Mathematics

Unit 6

Statistics and Probability

Acknowledgements:

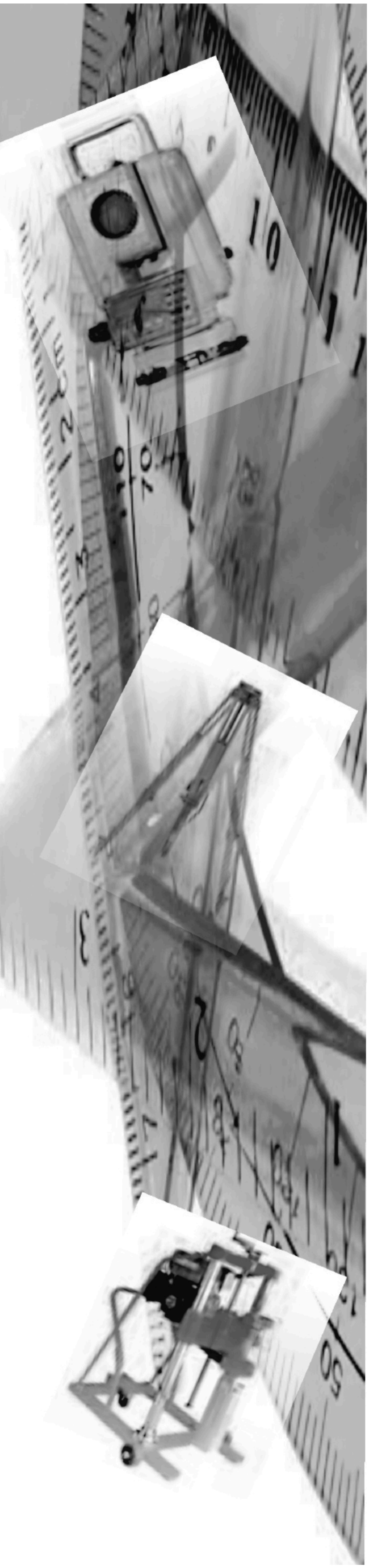
Editor: Allison Kitto
Writer: Allison Kitto, Rob Houghton-Rhodes
Layout: Lidia Kruger

This Study Unit is the property of the learner to whom it is given.

CONTENTS

LESSON 1

Statistics and the Calculator



About this lesson

This lesson deals with statistics and probability. Statistics is a branch of mathematics that deals with the collection, presentation, analysis and interpretation of numerical data. It is a kind of mathematics that many of you need in the work that you do.

In real life we come across huge sets of statistical data like election results and population censuses making it impossible to analyse them without the use of a calculator or even a computer. Since many of you may not have access to a computer, we shall concentrate on the use of a calculator to analyse our data in this lesson.

Before we do that, we shall look at some of the symbols you see on your calculator. We shall also calculate standard deviation from a set of simple data without the use of a calculator. Standard deviation is used to describe the spread of data about the mean.

In this lesson you will:

- calculate the standard deviation for a set of simple data without the use of a calculator
- put statistical data in your calculator
- use your calculator to find totals, means, sums of squares and standard deviations.

The Σ Notation

You are familiar with the 'instruction signs' like +, −, × and ÷. The Σ sign can also be considered to be an instruction sign. Σ is pronounced sigma. It is the Greek capital letter for S.

In statistics we often need to add a long list of numbers. These numbers may be represented by $x_1, x_2, x_3, ...x_n$ if there are n of them.

To write out $x_1 + x_2 + x_3 + ... + x_n$ each time is annoying and time-consuming. There should be a simple way to write this out. This is exactly what the Σ sign does. For instance if there are 7 values for a set of data then $\Sigma x_i = x_1 + x_2 + x_3 + ... + x_7$. So Σx_i means add up all the x's that are there. The *'i'* of Σx_i is called a subscript. You can use any of the letters of the alphabet for the subscript. You may also leave the subscript out.

This means that Σx, Σx_i, Σx_j, Σx_r etc. all mean the same thing if we are certain of the number of values.

$$\text{Mean} = \frac{\text{Total of all values}}{\text{Total number of values}}$$

Do you remember how to find the mean of a set of data?

$$\text{Mean} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n}$$

$$= \frac{\Sigma x}{n}$$

which we say as 'the sum of all the values divided by the total number of values'. We often denote the mean by $\bar{x}$ (which is pronounced x bar). So:

$$\bar{x} = \frac{\Sigma x}{n}$$

where $\bar{x}$ is the mean,
Σx is the sum of all values of x and
n is the total number of values

What do you think Σx^2 means?
Σx^2 means add up the squares of x. That is if there are n values, then

$$\Sigma x^2 = x_1^2 + x_2^2 + x_3^2 + ... + x_n^2$$

Example 1
Suppose there are 6 x–values.
These are $x_1 = 5, x_2 = 4, x_3 = 10, x_4 = 1, x_5 = 11$ and $x_6 = 7$

Find
a) Σx
b) $\bar{x}$

Solution

a) Here there are 6 values. That is $n = 6$. Therefore
$\Sigma x = 5 + 4 + 10 + 1 + 11 + 7 = 38$

b) $$\bar{x} = \frac{\Sigma x}{6} = \frac{38}{6} = 6{,}\dot{3}$$

Now try some examples yourself.

ACTIVITY 1

1. For the set of data: 11, 5, 7, 2, 9, 13, 4, 8, find:
   a) Σx
   b) $\bar{x}$

2. What is the difference between each of the following?
   a) $\Sigma x - 2$ and $\Sigma(x - 2)$
   b) $\Sigma x - \bar{x}$ and $\Sigma(x - \bar{x})$

3. What does $\Sigma(x - \bar{x})^2$ mean?

ANSWERS ON PAGE 51

Standard Deviation

The range is a measure of spread of a set of data.

range:
difference between the largest and the smallest values

Sometimes, the range does not give a very good picture of the spread of a set of data. This happens when we have extreme values (a value too large or too small compared to all the other values) in the data set. The range in this case may give an impression that the data is widely spread when it is not. For example, the set of data 1, 2, 3, 4, 5, 6, 7, 8, 9, and 100 will have a range of 99 (100 – 1) giving the impression that the data is spread between 1 and 100 but in actual fact all of the values except the last are below ten.

The standard deviation is the most commonly used measure of spread. It describes how the values are spread around the mean. A small standard deviation means that the values are spread closely around the mean; a large standard deviation means many values are scattered away from the mean.

The standard deviation is denoted by σ. σ (pronounced sigma) is the Greek small letter for s. That is, σ is the small letter for Σ. The standard deviation of a set of data is defined as the 'root mean squared deviation'.

Using the σ notation it is defined as

$$\sigma = \sqrt{\frac{\Sigma\left(x - \bar{x}\right)^2}{n}}$$

Where σ is the standard deviation and *n* the number of values.

Example 2

Find the standard deviation for a set of data consisting of these numbers 45, 20, 60, 40 and 50.

Solution

$\bar{x} = \frac{\Sigma x}{n}$ (first find the mean $\bar{x}$)

$$\frac{45 + 20 + 60 + 40 + 50}{5}$$

$= 43$

We now need to find the differences between the individual data items and the mean $\left(x - \bar{x}\right)$ as well as the squares of these differences $\left(x - \bar{x}\right)^2$ The total of the $\left(x - \bar{x}\right)^2$ values will give $\sum\left(x - \bar{x}\right)^2$.

x	$\left(x - \bar{x}\right)$	$\left(x - \bar{x}\right)^2$
45	2	4
20	−23	529
60	17	289
40	−3	9
50	7	49
TOTAL (Σ)	(0)	880

Notice that adding the values $\left(x - \bar{x}\right)$ does not give any indication of the spread of the data items as the positive and negative differences between the individual items and the mean cancel each other out.

$$\sigma = \sqrt{\frac{\Sigma\left(x - \bar{x}\right)^2}{n}}$$

$$= \sqrt{\frac{880}{5}} = \sqrt{176} = 13,27 \text{ (to 2 decimal places)}$$

ACTIVITY 2

1. The wages of the eight people who do similar work but at different places are:

   R369, R267, R315, R312, R585, R315, R279, R510

   Calculate:
   a) the mean wage
   b) the standard deviation of the wages.

2. Find the mean and standard deviation of the following set of readings: 1, 3, 3, 5, 8

ANSWERS ON PAGE 51

The Calculator

Can you find the following on your calculator?

n, $\sum x$, $\bar{x}$ and σ

These are the results you can get from your calculator after putting your set of data in your calculator.

Do you still remember what they stand for?

n is the number of values (or x –values)
$\sum x$ is the sum of all the x –values
$\bar{x}$ is the mean of the x –values
σ is the standard deviation.

Putting data into the calculator

There are many different types of calculator. We have based the following instruction on the CASIO fx-82ES PLUS. If you have a different calculator from this, even if it is a different model number of the CASIO, you may find you have to press different keys in order to use the calculator for statistics. You must consult the instruction book which comes with the calculator.

If you have the CASIO fx-82ES PLUS you do the following to get into Stats mode:

Press [MODE] and choose [2] from the menu. Then choose [1] (for data with one variable) from the next menu that comes up. A table will appear which might have one or two columns. If there is a second (frequency) column, ignore it for the moment.

You are now ready to put your statistical date into your calculator.
Let's illustrate how to put data into a calculator with an example.

Example 3
Use your calculator to determine the number of data items, the mean and the standard deviation of the following data: 27, 5, 6, 4, 10, 9.

Solution
Get into the statistics mode by pressing [MODE], then [2] and then [1].

Put the given values 27, 5, 6, 4, 10, 9 into your calculator by pressing [2] [7] [=] [5] [=] [6] [=] [4] [=] [1] [0] [=] [9] [=]. Then press the orange [AC] button followed by [SHIFT] [1]. Then choose [4] from the menu which comes up.

The next menu offers 4 options: $n, \bar{x}, \sigma x$ and sx . The last option will not be used in this course. It gives an estimate of the standard deviation of a large population from a smaller sample.

Pressing [1][=] gives the value of n which is 6.
You have to press [SHIFT] [1] and then [4] again to get back to the menu where pressing [2] [=] gives the mean. In this example it is 10,16666667 or 10,2 (correct to 1 decimal place).

Notice the calculator displays 10,6666667. Pressing [SHIFT] [1] and then [4] again we can now choose the third option [3] [=] and the standard deviation will be displayed. $\sigma \approx 7{,}81913607 = 7{,}8$ (to 1 decimal).

ACTIVITY 3

Put the following sets of data into your calculator and find the mean and the standard deviation in each case.

1. 66, 64, 58, 70, 45

ANSWERS ON PAGE 53

2. 7, 10, 15, 2, 4, 6

Putting data from a frequency table into your calculator

Example 4
A frequency table of marks scored by a student is given below. Put the data into your calculator and then read off the mean and standard deviation.

Marks	Frequency
2	3
3	5
5	2
TOTAL	10

The table tells us that the student had 2 marks three times, 3 marks five times and 5 marks twice.

Solution:
This time you will need the frequency column in your data entry table. If you only have one column when you press [MODE] [2] [1], press [SHIFT] [MODE], scroll down to a second screen and choose [3] and then press [1] to insert a frequency column. This will take you to a blank table which should now have a frequency column.

Enter into the first column: [2] [=] [3] [=] [5] [=] and then move to the frequency column and enter [3] [=] [5] [=] [2] [=]. Now pressing [AC] [SHIFT] [1] [4] will once again bring up the menu from which answers can be read off by pressing [2] [=] and then [SHIFT] [1] [4] [3] [=]
The answers are: $\bar{x} = 3{,}1$ and $\sigma \approx 1{,}044030651$

ACTIVITY 4

The monthly salaries of workers of a small business were recorded in the table below. Put the data into your calculator and write down the mean monthly salary and the standard deviation.

Monthly salary	Number of workers/Frequency
5 000	1
3 000	2
1 025	3
760	2
TOTAL	8

ANSWERS ON PAGE 53

Distribution around the median

You will remember that the **median** is a number which divides a set of data which has been arranged in ascending (or descending) order, into two equal data sets. When there is an odd number of data items, the median is the middlemost item, but when there is an even number of data items, the median is often not actually a data item but the mean of the two middle data items.

median:
situated in or pertaining to the middle

So if the items are 3; 5; 7; 8; 9; 11; 13, then the median of these seven items is the fourth item, namely 8.

If there is an additional item in the set of data 3; 5; 7; 14, then the median of these eight items is the mean between the fourth and fifth items: $\frac{8+9}{2} = 8{,}5$

Measures of dispersion about the median are **quartiles** which divide the data into four equal, ordered subsets, **deciles** which divide the ordered data into ten equal sized ordered subsets and **percentiles** which divide ordered data into a hundred equal sized ordered subsets. In a large data set of 2 000 items for example, the seventh decile divides the data into the lower 1 400 values and the highest items. The seventy fifth percentile divides the data item into the lower items and the top 500 items in an ordered list of these 2 000 items.

quartiles:
one of the values of a variable that divides the distribution of the variable into four groups having equal frequencies

deciles:
one of the values of a variable that divides the distribution of the variable into ten groups having equal frequencies

percentiles:
one of the values of a variable that divides the distribution of the variable into 100 groups having equal frequencies

Quartiles are calculated in the same way as medians, but we work with the bottom half of the data for the lower quartile and with the upper half for the upper quartile.

In the example with seven items: 3; 5; 7; 8; 9; 11; 13, the lower quartile must divide the three lower items: 3; 5 and 7 into equal sized sets and so this number is the data item: 5. Similarly the upper quartile is 11.

When the extra item is added, the lower quartile is the mean of the second and third items: $\frac{5+7}{2}=6$ and the upper quartile is the mean between the 6th and 7th items: $\frac{11+13}{2}=12$.

These measures are generally only relevant when examining much larger data sets. We will discuss a more practical method of finding quartiles. deciles and percentiles later.

However, we usually use quartiles in a five number summary of a given data set which might not be that large.

Five number summaries

Example

A teacher recorded the following marks for the 25 learners in her class:
54; 72; 17; 40; 92; 67; 81; 63; 72; 70; 68; 65; 77; 85; 50; 75; 56; 90; 63; 66; 80; 78; 55; 70; 69.

Write down a five number summary of this data, explain the meaning of the summary and draw a box and whisker diagram to illustrate the data spread.

Solution

First the numbers must be arranged in ascending order (lowest to highest): 17; 40; 50; 54; 55; 56; 63; 63; 65; 66; 67; 68; 69; 70; 70; 72; 72; 75; 77; 78; 80; 81; 85; 90; 92.

The median is the thirteenth item in the ordered list above.

The lower quartile is the mean of the sixth and seventh items:

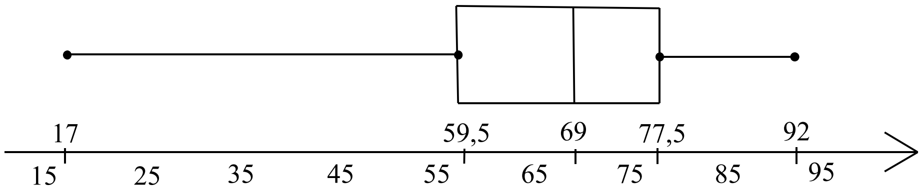
$\frac{56+63}{2}=59{,}5$ and the upper quartile is the mean of the nineteenth and twentieth items: $\frac{77+78}{2}=77{,}5$.

The five number summary consists of the lowest data item, the lower quartile, the median, the upper quartile and the highest item: 17; 59, 5; 69; 77,5 and 92.

> ***distorted:***
> *not truly or completely representing the facts or reality; misrepresented; false*
>
> ***outliers:***
> *something that lies outside the main body or group that it is a part of*

This means that the lowest mark was 17; the lowest quarter of the marks were between 17 and 59,5, the next quarter were between 59,5 and 69, the third quarter between 69 and 77,5 and the top quarter of the marks were between 77,5 and 92. The middle half of the learners got between 59,5 and 77,5. This interval is called the interquartile range. It is perhaps more useful than the top and bottom quarters as these could be **distorted** by **outliers.** An outlier is a data item which is significantly different from the rest. In this case the mark of 17 is less than half the next lowest mark and is an outlier.

A **box and whisker diagram** uses the five number summary in the following way: the lowest number is the start of the first 'whisker', the lower quartile marks the end of the left hand 'whisker' and the start of the 'box', a line somewhere between the start and the end of the 'box' shows the median, the upper quartile the end of the 'box' and the start of the right hand 'whisker' and the maximum value the end of the right hand 'whisker'.



box and whisker diagram:
A box-and-whisker diagram shows the distribution of a set of data along a number line, dividing the data into four parts using the median and quartiles.

ACTIVITY 5

For each of these sets of data which have been arranged in ascending order:

a) write down the five number summary and
b) draw a box and whisker diagram.

1. 13, 14, 15, 17, 18, 20, 22.

2. 5, 6, 7, 7, 8, 10, 10, 13, 15, 20.

3. 25, 28, 28, 29, 30, 31, 32, 34, 35, 36, 38, 39

ANSWERS ON PAGE 53

Summary

In this lesson you learnt that the sign Σ is an ‘instruction sign’ for adding up numbers. You also learnt that the standard deviation of a set of data describes the spread of the data around its mean. The bigger the value of the standard deviation the wider the spread of the data is about its mean. The standard deviation is denoted by σ and is defined as

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$$

You were shown how to calculate the standard deviation ‘manually’ calculating the individual differences between the mean and individual data items, squaring these differences, adding them together, dividing by the number of data items and then finding the square root of this result. You have used your calculator to enter data and to enter data given in a frequency table and then find the number of data items entered, the mean and the standard deviation of this data. Finally, you have seen that to calculate quartiles, you had to determine a five number summary for a data set and illustrate this with a box and whisker diagram.

CHECKLIST

Are you able to:

- calculate the standard deviation for a set of simple data without the use of a calculator;
- put in statistical data into your calculator
- use your calculator to find totals, means, sums of squares and standard deviations

SELF-CHECK EXERCISE

1. The sales (in R) of each branch of a chain of stores in KwaZulu Natal province were recorded for a particular day. The results were 340; 370; 300; 340; 380; 310; 350; 330. Calculate:

   a) the mean and
   b) the standard deviation of the sales.

2. The Mathematics marks for 27 workers who enrolled in an out-of-school youth programme in the Eastern Cape province were as follows:

   54; 56; 50; 44; 48; 58; 55; 49; 64; 50; 48; 59; 37; 47; 58; 54; 51; 56; 54; 40; 51; 54; 44; 53; 43; 38; 51

   Use your calculator to find:
   a) the total marks for the 27 workers
   b) the mean mark for the workers
   c) the standard deviation of the marks.

3. The rainfall in June for 6 consecutive years in a certain city was 48 mm, 50 mm, 62 mm, 47 mm, 56 mm and 49 mm. With the use of a calculator find the

   a) mean
   b) standard deviation of the rainfall in June and
   c) interpret your results.

This is a way of displaying data items 17, 19, 22, ...59 where the tens digit of each data item is the stem and the units digit is the leaf. The items are arranged in ascending order in 'bars' made of digits.

4. Given is a stem and leaf plot of the number of pages written each week by an author for a period of twenty four weeks:

Stem	Leaf
1	7 9
2	2 3 4 5 7
3	0 1 1 2 4 5 6 8
4	2 2 3 5 6 9
5	3 8 9

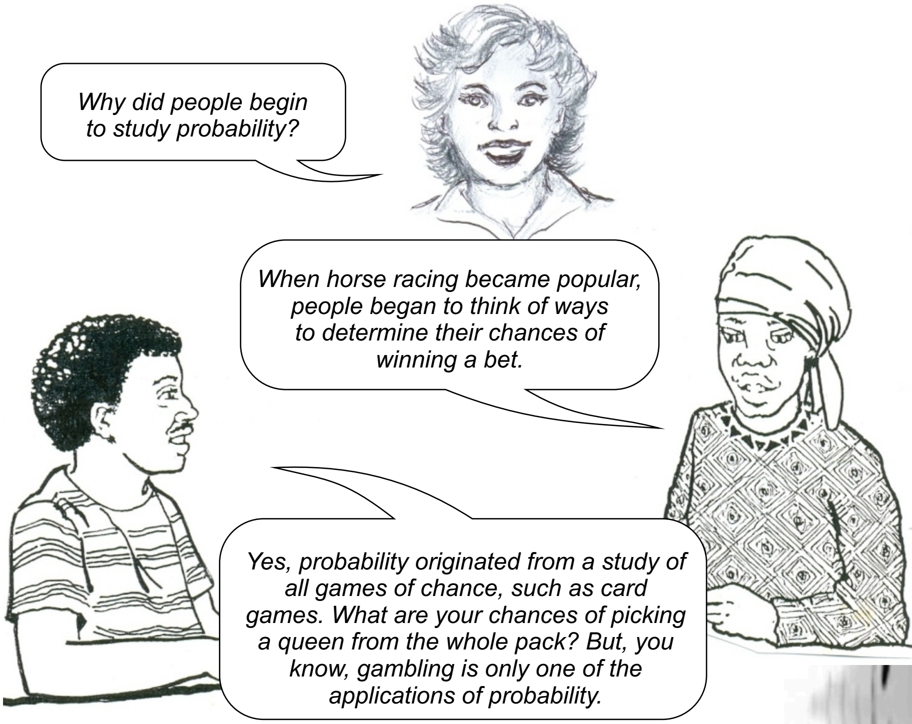
   a) Write down the five number summary of this data.
   b) Draw a box and whisker plot to illustrate the spread of the data around the median.

ANSWERS ON PAGE 63

LESSON 2

Elementary Probabilities

About this lesson

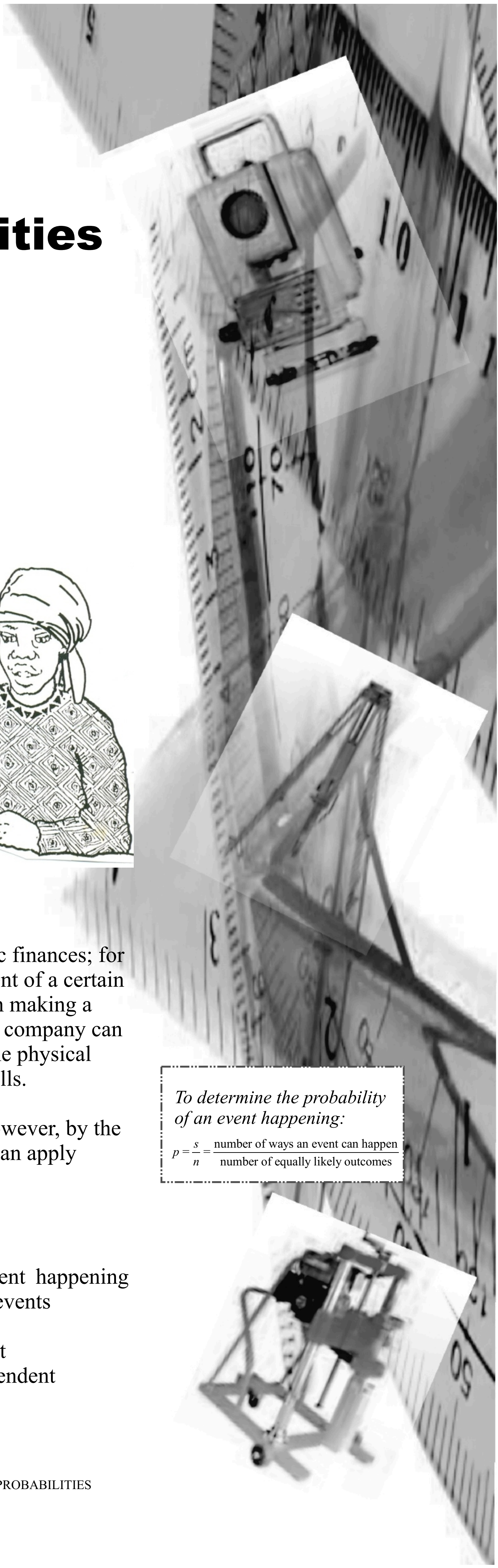


Probability provides statistical tools for managing public finances; for determining the chances whether a drug can cure a patient of a certain illness; for determining the chances of an insured person making a claim on his/her policy; to determine how much profit a company can make in marketing a certain product. It is also used in the physical sciences, economics and during elections for opinion polls.

We shall not consider all these applications in detail. However, by the end of the lesson, you will have a fair idea of how you can apply probability in these areas.

In this lesson you will:

- calculate the probabilities of more than one event happening and the relationship that can exist between any two events
- use tree diagrams to solve probability problems
- decide whether events are dependent or independent
- calculate the probabilities of dependent and independent events.



To determine the probability of an event happening:

$$p = \frac{s}{n} = \frac{\text{number of ways an event can happen}}{\text{number of equally likely outcomes}}$$

Using tree diagrams to solve probability problems

Example 1

Let's assume that we have two bags. The first bag contains 3 red balls and a blue ball. The second bag contains 5 red balls and 6 blue balls. Find the probability of picking:

a) a red ball from the first bag
b) a blue ball from the first bag
c) a red ball from the second bag
d) a blue ball from the second bag.

Solution

a) There are 4 balls in the first bag. 3 of these balls are red. Therefore the probability of picking a red ball is $\frac{3}{4}$.

b) Using the same reasoning as we did in a), the probability of picking a blue ball is $\frac{1}{4}$.

c) There are a total of 11 balls in the second bag. 5 of these are red and so the probability of picking a red ball is $\frac{5}{11}$.

d) The required probability is $\frac{6}{11}$.

We shall extend this idea to cover the cases where we have two or more bags with different numbers of balls.

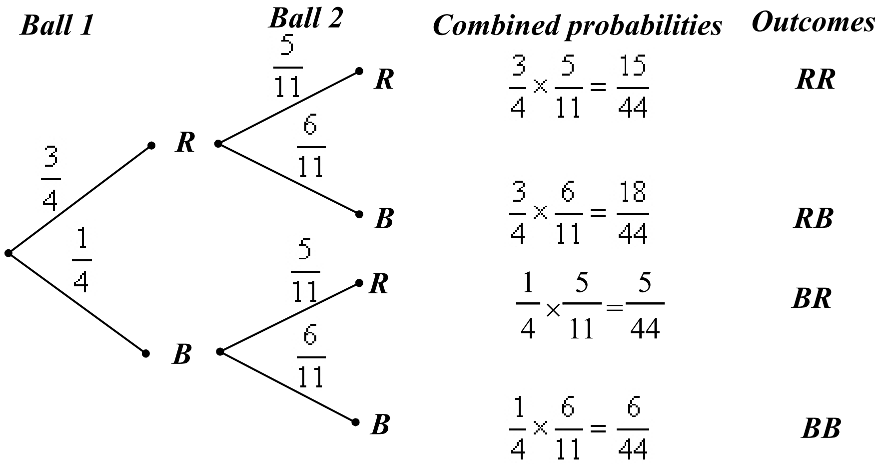
Combined Probabilities

Now suppose we wanted to find the probability of picking a red ball from the first bag and another red ball from the second bag or the probability of a red ball from the first bag and a blue one from the second bag, how do we go about it? One method of solving such a problem is to use the tree diagram.

Tree diagram

Example

We can draw a tree diagram for this problem like this:



The first fork tells us whether we pick a red ball or a blue ball from the first bag. The probabilities for each event of picking a ball are written along their branches. From the example on page 14, the probability of picking a red ball from the first bag is $\frac{3}{4}$ and is written along the top branch. Also the probability of picking a blue ball from the first bag is $\frac{1}{4}$ and is written along the lower branch.

The second fork corresponds to whether we pick a red ball or a blue ball from the second bag. The top fork gives the probabilities of picking a red or a blue ball from the second bag, if we picked a red ball from the first bag. The lower fork gives the probabilities for picking a red ball or a blue ball from the second bag if we picked a blue ball in the first bag. The probability of picking a red ball from the second bag is $\frac{5}{11}$ and is written along the top branch of the top fork. The probability of picking a blue ball from the second bag is $\frac{6}{11}$ and is written along the lower branch of the top fork.

Can you see that the probabilities written along the lower fork are the same as those on the top fork? This is because they both refer to the same events (picking balls from the second bag). The forks of the second bag are linked to those of the first because they describe the different ways in which you can pick two balls from the two bags, the first ball from the first bag and the second ball from the second bag.

You can see from the tree diagram that in all, there are four ways of picking the two balls. You can pick a red ball from the first bag and another red ball from the second bag (RR), or you can pick a red ball from the first bag and a blue ball from the second bag (RB), or you can pick a blue ball from the first bag and a red ball from the second bag (BR), or you can pick a blue ball from the first bag and another blue ball from the second bag (BB).

Now you can read the probabilities for the combined event from the diagram. You can find them by multiplying the probabilities along each branch. For example the probability of picking a red ball in the first bag and another red ball in the second bag is

$$P(RR) = \frac{3}{4} \times \frac{5}{11} = \frac{15}{44}$$

where P(RR) denotes the probability of picking a red ball in the first draw and another red ball in the second draw. In general, the notation P(A) stands for the probability of an event A occurring. You can calculate the other probabilities in the same way.

The probability of picking a blue ball from the first bag and a red ball from the second bag is:

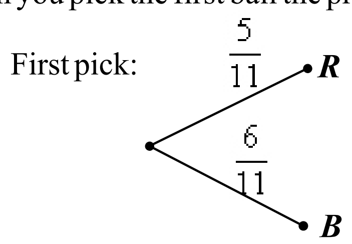
$$P(BR) = \frac{1}{4} \times \frac{5}{11} = \frac{5}{44}$$

Can you see that the four combined probabilities $\left(\frac{15}{44}; \frac{18}{44}; \frac{5}{44} \text{ and } \frac{6}{44}\right)$ add up to $\frac{44}{44} = 1$ because all possible combinations are included and no two combinations can happen at the same time? Also the tree applies only for picking the first ball from the first bag and the second ball from the second bag but can be extended if you want to calculate further probabilities.

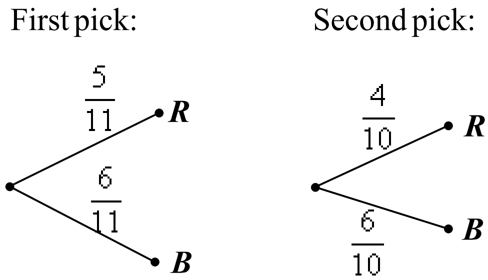
Example 2

Let us consider one more example. Suppose this time we have only one bag, but we pick up two balls from the bag one at a time without replacing the first ball. Let's use the bag with 5 red balls and 6 blue ones.

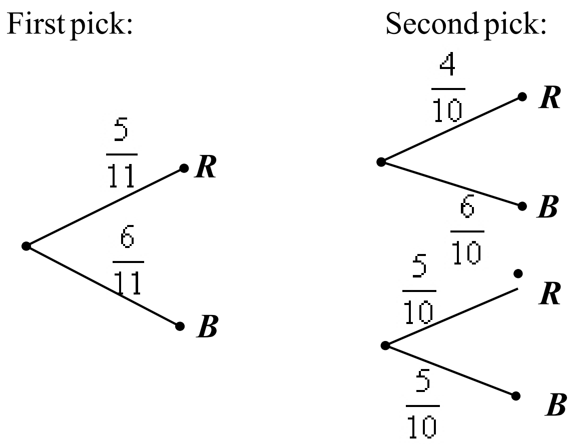
When you pick the first ball the probabilities can be summarised as



Once you have picked the first ball and it is red, how many balls will be left and how many of them are red? There will be 10 balls left and 4 of them will be red. So the chance of picking a red ball again from the same bag is $\frac{4}{10}$ and the chance of picking a blue ball is $\frac{6}{10}$ because there are still 6 blue balls in the bag.



On the other hand if you first picked a blue ball from the bag, you would be left with 5 blue balls and 5 red balls in the bag. So the probability of choosing a blue ball is $\frac{5}{10}$ and the probability of choosing a red ball is also $\frac{5}{10}$. This then gives us the following tree diagram for picking two balls from the bag.



The sum of the probabilities on each pair of branches is 1:

$\frac{5}{11} + \frac{6}{11} = 1$; $\frac{4}{10} + \frac{6}{10} = 1$ and $\frac{5}{10} + \frac{5}{10} = 1$

And the sum of the combined probabilities:

$$\frac{5}{11} \times \frac{4}{10} + \frac{5}{11} \times \frac{6}{10} + \frac{6}{11} \times \frac{5}{10} + \frac{6}{11} \times \frac{5}{10} = \frac{110}{110} = 1$$

We can then read off the four combined probabilities from the tree diagram.

Find the probabilities that

a) both balls are red
b) both balls are blue
c) both balls are of the same colour
d) the first ball is blue and the second ball is red
e) the first ball is blue
f) the second ball is red

Solution

a) The probability that both balls are red is the top branch. That is

$$P(\text{both are red}) = P(PR) = \frac{5}{11} \times \frac{4}{10} = \frac{20}{110}$$

b) The probability that both balls are blue is the last (bottom) branch.

$$P(BB) = \frac{5}{11} \times \frac{6}{10} = \frac{30}{110}$$

c) Both balls are of the same colour means we either have both balls red or both balls blue. Therefore the required probability is the sum of the probabilities of the first and the last branches.

$$\begin{aligned} P(\text{both same colour}) &= P(BB) + P(RR) \\ &= \frac{5}{11} \times \frac{4}{10} + \frac{5}{11} \times \frac{6}{10} \\ &= \frac{20}{110} + \frac{30}{110} \\ &= \frac{50}{110} \text{ or } \frac{5}{11} \end{aligned}$$

d) The first ball is blue and the second ball that is red is the third branch.

Therefore P (BR) $\frac{6}{11} \times \frac{5}{10} = \frac{30}{110}$ or $\frac{3}{11}$

e) The probability that the first ball is blue is simply $\frac{6}{11}$.

f) The probability that the second ball is red can happen in one of two ways: either the first ball is red and the second ball is red or the first ball is blue and the second ball is red. Therefore we have the first and the third branches.

$$\text{So P(second ball red)} = \frac{5}{11} \times \frac{4}{10} + \frac{6}{11} \times \frac{5}{10}$$
$$= \frac{20}{110} + \frac{30}{110}$$
$$= \frac{50}{110} \text{ or } \frac{5}{11}$$

ACTIVITY 1

1. A survey conducted by a social worker showed that the probability of a baby being a girl was $\frac{3}{5}$ and the probability of a baby being a boy was $\frac{2}{5}$.

   If a family has 2 children, use a tree diagram to find the probability that
   a) both are girls.
   b) both are boys.
   c) one is a boy and the other a girl.

2. Vusi has 8 left shoes and 11 right shoes in his cupboard. He picks 2 shoes without replacing the first in the cupboard. Find the probabilities that
   a) both are left shoes.
   b) he picks a left shoe and a right shoe.

ANSWERS ON PAGE 54

Independent and dependent events

Let us look back at the example where we had 5 red balls and 6 blue balls in a bag and we picked 2 balls from the bag without replacing the first ball. The probability of choosing a red is $\frac{5}{11}$. We said that if a red ball is chosen first, then the probability of the second ball being red is $\frac{4}{10}$.

This is because we will be left with 4 red balls and 6 blue balls in the bag. If a blue ball was chosen first then the probability of the second ball being red is $\frac{5}{10}$, but $\frac{4}{10} \neq \frac{5}{10}$. This means that the probability of choosing a second ball depends on what was chosen for the first ball. So we say that the two events are **dependent**.

Let us consider another example with a pack of cards. Normally, a pack of cards contains 52 cards excluding joker cards, among which are 4 Aces, 4 Jacks, 4 Queens and 4 Kings. If 2 cards are drawn from the pack without replacing the first, the probability that

the first card drawn is an Ace is $\frac{4}{52}$. The probability of the second card also being an Ace will now be $\frac{3}{51}$ because there will be only 3 Aces left and the pack will also be one card less. Now supposing the first card drawn was a Queen, then the probability of the second card drawn being an Ace is $\frac{4}{51}$ because there will be 4 Aces in the pack before the second draw.

Since the probability of the second card drawn being an Ace is not the same in both cases $\left(\frac{3}{51} \neq \frac{4}{51}\right)$ it means that the probability of drawing the second card depends on the first card drawn. So we say that the two events are **dependent**.

*Two events are **dependent** if the occurrence of one event makes the occurrence of a second event either more or less likely.*

Now let's look at the example where we had two bags. The first bag contained 3 red balls and 1 blue ball and the second bag contained 5 red balls and 6 blue balls. We picked one ball from each bag. The probability of choosing a red ball from the second bag was $\frac{5}{11}$. This probability was the same whether we picked a blue ball in the first bag or a red ball in the first bag. Therefore the probability of the second event was not influenced by the first. In such a case we say the two events are **independent**.

In general two events are said to be **independent** if one of them does not make the other event either more or less likely to occur. They are said to be **not independent** or **dependent** if one event is likely to make the other event either more or less likely to occur.

ACTIVITY 2

In each of the following situations, indicate whether the two events A and B are independent or not independent.

a) Two dice are rolled.
Event A: There is a 6 on the first dice.
Event B: There is a 6 on the second dice.

b) Two cards are drawn from a pack without the first being replaced.
Event A: There is an Ace on the first draw.
Event B: There is an Ace on the second draw.

c) A coin and a dice are tossed.
Event A: There is a tail on the coin.
Event B: There is a 5 on the dice.

d) A bag contains 8 red discs and 6 blue discs. Two discs are drawn at random one at a time without being replaced.

Event A: The first disc drawn is blue.
Event B: The second disc drawn is red.

ANSWERS ON PAGE 55

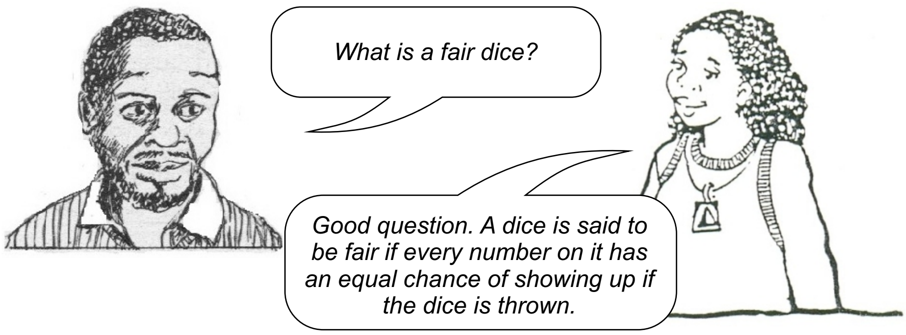
Calculating probabilities of independent events

If you know the probability of an event A occurring in one experiment (like coin throwing) and the probability of an event B occurring in another (like throwing a dice) and that the outcome of one experiment is totally unconnected to the outcome of the other, then the probability of A and B both happening = probability of A happening × probability of B happening.

In simple terms if A and B are independent then

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

Example
Two fair dice are thrown. What is the probability that the result is a double six?



Solution
For each single dice, the probability of a six is $\frac{1}{6}$. Since the two dice cannot influence each other in any way, the probability that both dice give sixes is $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$.

ACTIVITY 3

1. A football team has probabilities $\frac{1}{2}, \frac{1}{3}$ and $\frac{1}{6}$ of winning, losing and drawing a match respectively. If the team plays two matches, find the probability that:

   a) both matches are drawn
   b) the first match is won and the second lost

2. A marksman hits the bull's eye on a target board with probability 0,9. He fires twice. What is the probability that he gets 2 bulls?

ANSWERS ON PAGE 56

Summary

You learnt to calculate probabilities of combined events by using tree diagrams. You saw when two events were dependent and independent and that if the two events are independent, then the probabilities of them both happening is the product of their individual probabilities.

CHECKLIST

Are you able to:

- calculate the probability of more than one event happening
- use tree diagrams
- decide whether events are dependent or independent
- calculate the probability of dependent and independent events.

SELF-CHECK EXERCISE

1. You choose a committee of 2 people from a group of 6 women and 4 men. Use a tree diagram to calculate the probabilities of getting

   a) 2 women;
   b) 2 men;
   c) one woman and one man.

2. Nomonde gets up late one day in five on average. She knows her bus to work comes late three days in every five. Write down the probability that:

   a) Nomonde gets up late.
   b) her bus comes on time.

Nomonde misses the bus if she gets up late and the bus is on time.

c) Calculate the probability that she will miss the bus.

d) Show all the different possibilities on a tree diagram.

3. In determining the president, secretary general and treasurer for a certain organisation, the following nominations were made:

- for the president, 4 people one of whom is a women
- for the secretary general 5 people of whom 3 are older than 50 years and
- for the treasurer, 3 people of whom 2 are women.

a) Calculate the probabilities that the elected

i) president is a man
ii) secretary general is above 50 years and
iii) the treasurer is a woman

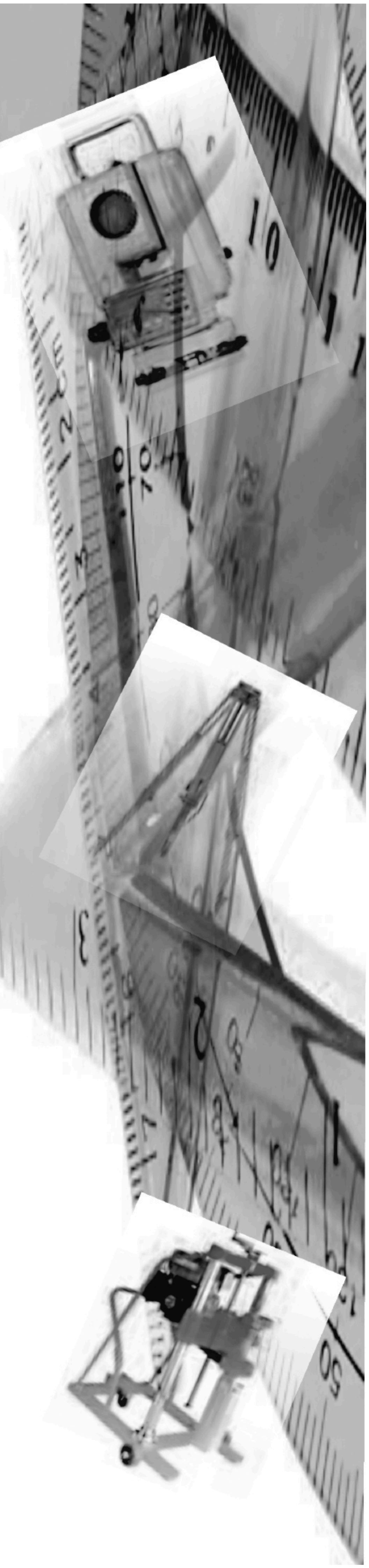
b) Assuming that none of the people nominated for one position was nominated for any other position find the probabilities that:

i) the president is a man and the secretary general is older than 50 years
ii) both the president and the treasurer are males
iii) the treasurer is a woman and the secretary general is older than 50 years

ANSWERS ON PAGE 64

LESSON 3

Venn diagrams and probability



About this lesson

Venn diagrams are sometimes useful in illustrating the possible outcomes in a statistical experiment and solving some probability problems. A Venn diagram is an illustration of the relationships between and among sets, and groups of objects that share something in common. Usually, Venn diagrams are used to depict intersections (denoted by an upside-down letter U). This type of diagram is used in scientific and engineering presentations, in theoretical mathematics, in computer applications, and in statistics.

In this lesson you will:

- learn about notation and illustration
- determine probabilities
- use Venn diagrams
- understand mutually exclusive, certain and impossible events

Notation and illustration

Example 1

Suppose we want to investigate the probabilities of drawing various kinds of numbers from a hat containing the numbers from 1 to 9. We call the choosing of a number from the hat an **event** or an **outcome**. A list of all the equally likely outcomes or events is called the **sample space**:

$$S = \{1; 2; 3; 4; 5; 6; 7; 8; 9\}.$$

The total number of outcomes in the sample space is written $n(S) = 9$

We might call the event or outcome of drawing the 6, A. Then A={6}, $n(A) = 1$ and the probability of drawing the 6, P(A)= $\frac{n(A)}{n(S)} = \frac{1}{9}$.

Some events can happen in more than one way:
Calling the event of drawing an even number B we have B = {2; 4; 6; 8}

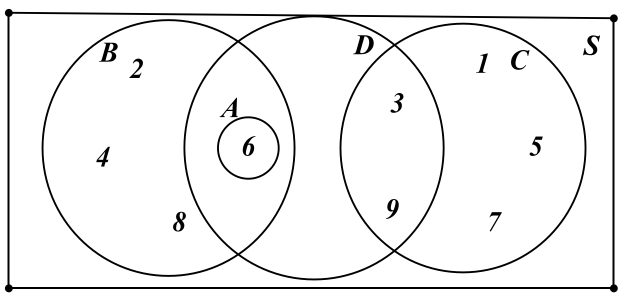
$n(B) = 4$ and $P(B) = \frac{n(B)}{n(S)} = \frac{4}{9}$.

If the event of drawing an odd number is C and that of drawing a multiple of 3 is D, we have:

C= $\{1; 3; 5; 7; 9\}$, $n(C) = 5$ and $P(C) = \frac{n(C)}{n(S)} = \frac{5}{9}$ and

D = {3; 6; 9), $n(D) = 3$ and $P(D) = \frac{n(D)}{n(S)} = \frac{3}{9} = \frac{1}{3}$

This situation can be illustrated as follows in what is called a Venn diagram:



The probability of events A or C occurring is denoted by $P(A \cup C)$.

$A \cup C = \{1; 3; 5; 6; 7; 9\}$, $n(P \cup C) = 6$ and $P(A \cup C) = \frac{6}{9} = \frac{2}{3}$

We read this as the probability of A **or** C or the probability of A **union** C is one third.

The probability of events D and C occurring is denoted by $P(D \cap C)$.

$D \cap C = \{3;9\}$, $n(D \cap C) = 2$ and $P(D \cap C) = \frac{2}{9}$.

We read this as the probability of D **and** C or D **intersection** C is two ninths.

The circle labelled A represents the event A and contains the number 6, circle B represents the event B and contains the even numbers, C contains the odd numbers and D the multiples of 3.

It is clear from the Venn diagram that if event B has occurred, C cannot also have occurred. If B occurs this excludes the possibility of C occurring and if C occurs, this excludes the possibility of B occurring. We call events which exclude each other **mutually exclusive**.

So P(B **and** C) = P(B ∩ C) = 0. Then events B and C are called mutually exclusive.

We can also see that P(A∩C) = 0, so events A and C are also mutually exclusive.

But P(D and C) = $\frac{n(D \cap C)}{n(S)} = \frac{2}{9}$. So D and C are not mutually exclusive events.

Now try this example.

ACTIVITY 1

1. An experiment consists of tossing a fair coin and throwing a fair dice. Describe the sample space S and find nS.

2. A box contains eight similar balls marked A to H. An experiment involves taking a ball at random from the box. Event A is taking a ball marked with a consonant, event B is the event of drawing a ball marked with a vowel and event C is taking the ball marked with the letter C.
   a). Write down the sample space.
   b) Which of the events A, B and C are mutually exclusive?
   c) Write down the probability of drawing a vowel.
   d) Write down $n\ (B \cup C)$.
   e) Write down $P\,(B \cap C)$.

ANSWERS ON PAGE 56

Example 2

An experiment consists of tossing a fair coin and throwing a fair dice. Describe the sample space S and find n(S). Determine the probability of throwing

a) a head and a number greater than three;
b) a head or a tail and a number less than or equal to six;
c) a tail and a number greater than 6.

Solution

Tossing a coin can yield two possible outcomes: heads (H) and tails (T). Also throwing a fair dice can yield six possible outcomes:1, 2, 3, 4, 5 and 6.

Each of the outcomes of the coin can be combined with any of the outcomes of the dice. Therefore there should be $2 \times 6 = 12$ possible outcomes for the experiment of throwing a coin and a dice.

To get all the possible outcomes you can make a table with one set of outcomes on one side and the other set on the other side and combine them.

	1	2	3	4	5	6
H	H1	H2	H3	H4	H5	H6
T	T1	T2	T3	T4	T5	T6

Hence S = {H1; H2; H3; H4; H5; H6; T1; T2; T3; T4; T5; T6}
and n(S) = 12.

a) Let the event of throwing head*s* and a number greater than 3 be A.

Then $A = \{H4; H5; H6\}$, $n(A) = 3$ and $P(A) = \frac{3}{12} = \frac{1}{4}$

b) Let the event of throwing heads or tails and a number less than or equal to six be B. Every event in the sample space satisfies the requirements so the event $B = S$ and $P(B) = \frac{12}{12} = 1$.

Such events, where the probability is one, are called **certain events**.

c) Let the event of throwing tails and a number greater than 6 be C. Since there are no numbers greater than six on a regular dice, this event is impossible. $P(C) = \frac{0}{12} = 0$. Such events, where the probability is zero, are called **impossible events**.

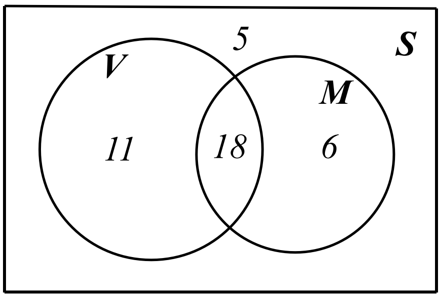
Example 3

a) Draw a Venn diagram to illustrate the following situation: of the 40 people at a staff party, 24 ate meat, 29 ate vegetables and 5 did not eat either meat or vegetables.

b) Determine the probability that a randomly chosen person at the party ate meat and vegetables.

Solution

a) Since 5 people did not eat meat or vegetables, the remaining 35 ate either meat or vegetables or both meat and vegetables, 29 of these 35 ate vegetables, so the other 6 ate only meat. Since 24 ate meat, the other $(24-6)=18$ who ate meat, also ate vegetables. This means that $(29-18)=11$ ate only vegetables.



In this case, we complete the Venn diagram showing the number of ways each event can occur, letting letting the event that a staff member ate meat be M and that s/he ate vegetables be V.

b) $P(M \cap V) = \frac{18}{40} = \frac{9}{20}$

In general, for any two events A and B,

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

In the given example:

$$P(V \cup M) = \frac{11+18+6}{40} = \frac{35}{40}$$

$$P(V) = \frac{29}{40}, \quad P(M) = \frac{24}{40} \text{ and } P(V \cap M) = \frac{18}{40}$$

$$\text{So } P(V) + P(M) - P(V \cap M) = \frac{29+24-18}{40} = \frac{35}{40}$$

and so $P(V \cup M) = P(V) + P(M) - P(V \cap M)$

The general case can be proved easily.

Notice that when the two events A and B, say, are mutually exclusive then $P(A \cup B) = P(A) + P(B)$, since $P(A \cap B) = 0$

Now try these examples:

ACTIVITY 2

1. A and B are events with P(A) = 0,2; P(B) = 0,16 and P(A and B) = 0,04.
   Notice that P(A and B) is often written $P(A \cap B)$

   a) Decide whether A and B are mutually exclusive events and justify your decision.
   b) Calculate P(A or B). This is also written $P(A \cup B)$.
   c) Calculate P(not A and B). This is also written $P(A' \cap B)$.

2. There are 25 students who are writing the National Senior Certificate. All these students are writing either English or Afrikaans or both languages. 21 are writing English and 8 are writing Afrikaans.

What is the probability that a randomly chosen student is writing:

a) both languages
b) English but not Afrikaans
c) neither English nor Afrikaans?

ANSWERS ON PAGE 56

Example 4
A fair dice is rolled. Find the probability of obtaining either a 1 or an even number.

Solution
Let A be the event of obtaining a 1, and B the event of obtaining an even number. A and B are mutually exclusive because the two cannot occur at the same time.

Now $P(A)=\frac{1}{6}$ and $P(B)=\frac{3}{6}$ and since A and B are mutually exclusive,

$$P(A\cup B)=P(A)+P(B)$$
$$=\frac{1}{6}+\frac{3}{6}$$
$$=\frac{2}{3}$$

ACTIVITY 3

1. A regular deck of cards has 52 cards. 26 are red and 26 are black. 13 of the red cards are hearts and the other 13 are diamonds. 13 of the black cards are spades and the other 13 are clubs. Each of the suits hearts, diamonds, spades and clubs has an ace, a king, a queen, a jack and the numbers 2 to 10 making up the 13 cards.

Consider the following events:
- A: drawing a red card;
- B: drawing a black card;
- C: drawing a spade;
- D: drawing a picture card.

a) Write down two events which are mutually exclusive but not complementary.
b) Write down a pair of complementary events.
c) Using the union or intersection of two of the given events write down :
   i) an impossible event
   ii) a certain event
d) Write down the probability of randomly drawing:
   i) a spade or a picture card
   ii) a picture card which is a spade.

ANSWERS ON PAGE 57

Summary

You have been introduced to concepts such as sample space, elementary events, union and intersection in probability theory. You learnt about the probability of events and the number of ways an event can occur. You have also learnt the identity for any events A and B and that this identity can be used to calculate other events. You have seen how two events are mutually exclusive if they have no outcome in common, and finally that some events are mutually exclusive and exhaustive.

CHECKLIST

Are you able to

- draw a Venn diagram to illustrate some probability problems;
- use the identity $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ and facts listed above to calculate probabilities;
- understand the idea of mutually exclusive events, certain events and impossible events.

SELF-CHECK EXERCISE

1. An experiment consists of tossing a fair dice.

   a) List the outcomes of the sample space S.
   b) Let A = {1, 3, 5}, B = {2, 4, 6}, C = {1, 6}and D = {4}.

      i) Find P(A), P(B), P(C) and P(D)
      ii) List the pairs of events which are mutually exclusive
      iii) Find $P(A \cap B)$, $P(B \cap C)$ and $P(A \cap C)$
      iv) Find $P(A \cup B)$ and $P(C \cup D)$

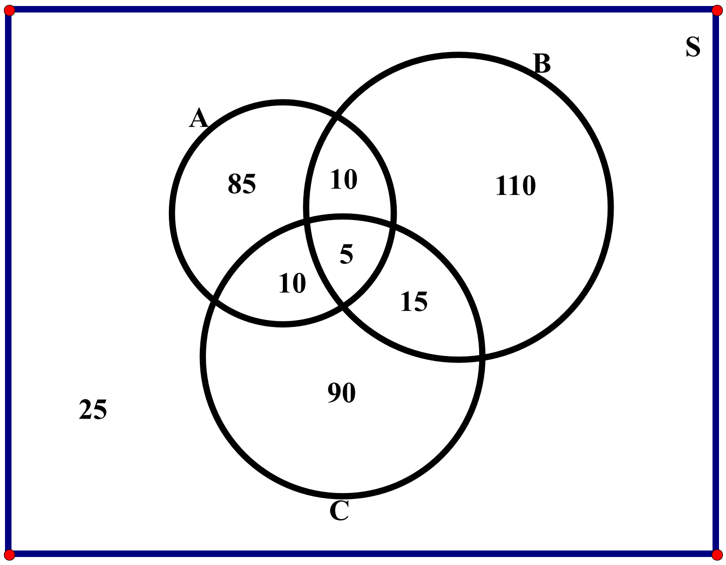
2. Out of 40 people, 30 own a car and 25 own their house. If you are told that every one of the 40 people has a car or a house or both, find the probability a person selected from the 40 people owns both a house and a car.

3. Consider an experiment of throwing two fair dice.

   a) Describe a sample space, S, for this experiment
   b) Find n(S)
   c) Determine the probability:
      i) that the sum of the two numbers showing on the two dice is 8;
      ii) that the numbers showing on both dice are the same;
      iii) of getting at least one six.

4. Calculate from the given Venn diagram in which the number of equally likely ways the events can occur has been filled in:

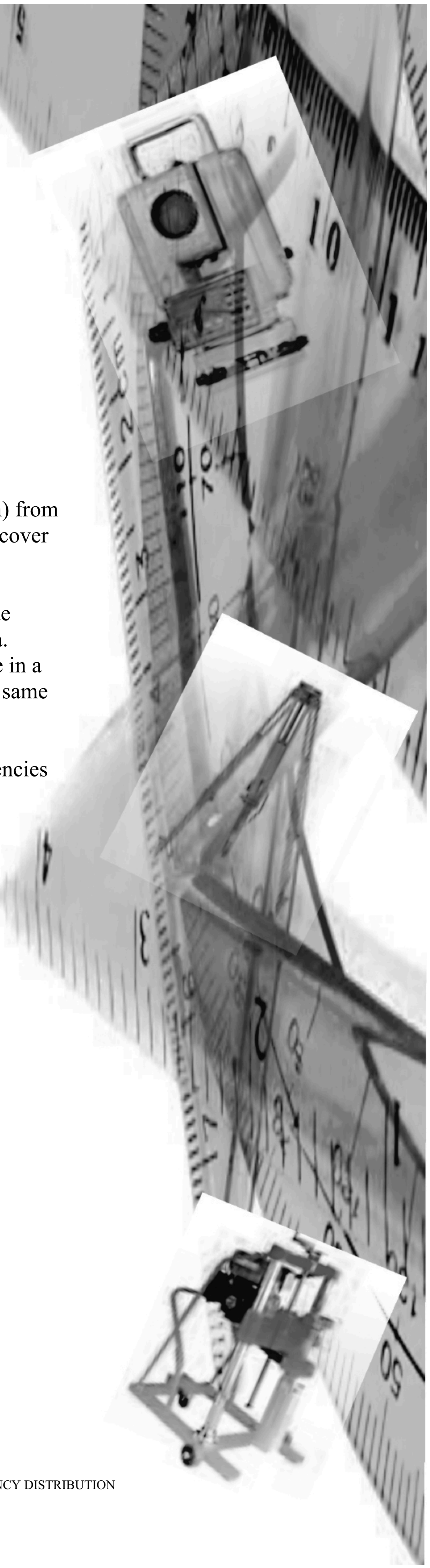
a) $P(A \cap B \cap C)$
b) $P(A \cap B \cap C')$
c) $P(A \cup B)$
d) $P(B \cup C \cap A')$



ANSWERS ON PAGE 65

LESSON 4

Frequency and cumulative frequency distribution



About this lesson

In this lesson you will learn to draw a frequency polygon (graph) from a frequency table. You will also extend the idea of frequency to cover cumulative frequency.

Ordinary frequency distribution tables describe a particular value according to how many times that value occurs in the set of data. However, we may need to know, for example, how many people in a certain company are older or younger than a particular age. The same considerations could apply to other sets of statistical data. The cumulative frequency takes care of such considerations. The cumulative frequency at a value is the running total of all frequencies up to that value.

In this lesson you will:

- draw a frequency polygon;
- construct a cumulative frequency table from a frequency table;
- draw and interpret cumulative frequency curves.

The frequency polygon

You will remember that a frequency table gives the frequency or number of times each event happens. Look at the frequency table in example 1 on the next page. You are going to learn how to draw a graph called a frequency polygon from the frequency table.

Steps in drawing a frequency polygon

1. Plot the vertical and horizontal axes. The vertical axis should be used for the frequency and the horizontal axis for the variable under discussion.
2. Use the frequency table to plot points on the grid..
3. Join these points with straight lines.
4. Join the end-points to the horizontal axis in order to form a polygon. You must also give a suitable heading to your frequency polygon.

Here is an example:

Example 1

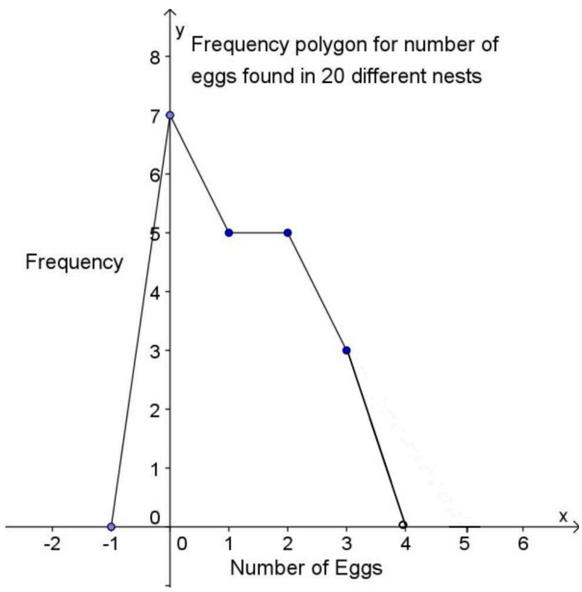
A zoologist wants to study the way birds in a certain area breed. She counts the eggs found in 20 different nests at the start of the breeding season and obtains the frequency table below. Draw a frequency polygon for the frequency table she obtained.

Number of eggs	Frequency
0	7
1	5
2	5
3	3
TOTAL	20

Solution

1. Draw the axes with the vertical axis 'frequency' and the horizontal axis 'number of eggs' and heading 'Frequency polygon for the number of eggs found in 20 different nests'.

polygon:
figure with any number of straight sides. So a four sided polygon is a quadrilateral and a six sided polygon is a hexagon.



2. Plot the points. For example, when the number of eggs is 0, the frequency is 7. Therefore plot the point (0,7). Using the same idea plot the points (1,5), (2,5) and (3,3) on the grid drawn.

3. Join these points with straight lines.

4. In order for our diagram to be a polygon it must be joined to the horizontal axis. Supposing we were to extend our data on both sides, then the number of eggs to consider are –1 and 4, but the zoologist did not find any nests with –1 (impossible) egg or 4 eggs. Therefore the frequencies for these numbers are 0. Plot these on your diagram [(–1; 0) and (4; 0)] and join them to get the frequency polygon.

Now try the following activity.

ACTIVITY 1

A Life Orientation teacher wants information about the home life of her pupils. To do so she needs to know what size their families are. She therefore asks her pupils to give the number of children in their families. She then made the following frequency table.

Number of children in a family	Frequency
1	3
2	7
3	6
4	2
5	1
6	1
Total	20

ANSWERS ON PAGE 58

Draw a frequency polygon for her data.

Frequency polygons for grouped data

The same procedure is used for drawing the frequency polygon for grouped data. The only difference is that when it comes to plotting the points (as in Step 3) we plot the points using the midmarks of the individual classes.

The midmark (or the classmark) of a class is the midpoint of the class. For example the midmark of the interval '1 – 5' is 3. In general we obtain the midmark of the class by adding the endpoints of the class and dividing by two. In the example given, the class mark is $\frac{1+5}{2} = 3$

Example 2
The final marks for 30 students of a Adult Education Institute were put on a frequency table:

Construct a frequency polygon of the marks.

Marks	Frequency
36 - 40	3
41 - 45	4
46 - 50	6
51 - 55	10
56 - 60	6
60 - 65	1
TOTAL	30

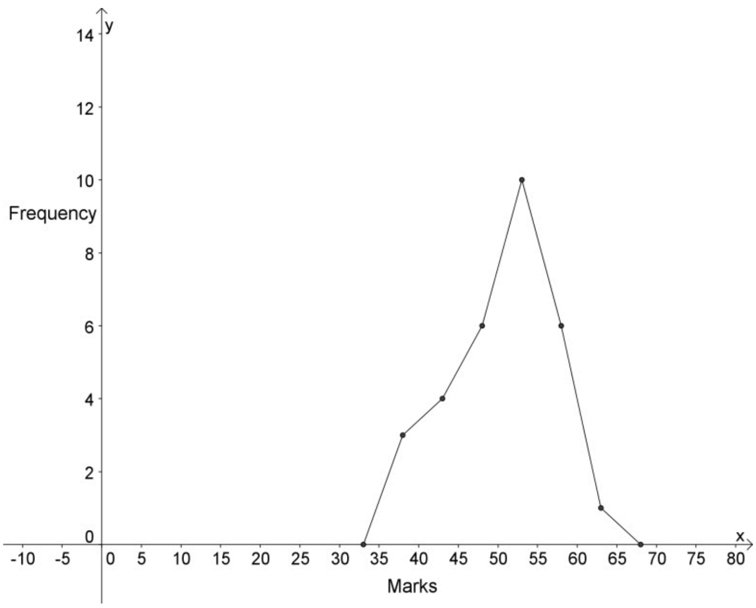
Solution
Before we draw the frequency polygon, we must find the class marks (midmark) for the groups. The class mark for a group is found by adding the two class end-points and dividing by 2. Therefore the class mark for the group '36 – 40' is $\frac{36+40}{2} = 38$

Continue in this way and you get:

Marks	**Midpoint**	**Frequency**
36 - 40	38	3
41 - 45	43	4
46 - 50	38	6
51 - 55	53	10
56 - 60	58	6
60 - 65	63	1
TOTAL		30

We then draw the polygon using the class marks for the horizontal axis.

Frequency polygon for the number of marks achieved by 30 students



Now try the question in the next activity.

ACTIVITY 2

The number of cars arriving at a petrol station each day over a period of 25 days was recorded as:

Number of cars per day	Number of days/ Frequency
30 - 39	5
40 - 49	10
50 - 59	6
60 - 69	3
70 - 79	1
TOTAL	25

Draw a frequency polygon for the data.

ANSWERS ON PAGE 58

Cumulative frequency

We can obtain the cumulative frequency at a value by adding all the frequencies up to that value. So, the cumulative frequency at a value is the running total of all the frequencies up to that value.

Example 3
Construct a cumulative frequency table for the frequency table in Example 1.

Solution

Number of eggs	Frequency	Cumulative frequency
0	7	7 = 7
1	5	7+5 =12
2	5	7+5+5 =17
3	3	7+5+5+3 =20
TOTAL	20	

The cumulative frequency for the last value should always be equal to the total frequency. In our example the cumulative frequency for 3 eggs is 20 which is the same as the total frequency.

Now try to do a cumulative frequency table yourself.

ACTIVITY 3

The table on the next page indicates the number of years 40 employees of a company in the Eastern Cape province have worked for the company. Obtain a cumulative frequency table from this table.

Number of years	Frequency
1	4
2	2
3	5
4	4
5	2
6	7
7	5
8	6
9	3
10	2
TOTAL	40

ANSWERS ON PAGE 59

Cumulative frequency graph

We shall concentrate on drawing the cumulative frequency graph for the grouped data. You can use the same idea for the individual data. Just as we did for the frequency polygon, we plot the points but in cumulative frequency graphs, a curved line is drawn through the points. Also in the cumulative frequency graph, the vertical axis is the cumulative frequency and the horizontal axis is the variable under discussion.

Example 4
Let us use Example 2 to draw a cumulative frequency graph.

The final marks of 30 students of an Adult Education Institute were put into a frequency table:

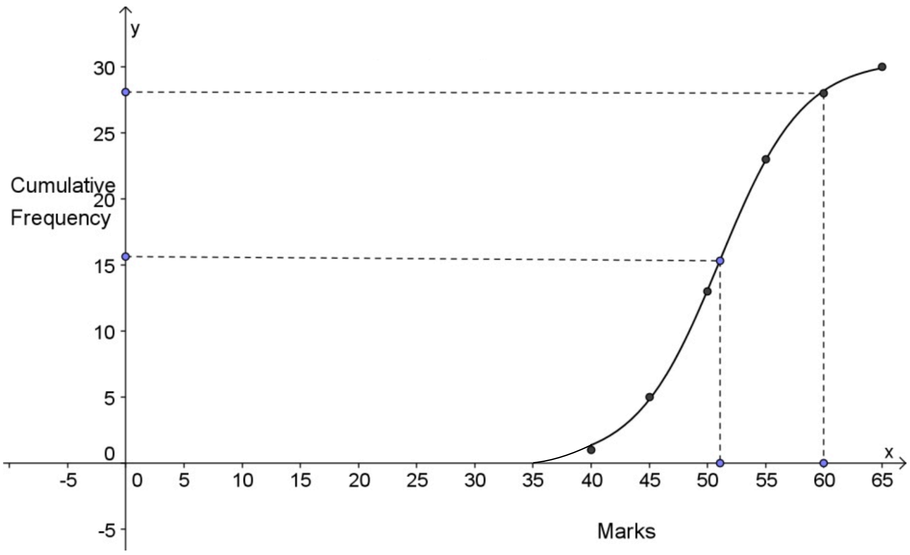
Marks	Frequency
36-40	3
41-45	4
46-50	6
51-55	10
56-60	6
61-65	1
TOTAL	30

Construct a cumulative frequency graph for the data.

Solution
In plotting the cumulative frequencies against the marks we use the highest marks for each class. This is because when we say that the cumulative frequency of the class '41 – 45' is 3 + 4 = 7, we can only say that 7 people had marks up to 45. In other words 7 people had marks less than or equal to 45.

Marks	Frequency	Cumulative frequency
36-40	3	3
41-45	4	7
46-50	6	13
51-55	10	23
56-60	6	29
61-65	1	30
TOTAL	30	

Cumulative Frequency Graph for Students Marks



You can see from the graph that the cumulative frequency for 35 is 0. This is because from the frequency table we can say that nobody had marks below 36.

We can use a cumulative frequency graph to answer questions, such as:

1. How many students obtained fewer than 60 marks?

2. What is the median mark?

To answer the first question we need to draw a vertical line starting from the 60 mark to meet the cumulative frequency graph. We then draw a horizontal line from this point to cut the cumulative frequency axis (or y –axis) and read off the value there. As you can see from the graph, the answer is '29 students'.

*The **median** is the number which divides data that is arranged in ascending order into two equal groups of data.*

From the second question we know that 30 students wrote the exams and the median mark is between the 15th and 16th mark so that there are 15 in the lower group and 15 in the upper group.

So we draw a horizontal line, starting from 15,5 on the cumulative frequency axis to cut the graph. We then draw a vertical line to cut the 'mark' or x –axis and read off the value there as illustrated on the graph. The answer is 'the median mark is approximately 52'.

ACTIVITY 4

100 companies in Gauteng were asked for the number of people they employed. The following table records the data.

Number of employees	Frequency
1 - 50	5
51 - 100	15
101 - 150	35
151 - 200	20
201 - 250	25
TOTAL	100

Draw a cumulative frequency graph for the data and use it to find how many companies employ fewer than 100 workers.

ANSWERS ON PAGE 59

Summary

You learnt how to draw frequency polygons from a frequency table and extended the idea of frequency to cover cumulative frequency. These frequency tables enabled you to determine sets of statistical data.

CHECKLIST

Are you able to:

- draw the frequency polygon for any frequency distribution (table)
- explain what cumulative frequency means
- draw a cumulative frequency graph
- answer simple questions based on cumulative frequency graphs

SELF-CHECK EXERCISE

1. A statistics student was asked to do a project of his choice. He decided to find the height of 20 people standing next to him. He obtained the following heights in cm.

151	162	174	168	185	156	172
167	144	162	148	176	166	171
160	168	181	157	174	174	

a) Complete the grouped frequency table for these statistics:

Height (cm)	Frequency	Cumulative frequency
141 - 150		
151 - 160		
161 - 170	7	
171 - 180		
181 - 190		20
	20	

b) Use the grouped frequency table obtained in a) to draw a frequency polygon for the data.

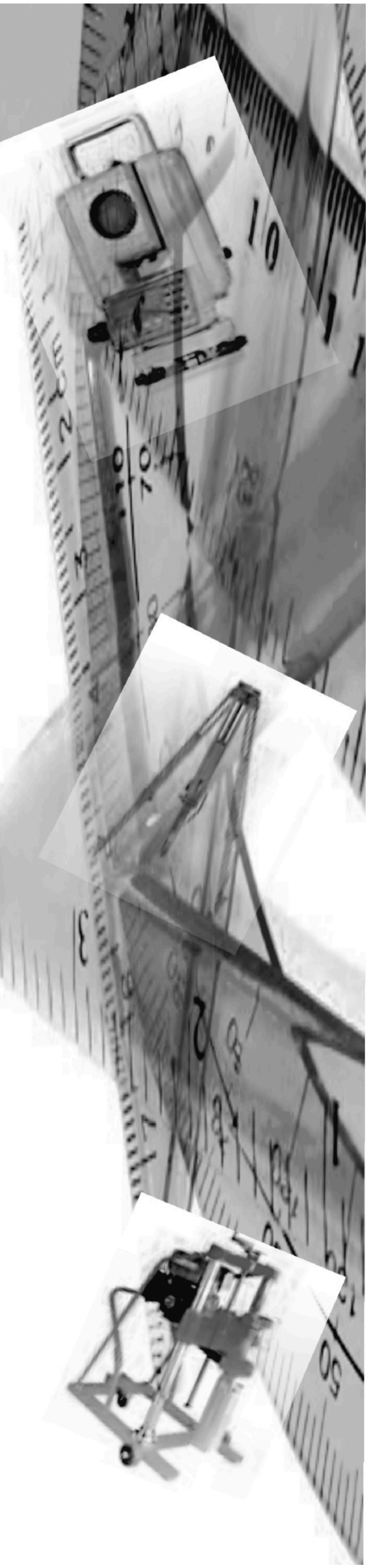
c) Draw a cumulative frequency graph for the data.

d) Use the cumulative frequency graph to find how many people had heights less than 180 cm.

ANSWERS ON PAGE 66

LESSON 5

Permutations and Combinations



About this lesson

Welcome to the final lesson of the Mathematics course. We hope you have enjoyed all the lessons so far.

You learned that the probability of an event happening is given by the formula:

$$\text{Probability} = \frac{\text{number of ways the event can happen}}{\text{number of equally likely outcomes}}$$

This means that before you can calculate probabilities well, your counting techniques should be good. Counting techniques give us a formal structure we can use to determine the number of ways in which an event can occur. This helps to reduce the trial and error or guesswork that otherwise might be involved in determining the occurrence of an event.

Permutations and combinations are two counting techniques often used in mathematics.

In this lesson, you will:

- state the difference between permutations and combinations
- use factorial notation
- solve permutation and combination problems using factorial notation
- use permutations and combinations to solve probability problems

Permutation, combination and the difference between them

Permutation

Permutation is defined as the variation (change) in the order of a set of things. For example if three pictures labelled A, B and C are to be hung in line on a wall, we may have them in this order A, B and C. Another order is A, C, B. Each of these is a permutation.

Can you list all the possible ways in which the pictures can be arranged on the wall?

The possible ways of hanging the three pictures are:
A, B, C, A, C, B, B, A, C, B, C, A, C, A, B and C, B, A

Therefore, there are six possible ways of hanging the pictures on the wall which means we have six permutations.

Combination

A combination, on the other hand, is an unordered (random) selection of a number of items from a given group of items (a set). For example, consider again the three pictures to be hung on a wall.

If these three pictures A, B, C were chosen from a set of six pictures A, B, C, D, E and F, then all six different permutations be regarded as only one combinations (a combination of the pictures A, B and C). In this case, what we want are the pictures A, B and C. So the order in which they are chosen, A B C, A C B etc is not important. Other combinations may be A B D, B C D, C D E, D E F, etc.

Note that each of these other combinations also have six permutations. For example the combination B C D has permutations B C D, B D C, C B D, C D B, D B C and D C B.

The difference between permutations and combinations

In this section, you are going to learn different methods for finding the total number of ways of arranging items or choosing groups of items from a set. But before you do this, it is important that you know the difference between permutations and combinations. Consider this example.

permutation:
ordering or placing of a number of objects where the order in which you place the objects is important

combination:
collection of a number of objects from a given group of objects where order is not important.

Example

A bookseller has 7 different books but only 4 spaces on his shelf to display the books so he can display only four books at a time. He clearly cannot display all 7 books at the same time. He must choose 4. The order in which he chooses the books is not important. The set of the 4 chosen books is only a combination. After making his choice, he can now place his 4 books in different order on the shelves. That is, each arrangement is a permutation.

ACTIVITY 1

Say (without working out the answers) whether each of the following problems involves a permutation or a combination.

1. A team of six members is chosen from a group of eight people. How many different teams can be selected?

2. The first, second and third prizes for a raffle are awarded by drawing tickets from a box of 100. In how many ways can the prizes be won?

3. A Cape Town telephone number (after the 021) is a seven digit number. How many Cape Town telephone lines are available?

4. A particular department in a large firm has 4 vacancies for a position. A shortlist of 8 applicants is drawn up. In how many ways can the 4 vacancies be filled from the list of 8 applicants?

ANSWERS ON PAGE 60

Permutations in different types of problems

We will now consider ways of finding a number of different permutations in various types of problems.

Example 1

How many arrangements of the letters A, B and C are there?

Solution

The first letter can be

A
or
B
or
C

} so there are three ways of choosing the first letter

After choosing the first letter we are now left with two letters to choose the second letter from:

1st	2nd	
A	B or C	For each of the three ways of choosing the first letter, there are two ways of choosing the second. So the number of ways of choosing the first two letters will be the product of the number of ways of choosing the first letter and the number of ways of choosing the second letter. Therefore, there are 3 × 2 ways of choosing the first two letters.
or B	A or C	
or C	A or B	

After choosing the first two there is only one letter left from which to choose, the third letter. Therefore there is only one way of choosing the third letter. Therefore there are $3 \times 2 \times 1 = 6$ ways of arranging the three letters A, B and C.

Let's take another example.

Example 2

How many three-digit-numbers can be made from the integers 1, 2, 3, 4, 5 if each integer is used only once?

Solution

The first digit can be any of the five numbers: 1, 2, 3, 4, 5. So there are five ways of choosing the first digit. After we have written the first digit, we are left with 4 digits from which to choose the second.

Therefore there are 4 ways to choose the second digit. The next digit can be taken from the remaining three digits. So there are 3 ways to choose the third digit. Therefore there are $5 \times 4 \times 3 = 60$ ways of choosing a 3 digit number from the five digits if each digit appears only once.

Now suppose the restriction 'each integer is used only once' is removed from the question, how many three-digit numbers can we get from the five integers? Since there is no restriction on the number of times an integer can be selected, it means each integer can be selected three times. This means there are 5 ways of choosing the integer for the first place. Because we can choose that integer again, we still have five integers from which to choose for the second place. In the same way, there are five ways of choosing the third integer. Therefore there are $5 \times 5 \times 5 = 125$ ways of choosing a three-digit number from 5 integers if there is no restriction on how many times an integer can be selected.

ACTIVITY 2

1. The first, second and third prizes for a raffle are awarded by drawing tickets from a box of 100 tickets. In how many ways can the prizes be won?
2. A South African telephone number is a three digit area code followed by a seven digit number. How many South African telephone lines are available?
3. In a group of five men one is a burglar. The 5 men take part in an identity parade, to see whether a witness can identify the burglar. In how many ways can the men be lined up if they can stand in any order?

ANSWERS ON PAGE 60

Combination problems

Let us look at the number of arrangements of four different books on a shelf. We know that the number of permutations of the four books is

$4 \times 3 \times 2 \times 1 = 24$

But if the order does not matter, there will only be one combination of the books. That is, the four books can be selected in 24 different ways if we consider the order in which they were picked. But each of the 24 ways will lead to choosing the same set of 4 books so we have only one set (combination) of 4 books chosen in 24 different ways (permutations).

Now suppose there are five different books available and only four of them can be placed on the shelf, then the total number of ways in which this can be done is $5 \times 4 \times 3 \times 2 = 120$ permutations.

But for each particular set of four books, $4 \times 3 \times 2 \times 1$ permutations = one combination

To get the number of different sets (combinations) of four books that can be chosen from the five available books we have to divide the total number of ways of selecting four books out of five (total number of permutations) by the number of ways a particular set can be selected (number of permutations within a set).

So the number of different sets (combinations) of four books that can be chosen from the five available books is given by

$$\frac{\text{total number of permutations}}{\text{number of permutations of each set}}$$

In general, if we have *n* objects from which we select groups of *r* objects, the total number of possible groups (combinations) is given by

$$\frac{\text{number of permutations of } r \text{ objects from } n \text{ objects}}{\text{number of permutations of } r \text{ objects}}$$

Example 3

A shop stocks ten different varieties of soup in packets. In how many ways can a shopper buy three packets of soup if each packet is a different variety?

Solution

Number of ways of selecting 3 packets of soup from 10 (total number of permutations) is $10 \times 9 \times 8 = 720$

Number of ways of arranging one set of 3 packets of soup (number of permutations within a set) is $3 \times 2 \times 1 = 6$

Therefore the number of ways the shopper can buy (choose) 3 different packets of soup out of 10 packets is

$$\frac{720}{6} = 120$$

Now try some problems yourself.

ACTIVITY 3

1. A team of six members is chosen from a group of eight people. How many different teams can be selected?

2. A team of 4 is chosen from 8 players. How many different teams can be chosen?

3. In how many ways can five boys be chosen from a class of twenty boys if the class prefect has to be included?

ANSWERS ON PAGE 61

The factorial notation

Let us consider the number of ways of arranging a set of 52 playing cards in a row. This gives us

52×51

which is a very large number, difficult to multiply out and takes a long time and much space to write out in full. We simplify this sum by writing it as 52! which means '52 factorial'.

In general, *n*! represents the number $n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1$.

That is $n!$ means the product of all the integers from 1 to n inclusive (including 1 and n) and is called ***n* factorial**.

There is a button (x! or n!) on your calculator that will produce this number for you.

n factorial:
$$n! = n(n-1)(n-2)\ldots 3.2.1$$

Example 4
Evaluate the following:

a) $10!$

b) $\frac{7!}{4!}$

c) $\frac{20!}{15!5!}$

Solution

a) Press the following on your calculator
[10] [2nd F] [x!] [=] and you will get 3 628 800

b) $\frac{7!}{4!} = \frac{7\times 6\times 5\times 4\times 3\times 2\times 1}{4\times 3\times 2\times 1} = 7\times 6\times 5 = 210$

c) You can rewrite the expression as $\frac{20\times 19\times 18\times 17\times 16\times 15!}{15!\times 5\times 4\times 3\times 2\times 1}$

Cancelling 15! in both the numerator and denominator gives
$$\frac{20\times 19\times 18\times 17\times 16}{5\times 4\times 3\times 2\times 1} = 15\ 504$$

There is one particular property of the factorial notation that you should always have in mind.

That property is $0! = 1$
Verify this using your calculator!

Sometimes it is cumbersome to evaluate an expression involving a long chain of multiplication. In such cases, it may be easier to rewrite the expression in factorial notation before using the calculator to evaluate.

Example 5
Write the following in factorial notation:

a) $12\times 11\times 10\times 9\times 8\times 7\times 6\times 5$

b) $\frac{9\times 8\times 7\times 6\times 5\times 4}{5\times 4\times 3\times 2\times 1}$

Solution

a) This can be written as

$$12\times 11\times 10\times 9\times 8\times 7\times 6\times 5\times \frac{4\times 3\times 2\times 1}{4\times 3\times 2\times 1}$$

$$=\frac{12!}{4!}$$

b) $$\frac{9\times 8\times 7\times 6\times 5\times 4}{5\times 4\times 3\times 2\times 1}\times \frac{3\times 2}{6}$$

$$=\frac{9!}{6!}$$

ACTIVITY 4

1. Write the following in factorial notation:
   a) $5 \times 4 \times 3$
   b) $\frac{10\times 9\times 8}{3\times 2\times 1}$

2. Evaluate:
   a) $\frac{8!}{(4!)^2}$
   b) $\frac{13!}{0!}$

ANSWERS ON PAGE 61

Now let's see how to use this notation in problems involving permutations and combinations.

Permutations and combinations in factorial notation

We saw earlier that the number of different **permutations** of three objects chosen from five different objects is $5 \times 4 \times 3$. Using factorial notation we can write this as

$$\frac{5\times 4\times 3\times 2\times 1}{2\times 1}=\frac{5!}{2!}=\frac{5!}{(5-3)!}$$

In general the number of permutations of r objects chosen from a set of n different objects is

$$\frac{n!}{(n-r)!}$$

Using ${}_nP_r$ as a symbol for the number of permutations of r objects chosen from a set of n different objects we write

$${}_nP_r=\frac{n!}{(n-r)!}$$

Example 6
How many two digit numbers can be made from the set $\{2; 3; 4; 5; 6; 7; 8; 9\}$, each number containing two different digits?

Solution
There are eight different digits in the set (so $n = 8$) from which we need to select two digit numbers (hence $r = 2$). Substituting the values of n and r into the above, we have

$$_8P_2 = \frac{8!}{(8-2)!} = \frac{8!}{6!} = 8 \times 7 = 56$$

We also saw earlier on that the number of different **combinations** of 3 that we can get from a set of 10 is $\frac{10 \times 9 \times 8}{3 \times 2 \times 1}$

This can be written using factorial notation as

$$\frac{10!}{3!7!} \text{ which is the same as } \frac{10!}{3!(10-3)!}$$

In general the number of combinations of r objects from a set of n objects is $\frac{n!}{r!(n-r)!}$

We usually denote the combinations of r objects from n objects by either

^nC_r or $_nC_r$ or $\binom{n}{r}$.

Therefore the combination of r objects from a set of n objects is given

$$^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Example 7
A sport club in Guguletu has 20 members

a) In how many ways can they choose an executive committee consisting of 3 members from the club?
b) In how many different ways can they choose a president, secretary and treasurer if all the nominations are eligible (or are equally suited) for the positions available?

Solution
a) Here it is just a matter of selecting. Order is not important. Therefore we have a combination problem.

The answer is $_{20}C_3 = \frac{20!}{3!17!} = \frac{20 \times 19 \times 18}{3 \times 2 \times 1} = 1\ 140$

b) Here order is important because choosing a president is different from choosing a secretary or the treasurer. So the problem involves permutation. We have

$$_{20}p_3 = \frac{20!}{(20-3)!} = \frac{20!}{17!} = 20 \times 19 \times 18 = 6\ 840$$

ACTIVITY 5

1. In how many ways can the manager of a small company choose any 3 of the 15 employees for promotion?
2. In a large class there are 10 students who always score the best marks in the class. In how many ways can the teacher place 3 of these 10 students in first, second and third place in the class?

ANSWERS ON PAGE 62

Multiplication of choices

Suppose a taxi company has 10 taxis and 5 drivers on duty. In how many ways can the company assign a taxi and a driver to a customer?

There are two steps involved in making this decision (choice).

1. Choose a taxi
2. Choose a driver

The number of ways in which the company can make the choice is the product of the number of ways the first step (n_1) can be made and the number of ways the second step (n_2) can be made. Since the company can choose any of the 10 taxis, the first step (choosing a taxi) can be made in 10 ways which means $n_1 = 10$. Similarly, the second step (choosing a driver) can be made in 5 ways or $n_2 = 5$. Therefore the number of ways the company can assign a taxi and a driver to a customer is $10 \times 5 = 50$.

In general, if a choice consists of K steps, then the first step can be made in n_1 ways, for each of these the second can be made in n_2 ways, ... and the Kth step can be made in n_K ways. The whole choice can therefore be made in $n_1 \times n_2 \times n_3 \times ... n_K$ ways.

Example 8
A choir has 30 members. 20 of them are women and the rest are men. In how many ways can the church choose an executive committee consisting of 5 members if 3of the members must be women and 2 men?

Solution
The number of ways the church can choose the 3 women from 20 women is ${}_{20}C_3$. The number of ways the church can choose 2 men from the 10 men is ${}_{10}C_2$.
Therefore the number of ways the church can choose the executive committee is

$$ {}_{20}C_3 \times {}_{10}C_2 = \frac{20 \times 19 \times 18}{3 \times 2 \times 1} \times \frac{10 \times 9}{2 \times 1} = 51\ 300 $$

ACTIVITY 6

1. A patient requires a prescription for an antibiotic and a painkiller. If there are 5 suitable antibiotics and 6 suitable painkillers and the doctor can prescribe any one of the antibiotics with any one of the painkillers in how many different ways can an antibiotic and a painkiller be prescribed?
2. A tennis club is to select a team of 4 people (two women and two men). The coach has to choose the team from 7 men and 5 women. In how many different ways can he select the team?

ANSWERS ON PAGE 62

Lastly we are going to see how we can use what we've learnt about permutations and combinations to deal with probability problems.

Application of permutations and combinations in probability theory

In Lesson 2 of this unit you used tree diagrams to solve probabilities of combined events. In this section of the lesson you will learn to use the idea of permutation and combination to solve similar problems. We shall use a few examples from lesson 2 to illustrate this.

Example 9
You choose a committee of 2 people from a group of 6 women and 4 men. Use the idea of permutation and combination to calculate the probabilities of getting

a) 2 women
b) 2 men
c) one woman and one man

Solution
You solved this problem in lesson 2 by using tree diagrams. Let's see how it can be solved differently.

a) Total number of ways of choosing 2 people from 10 people is

$${}_{10}C_2 = \frac{10 \times 9}{2 \times 1} = 45$$

Number of ways of choosing 2 women from 6 women is

$${}_{6}C_2 = \frac{6 \times 5}{2 \times 1} = 15$$

Therefore the probability of getting two women from the 10 people is

$$\frac{\text{number of ways of choosing two women}}{\text{total number of ways of choosing two people}} = \frac{15}{45} = \frac{1}{3}$$

b) Number of ways of choosing 2 men from the 4 men is

$${}_{4}C_2 = \frac{4 \times 3}{2 \times 1} = 6$$

Therefore the probability of choosing two men from the 10 people is

$$\frac{\text{number of ways of choosing two men}}{\text{total number of ways of choosing two people}} = \frac{6}{45} = \frac{2}{15}$$

c) Number of ways of choosing one woman from the six women is ${}_{6}C_1 = 6$.
In the same way the number of ways of choosing the man from the four men is ${}_{4}C_1 = 4$.

So for multiplication of choices, the number of ways of choosing one woman and one man is 6 × 4 = 24.

Therefore the required probability is $\frac{{}_{6}C_1 \times {}_{4}C_1}{{}_{10}C_2} = \frac{6 \times 4}{45} = \frac{8}{15}$.

Example 10

A bag contains 5 red balls and 6 blue balls. 2 balls are picked from the bag without replacement. Find the probabilities that

a) both balls are red

b) both balls are blue

c) one is red and the other blue.

Solution

Total number of ways of choosing 2 balls from the 11 balls is

$${}_{11}C_2 = \frac{11 \times 10}{2 \times 1} = 55.$$

a) Number of ways of choosing 2 red balls from 5 is

$${}_5C_2 = \frac{5 \times 4}{2 \times 1} = 10.$$

Therefore the probability that both balls are red is

$$\frac{{}_5C_2}{{}_{11}C_2} = \frac{10}{55} = \frac{2}{11}.$$

b) Number of ways of choosing 2 blue balls from 6 is

$${}_6C_2 = \frac{6 \times 5}{2 \times 1} = 15.$$

Therefore the probability that both balls are blue is

$$\frac{{}_6C_2}{{}_{11}C_2} = \frac{10}{55} = \frac{3}{11}.$$

c) Number of ways of choosing 1 red ball and 1 blue ball from the 11 balls (5 red and 6 blue) is ${}_5C_1 \times {}_6C_1 = 5 \times 6 = 30$

Therefore the probability of choosing 1 red ball and 1 blue ball is $\frac{{}_5C_1 \times {}_6C_1}{{}_{11}C_2} = \frac{5 \times 6}{55} = \frac{6}{11}$

ACTIVITY 7

In a batch of 10 tape decks 2 are defective. The shop assistant selects 3 of these 10 tape decks at random. What are the probabilities that

ANSWERS ON PAGE 62

a) only good tape decks will be chosen

b) 1 of the defective tape decks will be chosen?

Summary

You have learned that permutation is an ordering or pacing of a number of objects where the order in which the objects are placed is important and that the combination is a selection of a number of objects, where order is not important. You have also used factorial notations to solve permutation and combination problems both when order matters and when order does not matter.

Finally you learnt about the multiplication of choices which states that if a choice if a choice consists of K steps, the first can be made in n_1 ways, the second in n_2 ways, ... the Kth in n_K ways, the whole choice can be made in $n_1 \times n_2 \times n_3 \times ... \times n_K$ ways.

CHECKLIST

Are you able to:

- state the difference between permutations and combinations
- use factorial\solve notation
- use factorial notation to solve permutation and combination problems
- use permutations and combinations to solve probability problems.

SELF-CHECK EXERCISE

1. State the main difference between permutation and combination.

2. Evaluate the following:
   a) 7!
   b) $\frac{6!}{4!}$
   c) $\frac{8!}{2!6!}$

3. How many arrangements of the letters of the word BEGIN are there which start with a vowel?

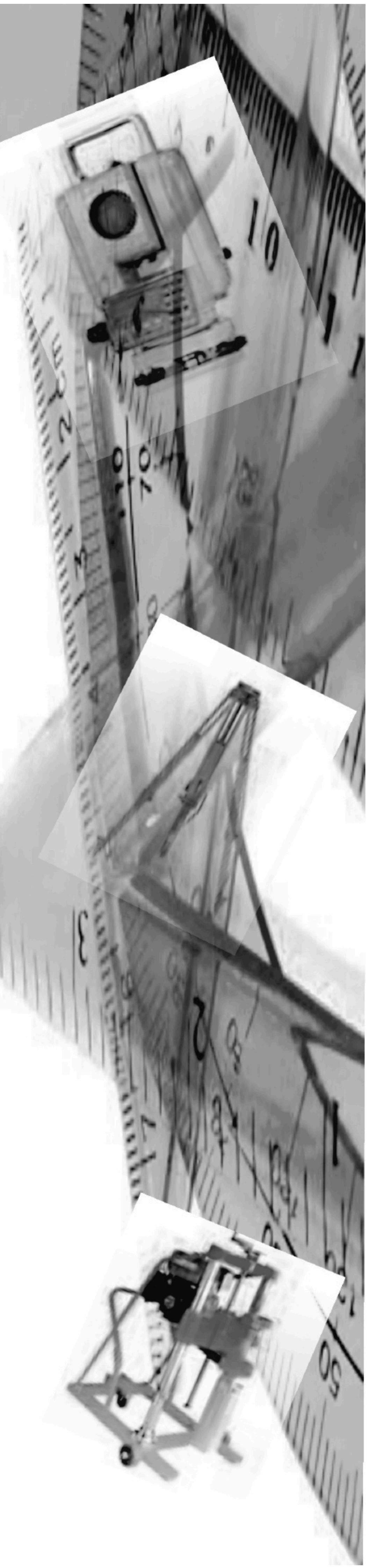
4. A coach chooses a team of two people consisting of a man and a woman to represent a club at a tennis match. If he chooses the team from 5 men and 4 women, in how many ways can he select the team?

5. The community wants to choose a committee of 3 people from 4 married couples. In how many ways can the community choose this committee if:
   i) all are equally eligible
   ii) the committee must consist of one woman and two men.

6. The community wants to choose a committee of 3 people from a group of 6 men and 4 women. Find the probabilities that the committee will consist of
   a) only men
   b) only women
   c) two men and one woman.

ANSWERS ON PAGE 68

Feedback to Activities



Lesson 1

Activity 1

1. a) $\sum x = 11 + 5 + 7 + 2 + 9 + 13 + 4 + 8 = 59$

   b) $\bar{x} = \frac{\sum x}{n} = \frac{59}{8} = 7{,}375$

2. a) $\sum x - 2$ means subtract 2 from the total $\sum x$ whilst $\sum (x - 2)$ means subtract two from each of the values of x before summing them up.

   If we use the values for question 1 then,
   $\sum x - 2 = (11 + 5 + 7 + 2 + 9 + 13 + 4 + 8) - 2 = 57$ and

   $$\sum (x - 2) = (11-2) + (5-2) + (7-2) + (2-2) + (9-2) + (13-2) + (4-2) + (8-2)$$
   $$= 43$$

   b) $\sum x - \bar{x}$ means subtract the mean from the total and $\sum (x - \bar{x})$ means find the total after subtracting the mean from each of the x –values. Try to work these out using the values from question 1.

3. $\sum (x - \bar{x})^2$ means subtract the mean from each of the x –values, square each of the results and then sum them up.

Activity 2

1. a) Mean wage = $\bar{x} = \frac{\sum x}{n}$

   $$= \frac{R(369 + 267 + 315 + 312 + 585 + 315 + 279 + 510)}{8}$$

   $= R2952 \div 8$
   $= R369$

   This means that on the average, a person receives R369 for doing the same type of work at different places.

b) Standard deviation $\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$ so first we need to find the totals of $(x-\bar{x})^2$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
369	0	0
267	– 102	10 404
315	– 54	2 916
312	– 57	3 249
585	216	46 656
315	– 54	2 916
279	– 90	8 100
510	141	19 881
TOTAL(Σ)		94 122

The standard deviation, $\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$

$$= \sqrt{\frac{94122}{8}} = \sqrt{11765,25} = 108,47 \quad \text{(to 2 decimals)}$$

This is the 'average' difference between individual wages and the mean wage. The large standard deviation in this example shows that the wages differ significantly at different places for similar work. A very small standard deviation in this context would mean that all the places paid an amount close to the mean for similar work.

2. The mean, $\bar{x} = \frac{1+3+3+5+8}{5} = \frac{20}{5} = 4$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
1	– 3	9
3	– 1	1
3	– 1	1
5	1	1
8	4	16
TOTAL (Σ)		28

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{28}{5}} = \sqrt{5,6} = 2,37 \quad \text{(to 2 decimals)}$$

Activity 3

1. Press [MODE][2][1] to clear any data already put into your calculator.

   Now press the following keys:

   [6][6] [=] [6] [4] [=] [5] [8] [=] [7] [0] [=] [4] [5] [=]

   To get the menu from which you can read the answers to the question press [AC] [SHIFT] [1] and then [4].Pressing [2] [=] gives the mean as 60,6 and then [SHIFT] [1] [4] [3] [=] gives the standard deviation as 8,708616423. This is not exact. It is limited by the calculator display. Answers will generally need to be given correct to one or two decimal places.

2. Clear the data already put into your calculator and get into the 'stat' mode again. Key in your new values:
   [7] [=] [1] [0] [=] [1] [5] [=] [2] [=] [4] [=] [6] [=]

   [AC] [SHIFT] [1] [4] will get you to the menu from which you can read your answers as before:

   $\bar{x} = 7{,}\dot{3}$ and $\sigma \approx 4{,}229525847 = 4{,}2$ (to 1 decimal)

   Errors in entering data can be corrected at once by simply re-entering the correct number. If you have already pressed the [AC] key when you discover your error, you can go back to the table by pressing [SHIFT][1] and then [2].

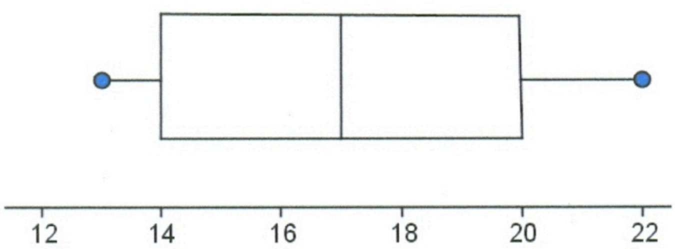
Activity 4

Press [MODE] [2] [1]to display a blank table with a frequency column. If no frequency column is displayed, press [SHIFT] [MODE], scroll to the next screen and choose [3] and then [1] to insert a frequency column. Then press [5] [0] [0] [0][=] [3] [0] [0] [0] [=] [1] [0] [2] [5] [=] [7] [6] [0] [=] and then move to the frequency column and press [1] [=] [2] [3] [4]. The press [SHIFT] [1] [4] and choose [2] from the menu for the mean and after again pressing [SHIFT] [1] [4] we choose [3] to get . The mean salary is R1 949,38 (to the nearest cent) and the standard deviation is R1 453,73 (to the nearest cent).

Activity 5

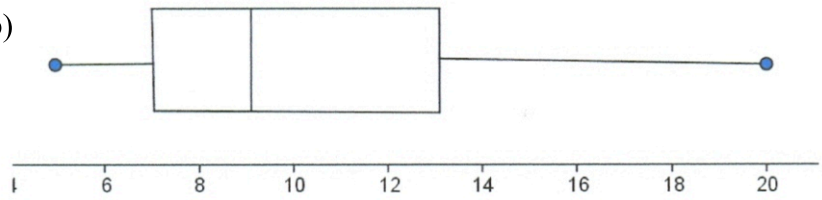
1. a) 13; 14; 17; 20; 22

   b)



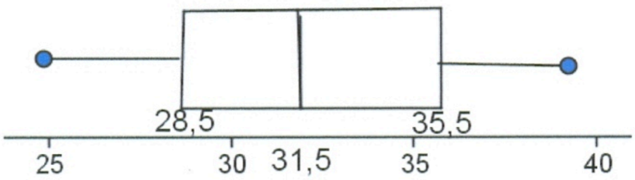
2. a) 5; 7; 9; 13; 20

b)



3. a) 25; 28,5; 31,5; 35,5; 39

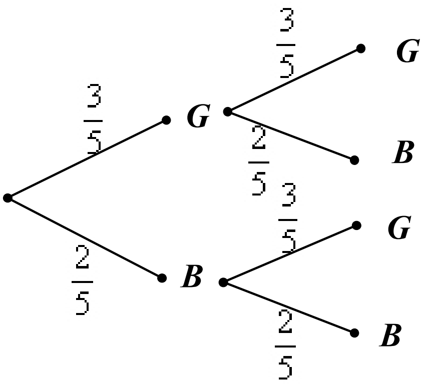
b)



Lesson 2

Activity 1

1.



a) P (both are girls) = P (GG) $\frac{3}{5}\times\frac{3}{5}=\frac{9}{25}$

b) P (both are boys) = P (BB) = $\frac{2}{5}\times\frac{2}{5}=\frac{4}{25}$

c) P (one is a boy and the other a girl =P (GB or BG)

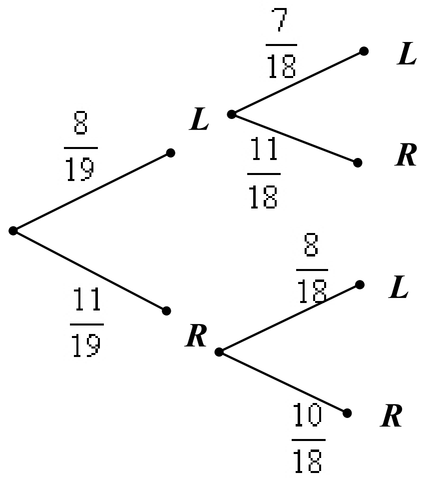
$$= P(GB) + P(BG)$$

$$=\frac{3}{5}\times\frac{2}{5}+\frac{2}{5}\times\frac{3}{5}$$

$$=\frac{6}{25}+\frac{6}{25}$$

$$=\frac{12}{25}$$

2.



a) The probability that both are left shoes is the first branch.

P (LL) = $\frac{8}{19} \times \frac{7}{18} = \frac{56}{342}$

b) The event 'picking a left shoe and a right shoe' can happen in one of these two ways:

i) either he picks a left shoe and then a right shoe or

ii) he picks a right shoe first and then a left shoe.

From the tree diagram, we have the second or the third branch.

Therefore P (a left shoe and a right shoes)
= P (LR or RL) = P (LR) + P (RL)

$= \frac{8}{19} \times \frac{11}{18} + \frac{11}{19} \times \frac{8}{18}$

$= \frac{88}{342} + \frac{88}{342}$

$= \frac{176}{342}$

Activity 2

a) The two events are independent because getting a 6 on the first die does not influence the chances of getting a 6 on the second dice.

b) The two events are not independent because if A is true, then B is less likely to be true. This is so because there is one fewer Ace in the pack.

Note that if the first card were replaced before the second card was taken, then the two events would be independent.

c) The two events are independent. Getting tails on the coin does not affect getting a 5 on the die in any way.

d) The two events are not independent because if A is true, then B is more likely to be true. This is so because the total number of discs will be one fewer but the number of red discs will remain the same.
If the first disc was replaced then the two events would be independent.

Activity 3

1. a) For each match the probability of the team drawing $\frac{1}{6}$.
Therefore the probability of drawing both matches is
$\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

b) The probability of winning the match is $\frac{1}{2}$ and the probability of losing a match is $\frac{1}{3}$. Since winning the first match is unconnected to losing the second match, the required probability is $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$.

2. For each firing there is a probability of 0,9 that the marksman will get a bull. So the probability of getting two bulls in 2 firings is $0{,}9 \times 0{,}9 = 0{,}81$.

Lesson 3

Activity 1

1. S = $\{H1; H2; H3; H4; H5; H6; T1; T2; T3; T4; T5; T6\}$, $nS = 12$

2. a) $S = \{A, B, C, D, E, F, G, H\}$

b) A and B as well as B and C are mutually exclusive

c) There are 2 vowels (A and E), so $P(B) = \frac{2}{8} = \frac{1}{4}$

d) $(B \cup C) = \{A, E, C\}$. So $n(B \cup C) = 3$

e) $P(B \cap C) = \frac{0}{8} = 0$

Activity 2

1. a) A and B are not mutually exclusive since $P(A \cap B) \neq 0$

b) $$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0{,}2 + 0{,}16 - 0{,}04 \\ &= 0{,}32 \end{aligned}$$

c) $P(A' \cap B) = P(B) - P(A \cap B)$
$= 0{,}16 - 0{,}04$
$= 0{,}12$

2. a) $P(E \text{ or } A) = P(E) + P(A) - P(E \text{ and A})$

[This can also be written] $P(E \cup A) = P(E) + P(A) - P(E \cap \text{A})$

So $\frac{25}{25} = \frac{21}{25} + \frac{8}{25} - P(E \cap A)$

$\therefore P(A \cap B) = \frac{4}{25}$

b) $P(E \cap A') = P(E) - P(E \cap A) = \frac{21}{25} - \frac{4}{25} = \frac{17}{25}$

c) $P(E' \cap A') = 0$

This is an impossible event as all students are writing English or Afrikaans.

Activity 3

a) A and C: It is impossible to draw both a red card and a spade in a single draw, but A and C do not cover all possibilities: one could also draw a club.

b) A and B: these are both mutually exclusive and complementary since all possibilities are covered by the two events. Also you need to understand that if A and B are complementary, P(A) + P(B) = 1 and $P(A \cap B) = 0$.

c) i) $(A \cap B)$ or $(A \cap C)$

ii) $(A \cup B)$

d) i) The event of drawing a spade or a picture card $(C \cup D)$ can happen in 16 + 9 ways: there are 16 picture cards (4 in each suit) plus the 2, 3, 4, …10 of spades.

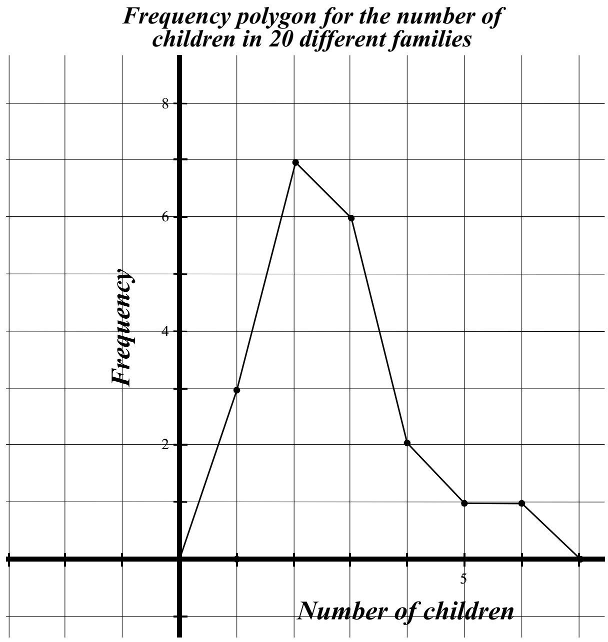
So $P(C \cup D) = \frac{25}{52}$

ii) There are 4 picture cards which are also spades, so

$P(C \cap D) = \frac{4}{52} = \frac{1}{13}$

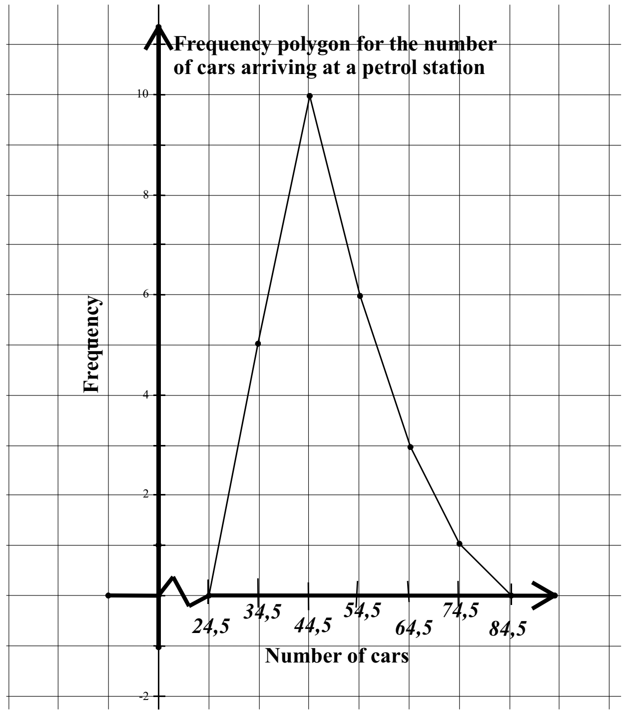
Lesson 4

Activity 1



Activity 2

Number of cars per day	Number of days/ Frequency	Mid-point
30-39	5	34,5
40-49	10	44,5
50-59	6	54,5
60-69	3	64,5
70-79	1	74,5
TOTAL	25	



The jagged line from the origin to the point where the polygon is fixed to the axis indicates that the scale on the horizontal axis is regular only from the midpoint of the interval before the first data is encountered.

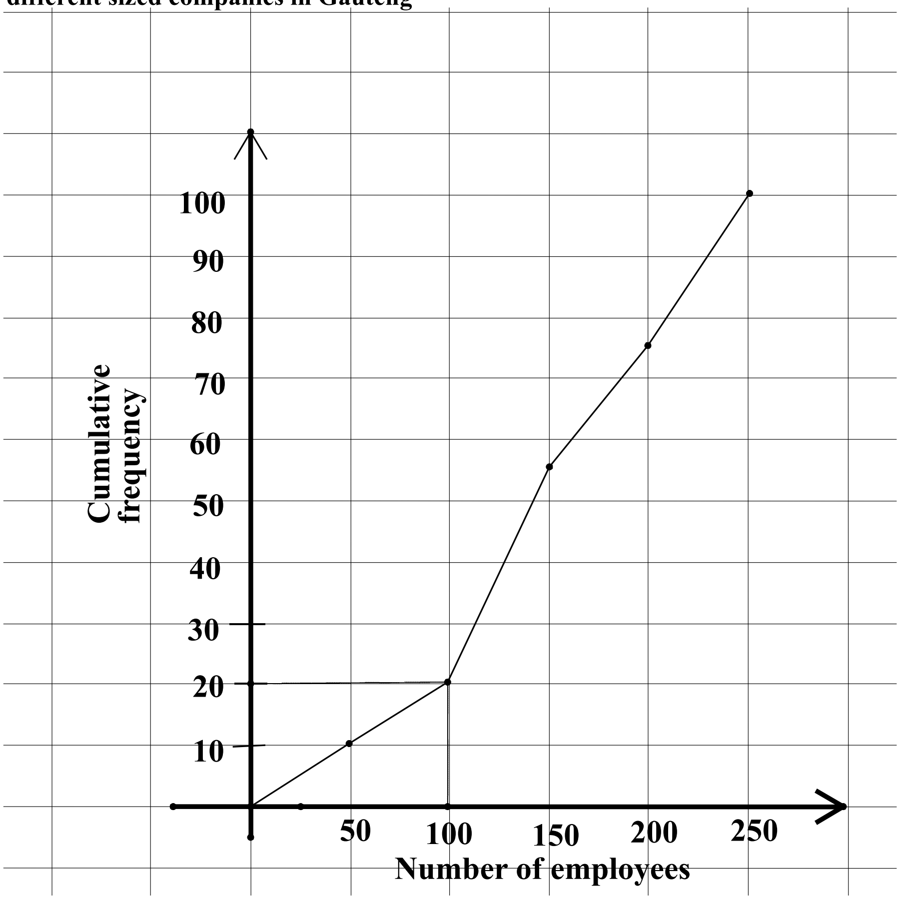
Activity 3

Number of years	Frequency	Cumulative frequency
1	4	4
2	2	4+2 =6
3	5	6+5 =11
4	4	11+4 =15
5	2	15+2 =17
6	7	17+7 = 24
7	5	24+5 =29
8	6	29+6 =35
9	3	35+3 =38
10	2	38+2 =40
Total	**40**	

Activity 4

Number of employees	Frequency	Cumulative frequency
1 - 50	5	5
51 - 100	15	20
101 - 150	35	55
151 - 200	20	75
201 - 250	25	100
TOTAL	100	

Cumulative frequency of number of workers employed by different sized companies in Gauteng



Lesson 5

Activity 1

1. Here we are only interested in choosing six people so the order in which the six people are chosen is not important. Therefore it is a combination.

2. Here we are not just interested in any three tickets winning a prize. Instead we are interested in the first ticket drawn winning the first prize, the second ticket drawn winning the second prize and the third ticket drawn winning the third prize. The order in which the tickets are drawn is therefore important so it involves permutation.

3. Here the order is important because a change in any one of the digits will result in a new number. Hence we have permutation.

4. This is just a matter of choosing. Order is not important. Therefore it is combination.

Activity 2

1. The first prize can be taken from any of the hundred tickets. That is there are 100 ways of choosing the first prize. There are 99 ways of choosing the second prize and there are 98 ways of choosing the third.

   Therefore the prizes can be won in $100 \times 99 \times 98 = 970\,200$ ways.

2. There are ten different digits (0 - 9) and a telephone number cannot start with zero. So the first number for the telephone can be any of the 9 digits 1 - 9. The second can be any of the 10 including zero. The third can be any of the 10 and so are the fourth, fifth, sixth and the seventh. Therefore there are:

   $9 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 = 9 \times 10^6$ = 9 million telephone lines available.

3. The first position can be taken by any of the 5 men. The second position can be taken by any of the remaining 4 men. The third position can be taken by any of the remaining 3. The fourth position can be taken by any of the remaining 2 and the last position can be taken by the last man. Therefore there are:
   $5 \times 4 \times 3 \times 2 \times 1 = 120$ different ways the men can be lined up.

Activity 3

1. Total number of permutations of 6 members from 8 people (the number of ways of selecting 6 members from 8 people) is:

   $8 \times 7 \times 6 \times 5 \times 4 \times 3$

   The number of permutations of the 6 members amongst themselves (number of ways of arranging 6 members within a team) is:

   $6 \times 5 \times 4 \times 3 \times 2 \times 1$

   Therefore the number of different teams of 6 members from 8 is

   $$\frac{8\times 7\times 6\times 5\times 4\times 3}{6\times 5\times 4\times 3\times 2\times 1} = 28$$

2. Number of permutations of 4 from 8 is $8 \times 7 \times 6 \times 5$.
   Number of permutations of 4 amongst themselves is $4 \times 3 \times 2 \times 1$
   Therefore the number of ways that the team of 4 can be selected from the eight players is:

   $$\frac{8\times 7\times 6\times 5}{4\times 3\times 2\times 1} = 70$$

3. As one particular boy (the prefect) must be chosen this leaves four more boys to be chosen from the remaining nineteen boys.
   Number of ways of arranging 4 boys from 19 is:

   $19 \times 18 \times 17 \times 16$

   Number of ways of arranging 4 boys among themselves is:

   $4 \times 3 \times 2 \times 1$

   Therefore the number of ways of choosing the four boys to join the prefect is:

   $$\frac{19\times 18\times 17\times 16}{4\times 3\times 2\times 1} = 3876$$

Activity 4

1. a) $$5\times 4\times 3 = 5\times 4\times 3\times \frac{2\times 1}{2\times 1} = \frac{5!}{2!}$$

   b) $$\frac{10\times 9\times 8}{3\times 2\times 1} = \frac{10\times 9\times 8}{3\times 2\times 1}\times \frac{7\times 6\times 5\times 4\times 3\times 2\times 1}{7\times 6\times 5\times 4\times 3\times 2\times 1}$$

   $$= \frac{10!}{3!7!}$$

2. a) $\frac{8!}{(4!)^2} = \frac{8\times 7\times 6\times 5\times 4!}{4!\times 4!} = \frac{8\times 7\times 6\times 5}{4\times 3\times 2\times 1} = 70$

b) Remember 0! = 1. Therefore 13!/0! is defined (has an answer) and is the same as:
$13! = 6\,227\,020\,800$

Activity 5

1. This is a matter of selecting 3 from 15. Therefore the number of ways is $_{15}C_3 = \frac{15!}{3!(15-3)!} = \frac{15!}{3!12!} = \frac{15\times 14\times 13}{3\times 2\times 1} = 455$

2. Since the positions (1st, 2nd and 3rd) are important, we have a situation where order is important. Therefore we have a permutation and the answer is:

$$_{10}P_3 = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 10\times 9\times 8 = 720$$

Activity 6

1. Number of ways the antibiotic can be chosen is 5.
Number of ways the painkiller can be chosen is 6.
Therefore by the multiplication of choices method the number of ways the antibiotic and the painkiller can be chosen is $5\times 6 = 30$.

2. 2 men can be chosen from the 7 men in $_7C_2 = \frac{7\times 6}{2\times 1} = 21$ ways.

Similarly, 2 women can be chosen from the 5 women in

$_5C_2 = \frac{5\times 4}{2\times 1} = 10$ ways.

Therefore by the multiplication of choices, the number of different ways of choosing the team is $_7C_2 \times {_5C_2} = 21\times 10 = 210$ ways.

Activity 7

Total number of ways of choosing 3 tape decks from 10 is:

$_{10}C_3 = \frac{10\times 9\times 8}{3\times 2\times 1} = 120$

a) Only good tapes will be chosen is the same as choosing all three tapes from the 8 good ones. That is $_8C_3 = \frac{8\times 7\times 6}{3\times 2\times 1} = 56$

Therefore the required probability is $\frac{_8C_3}{_{10}C_3} = \frac{56}{120} = \frac{7}{15}$.

b) 1 of the defective tape decks will be chosen is the same as choosing 1 of the defective tape decks from the 2 defective ones and 2 good tapes from the 8 good ones.
That is $_2C_1 \times {_8C_2} = 2\times 28 = 56$

Therefore the required probability is $\frac{_2C_1 \times {_8C_2}}{_{10}C_3} = \frac{2\times 28}{120} = \frac{7}{15}$

Feedback to Self-Check Exercises

Lesson 1

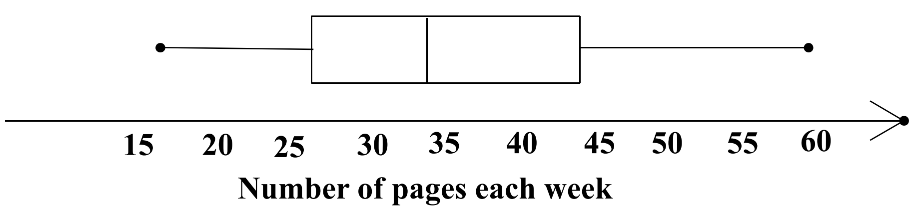
1. a) The mean sale is R340.
   This means that on average, a branch of a chain of stores in KwaZulu Natal made a sale of R340 that day.
   b) The standard deviation is R25,50 (to nearest cent).
   The small standard deviation means that the values are evenly spread around the mean and so the mean is representative of the sales in rands made by the branches that day.

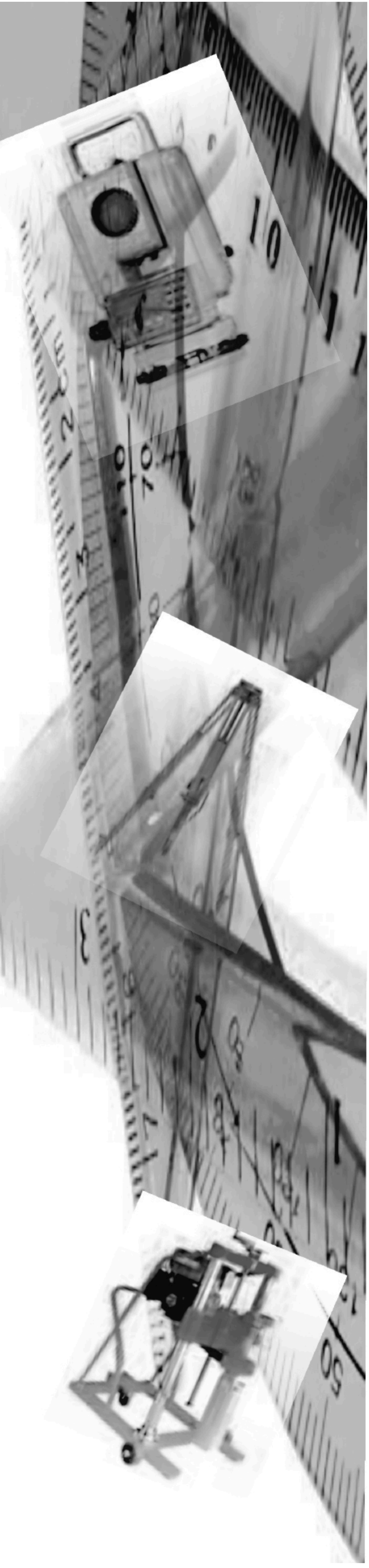
2. a) 1 366
   b) The mean mark is 50,59.
   This means that on average, a worked enrolled in a learning programme in the Eastern Cape province obtained 50,59 in their mathematics exams.
   c) The standard deviation is 6,44 (correct to 2 decimal places).
   The small standard deviation means that the values are evenly spread around the mean and so the mean is representative of the marks obtained by the workers enrolled in the learning programme.

3. a) The mean rainfall for June in the city is 52 mm.
   b) The standard deviation is 5,3 mm (correct to 1 decimal place).
   c) The result means that the average rainfall for the city in June is 52 mm and the small value of the standard deviation tells us that the mean rainfall is a good representative of the rainfall in June in the city.

4. a) The five number summary is 17; 26; 34,5; 44; 59 (The smallest, lower quartile, median, upper quartile and the highest value).

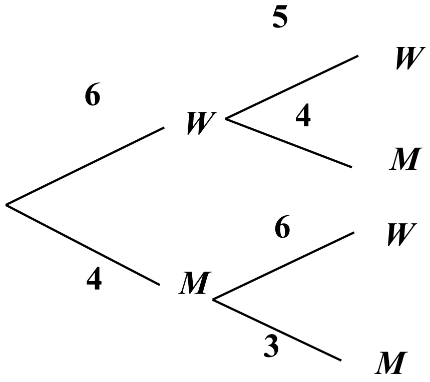
Box and whisker diagram of number





Lesson 2

1. The tree diagram should look like this:



a) The probability of getting both women is the top branch. So
$$P(WW) = \frac{6}{10} \times \frac{5}{9} = \frac{30}{90} = \frac{1}{3}$$

b) The probability of getting both men is the bottom branch. Therefore
$$P(MM) = \frac{4}{10} \times \frac{3}{9} = \frac{12}{90} = \frac{2}{15}$$

c) The second and third branches gives you one man and one woman. Therefore the required probability is:

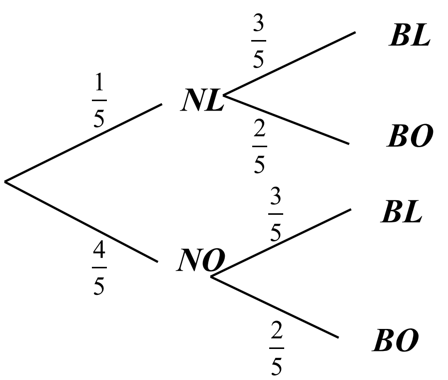
$$P(WM \text{ or } MW) = P(WM) + P(MW) = \frac{6}{10} \times \frac{4}{9} + \frac{4}{10} \times \frac{6}{9} = \frac{8}{15}$$

2. a) The probability that Nomonde gets up late is $\frac{1}{5}$

b) The probability that her bus comes on time is $\frac{2}{5}$

c) The probability that Nomonde misses the bus is
$$\frac{1}{5} \times \frac{2}{5} = \frac{2}{25}$$

d) NL - Nomonde gets up late
NO - Nomonde gets up on time
BL - The bus comes late
BO - The bus comes on time.



Nomonde catches (late) bus: $\frac{3}{25}$

Nomonde misses (on time) bus: $\frac{2}{25}$

Nomonde catches (late) bus: $\frac{12}{25}$

Nomonde catches (on time) bus: $\frac{8}{25}$

3. a) The probability that the elected

i) president is a man is $\frac{3}{4}$

ii) secretary general is older than 50 years is $\frac{3}{5}$

iii) treasurer is a woman is $\frac{2}{3}$

b) The required probabilities are

i) $\frac{3}{4} \times \frac{3}{5} = \frac{9}{20}$

ii) $\frac{3}{4} \times \frac{1}{3} = \frac{1}{4}$

iii) $\frac{2}{3} \times \frac{3}{5} = \frac{2}{5}$

Lesson 3

1. a) $S = \{1,2,3,4,5,6\}$

b) i) $P(A) = \frac{3}{6} = \frac{1}{2}$ $P(B) = \frac{3}{6} = \frac{1}{2}$

$P(C) = \frac{2}{6} = \frac{1}{3}$ $P(D) = \frac{1}{6}$

ii) A and B, A and D, C and D.

iii) $P(A \cap B) = 0$, $P(B \cap C) = \frac{1}{2}$, $P(C \cap D) = 0$

iv) $P(A \cup B) = \frac{6}{6} = 1$ and $P(C \cup D) = \frac{3}{6} = \frac{1}{2}$

2. Let C and H be the events that a person among the 40 people owns a car and a house respectively.

Then $P(C) = \frac{30}{40} = \frac{3}{4}$

Since each of the 40 people owns a car or a house or both, $P(C \cup H) = 1$

The required probability is $P(C \cap H)$

Now $P(C \cup H) = P(C) + P(H) - P(C \cap H)$

Therefore $P(C \cap H) = P(C) + P(H) - P(C \cup H)$

$$= \frac{30}{40} + \frac{25}{40} - 1 = \frac{15}{40} = \frac{3}{8}$$

3. a) Each dice can yield 6 possible outcomes. So to get all the possible outcomes we need to make a table with one set of outcomes on one side and the other set on the other side and combine them:

	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6
3	3,1	3,2	3,3	3,4	3,5	3,6
4	4,1	4,2	4,3	4,4	4,5	4,6
5	5,1	5,2	5,3	5,4	5,5	5,6
6	6,1	6,2	6,3	6,4	6,5	6,6

So S = {(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)}

b) Therefore n(S) = 36

c) i) Let A be the set such that the sum of the two numbers is 8
Then A {(2,6), (3,5), (4,4), (5,3), (6,2)}and n(A) = 5

So the required probability is $\frac{5}{36}$

ii) Let B be the set such that the two numbers are the same
Then B = {(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)}

Therefore n(B) = 6 and P(B) = $\frac{6}{36} = \frac{1}{6}$

iii) P (at least one 6) = 1 – P (no six) = 1 – ($\frac{5}{6} \times \frac{5}{6}$) = $\frac{11}{36}$

4. $n(S) = 350$

a) $P(A \cap B \cap C) = \frac{5}{350} = \frac{1}{70}$

b) $P(A \cap B \cap C') = \frac{10}{350} = \frac{1}{35}$

c) $P(A \cup B) = \frac{225}{350} = \frac{9}{14}$

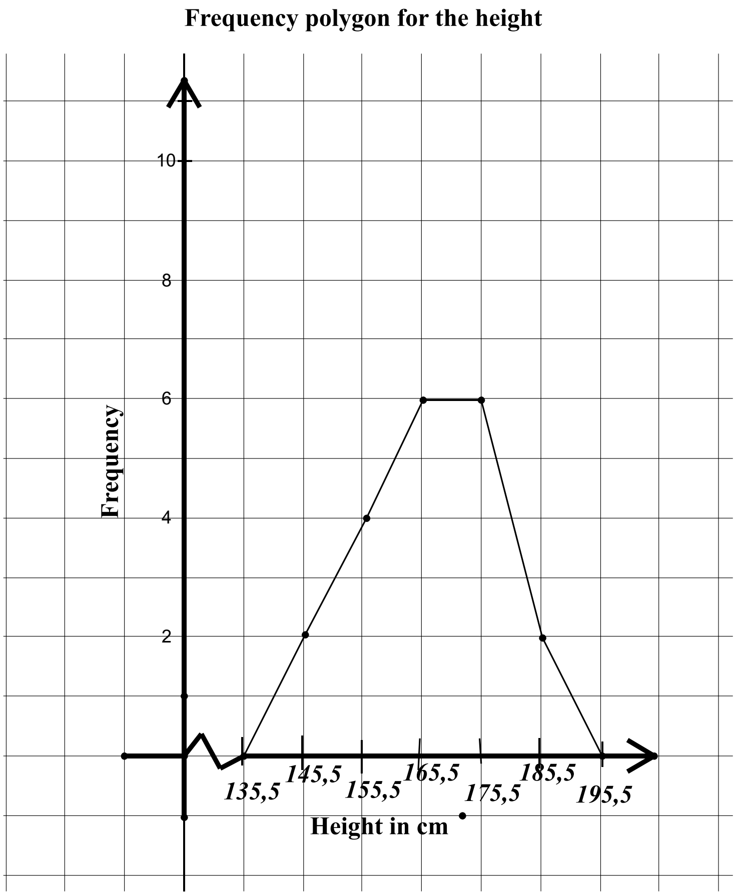
d) $P(B \cup B \cap A') = \frac{215}{350} = \frac{43}{70}$

Lesson 4

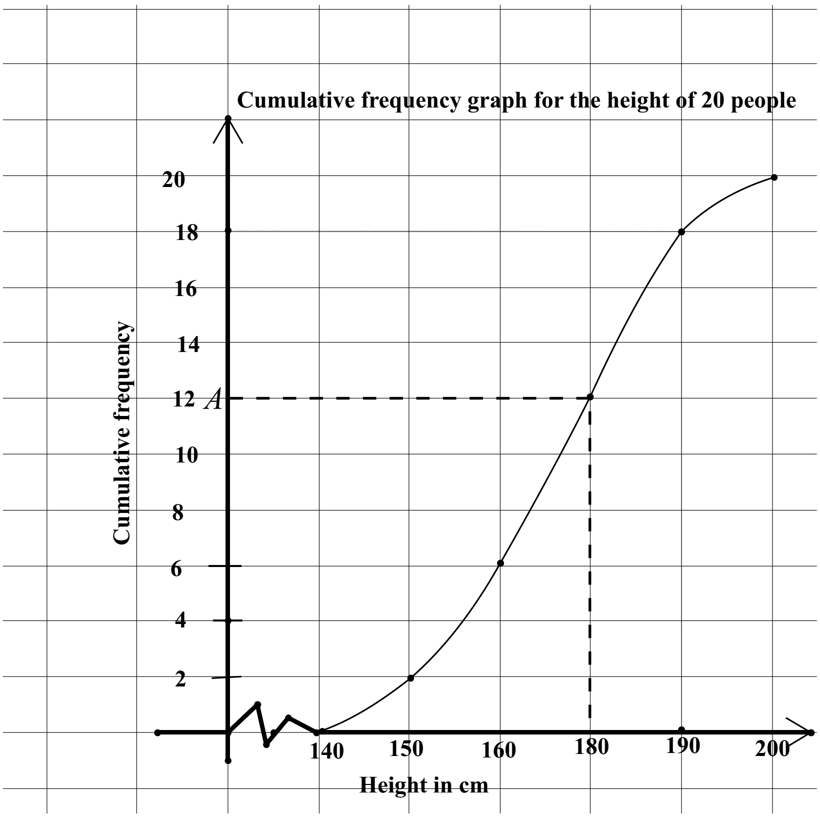
a)

Height (cm)	Frequency	Cumulative frequency
141 - 150	2	2
151 - 160	4	6
161 - 170	6	12
171 - 180	6	18
181 - 190	2	20
	20	

b)



c)



d) From the cumulative frequency graph 18 people had heights below 180 cm (Read at A).

Lesson 5

1. In permutation the order of the objects is important whilst in combination the order is not important.

2. a) 5 040 (using the calculator)
   b) 30 (cancelling through by 4!)
   c) 28 (cancelling through by 6!)

3. There are only 2 vowels in the word **BEGIN**, this means there can be only 2 ways of filling the first position. After the first letter has been selected, there will be 4 letters left to fill the remaining four positions. This can be done in 4! ways. The number of ways to arrange the letters so that it starts with a vowel will then be
   $2 \times 4! = 2 \times 4 \times 3 \quad 2 \times \ 1 = 48$ ways.

4. 20 $({}^5C_1 \times {}^4C_1)$

5. a) 56 $\left({}^8C_3\right)$
   b) 24 $\left({}^4C_1 \times {}^4C_2\right)$

6. a) $\dfrac{{}^6C_3}{{}^{10}C_3} = \dfrac{20}{120} = \dfrac{1}{6}$
   b) $\dfrac{{}^4C_3}{{}^{10}C_3} = \dfrac{4}{120} = \dfrac{1}{30}$
   c) $\dfrac{{}^6C_2 \times {}^4C_1}{{}^{10}C_3} = \dfrac{15 \times 4}{120} = \dfrac{1}{2}$