1 Representation of data

1.1 Terminologies

Note the following terms:

- qualitative data

  _________________ or _________________ are often used to show the information.

- quantitative data (numerical values)

  There are two types: _________________ and _________________.

- frequency distribution

  Case 1:

  Case 2:

- class boundry, interval width

Exercise 1

1. The table shows the areas of different categories of land use in a particular region.

Category	Urban	Woodland	Farmland	Reservoris	Total
Area(km^2)	615	660	1200	225	2700

   Illustrate the data diagramatically.

2. In a survey on the number of letters in the solutions of a crossword puzzle, the following data were obtained from the crossword puzzle in Monday's newspapaer.

9	5	7	7	5	12	6	6	12	5
7	7	5	5	3	7	3	7	9	6
3	5	7	3	6	7	7	6	3	7

   Draw a frequency distribution table for the data.

3. The following data were obtained in a survey of heights of 20 children in a sports club. Each height was measured to the nearest centimetre.

133	136	120	138	133	131	127	141	127	143
130	131	125	144	128	134	135	137	133	129

   Draw a frequency distribution table for the data.

1.2 Stem-and-leaf diagram

A way of grouping data into intervals while still retaining the original data is to draw a **stem-and-leaf diagram**.

For example, there are 20 students in an assignment, their marks are as follows:

84	17	38	45	47	53	76	54	75	32
66	65	55	54	51	44	39	19	54	72

A final plot of stem-and-leaf diagram would be like:

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

Key: $1|1$ means 17 marks

Table 1: Stem-and-leaf diagram of assignment marks

Important: You must always give a ____________ to explain what the stem and leaf represent.

Also, ________________________ must be chosen in a stem-and-leaf diagram.

For two sets of data, ____________________________________ is often used to compare the data.

Exercise 2

1. The lengths, in metres, of 20 measurements in a physics experiments are recorded as follows.

1.78	1.87	1.89	1.72	1.68	2.04	1.96	1.76	1.90	1.73
1.78	1.61	1.78	1.77	1.85	1.65	1.89	1.95	2.01	1.83

   (a) Represent this information on a stem-and-leaf diagram.

   (b) State the mode.

2. The maximum temperature in ˚C, measured to the nearest degree, was recorded each day during June in a particular city. The temperature were as follows:

19	23	19	19	20	12	19	22	22	16	18	16	19	20	17
13	14	12	15	17	16	17	19	22	22	20	19	19	20	20

   Draw a stem-and-leaf diagram to illustrate the temperatures and write down the mode.

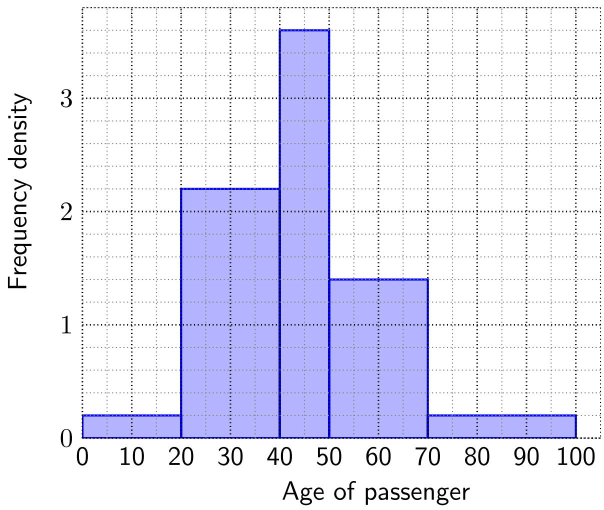
3. These are the examination marks for French and for English achieved by pupils in a particular class.

French	43	55	29	49	36	55	61	34	42	42
	54	60	48	23	44	31	55	45	37	57
English	80	65	74	59	79	92	52	71	43	86
	60	74	57	41	79	74	58	52	64	84

   Draw a back-to-back stem-and-leaf diagram to compare the two sets of marks.

1.3 Histograms

Grouped data can be displayed in a **histogram**, as in the following diagram:



The frequency distribution table would be like:

Age, x years	$0 \leqslant x < 20$	$20 \leqslant x < 40$	$40 \leqslant x < 50$	$50 \leqslant x < 70$	$70 \leqslant x < 100$
Frequency					

Notice the difference between histogram and ____________:

-

-

Frequency density:

$$\text{frequency density} = \frac{\text{frequency}}{\qquad\qquad\qquad\qquad}.$$

Modal class: the interval with the greatest ________________________.

Gaps treatment:

- Gropued continuous data (rounded values)

- Grouped discrete data (continuity correction, "0"):

The shape of a distribution:

Positive skew

Symmetrical

Negative skew

Exercise 3

1. A survey on the duration of telephone calls made to an office on a particular day gave the following results.

Duration, t minutes	$1 \leqslant t < 3$	$3 \leqslant t < 9$	$9 \leqslant t < 15$	$15 \leqslant t < 20$
Frequency	10	42	12	7

Draw a histogram to represent the data.

2. The grouped frequency table records the weights, to the nearest gram, of the letters delievered to an apartment block on a particular day.

Weight (gram)	$31-50$	$51-60$	$61-70$	$71-100$	$101-150$
Frequency	16	25	36	33	10

Draw a histogram to represent the data and state the modal class.

3. † One evening a waiter measured the amounts of water left by diners in the bottles on the table in a restaurant. The volume were measured to the nearest millimetre.

Volume (nearest ml)	$0-19$	$20-39$	$40-89$	$90-189$
Frequency	10	8	12	20

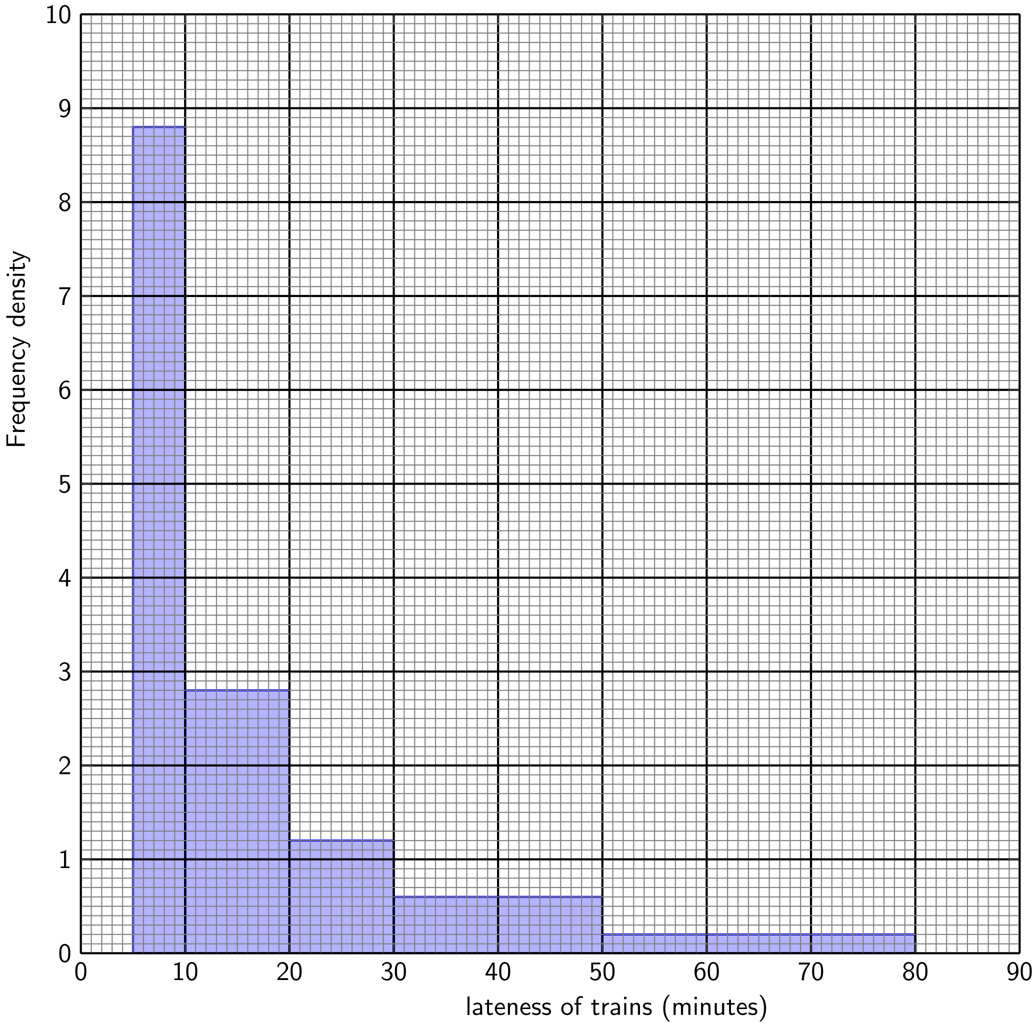
Draw a histogram to represent the data and state the modal class.

4. These are marks in a statistics test for a group of 120 A level students.

Mark	$0-9$	$10-19$	$20-29$	$30-49$	$50-79$
Frequency	8	21	53	28	10

Draw a histogram to represent the data.

5. A Passenger's Association conducted a survey on the lateness of trains arriving at a particular railway station. The results are illustrated in the histogram.



(a) Construct a frequency table

(b) What percentage of the train were less than 20 minutes late?

1.4 Average

An **average** value is useful when describing a set of data. This is a typical or representative value and is known as a ________________________________.

Mean: in general, the mean of the n numbers $x_1, x_2, \cdots, x_n$ is given by

$$\bar{x} =$$

Mean of a frequency distribution:

- Discrete data

- Continuous data (**Mid-interval value**)

Exercise 4

1. To obtain Grade A, Ben must achieve a mean mark of at least 70 in five test. His mean mark for the first four tests in 68, what is the lowest mark Ben could achieve in the fifth test to obtain Grade A.

2. The numbers of an orchestra were asked how many instruments each could play. These are their replies.

2	5	2	4	1	1	1	2	1	3	3	2	1	2	1
1	2	4	3	2	1	2	3	1	4	2	3	1	1	2

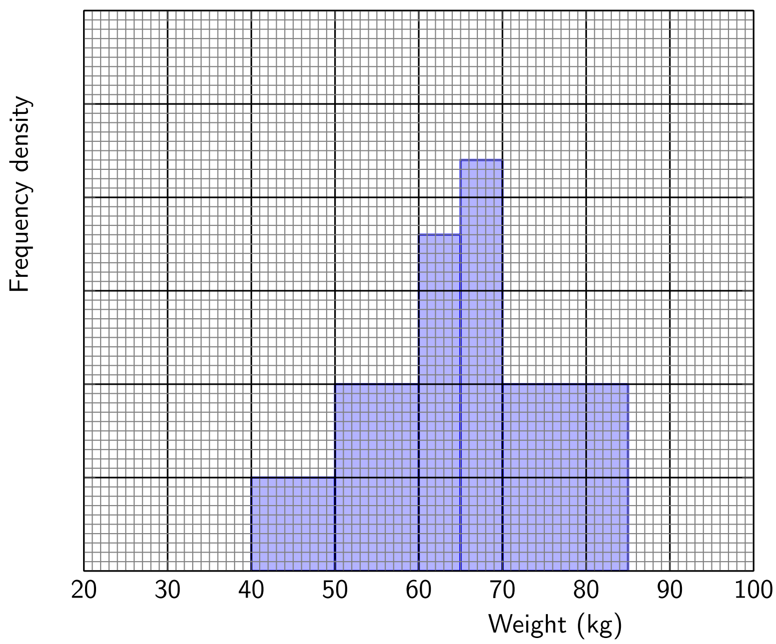
   Calculate the mean number of instruments played.

3. In a spot check, the speeds of 120 vehicles on a particular stretch of road through a village were noted. The results are shown in the table.

Speed x km h^{-1}	$21-25$	$26-30$	$31-35$	$36-45$	$46-60$
Frequency f	22	48	25	16	9

   Estimate the mean speed of these vehicles.

4. The diagram shows a histogram of distribution of the weights of 50 first-year students at a particular university. All the rectangles have been drawn, but the vertical scale is missing.



   Compile a grouped frequency table and find an estimate of the mean weight of these students.

1.5 Variability of data

Range:

Variation:

Standard deviation:

Comparing distributions: the ________________ the standard deviation, the ________________ there is and the ________________ the data are.

'Calculation' version of formula for standard deviation:

$$\text{s.d.} = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}, \quad \text{where} \quad \bar{x} = \frac{\sum x}{n}.$$

Standard deviation of a frequency distribution:

- Case 1

- Case 2

Exercise 5

1. The mean of the numbers $2, 3, 5, 6, 8$ is 4.8. Calculate the standard deviation.
2. The distribution shows the number of children in 20 families. The mean number of children in a family is 2.9. Calculate the range and the standard deviation.

Number of children, x	1	2	3	4	5
Number of families, f	3	4	8	2	3

3. An online test was taken by 115 students. The time spent on each question was recorded by the computer. The following table shows the time taken, in minutes, on the final question.

Time (mins)	$1 \leq x < 2$	$2 \leq x < 3$	$3 \leq x < 5$	$5 \leq x < 10$
Frequency	16	32	42	25

Calculate estimates of the mean and standard deviation of the time spent on the final question.

4. Becky plays a computer game where she fires at a target. Her score is 1 if she hits the target and 0 if she misses it.

She has 30 attempts and hits the target 18 times.

(a) Find her mean score for the 30 attempts.

(b) Find the variance of her scores for the 30 attempts.

Combining sets of data:
In general, for two sets of data, x and y:

$$\text{mean} = \frac{\sum x + \sum y}{n_1 + n_2}, \qquad \text{stand deviation} = \sqrt{\frac{\sum x^2 + \sum y^2}{n_1 + n_2} - (\text{mean})^2}.$$

Coding data
In general, if each data value is increased by a constant a,

- the mean is _________________ . That is

$$\bar{x} = \frac{\sum (x-a)}{n} + a.$$

- the standard deviation is ___________. That is

$$\text{s.d. of } x = \sqrt{\frac{\sum (x-a)^2}{n} - \left(\frac{\sum (x-a)}{n}\right)^2}.$$

Exercise 6

1. The ages, x years, of 18 people attending an evening class are summarized by the following totals:

$$\sum x = 745, \qquad \sum x^2 = 33951.$$

   (a) Calculate the mean and the standard deviation of the ages of this group of people.
   (b) One person leaves the group and the mean age of the remaining 17 people is exactly 41 years. Find the age of the person who left and the standard deviation of the ages of the remaining 17 people.

2. The following table shows the mean and standard deviation of the heights of 20 boys and 30 girls.

	Mean	Standard deviation
Boys	160 cm	4 cm
Girls	155 cm	3.5 cm

   Find the mean and standard deviation of the heights of the 50 children.

3. Sweets are packed into bags with a nominal weight of 75 grams. Ten bags are picked at random from the production line and weighted. Their weights, in gram, are

$$76.0 \quad 74.2 \quad 75.1 \quad 73.7 \quad 72.0 \quad 74.3 \quad 75.4 \quad 74.0 \quad 73.1 \quad 72.8$$

   (a) Use your calculator to find the mean and the standard deviation .
   (b) It is later discovered that the scales were reading 3.2 grams below the correct weight.
      i. What was the correct mean weight of the 10 bags?
      ii. What was the correct standard deviation of the 10 bags?

4. The time taken, x minutes, by Katy to do the Sudoku puzzle in a certain newspaper was observed on 20 occasions. The results are summarised below.

$$\sum (x-30) = -50 \qquad \sum (x-30)^2 = 562$$

   Find the mean and standard deviation of the time taken by Katy to solve the Sudoku puzzle.

5. A summary of 24 observations of x gave the following informations:

$$\sum(x-a) = -73.2 \qquad\qquad \sum(x-a)^2 = 2115$$

The mean of these values of x is 8.95.

(a) Find the value of the constant a.

(b) Find the standard deviation of these value of x.

6. Delip measured the speeds, x km per hour, of 70 cars on a road where the speed limit is 60 km per hour. His results are summarised by $\sum(x-60) = 245$.

(a) Calculate the mean speed of these cars.

His friend Sachim used values of $(x-50)$ to calculate the mean

(b) Find $\sum(x-50)$.

(c) The standard deviation of the speeds is 10.6 km per hour. Calculate $\sum(x-50)^2$.

7. It is known that, for 100 observations of x, the mean $\bar{x} = 25$.

Find

(a) $\sum x$, $\sum(x-20)$, $\sum(x-27)$.

Given also that the standard deviation of x is 3, find

(b) $\sum x^2$, $\sum(x-20)^2$, $\sum(x-27)^2$.

1.6 Cumulative frequency graph

When a set of data contains extreme values, known as ___________, the median is a more informative average and the interquartile range is more-useful measure of spread.

Median: for a set of n numbers arranged in _____________________.

- When n is odd,
- When n is even,

Quartiles:

- Lower quartile, Q_1,
- Upper quartile, Q_3,
- Interquartile range (IQR),

Cumulative frequency: a total frequency up to a particular item. It is particular useful when finding the median and quartiles.

- Ungrouped data
- grouped data, class boundaries.

Exercise 7

1. Find the median of each of these sets of numbers:
   (a) $7, 7, 2, 3, 4, 2, 7, 9, 31$
   (b) $36, 41, 27, 32, 29, 39, 39, 43$
2. To inverstigate hand-eye coordination in reacting to a stimulus, students took part in an experiment where a ruler was dropped and the distance it travelled before the student caught in was measured. The results of 21 girls and 27 boys are shown in the back-to-back stem-and-leaf diagram.

Distance, in cm

	Girls	Stem	Boys	
		4	8 9	(2)
		5	0 5 9	(3)
(1)	8	6	2 4	(2)
(5)	5 5 3 2 2	7	1 4 5	(3)
(2)	7 6	8	0 2 5 7 8 9	(6)
(6)	9 9 8 4 1 1	9	3 3 6	(3)
(3)	9 6 3	10	0 0 5 7 8	(5)
(3)	7 5 3	11	2 3	(2)
(1)	4	12	7	(1)

Key: $8|6|2$ represents a distance of 6.8 cm for the girls and 6.2 cm for the boys

Find the median and interquartile range for both sets of data. Comment on your answers.

3. In a survey on the number of absences in the term of the 32 children in a class, the data were recorded in a cumulative frequency table.

Time absent	0	$\leqslant 1$	$\leqslant 2$	$\leqslant 3$	$\leqslant 4$	$\leqslant 5$	$\leqslant 6$	$\leqslant 7$
Cumulative frequency	5	11	20	23	27	28	31	32

(a) Find the median number of absence.

(b) Find the interquartile range.

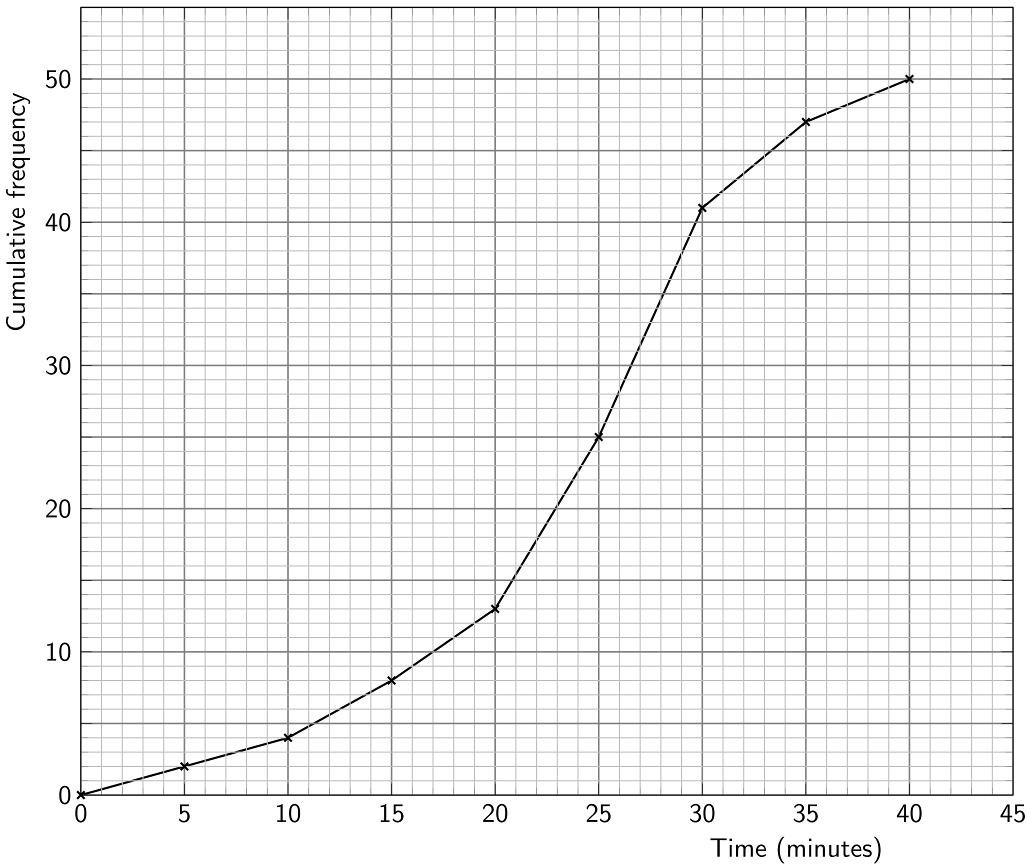
(c) Copy and complete this frequency table.

Time absent	0	1	2	3	4	5	6	7
Frequency								

(d) Calculate the mean number of absences per child.

(e) Calculate the standard deviation.

4. The cumulative frequency curve has been drawn from information about the amount of time spent by 50 people in a supermarket on a particular day.



Use the graph to estimate

(a) how many people spent at least 17 minutes but less than 27 minutes in the supermarket,

(b) the value of t, where 60% of the people spent less than t minutes in the supermarket,

(c) the value of s, where 60% of the people spent at least s minutes in the supermarket,

(d) the median time,

(e) the interquartile range.

5. A factory produces certain components. In a quality control test, 500 components were weighted and their weights recorded to the nearest gram. The table shows the results.

Weight (g)	$60-69$	$70-74$	$75-79$	$80-84$	$85-89$
Frequency	30	90	130	210	40

   (a) Construct a cumulative frequency table and draw a cumulative frequency graph.

   (b) Components that weight less than 64.5 grams or more than 87.5 grams were rejected. Use your graph to estimate the percentage of components that were accepted.

6. The cumulative frequency table shows the times taken by students to travel to college on a particular day.

Time (minutes)	<10	<15	<20	<30	<45
Cumulative frequency	35	79	157	350	400

   Construct a frequency table and use it to estimate the mean time taken to travel to college on that day.

7. The cumulative frequency table gives the heights of 400 children in a certain school.

Height, x cm	<100	<110	<120	<130	<140	<150	<160	<170
Cumulative frequency	0	27	85	215	320	370	395	400

   (a) Draw a cumulative frequency curve.

   (b) Use the curve to estimate the median height

   (c) Determine the interquatile range.

8. The arrival times of 204 trains were noted and the number of minutes, t, that each train was late was recorded.

Number of minutes late (t)	$-2 \leqslant t < 0$	$0 \leqslant t < 2$	$2 \leqslant t < 4$	$4 \leqslant t < 6$	$6 \leqslant t < 10$
Number of trains	43	51	69	22	19

   (a) Explain what $-2 \leqslant t < 0$ means about the arrival times of the trains.

   (b) Draw a cumulative frequency graph, and from it estimate the median and interquartile range of the number of minutes late of these trains.

9. The times, to the nearest minute, taken by 120 students to write a timed essay were recorded. The results are shown in the table.

Time (minutes) (t)	$40-44$	$45-49$	$50-54$	$55-59$	$60-64$
Frequency	8	24	32	30	26

   (a) Construct the cumulative frequency table and draw a cumulative frequency graph.

   (b) Use your graph to estimate the lower quartile and the median.

   Another group of 40 students wrote the same essay and all of them took at least 1 hour to complete it.

   (a) Use your graph to estimate the lower quartile of all 160 students.

   (b) Explain why it is not possible to estimate the interquartile range of the times spent by all 160 students.

1.7 Box-and-whisker plots

In a box-and-whisker plot the median and quartiles are shown, as well as the ___________ and ___________ values of a distribution.

For example, a survey on the heights of all the girls in a particular year group in a school gave the following information.

Minimum height	144 cm
Lower quartile	159 cm
Median	165 cm
Upper quartile	169 cm
Maximum height	181 cm

A box-and-whisker plot can be drawn as following:

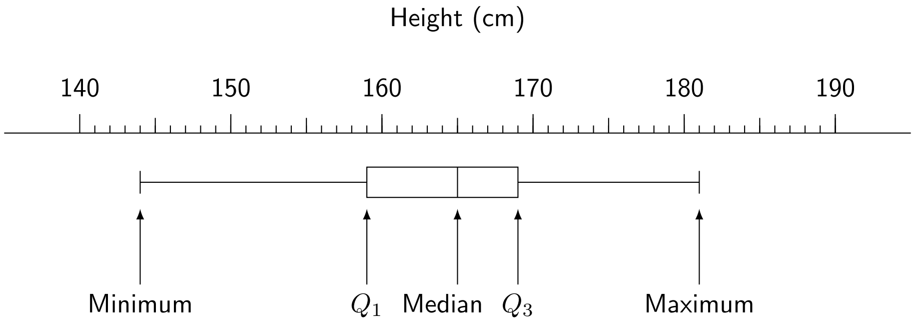


Figure 1: Box-and-whisker plot to show heights of 15-year-old girl

Tips on drawing the box-and-whisker plot:

- The scale must be ___________ and ___________.
- The whiskers must not be drawn through the box.

The shape of a distribution:

Positive skew   Symmetrical   Negative skew

Exercise 8

1. Two groups of people played a computer game which tested how quickly they reacted to a visual instruction to press a particular key. The computer measured reaction times in seconds, to the nearest tenth of a second. The following summary statistics were displayed for each group.

	Minimum	Lower quartile, Q1	Median, Q2	Upper quartile, Q3	Maximum
Group 1	0.6	0.8	1.0	1.5	1.9
Group 2	0.4	0.7	1.0	1.3	1.6

   Draw two box-and-whisker plots and compare the reaction time of the two groups.

2. The following back-to-back stem-and-leaf diagram show the cholesterol count for a group of 45 people who exercise daily and for another group of 63 who do not exercise. The figures in brackets show the number of people corresponding to each set of leaves.

	People who exercise		People who do not exercise	
(9)	9 8 7 6 4 3 2 2 1	3	1 5 7 7	(4)
(12)	9 8 8 8 7 6 6 5 3 3 2 2	4	2 3 4 4 5 8	(6)
(9)	8 7 7 7 6 5 3 3 1	5	1 2 2 2 3 4 4 5 6 7 8 8 9	(13)
(7)	6 6 6 6 4 3 2	6	1 2 3 3 3 4 5 5 5 7 7 8 9 9	(14)
(3)	8 4 1	7	2 4 5 5 6 6 7 8 8	(9)
(4)	9 5 5 2	8	1 3 3 4 6 7 9 9 9	(9)
(1)	4	9	1 4 5 5 8	(5)
(0)		10	3 3 6	(3)

Key: $2|8|1$ represents a cholesterol count of 8.2 in the group who exercise and 8.1 in the group who do not exercise

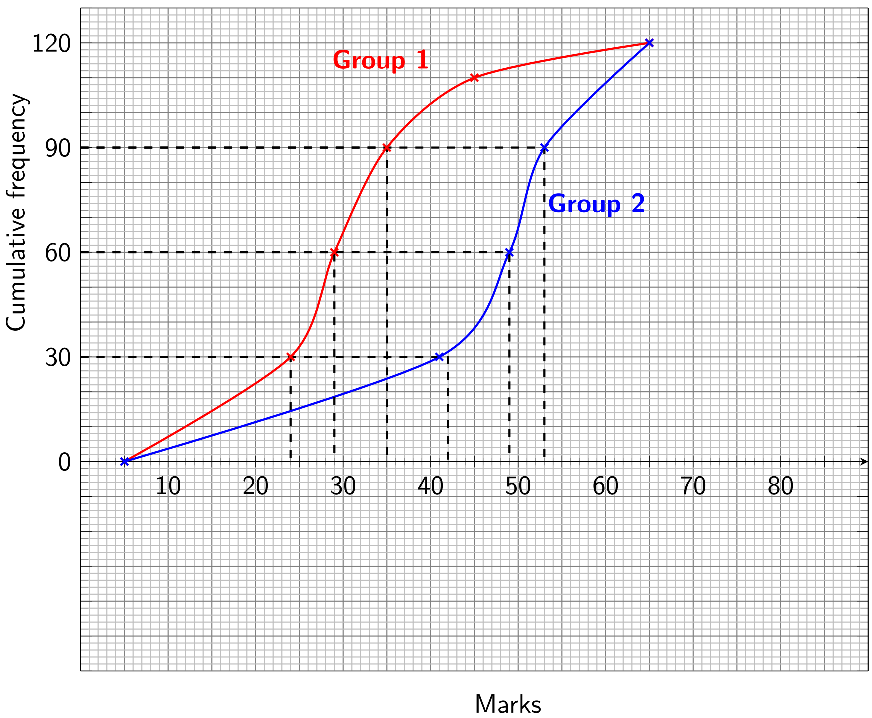
(a) Give one useful feature of a stem-and-leaf diagram.

(b) Find the median and quartiles of the cholesterol count for the group who do not exercise.

You are given that the lower quartile, median and upper quartile of the cholesterol count for the group who exercise are 4.25, 5.3 and 6.6 respectively.

(c) On a single diagram on graph, draw two box-and-whisker plots to illustrate the data.

3. The cumulative frequency graph below shows the marks of two groups of 120 people in a test.



Draw two box-and-whisker plots for each group on the same diagram, and compare the two sets of data, state why it is misleading by the cumulative frequency curves.

1.8 Choosing measures and diagrams

When representing data you need to consider the ways that would be most appropriate,

- What type of diagram should you draw?
- Which average best represent the data?
- Which measure of spread is most informative?

Measure of central tendency:

	Advantages	**Disadvantages**
Mode	It is useful when the post popular category is needed.	Very small data or more than two modes. There may not be a mode. Not representative. The modal class depends on the grouping of data.
Median	It is not affected by extreme values.	Not use the information of the whole data set
Mean	Use all the data and so represent every item	It is affected by one or two extreme values

Measure of spread:

	Advantages	**Disadvantages**
Range	Easy to calculate. Represent the complete spread of data.	It can be affected by extreme values.
Interquartile range	It is not affected by extreme values.	It depends only on particular values when the data was ranked.
Standard deviation	Use all the data and so represent every item. It is useful in comparing two sets of data, to show the consistence.	It is affected by one or two extreme values.

Diagrams:

	Advantages	**Disadvantages**
Stem-and-leaf diagram	It shows all the original data. It shows the shape of the distribution. The mode, median and quartile can be found. It is useful for comparing two sets of data.	It is not suitable for large amount of data.
Histogram	It can represent groups of different widths. It shows whether the distribution is symmetrical or skew The mean and the standard deviation can be estimated from the histogram.	The visual impact can be altered by choosing different groups. Two distributions cannot be shown on the same diagram.
Cumulative frequency graph	The median and quartiles can be estimated the graph. Sets of data can be compared by drawing graphs on the same diagram	The visual impact can be altered by using different scales.
Box-and-whisker plot	It is easy to see whether the distribution is symmetrical or whether there is a tail to the left or right. It can be used to find the extreme values. It is easy to see the range and interquartile range. You can compare two or more sets of data by drawing plots on the same diagram.	It does not show frequencies.

Miscellaneous exercise 1

1. A company manager, faced with the possibility of having to reduce staff during a recsssion, compiles a table of the ages of his employees.

Age, completed years	$20-30$	$31-35$	$36-40$	$41-45$	$46-50$	$51-60$
Number in age group	5	7	18	30	16	9

Draw:

(a) a cumulative frequency graph and so find estimates for the three quartiles,

(b) a histogram,

(c) a box-and-whisker plot,

to illustreate these figures.

2. The monthly salaries, w dollars, of 10 women are such that $\sum(w-300)=-200$.

The monthly salaries, m dollars, of 20 men are such that $\sum(m-4000)=120$.

(a) Find the difference between the mean monthly salary of the women and the mean monthly salary of the men.

(b) Find the mean monthly salary of all the women and men together.

3. Eighty candidates took an examination in Astronomy, for which no candidate scored more than 80%. The examiners suggest that five grades, A, B, C, D and E, should be awarded to these candidates, using upper grade boundaries 64, 50, 36 and 26 for grades B, C, D and E, respectively. In this case, grades A, B, C, D and E, will be awarded in the ratio $1:3:5:4:3$.

(a) Using the examiners' suggestion, represent the scores in a cumulative frequency graph and use it to estimate the median score.

(b) All of the grade boundaries are later reduced by 10%. Estimate how many candidates will be awarded a higher grade because of this.

4. A set of n pieces of data has mean $\bar{x}$ and standard deviation S. Another set of $2n$ pieces of data has mean $\bar{x}$ and standard deviation $\frac{1}{2}S$. Find the standard deviation of all these pieces of data together in term of S.

5. The daily journey times for 80 bank staff to get to work are given in the following table.

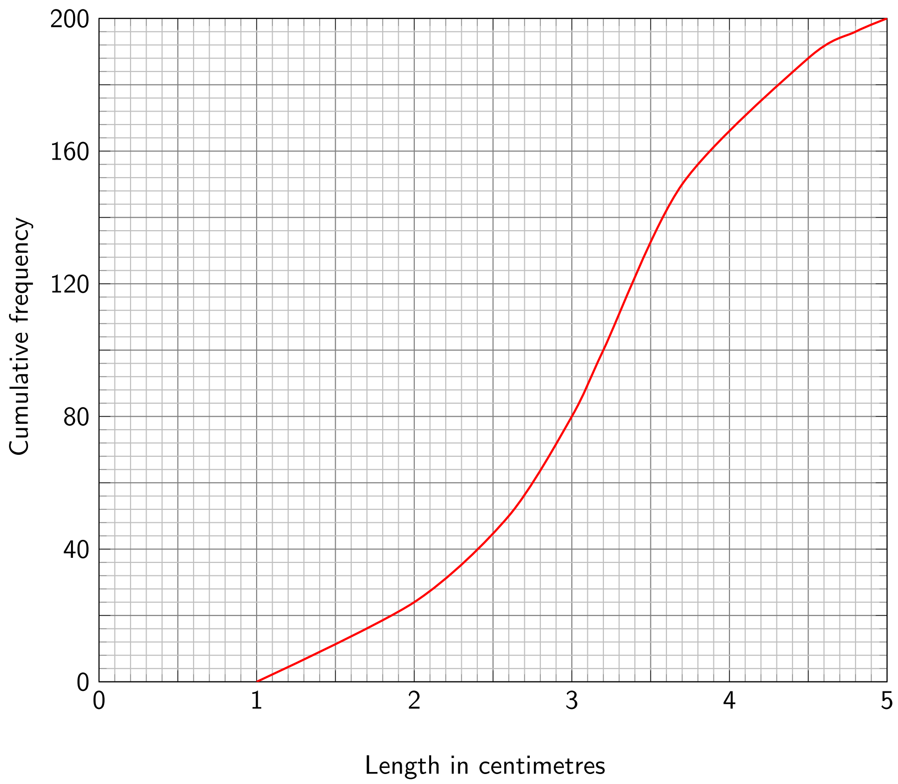
Time (t min)	$t<10$	$t<15$	$t<20$	$t<25$	$t<30$	$t<45$	$t<60$
No.staff (cf)	3	11	24	56	68	76	80

(a) How many staff take between 15 and 45 minutes to get to work?

(b) Find the exact number of staff who take $\frac{x+y}{2}$ minutes or more to get to work, given that 85% of the staff take less than x minutes and that 70% of the staff take y minutes or more.

Exam-style Questions 1

1. Anabel measured the lengths, in centimetres, of 200 caterpillars. Her results are illustrated in the cumulative frequency graph below.



   (i) Estimate the median and the interquartile range of the lengths. [3]

   (ii) Estimate how many caterpillars had a length of between 2 and 3.5 cm. [1]

   (iii) 6% of caterpillars were of length l centimetres or more. Estimate l. [2]

2. The heights, in cm, of the 11 members of the Anvils athletics team and the 11 members of the Brecons swimming team are shown below.

Anvils	173	158	180	196	175	165	170	169	181	184	172
Brecons	166	170	171	172	172	178	181	182	183	183	192

   (i) Draw a back-to-back stem-and-leaf diagram to represent this information, with Anvils on the left-hand side of the diagram and Brecons on the right-hand side. [4]

   (ii) Find the median and the interquartile range for the heights of the Anvils. [3]

   The heights of the 11 members of the Anvils are denoted by x cm. It is given that $\sum x = 1923$, and $\sum x^2 = 337221$. Suppose the Anvils are joined by 3 new members whose heights are 166 cm, 172 cm and 182 cm.

   (iii) Find the standard deviation of the heights of all 14 members of the Anvils. [4]

3. The speeds, in km h^{-1}, of 90 cars as they passed a certain marker on a road were recorded, correct to the nearest km h^{-1}. The results are summarised in the following table.

Speeds (km h^{-1})	$10-29$	$30-39$	$40-49$	$50-59$	$60-89$
Frequency	10	24	30	14	12

(i) On the grid, draw a histogram to illustrate the data in the table. [4]

(ii) Calculate an estimate for the mean speed of these 90 cars as they pass the marker. [2]

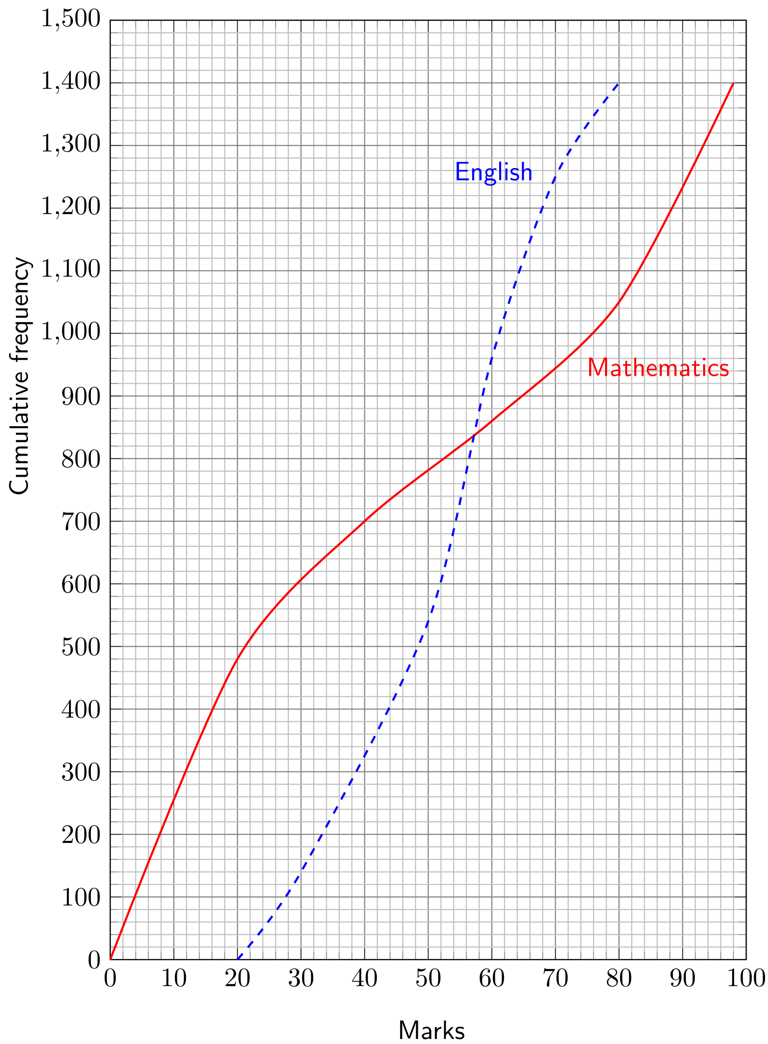
4. Farfield Travel and Lacket Travel are two travel companies which arrange tours abroad. The numbers of holidays arranged in a certain week are recorded in the table below, together with the means and standard deviations of the prices.

	Number of holidays	Mean price ($)	Standard deviation ($)
Farfield Travel	30	1500	230
Lacket Travel	21	2400	160

(i) Calculate the mean price of all 51 holidays. [2]

(ii) The prices of individual holidays with Farfield Travel are denoted by $\$x_F$ and the prices of individual holidays with Lacket Travel are denoted by $\$x_L$. By first finding $\sum x_F^2$ and $\sum x_L^2$, find the standard deviation of the prices of all 51 holidays. [5]

5. The Mathematics and English A-level marks of 1400 pupils all taking the same examinations are shown in the cumulative frequency graphs below. Both examinations are marked out of 100.



Use suitable data from these graphs to compare the central tendency and spread of the marks in Mathematics and English. [6]

6. The following back-to-back stem-and-leaf diagram shows the reaction times in seconds in an experiment involving two groups of people, A and B.

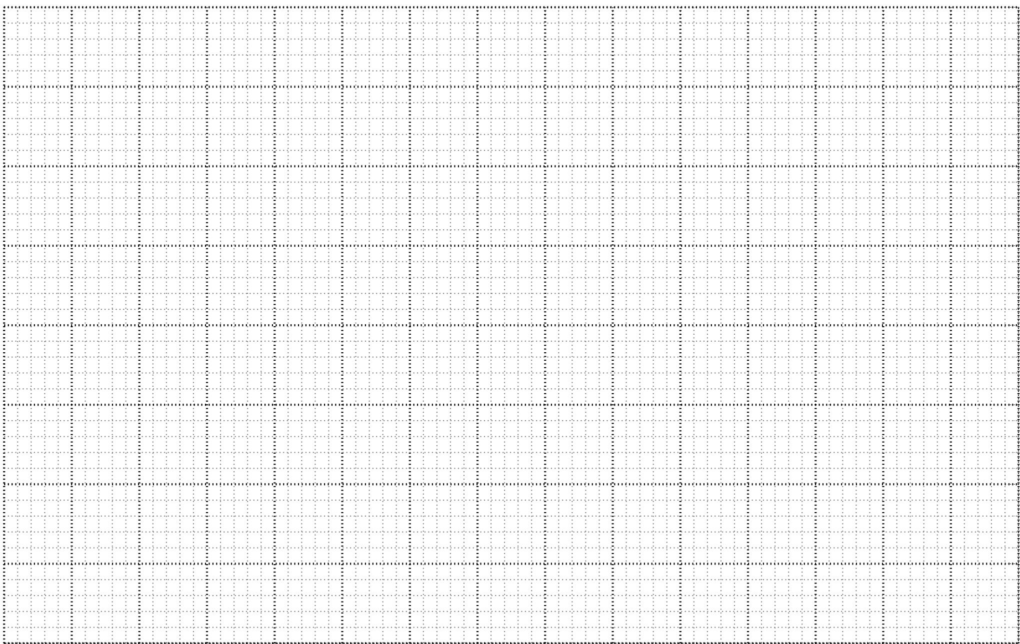
				A										B				
(4)					4	2	0	0	20	5	6	7						(3)
(5)				9	8	5	0	0	21	1	2	2	3	7	7			(6)
(8)	9	8	7	5	3	2	2	2	22	1	3	5	6	6	8	9		(7)
(6)			8	6	7	5	2	1	23	4	5	7	8	8	9	9	9	(8)
(3)						8	6	3	24	2	4	5	6	7	8	8		(7)
(1)								0	25	0	2	7	8					(4)

Key: $5|22|6$ means a reaction time of 0.225 seconds for A and 0.226 seconds for B

(i) Find the median and the interquartile range for group A. [3]

The median value for group B is 0.235 seconds, the lower quartile is 0.217 seconds and the upper quartile is 0.245 seconds.

(ii) Draw box-and-whisker plots for groups A and B on the grid. [3]



2 Permutations and combinations

2.1 Arrangements in a line

- Arrangements of **distinct** items

  The number of different arrangements of n distinct items is

$$n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1 = n!$$

- Arrangements when items are **not** distinct

  The number of different arrangements of n items of which p of one type are alike, q of another type are alike, r of another type are alike, and so on, is

$$\frac{n!}{p! \times q! \times r! \times \ldots}.$$

- Arrangements when there are restrictions
  - particular items have to be together, or must be separated.
  - not all $\&$ all not.
- Arrangemnets when **repetitions** are allowed

Exercise 9

1. Each of the letters of the word **CAMBRIDGE** is written on a card and the cards are placed in a line.
   (a) How many different arrangements are there?
   (b) How many arrangements begin with **CAM**.
2. Find the number of different arrangements using all ten letters of the word **STATISTICS**.
3. The word **ARGENTINA** includes the four consonants **R**, **G**, **N**, **T** and the three vowels **A**, **E**, **I**.
   (a) Find the number of different arrangements using all nine letters.
   (b) How many of these arrangements have a consonant at the beginning, then a vowel, then another consonant, and so on alternately?
4. Issam has 11 different CDs, of which 6 are pop music, 3 are jazz and 2 are classical. How many different arrangements of all 11 CDs on a shelf are there if the jazz CDs are all next to each other.
5. The eight sopranos in a choir are asked to stand in a line, but Ruby and Grace refuse to stand next to each other. How many different arrangements can there be?
6. How many 5-digit **odd** numbers can be made with the digits 2, 3, 6, 7, 8
   (a) if repetitions are not allowed, for example, $63\,287$.
   (b) if repetitions are allowed, for example, $88\,663$.
7. (a) Safebank requires its customers to use a four-digit PIN to access their account. Customers can choose any set of 4 digits from 0, 1, 2, $\ldots$, 9 and digits may be repeated. How many possible four-digit PINs are there?
   (b) Smartbank requires its customers to use a password consisting of four lower-case letters. Repetitions are allowed. How many possible passwords are there?
   (c) Excelbank requires its customers to use a pass-code consisting of four letters followed by four digits. Repititions are allowed. How many possible pass-codes are there?

Exercise 10

1. Find the number of ways to arrange the letters of each of the following words:
   (a) SPIDER
   (b) SANDWICH
   (c) PENGUIN
   (d) BLACKBERRY
   (e) MATHEMATICAL
2. Find the number of ways in which all eight letters of the word **ADVANCED** can be arranged if the arrangement must begin and end with an **A**.
3. Find the number of ways in which all eight letters of the word **NEEDLESS** can be arranged:
   (a) if there are no restrictions,
   (b) if the arrangement must end with **N**,
   (c) if the three letters **E** must be placed next to each other,
   (d) if the two letters **S** must not be placed next to each other,
   (e) if the three letters **E** must be placed together and the letter **S** must not be place together.
4. Five boys and four girls sit on a bench. In how many ways can they be seated if no two boys sit next to each other?
5. Three identical yellow ballons, two identical red ballons and two identical blue ballons are strung in a row to celebrate Shema's birthday. Calculate the number of arrangements if:
   (a) the ballon at each end is the same colour,
   (b) the yellow ballons are next to each other and the blue ballons are not next to each other.
6. Find how many arrangements there are of the nine letters in the word **GOLD MEDAL**
   (a) if there are no restrictions in the order of the letters,
   (b) if the two letters **D** come first and the two letters **L** come last.
7. Four identical tins of peaches and six identical tins of pears are arranged in a row on a shelf. Calculate the number of different arrangements if the tin at eache end contains the same type of fruit.
8. Three girls and seven boys stand in a line. Calculate the number of different arrangements if:
   (a) the two youngest pupils are separated,
   (b) all three girls stand together.
9. There are 10 seats in a row in a theatre. In how many ways can 5 couples be seated in a row if each couple sits together?
10. Find how many different arrangements there are of eleven letters of the word **PROBABILITY** if the two letters **B** are at the beginning and the two letters **I** are at the end.
11. $\boxed{1}$ $\boxed{3}$ $\boxed{3}$ $\boxed{8}$ $\boxed{8}$ $\boxed{8}$
    The above cards are placed in a line to form a 6-digit number. How many numbers can be made:
    (a) if there are no restrictions,
    (b) if the number ends in 3,
    (c) if the number is odd?

2.2 Permuations of r items from n items

Suppose A, B, C, D, E, F, G are put into 4 spaces,

number of ways of arrangements will be $7 \times 6 \times 5 \times 4$, which is also denoted by ${}_7\mathsf{P}_4$. Notice that

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

In general, the number of permutations, of r items taken from n **distinct** item is

$${}_n\mathsf{P}_r = \frac{n!}{(n-r)!}.$$

Special case, $0! = 1$.

Restrictions:

- Not distinct.

- Together.

- others.

Challenge: 8 students sit on 12 chairs in a row. Among these eight students, A and B must be together. Find the number of different arrangements.

Exercise 11

1. Find how many numbers bigger than $30\,000$ but smaller than $40\,000$ can be formed the digits $2, 3, 4, 5, 6, 7, 8$ if no digit is repeated and the number must be a multiple of 5.
2. A security code cosists of 4 letters chosen from A, B, C, D, E, F, G followed by 3 digits chosen from 0, 1, 2, 3, 4, 5.

   Examples are $BCDG102$ (without repetitions) and $CCDD225$ (with repetitions).

   Show that more than five times as many codes can be made when repetitions are allowed than when repetitions are not allowed.
3. Rory is playing a game in which he has to place coloured pegs into holes in a board. He has 6 identical red pegs and the board has 10 holes. How many different arrangements are there for placing 6 pegs and leaving 4 empty holes?
4. If repetitions are not allowed, how many numbers can be formed with the digits 3, 4, 5, 6 7
   (a) using three of the digits,
   (b) using one or more of the digits?
5. There are 10 seats in the front row at a theatre. Six people are shown to this row. In how many different ways can they be seated if
   (a) there are no restrictions,
   (b) two particular people in the group must sit next to each other?

2.3 Combinations of r items from n items

A combination is a selection of some items where the order of the selected item ______________________.

In general, the number of combinations of r items from n **distinct** items is given by

$$\binom{n}{r} = {}_n\mathsf{C}_r = \frac{n!}{r!(n-r)!}.$$

Notice, "**distinct**" is a very important conditions for combinations.

What happens when the selections are from items that are **NOT** distinct?

For example: Three letters are selected at random from the letters of the word **BIOLOGY**. Find the total number of selections.

Exercise 12

1. A team of 3 is to be chosen from 10 athletes. How many different teams could be chosen?
2. Without using a calculator, evaluate $\binom{12}{9}$ and $\binom{12}{3}$.
3. Issam has 11 different CDs of which 6 are pop music, 3 are jazz and 2 are classical. Issam makes a selection of 2 pop music CDs, 2 jazz CDs are 1 classical CD.

   How many different possible selections can be made?
4. A collection of 18 books contain one Harry Potter book. Linda is going to choose 6 of these books to take on holiday.
   (a) In how many ways can she choose 6 books?
   (b) How many of these choices will include the Harry Potter book?
5. A committee of 5 people is to be chosen from 6 men and 4 women. In how many ways can this be done:
   (a) if there must be 3 men and 2 women on the committee,
   (b) if there must be more than men than women on the committee,
   (c) if there must be 3 men and 2 women, and one particular woman refuses to be on the committee with one particular man?
6. In a mixed pack of coloured light bulbs there are three red bulbs, one yellow bulbs, one blue bulbs and one green bulbs. Four bulbs are selected at random from the pack. How many different selections are possible?
7. Four letters are to be selected from the letters in the word $RIGIDITY$. How many different combinations are there?
8. The letters of the word $POSSESSES$ are written on nine cards, one on each card. The cards are shuffled and four of them are selected and arranged in a straight line.
   (a) How many possible selections are there of four letters?
   (b) How many arrangments are there of four letters?

Miscellaneous exercise 2

1. (a) A plate of cakes holds 12 different cakes. Find the number of ways these cakes can be shared between Alex and James if each receives an odd number of cakes. [3]

   (b) Another plate holds 7 cup cakes, each with a different colour icing, and 4 brownies, each of a different size. Find the number of different ways these 11 cakes can be arranged in a row if no brownie is next to another brownie. [3]

   (c) A plate of biscuits holds 4 identical chocolate biscuits, 6 identical shortbread biscuits and 2 identical gingerbread biscuits. These biscuits are all placed in a row. Find how many different arrangements are possible if the chocolate biscuits are all kept together. [3]

2. A selection of 3 letters from the 8 letters of the word **COLLIDER** is made.

   (a) How many different selections of 3 letters can be made if there is exactly one **L**? [1]

   (b) How many different selections of 3 letters can be made if there are no restrictions? [3]

3. Find the number of different arrangements that can bemade of all 9 letters in the word **CAMERAMAN** in each of the following cases.

   (a) There are no restrictions. [2]

   (b) The **A**s occupy the 1st, 5th and 9th positions. [1]

   (c) There is exactly one letter between the **M**s. [4]

4. (a) Eight children of different ages stand in a random order in a line. Find the number of different ways this can be done if none of the three youngest children stand next to each other. [3]

   (b) David chooses 5 chocolates from 6 different dark chocolates, 4 different white chocolates and 1 milk chocolate. He must choose at least one of each type. Find the number of different selections he can make. [4]

   (c) A password for Chelsea's computer consists of 4 characters in a particular order. The characters are chosen from the following.

   - The 26 capital letters A to Z
   - The 9 digits 1 to 9
   - The 5 symbols $\# \sim * \, ? \, !$

   The password must include at least one capital letter, at least one digit and at least one symbol. No character can be repeated. Find the number of different passwords that Chelsea can make. [4]

5. A library contains 4 identical copies of book A, 2 identical copies of book B and 5 identical copies of book C. These 11 books are arranged on a shelf in the library.

   (a) Calculate the number of different arrangements if the end books are either both book A or both book B. [4]

   (b) Calculate the number of different arrangements if all the books A are next to each other and none of the books B are next to each other. [5]

6. (a) Find how many numbers between 3000 and 5000 can be formed from the digits 1, 2, 3, 4 and 5,

   i. if digits are not repeated, [2]

   ii. if digits can be repeated and the number formed is odd. [3]

   (b) A box of 20 biscuits contains 4 different chocolate biscuits, 2 different oatmeal biscuits and 14 different ginger biscuits. 6 biscuits are selected from the box at random.

   i. Find the number of different selections that include the 2 oatmeal biscuits. [2]

   ii. Find the probability that fewer than 3 chocolate biscuits are selected. [4]

Exam-style Questions 2

1. Find the number of different ways in which all 9 letters of the word **MINCEMEAT** can be arranged in each of the following cases.

   (i) There are no restrictions. [1]

   (ii) No vowel (**A**, **E**, **I** are vowels) is next to another vowel. [4]

   5 of the 9 letters of the word **MINCEMEAT** are selected.

   (iii) Find the number of possible selections which contain exactly 1 **M** and exactly 1 **E**. [2]

   (iv) Find the number of possible selections which contain at least 1 **M** and at least 1 **E**. [3]

2. (i) Find the number of ways in which all 9 letters of the word **AUSTRALIA** can be arranged in each of the following cases.

   (a) All the vowels (**A**, **I**, **U** are vowels) are together. [3]

   (b) The letter **T** is in the central position and each end position is occupied by one of the other consonants (**R**, **S**, **L**). [3]

   (ii) Donna has 2 necklaces, 8 rings and 4 bracelets, all different. She chooses 4 pieces of jewellery. How many possible selections can she make if she chooses at least 1 necklace and at least 1 bracelet? [4]

3. Find the number of ways the 9 letters of the word **SEVENTEEN** can be arranged in each of the following cases.

   (i) One of the letter **E**s is in the centre with 4 letters on either side. [2]

   (ii) No **E** is next to another **E**. [3]

   5 letters are chosen from the 9 letters of the word **SEVENTEEN**.

   (iii) Find the number of possible selections which contain exactly 2 **E**s and exactly 2 **N**s. [1]

   (iv) Find the number of possible selections which contain at least 2 **E**s. [4]

4. Freddie has 6 toy cars and 3 toy buses, all different. He chooses 4 toys to take on holiday with him.

   (i) In how many different ways can Freddie choose 4 toys? [1]

   (ii) How many of these choices will include both his favourite car and his favourite bus? [2]

   Freddie arranges these 9 toys in a line.

   (iii) Find the number of possible arrangements if the buses are all next to each other. [3]

   (iv) Find the number of possible arrangements if there is a car at each end of the line and no buses are next to each other. [3]

5. (i) A group of 6 teenagers go boating. There are three boats available. One boat has room for 3 people, one has room for 2 people and one has room for 1 person. Find the number of different ways the group of 6 teenagers can be divided between the three boats. [3]

   (ii) Find the number of different 7-digit numbers which can be formed from the seven digits 2, 2, 3, 7, 7, 7, 8 in each of the following cases.

   (a) The odd digits are together and the even digits are together. [3]

   (b) The 2s are not together. [4]

6. (i) Find the number of ways a committee of 6 people can be chosen from 8 men and 4 women if there must be at least twice as many men as there are women on the committee. [3]

   (ii) Find the number of ways a committee of 6 people can be chosen from 8 men and 4 women if 2 particular men refuse to be on the committee together. [3]

7. (i) A village hall has seats for 40 people, consisting of 8 rows with 5 seats in each row. Mary,Ahmad, Wayne, Elsie and John are the first to arrive in the village hall and no seats are taken before they arrive.
   (a) How many possible arrangements are there of seatingMary, Ahmad,Wayne, Elsie and John assuming there are no restrictions? [2]
   (b) How many possible arrangements are there of seatingMary, Ahmad,Wayne, Elsie and John if Mary and Ahmad sit together in the front row and the other three sit together in one of the other rows? [4]

   (ii) In how many ways can a team of 4 people be chosen from 10 people if 2 of the people, Ross and Lionel, refuse to be in the team together? [4]

8. (i) Find the number of different 3-digit numbers greater than 300 that can be made from the digits 1, 2, 3, 4, 6, 8 if
   (a) no digit can be repeated, [3]
   (b) a digit can be repeated and the number made is even. [3]

   (ii) A team of 5 is chosen from 6 boys and 4 girls. Find the number of ways the team can be chosen if
   (a) there are no restrictions, [1]
   (b) the team contains more boys than girls. [3]

9. A car park has spaces for 18 cars, arranged in a line. On one day there are 5 cars, of different makes, parked in randomly chosen positions and 13 empty spaces.

   (i) Find the number of possible arrangements of the 5 cars in the car park. [2]
   (ii) Find the probability that the 5 cars are not all next to each other. [5]

   On another day, 12 cars of different makes are parked in the car park. 5 of these cars are red, 4 are white and 3 are black. Elizabeth selects 3 of these cars.

   (iii) Find the number of selections Elizabeth can make that include cars of at least 2 different colours. [5]

10. 9 people are to be divided into a group of 4, a group of 3 and a group of 2. In how many different ways can this be done? [3]

11. In an orchestra, there are 11 violinists, 5 cellists and 4 double bass players. A small group of 6 musicians is to be selected from these 20.

   (i) How many different selections of 6 musicians can be made if there must be at least 4 violinists, at least 1 cellist and no more than 1 double bass player? [4]

   The small group that is selected contains 4 violinists, 1 cellist and 1 double bass player. They sit in a line to perform a concert.

   (ii) How many different arrangements are there of these 6musicians if the violinistsmust sit together? [3]

12. (i) How many different arrangements are there of the 11 letters in the word **MISSISSIPPI**? [2]
   (ii) Two letters are chosen at random from the 11 letters in the word **MISSISSIPPI**. Find the probability that these two letters are the same. [3]

13. (i) Find the number of different ways that 5 boys and 6 girls can stand in a row if all the boys stand together and all the girls stand together. [3]
   (ii) Find the number of different ways that 5 boys and 6 girls can stand in a row if no boy stands next to another boy. [3]

14. A group consists of 5 men and 2 women. Find the number of different ways that the group can stand in a line if the women are not next to each other. [3]

15. Out of a class of 8 boys and 4 girls, a group of 7 people is chosen at random.

    (i) Find the probability that the group of 7 includes one particular boy. [3]

    (ii) Find the probability that the group of 7 includes at least 2 girls. [4]

16. (i) Find the number of different ways in which all 12 letters of the word **STEEPLECHASE** can be arranged so that all four **E**s are together. [1]

    (ii) Find the number of different ways in which all 12 letters of the word **STEEPLECHASE** can be arranged so that the **S**s are not next to each other. [4]

    Four letters are selected from the 12 letters of the word **STEEPLECHASE**.

    (iii) Find the number of different selections if the four letters include exactly one **S**. [4]

17. (i) Find the number of different ways in which the 9 letters of the word **TOADSTOOL** can be arranged so that all three **O**s are together and both **T**s are together. [1]

    (ii) Find the number of different ways in which the 9 letters of the word **TOADSTOOL** can be arranged so that the **T**s are not together. [4]

    (iii) Find the probability that a randomly chosen arrangement of the 9 letters of the word **TOAD-STOOL** has a **T** at the beginning and a **T** at the end. [2]

    (iv) Five letters are selected fromthe 9 letters of the word **TOADSTOOL**. Find the number of different selections if the five letters include at least 2 **O**s and at least 1 **T**. [4]

18. (i) How many different arrangements are there of the 9 letters in the word **CORRIDORS**? [2]

    (ii) How many different arrangements are there of the 9 letters in the word **CORRIDORS** in which the first letter is **D** and the last letter is **R** or **O**? [3]

19. A sports team of 7 people is to be chosen from 6 attackers, 5 defenders and 4 midfielders. The team must include at least 3 attackers, at least 2 defenders and at least 1 midfielder.

    (i) In how many different ways can the team of 7 people be chosen? [4]

    The team of 7 that is chosen travels to a match in two cars. A group of 4 travel in one car and a group of 3 travel in the other car.

    (ii) In how many different ways can the team of 7 be divided into a group of 4 and a group of 3? [2]

3 Probability

3.1 Introduction

The probability of an event is a measure of the likehood that it will happen.

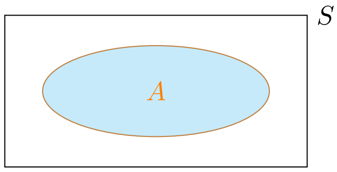
- A probability of 0 indicates that the event is ________________.
- A probability of 1 indicates that the event is ____________________________.
- All other events have a probability between ______ and ______.

Notation

The set of all possible outcomes is the **possibility space**, S, and the number of outcomes in the possibility space is written $n(S)$.

The event A consists of one or more of the outcomes in S. The number of outcomes resulting in event A is written $n(A)$.

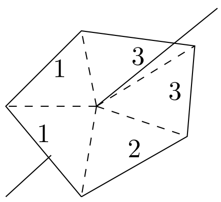
When outcomes are ________________________, the probability of event A is written $P(A)$ where



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

Exercise 13

1. A box contains 20 counters numbered 1,2,3,..., up to 20. A counter is picked at random from the box. Find the probability that the number on the counter is
   (a) a multiple of 5,
   (b) not a multiple of 5,
   (c) higher than 7.
2. A fair five-sided spinner has sides numbered 1, 1, 2, 3, 3. The spinner is spun twice. Find the probability that the spinner will stop at 1 at least once.



3. A card is dealt from a well-shuffled ordinary pack of 52 playing cards.
   (a) Find the probability that the card is
      i. the 3 of the spades,
      ii. the 3 of the spades or any diamond.

(b) The first card dealt is placed face-up on the table. It is the 3 of diamonds. What is the probability that the second card is from a red suit?

4. The table shows the results of all the driving tests taken at a particular test centre during the first week of September. A person is chosen at random from those who took their driving test that week.

	Male	**Female**
Pass	32	43
Fail	10	15

   (a) Find the probability that the person passed the driving test.

   (b) Find the probability that the person is a female who failed the driving test.

   (c) A male is chosen. What is the probability that he passed the driving test.

5. An ordinary tetrahedral die has four faces and they are labelled 1, 2, 3, 4. When the die is thrown, the score is the number on which the die lands. Two fair tetrahedral dice are thrown. By using a possibility space diagram, or otherwise, find the probability that

   (a) the sum of the scores is divisible by 4,

   (b) the product of the scores is an even number,

   (c) the scores differ by at least 2.

6. Two fair coins are tossed together. Find the probability that

   (a) exactly one tail is obtained,

   (b) at most one head is obtained.

7. Two ordinary fair cubical dice are thrown. Find the probability that

   (a) the sum of the numbers on the dice is 3,

   (b) the sum of the numbers on the dice exceeds 9,

   (c) the dice show the same number,

   (d) the numbers on the dice differ by more than 2.

8. Two ordinary fair cubical dice are thrown at the same time and the scores are multipled. $\mathsf{P}(N)$ denotes the probability that the number N will be obtained.

   (a) Find

      i. $\mathsf{P}(9)$,

      ii. $\mathsf{P}(4)$,

      iii. $\mathsf{P}(14)$,

      iv. $\mathsf{P}(37)$.

   (b) If $\mathsf{P}(N) = \frac{1}{9}$, find the possible values of N.

9. Three fair coins are tossed,

   (a) List all possible outcomes,

   (b) Find the probability that two heads and one tail are obtained.

3.2 Using permutations and combinations

In order to find the number of outcomes in a particular event and in the possibility space, you may need to use arrangements permutations and combinations.

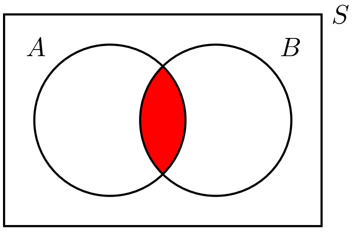
Exercise 14

1. Evan throws three fair dice.
   (a) List all possible scores on the three dice which give a total of 5, and hence show that the probability of Evan obtaining a total score of 5 is $\frac{1}{36}$.
   (b) Find the probability of Evan obtaining a total score of 7.
2. Each of the eleven letters of the word **MATHEMATICS** is written on a separate card and cards are laid out in a line.
   (a) Calculate the number of different arrangements of these letters.
   (b) Find the probability that all the vowels are placed together.
3. Four letters are chosen at random from the letters in the word **RANDOMLY**. Find the probability that all letters chosen are consonants.
4. A team of 5 pupils is chosen from a class of 7 girls and 8 boys. Find the probability that the team consists of 3 girls and 2 boys.
5. A staff park at a school has 13 parking spaces in a row. There are 9 cars to be parked.
   (a) How many different arrangements are there for parking the 9 cars and leaving 4 empty spaces?
   (b) How many different arrangements are there if the 4 empty spaces are next to each other?
   (c) If the parking is random, find the probability that there will not be 4 empty spaces next to each other?
6. Minerva is given a bag of 20 sweets of which 6 are apple flavoured , 6 are lemon flavoured and 8 are orange flavoured. Minerva takes out 5 sweets at random and eats them. Find the probability that she eats:
   (a) 5 orange flavoured sweets,
   (b) 3 apple flavoured and 2 lemon flavoured sweets,
   (c) exactly 2 apple flavoured sweets,
   (d) no lemon flavoured sweets.
7. A plate contains 15 cakes of which 6 have yellow icing, 5 have green icing and 4 have pink icing. Three cakes are taken at random from the plate.

   Find the probability that:
   (a) exactly two of the cakes have green icing,
   (b) one cake has green icing, one has pink icing and one has yellow icing,
   (c) none of the cakes has yellow icing.
8. Peter deals a hand of 10 cards from a well-shuffled pack of ordinary playing cards. Show the probability that she deals exactly 5 spades is less than 5%.

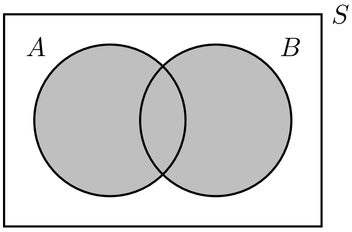
3.3 Two or more events

Now consider two events, A and B, in the possibility tree.



$A \cap B$

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



$A \cup B$

$\mathsf{P}(A \text{ or } B) = \mathsf{P}(A \cup B) = \mathsf{P}(\qquad\qquad\qquad\qquad)$

Addition rule for combined events:

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

Using set notation:

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B).$$

Exercise 15

1. Two events, X and Y, are such that $\mathsf{P}(X \text{ or } Y) = 0.8$, $\mathsf{P}(X \text{ and } Y) = 0.35$ and $\mathsf{P}(X) = 0.6$. Find $\mathsf{P}(Y')$.

2. Events A and B are such that $\mathsf{P}(A \text{ occurs}) = 0.6$, $\mathsf{P}(B \text{ occurs}) = 0.7$, $\mathsf{P}(\text{at least one of } A \text{ and } B \text{ occurs}) = 0.9$.

   Find

   (a) $\mathsf{P}(\text{both } A \text{ and } B \text{ occur})$.

   (b) $\mathsf{P}(\text{neither } A \text{ and } B \text{ occurs})$.

   (c) $\mathsf{P}(A \text{ occurs or } B \text{ occurs but not both } A \text{ and } B \text{ occurs})$

3. Some pupils did a survey on comics. They asked all 100 pupils in theri year group whether they read particular comics during the past week. They found that 65 had read Whizz, 55 had read Wham, 30 had read Whizz and Wham and some pupils in the year group had not read either comic.

   A pupil was selected at random from the year group to answer more questions in the survey. Find the probability that the pupil

   (a) had read Whizz or Wham,

   (b) had not read either comic.

Mutually exclusive events

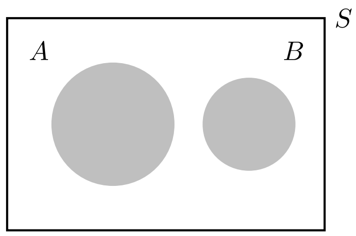
Events are mutually exclusive if they cannot occur at the same time.

For example, with one throw of a die, the events 'scoring a 3' and 'scoring a 5' are mutually exclusive, since the score cannot be 3 and 5 at the same time.

However, the events 'scoring an even number' and 'scoring a prime number' are not mutually exclusive, since a score of 2 is both even and prime.

Set notation and Venn diagram:

When A and B are mutually exclusive there is no overlap between A and B.



$A \cap B = \emptyset$

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B)$$

Note: if you are asked to show that event A and event B are mutually exclusive, you must give working to show either the following is satisfied:

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) \qquad\qquad \mathsf{P}(A \cap B) = 0.$$

Exercise 16

1. In a race where there can be only one winner, the probability that John will win is 0.3, the probability that Paul will win is 0.2 and the probability that Mark will win is 0.4.

   Find the probabilibty that

   (a) John or Mark wins,
   (b) John or Paul or Mark wins,
   (c) someone else wins.

2. A card is dealt from an ordinary pack of 52 playing cards. Find the probability that the card is

   (a) a club or a diamond,
   (b) a club or a king.

3. Two fair dice are shown.

   (a) Event A is 'the scores differ by 3 or more'. Find the probability of event A.
   (b) Event B is 'the product of the scores is greater than 8'. Find the probability of event B.
   (c) State with a reason whether events A and B are mutually exclusive.

4. Two fair cubical dice are thrown.

   Event A is 'the scores on the dice are the same',

   Event B is 'the product of the scores is a multiple of 3',

   Event C is 'the sum of the scores is 7'.

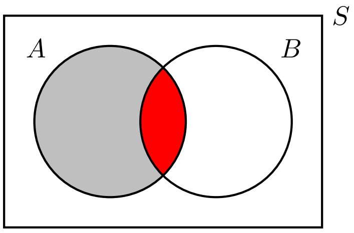
   State with a reason whether A and B, A and C, B and C are mutually exclusive.

3.4 Condtional probability

Condtional probability is used when the probability that an event will occur depends on whether another event has happened.

For event A and B, the conditional probability that event B occurs, given that A has already occured, i.e.

$$\mathrm{P}(B \text{ given } A) = \mathrm{P}(B|A)$$



A and B

$$\mathrm{P}(B|A) = \frac{n(A \text{ and } B)}{n(A)} = \frac{\frac{n(A \text{ and } B)}{n(A)}}{} = \frac{\mathrm{P}(A \text{ and } B)}{\mathrm{P}(A)}.$$

This gives the multiplication rule:

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

Exercise 17

1. There are 5 red counters and 7 blue counters in a bag. Darian takes a counter from the bag, puts it on the table and then takes another counter from the bag. Find the probability that he takes out
   (a) a red counter then a blue counter,
   (b) two counters that are the same colour,
   (c) at least one red counter.
2. Two events X and Y are such that $\mathrm{P}(X) = 0.2$, $\mathrm{P}(Y) = 0.25$, $\mathrm{P}(Y|X) = 0.4$.
   Find
   (a) $\mathrm{P}(X \text{ and } Y)$, (b) $\mathrm{P}(X|Y)$, (c) $\mathrm{P}(X \text{ or } Y)$.
3. Of the 120 first year students at a college, 36 study chemistry, 60 study biology and 10 study both chemistry and biology. A first year student is selected at random to represent the college at a conference. Find the probability that the student studies
   (a) chemistry, given that the student studie biology,
   (b) biology, given that the student studies chemistry.
4. Last month a consultant saw 60 men and 65 women suspected of having a particular eye condition. Tests were carried out and the following table shows the results. The totals are shown in bold.

	Had eye condition (C)	**Did not have eye condition (C')**	
Man (M)	25	35	**60**
Woman (W)	20	45	**65**
	45	**80**	**125**

   One of these patients was selected at random to take part in a survey. Find the probability that the patient selected
   (a) was a women, given that the patient had the eye condition
   (b) had the eye condition, given that the patient was a man.

Independent events

In general, the following rule is correct for any events A and B

$$\mathsf{P}(A\,\text{and}\,B) = \mathsf{P}(A) \times \mathsf{P}(B|A).$$

Now when either of these events can occur without being affected by the outcome of the other, the events are said to be _________________.

So, for **independent events** A and B,

$$\mathsf{P}(B|A) = \mathsf{P}(\quad)$$

Also

$$\mathsf{P}(A|B) = \mathsf{P}(\quad)$$

This gives the multiplication rule for **independent events**:

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

In set notation

$$\mathsf{P}(\qquad\quad) = \mathsf{P}(A) \times \mathsf{P}(B).$$

Exercise 18

1. There are 5 red counters and 7 blue counters in a bag. Eliza takes a counter from the bag, notes its colour and then puts it back into the bag. Freddie then takes a counter from the bag. Find the probability that
   (a) Eliza takes a red counter and Freddie takes a blue counter,
   (b) Freddie's counter is the same colour as Eliza's counter.
2. Two fair cubical dice are thrown. One is red and the other is blue. Find the probability that
   (a) the score is 3 on both dice,
   (b) neither die has a score of 3.
3. The probability that a certain type of machine will break down in the first month of operation is 0.1. Three machines of this type are installed at the same time. The performances of the three machines are independent. Find the probability that at the end of the first month
   (a) all three machines have broken down,
   (b) just one machine has broken down,
   (c) at least one machine is working.
4. In a large group of people it is known that 10% have a hot breakfast, 20% have a hot lunch and 25% have a hot breakfast or a hot lunch. Find the probability that a person chosen at random from this group
   (a) has a hot breakfast and a hot lunch.
   (b) has a hot lunch, given that the person chosen had a hot breakfast.
5. Two events A and B are such that $\mathsf{P}(A|B) = 0.4$, $\mathsf{P}(B|A) = 0.25$ and $\mathsf{P}(A\,\text{and}\,B) = 0.12$.
   (a) Are A and B independent? Give a reason for your answer.
   (b) Find $\mathsf{P}(A\,\text{or}\,B)$.

6. Data about employment for males and females in a small rural area are shown in the table.

	Unemployed	**Employed**
Male	206	412
Female	358	305

A person from this area is chosen at random. Let M be the event that the person is male and let E be the event that the person is employed.

(a) Find P(M).

(b) Find P$(M \text{ and } E)$.

(c) Are M and E independent events? Justify your answer.

(d) Given that the person chosen is unemployed, find the probability that the person is female.

7. Two ordinary fair dice, one red and one blue, are thrown.

Events A, B and C are defined as follows:

Event A: the number showing on the red die is 5 or 6

Event B: the total of the numbers showing on the two dice is 7

Event C: the total of the numbers showing on the two dice is 8

(a) State, with a reason, which two of the events A, B and C are mutually exclusive.

(b) Show that the events A and B are independent.

8. A school has 100 teachers. In a survey on the use of the school car park, the teachers were asked whether they had driven a car to school on a particular day. Of the 70 full-time teachers, 45 had driven a car to school and of the 30 part-time teachers, 12 had driven a car to school.

(a) Copy and complete the two-way table, where C denotes the event 'the teacher had driven a car to school that day'.

	C	C'	Total
Full-time teacher			
Part-time teacher			
Total			100

(b) Find the probability that a teacher chosen at random

i. is a part-time teacher who had driven a car to school,

ii. is a full-time teacher who had not driven a car to school,

iii. is a full-time teacher or had driven a car to school,

iv. is a part-time teacher, given that the teacher had driven a car to school.

(c) Are the events 'the teacher had driven a car to school' and 'the teacher is full-time' independent? Give a reason for your answer.

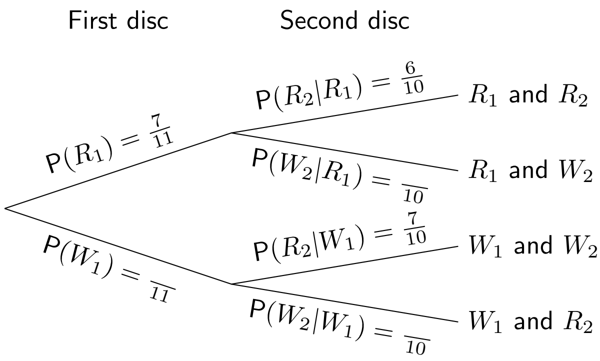
(d) Describe two events that are mutally exclusive.

3.5 Probability trees

A very useful method for tacking many probability problems is to draw a tree diagram.

For example, a jar contains seven red discs and four white discs. Two disc are selected without replacement.

A tree diagram can be formed as follows:



then you can find the probability that both of the discs are red by:

$$\mathsf{P}(R_1 \text{ and } R_2) = \mathsf{P}(R_1) \times \mathsf{P}(R_2|R_1) = \qquad \times \qquad = \qquad .$$

Meanwhile, the probability of the second discs is red is given by:

$$\mathsf{P}(R_2) = \mathsf{P}(\qquad \text{and } R_2) + \mathsf{P}(\qquad \text{and } R_2) =$$

Exercise 19

1. In a country A 30% of people who drink tea have sugar in it. In country B 65% of people who drink tea have sugar in it. There are 3 million in country A who drink tea and 12 million in country B who drink tea. A person is chosen at random from these 15 million people.
   (a) Find the probability that the person chosen is from country A.
   (b) Find the probability that the person chosen does not have sugar in their tea.
   (c) Given that chosen does not have sugar in their tea, find the probability that the person is from country B.

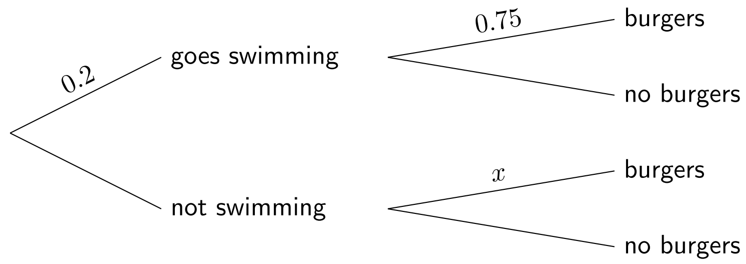
2. There are three sets of traffic lights on Karinne's journey to work. The independent probabilities that Karinne has to stop at the first, second and third sets of lights are 0.4, 0.8 and 0.3 respectively.
   (a) Draw a tree diagram to show this information.
   (b) Find the probability that Karinne has to stop at each of the first two sets of lights but does not have to stop at the third set of lights.
   (c) Find the probability that Karinne has to stop at exactly two of the three sets of lights.
   (d) Find the probability that Karinne has to stop at the first set of lights, given that she has to stop at exactly two sets of lights.

3. Dan is playing a game in which players pick counters at random, one at a time without replacement, from a bag. At the beginning of the game, the bag contains 6 red counters and 4 blue counters. Dan takes two counters from the bag.
   (a) Find the probability that both counters are blue.
   (b) Find the probability that the counters are the same color.

(c) Given that the counters are the same color, find the probability that they are both blue.

4. When a farmer's dog is let loose, it chases either ducks with probability $\frac{3}{5}$ or geese with probability $\frac{2}{5}$. If the dog chases the ducks, there is a probability of $\frac{1}{10}$ that they will attack the dog. If the dog chases the geese, there is a probability of $\frac{3}{4}$ that they will attack the dog. Given that the dog is not attacked, find the probability that it was chasing the geese.

5. The probability that Henk goes swimming on any day is 0.2. On a day that he goes swimming, the probability that Henk has burgers for supper is 0.75. On a day when he does not go swimming, the probability that he has burgers for his supper is x. The information is shown on the following tree diagram.



The probability that Henk has burgers for supper on any day is 0.5.

   (a) Find x.
   (b) Given that Henk has burgers for supper, find the probability that he went swimming that day.

6. When Don plays tennis, 65% of his first serves go into the correct are of the court. If the first serve goes into the correct area, his chance of winning the point is 90%. If his first serve does not go into the correct area, Don is allowed a second serve and, of these, 80% go into the correct area. If the second serve goes into the correct, his chance of winning the point is 60%. If neither serve goes into the correct area, Don loses the point.

   (a) Draw a tree diagram to represent this information.
   (b) Using the tree diagram, find the probability that Don loses the point.
   (c) Find the conditional probability that Don's first serve went into the correct area, given that he loses the point.

7. In an archery competition, Bill is allowed up to three attempts to hit the target. If he succeeds on any attempt, he does not make any more attempts. The probability that he will hit the target on the first attempt is 0.6. If he misses, the probability that he will hit the target on his second attempt is 0.7. If he misses on the second attempt, the probability that he will hit the target on his third attempt is 0.8.

   (a) Draw a fully labelled tree diagram.
   (b) Find the probability that Bill will hit the target.
   (c) Given that Bill hits the target, find the probability that he made at least two attempts.

8. Boxes of sweets contain toffees and chocolates. Box A contains 6 toffees and 4 chocolates, box B contains 5 toffees and 3 chocolates and box C contains 3 toffees and 7 chocolates. One of the boxes is chosen at random and two sweets are taken out, one after the other, and eaten.

   (a) Find the probability that they are both toffees.
   (b) Given that they are both toffees, find the probability that they both came from box A.

Miscellaneous exercise 3

1. A bag contains 10 pink balloons, 9 yellow balloons, 12 green balloons and 9 white balloons. 7 balloons are selected at random without replacement. Find the probability that exactly 3 of them are green. [3]

2. Last Saturday, Sarah recorded the colour and type of 160 cars in a car park. All the cars that were not red or silver in colour were grouped together as 'other'. Her results are shown in the following table.

		Type of Car		
		Saloon	Hatchback	Estate
Color of car	Red	20	40	12
	Silver	14	26	10
	Other	6	24	8

   (a) Find the probability that a randomly chosen car in the car park is a silver estate car. [1]

   (b) Find the probability that a randomly chosen car in the car park is a hatchback car. [1]

   (c) Find the probability that a randomly chosen car in the car park is red, given that it is a hatchback car. [2]

   (d) One of the cars in the car park is chosen at random. Determine whether the events 'the car is a hatchback car' and 'the car is red' are independent, justifying your answer. [2]

3. Ashfaq throws two fair dice and notes the numbers obtained. R is the event 'The product of the two numbers is 12'. T is the event 'One of the numbers is odd and one of the numbers is even'. By finding appropriate probabilities, determine whether events R and T are independent. [5]

4. Two identical biased triangular spinners with sides marked 1, 2 and 3 are spun. For each spinner, the probabilities of landing on the sides marked 1, 2 and 3 are p, q and r respectively. The score is the sum of the numbers on the sides on which the spinners land. You are given that P(score is 6) $= \frac{1}{6}$ and P(score is 5) $= \frac{1}{9}$. Find the values of p, q and r. [6]

5. A shop sells two makes of coffee, Café Premium and Café Standard. Both coffees come in two sizes, large jars and small jars. Of the jars on sale, 65% are Café Premium and $35\%$are Café Standard. Of the Café Premium, 40% of the jars are large and of the Café Standard, 25% of the jars are large. A jar is chosen at random.

   (a) Find the probability that the jar is small. [2]

   (b) Find the probability that the jar is Café Standard given that it is large. [3]

6. Vehicles approaching a certain road junction from town A can either turn left, turn right or go straight on. Over time it has been noted that of the vehicles approaching this particular junction from town A, 55% turn left, 15% turn right and 30% go straight on. The direction a vehicle takes at the junction is independent of the direction any other vehicle takes at the junction.

   (a) Find the probability that, of the next three vehicles approaching the junction from town A, one goes straight on and the other two either both turn left or both turn right. [4]

   (b) Three vehicles approach the junction from town A. Given that all three drivers choose the same direction at the junction, find the probability that they all go straight on. [4]

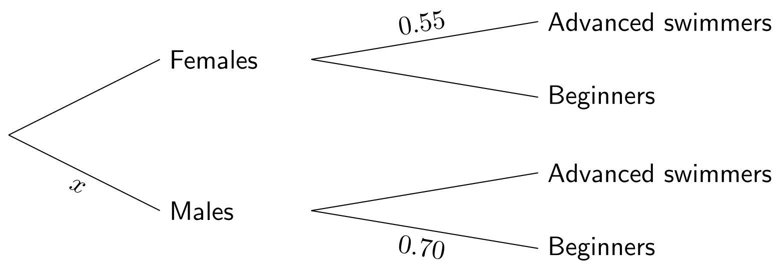
7. In a group of students, $\frac{3}{4}$ are male. The proportion of male students who like their curry hot is $\frac{3}{5}$ and the proportion of female students who like their curry hot is $\frac{4}{5}$. One student is chosen at random.

   (a) Find the probability that the student chosen is either female, or likes their curry hot, or is both female and likes their curry hot. [4]

   (b) Showing your working, determine whether the events 'the student chosen ismale' and 'the student chosen likes their curry hot' are independent. [2]

Exam-style Questions 3

1. The members of a swimming club are classified either as 'Advanced swimmers' or 'Beginners'. The proportion of members who are male is x, and the proportion of males who are Beginners is 0.7. The proportion of females who are Advanced swimmers is 0.55. This information is shown in the tree diagram.



For a randomly chosen member, the probability of being an Advanced swimmer is the same as the probability of being a Beginner.

(i) Find x [3]

(ii) Given that a randomly chosen member is an Advanced swimmer, find the probability that the member is male. [3]

2. Two ordinary fair dice are thrown and the numbers obtained are noted. Event S is 'The sum of the numbers is even'. Event T is 'The sum of the numbers is either less than 6 or a multiple of 4 or both'. Showing your working, determine whether the events S and T are independent. [4]

3. Over a period of time Julian finds that on long-distance flights he flies economy class on 82% of flights. On the rest of the flights he flies first class. When he flies economy class, the probability that he gets a good night's sleep is x. When he flies first class, the probability that he gets a good night's sleep is 0.9.

(i) Draw a fully labelled tree diagram to illustrate this situation. [2]

The probability that Julian gets a good night's sleep on a randomly chosen flight is 0.285.

(ii) Find the value of x. [2]

(iii) Given that on a particular flight Julian does not get a good night's sleep, find the probability that he is flying economy class. [3]

4. In a group of students, the numbers of boys and girls studying Art, Music and Drama are given in the following table. Each of these 160 students is studying exactly one of these subjects.

	Art	Music	Drama
Boys	24	40	32
Girls	15	12	37

(i) Find the probability that a randomly chosen student is studying Music. [1]

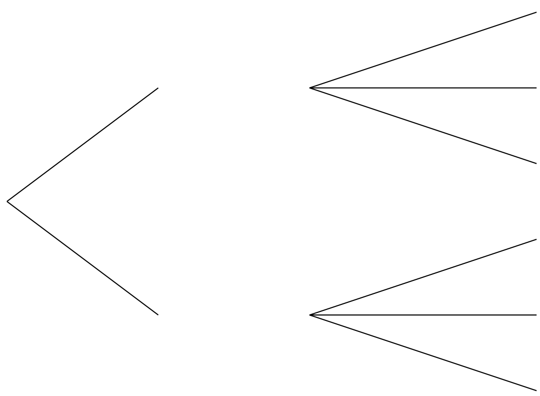
(ii) Determine whether the events 'a randomly chosen student is a boy' and 'a randomly chosen student is studying Music' are independent, justifying your answer. [2]

(iii) Find the probability that a randomly chosen student is not studying Drama, given that the student is a girl. [2]

(iv) Three students are chosen at random. Find the probability that exactly 1 is studying Music and exactly 2 are boys. [5]

5. A box contains 3 red balls and 5 blue balls. One ball is taken at random from the box and not replaced. A yellow ball is then put into the box. A second ball is now taken at random from the box.

   (i) Complete the tree diagram to show all the outcomes and the probability for each branch. [2]



   (ii) Find the probability that the two balls taken are the same colour. [2]

   (iii) Find the probability that the first ball taken is red, given that the second ball taken is blue. [3]

6. Deeti has 3 red pens and 1 blue pen in her left pocket and 3 red pens and 1 blue pen in her right pocket. 'Operation T' consists of Deeti taking one pen at random from her left pocket and placing it in her right pocket, then taking one pen at random from her right pocket and placing it in her left pocket.

   (i) Find the probability that, when Deeti carries out operation T, she takes a blue pen from her left pocket and then a blue pen from her right pocket. [2]

   The random variable X is the number of blue pens in Deeti's left pocket after carrying out operation T.

   (ii) Find $\mathsf{P}(X) = 1$. [3]

   (iii) Given that the pen taken from Deeti's right pocket is blue, find the probability that the pen taken from Deeti's left pocket is blue. [4]

7. A fair eight-sided die has faces marked 1, 2, 3, 4, 5, 6, 7, 8. The score when the die is thrown is the number on the face the die lands on. The die is thrown twice.

   - Event R is 'one of the scores is exactly 3 greater than the other score'.
   - Event S is 'the product of the scores is more than 19'.

   (i) Find the probability of R. [2]

   (ii) Find the probability of S. [2]

   (iii) Determine whether events R and S are independent. Justify your answer. [3]

8. Jason throws two fair dice, each with faces numbered 1 to 6. Event A is 'one of the numbers obtained is divisible by 3 and the other number is not divisible by 3'. Event B is 'the product of the two numbers obtained is even'.

   (i) Determine whether events A and B are independent, showing your working. [5]

   (ii) Are events A and B mutually exclusive? Justify your answer. [1]

9. When Joanna cooks, the probability that the meal is served on time is $\frac{1}{5}$. The probability that the kitchen is left in a mess is $\frac{3}{5}$. The probability that the meal is not served on time and the kitchen is not left in a mess is $\frac{3}{10}$. Some of this information is shown in the following table.

	Kitchen left in a mess	Kitchen left not in a mess	Total
Meal served on time			$\frac{1}{5}$
Meal not served on time		$\frac{3}{10}$	
Total			1

(i) Copy and complete the table. [3]

(ii) Given that the kitchen is left in a mess, find the probability that the meal is not served on time. [2]

10. Tom and Ben play a game repeatedly. The probability that Tom wins any game is 0.3. Each game is won by either Tom or Ben. Tom and Ben stop playing when one of them (to be called the champion) has won two games.

(i) Find the probability that Ben becomes the champion after playing exactly 2 games. [1]

(ii) Find the probability that Ben becomes the champion. [3]

(iii) Given that Tom becomes the champion, find the probability that he won the 2nd game. [4]

4 Discrete random variables

4.1 Introduction

A discrete random variables is a variable which can take individual values each with a given ___________.

Notation:

- Random variables are denoted by ________________
- Particular values of variables are denoted by ________________
- The probability that the variable X takes a particular value is written ________________

Probability distributions:z

A list of all possible values of the discrete random variable X, together with their associated probabilities.

x	x_1	x_2	x_3	$\dots$	x_n
$\mathsf{P}(X = x)$	p_1	p_2	p_3	$\dots$	p_n

Sum of probabilities:

$$\mathsf{P}(X = x_1) + \mathsf{P}(X = x_2) + \mathsf{P}(X = x_3) + \cdots + \mathsf{P}(X = x_n) = 1$$

Alternatively you can write

$$p_1 + p_2 + \cdots + p_n = 1$$

Exercise 20

1. Emma is playing a game with a blased five-sided spinner marked with the numbers 1,2, 3, 4 and 5.

   When she spins the spinner, her score, X, is the number on which the spinner lands. The probability distribution of X is shown in the table.

x	1	2	3	4	5
$\mathsf{P}(X = x)$	0.15	0.24	a	0.25	0.19

   (a) Find the value of a.
   (b) Find the probability that the score is at least 4.
   (c) Find the probability that the score is less than 5.
   (d) Find $\mathsf{P}(2 < x \leqslant 4)$.
   (e) Write down the most likely score.

2. Lancelot decides to replace the two used batteries in his torch with new ones. Unfortunately, when he takes them out, he mixes them up with three new batteries. All five batteries are identical in appearance.

   Lancelot selects two of the batteries at random. Draw up a probability distribution table for X, the number of **new** batteries that Lancelot selects.

3. A vegetable basket contains 12 peppers, of which 3 are red, 3 are green and 5 are yellow. Three peppers are taken at random, without replacement, from the basket.

   (a) Find the probability that the three peppers are all different colours.
   (b) Show that the probability that exactly 2 of the peppers taken are green is $\frac{12}{55}$.

(c) The number of **green** peppers taken is denoted by the discrete random variable X. Draw up a probability distribution table for X.

4. In a probability distribution the random variable X takes the value x with probability kx, where x takes the value 5, 10, 15, 20 and 25 only.

   Draw up a probability distribution for X, in terms of k, and find the value of k.

5. Sherry has two fair tetrahedral (four-sided) dice. The faces on each die are labelled 1, 2, 3 and 4. One die is red and the other is blue. Sherry throws each die once. The random variable X is the sum of the numbers on which the dice land.

   (a) Find the probability that $X = 4$.

   (b) Draw up a probability distribution table for X.

   (c) Given that $X = 6$, find the probability that the red die landed on 2.

6. Laura is playing a game in which he tries to throw tennis balls into a bucket. The probability that the tennis ball lands in the bucket is 0.4 for each attempt.

   Laura has three attempts.

   (a) By drawing a tree diagram, or otherwise, show that the probability that exactly one tennis ball lands in the bucket is 0.432.

   (b) Draw up a probability distribution table for X, the number of tennis balls that land in the bucket.

   Laura wins a prize if at least two tennis balls land in the bucket.

   (c) What is the probability that he wins a prize.

4.2 $\mathsf{E}(X)$, the expectation of X

The expectation of a random variable X is the result that you would expect to get if you took a very large number of values of X and found their mean.

It is written ___________ and is denoted by the symbol ___________.

If the discrete random variable X has the following probability distribution

x	x_1	x_2	x_3	$\dots$	x_n
$\mathsf{P}(X = x)$	p_1	p_2	p_3	$\dots$	p_n

The expectation is given by:

$$\mu = \mathbf{E}(X) = x_1p_1 + x_2p_2 + x_3p_3 + \cdots + x_np_n.$$

Note:

- A practical approach results is a frequency distribution and an experiment mean $\bar{x}$.
- A theoretical approach uses a probability distribution and results in an expected mean μ.

Exercise 21

1. Natasha plays a fairground game. She throws an unbiased tetrahedral die with faces numbered 1,2,3 and 4. If the die lands on the face marked 1 she has to pay \$1. If it lands on 3 she wins 30 cents. If it lands on 2 or 4 Natasha wins 50 cents.

   (a) Find the expected profit in a single throw.

   (b) If the fairground owner changes the rules so that Natasha has to pay \$1.30 if the die lands on 1, what will be Natasha's expected profit in a single throw?

2. The discrete random variable X has the following probability distribution

x	1	3	5	7
$\mathsf{P}(X = x)$	0.3	a	b	0.25

   (a) Write down an equation satisfied by a and b.

   (b) Given that $\mathsf{E}(X) = 4$, find a and b.

3. In a competition, people pay \$1 to throw a ball at a target. If they hit the target on the first throw, they receive \$5. If they hit it on the second or third throw they receive \$3, and if they hit it on the fourth or fifth throw, they receive \$1. People stop throwing after the first hit, or after 5 throws if no hit is made. Marlo has a constant probability of $\frac{1}{5}$ of hitting the target on any throw, independently of the results of the other throws.

   (a) Mario misses with his first and second throws and hits the target with his third throw. State how much profit he has made.

   (b) Show that the probability that Mario's profit is \$0 is 0.184, correct to 3 significant figures.

   (c) Draw up a probability distribution table for Mario's profit.

   (d) Calculate his expected profit.

4.3 Var(X), the variance of X

The variance of a discrete random variable is a measure of the ________ of X about the expected mean μ.

It is written ___________ and is denoted by ___________.

If the discrete random variable X has the following probability distribution

x	x_1	x_2	x_3	$\ldots$	x_n
P$(X = x)$	p_1	p_2	p_3	$\ldots$	p_n

The variance is given by:

$$\sigma^2 = \text{Var}(X) = \sum (x_i - \mu)^2 p_i.$$

Alternatively,

$$\sigma^2 = \text{Var}(X) = \sum x_i^2 p_i - \mu^2. \qquad \text{where} \quad \mu = \text{E}(X) = \sum x_i p_i.$$

Exercise 22

1. The table below shows the probability distribution of the discrete random variable X.

x	1	2	3	4
P$(X = x)$	0.1	0.3	0.45	0.15

(a) Calculate E(x). (b) Calculate Var(x). (c) Find the standard deviation.

2. A discrete random variable X has the following probability distribution.

x	1	2	3	4
P$(X = x)$	$3c$	$4c$	$5c$	$6c$

(a) Find the value of the constant c.
(b) Find E(X) and Var(X).
(c) Find P$(X > \text{E}(X))$.

3. A small farm has 5 ducks and 2 geese. Four of these birds are to be chosen at random. The random variable X represents the number of geese chosen.

(a) Draw up the probability distribution of X.
(b) Show that E$(X) = \frac{8}{7}$ and calculate Var(X).

4. Box A contains 5 red paper clips and 1 white paper clip. Box B contains 7 red papaer clips and 2 white paper clips. One paper clip is taken at random from box A and transfered to Box B. One paper clip is then taken at random from Box B.

(a) Find the probability of taking both a white paper clip from Box A and a red paper clip from Box B.
(b) Find the probability that the paper clip taken from Box B is red.
(c) Find the probability that the paper clip taken from Box A was red, given that the paper clip taken from Box B is red.
(d) The random variable B denotes the number of times that a red paper clip is taken. Draw up a table to show the probability distribution of X.

Miscellaneous exercise 4

1. Pack A consists of ten cards numbered 0, 0, 1, 1, 1, 1, 1, 3, 3, 3. Pack B consists of six cards numbered 0, 0, 2, 2, 2, 2. One card is chosen at random from each pack. The random variable X is defined as the sum of the two numbers on the cards.

   (a) Show that $\mathsf{P}(X = 2) = \frac{2}{15}$. [2]

   (b) Draw up the probability distribution table for X. [4]

   (c) Given that $X = 3$, find the probability that the card chosen from pack A is a 1. [3]

2. The discrete random variable X has the following probability distribution.

x	-2	0	1	3	4
$\mathsf{P}(X = x)$	0.2	0.1	p	0.1	q

   (a) Given that $\mathsf{E}(X) = 1.7$, find the values of p and q. [4]

   (b) Find $\mathsf{Var}(x)$. [2]

3. The random variable X takes the values -1, 1, 2, 3 only. The probability that X takes the value x is kx^2, where k is a constant.

   (a) Draw up the probability distribution table for X, in terms of k, and find the value of k. [3]

   (b) Find $\mathsf{E}(x)$ and $\mathsf{Var}(x)$. [3]

4. Andy has 4 red socks and 8 black socks in his drawer. He takes 2 socks at random from his drawer.

   (a) Find the probability that the socks taken are of different colours. [2]

   The random variable X is the number of red socks taken.

   (b) Draw up the probability distribution table for X. [3]

   (c) Find $\mathsf{E}(X)$. [1]

5. A game is played with 3 coins, A, B and C. Coins A and B are biased so that the probability of obtaining a head is 0.4 for coin A and 0.75 for coin B. Coin C is not biased. The 3 coins are thrown once.

   (a) Draw up the probability distribution table for the number of heads obtained. [5]

   (b) Hence calculate the mean and variance of the number of heads obtained. [3]

6. At a funfair, Amy pays $\$1$ for two attempts to make a bell ring by shooting at it with a water pistol.

   - If she makes the bell ring on her first attempt, she receives $\$3$ and stops playing. This means that overall she has gained $\$2$.
   - If she makes the bell ring on her second attempt, she receives $\$1.50$ and stops playing. This means that overall she has gained $\$0.50$.
   - If she does not make the bell ring in the two attempts, she has lost her original $\$1$.

   The probability that Amy makes the bell ring on any attempt is 0.2, independently of other attempts.

   (a) Show that the probability that Amy loses her original $\$1$ is 0.64. [2]

   (b) Construct a probability distribution table for the amount that Amy gains. [4]

   (c) Calculate Amy's expected gain. [1]

Exam-style Questions 4

1. The discrete random variable X has the following probability distribution.

x	1	2	3	6
$\mathsf{P}(X = x)$	0.15	p	0.4	q

   Given that $\mathsf{E}(X) = 3.05$, find the values of p and q. [4]

2. A fair five-sided spinner has sides numbered $1, 1, 1, 2, 3$. A fair three-sided spinner has sides numbered $1, 2, 3$. Both spinners are spun once and the score is the product of the numbers on the sides the spinners land on.

   (i) Draw up the probability distribution table for the score. [4]

   (ii) Find the mean and the variance of the score. [3]

   (iii) Find the probability that the score is greater than the mean score. [2]

3. Maryam has 7 sweets in a tin, 6 are toffees and 1 is a chocolate. She chooses one sweet at random and takes it out. Her friend adds 3 chocolates to the tin. Then Maryam takes another sweet at random out of the tin.

   (i) Draw a fully labelled tree diagram to illustrate this situation. [3]

   (ii) Draw up the probability distribution table for the number of toffees taken. [3]

   (iii) Find the mean number of toffees taken. [1]

   (iv) Find the probability that the first sweet taken is a chocolate, given that the second sweet taken is a toffee. [4]

4. A fair red spinner has 4 sides, numbered $1, 2, 3, 4$. A fair blue spinner has 3 sides, numbered $1, 2, 3$. When a spinner is spun, the score is the number on the side on which it lands. The spinners are spun at the same time. The random variable X denotes the score on the red spinner minus the score on the blue spinner.

   (i) Draw up the probability distribution table for X. [3]

   (ii) Find $\mathsf{Var}(X)$. [3]

   (iii) Find the probability that X is equal to 1, given that X is non-zero. [3]

5. A box contains 3 red balls and 5 white balls. One ball is chosen at random from the box and is not returned to the box. A second ball is now chosen at random from the box.

   (i) Find the probability that both balls chosen are red. [1]

   (ii) Show that the probability that the balls chosen are of different colours is $\frac{15}{28}$. [2]

   (iii) Given that the second ball chosen is red, find the probability that the first ball chosen is red. [2]

   The random variable X denotes the number of red balls chosen.

   (iv) Draw up the probability distribution table for X. [2]

   (v) Find $\mathsf{Var}(X)$. [3]