CHAPTER 10

Probability

CONTENTS

1 Revision **1**

1.1 Terminology . . . 1

1.2 Identities . . . 2

1.3 Venn Diagrams . . . 4

2 Dependent and independent events **10**

3 More Venn diagrams **17**

4 Tree diagrams **23**

5 Contingency tables **29**

6 Summary **32**

7 Exercises **34**

7.1 Exercise 1 . . . 34

7.2 Exercise 2 . . . 34

7.3 Exercise 3 . . . 35

7.4 Exercise 4 . . . 37

7.5 Exercise 5 . . . 37

7.6 Exercise 6 . . . 38

8 Answers for Exercises **39**

8.1 Exercise 2 . . . 39

8.2 Exercise 3 . . . 39

8.3 Exercise 4 . . . 40

8.4 Exercise 5 . . . 40

August 26, 2021

1 REVISION

1.1 Terminology

Outcome: a single observation of an uncertain or random process (called an experiment). For example, when you accidentally drop a book, it might fall on its cover, on its back or on its side. Each of these options is a possible outcome.

Sample space of an experiment: the set of all possible outcomes of the experiment. For example, the sample space when you roll a single 6 sided die is the set $\{1; 2; 3; 4; 5; 6\}$. For a given experiment, there is exactly one sample space. The sample space is denoted by the letter S.

Event: a set of outcomes of an experiment. For example, during radioactive decay of 1 gram of uranium-234 , one possible event is that the number of alpha-particles emitted during 1 microsecond is between 225 and 235.

Probability of an event: a real number between 0 and 1 that describes how likely it is that the event will occur. A probability of 0 means the outcome of the experiment will never be in the event set. A probability of 1 means the outcome of the experiment will always be in the event set. When all possible outcomes of an experiment have equal chance of occurring, the probability of an event is the number of outcomes in the event set as a fraction of the number of outcomes in the sample space.

Relative frequency of an event: the number of times that the event occurs during experimental trials, divided by the total number of trials conducted. For example, if we flip a coin 10 times and it landed on heads 3 times, then the relative frequency of the heads event is $\frac{3}{10} = 0,3$.

Union of events: the set of all outcomes that occur in at least one of the events,For 2 events called A and B, we write the union as "A or B".Another way of writing the union is using set notation:$A \cup B$.

Intersection of events: the set of all outcomes that occur in all of the events. For 2 events called A and B,we write the union as "A or B". .Another way of writing the union is using set notation:$A \cap B$.

Mutually exclusive events events with no outcomes in common, that is (A and B)= ∅.Mutually exclusive events can never occur simultaneously. For example the event that a number is even and the event that the same number is odd are mutually exclusive, since a number can never be both even and odd.

Complementary events two mutually exclusive events that together contain all the outcomes in the sample space. For an event called A we write the complement as "not A".Another way of writing the complement is as A'.

1.2 Identities

The **addition rule** (also called the sum rule) for any 2 events, A and B, A and B is.

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

This rule relates the probabilities of 2 events with the probabilities of their union and intersection.

The **addition rule** 2 **mutually exclusive events** is.

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

This rule is a special case of the previous rule. Because the events are mutually exclusive, $P(A \text{ and } B) = 0$.

The **complementary rule** is.

$$P(\text{ not } A) = 1 - P(A)$$

This rule is a special case of the previous rule. Since A and $(\text{not } A)$ are mutually exclusive, $P(A(\text{Not } A)) = 1$.

WORKED EXAMPLE 1: EVENTS

QUESTION

You take all the hearts from a deck of cards. You then select a random card from the set of hearts. What is the sample space? What is the probability of each of the following events?

1. The card is the ace of hearts.
2. The card has a prime number on it.
3. The card has a letter of the alphabet on it.

SOLUTION

Step 1: Write down the sample space

Since we are considering only one suit from the deck of cards (the hearts), we need to write down only the letters and numbers on the cards. Therefore the sample space is

$$S = \{\text{A}; 2; 3; 4; 5; 6; 7; 8; 9; 10; \text{J}; \text{Q}; \text{K}\}$$

Step 2: Write down the event sets

- ace of hearts: A
- prime number: $2; 3; 5; 7$
- letter of alphabet: $A; J; Q; K$

Step 3: Compute the probabilities
The probability of an event is defined as the number of elements in the event set divided by the number of elements in the sample space. There are 13 elements in the sample space. So the probability of each event is

- ace of hearts: $\frac{1}{13}$
- prime number: $\frac{4}{13}$
- letter of alphabet: $\frac{4}{13}$

WORKED EXAMPLE 2: EVENTS

QUESTION
You roll two 6-sided dice. Let E be the event that the total number of dots on the dice is 10. Let F be the event that at least one die is a 3

1. Write down the event sets for E and F.
2. Determine the probabilities for E and F.
3. Are E and F mutually exclusive? Why or why not?

SOLUTION
Step 1: Write down the sample space The sample space of a single 6-sided die is just $1; 2; 3; 4; 5; 6$. To get the sample space of two 6-sided dice, we have to take every possible pair of numbers from 1 to 6.

$$S = \left\{ \begin{array}{cccccc} (1;1) & (1;2) & (1;3) & (1;4) & (1;5) & (1;6) \\ (2;1) & (2;2) & (2;3) & (2;4) & (2;5) & (2;6) \\ (3;1) & (3;2) & (3;3) & (3;4) & (3;5) & (3;6) \\ (4;1) & (4;2) & (4;3) & (4;4) & (4;5) & (4;6) \\ (5;1) & (5;2) & (5;3) & (5;4) & (5;5) & (5;6) \\ (6;1) & (6;2) & (6;3) & (6;4) & (6;5) & (6;6) \end{array} \right\}$$

Step 2: Write down the events

For E the dice have to add to 10.

$$E = \{(4;6);(5;5);(6;4)\}$$

For F at least one die has to be 3.

$$F = \{(1;3);(3;1);(2;3);(3;2);(3;3);(4;3);(3;4);(5;3);(3;5);(6;3);(3;6)\}$$

Step 3: Compute the probabilities

The probability of an event is defined as the number of elements in the event set divided by the number of elements in the sample space. There are

- $6 \times 6 = 36$ outcomes in the sample space, S;
- 3 outcomes in event E; and
- 11 outcomes in event F.

Therefore

$$P(E) = \frac{3}{36} = \frac{1}{12}$$

and

$$P(F) = \frac{11}{36}$$

Step 4: Are they mutually exclusive

To test whether two events are mutually exclusive, we have to test whether their intersection is empty. Since E has no outcomes that contain a 3 on one of the dice, the intersection of E and F is empty: $(E \text{ and } F) = \emptyset$. This means that the events are mutually exclusive.

1.3 Venn Diagrams

A Venn diagram is used to show how events are related to one another. A Venn diagram can be very helpful when doing calculations with probabilities. In a Venn diagram each event is represented by a shape, often a circle or a rectangle. The region inside the shape represents the outcomes included in the event and the region outside the shape represents the outcomes that are not in the event.

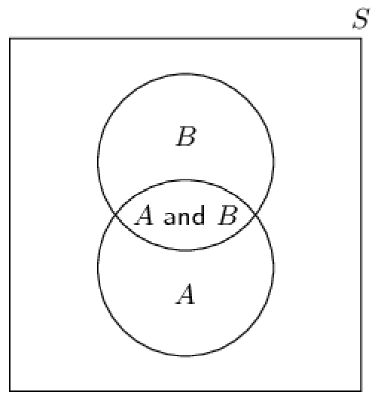


Figure 1: A Venn diagram representing a sample space, S , as a square; and two events, A and B, as circles. The intersection of the two circles contains outcomes that are in both A and B

Venn diagrams can be used in slightly different ways and it is important to notice the differences between them. The following 3 examples show how a Venn diagram is used to represent.

- the outcomes included in each event;
- the number of outcomes in each event; and
- the probability of each event.

WORKED EXAMPLE 3: VENN DIAGRAMS WITH OUTCOMES

QUESTION

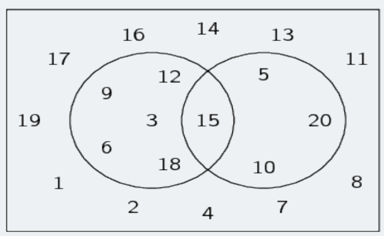
Choose a number between 1 and 20. Draw a Venn diagram to answer the following questions.

1. What is the probability that the number is a multiple of 3?
2. What is the probability that the number is a multiple of 5?
3. What is the probability that the number is a multiple of 3 or 5?
4. What is the probability that the number is a multiple of 3 and 5?

SOLUTION

Step 1: Draw a Venn diagram

The Venn diagram should show the sample space of all numbers from 1 to 20.It should also show an event set that contains all the multiples of 3, let $A = 3; 6; 9; 12; 15; 18$, and another event set that contains all the multiples of 5 , let $B = 5; 10; 15; 20$. Note that there is one shared outcome between these two events, namely 15.



Step 2: Compute probabilities The probability of an event is the number of outcomes in the event set divided by the number of outcomes in the sample space.There are 20 outcomes in the sample space.

1. Since there are 6 outcomes in the multiples of 3 event set, the probability of a multiple of 3 is $P(A) = \frac{6}{20} = \frac{3}{10}$.
2. Since there are 4 outcomes in the multiples of 5 event set, the probability of a multiple of 5 is $P(B) = \frac{4}{20} = \frac{1}{5}$.
3. The event that the number is a multiple of 3 or 5 is the union of the above two event sets. There are 9 elements in the union of the event sets, so the probability is $\frac{9}{20}$.
4. The event that the number is a multiple of 3 and 5 is the intersection of the two event sets. There is 1 element in the intersection of the event sets, so the probability is $\frac{1}{20}$.

MathU Teaching® Emporium

WORKED EXAMPLE 4: VENN DIAGRAMS WITH COUNTS

QUESTION

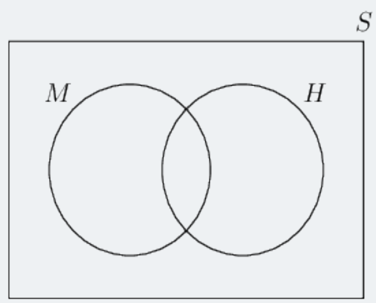
In a group of 50 learners, 35 take Mathematics and 30 take History, while 12 take neither of the two subjects. Draw a Venn diagram representing this information. If a learner is chosen at random from this group, what is the probability that he takes both Mathematics and History?

SOLUTION

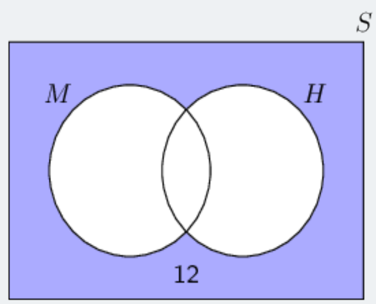
Step 1: Draw outline of Venn diagram There are 2 events in this question, namely

- M: that a learner takes Mathematics; and
- H: that a learner takes History.

We need to do some calculations before drawing the full Venn diagram, but with the information above we can already draw the outline.



Step 2: Write down sizes of the event sets, their union and intersection We are told that 12 learners take neither of the two subjects. Graphically we can represent this as:

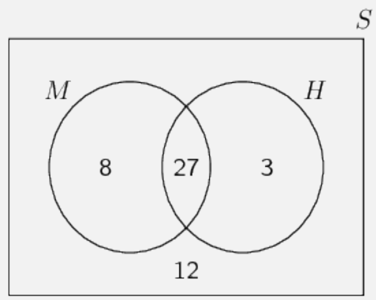


Since there are 50 elements in the sample space, we can see from this figure that there are 50−12=38 elements in (M or H). So far we know

- $n(M) = 35$
- $n(H) = 30$
- $n(M \text{ or } H) = 38$

$$n(M \text{ or } H) = n(M) + n(H) - n(M \text{ and } H)$$
$$\therefore n(M \text{ and } H) = 35 + 30 - 38$$
$$= 27$$

Step 3: Draw the final Venn diagram



WORKED EXAMPLE 5: VENN DIAGRAMS WITH PROBABILITIES

QUESTION

Draw a Venn diagram to represent the same information as in the previous example, except showing the probabilities of the different events, rather than the counts.

If a learner is chosen at random from this group, what is the probability that she takes both Mathematics and History?

SOLUTION

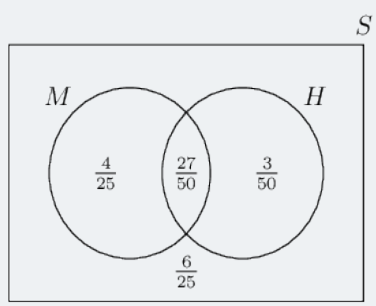
Step 1: Use counts to compute probabilities

Since there are 50 elements (learners) in the sample space, we can compute the probability of any event by dividing the size of the event set by 50. This gives the following probabilities:

- $P(M) = \frac{35}{50} = \frac{7}{10}$
- $P(H) = \frac{30}{50} = \frac{3}{5}$
- $P(M \text{ or } H) = \frac{38}{50} = \frac{19}{25}$
- $P(M \text{ and } H) = \frac{27}{50}$

Step 2: Draw the Venn diagram

Next we replace each count from the Venn diagram in the previous example with a probability.



Step 3: Find the answer

The probability that a random learner will take both Mathematics and History is $P(M \text{ and } H) = \frac{27}{50}$

2 DEPENDENT AND INDEPENDENT EVENTS

Sometimes the presence or absence of one event tells us something about other events. We call events dependent if knowing whether one of them happened tells us something about whether the others happened. Independent events give us no information about one another; the probability of one event occurring does not affect the probability of the other events occurring.

Definition

Independent events

Two events, A and B are independent if and only if

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

At first it might not be clear why we should call events that satisfy the equation above independent. We will explore this further using a number of examples.

Investigation

Independence

Roll a single 6-sided die and consider the following two events:

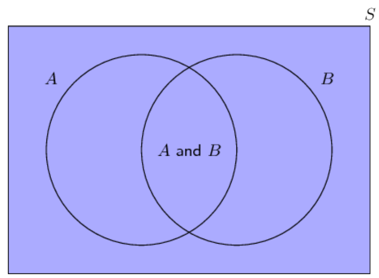
- E: you get an even number
- T: you get a number that is divisible by three

Now answer the following questions:

- What is the probability of E?
- What is the probability of getting an even number if you are told that the number was also divisible by three?
- Does knowing that the number was divisible by 3 change the probability that the number was even?

Are the events E and T dependent or independent according to the definition? (Hint: compute the probabilities in the definition of independence.)

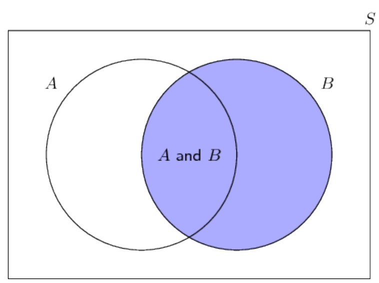
So, why do we call it **independence** when $P(A \text{ and } B) = P(A)+P(B)$? For two events, A and B, independence means that knowing the outcome of B does not affect the probability of A.
Consider the following Venn diagram. The probability of A is the ratio between the number of outcomes in A



and the number of outcomes in the sample space, S.

$$P(A) = \frac{n(A)}{n(S)}$$

Now, let's say that we **know** that event B happened. How does this affect the probability of A? Here is how the Venn diagram changes:



A lot of the possible outcomes (all of the outcomes outside B) are now out of the picture, because we know that they did not happen. Now the probability of A happening, given that we know that B happened, is the ratio between the size of the region where A is present (A and B) and the size of all possible events (B).

$$P(A \text{ if we know } B) = \frac{n(A \text{ and } B)}{n(B)}$$

If $P(A) = P(A$ if we know B) we call them independent, because knowing B does not change the probability of A.
With some algebra, we can prove that this statement of independence is the same as the definition of independence that we saw at the beginning of this section. For independent events

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

This is equivalent to

$$\begin{aligned} P(A) &= P(A \text{ and } B) \div P(B) \\ &= \frac{n(A \text{ and } B)}{n(S)} \div \frac{n(B)}{n(S)} \\ &= \frac{n(A \text{ and } B)}{n(B)} \\ &= P(A \text{ if we know } B) \end{aligned}$$

That is why we call events independent!

(For enrichment only):

The ratio

$$\frac{P(A \text{ and } B)}{P(B)}$$

is called a **conditional probability** and written using the notation $P(A|B)$. This notation is read as "the probability of A given B."

If (and only if) A and B are independent: P(A|B)=P(A) and P(B|A)=P(B). Try to prove this using the definition of independence.

WORKED EXAMPLE 6: INDEPENDENT AND DEPENDENT EVENTS

QUESTION

A bag contains 5 red and 5 blue balls. We remove a random ball from the bag, record its colour and put it back into the bag. We then remove another random ball from the bag and record its colour.

1. What is the probability that the first ball is red?
2. What is the probability that the second ball is blue?
3. What is the probability that the first ball is red and the second ball is blue?
4. Are the first ball being red and the second ball being blue independent events?

SOLUTION

Step 1: Probability of a red ball first Since there are a total of 10 balls, of which 5 are red, the probability of getting a red ball is

$$P(\text{ first ball red }) = \frac{5}{10} = \frac{1}{2}$$

Step 2: Probability of a blue ball second The problem states that the first ball is placed back into the bag before we take the second ball. This means that when we draw the second ball, there are again a total of 10 balls in the bag, of which 5 are blue. Therefore the probability of drawing a blue ball is

$$P(\text{ second ball blue }) = \frac{5}{10} = \frac{1}{2}$$

Step 3: Probability of red first and blue second When drawing two balls from the bag, there are 4 possibilities. We can get.

- a red ball and then another red ball;
- a red ball and then a blue ball;
- a blue ball and then a red ball;
- a blue ball and then another blue ball.

We want to know the probability of the second outcome, where we have to get a red ball first. Since there are 5 red balls and 10 balls in total, there are $\frac{5}{10}$ ways to get a red ball first. Now we put the first ball back, so there are again 5 red balls and 5 blue balls in the bag. Therefore there are $\frac{5}{10}$ ways to get a blue ball second if the first ball was red. This means that there are

$$\frac{5}{10} \times \frac{5}{10} = \frac{25}{100}$$

ways to get a red ball first and a blue ball second. So, the probability of getting a red ball first and a blue ball second is $\frac{1}{4}$.

Step 4: Dependent or independent?

According to the definition, events are independent if and only if

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

In this problem:

- $P(\text{first ball red}) = \frac{1}{2}$
- $P(\text{second ball red}) = \frac{1}{2}$
- $P(\text{first ball red and second ball blue}) = \frac{1}{4}$

Since $\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$, the events are independent.

WORKED EXAMPLE 7: INDEPENDENT AND DEPENDENT EVENTS

QUESTION

In the previous example, we picked a random ball and put it back into the bag before continuing. This is called **sampling with replacement**. In this example, we will follow the same process, except that we will not put the first ball back into the bag. This is called **sampling without replacement**.

So, from a bag with 5 red and 5 blue balls, we remove a random ball and record its colour. Then, without putting back the first ball, we remove another random ball from the bag and record its colour.

1. What is the probability that the first ball is red?
2. What is the probability that the second ball is blue?
3. What is the probability that the first ball is red and the second ball is blue?
4. Are the first ball being red and the second ball being blue independent events?

SOLUTION

Step 1: Count the number of outcomes We will look directly at the number of possible ways in which we can get the 4 possible outcomes when removing 2 balls. In the previous example, we saw that the 4 possible outcomes are

- a red ball and then another red ball;
- a red ball and then a blue ball;
- a blue ball and then a red ball;
- a blue ball and then another blue ball.

For the first outcome, we have to get a red ball first. Since there are 5 red balls and 10 balls in total, there are $\frac{5}{10}$ ways to get a red ball first. After we have taken out a red ball, there are now 4 red balls and 5 blue balls left. Therefore there are $\frac{4}{9}$ ways to get a red ball second if the first ball was also red. This means that there are

$$\frac{5}{10} \times \frac{4}{9} = \frac{20}{90}$$

ways to get a red ball first and a red ball second. The probability of the first outcome is $\frac{2}{9}$.

For the second outcome, we have to get a red ball first.As in the first outcome, there are $\frac{5}{10}$ ways to get a red ball first; and there are now 4 red balls and 5 blue balls left. Therefore there are $\frac{5}{9}$ ways to get a blue ball second if the first ball was red. This means that there are

$$\frac{5}{10} \times \frac{5}{9} = \frac{25}{90}$$

ways to get a red ball first and a blue ball second. The probability of the second outcome is $\frac{5}{18}$.

We can compute the probabilities of the third and fourth outcomes in the same way as the first two, but there is an easier way. Notice that there are only 2 types of ball and that there are exactly equal numbers of them at the start. This means that the problem is completely symmetric in red and blue. We can use this symmetry to compute the probabilities of the other two outcomes.

In the third outcome, the first ball is blue and the second ball is red. Because of symmetry this outcome must have the same probability as the second outcome (when the first ball is red and the second ball is blue). Therefore the probability of the third outcome is $\frac{5}{18}$.

In the fourth outcome, the first and second balls are both blue. From symmetry, this outcome must have the same probability as the first outcome (when both balls are red). Therefore the probability of the fourth outcome is $\frac{2}{9}$.

To summarise, these are the possible outcomes and their probabilities:

- first ball red and second ball red: $\frac{2}{9}$;
- first ball red and second ball blue: $\frac{5}{18}$;
- first ball blue and second ball red: $\frac{5}{18}$;
- first ball blue and second ball blue: $\frac{2}{9}$;

Step 2: Probability of a red ball first

To determine the probability of getting a red ball on the first draw, we look at all of the outcomes that contain a red ball first. These are

- a red ball and then another red ball;
- a red ball and then a blue ball.

The probability of the first outcome is $\frac{2}{9}$ and the probability of the second outcome is $\frac{5}{18}$. By adding these two probabilities, we see that the probability of getting a red ball first is

$$P(\text{ first ball red }) = \frac{2}{9} + \frac{5}{18} = \frac{1}{2}$$

This is the same as in the previous exercise, which should not be too surprising since the probability of the first ball being red is not affected by whether or not we put it back into the bag before drawing the second ball.

Step 3: Probability of a blue ball second

To determine the probability of getting a blue ball on the second draw, we look at all of the outcomes that contain a blue ball second. These are

- a red ball and then a blue ball;
- a blue ball and then another blue ball.

The probability of the first outcome is $\frac{5}{18}$and the probability of the second outcome is $\frac{2}{9}$. By adding these two probabilities, we see that the probability of getting a blue ball second is

$$P(\text{ second ball blue }) = \frac{5}{18} + \frac{2}{9} = \frac{1}{2}$$

This is also the same as in the previous exercise! You might find it surprising that the probability of the second ball is not affected by whether or not we replace the first ball. The reason why this probability is still $\frac{1}{2}$is that we are computing the probability that the second ball is blue without knowing the colour of the first ball. Because there are only two equal possibilities for the second ball (red and blue) and because we don't know whether the first ball is red or blue, there is an equal chance that the second ball will be one colour or the other.

Step 4: Probability of red first and blue second

We have already calculated the probability that the first ball is red and the second ball is blue. It is $\frac{5}{18}$.

Step 5: Dependent or independent?

According to the definition, events are independent if and only if

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

In this problem:

- $P(\text{ first ball red }) = \frac{1}{2}$
- $P(\text{ second ball blue }) = \frac{1}{2}$
- $P(\text{ first ball red and second ball blue }) = \frac{5}{18}$

Since $\frac{5}{18} \neq 12 \times 12$, the events are dependent.

WARNING

Just because two events are mutually exclusive does not necessarily mean that they are independent. To test whether events are mutually exclusive, always check that $P(AandB) = 0$. To test whether events are independent, always check that $P(AandB) = P(A) \times P(B)$. See the exercises below for examples of events that are mutually exclusive and independent in different combinations.

3 MORE VENN DIAGRAMS

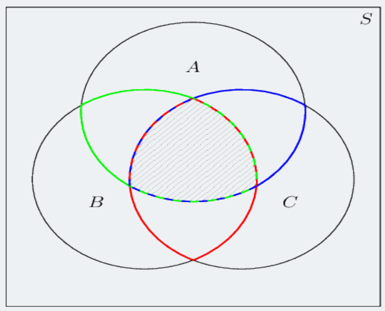
In the rest of this chapter we will look at tools and techniques for working with probability problems.

When working with more complex problems, we can have three or more events that intersect in various ways. To solve these problems, we usually want to count the number (or percentage) of outcomes in an event, or a combination of events. Venn diagrams are a useful tool for recording and visualising the counts.

INVESTIGATION

Venn diagram for 3 events

The diagram below shows a general Venn diagram for 3 events.



Write down the sets corresponding to each of the three coloured regions and also to the shaded region. Remember that the intersections between circles represent the intersections between the different events.

What is the event for

- the red region;
- the green region;
- the blue region; and
- the shaded region?

WORKED EXAMPLE 8: VENN DIAGRAM FOR 3 EVENTS

QUESTION

Draw a Venn diagram that shows the following sample space and events:

- S: all the integers from 1 to 30
- P : prime numbers
- M : multiples of 3
- F : factors of 30

SOLUTION

Step 1: Write down the sample space and event sets

The sample space contains all the positive integers up to 30.

$$S = \{1; 2; 3; \ldots; 30\}$$

The prime numbers between 1 and 30 are

$$P = \{2; 3; 5; 7; 11; 13; 17; 19; 23; 29\}$$

The multiples of 3 between 1 and 30 are

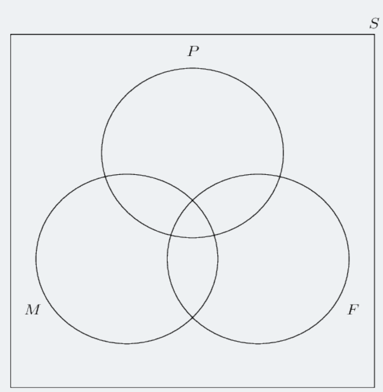
$$M = \{3; 6; 9; 12; 15; 18; 21; 24; 27; 30\}$$

The factors of 30 are

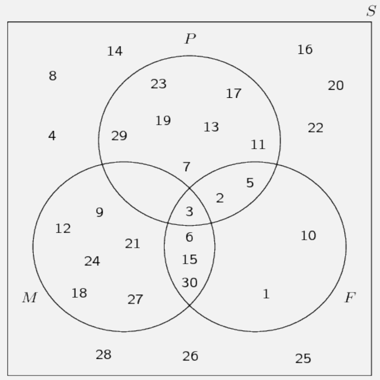
$$F = \{1; 2; 3; 5; 6; 10; 15; 30\}$$

Step 2: Draw the outline of the Venn diagram

There are 3 events, namely P, M and F, and the sample space, S. Put this information on a Venn diagram:



Step 3: Place the outcomes in the appropriate event sets



WORKED EXAMPLE 9: VENN DIAGRAM FOR 3 EVENTS

QUESTION

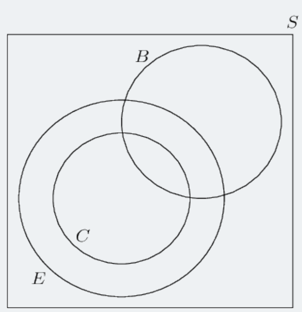
At Dawnview High there are 400 Grade 11 learners. 270 do Computer Science, 300 do English and 50 do Business studies. All those doing Computer Science do English, 20 take Computer Science and Business studies and 35 take English and Business studies. Using a Venn diagram, calculate the probability that a pupil drawn at random will take:

1. English, but not Business studies or Computer Science
2. English but not Business studies
3. English or Business studies but not Computer Science
4. English or Business studies

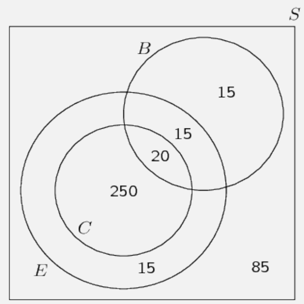
SOLUTION

Step 1: Draw the outline of the Venn diagram

We need to be careful with this problem. In the question statement we are told that all the learners who do Computer Science also do English. This means that the circle for Computer Science on the Venn diagram needs to be inside the circle for English.

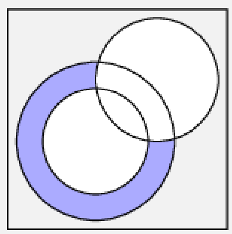


Step 2: Fill in the counts on the Venn diagram



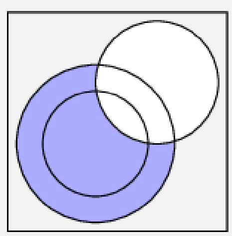
Step 3: Compute probabilities

To find the number of learners taking English, but not Business studies or Computer Science, we need to look at this region of the Venn diagram:



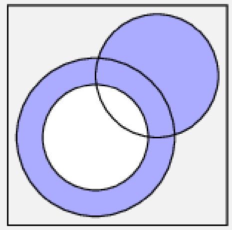
The count in this region is 15 and there are a total of 400 learners in the grade. Therefore the probability that a learner will take English but not Business studies or Computer Science is $\frac{15}{400} = \frac{3}{80}$.

To find the number of learners taking English but not Business studies, we need to look at this region of the Venn diagram:



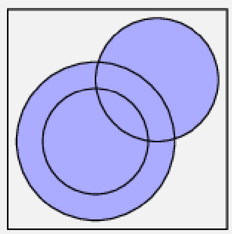
The count in this region is 265. Therefore the probability that a learner will take English but not Business studies is $\frac{265}{400} = \frac{53}{80}$.

To find the number of learners taking English or Business studies but not Computer Science, we need to look at this region of the Venn diagram:



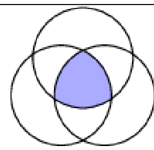
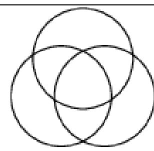
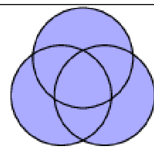
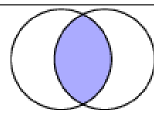
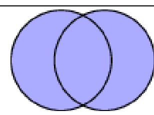
The count in this region is 45. Therefore the probability that a learner will take English or Business studies but not Computer Science is $\frac{45}{400} = \frac{9}{80}$.

To find the number of learners taking English or Business studies, we need to look at this region of the Venn diagram:



The count in this region is 315. Therefore the probability that a learner will take English or Business studies is $\frac{315}{400} = \frac{63}{80}$.

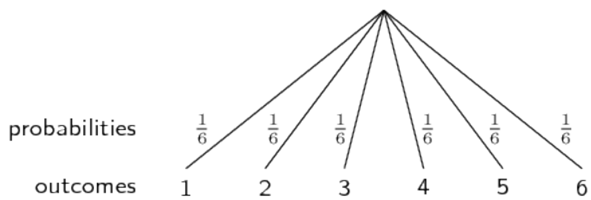
There are some words that tell you which part of the Venn diagram should be filled in. The following table summarises the most important ones:

Words	Symbols	Venn diagram
"all"	A and B and $C / A \cap B \cap C$	
"none"		
"at least one"	A or B or $C / A \cup B \cup C$	
"both A and B"	A and $B / A \cap B$	
"A or B"	A or $B / A \cup B$	

4 TREE DIAGRAMS

Tree diagrams are useful for organising and visualising the different possible outcomes of a sequence of events. For each possible outcome of the first event, we draw a line where we write down the probability of that outcome and the state of the world if that outcome happened. Then, for each possible outcome of the second event we do the same thing.

Below is an example of a simple tree diagram, showing the possible outcomes of rolling a 6-sided die. Note



that each outcome (the numbers 1 to 6) is shown at the end of a line; and that the probability of each outcome (all $\frac{1}{6}$ in this case) is shown shown on a line. The probabilities have to add up to 1 in order to cover all of the possible outcomes. In the examples below, we will see how to draw tree diagrams with multiple events and how to compute probabilities using the diagrams.

Earlier in this chapter you learned about dependent and independent events. Tree diagrams are very helpful for analysing dependent events. A tree diagram allows you to show how each possible outcome of one event affects the probabilities of the other events.

Tree diagrams are not so useful for independent events since we can just multiply the probabilities of separate events to get the probability of the combined event. Remember that for independent events:

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

So if you already know that events are independent, it is usually easier to solve a problem without using tree diagrams. But if you are uncertain about whether events are independent or if you know that they are not, you should use a tree diagram.

WORKED EXAMPLE 10: DRAWING A TREE DIAGRAM

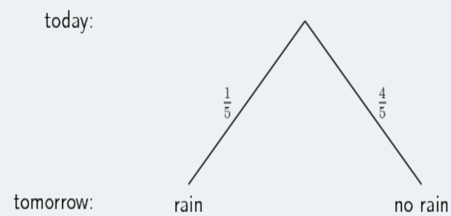
QUESTION

If it rains on a given day, the probability that it rains the next day is $\frac{1}{3}$. If it does not rain on a given day, the probability that it rains the next day is $\frac{1}{6}$. The probability that it will rain tomorrow is $\frac{1}{5}$. What is the probability that it will rain the day after tomorrow? Draw a tree diagram of all the possibilities to determine the answer.

SOLUTION

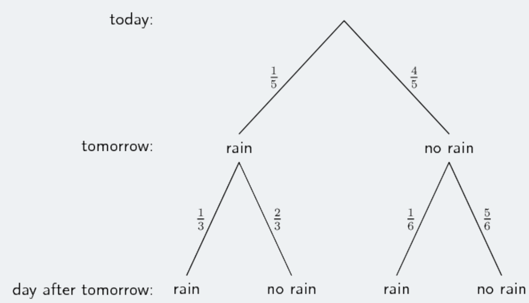
Step 1: Draw the first level of the tree diagram

Before we can determine what happens on the day after tomorrow, we first have to determine what might happen tomorrow. We are told that there is a $\frac{1}{5}$ probability that it will rain tomorrow. Here is how to represent this information using a tree diagram:



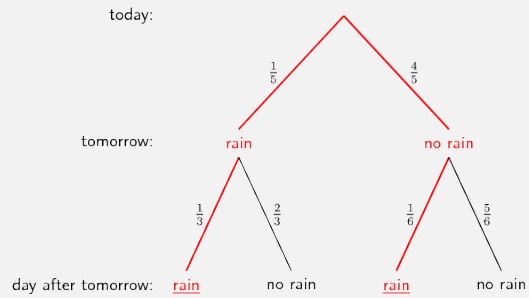
Step 2: Draw the second level of the tree diagram

We are also told that if it **does** rain on one day, there is a $\frac{1}{3}$ probability that it will also rain on the following day. On the other hand, if it **does not** rain on one day, there is only a $\frac{1}{6}$ probability that it will also rain on the following day. Using this information we complete the tree diagram:



Step 3: Compute the probability

We are asked what the probability is that it will rain the day after tomorrow. On the tree diagram above we can see that there are 2 situations where it rains on the day after tomorrow. They are marked in red below.



To get the probability for the first situation (that it rains tomorrow and the day after tomorrow) we have to multiply the probabilities along the first red line.

$$P(\text{ rain tomorrow and rain day after tomorrow}) \\ = \frac{1}{5} \times \frac{1}{3} \\ = \frac{1}{15}$$

To get the probability for the second situation (that it does not rain tomorrow, but it does rain the day after tomorrow) we have to multiply the probabilies along the second red line.

$$P(\text{ not rain tomorrow and rain day after tomorrow}) \\ = \frac{4}{5} \times \frac{1}{6} \\ = \frac{2}{15}$$

Therefore the total probability that it will rain the day after tomorrow is the sum of the probabilities along the two red paths, namely

$$\frac{1}{15} + \frac{2}{15} = \frac{1}{5}$$

WORKED EXAMPLE 11: DRAWING A TREE DIAGRAM

QUESTION

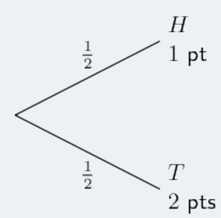
You play the following game. You flip a coin. If it comes up tails, you get 2 points and your turn ends. If it comes up heads, you get only 1 point, but you can flip the coin again. If you flip the coin multiple times in one turn, you add up the points. You can flip the coin at most 3 times in one turn. What is the probability that you will get exactly 3 points in one turn? Draw a tree diagram to visualise the different possibilities.

SOLUTION

Step 1: Write down the events and their symbols

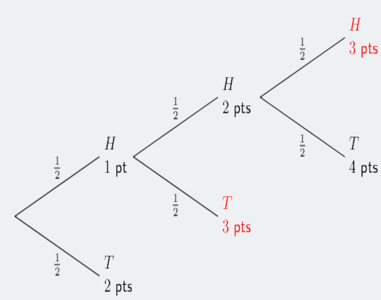
Each coin toss has on of two possible outcomes, namely heads (H) and tails (T). Each outcome has a probability of $\frac{1}{2}$. We are asked to count the number of points, so we will also indicate how many points we have for each outcome.

Step 2: Draw the first level of the tree diagram



This tree diagram shows the possible outcomes after 1 flip of the coin. Remember that we can have up to 3 flips, so the diagram is not complete yet. If the coin comes up heads, we flip the coin again. If the coin comes up tails, we stop.

Step 3: Draw the second and third level of the tree diagram



In this tree diagram you can see that we add up the points we get with each coin flip. After three coin flips, the game is over.

Step 4: Find the relevant outcomes and compute the probability

We are interested in getting exactly 3 points during the game. To find these outcomes we look only at the tips of the tree. We end with exactly 3 points when the coin flips are

- $(H;T)$ with probability $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$;
- $(H;H;H)$ with probability $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

Notice that we compute the probability of an outcome by multiplying all the probabilities along the path from the start of the tree to the tip where the outcome is. We add the above two probabilities to obtain the final probability of getting exactly 3 points as $\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$

WORKED EXAMPLE 12: DRAWING A TREE DIAGRAM

QUESTION

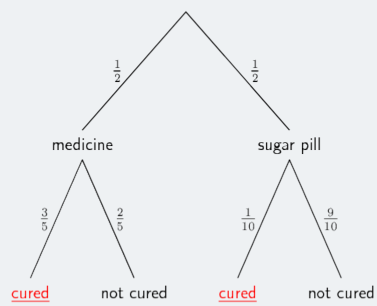
A person takes part in a medical trial that tests the effect of a medicine on a disease. Half the people are given medicine and the other half are given a sugar pill, which has no effect on the disease. The medicine has a 60% chance of curing someone. But, people who do not get the medicine still have a 10% chance of getting well. There are 50 people in the trial and they all have the disease. Talwar takes part in the trial, but we do not know whether he got the medicine or the sugar pill. Draw a tree diagram of all the possible cases. What is the probability that Talwar gets cured?

SOLUTION

Step 1: Summarise the information in the problem

There are two uncertain events in this problem. Each person either receives medicine (probability $\frac{1}{2}$) or a sugar pill (probability $\frac{1}{2}$). Each person also gets cured (probability $\frac{3}{5}$ with medicine and $\frac{1}{10}$ without) or stays ill (probability $\frac{2}{5}$ with medicine and $\frac{9}{10}$ without).

Step 2: Draw the tree diagram



In the first level of the tree diagram we show that Talwar either gets the medicine or the sugar pill. The second level of the tree diagram shows whether Talwar is cured or not, depending on which one of the pills he got.

Step 3: Compute the required probability

We multiply the probabilites along each path in the tree diagram that leads to Talwer being cured:

$$\frac{1}{2} \times \frac{3}{5} = \frac{3}{10}$$
$$\frac{1}{2} \times \frac{1}{10} = \frac{1}{20}$$

We then add these probabilities to get the final answer. The probability that Talwar is cured is $\frac{7}{20}$.

5 CONTINGENCY TABLES

A contingency table is another tool for keeping a record of the counts or percentages in a probability problem. Contingency tables are especially helpful for figuring out whether events are dependent or independent.

We will be studying two-way contingency tables, where we count the number of outcomes for 2 events and their complements, making 4 events in total. A two-way contingency table always shows the counts for the 4 possible combinations of events, as well as the totals for each event and its complement. We can use a contingency table to compute the probabilities of various events by computing the ratios between counts, and to determine whether the events are dependent or independent. The example below shows a two-way contingency table, representing the outcome of a medical study.

WORKED EXAMPLE 13: CONTINGENCY TABLES

QUESTION

A medical trial into the effectiveness of a new medication was carried out. 120 females and 90 males took part in the trial. Out of those people, 50 females and 30 males responded positively to the medication. Given below is a contingency table with the given information filled in.

	Female	**Male**	**Total**
Positive	50	30	
Negative			
Totals	120	90	

1. What is the probability that the medicine gives a positive result for females?
2. What is the probability that the medicine gives a negative result for males?
3. Was the medication's success independent of gender? Explain.

SOLUTION

Step 1: Complete the contingency table

The best place to start is always to complete the contingency table. Because the each column has to sum up to its total, we can work out the number of females and males who responded negatively to the medication. Then we can add each row to get the totals on the right hand side of the table.

	Female	**Male**	**Total**
Positive	50	30	80
Negative	70	60	130
Totals	120	90	210

Step 2:Compute the required probabilities

The way the first question is phrased, we need to determine the probability that a person responds positively if she is female. This means that we do not include males in this calculation. So, the probability that the medicine gives a positive result for females is the ratio between the number of females who got a positive response and the total number of females.

$$\begin{aligned} P(\text{ positive if female }) &= \frac{n(\text{ positive and female })}{n(\text{ female })} \\ &= \frac{50}{120} \\ &= \frac{5}{12} \end{aligned}$$

Similarly, the probability that the medicine gives a negative result for males is:

$$\begin{aligned} P(\text{ negative if male }) &= \frac{n(\text{ negative and male })}{n(\text{ male })} \\ &= \frac{60}{90} \\ &= \frac{2}{3} \end{aligned}$$

Step 3: Independence

We need to determine whether the effect of the medicine and the gender of a participant are dependent or independent. According to the definition, two events are independent if and only if

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

We will look at the events that a participant is female and that the participant responded positively to the trial.

$$\begin{aligned} P(\text{ female }) &= \frac{n(\text{ female })}{n(\text{ total trials })} \\ &= \frac{120}{210} \\ &= \frac{4}{7} \end{aligned}$$

$$\begin{aligned} P(\text{ positive }) &= \frac{n(\text{ positive })}{n(\text{ total trials })} \\ &= \frac{80}{210} \\ &= \frac{8}{21} \end{aligned}$$

$$\begin{aligned} P(\text{ female and positive }) &= \frac{n(\text{ female and positive })}{n(\text{ total trials })} \\ &= \frac{50}{210} \\ &= \frac{5}{21} \end{aligned}$$

From these probabilities we can see that

$$P(\text{ female and positive }) \neq P(\text{ female }) \times P(\text{ positive })$$

and therefore the gender of a participant and the outcome of a trial are dependent events.

WORKED EXAMPLE 14: CONTINGENCY TABLES

QUESTION

Use the contingency table below to answer the following questions.

	Grade 11	**Grade** 12	**Totals**
Has cellphone	59	50	109
No cellphone	6	3	9
Totals	65	53	118

1. What is the probability that a learner from Grade 11 has a cellphone?
2. What is the probability that a learner who does not have a cellphone is from Grade 11.
3. Are the grade of a learner and whether he has a cellphone or not independent events? Explain your answer.

SOLUTION

1. There are 65 learners in Grade 11 and 59 of them have a cellphone. Therefore the probability that a learner from Grade 11 has a cellphone is $\frac{59}{65}$.
2. There are 9 learners who do not have a cellphone and 6 of them are in Grade 11. Therefore the probability that a learner who does not have a cellphone is from from Grade 11 is $\frac{6}{9} = \frac{2}{3}$.
3. To test for independence, we will consider whether a learner is in Grade 11 and whether a learner has a cellphone. The probability that a learner is in Grade 11 is $\frac{65}{118}$. The probability that a learner has a cellphone is $\frac{109}{118}$. The probability that a learner is in Grade 11 and has a cellphone is $\frac{59}{118} = \frac{1}{2}$. Since $\frac{1}{2} \neq \frac{65}{118} \times \frac{109}{118}$ the grade of a learner and whether he has a cellphone are dependent.

6 SUMMARY

- Terminology
  - **Outcome**: a single observation of an experiment.
  - **Sample** space of an experiment: the set of all possible outcomes of the experiment.
  - **Event**: a set of outcomes of an experiment.
  - **Probability** of an event: a real number between 0 and 1 that describes how likely it is that the event will occur.
  - **Relative frequency** of an event: the number of times that the event occurs during experimental trials, divided by the total number of trials conducted.
  - **Union** of events: the set of all outcomes that occur in at least one of the events, written as "A or B".
  - **Intersection** of events: the set of all outcomes that occur in all of the events, written as "A and B".
  - **Mutually exclusive events**: events with no outcomes in common, that is $(A \text{ and } B) = \emptyset$.
  - **Complementary events**: two mutually exclusive events that together contain all the outcomes in the sample space. We write the complement as "not A".
  - **Independent events**: two events where knowing the outcome of one event does not affect the probability of the other event. Events are independent if and only if $P(A \text{ and } B) = P(A) \times P(B)$.
- Identities
  - The **addition rule**: $P(A \text{ or } B) = P(A) + P(B)$–$P(A \text{ and } B$
  - The **addition rule** for 2 **mutually exclusive events**: $P(A \text{ or } B) = P(A) + P(B)$
  - The **complementary rule**: $P(\text{ not } A) = 1 - P(A)$
- A **Venn diagram** is a visual tool used to show how events overlap. Each region in a Venn diagram represents an event and could contain either the outcomes in the event, the number of outcomes in the event or the probability of the event.
- A **tree diagram** is a visual tool that helps with computing probabilities for dependent events. The outcomes of each event are shown along with the probability of each outcome. For each event that depends on a previous event, we go one level deeper into the tree. To compute the probability of some combination of outcomes, we
  - find all the paths that contain the outcome of interest;
  - multiply the probabilities along each path;
  - add the probabilities between different paths.

- A **two-way contingency table** is a tool for organising data, especially when we want to determine whether two events, each with only two outcomes, are dependent or independent. The counts for each possible combination of outcomes are entered into the table, along with the totals of each row and column.

7 EXERCISES

7.1 Exercise 1

1. Since the events "Jane will not lose all her money" are complimentary events, their probabilities sum to . Therefore the probability that Jane will lose all her money is $1 - 0,32 = 0,68$

2. If D and F are mutually exclusive events, with P(not D)$= 0,3$ and P(D or F)$= 0,94$, find $P(F)$.

7.2 Exercise 2

1. Given the following information:

   - $P(A) = 0,3$
   - $P(A$ and $B) = 0,2$
   - $P(B) = 0,7$

   First draw a venn diagram to represent this information. Then compute the value of $P(B$ and not $A)$.

2. You are given the following information:

   - $P(A) = 0,5$
   - $P(A$ and $B) = 0,2$
   - $P($not $B) = 0,6$

   Draw a Venn diagram to represent this information and determine $P(A$ or $B)$.

3. A study was undertaken to see how many people in Port Elizabeth owned either a Volkswagen or a Toyota. 3% owned both, 25% owned a Toyota and 60% owned a Volkswagen. What percentage of people owned neither car?

4. Let S denote the set of whole numbers from 1 to 15 , X denote the set of even numbers from 1 to 15 and Y denote the set of prime numbers from 1 to 15 . Draw a Venn diagram depicting S, X and Y .

5. Given: $P(A) = 0,57, P(B) = 0,23$ and $P(A$ or $B) = 0,8$. Are events A and B mutually exclusive? Justify your answer by calculations or by drawing an appropriate diagram.

6. For two events A and $B,$ it is given that $P(A) = 0,6, P(B) = 0,45$ and $P(A$ and $B) = 0,2$.

   6.1 Draw a Venn diagram representing the given the information.

   6.2 Determine $P(A$ or $B)$.

   6.3 Determine $P(A$ and $B\prime)$.

6.4 Determine $P(A$ or $B\prime)$

7. In the Grade 11 class, 60% of learners take mathematics (M), 30% take science (S) and 15% takes mathematics and science.

   7.1 Draw a Venn diagram to represent this data.

   7.2 Determine $P(M\prime$ or $S\prime)$.

7.3 Exercise 3

1. Of the 30 learners in a class 17 have black hair, 11 have brown hair and 2 have red hair. A leaner is selected from the class at random.

   1.1 What is the probability that the learner has black hair?

   1.2 What is the probability that the learner has brown hair?

   1.3 Are these two events mutually exclusive?

   1.4 Are these two events independent?endenumerate

2. $P(M) = 0,45; P(N) = 0,3$ and $P(M or N) = 0,615$. Are the events M and N mutually exclusive, independent or neither mutually exclusive nor independent?

3. Prove that if event A and event B are mutually exclusive with $P(A) = 0$ and $P(B) = 0$, then A and B are always dependent.

4. Given $P(A) = 0,25, P(B) = 0,5,$ and $P(A$ and $B) = 0,125$. Determine whether events A and B are independent or not. Show all your calculations.

5. Given $P(A) = 0,65, P(B) = 0,45,$ and $P(A$ or $B) = 0,8$.

   5.1 Determine $P(A$ and $B)$.

   5.2 Determine whether events A and B are independent or not. Show all your calculations.

6. P and Q are two events. $P(P) = 0,45, P(Q) = 0,75$ and $P(P$ or $Q) = 0,9$.

   6.1 Draw a Venn diagram to represent the events.

   6.2 Determine whether events P and Q are independent or not. Show all your calculations.

7. A six-sided die is rolled and the number landing face up is noted. Consider the following events:

   - Event A: The number landing face up is smaller than 3.
   - Event B: The number landing face up is a 4.
   - Event C: The number landing face up is an even number.

   7.1 Calculate $P(A$ or $B)$

7.2 Which events are mutually exclusive?

7.3 Are A and C independent events? Use calculations to substantiate your answer.

7.4 Exercise 4

1. Use the Venn diagram below to answer the following questions. Also given: $n(s) = 120$.

   1.1 Compute $P(F)$

   1.2 Compute $P(G \text{ or } H)$

   1.3 Compute $P(F \text{ and } G)$

   1.4 Are F and G dependent or independent?

2. The Venn diagram below shows the probabilities of 3 events. Complete the Venn diagram using the additional information provided.
   $P(Z \text{ and } (\text{not } Y)) = \frac{31}{100}$
   $P(Y \text{ and } X) = \frac{23}{100}$
   $P(Y) = \frac{39}{100}$
   After completing the Venn diagram, compute the following:
   $P(Z \text{ and not } (X \text{ or } Y))$

3. In a school, there are 230 learners in total. The learners play cricket (C), rugby (R) and football (F). Of this group of learners:
   35 plays cricket only
   50 plays rugby only
   61 plays football only
   20 plays cricket and rugby
   15 plays rugby and cricket
   10 plays cricket, rugby and football
   47 do not play either of the three sports

   3.1 Draw a Venn diagram representing this information.

   3.2 How many learners play rugby and football only?

   3.3 What is the probability that a learner plays cricket and football?

   3.4 What is the probability that a learner plays at least 2 of the sports?

7.5 Exercise 5

1. You role a die twice and add up the dots to get a score. Draw a tree diagram to represent this experiment. What is the probability that your score is a multiple of 5?

2. What is the probability of throwing at least one five in four rolls of a regular 6 -sided die? Hint: do not show all possible outcomes of each roll of the die. We are interested in whether the outcome is 5 or not 5 only.

3. You flip one coin 4 times.

   3.1 What is the probability of getting 3 heads?

   3.2 What is the probability of getting at least 3 heads?

4. You flip 4 different coins at the same time.

   4.1 What is the probability of getting 3 heads?

   4.2 What is the probability of getting at least 3 heads?

5. In his refrigerator, Peter has 4 apples and 6 tomatoes.Peter takes out two pieces of fruit at random, one after the other and eats the fruit.

   5.1 Draw a tree diagram to represent this

   5.2 Determine the probability that Peter will first take out a tomato and then an apple.

   5.3 Determine the probability that Peter will take out two apples.

   5.4 Determine the probability that none of the fruit peter will take out is an apple.

6. At a gold day, there are 3 prizes to be won at the end of the day from a lucky draw. The prizes are drawn one after the other. 200 tickets are sold for the lucky draw of which Stephan buys 7 . Winning any of the prizes does not disqualify you from winning another prize.

   6.1 Draw the tree diagram.

   6.2 Determine the probability that Stephan will win one prize. Give your answer correct to 3 decimal places.

7.6 Exercise 6

1.

8 ANSWERS FOR EXERCISES

8.1 Exercise 2

1. $P(B$ and(not $A) = 0,5$
2. $P(A$ or $B) = 0,7$
3. $P($not $(T$ or $V)) = 0,18$
4.
5. Yes, mutually exclusive since $P(A) + P(B) = P(A$ or $B)$
6. 6.1
   6.2 $P(A$ or $B) = 0,85$
   6.3 $P(A$ and $B\prime = 0,4$
   6.4 $P(A$ or $B\prime = 0,75$
7. 7.1
   7.2 $P(M\prime$ or $S\prime)$

8.2 Exercise 3

1. 1.1 $P($Blackhair$) = \frac{17}{30}$
   1.2 $P($Brownhair$) = \frac{11}{30}$
   1.3 Yes, since each learner has only one hair colour, so $P($ blackandbrown $) = 0$
   1.4 No, since $P($ blackandbrown $) = 0$ $P($ black $) \times P($ brown $)$
2. Independent and not mutually exclusive
3.
4. Events A and B are independent since $P(A) \times P(B) = P($ A and $B)$
5. 5.1 $P(A$ and $B) = 0,3$
   5.2 Events A and B are not independent since $P(A) \times P(B)$ $P(A$ and $B)$
6. 6.1
   6.2 Events P and Q are not independent since $P(P) \times P(Q)$ $P(P$ and $Q)$
7. 7.1 $P(A$ or $B) = \frac{1}{2}$
   7.2 A and B
   7.3 Events A and C are independent since $P(A) \times P(C) = P(A$ and $C)$

8.3 Exercise 4

1. 1.1 $\frac{9}{40}$

   1.2 $\frac{3}{5}$

   1.3 $\frac{1}{10}$

   1.4 Dependent

2. $\frac{4}{25}$

3. 3.1

   3.2 12

   3.3 $\frac{3}{46}$

   3.4 $\frac{37}{230}$

8.4 Exercise 5

1. $\frac{7}{36}$

2. $\frac{671}{1296}$

3. 3.1 $\frac{1}{4}$

   3.2 $\frac{5}{16}$

4. 4.1 $\frac{1}{4}$

   4.2 $\frac{5}{16}$

5. 5.1

   5.2 $\frac{4}{15}$

   5.3 $\frac{2}{15}$

   5.4 $\frac{1}{3}$

6. 6.1

   6.2 0,099