impossible** event has a probability of zero and it occurs when there are no events favourable to the outcome. For example, if A is the event - choosing a 7 on a die, then $P(A) = 0$. A probability of zero is the smallest possible value of the probability of an event. So, $P(A) \ge 0$ for any event A.

The largest value of *P(A)* is one. Such an event is one that is sure or certain to happen and has a probability of one. In this case, the number of outcomes favourable to the event is the same as the number of possible outcomes. For example, the probability of either a Head or a Tail occurs when a coin is tossed is one. A probability of one is the largest possible value of the probability of an event. So, $P(A) \le 1$ for any event A. So $0 \le P(A) \le 1$.

The second law in probability is concerned with the sample space and the set of all outcomes within it. It is stated below.

2. The sum of the probabilities of all the outcomes in a sample space must total one. $P(S) = 1$.

A Venn diagram is useful to represent this law.



If we have a collection of disjoint events, say,

$\{A_1, A_2, \ldots, A_8\}$, whose union make up the sample space, then

$$A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6 \cup A_7 \cup A_8 = S$$

$$P(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6 \cup A_7 \cup A_8) = P(S) = 1$$

In other words, the sum of the probabilities of each possible outcome must total to one.

Let us apply this law to an experiment in which three coins are tossed. The set of all outcomes are:

{HHH, TTT, HTH, HTT, HHT, THT, THH, TTH}.

P(3 heads) = $\frac{1}{8}$ P(2 Heads, 1 Tail) = $\frac{3}{8}$

P(3 Tails) = $\frac{1}{8}$ P(2 Tails, 1 Head) $\frac{3}{8}$

Notice that each of these outcomes is distinct which means that no two of them can ever occur at the same time.

The sum of all the probabilities

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

The third law is really a special case of the second law when there are only two events in the sample space.

If two events, A and B make up a sample space, then we can consider $B = A'$, that is, one event is the complement of the other. This means that either A occurs or that it does not occur.

$P(A) + P(A') = P(S)$

$P(S) = 1$ (from axiom 2)

$P(A) + P(A') = 1$

$P(A') = 1 - P(A)$ or $P(A) = 1 - P(A')$

> If two events, A and A' make up a sample space, then $P(A') = 1 - P(A)$ where $P(A')$ is the probability that event A does not occur.

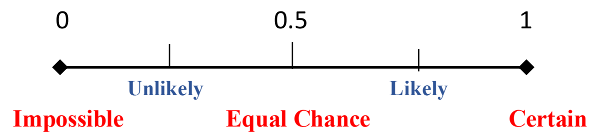
For example, if the probability, $P(R)$, that it will rain tomorrow is given by $P(R) = \frac{5}{7}$, then, $P(R')$ is the probability that it will not rain tomorrow, and

$$P(R') = 1 - P(R) = 1 - \frac{5}{7} = \frac{2}{7}$$

Probability scale

We know that the probability of an event lies between 0 and 1. We can now create a number line to plot certain probabilities and interpret their values based on the position on the scale. The Probability Scale is shown below with some critical values.

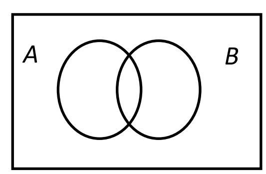
Events whose probabilities are closer to one are considered 'more likely', while events, whose probabilities are closer to zero, are considered 'less likely'. An event with probability equal to one-half has a 50-50 (or equal) chance of occurring.



The Addition Rule

We know from set theory that for two intersecting sets A and B, we can write their union as shown below.

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



If we divide each term by the total number of outcomes in the sample space, we obtain

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

This is the **addition rule of probability.** Recall that in set theory AUB, refers to A or B or both A and B.

Similarly, we interpret the probability of A or B, as the probability of A or B or both A and B. The addition rule may be stated as follows:

If A and B are two events then,

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

We use this rule in situations where both events can occur at the same time.

Example 1

A single card is drawn out of a deck of 52 playing cards. What is the probability of drawing a red card or a face card?

Solution

We define the events R and F as follows:
R – Selecting a red card
F – Selecting a face card
It is possible that both R and F could have occurred at the same time when a single card is selected.

$P(R) = \frac{26}{52} = \frac{1}{2}$ (26 of the cards are red)

$P(F) = \frac{12}{52} = \frac{3}{13}$ (12 of the cards are face cards)

$P(R \cap F) = \frac{6}{13}$ (6 of the face cards are red)

The probability of drawing a Red Card or a Face Card is $P(R \cup F)$. Applying the addition rule:

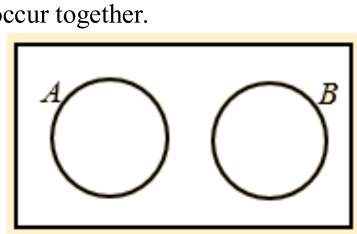
$P(R \cup F) = P(R) + P(F) - P(R \cap F)$

$= \frac{1}{2} + \frac{3}{13} - \frac{6}{13} = \frac{7}{26}$

Addition Rule for mutually exclusive events

From set theory, two sets are disjoint if there are no elements in common. Similarly, if two events are disjoint they cannot occur together. If there are only two **disjoint** sets, A and B, in the sample space, A and B cannot occur together, we refer to A and B as mutually exclusive events.

The events A and B are mutually exclusive if they cannot occur together.



$A \cap B = \phi$

For example, when a die is tossed, the events A and B are mutually exclusive, if
Event A- a 3 turns up
Event B- a 4 turns up
Note that it is NOT possible that both A and B could have occurred at the same time.

From a deck of cards, the events A and B are mutually exclusive, if
Event A- selecting a King
Event B - selecting a Queen
Note that it is NOT possible that both A and B could have occurred at the same time.

The addition rule for mutually exclusive events is obtained from the addition rule of probability by letting $A \cap B \neq \phi$. It is stated as follows

$P(A \cup B) = P(A) + P(B)$.

Both cases are summarized as follows:

Addition Rule
If A and B are not mutually exclusive
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
If A and B are mutually exclusive, $P(A \cap B) = 0$,
$P(A \cup B) = P(A) + P(B)$

We may further extend the addition rule for any number of mutually exclusive events in a sample space. If $E_1, E_2, E_3, \ldots, E_n$ are mutually exclusive events, then,

$P(E_1 \text{ or } E_2 \text{ or } E_3 \text{ or} \ldots \text{or } E_n)$

$= P(E_1) + P(E_2) + P(E_3) + \ldots + P(E_n)$

For example, if we toss a coin, the outcomes $P(1)$, $P(2)$, $P(3)$, $P(4)$, $P(5)$ and $P(6)$ are all mutually exclusive, so

$P(1 \text{ or } 2 \text{ or } 3 \text{ or } 4 \text{ or } 5 \text{ or } 6) = P(1) + P(2) + P(3) + P(4) + P(5) + P(6)$

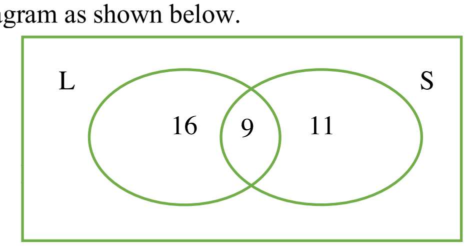
Example 2

A class has 36 students. Each student studies either Literature or Spanish or both. Twenty students study Spanish and 25 study Literature. A student is selected at random, what is the probability that the student studies
(a) Both subjects
(b) Literature only?

Solution

The information can be represented in a Venn Diagram as shown below.

(a)



(b) P(Literature only)= $\frac{16}{36} = \frac{4}{9}$

Example 3

The probabilities of three teams A, B and C winning a badminton competition are $\frac{1}{3}$, $\frac{1}{5}$ and $\frac{1}{9}$ respectively. Calculate the probability that:

a) Either A or B will win
b) Either A or B or C will win
c) None of these teams will win
d) Neither A nor B will win

Solution

a) $P(\text{A or B will win}) = \frac{1}{3} + \frac{1}{5} = \frac{8}{15}$

b) $P(A \text{ or } B \text{ or } C \text{ will win}) = \frac{1}{3} + \frac{1}{5} + \frac{1}{9} = \frac{29}{45}$

c)

$P(\text{None will win}) = 1 - P(\text{A or B or C will win})$

$= 1 - \frac{29}{45} = \frac{16}{45}$

d) P(Neither A nor B will win)

= 1− P(Either A or B will win)

$= 1 - \frac{8}{15} = \frac{7}{15}$

Sample Space for a discrete random variable

We can use the laws of probability to solve problems involving random variables. A random variable is a variable whose values are outcomes of a random experiment. In tossing a coin, there are two possible outcomes while in tossing two coins there are four possible outcomes. The sample space enables us to calculate probabilities for any event.

In tossing two coins, the sample space is {HH, HT, TH, TT}. The following probabilities are computed:

P(Two heads) $= \frac{1}{4}$

P(Two Tails) $= \frac{1}{4}$

P(A head and a tail in any order) $= \frac{1}{2}$

Notice that the sum of these probabilities is one because this set of outcomes are mutually exclusive and take into consideration all the possibilities.

Let us construct a sample space for tossing two dice, a green and a blue. As shown below there are 36 outcomes.

6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)
5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	1	2	3	4	5	6

We now define a random variable, S which represents the sum of the scores obtained when the dice are tossed. We can construct a sample space for this random variable as follows:

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

Each cell in the table gives the sum of the scores on both dice. If S represents the sum of the scores on both faces, then we can construct a table which shows the probability distribution for all possible values of the total, s.

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

This distribution enables us to calculate probabilities such as:

P(S< 5)= $\frac{4+3+2+1}{36} = \frac{10}{36}$, P(S≥ 10)= $\frac{3+2+1}{36} = \frac{6}{36}$

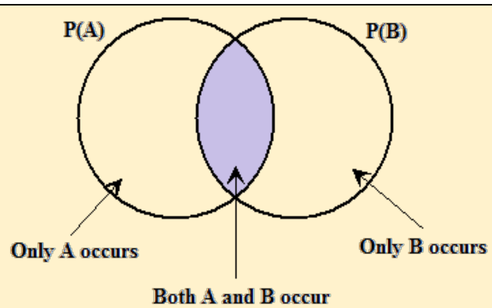
The Multiplication Rule for Probability

The multiplication rule can also be used to combine probabilities, but in such a case, the events must be independent. For example, in rolling **two** dice, we may be interested in the probability of getting a 6 on the first die and a 2 on the second die. The outcome of the first event does not influence the outcome of the second event.

Independent events

Two events, A and B are said to be independent if the probability of event B is not influenced or changed by the occurrence of event A and vice versa.

In tossing two coins, the event that a head occurs on the first coin and a tail on the second coin are independent events. Notice the outcome of the first coin (getting an H or not) does not influence the outcome of the second coin (getting a T or not). We can represent this situation using a Venn Diagram.



The shaded area, $A \cap B$, indicates when A and B both occur.

$P(A \text{ and } B) = P(A \cap B)$

$= P(A) \times P(B)$

Independent events cannot be mutually exclusive, because for mutually exclusive events $P(A \cap B) = 0$.

We can now state the multiplication rule for probability.

Multiplication Rule - Independent events

If A and B are two independent events, then

$P(A \cap B) = P(A) \times P(B)$.

This rule can be extended to two, three or any number of independent events. If A, B and C are all independent events, then $P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$

Example 4

A coin is flipped and a die is tossed. Let H represent the event that a head is obtained when the coin is flipped and S represent the event that a six (6) is obtained when the die is tossed.

Calculate the probability that both a head and a six are obtained.

Solution

Let P(H) = probability that a head occurs.

Let P(6) = probability that a six occurs

The sample space for the experiment has 12 outcomes;

{(H, 1), (H, 2), (H, 3), (H, 4), (H, 5), (H, 6), (T, 1), (T, 2), (T, 3) (T, 4), (T, 5), (T, 6)}

We are interested in combining the probability P(H) and P(S).

$P(\text{H}) = \frac{6}{12} = \frac{1}{2}$,

there are 6 favourable outcomes,

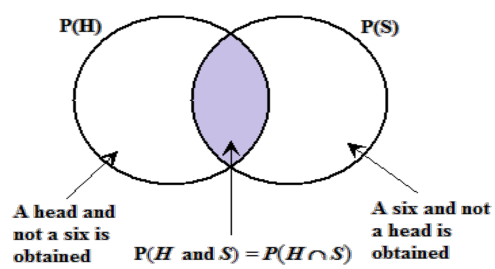
(H, 1), (H, 2), (H, 3), (H, 4), (H, 5), (H, 6),

$P(6) = \frac{2}{12} = \frac{1}{6}$,

there are 2 favourable outcomes, (H, 6) and (T, 6)

To obtain the probability that both a head and a six are obtained we must calculate:

$P(H \text{ and } S) = P(H \cap S)$ The Venn diagram below illustrates how the solution is obtained.



These are independent events since the outcome of one does not affect the outcome of the other, so we can multiply the probabilities. The intersection of H and S, shown shaded, indicates that H and S have

$P(H \cap S)$

$= P(H) \times P(S)$

$= \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$

Example 5

Three events are defined as follows:
H –a head is obtained when a coin is tossed.
F – a five (5) is obtained when a die is tossed.
A – an ace is obtained when a card is picked for a pack.
Calculate the probability that all three events occur.

Solution

We may think of three separate sets and we are interested in the intersection of the three regions.
P(a head and a five and ace)
$= P(H \text{ and } F \text{ and } A) = P(H \cap F \cap A)$

$P(H \cap F \cap A) = P(H) \times P(F) \times P(A)$

$= \frac{1}{2} \times \frac{1}{6} \times \frac{4}{52} = \frac{1}{156}$

Example 6

A bag contains 12 balls, of which 5 are red (R) and 7 are blue (B). The balls are indistinguishable apart from colour. We are interested in choosing 2 balls from the bag, one at a time.
A ball is chosen at random from the bag, it is replaced and a second ball is chosen.
(a) If the first ball is red, what is the probability that the second is also red?
(b) If the first ball is not red, what is the probability that the second is also not red?

Solution

As the ball is replaced, the second pick would now be independent of the result of the first pick. By replacing the ball, the probability of choosing Red remains the same on both occasions.

(a) $P(1^{st}$ is Red) = $\frac{5}{12}$ $P(2^{nd\ t}$ is Red) = $\frac{5}{12}$

P(Red, then Red) = $\frac{5}{12} \times \frac{5}{12} = \frac{25}{144}$

(ii) If the first ball is not red, what is the probability that the second is also red?
As the ball is replaced, the number of balls in the bag remains unchanged.

(b) $P(1^{st}$ is Not Red) = $\frac{7}{12}$ P(2^{nd} is Red) = $\frac{5}{12}$

P(Not Red, Red) = $\frac{7}{12} \times \frac{5}{12} = \frac{35}{144}$

If there is a replacement of the ball after the second pick, third pick or any subsequent picks, then such events are **independent** and we apply the multiplication law of probability.

Dependent events

Dependent events are commonly found in experiments involving selection from a set without replacements. In the above example, where two balls were chosen from a set, the first ball was replaced and this did not affect the outcome of the second pick.
If we were to repeat this exercise without replacing the first ball then this will have an effect on the outcome of choosing the second ball.
We now restate the problem as follows:

Example 7

A bag contains 12 balls, of which 5 are red (R) and 7 are blue (B). The balls are indistinguishable apart from colour. We are interested in choosing 2 balls from the bag, one at a time.
A ball is chosen at random from the bag, it is not replaced and a second ball is chosen.
(a) If the first ball is red, what is the probability that the second is also red?
(b) If the first ball is not red, what is the probability that the second is also not red?

Solution

(a) If the first ball is red, what is the probability that the second is also red?
If the 1^{st} ball chosen is R, then there will now be one less red ball from which to choose.
There will also be one less ball in total from which to choose since the first ball was not replaced.
P(Second is Red given that the first ball is Red)
$= \frac{5-1}{12-1} = \frac{4}{11}$
(b) If the first ball is not red, what is the probability that the second is red?
If the 1^{st} ball chosen is **not** red, then the number of red balls remains the same (5), though the number of balls from which to choose is still one less than before.
P(Second is red given that the first is not red)
$= \frac{5}{12-1} = \frac{5}{11}$

Hence, we see the result of the 1st event affected the result of the next event.

This is an example of dependent events.

Dependent Events

Two events are said to be dependent if the result (whether it be an occurrence or a non-occurrence) of one affects the result of the other. In other words, they are not independent.

When solving problems involving dependent events, we must note that the sample space was reduced when we calculated the probability of an event when we had information on what happened before (for example when we selected the second ball we knew that a red was selected before).

Conditional probability

Before explaining this concept, let us revisit the concept of a sample space.

In an experiment, two dice are tossed (a red and a yellow). The sample space is shown below.

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

Let A = sum of the numbers is at least 9.
Let B = the two scores are the same (doubles)

The number of elements in the sample space, $S = 36$.
The numbers in the blue region of the table show the 10 outcomes in which the sum is 9 or more.

$P(\text{A}) = \dfrac{10}{36}$

The numbers highlighted in green are those whose scores on both dice are the same - (1,1), (2,2), (3,3), (4,4), (5,5) and (6,6).

$P(\text{B}) = \dfrac{6}{36}$

We now define **the conditional probability of B given A has occurred**, written as $P(\text{B}|\text{A})$ as the occurrence of the event B, knowing that A has occurred.

In other words, we are interested in the probability that the two scores are the same, given that their sum is at least 9. By simply looking at the sample space, we observe that two of the 10 possible scores whose sum is at least 9 belong to the set of repeated scores. Hence, the diagram above clearly shows this proportion as 2 out of 10.

In calculating the conditional probability of B given A has occurred, we used a reduced sample space from 36 to 10 by removing all the numbers whose sum is less than 10. (the numbers in the blue region of the table) because we applied a given condition - A has occurred.

We write,

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{2}{10}$$

We may think of the conditional probability of B, given A has occurred, as the proportion of A that is common to B.

Without the diagram, we can apply this formula to calculate the conditional probability as follows:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{2/36}{10/36} = \frac{2}{10}$$

where

$P(B \cap A) = \dfrac{2}{36}$ and $P(A) = \dfrac{10}{36}$

The conditional probability of an event *B* in relation to an event A is the probability that event B occurs given that event *A* has already occurred. The notation for conditional probability is *P(B|A),* read as *the probability of B given A*.

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

It is sometimes convenient to express the rule as follows:

$$P(B \cap A) = P(B|A) \times P(A)$$

If we apply this result to independent events, we obtain the following:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A) \times P(B)}{P(A)}$$

[For independent events, $P(B \cap A) = P(A) \times P(B)$]

Hence,

$$P(B|A) = P(B)$$

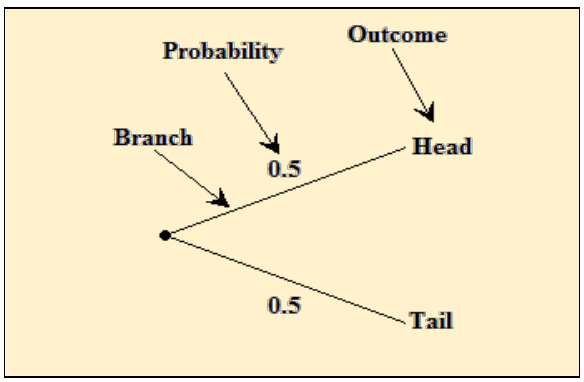
Since the occurrence or non-occurrence of A does not affect the probability of event B, this result makes sense.

Sample Space for a random variable

Probability tree diagrams

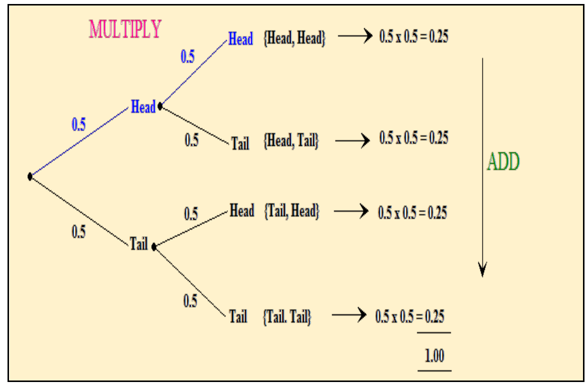
Calculating probabilities can often be simplified by using tree diagrams. Tree diagrams, as the name suggests, look like a tree, as they 'branch out' symmetrically. They are used to help us to visualise probability problems. The diagram is expected to show all the possible events. The first event is represented by a dot and from this dot branches are drawn to represent all the outcomes of the subsequent events.

The probability of each outcome is written on its branch. Here is a tree diagram for a single toss of a fair coin.



There are two "branches", heads and tails. The probability of each branch is written on the branch. The outcome is written at the end of the branch.

When a tree diagram is constructed, we calculate the overall probabilities quickly and simply. If we were to toss a coin twice, we can now extend the tree to include all events in the sample space, as shown below. The first branch has two outcomes, H or T with $P(H) = 0.5$ and $P(T) = 0.5$.



The second branch represents the toss of the second coin. There are four branches, each representing the following conditional probabilities - P(H|H), P(H|T), P(T|H) and P(T|T). The probability of each of the four outcomes occurring is computed on the right of the tree diagram by using the rule:

$$P(B \cap A) = P(B|A) \times P(A)$$

By looking at the diagram, one can observe the following:

- P(Head and Head) = 0.5 ×0.5 = 0.25.
- All probabilities add to 1.0. It is always wise to confirm this addition when a tree diagram is constructed.
- The probability of getting at least one Head from two tosses can be calculated in two ways as shown below.

$$P(T\ and\ H) + P(H\ and\ T) + P(H\ and\ H)$$
$$= 0.25 + 0.25 + 0.25 = 0.75$$

OR

$$1 - P(T\ and\ T) = 1 - (0.5 \times 0.5) = 0.75$$

General principles in working with tree diagrams

In working with tree diagrams

- We **multiply** probabilities **along the branches**.
- We **add** probabilities down **the columns**.

Example 8

A bag contains 3 black balls and 5 white balls. A ball is chosen at random and replaced. A second ball is then chosen.

a) Construct a tree diagram to illustrate the data.

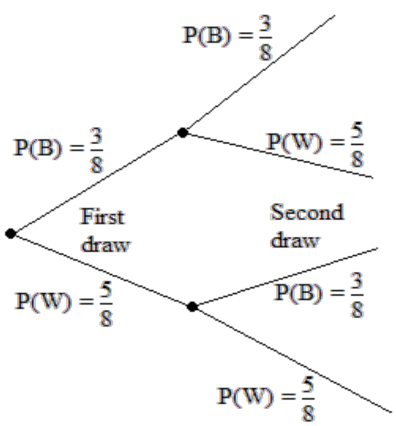
b) Calculate the probability that

   i. two black balls are chosen.

   ii. a black ball is chosen on the second pick.

Solution

a) Since the ball is replaced after the first pick, the events are independent.



(b) The probabilities of the possible outcomes are:

$P(B, B)=\frac{3}{8}\times\frac{3}{8}=\frac{9}{64}$ $\quad P(B, W)=\frac{3}{8}\times\frac{5}{8}=\frac{15}{64}$

$P(W, B)=\frac{5}{8}\times\frac{3}{8}=\frac{15}{64}$ $\quad P(W, W)=\frac{5}{8}\times\frac{5}{8}=\frac{25}{64}$

Total probability $=\frac{9}{64}+\frac{15}{64}+\frac{15}{64}+\frac{25}{64}=1$

Notice, the sum of the probabilities = 1, according to the law of total probabilities.

i) $P(B \text{ and } B)=\frac{3}{8}\times\frac{3}{8}=\frac{9}{64}$

ii) $P(\text{2nd is black})=P(B \text{ and } B) \; or \; P(W \text{ and } B)$

$=\frac{9}{64}+\frac{15}{64}=\frac{24}{64}=\frac{3}{8}$

Example 9

The probability of hiring a taxi from garage *A*, *B* or *C* is $P(A)=0.3$, $P(B)=0.5$ and $P(C)=0.2$. The probability that the taxi arrives late from each garage is given as: $P(\text{Late from } A)=0.07$,

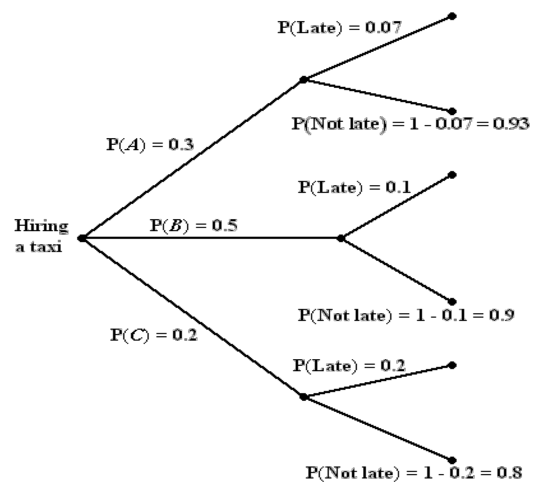
$P(\text{Late from } B)=0.1$ and

$P(\text{Late from } C)=0.2$

(i) Illustrate the information given on a tree diagram

(ii) Determine the probability that a taxi chosen at random will come from garage *C* given that it is late.

Solution



(i) The probability that a taxi, chosen at random, will come from garage *C*, given that it is late.

$P(\text{Taxi will arrive late})$

$=P(A \text{ is chosen and taxi is late})$

or $P(B \text{ is chosen and taxi is late})$

or $P(C \text{ is chosen and taxi is late})$

$=(0.3\times0.07)+(0.5\times0.1)+(0.2\times0.2)$

$=\frac{111}{1000}$

(ii) $P(\text{Taxi comes from } C \text{ given that it is late})$

Let *C* be the event that a taxi comes from garage *C*.
Let *L* be the event that the taxi arrives line.
Required to calculate $P(C/L)$

$$P(C/L)=\frac{P(C\cap L)}{P(L)} \qquad \text{(Conditional probability)}$$

$$P(C\cap L)=P(\text{Taxi is late and comes from } C)$$

$$=0.2\times 0.2$$

$$P(L)=0.111 \text{ (from above)}$$

$$P(C/L)=\frac{0.2\times 0.2}{0.111}$$

$$=\frac{40}{111}$$

The probability that a taxi chose at random, will come from garage C given that it is late is $\frac{40}{111}$.